REFERENCE ONLY

UNIVERSITY OF LONDON THESIS

Degree PhD Year 2000 Name of Author IBRAHIM A.

COPYRIGHT

This is a thesis accepted for a Higher Degree of the University of London. It is an unpublished typescript and the copyright is held by the author. All persons consulting the thesis must read and abide by the Copyright Declaration below.

COPYRIGHT DECLARATION

I recognise that the copyright of the above-described thesis rests with the author and that no quotation from it or information derived from it may be published without the prior written consent of the author.

LOANS

Theses may not be lent to individuals, but the Senate House Library may lend a copy to approved libraries within the United Kingdom, for consultation solely on the premises of those libraries. Application should be made to: Inter-Library Loans, Senate House Library, Senate House, Malet Street, London WC1E 7HU.

REPRODUCTION

University of London theses may not be reproduced without explicit written permission from the Senate House Library. Enquiries should be addressed to the Theses Section of the Library. Regulations concerning reproduction vary according to the date of acceptance of the thesis and are listed below as guidelines.

A. Before 1962. Permission granted only upon the prior written consent of the author. (The Senate House Library will provide addresses where possible).

B. 1962 - 1974. In many cases the author has agreed to permit copying upon completion of a Copyright Declaration.

C. 1975 - 1988. Most theses may be copied upon completion of a Copyright Declaration.

D. 1989 onwards. Most theses may be copied.

This thesis comes within category D.

☑ This copy has been deposited in the Library of UCL

☐ This copy has been deposited in the Senate House Library, Senate House, Malet Street, London WC1E 7HU.

THE DISTRIBUTION OF COMPOUNDS BETWEEN BLOOD OR THE GAS PHASE AND VARIOUS BIOLOGICAL TISSUES

A THESIS PRESENTED TO THE UNIVERSITY OF LONDON FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN THE FACULTY OF MATHEMATICAL AND PHYSICAL SCIENCES (MAPS)

BY ADAM IBRAHIM

SUPERVISOR: PROFESSOR MICHAEL H ABRAHAM

UNIVERSITY COLLEGE LONDON
20 GORDON STREET
LONDON
WC1H 0AJ

2006

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

THE DISTRIBUTION OF COMPOUNDS BETWEEN BLOOD OR THE GAS PHASE AND VARIOUS BIOLOGICAL TISSUES

TABLE OF CONTENTS

Title page 1

Contents page 2

List of tables 11

Abstract 25

Acknowledgements 26

CHAPTER 1 **Introduction to the blood tissue distribution of volatile organic compounds (VOCs)** 27

References 29

CHAPTER 2 **Volatile organic compounds (VOCs) in general and their environmental importance** 30

The sources of indoor air pollutant 32

Measurement of volatile organic compounds in indoor air 34

Biological monitoring 35

The amount of indoor air pollutants 36

The health effects of VOCs, both indoors and outdoors 38

VOCs that acts as anaesthetic vapours 40

The oil-gas partition coefficient 40

What determines how quickly an anaesthetic wears off? 41

References 42

CHAPTER 3 **Air to blood and air to tissue distribution of volatile organic compounds (VOCs)** 44

Experimental methods for obtaining air to blood (or air to tissue) partition coefficient for rats and humans using head space chromatography 50

References 52

CHAPTER 4 **Introduction to blood-tissue distribution of drugs** 53

Aim and plan 57

How important is drug distribution data and how is this obtained 59

References 67

CHAPTER 5 **Introduction to physicochemical considerations** 69

Brønsted acidity and basicity 69

Henderson-Hasselbach equation 71

Hydrophobic effect and the influence of lipophilicity 72

Hydrogen bonding 73

Index of lipophilicity from partition coefficients 73

How the effect of partition is affected by ionisation of the solute 75

References 78

CHAPTER 6 **Linear free-energy relationships (LFERs)** 79

Quantitative structure-activity relationships (QSAR) 79

Multiple linear regression analysis (MLRA) 81

Statistics 81

The Limitations of MLRA 84

References 86

CHAPTER 7 **The method of Abraham** **87**

The cavity model of solvation **89**

The general solvation equation **91**

The excess molar refraction, E (R_2) **92**

Solute dipolarity/ polarisability scale, S (Π_2^H) **93**

Hydrogen bond acidity of a solute, A **94**

Hydrogen bond basicity for solute, B **96**

Descriptor L or Log L^{16} **97**

The McGowan characteristic volume V (V_x) **98**

The coefficients used in Abraham equations **99**

References **101**

CHAPTER 8 **Methods of determination of descriptors** **103**

Absolv **104**

Unix **105**

Estimation of descriptors based on the analogy method **107**

Solver descriptor calculation **107**

Descriptor calculation for terbinafine (via the analogy method) **108**

Descriptor calculation for erythromycin (via the analogy method) **110**

References **113**

CHAPTER 9 **Combining Rat and Human Log K data for VOCs** **114**

Air to blood distribution in humans and rats **115**

Experimental error for three different VOCs **116**

Total regression of blood log K for humans and rats **120**

Air to tissue distribution for humans and rats **124**

Conclusion **127**

References **129**

CHAPTER 10 **Air to blood distribution for VOCs (rat and human average data)** 132

Discussion on air to blood distribution for VOCs in human and rat combined 134

References 139

CHAPTER 11 **Air to plasma, air to blood and plasma to blood distribution for VOCs** 140

Air to plasma distribution for VOCs 140

Discussion on air to plasma distribution for VOCs 141

Air to blood and plasma distribution for VOCs 144

Discussion on air to blood and plasma distribution for VOCs 144

Comparison between air to plasma and air to blood distribution data for common VOCs 148

Blood to plasma distribution for VOCs 150

Discussion on blood to plasma distribution for VOCs 151

References 154

CHAPTER 12 **Air to brain, blood to brain and plasma to brain distribution (rat and human) for VOCs** 155

Air to brain distribution for VOCs 155

Discussion on air to brain distribution for VOCs (humans and rats combined) 157

Blood to brain distribution for VOCs 161

Discussion on blood to brain distribution for VOCs (humans and rats combined) 163

Plasma to brain distribution for VOCs 167

Discussion on plasma to brain distribution for VOCs (humans and rats combined) 167

Blood to brain combined with plasma to brain distribution data for VOCs 170

Discussion on blood to brain and plasma to brain distribution for VOCs 170
References 175

CHAPTER 13 **Air to kidney and blood to kidney distribution for VOCs (human and rats)** 176
Air to kidney distribution for VOCs 176
Discussion on air to kidney distribution for VOCs (humans and rats combined) 178
Blood to kidney distribution for VOCs 182
Discussion on blood to kidney distribution for VOCs (humans and rats combined) 184
References 187

CHAPTER 14 **Air to fat and blood to fat distribution for VOCs (human and rats)** 188
Air to fat distribution for VOCs 188
Discussion on air to fat distribution for VOCs (humans and rats combined) 190
Blood to fat distribution for VOCs 194
Discussion on blood to fat distribution for VOCs (humans and rats combined) 196
References 199

CHAPTER 15 **Air to liver and blood to liver distribution for VOCs** 201
Air to liver distribution for VOCs 201
Discussion on air to liver distribution for VOCs (humans and rats combined) 203
Blood to liver distribution for VOCs 207
Discussion on blood to liver distribution for VOCs (humans and rats combined) 209
References 213

CHAPTER 16	**Air to muscle and blood to muscle distribution for VOCs**	**215**
	Air to muscle distribution for VOCs	**215**
	Discussion on air to muscle distribution for VOCs (humans and rats combined)	**217**
	Blood to muscle distribution for VOCs	**221**
	Discussion on blood to muscle distribution for VOCs (humans and rats combined)	**223**
	References	**227**
CHAPTER 17	**Air to urine distribution for VOCs (human only)**	**229**
	Discussion on air to urine distribution for VOCs in humans	**230**
	References	**235**
CHAPTER 18	**Air to olive oil distribution for VOCs**	**236**
	Air to olive oil distribution for VOCs	**236**
	Discussion on air to olive oil distribution for VOCs	**239**
	References	**243**
CHAPTER 19	**Air to saline and saline to olive oil distribution for VOCs**	**244**
	Air to saline distribution for VOCs	**244**
	Discussion on air to saline distribution for VOCs	**245**
	Saline to olive oil distribution for VOCs	**249**
	Discussion on saline to olive oil distribution for VOCs	**249**
	References	**253**
CHAPTER 20	**Conclusion on air to tissue and blood to distribution for VOCs**	**254**
	Coefficients e, s, a, b and l versus the tissue composition for air to tissue distribution of VOCs	**254**

Coefficients a, b and v versus the tissue composition for blood to tissue distribution of VOCs 262
References 268

CHAPTER 21 **Blood/plasma/serum to brain distribution for drugs and VOCs combined** 269
Blood/plasma/serum to brain distribution for drugs 269
Blood to brain/tissues distribution combined with plasma or serum to brain/tissues distribution data for drugs only 282
Discussion on blood/plasma/serum to brain distribution for drugs only 284
Combining the *in vivo* data of drugs with the *in vitro* data of VOCs 288
Discussion on blood/plasma/serum to brain distribution for drugs and VOCs combined 288
References 294

CHAPTER 22 **Blood/plasma to fat distribution for drugs and VOCs combined** 297
Discussion on blood/plasma to fat distribution for drugs 298
References 304

CHAPTER 23 **Blood/plasma to heart distribution for drugs and VOCs combined** 305
Discussion on blood/plasma to heart distribution for drugs only 306
Combining the *in vivo* data of drugs with the *in vitro* data of VOCs 310
Discussion on blood/plasma to heart distribution for drugs and VOCs combined 310
References 316

CHAPTER 24 **Blood/plasma/serum to lung distribution for drugs and VOCs combined** 318

Discussion on blood/plasma/serum to lung distribution for drugs only 320

Combining the *in vivo* data of drugs with the *in vitro* data of VOCs 324

Discussion on blood/plasma/serum to lung distribution for drugs and VOCs combined 324

References 330

CHAPTER 25 **Blood/plasma to muscle distribution for drugs and VOCs combined** 332

Discussion on blood/plasma to muscle distribution for drugs only 334

Combining the *in vivo* data of drugs with the *in vitro* data of VOCs 337

Discussion on blood/plasma to muscle distribution for drugs and VOCs combined 337

References 343

CHAPTER 26 **Blood/plasma/serum to liver distribution for drugs and VOCs combined** 345

Discussion on blood/plasma/serum to liver distribution for drugs only 347

Combining the *in vivo* data of drugs with the *in vitro* data of VOCs 350

Discussion on blood/plasma/serum to liver distribution for drugs and VOCs combined 350

References 356

CHAPTER 27 **Blood/plasma to skin distribution (rat only)** 358

Discussion on blood/plasma to skin distribution for drugs 359

References 363

CHAPTER 28 **Blood/plasma/serum to kidney distribution for drugs and VOCs combined** 364

Discussion on blood/plasma/serum to kidney distribution for drugs only 366

Combining the *in vivo* data of drugs with the *in vitro* data of VOCs 369

Discussion on blood/plasma/serum to kidney distribution for drugs and VOCs combined 369

References 375

CHAPTER 29 **Conclusion on blood to tissue distribution for VOCs and drugs combined** 377

Coefficients e, s, a, b, v and Ia versus the tissue composition for blood to tissue distribution of VOCs and drugs combined 378

References 390

CHAPTER 30 **Conclusion and future work** 391

Conclusion 391

Future work 392

References 393

APPENDIX **Tables for distribution data** 394

Structures for trivial names 517

VOCs 518

Drugs 519

LIST OF TABLES

CHAPTER 2		PAGE NO.
TABLE 2.0	Sources of VOCs, excluding methane, in the outdoor air of UK during 1995.	31
TABLE 2.1	Sources of indoor VOCs	33
TABLE 2.2	Summary of indoor and outdoor air concentrations and exposures for common air pollutants.	37
TABLE 2.3	Health effects of selected volatile organic compounds	39
TABLE 2.4	Blood-tissue partition coefficients (P)	41

CHAPTER 3		PAGE NO.
TABLE 3.0	Literature survey on VOCs air-blood (or air to tissue) distribution	46
TABLE 3.1	Literature survey on VOC blood to tissue distribution	48
TABLE 3.2	Literature survey on VOCs air to olive oil distribution (310K)	50

CHAPTER 4		PAGE NO.
TABLE 4.0	Shows how log BB changes with the functional group X on an aromatic ring system	55
TABLE 4.1	Summary of models for blood to brain distribution of drugs and VOCs	60
TABLE 4.2	Concentration of drug (adenosine) versus time	63
TABLE 4.3	Ratio for drug blood to tissue distribution obtained from data in Table 4.2	63

TABLE 4.4	Ratio of blood to tissue at steady state distribution	65

CHAPTER 7		*PAGE NO.*
TABLE 7.0	Abraham solute descriptors used today	92
TABLE 7.1	The solute comparison between α_2^H and A values	96
TABLE 7.2	Comparison between β^H_2 and B values	97
TABLE 7.3	Atom contributions for calculation of V (in $cm^3\ mol^{-1}$)	99
TABLE 7.4	Interpretation of the coefficients used by the Abraham solvation equations	99

CHAPTER 8		*PAGE NO.*
TABLE 8.0	Aliphatic functional group structural constants	105
TABLE 8.1	Aromatic functional group structural constants	106
TABLE 8.2	Calculated values for fluorescein obtained from solver	108
TABLE 8.3	Estimated descriptors for terbinafine	109
TABLE 8.4	Calculated values for terbinafine obtained from solver	109
TABLE 8.5	Estimated descriptors for erythromycin	110
TABLE 8.6	Comparison of the calculated descriptors obtained from the three methods for erythromycin	112

CHAPTER 9		*PAGE NO.*
TABLE 9.0	Air to blood distribution for VOCs (see appendix)	114
TABLE 9.1	Air to tissue (brain, fat, kidney, liver, lung, muscle and plasma) distribution for VOCs (see appendix)	114
TABLE 9.2	Plot for the common compounds between rat and human log K blood	115

TABLE 9.3	Statistics on common compounds between rat and human log K blood, this work	115
TABLE 9.4	Experimental data for air to blood distribution for acetone in humans	119
TABLE 9.5	Experimental data for air to blood distribution for chloroform in humans	119
TABLE 9.6	Experimental data for air to blood distribution for Trichloroethylene in humans	119
TABLE 9.7	Average standard deviation for three VOCs (air to blood in human)	120
TABLE 9.8	Human and Rat log K blood (blood log K verses descriptors)	120
TABLE 9.9	The range for the standard error (95% confidence interval) of the coefficient	121
TABLE 9.10	Equation for the common compounds between rat and human tissue log K (brain, fat, kidney, liver, lung, muscle and plasma)	124
TABLE 9.11	Statistics on the log K values for the common compounds between rat and human tissues	124
TABLE 9.12	Human and Rat log K tissue (fat log K verses descriptors	126
TABLE 9.13	The range for the standard error (95% confidence interval) of the coefficient (for fat log K)	126

CHAPTER 10		***PAGE NO.***
TABLE 10.0	Air to blood distribution for VOCs (see appendix)	134
TABLE 10.1	Summary of the equations for VOC air to blood distributions and statistics for test sets of 98 solutes	135
TABLE 10.2	The inter-correlation of descriptors for 196 solutes using descriptors E, S, A, B and L.	136

CHAPTER 11		PAGE NO.
TABLE 11.0	Data set used to obtain a model to correlate air to plasma distributions for VOCs in humans and rats combined (see appendix)	140
TABLE 11.1	Summary of the equations for VOC air to plasma distributions	141
TABLE 11.2	The inter-correlation of descriptors for 36 solutes using descriptors E, S, A, B and L	142
TABLE 11.3	Summary of the equations for VOC air to blood and plasma distributions and statistics for training sets and test sets of 116 solutes	145
TABLE 11.4	The inter-correlation of descriptors for 232 solutes using descriptors E, S, A, B and L	146
TABLE 11.5	Data set used in this model to obtain model to predict blood to plasma distributions for VOCs (see appendix)	148
TABLE 11.6	Statistics used to compare the common compounds for log K between air to plasma and air to blood distribution	148
TABLE 11.7	The comparison of the various coefficients	149
TABLE 11.8	Statistics used to compare the difference between observed and predicted values for plasma log K; the predicted values are from the air to blood equation	149
TABLE 11.9	Summary of the equations for VOC blood to plasma distributions	151
TABLE 11.10	The inter-correlation of descriptors for 35 solutes using descriptors E, S, A, B and V	152

CHAPTER 12		PAGE NO.
TABLE 12.0	Air to brain distribution for VOCs (see appendix)	155
TABLE 12.1	Summary of the equations for VOC air to brain distributions and statistics for test sets of 40 solutes	157

TABLE 12.2	The inter-correlation of descriptors for 81 solutes using descriptors E, S, A, B and L	158
TABLE 12.3	Blood to brain distribution for VOCs (see appendix)	163
TABLE 12.4	Summary of the equations for VOC blood to brain distributions and statistics for test sets of 39 solutes	163
TABLE 12.5	The inter-correlation of descriptors for 78 solutes using descriptors E, S, A, B and V	164
TABLE 12.6	Plasma to brain distribution for VOCs (see appendix)	167
TABLE 12.7	Summary of the equations for VOC plasma to brain distributions	168
TABLE 12.8	The inter-correlations of descriptors used for VOC plasma to brain distribution for 21 solutes	168
TABLE 12.9	Plasma to brain and blood to brain distribution combined for VOCs (see appendix)	170
TABLE 12.10	Summary of the equations for VOC blood (and plasma) to brain distributions and statistics for test sets of 99 solutes	171
TABLE 12.11	The inter-correlation of descriptors for 99 solutes using descriptors E, S, A, B and V	172
TABLE 12.12	19 common compounds, for plasma to brain and blood to brain distribution, which includes the statistics between the two columns of data	172

CHAPTER 13		*PAGE NO.*
TABLE 13.0	Air to kidney distribution for VOCs (see appendix)	176
TABLE 13.1	Summary of the equations for VOC air to kidney distributions for humans and rats	178
TABLE 13.2	The inter-correlation of descriptors for 76 solutes using descriptors E, S, A, B and L for air to kidney distribution model	179
TABLE 13.3	Blood to kidney distribution for VOCs (see appendix)	183

TABLE 13.4	Summary of the equations for VOC blood to kidney distributions for humans and rats	184
TABLE 13.5	The inter-correlation of descriptors for 73 solutes using descriptors E, S, A, B and V	185

CHAPTER 14		*PAGE NO.*
TABLE 14.0	Air to fat distribution for VOCs (see appendix)	188
TABLE 14.1	Summary of the equations for VOC air to fat distributions for rats and humans.	191
TABLE 14.2	The inter-correlation of descriptors for 129 solutes using descriptors E, S, A, B and L for air to fat distribution model	191
TABLE 14.3	Blood to fat distribution for VOCs (see appendix)	195
TABLE 14.4	Summary of the equations for VOC blood to fat distributions	196
TABLE 14.5	The inter-correlation of descriptors for 126 solutes using descriptors E, S, A, B and V	197

CHAPTER 15		*PAGE NO.*
TABLE 15.0	Data set used to obtain a model to correlate air to liver distributions for VOCs in humans and rats combined (see appendix)	201
TABLE 15.1	Summary of the equations for VOC air to liver distributions	204
TABLE 15.2	The inter-correlation of descriptors for 124 solutes using descriptors E, S, A, B and L for air to liver distribution model	204

TABLE 15.3	Data set used to obtain a model to correlate blood to liver distributions for VOCs in humans and rats combined (see appendix)	207
TABLE 15.4	Summary of the equations for VOC blood to liver distributions	209
TABLE 15.5	The inter-correlation of descriptors for 120 solutes using descriptors E, S, A, B and V	210

CHAPTER 16		*PAGE NO.*
TABLE 16.0	Data set used to obtain a model to correlate air to muscle distributions for VOCs in humans and rats combined (see appendix)	215
TABLE 16.1	Summary of the equations for VOC air to muscle distributions for humans and rats combined	218
TABLE 16.2	The inter-correlation of descriptors for 114 solutes using descriptors E, S, A, B and L for air to muscle distribution model	218
TABLE 16.3	Data set used to obtain a model to correlate air to liver distributions for VOCs in humans and rats combined (see appendix)	223
TABLE 16.4	Summary of the equations for VOC blood to muscle distributions for humans and rats combined	223
TABLE 16.5	The inter-correlation of descriptors for 110 solutes using descriptors E, S, A, B and V	224

CHAPTER 17		PAGE NO.
TABLE 17.0	Data set used in this model to obtain model to correlate air to urine distributions in humans for VOCs (see appendix)	230
TABLE 17.1	Equation for air to urine distribution for VOCS in humans	230
TABLE 17.2	Summary of the two equations for VOC air to urine distributions	231
TABLE 17.3	The inter-correlation of descriptors for 47 solutes using descriptors E, S, A, B and L.	232
TABLE 17.4	VOC air to water distributions for 408 solutes	232

CHAPTER 18		PAGE NO.
TABLE 18.0	Data set used to obtain a model to correlate air to olive oil distributions for VOCs (see appendix)	237
TABLE 18.1	Summary of the equations for VOC air to olive oil distributions	239
TABLE 18.2	The inter-correlation of descriptors for 213 solutes using descriptors E, S, A, B and L for air to oil distribution model	240

CHAPTER 19		PAGE NO.
TABLE 19.0	Data set used to obtain a model to correlate air to saline distributions for VOCs (see appendix)	244
TABLE 19.1	Summary of the equations for VOC air to saline distributions	245

TABLE 19.2	The inter-correlation of descriptors for 134 solutes using descriptors E, S, A, B and L for air to saline distribution model	246
TABLE 19.3	Data set used to obtain a model to correlate saline to olive oil distributions for VOCs (see appendix)	249
TABLE 19.4	Summary of the equations for VOC saline to olive oil distributions	249
TABLE 19.5	The inter-correlation of descriptors for 131 solutes using descriptors E, S, A, B and V	250

CHAPTER 20		***PAGE NO.***
TABLE 20.0	Summary of equations for air to tissue (or phase) distribution, and the tissue composition (lipid, protein and water) in humans	255
TABLE 20.1	Summary of equations for blood to tissue distribution, and the tissue composition (lipid, protein and water) for humans	263

CHAPTER 21		***PAGE NO.***
TABLE 21.0	Blood/plasma/serum to brain distribution for drugs only (see appendix)	270
TABLE 21.1	Correspondence between blood to brain, plasma to brain and serum to brain distribution coefficients, as log P values	282
TABLE 21.2	Summary of the *in vivo* equations for blood/serum/plasma to brain distributions for drugs only	285
TABLE 21.3	The inter-correlation of descriptors for 120 compounds	286

TABLE 21.4	Summary of the equations for blood/serum/plasma to brain distributions for VOCs and drugs combined.	290
TABLE 21.5	The inter-correlation of descriptors for 120 compounds	291

CHAPTER 22		*PAGE NO.*
TABLE 22.0	Drug data set used in this model to obtain model to predict blood (or plasma) to fat distributions in rats (see appendix)	298
TABLE 22.1	Summary of the two equations for blood/plasma to fat distributions for VOCs and drugs combined.	299
TABLE 22.2	The inter-correlation of descriptors for 141 (142 data points) compounds	301

CHAPTER 23		*PAGE NO.*
TABLE 23.0	Drug data set used in this model to obtain model to predict blood (or plasma) to heart distributions in rats (see appendix)	306
TABLE 23.1	Summary of the two equations for blood/plasma to heart distributions for drugs only	307
TABLE 23.2	The inter-correlation of descriptors for 51 compounds	308
TABLE 23.3	Summary of the two equations for blood/plasma to heart distributions for VOCs and drugs combined	312
TABLE 23.4	The inter-correlation of descriptors for 74 compounds	313

CHAPTER 24		PAGE NO.
TABLE 24.0	Drug data set used in this model to obtain a model to predict blood (plasma or serum) to lung distributions in rats (see appendix)	320
TABLE 24.1	Summary of the two equations for blood/plasma/serum to lung distributions for drugs only	321
TABLE 24.2	The inter-correlation of descriptors for 50 compounds	322
TABLE 24.3	Summary of the two equations for blood/plasma/serum to lung distributions for VOCs and drugs combined	326
TABLE 24.4	The inter-correlation of descriptors for 91 compounds	327

CHAPTER 25		PAGE NO.
TABLE 25.0	Drug data set used in this model to obtain a model to predict blood (or plasma) to muscle distributions in rats (see appendix)	333
TABLE 25.1	Summary of the two equations for blood/plasma to muscle distributions for drugs only	334
TABLE 25.2	The inter-correlation of descriptors for 59 compounds	335
TABLE 25.3	Summary of the two equations for blood/plasma to muscle distributions for VOCs and drugs combined	339
TABLE 25.4	The inter-correlation of descriptors for 163 compounds	340

CHAPTER 26		PAGE NO.
TABLE 26.0	Drug data set used in this model to obtain a model to predict blood (plasma or serum) to liver distributions in rats (see appendix)	346
TABLE 26.1	Summary of the two equations for blood/plasma/serum to liver distributions for drugs only	347

TABLE 26.2	The inter-correlation of descriptors for 52 compounds	348
TABLE 26.3	Summary of the two equations for blood/plasma/serum to liver distributions for VOCs and drugs combined	352
TABLE 26.4	The inter-correlation of descriptors for 169 compounds	353

CHAPTER 27		*PAGE NO.*
TABLE 27.0	Data set used in this model to obtain model to predict drug blood (and plasma) to skin distributions in rats (see appendix)	359
TABLE 27.1	Summary of the two equations for drug blood/plasma skin distributions	360
TABLE 27.2	The inter-correlation of descriptors for 35 compounds	361

CHAPTER 28		*PAGE NO.*
TABLE 28.0	Drug data set used in this model to obtain a model to predict blood (plasma or serum) to kidney distributions in rats (see appendix)	365
TABLE 28.1	Summary of the two equations for blood/plasma/serum to kidney distributions for drugs only	366
TABLE 28.2	The inter-correlation of descriptors for 50 compounds	367
TABLE 28.3	Summary of the two equations for blood/plasma/serum to kidney distributions for VOCs and drugs combined	371
TABLE 28.4	The inter-correlation of descriptors for 120 compounds	372

CHAPTER 29		PAGE NO.
TABLE 29.0	Summary of equations for blood to tissue distribution (drugs only), and the tissue composition (lipid, protein and water) in humans	378
TABLE 29.1	Summary of equations for blood to tissue distribution (VOCs and drugs combined), and the tissue composition (lipid, protein and water) for humans	379

APPENDIX		PAGE NO.
TABLE 9.0	Air to blood distribution for VOCs	395
TABLE 9.1	Air to tissue (brain, fat, kidney, liver, lung, muscle and plasma) distribution for VOCs	406
TABLE 10.0	Air to blood distribution for VOCs	412
TABLE 11.0	Data set used to obtain a model to correlate air to plasma distributions for VOCs in humans and rats combined	418
TABLE 11.5	Data set used in this model to obtain model to predict blood to plasma distributions for VOCs	420
TABLE 12.0	Air to brain distribution for VOCs	422
TABLE 12.3	Blood to brain distribution for VOCs	425
TABLE 12.6	Plasma to brain distribution for VOCs	428
TABLE 12.9	Plasma to brain and blood to brain distribution combined for VOCs	429
TABLE 13.0	Air to kidney distribution for VOCs	433
TABLE 13.3	Blood to kidney distribution for VOCs	436
TABLE 14.0	Air to fat distribution for VOCs	439
TABLE 14.3	Blood to fat distribution for VOCs	445
TABLE 15.0	Data set used to obtain a model to correlate air to liver distributions for VOCs in humans and rats combined	450
TABLE 15.3	Data set used to obtain a model to correlate blood to liver distributions for VOCs in humans and rats combined	457

TABLE 16.0	Data set used to obtain a model to correlate air to muscle distributions for VOCs in humans and rats combined.	462
TABLE 16.3	Data set used to obtain a model to correlate air to liver distributions for VOCs in humans and rats combined	468
TABLE 17.0	Data set used in this model to obtain model to correlate air to urine distributions in humans for VOCs	473
TABLE 18.0	Data set used to obtain a model to correlate air to olive oil distributions for VOCs	475
TABLE 19.0	Data set used to obtain a model to correlate air to saline distributions for VOCs	486
TABLE 19.3	Data set used to obtain a model to correlate saline to olive oil distributions for VOCs	491
TABLE 21.0	Blood/plasma/serum to brain distribution for drugs only	495
TABLE 22.0	Drug data set used in this model to obtain model to predict blood (or plasma) to fat distributions in rats	503
TABLE 23.0	Drug data set used in this model to obtain model to predict blood (or plasma) to heart distributions in rats	504
TABLE 24.0	Drug data set used in this model to obtain a model to predict blood (plasma or serum) to lung distributions in rats	506
TABLE 25.0	Drug data set used in this model to obtain a model to predict blood (or plasma) to muscle distributions in rats	508
TABLE 26.0	Drug data set used in this model to obtain a model to predict blood (plasma or serum) to liver distributions in rats	511
TABLE 27.0	Data set used in this model to obtain model to predict drug blood (& plasma) to skin distributions in rats	513
TABLE 28.0	Drug data set used in this model to obtain a model to predict blood (plasma or serum) to kidney distributions in rats	515

Abstract

Distribution coefficients, K_{blood} and $K_{tissues\ (or\ biophase)}$, from the gas phase to blood and gas phase to tissues (plasma, brain, fat, heart, liver, lung, kidney, muscle, urine, saline and olive oil) have been collected for large number of volatile organic compounds (VOCs). For these datasets of VOCs, linear free energy relationships (LFERs) have been established and Abraham equations have successfully been constructed to predict these distributions. It has also been shown that human and rat data for the air to blood and air to tissue distribution of VOCs can be combined. The differences in the two data sets, for the common compounds are smaller than the estimated inter-laboratory experimental error. The combination of the log K_{tissue} values with values for air to blood yields distribution coefficients from blood to tissue, as log P_{tissue}. Equations have successfully been constructed to predict these distributions.

From a large amount of collected data on the distribution of drugs from blood, plasma or serum to tissue (brain, fat, heart, kidney, lung, liver, muscle and skin) at steady-state concentration in rats, it is shown that the three datasets of data can be combined. Predictive LFER equations for blood/plasma/serum to tissue for a large number of drugs, has been achieved and their predictive capability has been assessed. Finally, it has been shown that the *in vitro* data on VOCs and the *in vivo* data on drugs can be combined; LFERs on the total data yield correlative and predictive Abraham equations. Because the descriptors used in the LFERs can be calculated from structure, distribution coefficients for air to blood or tissue for VOCs and for blood/plasma/serum to tissue for VOCs and drugs can be predicted directly from the molecular structures of the VOCs and drugs.

Acknowledgements

My sincere thanks go to Professor Michael Abraham for his help over the last four years of this project. Especially for his support, enthusiasm, his wealth of knowledge and expertise in the area of chemistry, that has helped me to accomplish my Ph.D. I am extremely grateful for his unfailing encouragement, patience and guidance.

Thank you to Professor William Acree, Jr (University of Texas) for working in collaboration with us on the air to blood and blood to brain distribution model. I would like to thank Dr Yuan Zhao with his commitment, help and the continuous support on the drug blood to brain project. Special thanks also go to Dr Chris Luscome at GSK. Miss Elizabeth Draper has been very supportive and encouraging as a friend, during the build up to the completion of this thesis.

Special thanks to EPSRC for their financial support during part of the time of this project. My gratitude goes out to Dr Andreas Zissimos for his support in the first two years of my Ph.D., with the development of a computer program called ADEMA and PharmaAlgorithm (Absolv).

I am very grateful to other research group members, both past and present, Dr Joelle Gola, Dr Andreas H, Dr Kei Enomoto, Dr Joelle Le, Dr Mayte Gil, Mr Ricardo Sanchez, Mr Javier Gil, Mr Jorge Sánchez, Mr Vikas Gupta and Mr Mital Amin, and also to Dr Derek Reynolds and Dr Eric Clark.

I would like to thank close family and friends that include: Mr Alnoor Kara, Mrs Asma Kara, Mr Jairam Heer, Mrs Mary Jabore, Mrs Julie Davies, Mr Tony Cheung, Mr Deepak Datta, Mr Menesh Patel, Mrs Kulsum Musa, Miss Raissa Ibrahim, Mr Shabaz Ibrahim, Miss Shaheen Ibrahim, Dr Samudra Sarka, Dr Elbertus Kruiswijk, Mrs Jill Maxwell and others for their tireless understanding and for always being there when I needed them – Thank you!

Finally, my heart goes out to my special wife (Mrs Nazneen Ibrahim) for her constant love and support during my Ph.D. Special thanks to my children (Amber and Danyaal) who have helped me to keep my spirit high and aim towards to finishing this Ph.D. Warm thanks, goes to my father (Mr Ismail Kara) and mother for their dedication in encouraging me in my academic studies.

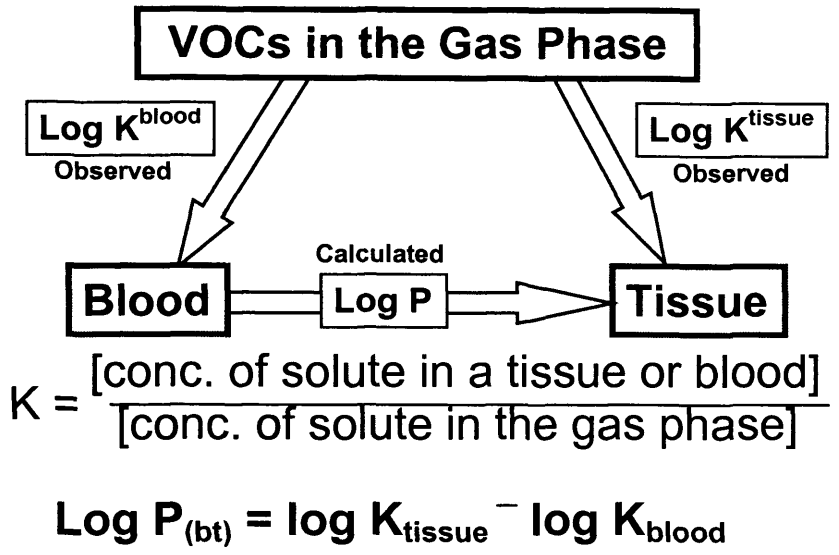
Chapter 1

Introduction to the blood tissue distribution of volatile organic compounds (VOCs)

The solubility of gases and vapours in biological fluids (e.g. blood, plasma, serum and urine) and tissues (e.g. brain, muscle, lung, liver, kidney, heart, skin and fat) is of importance in areas such as anaestheology and environmental health especially with regard to indoor and outdoor air pollution.

The purpose of this work is to construct predictive equations for VOC distribution from air to tissue and from blood to tissue. This data will than be combined with drug data to give new Abraham equations for distribution of compounds between blood and various biological tissues.[1] An extensive literature search has been be carried out in order to obtain the data that is needed for modelling of air to blood and blood to tissue distribution for various types of VOCs. The large amounts of data obtained from the literature are for compounds such as alcohols (e.g. 1-hexanol),[2] anaesthetics (e.g. halothane),[1] agrochemicals (e.g. hexachlorobenzene),[3] saturated alkanes (e.g. nonane),[4] unsaturated alkenes (e.g. propylene)[5] and many substituted cyclic alkanes (e.g. methylcyclohexane)[6] and aromatic compounds (e.g. pentaflurobenzene).[2]

The data obtained for VOCs is expressed as the Ostwald solubility coefficient (gas-liquid partition), K, for solutes in a biological phase (e.g. blood) at 310 K. This distribution data is obtained using analytical methods known as head-space chromatography.[3] Scheme 1.0 outlines the importance of obtaining the direct Log K values from the literature to obtain air to blood and air to tissue distribution data. From this direct method, the log K tissue and log K blood values, lead to an indirect method for obtaining solute blood-tissue distribution data expressed as log P.



Scheme 1.0. Determination of blood-tissue distribution for VOCs[1]

Data for compounds that have been measured several times will then be combined in order to obtain an average log K value for any given set of compounds for air/blood and air/tissue distribution of VOCs. Human and rat log K data for VOCs shall than be compared for common compounds for air/blood and air/tissue, to see if there are similarities or differences between the two species.

Models will be made based on the Abraham equations for air/blood distribution as well as air/tissue distributions. Finally the difference in the two will give rise to log P, for distribution of compounds between blood and various biological tissues. Later, this data shall be gathered together with drug data to give rise to overall equations that can be used to predict log P values.

References

1. Abraham, M. H. and Weathersby, P. K (1994). J. Pharm. Sci., **83**, 1450-1455.
2. Meulenberg, C. J. W. and Vijverberg, H. P. M (2000). Toxicology and Applied Pharmacology, **165**, 206-216.
3. Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H. and Anderson, M. E (1989). Toxicology and Applied Pharmacology, **98**, 87-99.
4. Zahlsen, K. and Eide, I (1996). Arch. Toxicol, **70**, 397-404.
5. Filser, J. G., Schmidbauer, R., Rampf, F., Baur, C. M., Putz, C. and Csanady, G. A (2000). Toxicology and Applied Pharmacology, **169**, 40-51.
6. Imbriani, M., Ghittori, S., Pezzagno, G. and Capodagio, E (1985). G. Ital. Med. Lav., **7**, 133-140.

Chapter 2

Volatile organic compounds (VOCs) in general and their environmental importance

There is a perception that many individuals believe that the risk associated with out door air pollution appears to be higher than for indoor air pollution. However, exposure to indoor air pollutants like volatile organic compounds (VOCs) appears to be a serious public health problem in a wide variety of non industrial settings, for example, residences, offices, schools and vehicles.[1] Recent research conducted in the US and Europe shows that individual persons in industrialised nations spend 90% or more of their time indoors.[2] The young, elderly and those with disability problems or suffering from diseases are most likely to spend their time at home. The experience of an individual who is exposed to VOCs is dependent upon the function of time and concentration involved.[3]

Although a large deal of time and money has been spent on the interest of environmental pollution (i.e. the effects of outdoor air notably from traffic pollution), recently scientific interest on indoor pollution has been taken very seriously. The understanding on how VOCs affects human beings (and animals) is of prime importance in order to improve the quality of indoor air and reduce illnesses (and discomfort). Recent analysis of common 30 VOCs in Germany (belonging to the groups of alkanes, cycloalkanes, aromatics, volatile halogenated hydrocarbons and terpenes) have shown that the concentrations of VOCs in indoor air is, on average, 10 times higher than outdoors.[4] Attempts to reduce the exposure to VOCs have been proposed with strict guidelines on the standard levels that should be acceptable.

VOCs have a wide range of physical and chemical characteristics. The

chemical groups typically include hydrocarbons, halogenated hydrocarbons, aromatic hydrocarbons, alcohols, esters, ketones and aldehydes. These compounds have a wide range of boiling points and have been classed as very volatile (<0 to 50-100°C), volatile (50-100 to 240-260°C) and semi-volatile (250-260 to 400°C).[3] According to a definition given by the European Communities the expression "volatile organic compounds" means any compound having at 293.15 K a vapour pressure of 0.01 kPa or more, or having a corresponding volatility under the particular condition of use.[5]

VOCs contribute to a broad scale of chemicals with production levels all over the world and widespread applications in industry, trade and private households. One of the most important industrial uses of the VOCs is their supply as solvents. In Table 2.0, sources and tonnage of VOCs found in air of the United Kingdom, in 1995, are displayed. The identified numbers of VOCs present in ambient air has increased steadily in recent years, from 250 to more than 900 in 1989 and now currently over 1000.[6] VOCs have been of increasing concern since the 1970s because of their potential to cause health effects similar to those reported in the sick building syndrome and to contribute to respiratory problems and other diseases including cancer.[3]

Source category	Estimated Emissions (thousand of tonnes)	% Total
Power station	5	-
Domestic	30	1.3
Commercial/public service	2	-
Refineries	2	-
Iron and steel	4	-
Other industrial combustion	9	0.4
Non-combustion sources	335	14.3
Extraction and distillation of fossil fuel	334	14.3
Solvent use	700	29.9
Road transport	690	29.5
Off road sources	96	4.1
Military	1	-

Railways	8	0.3
Civil aircraft	4	-
Shipping	12	0.5
Waste treatment	26	1.1
Forest	80	3.4
Total	**2338**	**100**

Table 2.0. Sources of VOCs, excluding methane, in the outdoor air of UK during 1995.[3]

The sources of indoor air pollutant

There are number sources of indoor air pollution at home. This includes: combustion sources, such as oil, gas, kerosene, coal, wood, and tobacco products (e.g. benzene), a wide variety of building materials (e.g. toluene, xylenes, decane), cleaners, office products and machines, paints, and furnishings. Others come from bathing (e.g. chloroform is formed when chlorine interacts with organic substances found in hot water), cooking, cosmetics, hygienic products, plants, as well as human biological processes give rise to indoor air pollution.[6] Table 2.1 gives examples of sources of VOCs known to contribute to indoor air pollution.[3]

Outdoor air may also contribute to indoor air contamination, particularly when air intakes are positioned near parking areas, roads, or other locations where contaminated air can be entered into the buildings. Example, car exhaust gases and particles can gain access into the homes, especially if windows remain open.[1]

Volatile organic compounds (VOCs) pose an air pollution hazard because of their potential for generating ozone. Some VOCs such as benzene, 1,3-butadiene, polycyclic aromatic hydrocarbons (PAH) and dioxins are hazardous in themselves to human health. Benzene comes mainly from petrol combustion, but some from cigarette smoke. There is an increased risk of leukaemia from occupational exposure but no evidence of risk below 25-30 ppm. 1, 3-butadiene comes mainly from vehicle exhausts, but cigarette smoke is a major source indoors. There is an increased risk of leukaemia and lymphoma from occupational exposure in the rubber production industry (around 1-10 ppm). Finally polycyclic aromatic hydrocarbons (PAHs)

come mainly from incomplete combustion of vehicle emissions in urban areas and coal and coke burning elsewhere. Total PAH levels are around 10-150 ng/m^3 in the UK. There is an increased risk of lung cancer from occupational exposure in coke-oven and coal gas workers (around 30 μ g/m^3).[7]

VOCs	Source
p-Dichlorobenzene	Moth crystals and room deodorants.
Styrene	Insulation, textiles, disinfectants, plastics and paints.
Benzyl chloride	Vinyl tiles.
Benzene	Smoking.
Tetrachloroethylene	Dry cleaned clothes.
Chloroform	Chlorinated water.
1,1,1-Trichloroethane	Dry cleaned clothes, aerosol sprays and fabric protectors.
Carbon tetrachloride	Industrial strength cleaners.
Aromatic hydrocarbons (toluene, xylenes, ethylbenzene, trimethylbenzenes)	Paints, adhesives, gasoline and combustion products.
Aliphatic hydrocarbons	Paints, adhesives, gasoline and combustion products.
Terpenes	Scented deodorisers, polishes, fabrics, fabric softeners, cigarettes and food.
Polyaromatic hydrocarbons (PAHs)	Combustion products (smoking, wood-burning, kerosene heaters).
Alcohols	Aerosols, window-cleaners, paints, paint thinning, cosmetics and adhesives.
Ketones	Lacquers, varnishes, polish removers and adhesives.
Ethers	Resin, paint, varnishes, lacquers, dyes, soaps and cosmetics.
Esters	Plastics, resins, plasticizers, lacquer solvents, flavours and perfumes.

Table 2.1. Sources of indoor VOCs.[3]

Environmental tobacco smoke (ETS) is a source of particular concern because of the displeasure and irritation it can cause to many users of multi-occupied buildings and the risks of disease to those exposed long term smoke. Its complexity can be illustrated by the fact that some 4000 various components have been identified from tobacco smoke alone. The main VOCs released in side stream smoke in quantities exceeding 1 mg per cigarette are nicotine, acetaldehyde, acrolein, isoprene, and acetonitrile.[16] Very few of the constituents are unique to ETS and this has caused difficulties in apportioning the contribution that ETS makes to concentrations of particular pollutants within buildings.[3]

There are three fundamental processes control the rate of VOC emissions from building sources: (1) evaporation; (2) desorption of absorbed compounds and (3) diffusion of a material. How fast the VOCs are produced depends on the process and the source characteristics. The sources can be divided into those with continuous emissions, and discontinuous emissions.[3, 8] Some sources, such as building materials and furnishings release pollutants more or less continuously. The nature of emission and the variability of indoor spaces and ventilation conditions result in a dynamic behaviour of air pollutants in indoor environment.[9]

Measurement of volatile organic compounds in indoor air

A large range of sampling and analytical methods can be applied to determine the nature and concentration of many VOCs.[10-12] The most common methods are based on collection using absorbents, e.g. Tenax TA, graphite carbon black and molecular carbon sieves contained in a sampling tubes or badge. These are used in multiple bed form to match targeted samples. The absorbent can be thermally desorbed and the VOCs can then be determined by gas chromatography coupled with various detection systems like mass spectrometry. Other absorbents are available and are chosen according to the polarity of the VOCs investigated.

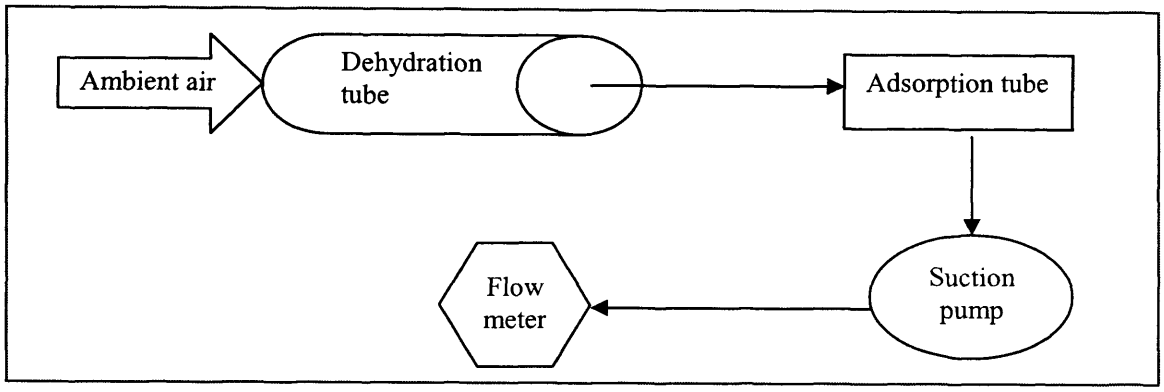


Figure 2.0. Sampling with tube collection method.[10-12]

Biological monitoring

The definition of biological monitoring (BM) by the International Union of Pure and Applied Chemistry[13] (IUPAC), is a 'systematic continuous or repeated measurement and assessment of workplace agents or their metabolites either in tissues secreta, excreta or any combination of these to evaluate exposure or health risk compared to appropriate reference'. Basically, BM allows one to assess the integrated exposure by different routes, including ingestion, inhalation, dermal absorption, blood, exhaled air and urine. The biological marker or biomarkers for exposure is an endogeneous substance or its metabolite or the product of interaction between a xenobiotic and some target molecule or cell that is measured in a compartment with an organism.[13] The purpose of biomarkers are mainly focused on the amount of the pollutant penetrating into the organism. There are several biomarkers that are relevant to indoor air pollution: e.g. urinary excreted nicotine is used for exposure to ETS, the carboxyhaemoglobin level in blood is used to characterise exposure to CO, and the presence of VOCs in exhaled air breath is used to mark these compounds.[14] Examples of the use of BM are available in the literature. Imbriani and co-workers[15] developed a method for the BM of exposure to enflurane in operating room personnel, based on the measurement of the unchanged anaesthetic in urine. Jo and Pack[16] employed a breath analysis for exposure to benzene associated with active smoking. Mathews et al.[17] studied endogenous

VOCs found in breath. Biological monitoring has extensively been applied in practical occupational medicine found in many developed countries. From this, as a result the biological threshold values have been evaluated in order to control the worker's exposure.

The amount of indoor air pollutants

The amount of indoor air pollutants depends on many factors. The amount (levels) of indoor air pollutant are also dependent on human activity.[18] Over recent years, the combination of reduced ventilation rates, warmer (temperature) and more humid conditions indoors, together with the greater use and diversity of materials, furnishing and consumer products, has resulted in accumulation of a wide range of pollutants occurring indoors at levels often exceeding those outdoors.[1,3] The age of the buildings plays an important factor in concentration of pollutants. The US EPA studies have shown that a new building indicated to have eight of thirty-two target chemicals measured within days after completion of the building, was found to have 100 fold higher levels compared to outdoor levels: xylenes, ethylbenzene, ethyltoluene, trimethylbenzene, decane and undecane.[19] There are many VOCs that are commonly found as mixtures, with a mean concentration below 50 microgram per cubic meter ($\mu g.m^{-3}$) in established buildings, but much higher in new buildings.[3]

Recently the EPA conducted a study on indoor and outdoor air concentrations of twenty-seven hazardous air pollutants.[20] Indoor and outdoor exposures (i.e. concentrations breathed multiplied by duration of time breathed) was estimated from data of literature. The results that have emerged from this study showed that the indoor air concentration of these hazardous compounds are generally one to five times more than the outdoor concentration, and indoor exposures are ten to fifty times more than the outdoor exposures. The summary of results is shown in Table 2.2.[20]

SOLUTE (VOCs)	Typical concentration [a]			Typical Daily Exposure [b]		
NAME	Indoor	Outdoor	I/O Ratio	Indoor	Outdoor	I/O Ratio
Acetaldehyde [c]	<10	<3	*[5]*	<216	<7.2	*[50]*
Benzene [c]	5	5	*[2]*	108	12	*[20]*
1,2,3,6-Tetrahydro-N(trichloromethylthio)-phthalimide	<0.001	Na	10	<0.02	<0.0002	~100
Carbon tetrachloride[c]	<5	1	*[2]*	<108	2.4	*[20]*
1,2,4,5,6,7,8,8-Octachloro-3a,4,7,7a-tetrahydro-4,7-endo-methanoindan	0.2	0.01	20	4.32	0.024	200
Chloroform[c]	1	0.2	*[5]*	21.6	0.48	*[50]*
Cumene	1	0.2	5	21.6	0.48	50
1,1-dichloro-2,2-bis(4-chlorophenyl)ethylene	0.0005	<0.002	>0.2	0.0108	<0.005	>2
1,1-dichloro-2-dimethoxyphosphoryloxyethene	0.05	0.001	50	1.08	0.0024	400
Ethylbenzene	5	1	*[3]*	108	2.4	*[30]*
Formaldehyde	50	4	10	1080	9.6	100
Formaldehyde[c]	0.1	0.002	50	2.16	0.0048	400
Hexachlorobenzene	0.0005	0.0001	5	0.0108	0.00024	50
Hexane	5	4	*[10]*	108	9.6	*[90]*
2,2-bis(p-methoxyphenyl)-1,1,1-trichloroethane	0.0001	0.00003	3	0.00216	0.000072	30
Methyl Ethyl Ketone	10	<1	*[4]*	216	<2.4	*[40]*
Methylene chloride	10	1	*[10]*	216	2.4	*[90]*
Naphthalene	1	1	*[4]*	21	2.4	*[40]*
Paradichlorobenzene	1	<0.05	*[5]*	21.6	<0.12	*[50]*
2-(1-methylethoxy)-phenylmethylcarbamate	0.1	0.01	10	2.16	0.024	90
Styrene [c]	2	0.6	*[5]*	43.2	1.44	*[50]*
Tetrachloroethylene[c]	5	2	*[3]*	108	4.8	*[30]*
Toluene	20	5	*[5]*	432	12	*[50]*
Trichloroethylene[c]	5	0.5	*[5]*	108	1.2	*[50]*
Xylenes (o+m+p)	15	10	*[2]*	324	24	*[20]*

[a] Indoor and outdoor concentrations in $\mu g.m^{-3}$. Typical values from the literature for U.S. locations. I/O ratios based on typical concentrations *[or reported ratios as indicated by values in italics and brackets]*.

[b] Exposure, in $\mu g.m^{-3}.h$, based on assumption that the typical person spends 90% of the typical day indoors.

[c] Urban air toxic substances, [d] Radionuclides/ Rn in $pCi.dm^{-3}$

Table 2.2. Summary of indoor and outdoor air concentrations and exposures for common air pollutants.[20]

The health effects of VOCs, both indoors and outdoors

The health effects from indoor air pollutants (IAP) may be experienced soon after exposure or possibly at a later time (years) to come. The most frequently reported health complaints include eye, nose, and throat irritation, headache, nausea, hoarseness, sore throat, cough, chest tightness, nasal congestion, palpitations, shortness of breath, stress, drowsiness, and alterations in mood found from single/ or continuous exposure of indoor air pollutants (includes VOCs).

Most of the immediate effects are usually short term and are treatable. The likelihood of immediate reactions to IAP depends on several factors. Age and any pre-existing medical conditions are two important influences. The effects on whether a person reacts to a pollutant depend on individual sensitivity, which varies from one individual to the next. Secondly, other health effects may show up only after long or repeated periods of exposure. These effects which include some respiratory diseases, heart disease, and cancer, can be more fatal. [1, 3, 21-23]

There are three distinct classes of insidious poison, all of which display what can be termed as "delayed effects." They are carcinogens, mutagens, and teratogens. Chemical carcinogens are substances which induce cancer in humans and animals. Vinyl chloride has been identified as a causative agent in cancer of the liver. Mutagens are substances which are capable of altering the body's genetic material i.e. the genes and chromosomes of the cells. Ethylene-oxide, a non-carcinogen, has been found to have mutagenic properties. This mono-epoxide is used in the manufacture of numerous non-ionic surfactants. Examples of teratogens are phthalate-esters, frequently used as plasticizers in coating formulations. These agents can cause congenital defects in a developing embryo (or fetus).

Toxic hazards can impact the human body either by affecting the respiratory system, the eyes and/or the skin; the respiratory system being the most significant route for poisonings as far as the paint industry is concerned.[24] Substances used in the formulations of coatings, which can be inhaled, are primarily the vapours of volatile liquids and solids. Toxic vapours can be irritants, asphyxiants, anaesthetics and systemic poisons. Volatile irritants are capable of preferentially effecting specific areas of the respiratory system, such as the respiratory tract, the lung tissue of the terminal air passages and air sacs.

The most common eye disorder experienced in the handling of paint is undoubtedly eye irritation from volatile paint solvents. Sometimes, however, the irritation is primarily due to the presence of minor amounts of volatile chemicals, such as un-reacted monomers (e.g. acrylates) or pungent additives (e.g. allyl compounds). Formaldehyde is a very common indoor air contaminant. It is found in many products and materials; it has many uses. For instance, formaldehyde binds wood chips in particleboard, it is a solvent in dyes for cloth or paper, and it is used in wrinkle-resistant material and as a water-repellent in floor coverings. Among VOCs and their health effects, most is known concerning exposure to formaldehyde. Formaldehyde may cause a burning or tingling sensation in the eyes, nose and throat that goes away when the exposure is removed. It may trigger asthmatic symptoms in susceptible infants and children. Formaldehyde is the agent alleged to be responsible for blindness in methylated spirit drinkers. Methyl alcohol (methanol) taken internally is oxidized to formaldehyde and routed via the bloodstream to the optic nerve, which is ultimately destroyed. In table 2.3 is a summary of other selected VOCs and their various impacts on human beings.[21, 24]

VOCs	Health Effects
Benzene	Carcinogen; respiratory tract irritant
Xylenes	Narcotic; irritant; affects heart, liver, kidney, and nervous system
Toluene	Narcotic; possible cause of anemia
Styrene	Narcotic; affects control of nervous system; probable human carcinogen
Toluene di-isocyanate	Sensitizer; probable human carcinogen
Trichloroethane	Affects central nervous system
Ethyl benzene	Severe irritation of eyes and respiratory tract; affects central nervous system

Dichloromethane	Narcotic; affects control of nervous system; probable human carcinogen
1,4-Dichlorobenzene	Narcotic; eyes and respiratory tract irritant; affects heart, liver, kidney, and nervous system
Benzyl chloride	Central nervous system irritant depressant, affects liver and kidney; eye and respiratory tract irritant
2-Butanone	Irritant; central nervous system depressant
Petroleum distillate	Affects central nervous system, liver and kidneys
4-Phenycyclohexene	Eye and respiratory tract irritant; central nervous system effects

Sources adapted from ref. US EPA 'introduction to IAQ' report no. EPA/400/3-91/003, Washington, DC, 1991.

Table 2.3. Health effects of selected volatile organic compounds[21, 24]

VOCs that acts as anaesthetic vapours

The solubility of an inhalation anaesthetic determines how fast it works and how fast it wears off. The more insoluble a drug (i.e. VOCs) is in blood, the quicker it works. If one breathes in a highly soluble anaesthetic vapour, such as ether, as soon as it arrives in the alveoli it dissolves into the bloodstream and ends up in muscle, fat, bone – everywhere, and so, very little ether arrives in the brain. Consequently it takes "all day" to go to sleep and wake up. Solubility is related to, and described by, the blood-gas partition coefficient (BGPC). The amount of solute (VOCs) that dissolves in a biological fluid is related to the Ostwald solubility coefficient (K or L) for that substance, and, for blood it is known as the blood-gas partition coefficient.

The oil-gas partition coefficient

Many volatile agents are highly lipophilic, and there is a lot of lipid in cell membranes. Meyer and Overton hypothesised, at the beginning of the century, that anaesthetic agents acted at the level of the lipid membrane and, consequently potency is related to lipophilicity. Halothane is highly lipid soluble, isoflurane less so, and nitrous oxide is comparatively insoluble in fat. Lipid solubility is represented by the oil-gas partition coefficient (OGPC): again, the larger the number,

the higher the solubility. An ideal anaesthetic would act extremely quickly (be highly insoluble in blood: have a low blood gas partition coefficient), and achieve anaesthesia at very low concentrations (be very potent-highly soluble in fat [brain]: have a high oil gas partition coefficient.[25])

What determines how quickly an anaesthetic wears off?

The blood-gas partition coefficient (BGPC) is important, as once the drug is in the blood, the lower the BGPC, the quicker the drug is excreted through the lungs. However, every tissue in the body has its own blood-tissue partition coefficient and as the tissues act as a reservoir for anaesthetic gas, the rate at which the anaesthetic diffuses back from the tissues ultimately determines how quickly the patient wakes up.

TISSUE	ISOFLURANE	SEVOFLURANE	HALOTHANE
FAT	44.90	47.50	51.10
MUSCLE	2.92	3.13	3.38

Table 2.4. Blood-tissue partition coefficients (P).[26]

The combination of the blood-tissue and gas-blood partition coefficients determines how quickly the patient wakes up. One can imagine that a patient wakes up very quickly from isoflurane, and quite slowly from halothane. Sevoflurane is quicker than halothane because of its much lower blood-tissue partition coefficient. Finally, volatile agents dissolve much better in fat than in any other tissue. The more fat, the longer the wash out time: obese patients wake up much more slowly.[26]

References

1 P.T.C. Harrison, Chemistry and Industry (1997). 677.

2. J. Robbinson, W.C. Nelson, National Human Activity Pattern Survey Database, EPA/600/R-96/074, USEPA, Research Triangle Park, NC, (1995).

3. L. Sheldon, H. Zeton, J. Sickes, C. Eaton, T. Hartwell, 'Indoor air quality in public buildings, vol.II', Research Triangle NC: EPA/600/6-88/009b, 1988b.

4. Rehwagen, M., Schlink, U. and Herbarth, O (2003). Indoor Air, **13**, 283-291.

5. Commission of the European communities, proposal for a council directive on limitation of emissions of volatile organic compounds due to the use of organic solvents in certain industrial activities (1996). 96/0276 (SYN) Article 2.

6. W.E. Lambert, J.M. Samet, 'occupational and environmental disease', Eds. Harber, P., Schenker, M. B., Balmes, J. R, Mosby, St Louis, (1996) 784.

7. Department of health. Committee on medical effects of air pollutants. Handbook on Air Pollution and Health. London (1997).

8. World Health Organisation (WHO) 'air quality guidelines', WOH regional publication Europe, Copenhagen, (1999).

9. Berglund, B. and Johansson, I (1987). Proceeding of the 4th international conference on indoor air quality and climate, **1**, 16.

10. Wallace, L (1987). Total exposure assessment methodology (TEAM) study. Summary and Analysis', vol.1, US EPA, (1987).

11. Ozkaynak, H (1999). Air Pollution and Health, Eds. Holgate, S. T., Samet, J. W., Koren, H. S. and Maynard, R. L (1999) Academic Press, 148.

12. Clement, R. E., Yang, P. W. Koester, C. J (1999) Anal. Chem., 71, 257-R.

13. Dufus, J. H (1993) Pure Appl. Chem, **65**, 2003.

14. Folinsbee, L. J., Kim, C. S., Kerl, H. J., Prah, J. D. and Devlin, R. B. Handbook of Toxicology.

15. Imbriani, M., Ghittori, S., Pezzagno, G., Capodaglio, E (1994). Arch. Environ. Health, **49**, 135.

16. Jo, W. J. and Pack, K. W (2000) Environ. Res. Sec. A, **83**, 180.

17. Mathews, J. M., Raymer, J. H., Etheridge, A. S., Velez, G. R. and Bucher, J. R (1997). Toxicol. App. Pharmacol., **146**, 255.

18. Rushton, L., Cameron, K (1999) Air pollution and health, eds. Holgate, S. T.,

Samet, J. M., Koren, H. S. and Maynard, R. L. Acad. press, London, 813.

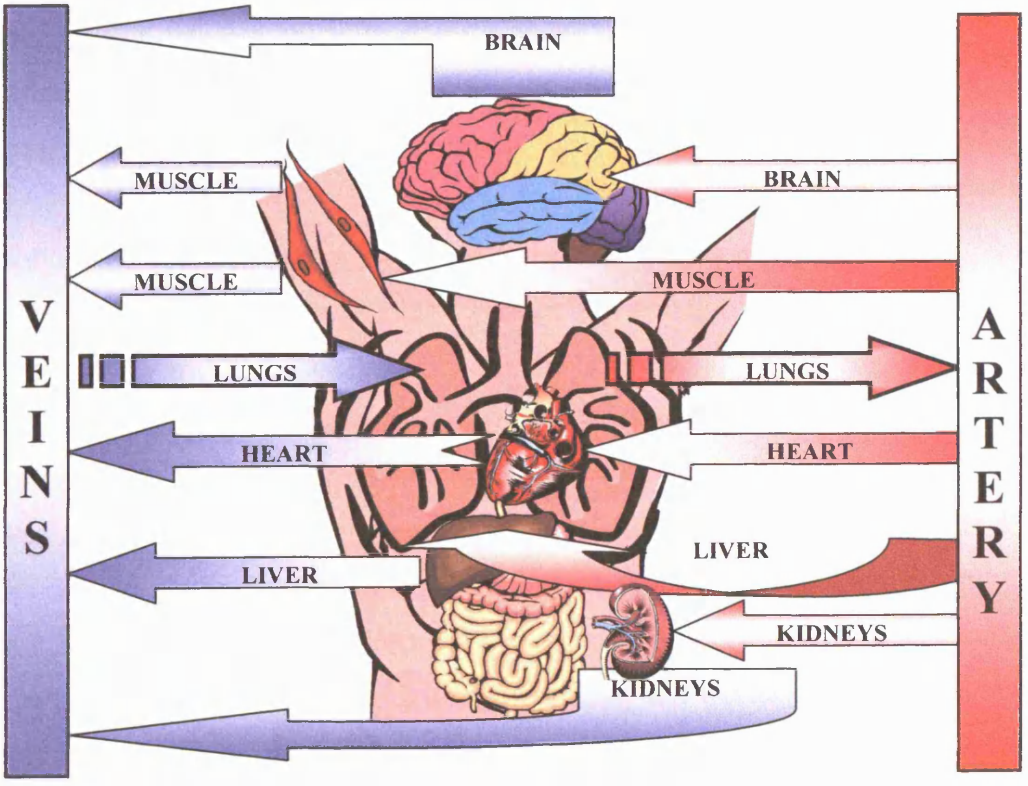
19. US EPA, indoor air quality in public building (1998). Vol II EPA600/6-88/009b.
20. US EPA, inside IAQ (1998). EPA/600/N-98/002, 1.
21. Sexton, K, Dyer, R. S (1996). Indoor air quality and human health, 2nd edition, Eds. Gammage, R. B. and Berven, B. A. CRC Press, Boca Raton, 1.
22. Rose, C (1996). Indoor air quality and human health, 2nd edition, Eds. Gammage, R. B. and Berven, B. A. CRC Press, Boca Raton, 127.
23. US EPA (1993). A guidebook to comparing risks and setting environmental priorities, EPA 230-8-93-003, Washington D.C.
24. Minnesota department of health fact sheet (2003). Volatile organic compounds-VOCs, 1-2.
25. Abraham, M. H. and Weatherby, P. K (1994). J. Pharm. Sci., **83**, 1450-1455.
26. Yasuda, N., Targ, A. G. and Eger, E. I (1989). Anesth. Analg., **69**, 370.

Chapter

3

Air to blood and air to tissue distribution of volatile organic compounds (VOCs)

The solubility of gases and vapours in biological fluids (plasma, blood and urine) and tissues (brain, muscle, lung, liver, kidney, heart and fat) is important in areas of research such as toxicology, environmental health and medicine especially anaesthetics. Knowledge of the distribution of VOCs in various body compartments is valuable and contributes to the understanding of the risk of toxic effects.



Scheme 3.0. Identifies the important compartments in human body where VOCs may distribute to and from the blood

Reported experimental values can be obtained from various journal articles in the literature. The distribution of VOCs from air to blood and air to tissue for a given temperature (310 K) is referred to as a partition coefficient (K or L). The partition coefficient can be calculated by dividing the concentration of solute in the tissue (or blood) phase over the concentration of solute in the gas phase at equilibrium (see scheme 1.0, chapter 1). The measure of K has units of gas volume, measured at ambient temperature and at 1 ATA partial pressure, which dissolves in a unit volume (V) of fluid. Other common units for concentration of the solute in the tissue phase is represented by μmol/kg.[1]

This distribution constant K (or L) is known as the Ostwald solubility coefficient. In scheme 1.0, chapter 1, the K^b and K^t are the gas-to-blood or gas-to-tissue partition coefficients, and P is the blood-to-tissue partition coefficient. Most air to blood distributions have been determined by the technique of headspace vial-equilibrium, which shall be discussed latter in this chapter.[2-3]

Data for air to olive oil and air to saline (9% salt concentration) have been obtained in order to construct models that resemble membrane permeability. The tissue partitioning of volatile organic compounds (VOCs) is due to lipophilic and hydrophilic interactions with tissue components, and empirical relations have been established between olive oil ($K_{oil:air}$), saline ($K_{saline:air}$), and tissue partition coefficients ($K_{tissue:air}$) for human and rat tissues.

The literature sources for air to blood (or air-tissue) distribution are from the following authors: Abraham and Weathersby,[2] Meulenberg,[3] Krishnan,[4-5] Poulin,[5] Gargas,[6-7] Fiserova-Bergerova,[8-9] Filser,[10] Pezzagno,[11] Zahlsen,[12] Balaz[13] and Zhang.[14] Various model (equations) have been formulated to correlate air-to-tissue distribution, and are summarised in table 3.0. Meulenberg[3] has obtained a reasonable amount of data in his training set as shown in table 3.0; this model has no test sets to verify predictions. Gargas et al have also produced models for air to blood and air tissue distribution of VOCs.[6-7] Unfortunately their training set does not include standard deviations and also statistical information is not provided about the test set. From table 3.0 it appears that none of the models have any test sets to verify predictions. As shown by this table, it appears that the number of compounds is small for the majority of tissues studied and it seems that standard deviations are absent for some of the training sets.

An improvement can be made on air to blood (or air to tissue) distribution by using training sets and test sets containing a large number of compounds. It would be very interesting to see if the diversity of compounds helps to improve models that are currently available. Most of the models found in table 3.0 are outdated and new measurements have been made for VOCs, which should improve the standard deviation and the square of the correlation coefficient for the training sets.

AIR-TISSUE DISTRIBUTION		TRAINING SET				TEST SET	
Reference	**Tissue**	**No.**	**Var.**	**r^2**	**S.D.**	**No.**	**A. E.**
Gargas, M. L et al.[7]	Blood (Hum)	55	2	0.928	-	-	-
Gargas, M. L et al.[7]	Blood (Hum)	36	2	0.875	-	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Blood (Hum)	82	5	0.977	0.20	-	-
Meulenberg, C. J.W. and Vijverberg, H. P. M[3]	Blood (Rat)	92	2	0.930	-	-	-
Meulenberg, C. J.W. and Vijverberg, H. P. M[3]	Blood (Hum)	109	2	0.990	-	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Plasma (Hum)	32	5	0.984	0.23	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Brain (Hum)	41	5	0.968	0.27	-	-
Meulenberg, C. J.W. and Vijverberg, H. P. M[3]	Brain (Rat)	19	2	0.900	-	-	-
Meulenberg, C. J.W. and Vijverberg, H. P. M[3]	Brain (Hum)	35	2	0.980	-	-	-
Gargas, M. L et al.[7]	Liver (Rat)	55	2	0.903	-	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Liver (Hum)	29	5	0.981	0.10	-	-
Meulenberg, C. J.W. and Vijverberg, H. P. M[3]	Liver (Rat)	77	2	0.920	-	-	-
Meulenberg, C. J.W. and Vijverberg, H. P. M[3]	Liver (Hum)	28	2	0.880	-	-	-
Gargas, M. L et al.[7]	Muscle (Rat)	55	2	0.879	-	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Muscle (Hum)	41	5	0.966	0.26	-	-
Meulenberg, C. J.W. and Vijverberg, H. P. M[3]	Muscle (Rat)	76	2	0.820	-	-	-

Meulenberg, C. J.W. and Vijverberg, H. P. M[3]	Muscle (Hum)	35	2	0.990	-	-	-
Gargas, M. L et al.[7]	Fat (Rat)	55	2	0.947	-	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Fat (Hum)	36	5	0.988	0.12	-	-
Meulenberg, C. J.W. and Vijverberg, H. P. M[3]	Fat (Rat)	76	2	0.860	-	-	-
Meulenberg, C. J.W. and Vijverberg, H. P. M[3]	Fat (Hum)	41	2	0.920	-	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Lung (Hum)	36	5	0.976	0.23	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Kidney (Hum)	36	5	0.951	0.27	-	-
Meulenberg, C. J.W. and Vijverberg, H. P. M[3]	Kidney (Rat)	16	2	0.910	-	-	-
Meulenberg, C. J.W. and Vijverberg, H. P. M[3]	Kidney (Human)	34	2	0.980	-	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Heart (Hum)	25	5	0.957	0.17	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Urine (Hum)	40	5	0.876	0.40	-	-

No. = the number of compounds used in the training set

Var. = the number variables used in the training set

r^2 = the overall square of the correlation coefficient (explained in detail from chapter 6)

S.D. = standard deviation (explained in detail from chapter 6)

A.E. = average error (explained in detail from chapter 6)

Hum = Human data only

Table 3.0. Literature survey on VOCs air-blood (or air to tissue) distribution

A number of equations have also been reported for the correlation of blood-tissue distribution of VOCs. These equations are summarised in table 3.1. From table 3.1., the number of compounds represented by the training sets can be seen to be very small. This may cause problems if one needs to predict the distribution of compounds between blood and tissue based on structural diversity. Unfortunately, as shown in table 3.1, there are no test sets to verify the model predictions and this would be very useful to incorporate into new models in the near future, especially with larger data sets. Nothing has been mentioned with regards to statistics to the

test sets to verify the training equations used to predict solute blood to tissue distributions. There is no indication from the table that the VOCs and drug data can be combined to enable equations to predict the distribution for VOCs and drugs in general.

VOC BLOOD-TISSUE DISTRIBUTION		TRAINING SET				TEST SET	
Reference	**Tissue**	**No.**	**Var.**	**r^2**	**S.D.**	**No.**	**A. E.**
Fiserova-Bergerova, V. Diaz, M. L[9]	Brain (Hum)	35	1	0.940	0.78	-	-
Zhang, H.[14]	Brain (Hum)	35	2	0.920	0.12	-	-
Balaz, S. and Lukacova, V [13]	Brain (Hum)	35	7	0.956	0.08	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Brain (Hum)	39	5	0.865	0.16	-	-
Fiserova-Bergerova, V. Diaz, M. L[9]	Kidney (Hum)	35	1	0.945	0.42	-	-
Zhang, H.[14]	Kidney (Hum)	34	2	0.897	0.11	-	-
Balaz, S. and Lukacova, V [13]	Kidney (Hum)	33	7	0.924	0.09	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Kidney (Hum)	35	4	0.774	0.17	-	-
Fiserova-Bergerova, V. Diaz, M. L[9]	Liver (Hum)	27	1	0.920	1.00	-	-
Zhang, H.[14]	Liver (Hum)	29	2	0.872	0.13	-	-
Balaz, S. and Lukacova, V [13]	Liver (Hum)	28	5	0.931	0.10	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Liver (Hum)	28	5	0.865	0.16	-	-
Fiserova-Bergerova, V. Diaz, M. L[9]	Lung (Hum)	30	1	0.541	3520	-	-
Zhang, H.[14]	Lung (Hum)	29	2	0.653	0.10	-	-
Balaz, S. and Lukacova, V [13]	Lung (Hum)	28	4	0.709	0.07	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Lung (Hum)	35	2	0.389	0.14	-	-
Zhang, H.[14]	Muscle (Hum)	36	2	0.884	0.13	-	-
Fiserova-Bergerova, V. Diaz, M. L[9]	Muscle (Hum)	34	1	0.927	0.70	-	-

Balaz, S. and Lukacova, V [13]	Muscle (Hum)	34	7	0.908	0.11	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Muscle (Hum)	40	3	0.790	0.18	-	-
Zhang, H.[14]	Fat (Hum)	36	2	0.972	0.18	-	-
Balaz, S. and Lukacova, V [13]	Fat (Hum)	36	7	0.960	0.20	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Fat (Hum)	35	5	0.967	0.19	-	-
Zhang, H.[14]	Heart (Hum)	16	2	0.599	0.20	-	-
Balaz, S. and Lukacova, V [13]	Heart (Hum)	14	3	0.701	0.08	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Heart (Hum)	23	4	0.849	0.12	-	-
Weathersby, P. K. and Abraham, M. H.[2]	Plasma (Hum)	32	4	0.656	0.09	-	-

No. = the number of compounds used in the training set

Var. = the number variables used in the training set

r^2 = the overall square of the correlation coefficient (explained in detail from chapter 6)

S.D. = standard deviation (explained in detail from chapter 6)

A.E. = average error (explained in detail from chapter 6)

Hum = Human data only

Table 3.1. Literature survey on VOC blood to tissue distribution

The air to olive oil distribution model can also be incorporated in pharmacokinetic processes for predicting blood-tissue distribution of VOCs. Weathersby[2], Fuchs[15] and Klopman[16] have made models to predict air to olive oil distribution (see table 3.2). It appears from his report, that Klopman[16] has made his models based on training data sets containing both predictive and experimental values (see table 3.2). It would be better to devise a model based on true experimental values for solutes (VOCs) that have been accurately measured. So his overall equation may not be as valuable as one might expect in order to assess reliability for the predictions of air to olive oil distribution, for VOCs in general. Finally, this model shown in table 3.2 is outdated and new compounds have recently been measured. This can be used to devise new models based on larger training data

sets. This would allow one to use larger test sets, which would be beneficial in determining how reliable the model is for such predictions.

AIR-SOLVENT DISTRIBUTION		TRAINING SET				TEST SET	
Reference	**PHASE**	**No.**	**Var.**	**r^2**	**S.D.**	**No.**	**A. E.**
Weathersby, P. K. and Abraham, M. H.[2]	Olive oil	88	4	0.997	0.08	-	-
Fuchs, R. and Abraham, M. H[15]	Olive oil	52	3	0.947	0.23	-	-
Klopman, G et al.[16]	Olive oil	159[a]	23	0.938	0.29	36	0.271
Klopman, G et al.[16]	Olive oil	192[a]	26	0.968	0.21	-	-

[a] Indicates that this data includes a large amount of calculated data, as opposed to experimental data.

No. = the number of compounds used in the training set

Var. = the number variables used in the training set

r^2 = the overall square of the correlation coefficient

S.D. = standard deviation (explained in detail from chapter 6)

A.E. = average error (explained in detail from chapter 6)

Table 3.2. Literature survey on VOCs air to olive oil distribution (310K)

Experimental methods for obtaining air to blood (or air to tissue) partition coefficient for rats and humans using head space chromatography

Measurements of air to blood distribution for VOCs in rats, has been studied by many scientist including Michael Gargas and his team.[6-7] His experiments involved using rats to measure the distribution of VOCs from air to blood (or tissue), using the method of head space chromatography. Three or more rats are contained in a glass chamber with oxygen being maintained at 21%. Carbon dioxide was removed by soda lime. Appropriate amounts of known VOC were injected until a suitable concentration was reached. The concentration of VOC in the gas phase was monitored over a period of time using gas chromatography (with flame ionization) until equilibrium had been reached. At various fixed times, the rats were anesthetised with carbon dioxide and blood was removed from the inferior vena

cava, using a heparinised gas-tight syringe. Due to the volatility and rapid metabolism of VOCs, the removal of tissues was carefully timed. One-millilitre aliquots of blood were immediately transferred to chilled 5.5 ml vials sealed with Teflon lined septa. The blood samples were incubated, with shaking for 1 hour at 37°C in a vortex evaporator and 1 ml aliquot of the headspace was analysed for VOC in blood. The concentrations of VOCs were then determined from the standard curves prepared in blood, under the condition described.

Immediately after blood sampling, the tissue e.g. liver of each rat was excised, a 4 g section transferred to chilled pre-weighed 20 ml vial containing buffer, the vial was capped and reweighed. The liver was than briefly homogenised (approx 10 seconds) by using a Tissue-miser and 1 ml aliquots of the homogenates were transferred to chilled 5.5 ml vials. The vials were sealed immediately and incubated with shaking for 1 hour at 37°C in a vortex evaporator; equilibrium was established within an hour (between the liver homogenate and the headspace). Following the incubation, the VOC in the vial headspace was analysed by GC or by GC-mass spectrometry. The advantage of GC-MS is that it can used to determine the actual compound and also any metabolites formed from the sample. The GC injector, detector and oven temperatures were held at constant temperature. Helium was used as a carrier gas for this given sample. The concentrations of VOCs were then determined from standard curves prepared in liver under the condition described.

For human (or rat) subjects, the experimental method was the same but the dead tissue of the human (or rat) was homogenised and placed into air tight vial where the VOCs was injected and were allowed to equilibrate over period of time. Following incubation, the VOC in the vial headspace was analysed by GC or by GC-mass spectrometry.

References

1. Weathersby, P. K. and Homer, L. D (1980). Undersea Biomedical Research, **7**, 277-296.
2. Abraham, M. H. and Weathersby, P. K (1994). J. Pharm. Sci, **83**, 1450-1455.
3. Meulenberg, C. J. W. and Vijverberg, H. P. M (2000). Toxicology and Applied Pharmacology, **165**, 206-216.
4. Beliveau, M., Charest-Tardif, G. and Krishnan, K (2001). Chemosphere, **44**, 477-381.
5. Poulin, P. and Krishan, K (1996). Toxicology and applied pharmacology, **136**, 131-137.
6. Reitz, R. H., Gargas, M. L., Anderson, M. E., Provan, W. M. and Green, T. L (1996). Toxicology and Applied Toxicology, **137**, 253-267.
7. Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H. and Anderson, M. E (1989). Toxicology and Applied Pharmacology, **98**, 87-99.
8. Tichy, M., Fiserova-Bergerova, V. and Di-Carlo, F. J (1985). Pharmaco-chemistry, **8**, 225-231.
9. Fiserova-Bergerova, V. and Diaz, M. L (1986). International Archives of Occupational and Environmental Health, **58**, 75-87.
10. Filser, J. G., Csanady, GY. A., Hartmann, M., Denk, B., Kauffmann, A., Kessler, W., Kreuzer, P. E., Putz, C., Shen, J. H. and Stei, P (1996). Toxicology, **113**, 278-287.
11. Imbriani, M., Ghittori, S., Pezzagno, G. and Capodagio, E (1985). G. Ital. Med. Lav., **7**, 133-140.
12. Zahlsen, K., Eide, I., Nilsen, A. M. and Nisen, O. G (1993). Pharmacology & Toxicology, **73**, 163-168.
13. Balaz, S. and Lukacova, V (1999). QSAR, **18**, 361-368.
14. Zhang, H (2004). Journal of Pharmaceutical Sciences, **93**, 1595-1604.
15. Fuchs, R. and Abraham, M. H (1988). J. Chem. Soc. Perkin Trans. II, 523-527.
16. Klopman, G., Ding, C. and Macina, O. T (1997). J. Chem. Information and Computer Sciences, **37**, 569-575.

Chapter

4

Introduction to blood-tissue distribution of drugs

The single largest cause of attrition in drug development is drug absorption, distribution, metabolism and elimination (ADME).[1]

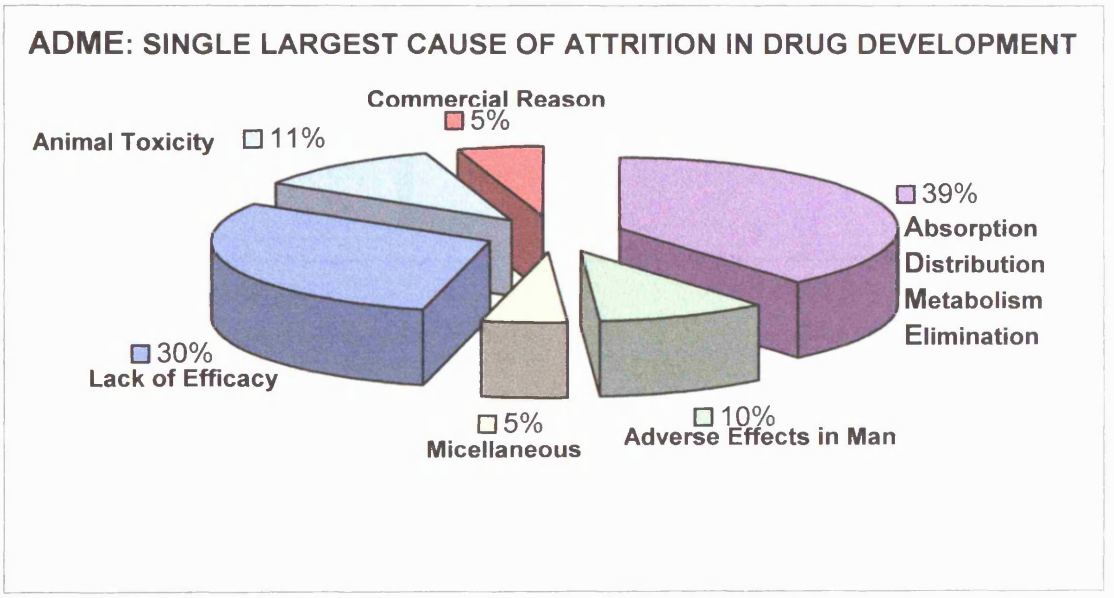


Fig. 4.0. Schematic representation of a pie chart displaying the biggest problems in drug development[1]

The purpose of drug therapy is to prevent, cure, or control many different disease states. To obtain this goal, sufficient drug doses must be delivered to the target tissues so that therapeutic and non-toxic levels are achieved. The clinician must be aware that four fundamental pathways control the speed of onset of drug action, the intensity of the drug's effect, and the duration of the drug action. The four main steps involve the movement and the modification of the drug within the body (Figure 4.1).

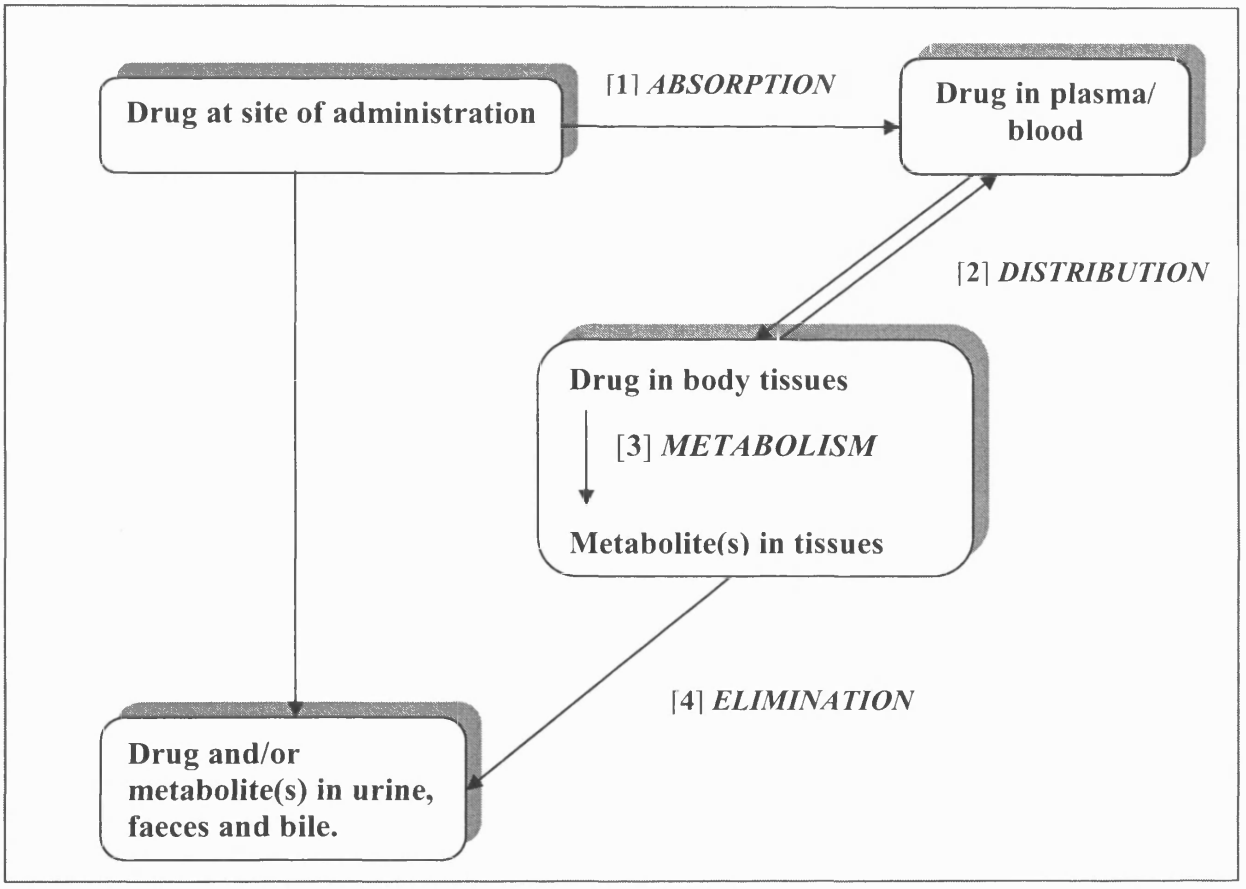


Fig. 4.1. Schematic representation of drug absorption, distribution, metabolism and elimination (ADME).[2]

First, drug absorption from the site of administration authorises the entry of the therapeutic agent (either directly or indirectly) in to plasma or blood (input). Second, the drug may then reversibly leave the blood stream and distribute between the various body tissues. Third, the liver, kidney and body tissue help to metabolise the drug from the body. Finally, the drug and its metabolites are eliminated from the body (output) in the urine, bile, sweat or faeces.[2]

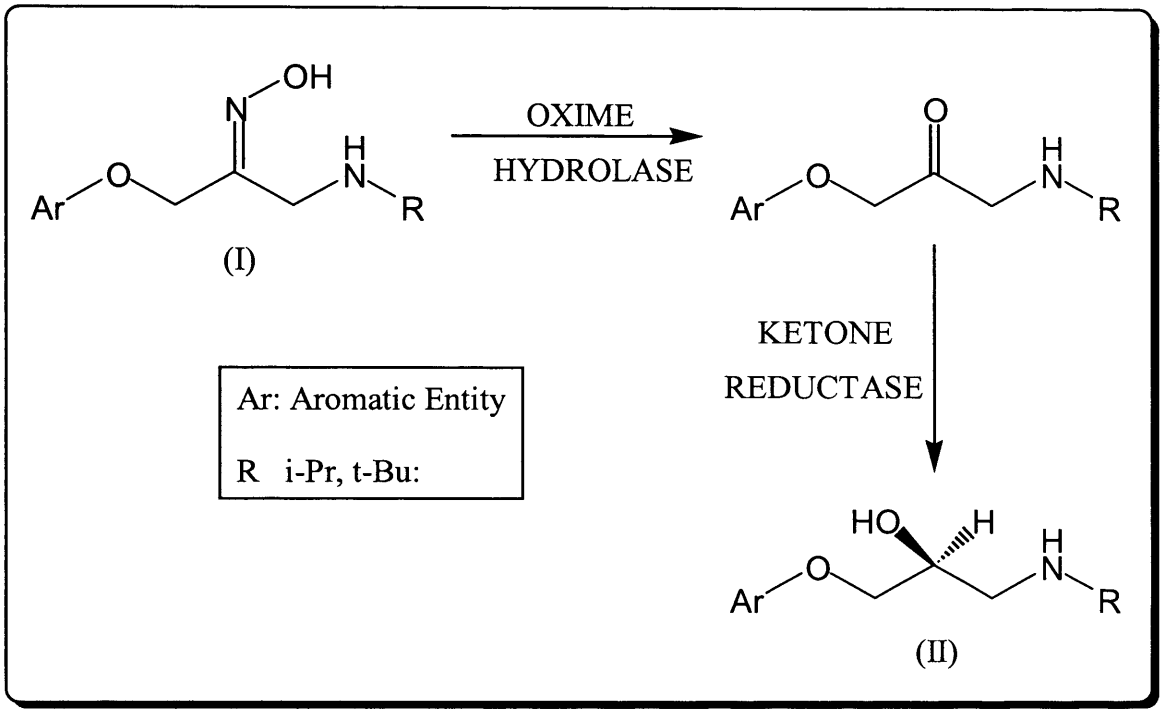
The aim of the present project is to concentrate on the second point (Figure 4.1) especially the distribution of drugs between blood, plasma and tissues. It is very important to know that ratio, and how it may be altered. For example Chadha and Abraham, showed that the change in functional group could alter the blood-brain ratio (BB). A non-polar functional group as a substituent, i.e. methyl or ethyl, would help a compound to distribute more to the brain than in the blood. Whereas polar groups containing OH or OMe would lead to distribution more in the blood than in the brain.[3]

FUNCTIONAL GROUP X	$LOGBB_{(CALCULATED)}$
H	0.42
Me	0.53
Et	0.67
Cl	0.54
OMe	0.35
NH_2	-0.11
OH	-0.28

Table 4.0. Shows how log BB changes with the functional group X on an aromatic ring system.[3]

The majority of potent drugs that are administered to patients are pharmacologically active; i.e., they have the capability to bind to a specific receptor and produce their pharmacological effects. When absorbed into the circulation, the drug is distributed into various body tissues. This depends on the characteristic of the drug as it may bind to plasma proteins, accumulate in fat, adipose or other tissues and then bind specifically to the receptor site. The drug interaction with the receptor site is important in terms of the drug's pharmacological effects.

The biochemical transformation of a drug into another substance by an organism is known as drug metabolism. Higher concentrations of drug-metabolising enzymes are found in many body tissues including the intestines, lungs, liver and the kidneys. Drug metabolites are usually more water-soluble and, consequently, more readily excreted from the body. The pharmacological activities for these metabolites are mostly reduced. The majority of drugs that are taken by patients are pharmacologically active in the administered form.[3] Though, not all drugs that administered are pharmacologically active; some must be metabolically converted to the pharmacologically active form and these specialised drugs are known as prodrugs.[5] For example **II** is not readily transferred to the eye, but the prodrug **I** is more readily transferred. When it gets to the eye it is enzymatically converted to the active species **II** (Scheme 4.0).



Scheme 4.0. Shows the sequential activation of oximes in the site and stereospecific delivery of a drug (e.g. alprenolol is a β blocker that has a protective effect after a heart attack by reducing high blood pressure) to the eye.[6]

Both of the above methods, which involve functional group interconversion, as shown in Table 4.0 and Scheme 4.0, are important to help the transport of drugs to the active site.

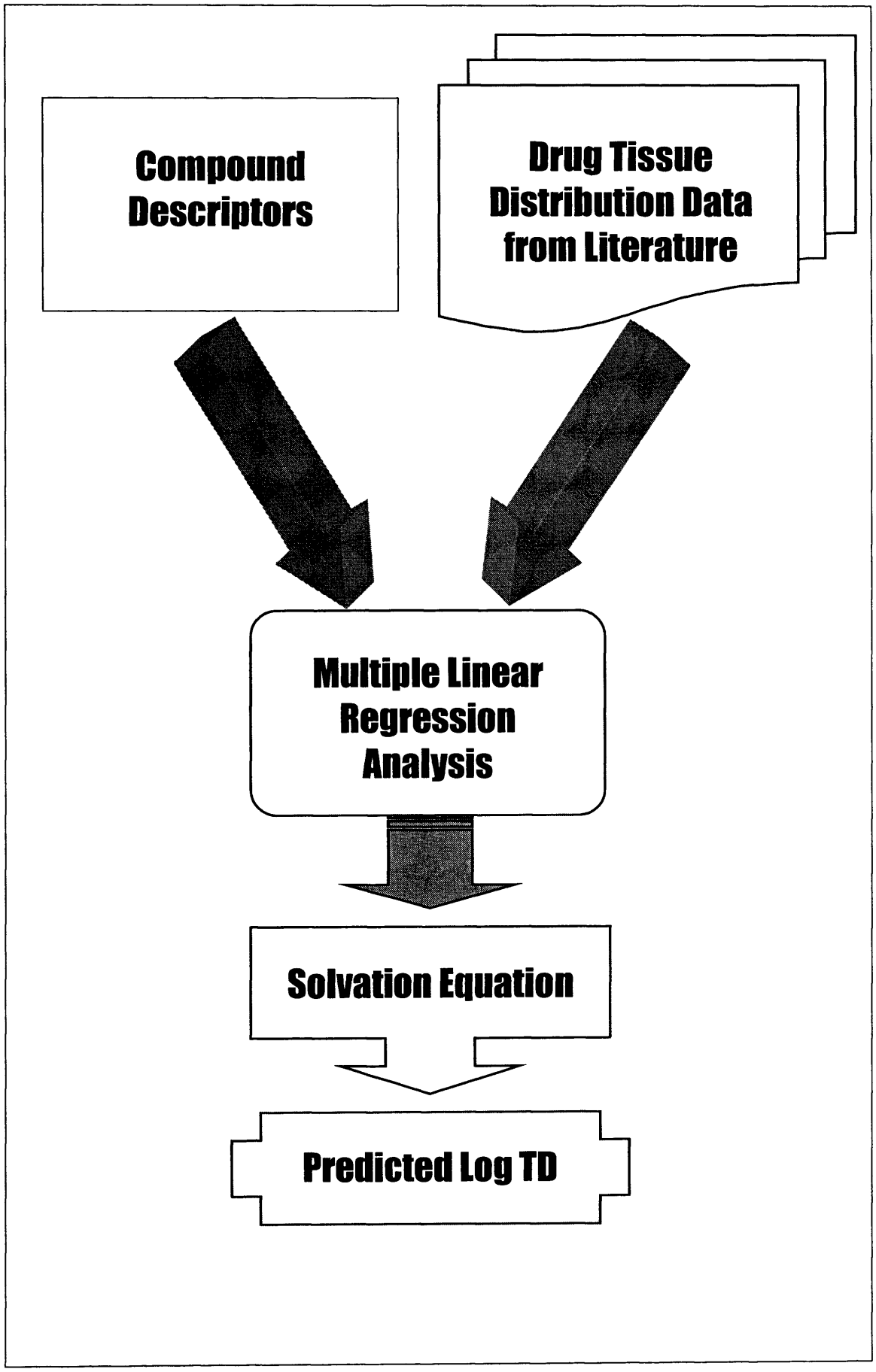
Aim and plan

The present project involves the collection of drug distribution data obtained from a wide range of literature papers published by various types of journals. These published data will be analysed carefully in order to obtain values of the distribution coefficients of drugs between blood/plasma/serum and various tissues.

Once enough data for distribution to a given tissue have been collected, they will be used to obtain predictive equations using the Abraham solvation equation. A necessary requirement for this is the determination of descriptors for the drugs in question. These predictive methods using the Abraham equation may then be used in the first stages of drug design. These predictions can be made from structure before compounds are synthesised. Compounds with poor tissue distribution, can then be identified.

The drug distribution data will be collected into a database, and the descriptors for the compounds in the database will be found at a later stage. The drug distribution data and the descriptors will be combined into a multi-linear regression (statistical calculation) to give provisional solvation equations. These will then be evaluated to then give final predictive equations for drug blood/plasma/serum to tissue distribution. The basic steps are outlined by a flow diagram in Figure 4.2.

Figure 4.2. Protocol used to obtain predicted equations.



How important is drug distribution data and how is this obtained

The prediction of the specific behaviour and effects which a solute (or drug) will demonstrate in a given biological environment, is desirable in a number of current research areas including: medicine (e.g. the study of anaesthetics), pharmacy (e.g. formulation for various drugs) and toxicology for the study of environmental toxicants (e.g. PCB's) etc. There are many factors affecting the biological activity and availability, ranging from the chemical structure of the solute molecules (or drugs) themselves to the uptake process by which they distribute into various tissue sites. For example, in the pharmacokinetics of a drug molecule, it is important to quantify how much drug is being distributed to various body tissue sites.

The main workers' in the field of drug distribution data for drug compounds are Patrick Poulin et al[7-9] and Sven Bjorkman.[10] The objectives of Poulin (et al) studies were to develop and validate mechanistic equations for predicting the rabbit, rat, and mouse $P_{t:p}$ (ratio of drug tissue concentration over the drug plasma concentration at steady state) of non-adipose and non-excretory tissues (bone, brain, heart, intestine, lung, muscle, skin, spleen) for 65 structurally unrelated drugs and to evaluate the adequacy of using $P_{p:t}$ of muscle as predictors for $P_{p:t}$ of other tissues.[7] Bjorkman derived equations to predict volume of distribution based on data obtained from various literatures sources containing plasma-tissue coefficients (K values at steady state) of 43 drugs. His model has benefits for people who have interest primarily in pharmacokinetics where the predictive values for volume of distribution (Vd) based on drugs alone is important.[10]

The main workers in the field of drug blood-brain distribution are: Cabrera et al,[11] Rose et al,[12] Clark,[13] Keseru and Molnar,[14] Feher et al,[15] Lombardo et al[16], Luco,[17] Norinder et al,[18] Lobell et al,[19] Platts et al,[20] Kelder et al,[21]Chadha et al,[22, 23] Hutter,[24] Liu et al,[26] Kaznessis et al,[27] Ertl et al,[28] Ooms et al,[29] Salminen et al,[30] Kaliszan and Markuszewski,[31] Kaliszan et al[31] and Hou and Xu[32,33] Although there has been a large number of reported work on blood to brain distribution for drugs (as summarised by table 4.1), but there still is a large scope of work to be done on drug blood to tissue distribution.

Cabrera et al method of using TOPS MODE is complicated.[11] This method involves calculating descriptors using a software (MODES labs) using spectral moment of each bond weight and number bonds in the molecules also considering reported influence of polar surface, atomic mass and hydrophobicity. Cabrera et al training set appears to have large data set with compounds containing drugs and VOCs.

DRUG BLOOD-BRAIN DISTRIBUTION	**TRAINING SET**				**TEST SET**	
Reference	**No.**	**Var.**	**r^2**	**S.D.**	**No.**	**AAE.**
Cabrera et al.[11]	114	2	0.697	0.422	28	0.33[b]
Cabrera et al.[11]	81	2	0.740	-	33	0.33[b]
Keseru and Molnar.[14]	55	1	0.723	0.370[d]	5, 25	0.14,[c] 0.37[c]
Clark[13]	55	2	0.787	0.354	5, 5	0.14, 0.24
Luco.[17]	58	18	0.850	0.318	12, 25	0.24,[a] 0.54,[a]
Feher et al.[15]	61	3	0.723	0.420	14,25	0.76, 0.80
Hutter.[24]	90	12	0.865	0.309[d]	-	-
Lombardo et al.[16]	55	1	0.670	0.410[d]	6	0.62
Norinder et al.[18]	56	14	0.834	0.312	6	0.52
Rose et al.[12]	106	3	-	0.450	20	0.47
Rose et al.[12]	102	3	0.660	-	-	-
Platts et al.[20]	148	6	0.745	0.343	-	-
Platts et al.[20]	74	6	-	0.340	74	0.38
Lobell et al.[19]	48	5	0.837	0.260[b]	17	0.41[b]
Kelder et al.[21]	45	1	0.841	0.360	-	-
Chadha et al.[22, 23]	57	5	0.906	0.197	-	-
Liu et al[26]	55	2	0.790	0.350	11	0.30[c]
Kaznessis et al[27]	76	5	0.941	0.173	4	0.48[a]
Ertl et al[28]	45	1	0.840	-	-	-
Ooms et al[29]	79	31	0.760	-	-	-
Salminen et al[30]	21	3	0.848	0.270	-	-
Salminen et al[30]	23	3	0.848	0.320	-	-
Kaliszan et al[31]	20	2	0.642	0.486[d]	-	-
Kaliszan et al[31]	20	2	0.845	0.321[d]	-	-
Kaliszan et al[31]	20	2	0.889	0.271[d]	-	-
Kaliszan et al[31]	20	2	0.897	0.126[d]	-	-
Hou and Xu[32]	72	3	0.785	0.358	35	0.41[c]
Hou and Xu[33]	78	3	0.743	0.375	35	0.44[c]

[a] is the root mean square value

[b] is the mean absolute error

[c] is the standard deviation

[d] is the standard error

Table 4.1. Summary of models for blood to brain distribution of drugs and VOCs

This training set appeared to have relative large standard deviations (S.D. 0.422) compared to other models shown in table 4.1.[11] Cabrera et al's correlation coefficient appears to be reasonable, although, there have better reported models with larger data sets for example; Platts et al.[20] and Rose et al[12] also have a large data set for drug blood brain distribution model. Rose et al's equation unfortunately does not have a very good correlation coefficient ($r^2 = 0.66$) for the large data set in the training equation. It appears that for this blood to brain model that the training set has the largest standard deviation, as compared to other models shown in table 4.1. The number of compounds used in the test sets by Rose (et al), are too small.[12]

With regards to the drug blood-brain model, new models would be of interest to pharmaceuticals industry and the environmental health providing that the model consist of larger data sets along with bigger test sets for the validation of predictive model.

The present work expands on the diversity of the drugs and also includes other small solutes (mostly general anaesthetics) to enable the construction of more general predictive equations.

The purpose of this project is to collect large amounts of data from various sources for blood: tissue distribution coefficients (either directly or indirectly) for drugs (or small solutes) to facilitate the use of physiological-based pharmacokinetic models (**PBPK**) in drug discovery used to estimate the disposition of drugs in various species under steady state condition. Until now, the use of PBPK models in early stages of the drug discovery processes has not been possible, since the estimation of $\mathbf{P}_{\text{blood: tissue}}$ (or $\mathbf{P}_{\text{plasma: tissue}}$ and $\mathbf{P}_{\text{serum: tissue}}$) of new drug candidates by using conventional *in vitro* and /or *in vivo* methods is too time and cost intensive especially with animal testing. The objectives of the present study are to develop and validate mechanistic equations for predicting the disposition of solutes (drugs) for various species (including rats, pregnant rats, rabbits, mouse and most

importantly humans) and to see if there are any similarities for drug distribution data. If so, then data for two or more species may be combined to generate one general equation for predicting drug blood to tissue distribution. The different tissues considered in this study when validating $P_{\text{blood: tissue}}$ ($P_{\text{plasma: tissue}}$ & $P_{\text{blood: tissue}}$) are adrenal, adipose, brain (including various brain compartments), bile, bone, cerebrospinal fluid, fat, gut, gastrointestinal tract, heart, kidney, lung, muscle, ovary, placenta, pancreas, spinal fluid, skin, spleen, testes, trachea, large intestine and small intestines. Ultimately the data for drug blood to tissue distribution for various tissues will be compared by statistical methods to assess if there are any similarities between the different types of data of the same solutes (or drugs). The drug blood to tissue distribution between two phases, at steady-state, can be calculated from the following equation:

P = Total drug concentration in tissue/ Total drug concentration in blood Equation 4.0

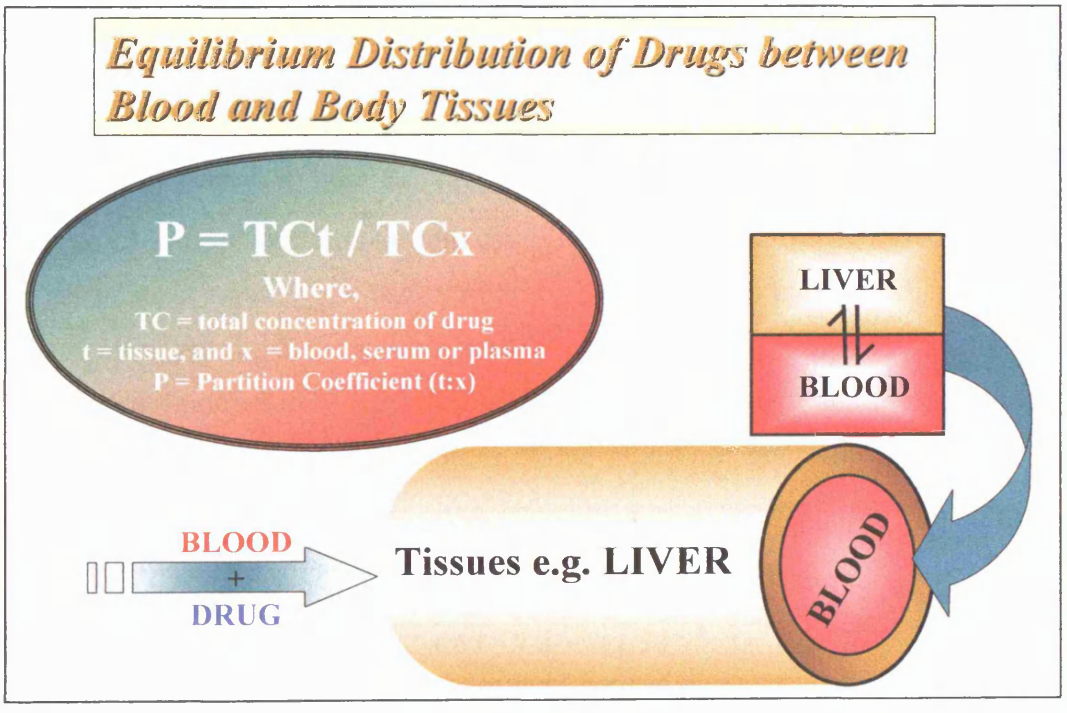


Figure 4.3. Diagram for drug blood to tissue distribution at steady state (equilibrium)

Careful interpretation of various recorded data has to be done in order to collect reliable data for drug blood to tissue distribution. Some pharmacokinetic models for drug disposition only produce data for the concentration of drug blood to tissue distribution versus time (Table 4.2) without displaying any direct results for any drug blood to tissue partition between various tissues.

TISSUE	5 MIN	15 MIN	30 MIN	60 MIN	90 MIN
Blood	0.13	0.09	0.06	0.03	0.05
Plasma	0.19	0.13	0.08	0.07	0.07
Heart	0.26	0.15	0.09	0.04	0.04
Trachea	0.17	0.12	0.09	0.04	0.03

Table 4.2. Concentration of drug (adenosine) versus time.[25]

From the data in Table 4.2, a table of results for drug to tissue distribution data can be obtained by using equation 4.0 and this ratio is summarised by table 4.3. This can then be used to plot graphs to determine the drug blood to tissue distribution at steady state before finally adding this data to the database. Note the data for drug blood to tissue distribution can only be obtained if the drug remains at a constant equilibrium over a period of time as shown by Figures 4.4, 4.5 and 4.6. This would allow consistent comparison to be made with other tissue distribution values at steady state. The squared points from the graph indicate where the values can be averaged for the steady state equilibrium. The diamond plots indicate where the concentration of drug has changed over a period of time and the drug does not appear to be at steady state equilibrium (these values are ignored).

TISSUE	5 MIN	15 MIN	30 MIN	60 MIN	90 MIN
Plasma	1.46	1.44	1.33	2.33	1.40
Heart	2.00	1.67	1.50	1.33	0.80
Trachea	1.31	1.33	1.50	1.33	0.60

Table 4.3. Ratio for drug blood to tissue distribution obtained from data[25] in Table 4.2.

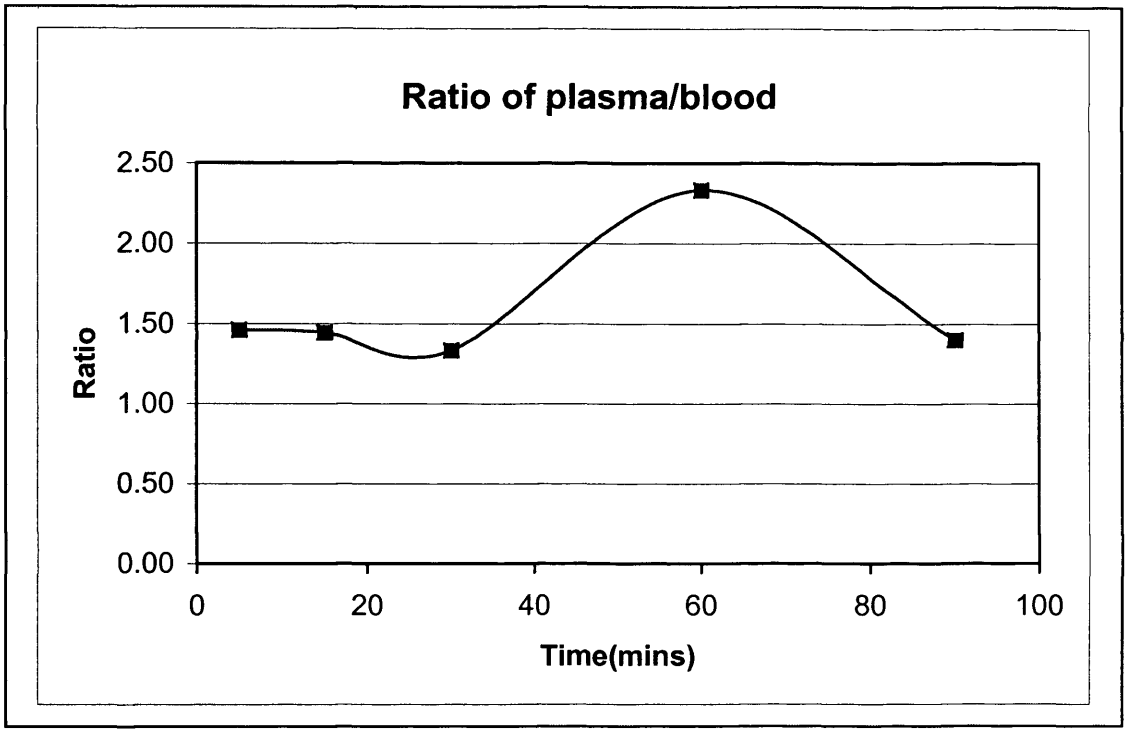


Figure 4.4. Graph showing the blood to plasma ratio vs time,[25] from Table 4.3.

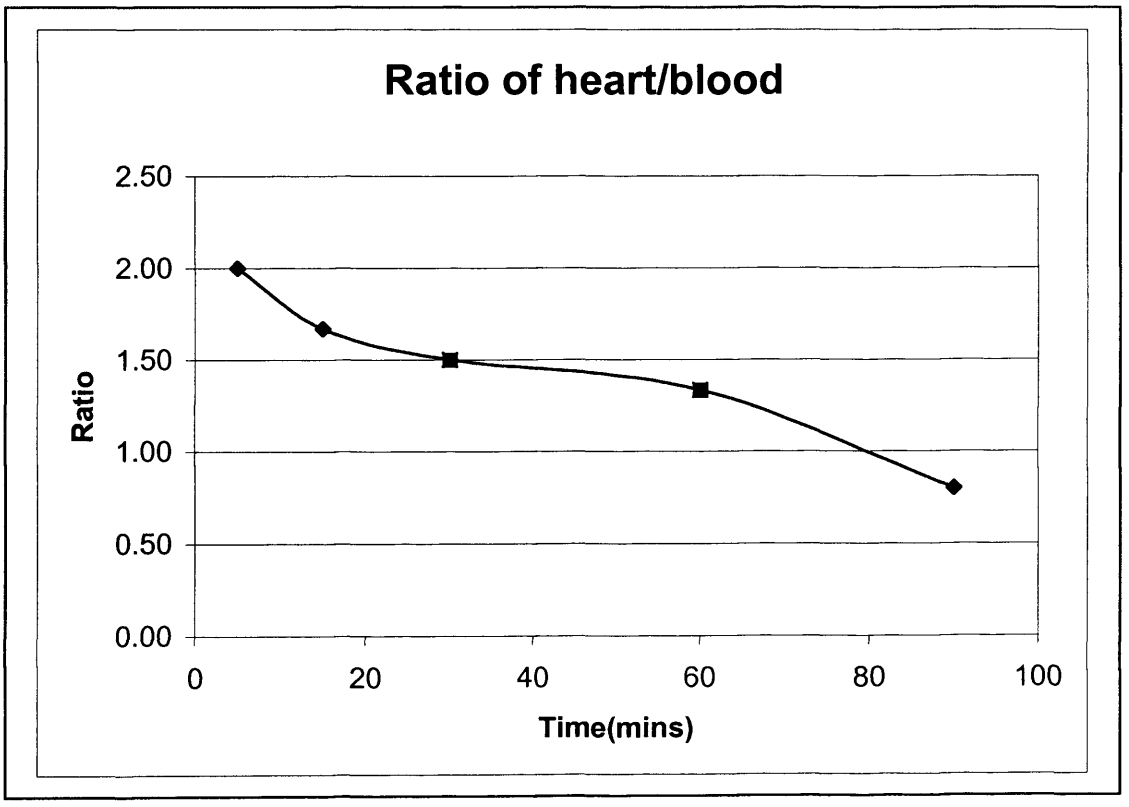


Figure 4.5. Graph showing the blood to heart ratio vs time,[25] from Table 4.3.

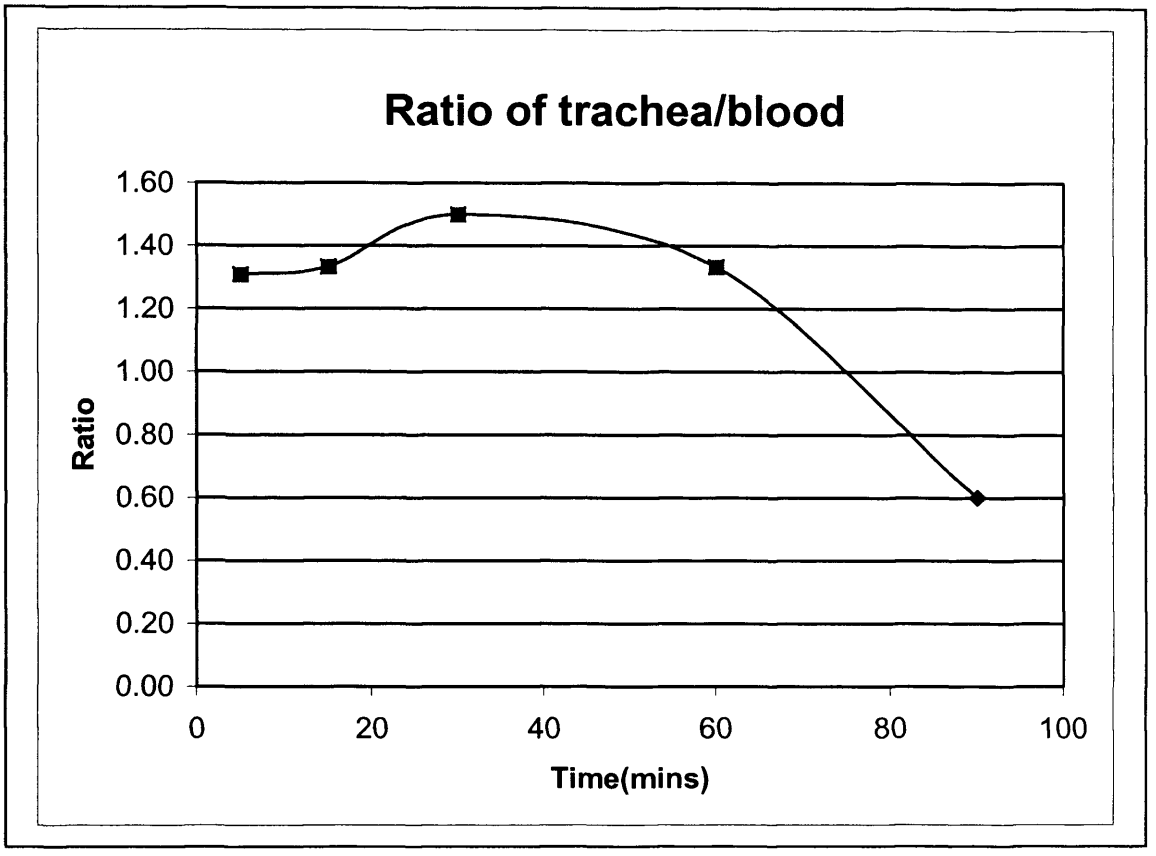


Figure 4.6. Graph showing the blood to trachea ratio vs time,[25] from Table 4.3.

From figures 4.4, 4.5, and 4.6 it is possible to obtain steady state distribution from the data corresponding to the blue square (■) points. These are summarised in Table 4.4.

TISSUE SAMPLES	BLOOD TO TISSUE RATIO TAKEN
Plasma	1.60
Heart	1.42
Trachea	1.37

Table 4.4. Ratio of blood to tissue at steady state distribution

Data from literature for drugs that do not give a steady state for blood to tissue distribution has to be avoided, as this data would not be reliable for comparisons

with other tissues at steady state (see Figure 4.7). It is important to ensure that plots are constructed before adding data to the database. In the case of Figure 4.7, this data can not be used for drug blood to tissue distribution, as no steady state is reached.

When collecting drug distribution data, it is very important to separate out data within one species. For example, the drug blood to tissue distribution data for one particular drug within rats and pregnant rats cannot be compared as they may have different distribution profiles. Later, comparison will be made to see if the overall data for several drugs are the same or different. Post mortem data for human subjects should also be kept separate from data on living human beings.

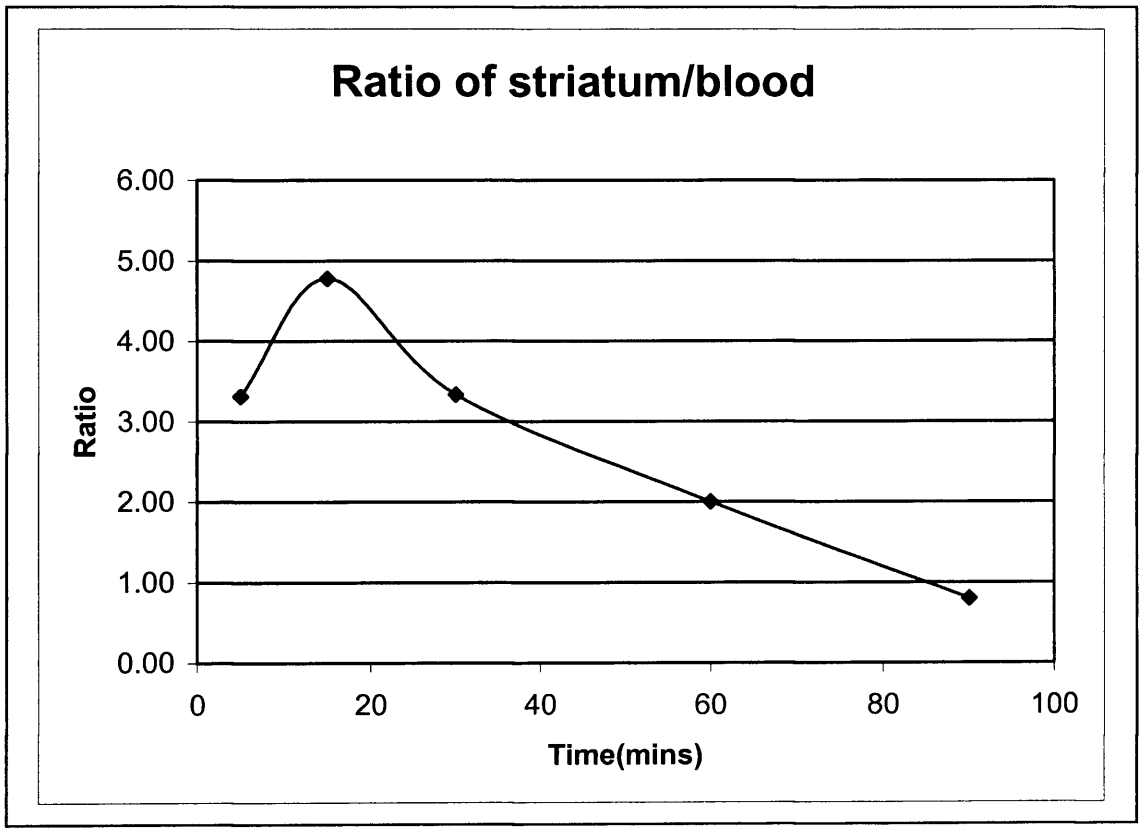


Figure 4.7. Graph showing the corpus blood to striatum (region of the brain) ratio vs time.[25]

References

1. Hill, S. A. (Review 2001) Modern medicinal chemistry. 1-4.
2. Mycek, M. J., Harvey, R. A. and Champe, C. P. (1997) Pharmacology. Lippincott illiams & Wilkins. 1-15.
3. Abraham, M. H. and Harpreet, S. C. (1994). Bioorganic & Medicinal Chemistry Letters. **4**, 2511-2516.
4. Backes, W. L. (2001). Clinical Journal of Oncology Nursing. **5**, 35-38.
5. Neal, M. J. (1993) Medical pharmacology at a glance. Blackwell scientific publications. 14-15.
6. Bodor, N. and Buchwald, P. (1997). Pharmacology. Ther. **76**, 1-27.
7. Poulin, P. and Theil, F. P. (2000). J. Pharm. Sci, **89**, 16-35.
8. Poulin, P., Schoenlein, K. and Theil, F. P. (2001). J. Pharm. Sci, **90**, 436-447.
9. Poulin, P. and Theil, F. P. (2002). J. Pharm. Sci, **91**, 1358-1367.
10. Bjorkman, S (2002). J. Pharmacy and Pharmacology, **54**, 1237-1245.
11. Cabrera, M. A., Bermejo, M., Perez, M. and Ramos, R (2004). Journal of Pharm. Sci., **93**, 1701-1717.
12. Rose, K., Hall, L. H. and Kier, L (2002). J. Chem. Inf. Comput. Sci., 42, 651-666.
13. Clark, D, E (1999). J. Pharm. Sci., **88**, 815-821.
14. Kerseru, G. M. and Molnar, L (2001). J. Chem. Inf. Comput. Sci., **41**, 120-128.
15. Fehur, M., Sourial, E. and Schimidt, J. H (2000). Int. J. Pharm., **201**, 239-247.
16. Lombardo, F., Blake, J. F. and Curatolo, W. J (1996). J. Med. Chem., **39**, 4750-4755.
17. Luco, J. M (1999). J. Chem. Inf. Comput. Sci., **39**, 396-404.
18. Norinder, U., Sjoberg, P., Osterberg, T (1998). J. Pharm. Sci., **87**, 952-959.
19. Lobell, M., Molnar, L. and Keseru, G. M (2003). J. Pharm Sci., **92**, 360-370.
20. Platts, J. A., Abraham, M. H., Zhao, Y. H., Hersey, A., Ijaz, L. and Butina, D (2001). Eur. J. Med. Chem., **36**, 719-730.
21. Kelder, J., Grootenhuis, P. D. J., Bayada, D. M., Deibressine, L. P. C. and Ploeman, J. P (1999). Pharm. Res., **16**, 1514-1519.
22. Abraham, M. H., Chadha, H. S. and Mitchell, R. C (1994). J. Pharm. Sci., **83**, 1257-1268.

23. Abraham, M. H., Chadha, H. S., Martins, Filomena, M., Mitchell, R. C., Bradbury, M. W and Gratton, J. A (1999). Pesticide Sci., **55**, 78-88.
24. Hutter, M. C (2003). J. Computer-Aided Molecular Design., **17**, 415-433.
25. Fazio, F., Todde, S., Moresco, R. M., Simonelli, P., Baraldi, P. G., Cacciari, B., Spalluto, G., Varani, K., Monopoli, A., Matarrese, M., Carpinelli, A., Magni, F., Kienle, G. K. (2000). J. Med. Chem., **43**, 4359-4362.
26. Liu, R., Sun, H. and So, S-S (2001). J. Chem. Inf. Comput. Sci., **41**, 1623-1632.
27. Kaznessis, Y. N., Snow, M. E. and Blankley, C. J (2001). J. Computer-Aided Molecular Design, **15**, 697-708.
28. Ertl, P., Rohde, B. and Selzer, P (2000). J. Med. Chem., **43**, 3714-3717.
29. Ooms, F., Weber, P., Carrupt, P-A. and Testa, B (2002). Biochimica et Biophysica Acta, **1587**, 118-125.
30. Salminen, T., Pulli, A. and Taskinen, J (1997). J. Pharmaceutical and Biochemical Analysis, **15**, 469-477.
31. Kaliszan, R. and Markuszewski (1996). International Journal of Pharmaceutics, **145**, 9-16.
32. Hou, T. J. and Xu, X. J (2003). ADME evaluation in drug discovery. 3. Modelling blood-brain barrier partitioning using simple molecular descriptors. J. Chem. Inf. Comput. Sci., **43**, 2137-2152.
33. Hou, T. J. and Xu, X. J (2004). ADME evaluation in drug discovery. 3. Modelling blood-brain barrier partitioning using simple molecular descriptors. J. Chem. Inf. Comput. Sci., **44**, 766-770.

Chapter

5

Introduction to physicochemical considerations

The physicochemical properties (e.g. lipophilicity, solubility) of a drug compound may be related directly or indirectly to its biological properties. Thus, by understanding the relationship between chemical structure and biological activity it is possible to gain knowledge on what is needed to attempt to predict biological properties, such as transport and the activity of new possible compounds. This chapter will outline a number of relevant physiochemical properties of compounds.

Brønsted acidity and basicity

To be able to appreciate the biological activity of a compound it is important to understand the transport processes of a compound to the site of action (i.e. body tissues) and also the binding of the compound to its target site (i.e. drug receptor). The acidity and basicity properties of a compound are important in order to control the transport and binding mechanism of that compound. It is important to notice the difference in Brønsted acidity and basicity, which refers to the loss or gain of a proton, and hydrogen-bond acidity and basicity, which refers to the tendency of a compound to donate or to receive a hydrogen bond. For protonic (Brønsted) acids and bases the following equations apply.

$$RCOOH \overset{K_a}{\rightleftharpoons} RCOO^- + H^+$$ Equation 5.0

$$K_a = \frac{[RCOO^-][H^+]}{[RCOOH]}$$ Equation 5.1

Here $\boldsymbol{K_a}$ is the acid dissociation constant for an acidic solute. Note the stronger the acid the greater the tendency for the solute to give up its proton. For a typical Brønsted base, the base denotes constant, $\boldsymbol{K_b}$ can be expressed as,

$$RNH_2 + H_2O \underset{}{\overset{K_b}{\rightleftharpoons}} RNH_3^+ + OH^-$$ Equation 5.2

$$K_b = \frac{[RNH_3^+][OH^-]}{[RNH_2]}$$ Equation 5.3

Very often, values of K_a for the conjugate acid are given and defined as,

$$RNH_3^+ \overset{K_a}{\rightleftharpoons} RNH_2 + H^+$$ Equation 5.4

$$K_a = \frac{[RNH_2][H^+]}{[RNH_3^+]}$$ Equation 5.5

From equations 5.3 and 5.5,

$K_a \times K_b = [H^+][HO^-] = 10^{-14}$ in water Equation 5.6

Then $\mathbf{pK_a + pK_b = 14}$ in water at 25°C.

The **pH** of a solution indicates the concentration of positively charged hydrogen ions in the solution (concentration is expressed as $[H^+]$). The definition of pH is:

$$pH = -\log [H^+] = a_{H+}$$ Equation 5.7

For convenience, the strength of an acid is generally indicated by its $\mathbf{pK_a}$ value rather than its K_a value. Then for a Brønsted acid, the concentration of neutral and ionised forms is given by the following equation.

$$pKa = pH + \log ([HA]/[A-])$$ Equation 5.8

$pK_a = -\log K_a$ Equation 5.9

For acidic solute the smaller the pK_a, the stronger is the acid.

Very strong acids	$pK_a < 1$
Moderately strong acids	$pK_a = 1\text{-}5$
Weak acids	$pK_a = 5\text{-}15$
Extremely weak acids	$pK_a > 15$

For a Brønsted base, the concentration of ionised and neutral forms is given by the following equation.[1] Here the concentration is taken as an approximation of activity. Note it is the ionisation that has the potential to limit the passage through membranes.

$pK_a = pH + \log[RNH_3^+] - \log[RNH_2]$ Equation 5.10

For basic solute the smaller the $\mathbf{pK_b}$, the stronger is the conjugate base.

Very strong bases	$pK_a > 9$	$pK_b < 5$
Moderately strong bases	$pK_a = 8\text{-}5$	$pK_b = 6\text{-}9$
Weak bases	$pK_a \leq 4$	$pK_b \geq 10$

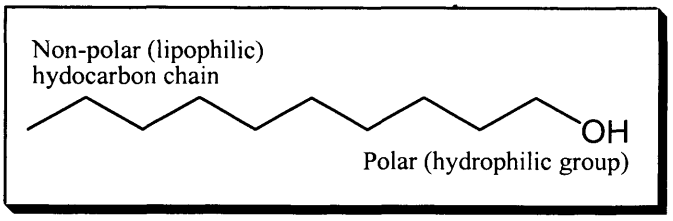
Henderson-Hasselbach equation

Whether a given acid will lose a proton in an aqueous solution depends on the pK_a of the acid and the pH of the solution.[1] The relationship between the two is given by the Henderson-Hasselbach equation. This is an extremely useful equation because it shows us whether a compound will exist in its acidic form (with its proton retained) or in its basic form (with its proton removed) at a particular pH.

An acid will exist primarily in its acidic form if the pH of the solution is less than its pK_a. An acid will exist primarily in its basic form if the pH of the solution is greater than its pK_a. When $pK_a = pH$ then the concentration of HA and A^- present are equal.

Hydrophobic effect and the influence of lipophilicity

The original definition of the hydrophobic effect was defined as the tendency of non-polar molecules to dissolve in non-aqueous solvents by comparison to water as solvent. Recently this definition has been refined to **lipophilicity**. The measure of lipophilicity of any compound is found by measuring the water-solvent (e.g. octanol) partition coefficient, **P**, and is usually expressed as **log P** (as discussed latter).



Scheme 5.0. Decan-1-ol[2]

The partition coefficient is one of the physicochemical parameters influencing drug transport and distribution, because the drug must distribute between aqueous like phases (blood or plasma) and lipophilic like phases (cell membranes of various tissues) to be absorbed and reached the sites of action. Following Testa et al.,[3] the lipophilicity of a compound can be divided two parts, (a) a polar contribution that drives the compound into water and (b) a size effect that drives the compound into the organic phase. It is this latter effect that is now usually termed **hydrophobicity**. It should be noted that neither hydrophobicity nor lipophilicity is the same as the hydrophobic interaction, which is the tendency for nonpolar molecules to interact with each other (to associate) in water.[2]

Note that it is important to consider lipophilicity, as it is a key factor affecting how drug molecules are distributed in biological organisms. Often there is a correlation between compound (e.g. drugs) with high lipophilicity and increased biological activity, poorer aqueous solubility (expressed as the maximum amount of solute that dissolves in a given volume of solvent), increased detergency (the ability of a surfactant to remove particles from a surface)/ cell lysis, increased storage in tissues, more rapid metabolism

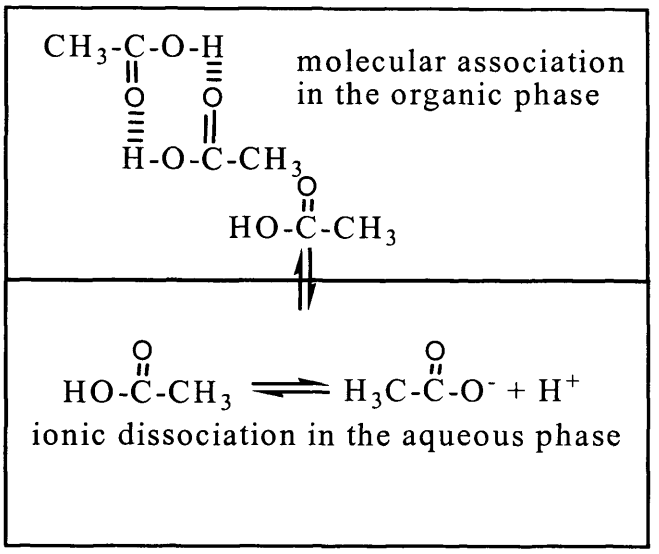
and elimination, increased rate of onset of action and also sometimes a case of shorter duration of action.

Hydrogen bonding

Most short-range interaction are caused by hydrogen bonds, this is important in biological systems and responsible for maintaining the tertiary structure of proteins (e.g. enzyme receptors).[4] Besides other weak interaction, hydrogen bonds are dependent on directional component and the hydrogen bonding strength decreases as the geometry becomes less optimal. The majority of the hydrogen bond donors are in general are: OH, SH (aromatic), NH, and activated CH groups while hydrogen bond acceptors are electronegative groups that contains of nitrogen, oxygen, sulphur and fluorine.[5]

Index of lipophilicity from partition coefficients

The partition coefficient (P) is a measure of the affinity of a molecule for a solvent phase versus that for water. This can be obtained by dividing the equilibrium concentration of a solute species in the solvent phase by the equilibrium concentration of the same solute species in the water phase. Note that it important to relate partition coefficient (P) to a particular species, because a solute may exist in more than one form in either the organic phase or in water, or even in both phases. For example carboxylic acids may be dimerised in organic phases, but remain monomeric in the water phase (see scheme 5.1). Thus, in order to calculate a partition coefficient for the species in the monomeric form, the concentration of the monomer in the organic phase has to be determined; this will not be the same as the total (formal) concentration in the organic phase. Partition coefficient is a dimension-less quantity (it is a ratio), and its logarithm (log P) is widely used to establish correlations with many biological and other physiological properties.



Scheme 5.1. Partitioning coefficient affected by two processes [2]

$$P = \frac{[C]_{solvent}}{[C]_{water}}$$ Equation 5.11

Original studies by Meyer and Overton used olive oil as the organic phase, but now octan-1-ol has been used as a standard solvent for the determination of partition and log P measurements.[6-9]

Decan-1-ol (Scheme 5.0) being like a lipid molecule has a long aliphatic alkyl chain, plus a polar functional group which gives rise to hydrogen bond accepting and donating characteristic due to the presence of lone pair electrons. Suppose P > 1 then the molecule is lipophilic. If P < 1 the molecule is hydrophilic. The log P measurement is directly carried out by the 'shake flask' method or the 'stir flask' method.[10]

Even though log P values do not vary very greatly with temperature, it is still worthwhile to maintain a record of the temperature for the system.[11] For compounds that are ionisable, the water (aqueous) phase must either be buffered to suppress ionisation, or an appropriate correction for pH can be made as shown later. If the solute associates in one or another phase, then an appropriate adjustment must be made to avoid this problem.

How the effect of partition is affected by ionisation of the solute

The measurement of the partition coefficient (P) at a certain pH is important, because for Brønsted acids and bases, the ionised form either does not partition into the organic phase, or partitions much less readily than the neutral form.

The equilibrium of compounds (i.e. drugs) between ionised and un-ionised form depends on the ambient pH and pK_a of the drug compound (i.e. drug). When pH is less than pK_a, the protonated forms HA and BH^+ predominate. When pH is greater than pKa, the depronated forms A^- and B predominate. When pH = pK_a then [HA] = $[A^-]$ and $[BH^+] = [B]$.[12]

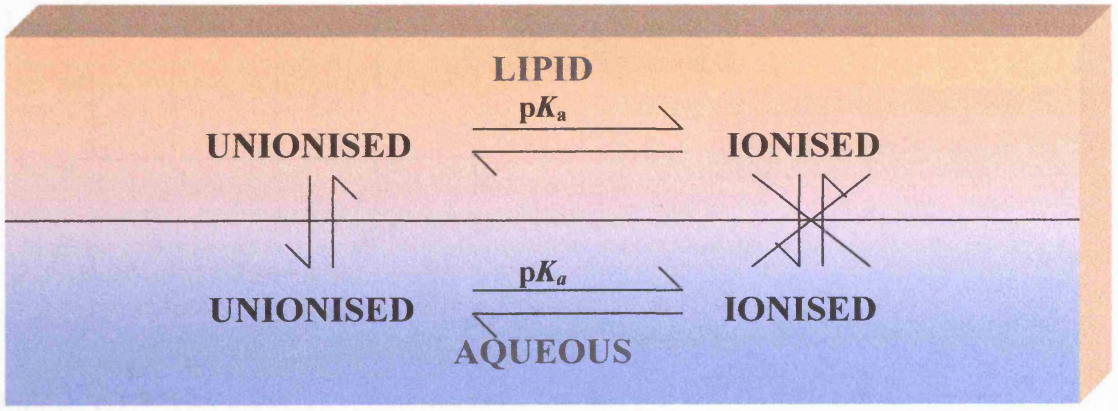


Figure 5.0. A diagram showing that lipophilicity depends on the degree of ionisation.[11]

It is important to take note of the degree of ionisation. The effective or net lipophilicity of a compound at a given stated pH is defined using the distribution coefficient log D. If the ionised form does not partition at all, then for a neutral compound (AH) will.

$$AH \rightleftharpoons H^+ + A^- \qquad \text{Equation 5.12}$$

$$D = \frac{[AH]org}{[AH]aq + [A\text{-}]aq} \qquad \text{Equation 5.13}$$

and for an organic base (BH^+).

$$BH^+ \rightleftharpoons H^+ + B$$ Equation 5.14

$$D = \frac{[B]org}{[B]aq + [BH^+]aq}$$ Equation 5.15

The pH partition diagram for weak acids is shown by a simple curve.

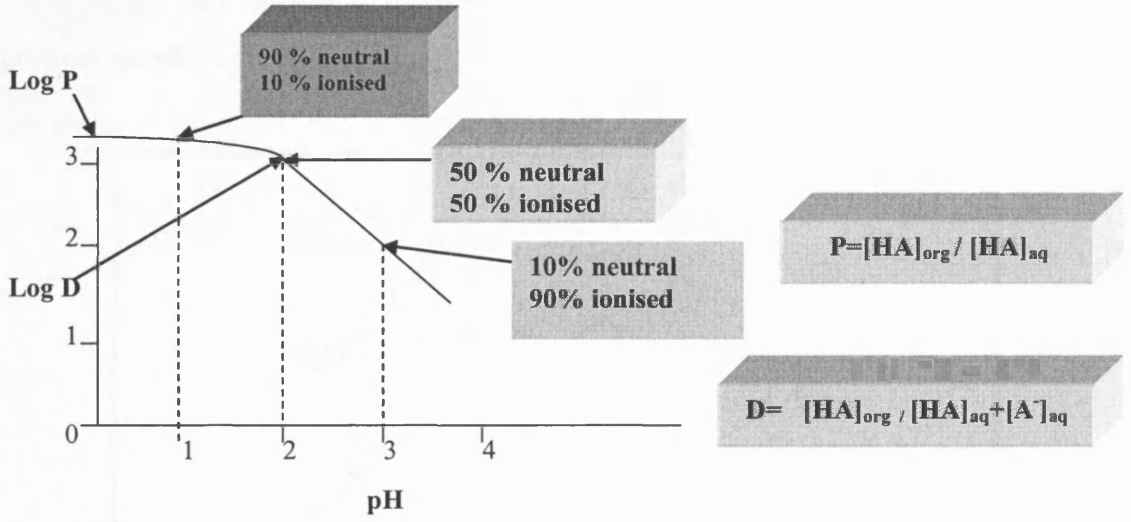


Figure 5.1.[9] The graph showing log D vs pH and the % neutral (un-ionised) form of an acidic compound of log P = 3.2 and pK_a = 2

If, as in the case of monoacids, only the neutral form partitions, the distribution coefficient and the partition coefficient are related through,

$$\text{Log D} = \log P - \log(1 + 10^{pH - pKa})$$ Equation 5.16

and for a mono-base where only the neutral form partitions.

$$\text{Log D} = \log P - \log(1 + 10^{-pH + pKa})$$ Equation 5.17

The partitioning of the charged species into the organic phase is restricted because the solute species possesses a charge.[13]

Very often log D values are carefully measured at the standard 'physiological' pH of 7.4 and are not corrected to log P values. If log P value is required, the larger the difference pH - pK_a, the larger will be the correction factor. The partitioning of zwitterions appears to be more complex. The concentrations of the anionic (-ve), the cationic (+ve), and the zwitterionic (- /+ ve) forms will depend on the pH, but it is the electronically neutral species, which co-exists with the zwitterions (- / + ve), that has the greatest lipophilicity.[14]

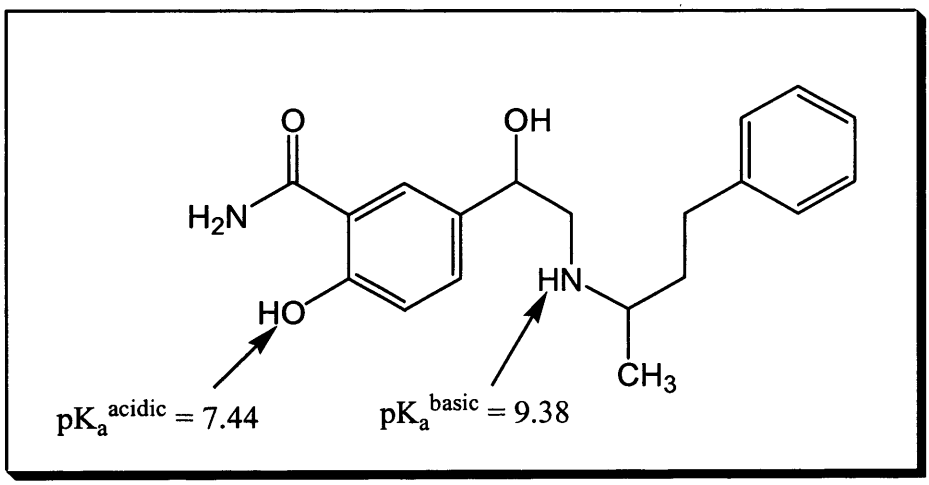
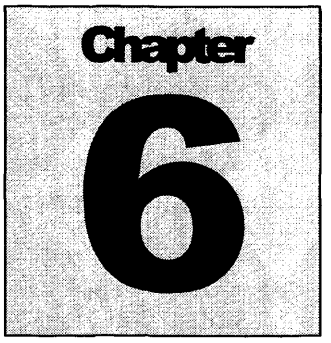


Figure 5.2. The structure of drug labetalol (beta blocker) that exists as a zwitterion.[14]

References

1. Wade, L. G. (2003) Organic chemistry 5th edition. 22-26.
2. Bailey, D. J. and Dorsey, J. G. (2001). Journal of Chromatography, **919**, 181-194.
3. Testa, B., Carrupt, P. A., Gaillard, P. and Tsai, R. S. (1996) In lipophilicity in drug action and toxicology. Ed. Pliska, V., Testa, B. and Waterbeemd, H. V. D. VCH, Weinheim.
4. Chiras, D. D. (1995). Chemistry of life. Human Biology, pp 45-46.
5. Loudon, G. M. (1988). Organic Chemistry, pp 284-287.
6. Overton, E (1901). Studien uber die Narkose, Fischer, Jena, Germany, 1901.
7. Meyer, H (1899). Arch.Exp.Pathol.Pharmakol. (Naunyn-Schmiedebergs), **42**, 109; 46 (1901) 338.
8. Meyer, K. H. and Gottlieb-Billroth, H (1920). Z.Physiol.Chem., **112**, 55.
9. Meyer, K. H. and Hemmi, H (1935). Biochem.Z., **277**, 39.
10. Dearden, J. C. and Bresnen, G. M. (1988). Quantitative Structure-Activity Relationships, **7**, 133-144.
11. Edited by King, F. D. (1994) Medicinal chemistry, Cambridge, Royal Society of Chemistry.
12. Mycek, M. J., Harvey, R. A. and Champe, C. P. (1997) Pharmacology. Lippincott Williams & Wilkins. 1-15.
13. Avdeef, A. (1996) In lipophilicity in drug action and toxicology, Ed. Pliska, V., Testa, B. and Waterbeemd, H. V. D. VCH, Weinheim.
14. Bouchard, G., Pagliara, A., Carrupt, P-A., Testa, B., Gobry, V. and G (2002). Pharmaceutical Research, **19**, 1150-1159.



Linear free-energy relationships (LFERs)

Many case studies have been carried out on the relationship between changes in reaction conditions and the result they have on rates of chemical reactions, and on chemical equilibrium. It has been found that if a series of changes in conditions or reactant structure affects the rate or equilibrium of a second reaction in a similar way to the way it affected the first reaction, there exists a linear free-energy relationship (**LFER**) between the two sets of data. This type of relationship is known as linear free-energy because at constant temperature the rate of reaction is related to the free energy of activation and the equilibrium constant is related to the standard free energy change. When relationships like this exist, they can be useful in helping to elucidate mechanism, in predicting reaction rates and the extent of a reaction at equilibrium. The best-known example of LFERs is the Hammett equation. When applied to solvation processes, LFERs are sometimes denoted as linear solvation energy relationships (**LSERs**).[1, 2]

Quantitative structure-activity relationships (QSAR)

The enormous cost involved in synthesising and testing thousands of modified compounds in an attempt to find an active drug led scientists to develop a more rational approach-called rational drug design of biologically active molecules. Scientists realised that if a physical or chemical property of a series of compounds could be correlated with biological activity, they would know what property of the drug was related to that particular activity. With this wealth of knowledge, scientists could design compounds that would have a good chance of exhibiting the desired

activity. This strategy would be a great improvement over the random approach to molecular modification that traditionally had been employed.

The first hint that a physical property of a drug could be related to biological activity, appeared almost 100 years ago, when scientists recognised that chloroform ($CHCl_3$), diethyl ether, cyclopropropane and nitrous oxide (N_2O) were all useful general anesthetics. Clearly, the chemical structures of these diverse compounds could not account for their similar pharmacological effects. Instead, some physical property must explain the similarity of their biological activities.

In the early 1960s, Cowin Hansch[3] postulated that the **biological activity** of a drug depended on two processes. The first is *distribution*: A drug must be able to get from the point where it enters the body to the receptor where it exerts its effects. For example, an anaesthetic must be able to cross the aqueous milieu (blood) and penetrate the lipid barrier of nerve cell membranes. The second process is *binding*: When a drug reaches its receptor, it must interact properly with it.

Various types of solutes can be used to see how they would distribute between water and octanol phases (for e.g. anaesthetics like chloroform, diethyl ether and cyclopropane). Octanol properties include a hydrophobic long chain and a hydrophilic polar head group. The water/octanol system is used as a model of a biological membrane in general. The anaesthetics above appear to have similar **distribution coefficients** (the ratio of the amount dissolving in octanol to the amount dissolving in water). In other words, the distribution coefficient could be related to biological activity. Drug compounds with lower distribution coefficients appear to have more polar properties and therefore could not penetrate the cell membrane. Whereas drug compounds with greater distribution coefficients appear to have non polar properties and therefore can penetrate cell membranes. The distribution coefficient of a compound could be used to determine whether a compound should be tested *in vivo* and could indicate the transport property of the drug. The technique of relating a property of a series of drug compounds to biological activity is known as a **quantitative structure-activity relationship (QSAR)**.[3]

Multiple linear regression analysis (MLRA)

Abraham and co-workers used the traditional multiple linear regression analysis to generate coefficients in various equations. This is a frequent technique used in statistics where the dependent variable (y) is linearly correlated to two or more independent variables to produce equation coefficients specific to the data set under analysis. When the coefficients are obtained, it is very desirable to predict values of y based on an independent (test) set of values of the dependent variable. The reliability of the predicted y variable is dependent on the degree of scatter (or spread) in the data. The least square method[3] is usually used for determining the best fit for the linear line through the data.

When a multiple linear regression analysis output is obtainable, it is important to assess how reliable the relationship is. The statistical methods used to do this include the standard deviation, (sd), correlation coefficient (r), t-test and the Fisher F-statistic (F).

Statistics

The standard deviation, sd

The standard deviation, (sd), is defined as the square root of the sample variance (sum of squares of deviations of individual results from the mean, divided by (n-p-1), where n is the number of data points, and p is the number of independent variables, as shown by Equation 6.0. For a simple comparison of two sets of data, p = 0, and (n-p-1) reduces to (n-1):

$$sd = (\Sigma (y_i - \bar{y})^2 / (n - p - 1))^{1/2} \qquad \text{Equation 6.0}$$

The property being measured has the same units as the standard deviation. The standard deviation measures the spread of distribution of data around the mean. A small standard deviation suggests a low spread of data with a good relationship, and a large standard deviation suggests a high spread of data with a poor relationship, thus not favourable for multi-linear relationship analysis.

Average Absolute Error, AAE

The average absolute error is sometimes used, and is a measure of the average absolute difference between predicted and measured responses. It is always smaller than the corresponding sd value, and is defined as

$$AAE = \left(\sum \left| \hat{y}_i - y_i \right| \right) / n \qquad \text{Equation 6.1}$$

$\hat{y}_i$ is the calculated (or predicted) value and y_i is the observed value (experimental data).

Root Mean Square Error, RMSE

The RMSE is the square root of the average of the squared differences between predicted and measured responses and is another measure of the prediction error of the model given in the same units.

$$RMSE = \left[\frac{\sum \left(\hat{y}_i - y_i \right)^2}{n} \right]^{1/2} \qquad \text{Equation 6.2}$$

$\hat{y}_i$ is the calculated (or predicted) value and y_i is the observed value (experimental data).

Average error, AE

Average error gives an indication of whether an equation is biased i.e. whether it systematically over- or under- predicts the desired response variable. The nearer the average error is to zero, the more unbiased is the model.

$$AE = \frac{\sum \left(\hat{y}_i - y_i \right)}{n} \qquad \text{Equation 6.3}$$

$\hat{y}_i$ is the calculated (or predicted) value and y_i is the observed value (experimental data).

The correlation coefficient, r

The 'r' term is defined as the correlation coefficient, and gives a measure of the success of the correlation of the dependent variable y against the independent variables x. Equation 6.4 shows how the correlation coefficient can be calculated:

$$r = \frac{\sum_i \{(y_i - \bar{y})(y_i - \bar{y})\}}{\{\sum_i (y_i - \bar{y})^2 \sum_i (y_i - \bar{y})^2\}^{1/2}} \qquad \text{Equation 6.4}$$

Where $\bar{y}$ is the mean observed value whereas y_i is the predicted values. The correlation coefficient (r) determines how closely the data set fits the relationship given by the multi-linear regression analysis and can range from -1 to 1. The 'r' value of –1 or 1 indicates that the data set is clarified by the correlation equation perfectly,[6] while a value of zero means there is no relationship at all between the data set and the analysis for multiple linear regression. A negative regression coefficient can sometimes often be interpreted as a poor correlation by an inexperienced eye so more often than not it is r^2 (square of the overall correlation coefficient) that is quoted in relation to multiple linear regressions.

The standard deviation (sd) and correlation coefficient (r) do not provide any mathematical evidence that any relationship observed between the dependent variable and the independent variables does not occur by chance alone. The statistical tests that can be used study the significance of the regression coefficients are the Fischer F-statistic (F) and t-test.

A normal distribution of errors is assumed by a t-test calculation that consist of a set of confidence levels, usually 95 or 99%. This allows a range of values acceptable at a specific confidence limit. T-test is performed on each every individual variable in the multiple linear regression analysis to test their significance. Occasionally not all the variables are necessary and this would be indicated by the level of significance, and so may be eliminated.

The Fisher's F-statistic, F

The F-statistic gives an indication of the quality of the regression, the higher value of F, the more significant it will be for given degrees of freedom and the better is the regression:

$$F = r^2 (n - \nu - 1) / (1 - r^2)\nu \qquad \text{Equation 6.5}$$

The r^2 is the square of the overall correlation coefficient; n is the number of data points and ν is the number of degrees of freedom ($\nu = p - 1$, where p is the number of variables). From equation 6.5 it is clear that as the number of data points and the correlation coefficient increases, so does the F value, and the larger the F the better the regression.[4-7]

The limitations of MLRA

In a multiple regression if two or more of the explanatory variables are nearly linear combinations of each other, the variables are multicollinear. The existence of such dependence implies that it is almost impossible to vary one of these variables while holding the others constant. Multicollinearity can make the regression equation unstable. Although, it may be possible to find a good multiple linear fit for the Y variable, the values of the individual coefficients may be highly variable in that unexpectedly large estimated standard errors for the coefficients of the independent variables may be produced. Thus, Y may be predicted with reasonable accuracy, but it would not be possible to draw any reliable conclusions about the coefficients. Furthermore, fitted coefficients could vary widely from sample to sample of the data, or if a single independent variable is added or deleted from the equation.

In terms of the data used for multiple linear regression analysis, both the quality and quantity needs to be taken into account. To obtain meaningful and statistically significant coefficients, a wide spread of explanatory variables is required. Furthermore, as a rule of thumb, there should be no more than one descriptor per 5 compounds in the regression. If more descriptors are used, the predictive capability of

the model is liable to be significantly affected since the extra descriptors are more likely to be describing random error in the Y variable.

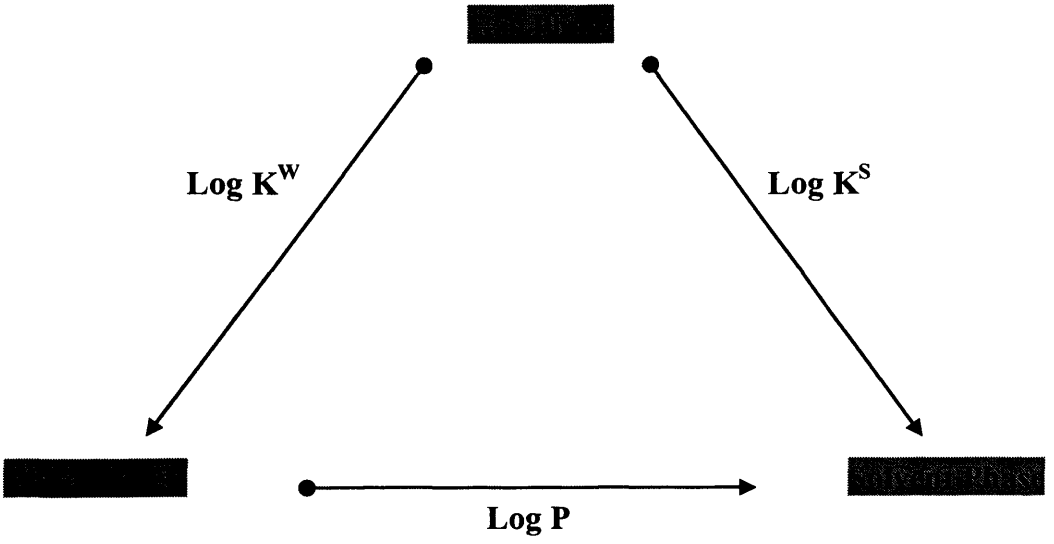
References

1. Hammett, L. P. Physical Organic Chemistry, 2nd ed, McGraw-Hill: New York, 1970.
2. Hammett, L. P (1935). Chem. Rev., **17**, 125.
3. Hansch, C. and Leo, A. (1995) Fundamental and application in chemistry and biology. Exploring QSAR, pp 172-395.
4. Solomon, R. C. (1998) Correlation and regression. Mathematics, pp 333-334.
5. Miller, J. N. and Miller, J. C. (2000) Calibration methods in instrumental analysis. Statistics and Chemometrics for Analytical Chemistry, pp 107-145.
6. Abraham, M. H. and Barker, M. (2000) QSAR, LFERs and regressions, Absolv Manuel, pp 16-17.
7. Microsoft excel help sheet (2003).

Chapter 7

The method of Abraham

The Abraham method is mainly concerned with transport-related properties of solutes.[1] These processes are defined as either equilibrium transfer **(K or P)** or the rate of transfer of a solute (***k***) between two phases. In Scheme 7.0, K^W and K^S are the gas-to-water or gas-to-solvent partition coefficients, and P is the water-to-solvent partition coefficient.



$$\mathbf{LogP = LogK^S - LogK^W}$$

Scheme 7.0. The relationship between log K (air to solvent and air to water) to give the calculated value for partition of a solute between water and the solvent phase.

A large number of transport-related processes are of importance in pharmaceutical and environmental chemistry. Equilibrium transfer is controlled by

the standard Gibbs free energy of the compound in the two phases, which in turn is related to the Gibbs free energy of solvation, $\Delta G^{\circ}s$ and $\Delta G^{\circ}w$, in a solvent and water (Equation 7.0).

$$\Delta G_s^{\circ} = -RT \ln K_s \text{ and } \Delta G_w^{\circ} = -RT \ln K_w \qquad \text{Equation 7.0}$$

From Equation 7.0 (where K or L is the gas-liquid partition coefficient or so called Ostwald solubility coefficient) and from equation 7.1, the standard Gibbs free energy of transfer (ΔG°_t) is found by subtracting ΔG°_s from ΔG°_w (Equation 7.2).

$$\Delta G_t^{\circ} = -RT \ln P \qquad \text{Equation 7.1}$$

$$\Delta G_t^{\circ} = \Delta G_s^{\circ} - \Delta G_w^{\circ} \qquad \text{Equation 7.2}$$

Accordingly **log P** can be calculated by using the following equation.[2]

$$\text{Log P} = \text{Log K}^{S} - \text{Log K}^{W} \qquad \text{Equation 7.3}$$

The general method of approach is based on the assumption that certain properties (or descriptors) of a given solute, will be important in solvation of the compound in different solvent phases. In addition, the solute properties will be important in determining the transfer of solute from one phase to another. Thus if the polarisability[3] for example, of a compound is an important factor in solvation in ethanol, we expect it to be relevant as regards solvation in hexane, or dichlorobenzene, or other solvents. Naturally, a compound property such as acidity will influence solubility in basic solvents, but will be redundant in non-basic solvents.

It is suggested that similar solute properties will be of importance in the solvation of the solute, not just in one particular solvent phase, but in solvent phases in general, and therefore similar solute properties will be important factors in the transfer of a solute between two phases.

The aim of Abraham and co-workers was to attempt to identify the general solute properties expected to be relevant to transport processes, and to identify a set of solute descriptors, that could be used in algorithms for transport processes. Hansch equations have been very important in this context, but they are restricted to sets of related solutes, whereas the Abraham equation should be applicable to varied sets of solutes (or compounds).

The cavity model of solvation

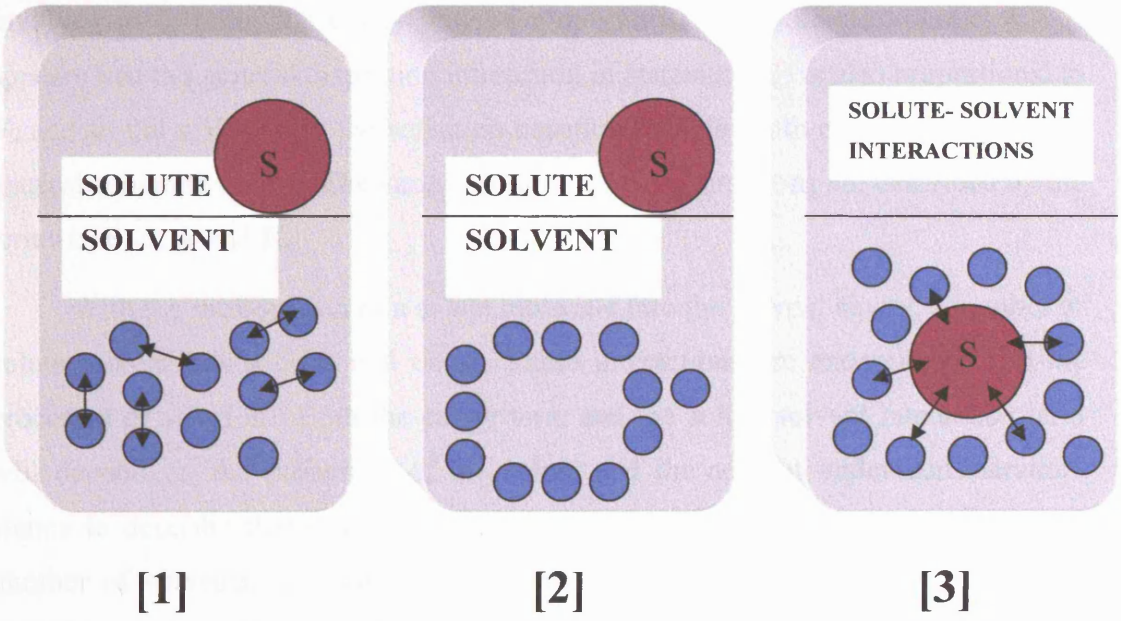


Figure 7.0. The cavity model of solvation involves three steps.

The Abraham solvation equation is based on a simple solvation model used by Abraham, Kamlet and Taft (**AKT**),[4-8] see Figure 7.0. This model describes the solvation of a gaseous solute into a liquid solvent and can be broken down into the following stages:

1) A cavity of suitable size (i.e. same size as the solute) is made in the solvent, this involves the endoergic breaking of solvent-solvent bonds and hence the Gibbs free energy change will be positive (ΔG = +ve). This is energetically unfavorable.

2) The bulk solvent molecules are reorganised into their equilibrium position around the cavity, for which the Gibbs free energy change is assumed to be negligible.
3) The solute is inserted into the cavity, where various solute-solvent interactions all take place. This process is exoergic and liberates energy (ΔG = -ve). This is energetically favorable.

Note that that the solvent phase is constant. The only variable that changes is the solute and hence the process can be described by solute properties. The work involved in forming the cavity will be proportional to the solute volume, **V**. It appears that the general dispersion interaction in statement (3) is also proportional to **V**, and so the v.**V** term in the solvation equation includes both cavity effect and the general solvation effect. The various solute-solvent interactions are described by the terms in **E**, **S**, **A** and **B**.

With the introduction of a solute molecule into the solvent cavity, a number of solute-solvent interactions will occur. These interactions are exoergic and aid the processes of solution. Both the cavity term and the solute-solvent interaction term will depend on the properties of the solute and the solvent under consideration. Hence to describe these effects, for the general case of a number of solutes in a number of solvents, it is necessary to construct a linear free energy relationship (**LFER**) equation, that includes the relevant properties of both the solutes and the solvents. The Hammett equation[9] is the best-known example of a LFER and his work has formed a sound basis for many other studies to be carried out involving LFERs. Considerable simplification is possible for the general case of a number of solutes in a number of solvents, by considering only one solvent, then the properties of the solvent remain constant, and it is necessary only to consider properties or the descriptors of the solute.

The general solvation equation

Abraham, Kamlet and Taft pointed out the necessity to consider both non-specific and specific solute/solvent interactions separately. Their equation has the following general form:

Solute Property = constant + cavity term + dipolarity/polarisabilty--term + hydrogen bonding term Equation 7.4

As Kamlet-Taft solute parameters were partly based on the solvent properties, Abraham and co-workers developed a new dipolarity/polarisability scale, **S** and new hydrogen bond acidity and basicity parameters, **A** and **B** respectively.[10-12] McGowan characteristic volume[13,14] was also represented as symbol **V**. The newly devised descriptors were combined linearly to give the following equations.[15]

$$SP = c + e.E + s.S + a.A + b.B + v.V \quad \text{Equation 7.5}$$

$$SP = c + e.E + s.S + a.A + b.B + l.L \quad \text{Equation 7.6}$$

Important applications of the general solvation equation in drug design are discussed in latter chapters.

SP is a biological or chemical property of a series of solutes in a system, usually as log of some properties, although sometimes other dependent variables are used, for example, the % intestinal absorption of drugs.[16] Mostly, however, distribution or partition is considered as in blood-brain distribution[17] or the partition of compounds in a particular water solvent system. The equation constant is defined by '**c**.' The equation descriptor coefficients (i.e. **e, s, a, b** and **v** or **l)** reflect on the properties of the system phase under investigation and are obtained by multiple linear regression using Microsoft excel program. The old and new notation for solute descriptors is shown Table 7.0.

SYMBOL NEW	SYMBOL OLD	SOLUTE DESCRIPTORS
E	R_2	The excess molar refraction in units of $(cm^3\ mol^{-1})/10$, which reflects solute polarisabilty. It provides a quantitative measure of the ability of a solute to interact with a solvent through n- and π- electron pairs.
S	π^H_2	The solute dipolar/ polarisabilty parameter. This measures the ability of the solute to stabilise a charge or dipole.
A	$\Sigma\alpha^H_2$	The overall hydrogen bond acidity parameter. A measure of hydrogen bonding by the solute in a basic solvent.
B	$\Sigma\beta^H_2$	The overall hydrogen bond basicity parameter. A measure of the extent of hydrogen bonding by the solute in an acidic solvent.
L	$\log L^{16}$	The logarithmic of the gas-hexadecane partition coefficient (or the Ostwald solubility coefficient) at 298K. Accounts for cavity size and dispersion interactions.
V	Vx	McGowan's characteristic volume in units of $(cm^3\ mol^{-1})/100$ Represents the three-dimensional space occupied by the solute.

Table 7.0. Abraham solute descriptors used today.[18]

The excess molar refraction, E (R_2)

The polarisability-correction descriptor (δ_2) used by Kamlet and Taft is only an empirical factor limited to one of three values, 0.5 for halogenated aliphatics, 1.0 for aromatics or 0.0 for all other compounds. A number of possible alternatives for δ_2 were considered by Abraham,[19] solute molar refraction ($\mathbf{MR_x}$) being one of them. This is latter defined as:

$$MR_X = 10[(n^2-1)/(n^2+2)]\ V \qquad \text{Equation 7.7}$$

Here **n** is the refractive index of a liquid solute taken at 293K with the sodium-D-line, and **V** is the McGowan's characteristic volume in $(cm^3mol^{-1})/100$.[13,14] For solids, the refractive index of the hypothetical liquid at 20°C can be also calculated by using the ACD software.[20] The $\mathbf{MR_x}$ parameter was found to be of little use.

Because of the presence of the volume term in molar refraction, this latter always increases with increasing size, and is very co-linear with volume itself. The refractive index function itself is an indication of the polarisable electrons for a molecule that is either aromatic or halogenated aliphatic compounds. However, an excess molar refraction can be defined as the solute molar refraction less the molar refraction of an alkane of the same characteristic volume. The parameter **E** ($10^{-1}cm^3mol^{-1}$) is given by;

$$E = MR_X(\text{observed}) - MR_X(\text{for alkane of the same V}) \qquad \text{Equation 7.8}$$

The advantage of **E** is that it is almost independent of volume and also provides a quantitative indication of polarisable **n** and **JI** electrons. **E** is an almost additive quantity that can easily be estimated for solids and for structures in general (from molecular fragment or substructure values for any compound).[21]

Solute dipolarity/polarisability scale S (Π_2^H)

Initially π_2^* was taken as the solvent parameter π_1^* for non-associated liquids set out by Kamlet and Taft.[4,8] Where π_1^* is experimentally accessible only for compounds that are liquid at 298 K, values of π_2^* had to be estimated for associated compounds such as acids, phenols, alcohols and amides as well as gaseous and solid solutes. Also in addition, there is present the inherent idea that π_1^* is the same as π_2^* for non-associated liquids, but this may not always be the situation. In addition, because of its spectroscopic origin, this parameter fails to be Gibbs energy related. It therefore seemed necessary to develop a method that would allow the determination of a dipolarity / polarisability scale, (π_2^H), that would be free energy related and include all types of solute molecule.

Abraham and his co-workers constructed the new dipolar / polarisability parameter, π_2^H (or descriptor **S** as known now), from extensive sets of gas liquid chromatographic (GLC) data. This provided π_2^H (S) values for hundreds of solutes. The π_2^H (S) values for halogenated or polyhalogenated solutes were again obtained by the same method using retention data for various other stationary phases.[22]

Hydrogen bond acidity of a solute, A

A new hydrogen bonding acidity scale (α^H_2)[23] was set up by Abraham, using log $K_{1:1}$ (equilibrium constants) for 1:1 complexation of a series of monomeric acids (e.g. pyridine) against a given reference base in tetrachloromethane (CCl_4) at 298K. The use of equilibrium constants means the scale is Gibbs free energy related. If the acid and base are present in low concentrations, both will be in monomeric and non-associated forms and will undergo interactions as shown by Figure 7.1.

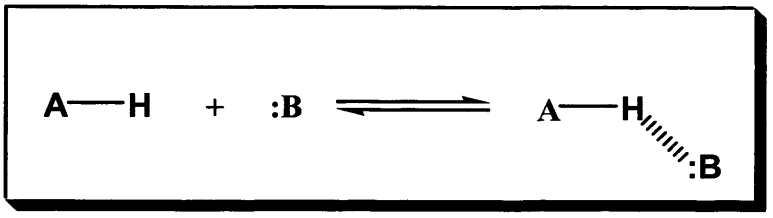


Figure 7.1. Hydrogen-bond complexation reaction.

Abraham and co-workers obtained forty-five equations in which log $K_{1:1}$ is the dependent variable (see equation 7.9). The value of log K_A^H describes the acid, and L_B and D_B characterise the base.

Log $K_{1:1}$ (series of acids against reference base B) = L_B log K^H_A + D_B Equation 7.9

The general hydrogen-bond acidity scale was set up by plotting a series of acid log K (against reference base) verses a series of log K (against any other reference base), to give rise to a series of straight lines, with an intersection point at log K^H_A = log K^H_A = -1.1 (equilibrium constant expressed in molar concentration units). Different types of log $K_{1:1}$ plots must show family-independent behaviour, so that it is possible to obtain the average hydrogen bond acidity (with some exceptions) for solutes in tetrachloromethane (CCl_4), given as Log K^H_A. The origin was then shifted from -1.1 to 0 and compressed by converting Log K^H_A to α_2^H in order that all compounds with zero hydrogen bond-acidity have a α_2^H value of zero.

$\alpha_2^H = (\log K^H_A + 1.1)/4.636$ Equation 7.10

The empirical value of 4.636 is used to compress the acidity scale into a suitable working range.

Within a bulk solvent the solute can form numerous hydrogen bonds with the neighbouring molecules, therefore the 1:1 complexation no longer applies and the α_2^H values may not be relevant and a scale of solute overall or effective hydrogen-bond acidity descriptor (**A**) was needed. To obtain **A** values a initial version of equation 7.6 was set up, using α_2^H as the hydrogen-bond acid descriptor and was applied to various water-solvent partitions coefficient (log SP).[24] The descriptor, α_2^H was then tailored where necessary, in order to obtain the effective value, **A**. The new set of equations was then constructed, and the same process repeated until a self-consistent set of equations and **A** values was given. Whilst solvent molecules surround the solutes in the water-solvent partitions, the overall hydrogen-bond acidity scale is the actual scale required. It was observed that values of **A** were constant along any homologous series, except perhaps for the first one or two members, so once a few values are found, values for the remainder of the homologous series can be established. The multiple hydrogen bonding of solute gives rise to a higher **A** value than α_2^H, obtained from a simple complexation constant.

Note that the work carried out to obtain on the whole **A** scale from the 1:1 α_2^H scale proceeded side-by-side with the calculation of **A** that took place during the **S** calculation. This suggests that a regular check had to be made on the self-consistency of the derived **A** values. It is important to clarify that hydrogen bond acidity, which indicates the ability of a compound to donate a hydrogen bond is not related to Bronstead acidity of a compound, which refers to loss of a proton. Table 7.1 indicates solutes with hydrogen bond donors. It can be seen from their A values that ethanol and pyrrole both have similar ability to participate in the hydrogen bond interaction.

SOLUTE	α_2^H	A
Alkanes	0.00	0.00
Triethylamine	0.00	0.00
Ethanol	0.33	0.37
Pyrrole	0.41	0.41
Water	0.35	0.82
Acetic Acid	0.55	0.61
Phenol	0.60	0.60

Table 7.1. The solute comparison between α_2^H and A values.[24]

Hydrogen bond basicity for solute, B

Abraham, Kamlet and Taft[8,25] were the first to propose the β_2 descriptor (the solute basicity parameter), which was derived from solvent β_1 values. The new β^H_2 scale was established by Abraham et al,[12,26] using 1:1 hydrogen-bond complexation constants in CCl_4 (tetrachloromethane), β^H_2. The hydrogen-bond basicity scale was established by plotting log K values of series of bases (against a given reference acid) versus a series of bases (against other acids). These plots gave straight lines passing through log K = -1.1, and allowed the average hydrogen-bond basicity, log K^H_B for solutes in CCl_4 to be obtained. These values were then converted into the **B** scale by applying equation.

$$\beta^H_2 = (\log K_B^H + 1.1)/4.636 \qquad \text{Equation 7.11}$$

The empirical factor 4.636 was selected so that $\beta^H_2 = 1.00$ for the hydrogen bond base hexamethylphosphortriamide, and has no physical importance other than to serve as a suitable working range of **B** values.

As for the acidity term, there is no guarantee that β^H_2 is the suitable scale to use when a solute is surrounded by the solvent species.[27] Though with few exceptions β^H_2 can be used as **B** for mono-bases. However, in the case of poly-bases,

there seems to be no alternative than to calculate **B** by back-calculation. Table 7.2 shows a comparison of β^H_2 and **B** values for some solutes.

An additional hydrogen bond basicity term was introduced by Abraham, ΣB^0, for which solutes such as pyridines, anilines, sulfoxides, and some heterocyclic compounds in water/solvent partitioning systems, where the solvent phase is quite aqueous. The latter include n-octanol, diethyl ether, di-n-butyl ether, ethyl acetate and n-butyl acetate. Whereas, for non-aqueous phases such as benzene, alkanes, chloroform and the gas phase, the original **B** term can be used for all solutes.

Solute	β^H_2	B
Alkanes (e.g. n-Hexane)	0.00	0.00
Dichloromethane	0.05	0.05
Benzene	0.15	0.14
Phenylcyanide	0.42	0.33
Diethyl ether	0.45	0.45
Ethanol	0.44	0.48
Pyridine	0.62	0.52
Dimethyl sulfoxide	0.78	0.88
Phenol	0.22	0.30
4-Methoxyaniline	0.45	0.65

Table 7.2. Comparison between β^H_2 and B values.[24]

Descriptor L or Log L^{16}

From the stage one in the cavity theory of solvation, the larger the solute the larger will be the cavity. However from stage 3, the larger the solute, the larger will be its tendency to take part in solute-solvent interactions of the general London dispersion type. A useful measure that includes a cavity effect and general London dispersion effect which is related to the log L^{16} Ostwald solubility coefficient, which was defined as solute gas-hexadecane partition coefficient at 298K.[28]

L_{16} = (Conc. of solute in hexadecane)/(Conc. of solute in gas phase) Equation 7.12

Hexadecane was chosen by Abraham and co-workers,[28] as a reference solvent, as it is a readily available non-polar liquid of well-defined structure. From this **L** can be readily obtained from GLC measurements. For volatile organic compounds, the **L** descriptor can be obtained by direct GLC measurements using packed columns coated with hexadecane at 298K and a number of n-alkanes used as standards. The **L** descriptor of a test solute can then be determined from its retention time.

The remaining four descriptors, **E**, **S**, **A** and **B**, can be regarded as measures of tendency of a solute to undergo various solute/ solvent interactions, all of which are energetically favourable, i.e. exoergic. Where as, **L** and **V** are both a measure of the size of a solute, so will be a measure of the cavity term. Additionally since the size of the solute is related to general dispersion interaction, both **L** and **V** describe solute/ solvent dispersion interactions.[22]

The McGowan characteristic volume V (V_x)

The cavity term **V** (in $cm^3\ mol^{-1}/100$)[13,28] was chosen by Abraham because it is so straightforward to calculate by simple summation of bonds and atoms found in the molecule. The method chosen by Abraham has replaced Leahy's computer program, which was used previously to calculate the intrinsic volume. The bonds in a solute are all considered equal, so that a double bond such as **C=O** or **C=C** or a triple bond such as **C≡C** are regarded as 'one bond'. An algorithm proposed by Abraham[13], allows the number of bonds in a molecule to be obtained simply;

$B = N - 1 + R$ Equation 7.13

B represents the number of bonds, **N** is the total number of atoms and **R** is the total number of ring structures. **V** is defined as,

$V = (\sum \text{atom contributions} - (6.56 \times B))/100$ Equation 7.14

Values for atom contributions required for the calculation of **V**, are given in Table 7.3.

C = 16.35	**P** = 24.78	**Se** = 27.81
N = 14.39	**S** = 22.91	**Br** = 26.21
O = 12.43	**Cl** = 20.95	**Sn** = 39.35
F = 10.48	**B** = 18.32	**Sb** = 37.74
H = 8.71	**Ge** = 31.02	**Te** = 36.14
Si = 26.83	**As** = 29.42	**I** = 34.35

For each bond between atoms, 6.56 $cm^3.mol^{-1}$ is to be subtracted.

Table 7.3. Atom contributions for calculation of **V** (in cm^3 mol^{-1})

The coefficients used in Abraham equations

As the descriptors correspond to the solute effect from the various solute-solvent interactions, the coefficients encode the interaction properties of the solvent phase. For a transfer between two phases: the coefficients indicate the difference in interaction properties of the two phases (see table 7.4):

SYMBOL	INTERPRETATION OF THE COEFFICIENTS
e	Tendency of the solvent to interact through π and σ electron pairs. The value of 'e' is usually positive but may be negative for phases that contain fluorine atoms.
s	Tendency of the phase to interact with polarisable/ dipolar solutes.
a	Measure of the hydrogen bond basicity of the phase (because acidic solutes will interact with basic phases).
b	Measure of the hydrogen bond acidity of the phase (because basic solutes will interact with acidic phases).
l	Combination of exoergic dispersion forces that make a positive contribution to the l-coefficient and an endoergic cavity term that makes a negative contribution. The dispersion interaction usually dominates, so that the l-coefficient is positive, except for solution of gases and vapours into water.
v	Measure of phase hydrophobicity, resultant of dispersion and cavity effects.

Table 7.4. Interpretation of the coefficients used by the Abraham solvation equations.

In a gas-to-solution processes the s-, a- and b-coefficients must always be positive (or zero), because of the interactions between the phase and a solute will increase the solubility of a gaseous solute. Whereas this e-coefficient is an exception, because it is tied to hydrocarbons as a zero; hence fluoro compounds as phases may give rise to a negative e-constant. The coefficients in the solvation parameter equation are therefore not just fitting constants but must obey general chemical principles.

As an example of how the coefficient of the general equation can be used to define the system: SP = log P_{oct} (octanol-water partition coefficient).[18]

$$\text{Log } P_{oct} = 0.088 + 0.562\mathbf{E} - 1.054\mathbf{S} + 0.034\mathbf{A} - 3.460\mathbf{B} + 3.814\mathbf{V} \qquad \text{Equation 7.15}$$

$$n = 613,\ r^2 = 0.9974,\ SD = 0.116 \text{ and } F = 23162$$

The **e**-coefficient has a value of 0.562, although its effect will not be as significant as the other interactions, solutes with larger excess molar refraction value will slightly favour the organic phase. A negative value of **s** (-1.054), implies that the more polar the molecule, the more it will favour the aqueous phase. The hydrogen bond acidity of the compound (hydrogen bond basicity of octanol) does not influence the partitioning, however, hydrogen bond bases will strongly favour water (**b** = -3.46). With a significant value of 3.814, the **v** coefficient shows that the larger the solute, the more it will be attracted to octanol.[18]

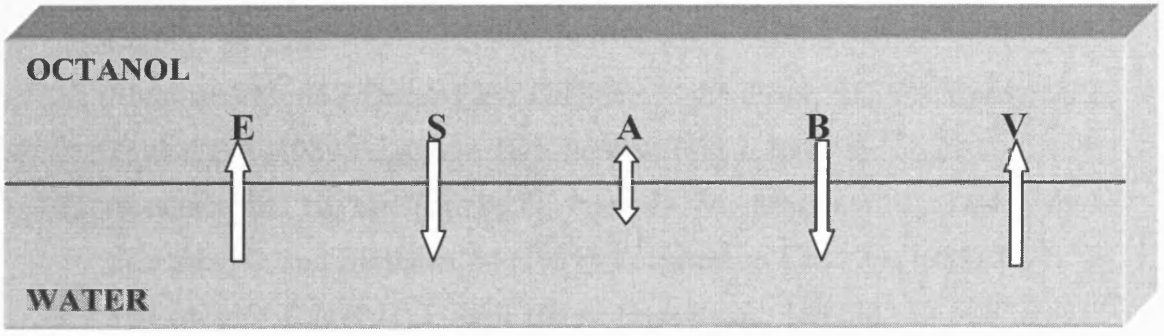


Figure 7.2. Properties of the octanol-water system

References

1. Abraham, M. H. and Platts, J. A. (2001). J. Org. Chem. **66**, 3484-3491.
2. Abraham, M. H., Chadha, H. S., Martins, F., Mitchell, R. C., Bradbury, M. W. and Gratton, J. A. (1999). Pesticide. Sci. **55**, 78-88.
3. Lamarche, O., Platts, J. A. and Hersey, A. (2001). Phys. Chem. Chem. Phys. **3**, 2747-753.
4. Kamlet, M. J., Doherty, R. M., Abboud, J-L. M., Abraham, M. H. and Taft, R.W (1986). Chem. Tech, 566.
5. Kamlet, M. J., Doherty, R. M., Abraham, M. H. and Taft, R.W (1986). Chem. Br., **22**, 551.
6. Kamlet, M. J., Abboud, J-L. M. and Taft, R.W (1977). J. Am. Chem. Soc., **91**, 8325.
7. M. H. Abraham, R. M. Doherty, M. J. Kamlet and R. W. Taft (1986). Chem. Br., **22**, 551.
8. Kamlet, M. J., Doherty, R. M., Abboud, J-L. M., Abraham, M. H. and Taft, R. W., Carr, P. W (1987). J. Phys. Chem, **91**, 1996-2004.
9. Hansch, C. and Leo, A. (1995) Fundamental and application in chemistry and biology. Exploring QSAR. 1-22.
10. Abraham, M. H., Whiting, G. S., Doherty, R. S. and Shuely, W. J (1991). J. Chrom., **587**, 213.
11. Abraham, M. H., Duce, P.P., Grellier, P. L., Prior, D. V., Morris, J. J., Taylor, P. J (1989). J. Chem. Soc. Perkin Trans 2, 699.
12. Abraham, M. H., Grellier, P. L., Prior, D. V., Morris, J. J., Taylor, P. J., Laurence, C. and Berthelot, M (1989). Tetrahedron Lett., **30** (1989) 2571.
13. McGowan, J. C (1984). J. Appl. Chem. Biotechnol., **34A**, 38.
14. Abraham, M. H. and McGowan, J. C (1987). Chromatographia, **23**, 243.
15. Abraham, M. H (1993). Chem. Soc. Revs., **22**, 73.
16. Zhao, Y. H., Le, J., Abraham, M. H., Hersey, A., Eddershaw, P. J., Luscombe, C. N., Boutina, D., Beck, G., Sherborne, B., Cooper, I. and Platts, J. A. (2001). J. Pharm. Sci. **90**, 749-784.

17. Platt, J. A., Abraham, M. H., Zhao, Y. H., Hersey, A., Ijaz, L. and Butina, D. (2001). J. Med. Chem. **36**, 719-730.
18. Abraham, M. H. and Barker, M (2000). Sirius solute property prediction manual. 19-30.
19. Abraham, M. H., Whiting, G. S., Doherty, R. M. and Shuely, W. J (1990). J. Chem. Soc. Perkin Trans 2, 1451.
20. ACD Software Chem. Sketch (version 6.00), 2002 edition.
21. Platts, J. A., Butina D., Abraham, M. H. and Hersey, A. (1999). J. Chem. Inf. Comput. Sci. **39**, 835-845.
22. Kamlet, M. J., Abboud, J.-L.M., Abraham, M. H. and Taft, R. W (1983). J. Org. Chem. **48**, 2877.
23. Abraham, M.H. (1993). Chem. Soc. Rev. **22**, 73-83.
24. Abraham, M. H., Kamlet, M. J., Dotherty, R. M., Abboud, J. L. M. and Taft, R.W. (1986). J. Pharm. Sci.**75**, 338-349.
25. Abraham, M. H., Grellier, P. L., Prior, D. V., Morris, J. and Taylor, P. J (1990). J. Chem. Soc. Perkin Trans. 2, 521.
26. Norinder, U. and Haeberlein, M. (2002). Advanced Drug Delivery Rev. **54**, 291- 313.
27. Abraham, M. H., Grellier, P. L. and McGill, R. A (1987). J. Chem. Soc. Perkin Trans. 2, 797.
28. Platt, J. A., Valko, K. and Abraham, M. H. (1998). J. Chromatography A. **803**, 51-60.

Chapter

8

Methods of determination of descriptors

One of the major parts of the present work is the determination of descriptors for the variety of compounds, especially drugs, for which blood-tissue distribution are available. The Abraham descriptors **E** can be obtained from the refractive index in conjunction with structural information, as discussed before in chapter 7, and **V** can always be calculated. There are a number of ways of obtaining values for descriptors (**S, A** and **B**), whether there are experimental values available (i.e. log P for partition of a solute in a given solvent system) or not. The experimental methods include RP-HPLC measurements on various stationary phases, water/solvent partitions, and even solubility data. In recent studies, Valko[1-3] and co-workers have used a fast gradient elution RP-HPLC in order to obtain Abraham descriptors by a reasonably rapid procedure. Valko (et al) used 9 solvation equations for the different HPLC systems to obtain the Abraham descriptors. They have set up their method by choosing the most orthogonal HPLC systems by non-linear mapping.[4] Zissimos et al have measured drug compounds to obtain log P values using the GlpKa instrument, utilising potentiometric methods developed by Sirius Analytical. This instrument used is designed specifically to determine ionisation constants (pK_a) and partition coefficients (log P) of weak proton acids and bases in various water-solvent systems.[5] The Abraham descriptors **S, A** and **B** were obtained by using the 'Solver' method for four equations corresponding to four water-solvent system. The systems were chosen so that the coefficients for the equations are as different as possible. The solvents used in Zissimo's log P method are: octanol, chloroform, cyclohexane and toluene.

When no experimental data are available, methods based on a fragment addition approach (**ABSOLV** as developed by Pharma Algorithm[6-7], **UNIX** based on

structure developed by Platts[8-9] and estimation by analogy) can give a quick estimation of a compound's descriptors based on its 2D structure. Accurate estimation of descriptor values (**S, A, B**) can be obtained by using experimental data in a program called **Solver** (a function of Microsoft Excel). This chapter is an overview of the methods available to obtain descriptor value for small compounds like VOCs and large compounds such as drugs or pesticides.

Absolv

Absolv is a computer program that has been developed by Pharma Algorithms. The program calculates various solvation-associated properties from Abraham type equations and predicts Abraham's solvation parameters necessary for such calculations. The parameters calculated are:

- Excess molar refraction - **E**
- Polarity/polarisability - **S**
- Hydrogen bonding acidity parameter - **A**
- Hydrogen bonding basicity parameter - **B**
- McGowan volume - **V**
- Partition coefficient between gaseous phase and hexadecane - **L**

These parameters have been discussed in detail in chapter 7 (the method of Abraham). The advantage of using Absolv is that for a given calculation for a given solute, it also lists up to 5 most similar structures from the Absolv database (experimental data for compounds found at UCL) along with their experimental values of the Abraham parameters and literature references. Absolv also allows one to calculate various solvent-solvent or gas-solvent partition coefficients from more than 100 predefined Abraham type equations, using predicted Abraham solvation parameters of a solute. The equation manager contained within the program allows one to define a new equation as well as view existing equations. Finally, Absolv has a fully searchable database, containing a list of more than 5000 compounds providing their names, CAS registry numbers, literature values of Abraham solvation

parameters and references. This data compilation includes training and test sets used in building and evaluating Absolv predictive algorithms.[6-7]

Unix

The first fragmentation algorithm was devised by Platts et al[8] who introduced an additive model (UNIX method) for the estimation of Abraham's molecular descriptors **E**, **S**, **A**, **B**, and **V**. This model was developed from a set of 81 atomic and functional group fragments and intra-molecular interactions for which an evaluation of their contribution to each descriptor was carried out through a process of multiple linear regressions. They proceeded to apply this group contribution model to sets of molecules for which partitioning data was predicted using Abraham's solvation equations.[9] The method gives good results for predicting the molecular descriptors in question and also produces partitioning data for a number of compounds.

Abraham and Platts[10] have set out characteristic parameters (s**A** and s**B**) for common functional groups, as given in Table 8.0 and 8.1. These are really fragment constants, but for functional groups, rather than the very small fragments considered by Platts et al. These 'structural' constants can be used very simply to estimate the effect of a substituent on overall **A** and **B** values. The main advantage of this method is that it can be set as a high throughput automated system, estimating descriptors for approximately 600 molecules per minute. The descriptors can be estimated from molecular structure entered as a SMILES string.[9]

Substituent	sE	sS	sA	sB	sV
-H	0.00	0.00	0.00	0.00	0.0000
-F	0.03	0.35	0.00	0.10	0.0177
-Cl	0.22	0.40	0.00	0.10	0.1224
-Br	0.37	0.40	0.00	0.12	0.1750
-I	0.63	0.40	0.00	0.15	0.2582
-OMe	0.06	0.25	0.00	0.44	0.1996
-CHO	0.19	0.65	0.00	0.45	0.1566

-CN	0.19	0.90	0.00	0.36	0.1547
$-NH_2$	0.23	0.35	0.16	0.61	0.0998
$-NO_2$	0.24	0.95	0.00	0.31	0.1742
-SH	0.39	0.35	0.00	0.24	0.1635
-OH(1^o)	0.24	0.42	0.37	0.48	0.0587
$-CO_2H$	0.21	0.62	0.60	0.45	0.2153
$CONH_2$	0.42	1.30	0.56	0.68	0.2564

[a] Derived from propyl-H, where E = 0.00, S= 0.00, A= 0.00, B= 0.00 and V= 0.5313.

Table 8.0. Aliphatic functional group structural constants[a]

Substituent	sE	sS	sA	sB	sV
-H	0.00	0.00	0.00	0.00	0.0000
-Me	-0.01	0.00	0.00	0.00	0.1409
-F	-0.13	0.05	0.00	-0.04	0.0177
-Cl	0.11	0.13	0.00	-0.07	0.1224
-Br	0.27	0.21	0.00	-0.05	0.1750
-I	0.58	0.30	0.00	-0.02	0.2582
-OMe	0.10	0.23	0.00	0.15	0.1996
-CHO	0.21	0.48	0.00	0.25	0.1566
-CN	0.13	0.59	0.00	0.19	0.1547
$-NH_2$	0.35	0.44	0.00	0.27	0.0998
$-NO_2$	0.26	0.59	0.00	0.14	0.1742
-SH	0.39	0.28	0.09	0.02	0.1635
-OH(1^o)	0.20	0.37	0.60	0.16	0.0587
$-CO_2H$	0.12	0.38	0.59	0.26	0.2153
SO_2NH_2	0.52	1.03	0.55	0.66	0.3807
$CONH_2$	0.38	0.98	0.35	-0.05	0.9728

[a] Derived from phenyl-H (benzene), where E = 0.61, S= 0.52, A= 0.00, B= 0.14 and V= 0.7164.

Table 8.1. Aromatic functional group structural constants[a]

Estimation of descriptors based on the analogy method

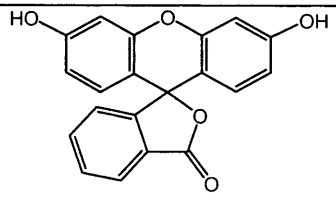
The descriptors **S**, **A** and **B** are approximately constant for a given functional group (see table 8.0/8.1). They can be estimated by comparison with known descriptors of similar compounds. In addition, **S**, **A** and **B** are usually constant along a homologous series. This method can provide a rapid estimation of the descriptor values but is more time consuming than automatic programs like Absolv and UNIX. However, both the latter methods do not take into account all intra-molecular interactions such as hydrogen bonding. Hence, estimating the descriptors manually by analogy may lead to more reliable values.

Solver descriptor calculation

Solver is a tool in Microsoft Excel, which can be used to determine the maximum or minimum value of one cell by changing other cells. The LFER equations are entered in a Microsoft Excel spreadsheet, as well as the corresponding measured/literature data. Solver minimises the sum of squares on the required equations to fit the targeted cells **S**, **A** and **B** and the values are accepted when the overall sum of squares reaches a minimum. Solver uses the Generalised Reduced Gradient (GRG2) non-linear optimisation code developed by Leon Lasdon, University of Texas at Austin, and Allan Waren, Cleveland State University.

As an example, the descriptors for fluorescein **S**, **A** and **B** can be calculated by using the Excel program called Solver. The descriptor **E** and **V** can be calculated by using a VR calculator program based on equations discussed in chapter 7. From table 8.2, **S**, **A** and **B** for fluorescein can be calculated based on seven experimental values for log P measured by Ricardo Sanchez at UCL. It was found that the system containing 7 equations of different water/solvent system could be used to minimise and optimise the three Abraham descriptors accurately, without any other combinations. This is why it is necessary to use as many systems as possible. In addition, the system coefficients have to be as different as possible, in order to obtain good fits by the solver method. In the present case, it was found that calculated values for log P were very close to experimental values, that suggest the fits for descriptor **S**, **A** and **B** are the best values for fluorescein. The descriptors for

fluoroscein are: **S = 3.28**, **A = 1.41** and **B = 1.00**. The standard deviation between observed and calculated log P is 0.172. Finally, I would like to thank Ricardo Sanchez for the permission to use his results.

	**Log P values**	
Solvent	**Observed values**	**Calculated values**
Octanol	3.367[a]	3.333
Dichloromethane	1.472[a]	1.224
Trichloromethane	0.977[a]	0.925
1,2-Dichloroethane	1.535[a]	1.569
Toluene	0.409[a]	0.344
Chlorobenzene	0.647[a]	0.544
Bromobenzene	0.652[a]	0.962
Solver descriptors	**E=2.75 S=3.28 A=1.41 B=1.00 V=2.25 Standard deviation = 0.172**	

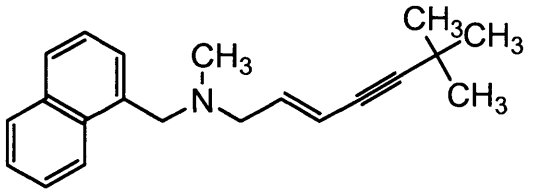
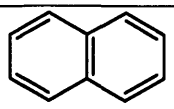
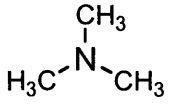
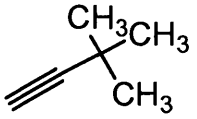
[a]Direct experimental values for log P obtained for fluorescein (Ricardo Sanchez, UCL)

Table 8.2. Calculated values for fluorescein obtained from solver.

Descriptor calculation for terbinafine (via the analogy method)

The descriptors **S**, **A** and **B** are constant values for a given functional groups. They can also be estimated by comparison with known descriptors of similar compounds, for which the descriptors are already known. In this case, the fragments for terbinafine are shown in table 8.3. The descriptors obtained for the fragments are added up to give an estimated value for the compound as a whole. For terbinafine, **E** was obtained from the hypothetical refractive index, which was calculated using ACD Chemsketch, version 6.0. Descriptor **V**, is obtained by molecular formula, the

number of rings present within the compound structure and the chemical formula as discussed in detail from chapter 7. From table 8.4, it seems that the calculated value for water/octanol partition agrees with the observed value, which suggests that the descriptors are good as one can get for this drug compound.

TERBINAFINE ($C_{21}H_{25}N$)	DATA			Vr Calculator	
	Number of rings = 2 n^{20}=1.586 (ACD) $C_{21}H_{25}N$			**E** = 1.8910 **V** = 2.6061	
ADDING THE VARIOUS FRAGMENTS		**S**	**A**	**B**	
 Napthalene		0.92	0.00	0.20	
 Trimethylamine		0.20	0.00	0.67	
 Ethylene		0.10	0.00	0.07	
		0.23	0.12	0.10	
	E	**S**	**A**	**B**	**V**
Estimated descriptors:	1.89	**1.45**	**0.12**	**1.04**	2.61

[b]The descriptor E and V were calculated by using a VR calculator program based on equations discussed in chapter 7

Table 8.3. Estimated descriptors for terbinafine

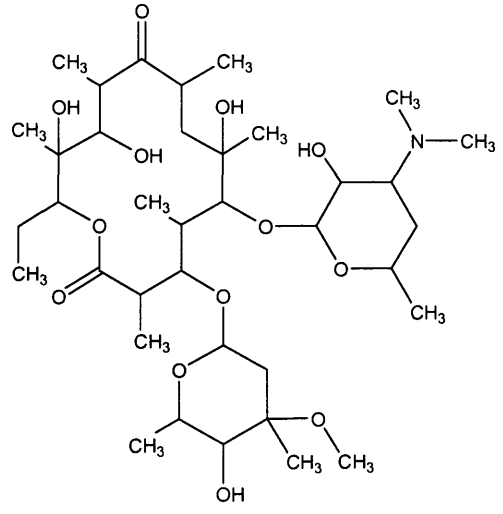
Log P values		
Solvent	**Observed values**	**Calculated values**
Octanol	6.00[a]	6.02

[a]Log P star (octanol/water) obtained from the MedChem database[11]

Table 8.4. Calculated values for terbinafine obtained from Solver.

Descriptor calculation for erythromycin (via the analogy method)

From the literature it was possible to obtain two partition coefficients for this compound. From this data one can check by using solver if the estimated descriptors are good or not. The descriptors **S**, **A** and **B** are estimated by adding fragments with known descriptors, of similar compounds of which the descriptors are already known. In this case, the fragments for erythromycin are shown in table 8.5. Descriptors **E** and **V** were calculated in the same way as previously described for terbinafine. From table 8.6, it seems that the calculated values for the partition for water/octanol and water/trichloromethane do not agree with the observed value. This suggests that the descriptors obtained by analogy method are not very good. The estimation for descriptors may be improved by using the Pharma Alogorithm method.

ERYTHROMYCIN ($C_{37}H_{67}N_1O_{13}$)	**DATA**	**Vr Calculator**
	Number of rings = 3 n^{20}=1.535 (ACD) $C_{37}H_{67}N_1O_{13}$	**E** = 2.150 **V** = 5.773

ADDING THE VARIOUS FRAGMENTS		**S**	**A**	**B**	
 2-Methoxyethanol		0.50	0.30	0.84	
 Dimethoxymethane		0.46	0.00	0.52	

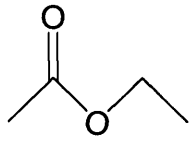
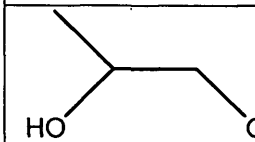
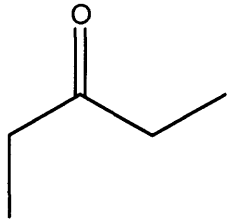
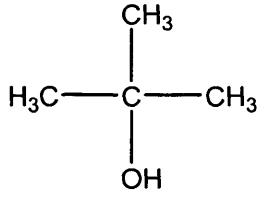
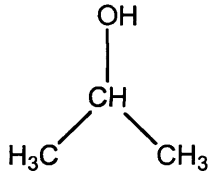
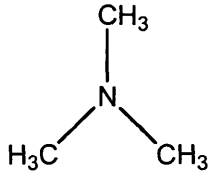
 Ethyl acetate		0.62	0.00	0.45	
 Propan-1,2-diol		0.90	0.58	0.80	
 Pentan-3-one		0.66	0.00	0.51	
 2-Methylpropanol-2-ol		0.30	0.31	0.60	
 Dimethoxymethane		0.46	0.00	0.52	
 Propan-2-ol		0.36	0.33	0.56	
		0.20	0.00	0.67	
	E	**S**	**A**	**B**	**V**
Estimated descriptors:	2.15	**4.46**	**1.52**	**5.47**	5.77

Table 8.5. Estimated descriptors for erythromycin

It appears that when the analogy method is used to estimate descriptors (**S**, **A** and **B**) for erythromycin, the two equations used to calculate log P for octanol and trichloromethane give values that differ greatly from the observed values. The Pharma Algorithm's[5] method for obtaining estimated descriptors for erythromycin,

gives more accurate predictions for **S**, **A** and **B** than the analogy method. From analysis of the two equations used by Solver to calculate **S**, **A** and **B** it appears that there many possible combinations of descriptors, because there are two equations but three unknowns. To limit this problem, three Solver calculations were obtained based on descriptors estimated for **S**, **A** and **B** from the Pharma Algorithm. For Solver (1) calculations, **B** was fixed and descriptors **S** and **A** were minimised until a suitable fit was obtained. For Solver (2) calculations, **A** was fixed and descriptors **S** and **B** were minimised until a suitable fit was obtained. Finally, **S** was fixed and descriptors **A** and **B** were minimised until a suitable fit was obtained based on the two equations mentioned earlier. The averages of descriptor values **S**, **A** and **B** obtained using solver (1 and 3) were taken, are shown in table 8.6. This average data for Solver supported by the Pharma Algorithms method gives log P values closer to the observed values, which suggests that using Solver in conjunction with the Pharma Algorithm calculated values is the best method for obtaining descriptors when only two equations are involved. If more equations are available, then there is only one combination to obtain descriptors, that can be found from Solver itself.

METHODS	S	A	B	OCTANOL LOG P		TRICHLOROMETHANE LOG P	
				CALC	OBS[a]	CALC	OBS[b]
ANALOGY	4.46	1.52	5.47	-0.26	2.54	-0.54	2.46
PHARMA ALGORITHM	3.80	1.16	4.88	2.46	2.54	2.90	2.46
SOLVER (1)	3.73	1.31	4.88	2.54	2.54	2.45	2.46
SOLVER (2)	3.02	1.16	5.10	2.53	2.54	2.45	2.46
SOLVER (3)	3.80	1.32	4.86	2.54	2.54	2.46	2.46
AVERAGE TAKEN	**3.77**	**1.31**	**4.87**	**2.54**	**2.54**	**2.47**	**2.46**

CALC = calculated values obtained from Solver

OBS = observed values

[a]Log P star (octanol/water) obtained from the MedChem database[10]

[b]Log P star (trichloromethane/water) obtained from the MedChem database[10]

Table 8.6. Comparison of the calculated descriptors obtained from the three methods for erythromycin.

References

1. Valko, K., Espinosa, S., Du, C. M., Bosch, E., Roses, M., Bevan, C., Abraham, M. H. (2001). J. Chromatogr. A., **933**, 73-81.
2. Du, C. M., Valko, K., Bevan, C., Reynolds, D., and Abraham, M. H. (2000). J. Chromatogr. Sci., **38**, 503.
3. Valko, K. Plass, M., Bevan, C., Reynolds, D. and Abraham, M. H. (1998). J. Chromatogr. A., **797**, 41-55.
4. Zissimos, A., Abraham, M. H., Du, C. M., Valko, C., Bevan, C., Reynolds, D., Wood, J. and Tam, K. Y (2002). J. Chem. Soc., Perkin Trans (2), 2001.
5. Zissimos, A. M., Abraham, M. H., Barker, K. J., Box, K. J. and Tam, K. Y (2002). J. Chem. Soc., Perkins Trans (2), 470.
6. ADME Boxes v2.2 (2004 edition), http://www.ap-algorithms.com (Accessed: 30/07/2004).
7. ADME Boxes/ Absolv 1.0 (2004 edition), http://www.sirius-analytical.com (Accessed: 30/07/2004).
8. Platts, J. A., Butina, D., Abraham, M. H., Hersey, A (1999). J. Chem. Inf. Comput. Sci., **39**, 835-845.
9. Platts, J. A., Butina, D., Abraham M. H., Hersey A (2000), J. Chem. Inf. Comput. Sci., **40**, 71-80.
10. Abraham, M. H. and Platts, J. A. (2001), J. Org. Chem. **66**, 3484-3491.
11. Thor database MedChem 2003, Biobyte, Inc. and Daylight CIS, Inc. Version 4.71-4.8x.

Chapter 9

Combining Rat and Human Log K data for VOCs

There has been a large number of sources of VOC data reported by various authors, on air to blood and air to tissue distribution.[1-38] The reported data have been obtained by measurements on known volatile organic compounds at steady state equilibrium (37°C) using a head space chromatography technique. Within this reported literature data there are two important groups of species: humans and rats.

The purpose of this work is firstly to collect literature data for air to blood distribution for humans and rats and secondly to collect VOC data for air to tissue partition and then to filter the log K data into these two main classes. If any new reported literature data is obtained for the same compound, this can then be averaged with the current data to give rise to an average value for log K. These constructed sets can then be used to make a direct comparison between human and rat values for both blood and tissue data. Tables are displayed for air to blood and air to tissue distribution for humans and rats as table 9.0 and 9.1. The tables show data that are common compounds to both species and also other non common compounds for humans and rats. Human and rat log K data for VOCs will then be compared for common compounds for air/blood and air/tissue distribution to see if there are similarities or differences between species. Table 9.0 (see appendix) is the up to date literature data for air to blood distribution for VOCs in general. Finally, table 9.1 (see appendix) is the up to date literature data for air to tissue (brain, fat, kidney, liver, lung, muscle and plasma) distribution for VOCs (see appendix).

Air to blood distribution in humans and rats

Gargas et al[6] made an attempt to look at the similarities between human and rat air to blood distribution for a group of VOCs. He compared 36 common compounds between both species to see if the air to blood distribution (blood:air) is the same or not.[6] His table of 36 common compounds contained blood log K values for human and rat that are not averaged literature values. From his table he plotted human blood:air versus rat blood air to reveal that rat values tend to be 1.5 to 2.0 times higher than the corresponding human values in some instances (equation for his plot, see table 9.2).

MODEL	GRADIENT [m]	STD ERROR [m]	INTERCEPT [c]	STD ERROR [c]	REG COEFF [r^2]	n	RMSE
THIS WORK (2004)[a]	1.00	0.028	-0.12	0.047	0.94	86	0.279
GARGAS (1989)[b, 6]	1.01	0.037	-0.23	0.051	0.96	36	0.132

[a] Straight line equation obtained from a logarithmic scale plot

[b] Straight line equation obtained from a numerical scale plot

Table 9.2. Plot for the common compounds between rat and human log K blood

RMSE =	0.279
AAE =	0.210
AE (rat-hum) =	0.124
SD (n-1) =	0.280
n =	86

Table 9.3. Statistics on common compounds between rat and human log K blood, this work

From his equation, the gradient is 1.01 and the intercept is -0.23 which suggest there is a direct linear relationship between the two sets of data for both species. The intercept is close to zero and Gargas et al suggest that rat and human log K values are not the same and they cannot be combined. In Table 9.3., this shows that with

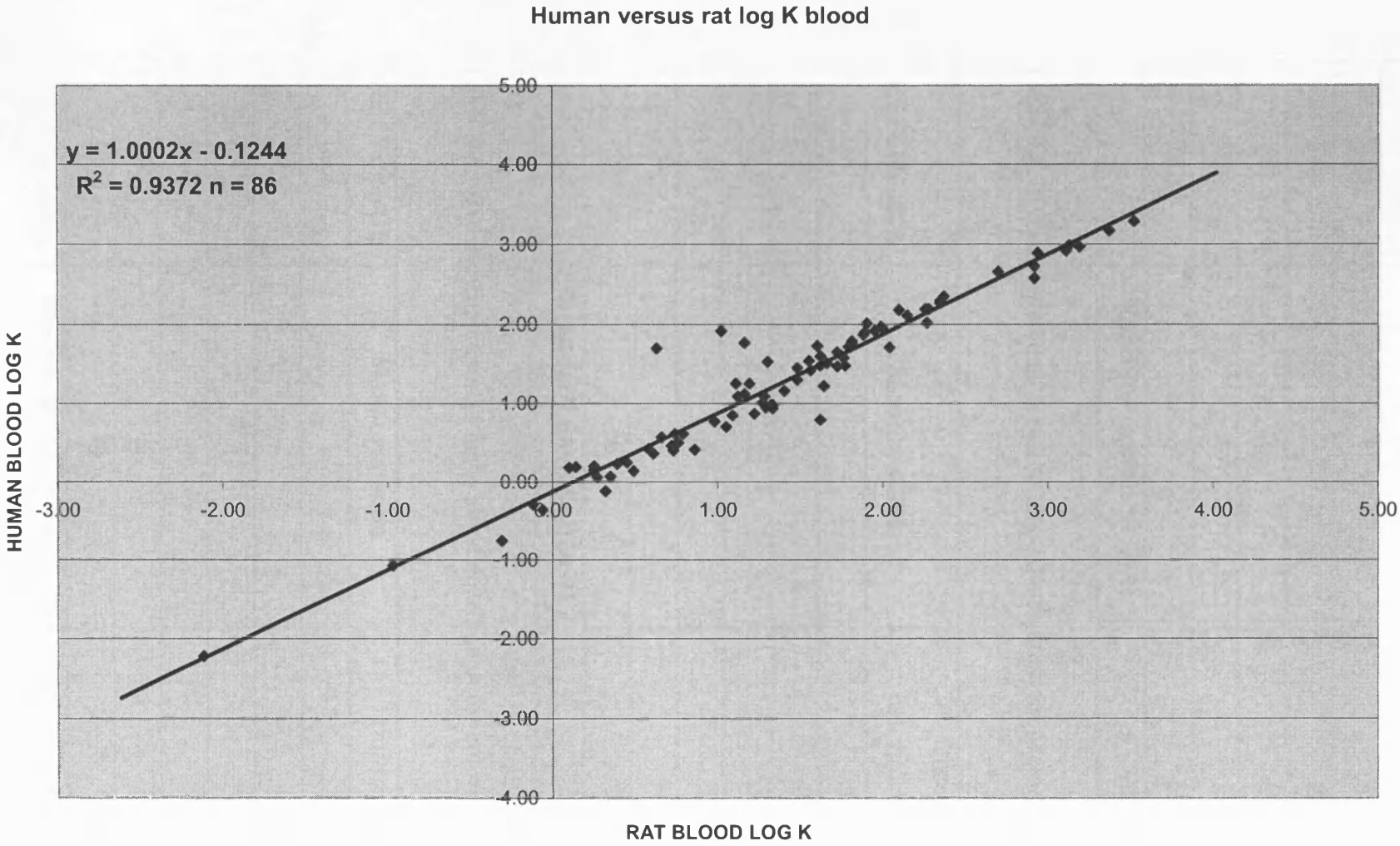
larger data sets of common compounds, the average absolute error to be 0.21 for the two sets of data. The standard deviation of 0.28 also indicates that there very little difference between the two groups of species.

This model takes this method further where the average values for log K blood are obtained for humans and rats. The model has 86 common compounds for VOCs and also the model has greater structural diversity than was used by Gargas. Using the 86 common compounds it is important to find out if there are similarities or differences between humans and rats. A plot was made for the 86 common compounds (see the human verses rat plots) and obtained a straight line equation (equation for the plot, see table 9.2). From this equation, the gradient is 1.00 and intercept of -0.12 which suggest there is a direct linear relationship between the two sets of data for both species. The intercept is very close to zero. This equation supports the conclusion of Gargas that there is a slight variation in data between both species, but this is extremely small as compared with the Gargas model; it appears from the plot that rat and human data are the same

Experimental error for three different VOCs

Fiserova-Bergerova et al[5] was one of the first to report large discrepancies in air to blood partition. She found that the values for gas to blood partition for acetone and 2-butanone appeared different from the other investigators using methods based on the same principle as used in her study. She had no explanation for this discrepancy, and suggested that it is possibly experimental error in the measurements.[5] It is important to find out what is the experimental error in obtaining VOC air to blood (or air to tissue) distribution at several laboratories. From table 9.4 (acetone), 9.5 (chloroform) and 9.6 (trichloroethylene) the partition coefficient appears to vary among the five to six independent laboratories. The standard deviation has been calculated for each of these three compounds, to gain some insight into what is the experimental error for these VOCs. The smallest experimental error is observed for trichloroethylene (standard deviation = 0.06), whilst the highest is for acetone (standard deviation = 0.34). There are many reasons why this variation of partition coefficients (air to blood) for VOC can take place. For example: not enough time to allow steady state concentration to be achieved

before measurements are taken, error measurement of sample of sample, error in the preparation sample, variation in the blood sample (or tissue sample), and many more.


Human versus rat log K blood
y = 1.0002x - 0.1244
R2 = 0.9372 n = 86
HUMAN BLOOD LOG K
RAT BLOOD LOG K
5.00
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00
5.00

REFERENCE	Measurements	Acetone (propan-2-one)
[26]	1	1.68
[4]	2	2.50
[5]	3	2.27
[34]	4	2.39
[35]	5	2.50
	Standard deviation (n-1)	**0.34**

Table 9.4. Experimental data for air to blood distribution for acetone in humans

REFERENCE	Measurements	Chloroform (trichloromethane)
[4]	1	1.09
[36]	2	0.91
[32]	3	1.03
[37]	4	1.01
[6]	5	0.84
[39]	6	0.93
	Standard deviation (n-1)	**0.09**

Table 9.5. Experimental data for air to blood distribution for chloroform in humans

REFERENCE	Measurements	Trichloroethylene
[38]	1	1.08
[4]	2	0.94
[36]	3	0.95
[37]	4	0.98
[6]	5	0.91
[39]	6	0.96
	Standard deviation (n-1)	**0.06**

Table 9.6. Experimental data for air to blood distribution for Trichloroethylene in humans

Compound name	Standard deviation (n-1)
Acetone	0.34
Chloroform	0.09
Trichloroethylene	0.06
Average standard deviation (n-1)	**0.16**

Table 9.7. Average standard deviation for three VOCs (air to blood in human)

The average standard deviation for the three compounds can be taken, to gain a better idea of what is the average experimental error obtained. In this case from Table 9.7, the average standard deviation is found to be 0.16, which could represents the average experimental error for air to blood (or air tissue to tissue) distribution for VOCs in general.

Total regression of blood log K for humans and rats

In table 9.0, a full list of blood K data is given from various literature sources. Here it can be seen that there is a large data set of solutes for air to blood distribution for humans and rats. It is important to compare both of the Abraham equations for humans and rats for air to blood distributions, to verify if the equations for both species are the same. From table 9.8 the two equations have been obtained by multi linear regression analysis. By comparing the standard errors for each of the coefficients in the equations in table 9.8, it can be seen that the coefficients **e**, **s**, **a**, **b** and **l** are same within the standard error (see Table 9.9) at 95% confidence interval.

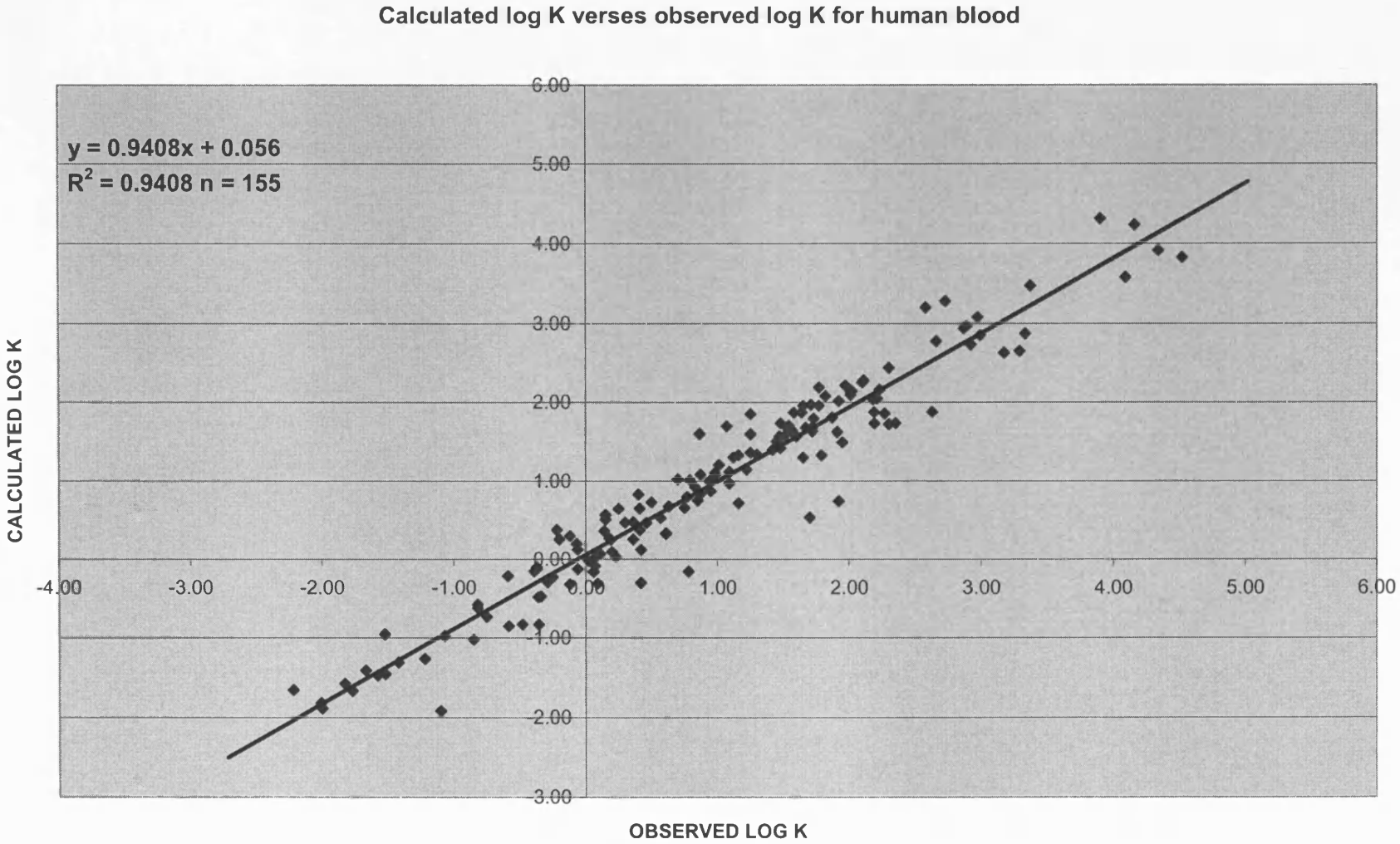
Species	c	e	s	a	b	l	n	r^2	SD (n-1)	F
HUMAN	-1.18	0.39	0.97	3.80	2.69	0.41	155	0.94	0.34	474.00
RAT	-0.75	0.56	1.06	3.64	2.41	0.29	127	0.91	0.29	241.63

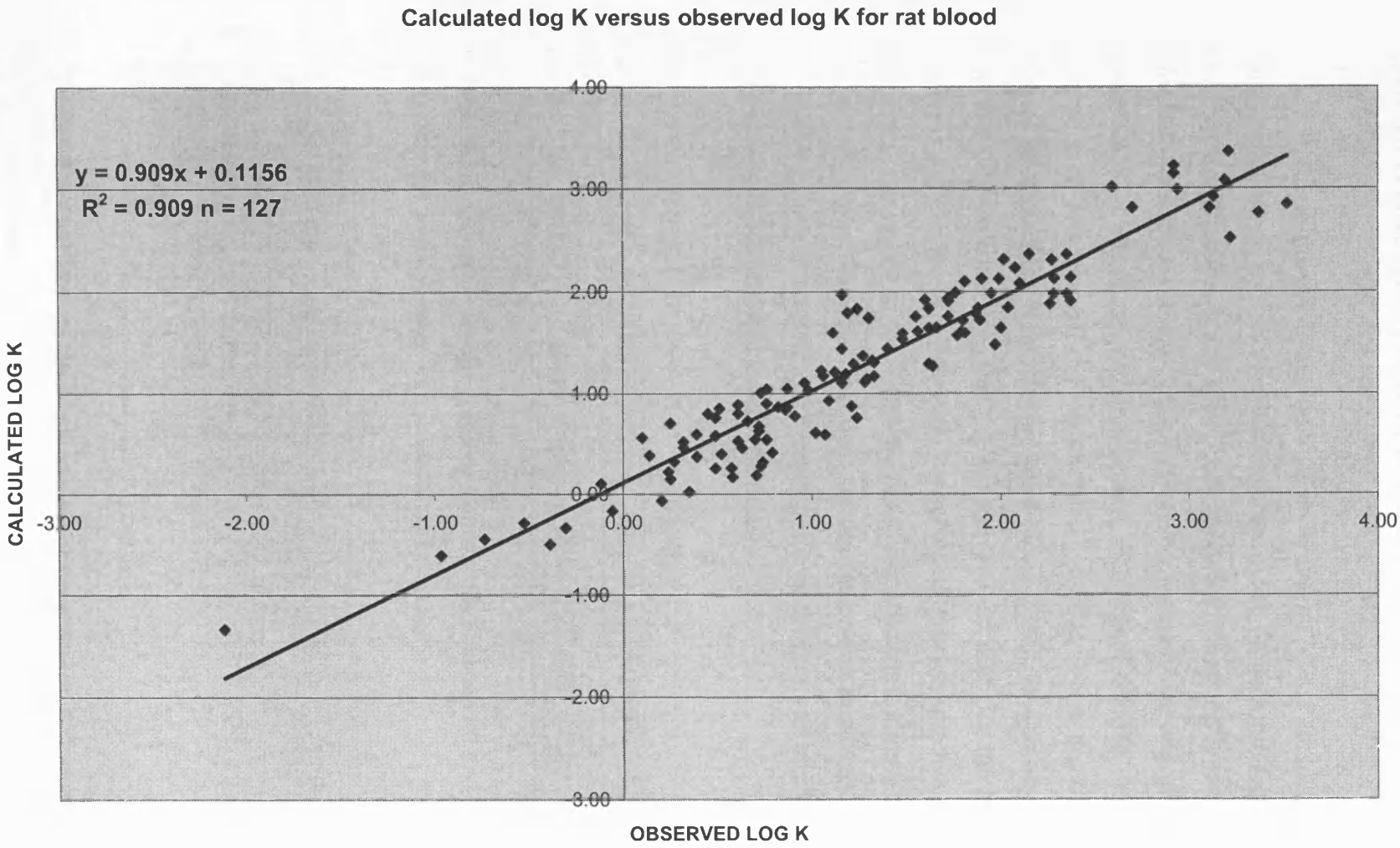
Table 9.8. Human and Rat log K blood (blood log K verses descriptors)

Coefficients	HUMAN 95% limits	HUMAN SD	RAT 95% limits	RAT SD
c	-1.29 to 1.06	0.06	-0.93 to -0.58	0.09
e	0.17 to 0.62	0.11	0.29 to 0.84	0.14
s	0.71 to 1.24	0.13	0.78 to 1.33	0.14
a	3.24 to 4.36	0.28	3.08 to 4.19	0.28
b	2.38 to 3.00	0.16	2.07 to 2.76	0.17
l	0.36 to 0.46	0.03	0.23 to 0.35	0.03

Table 9.9. The range for the standard error (95% confidence interval) of the coefficient

The coefficient **c** is just outside this range. However, it can be concluded that rat and human air to blood tissue distribution are the same for both species (see the plots for rats and humans verses Abraham's descriptors over the next two pages).


Calculated log K verses observed log K for human blood
y = 0.9408x + 0.056
R2 = 0.9408 n = 155
CALCULATED LOG K
OBSERVED LOG K
6.00
5.00
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00
5.00
6.00


Calculated log K versus observed log K for rat blood
y = 0.909x + 0.1156
R2 = 0.909 n = 127
CALCULATED LOG K
OBSERVED LOG K
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00

Air to tissue distribution for humans and rats

It is important to determine if human and rat VOCs air to tissue distribution are the same or not. Table 9.1 contains a list of data for log K air to tissue distribution for humans and rats. The tissue data where log K have been compiled for both species are: brain, fat, kidney, liver, lung, muscle and plasma. From this data set an overall list for air to tissue distribution for 108 compounds that is common to both humans and rats has been compiled. From this list a plot (humans verses rat tissues log K) was constructed to see if there is a similarity between both species. From table 9.10 that the straight line equation for the plot suggests that there is a direct linear relationship between both humans and rats. The intercept is almost 0, the square of the overall correlation coefficient is almost 1 and the gradient is almost 1 which suggest that they are the same. The standard deviation is small suggesting there is very little difference between the two sets of data (see plot of air to tissue distribution for humans verses rats).

MODEL	GRADIENT [m]	STD ERROR [m]	INTERCEPT [c]	STD ERROR [c]	REG COEFF [r^2]	n	RMSE
THIS WORK (2004)	0.91	0.019	0.10	0.037	0.96	108	0.27

Table 9.10. Equation for the common compounds between rat and human tissue log K (brain, fat, kidney, liver, lung, muscle and plasma)

RMSE =	0.265
AAE =	0.193
AE (rat-hum) =	0.030
SD (n-1) =	0.266
n =	108

Table 9.11. Statistics on the log K values for the common compounds between rat and human tissues.

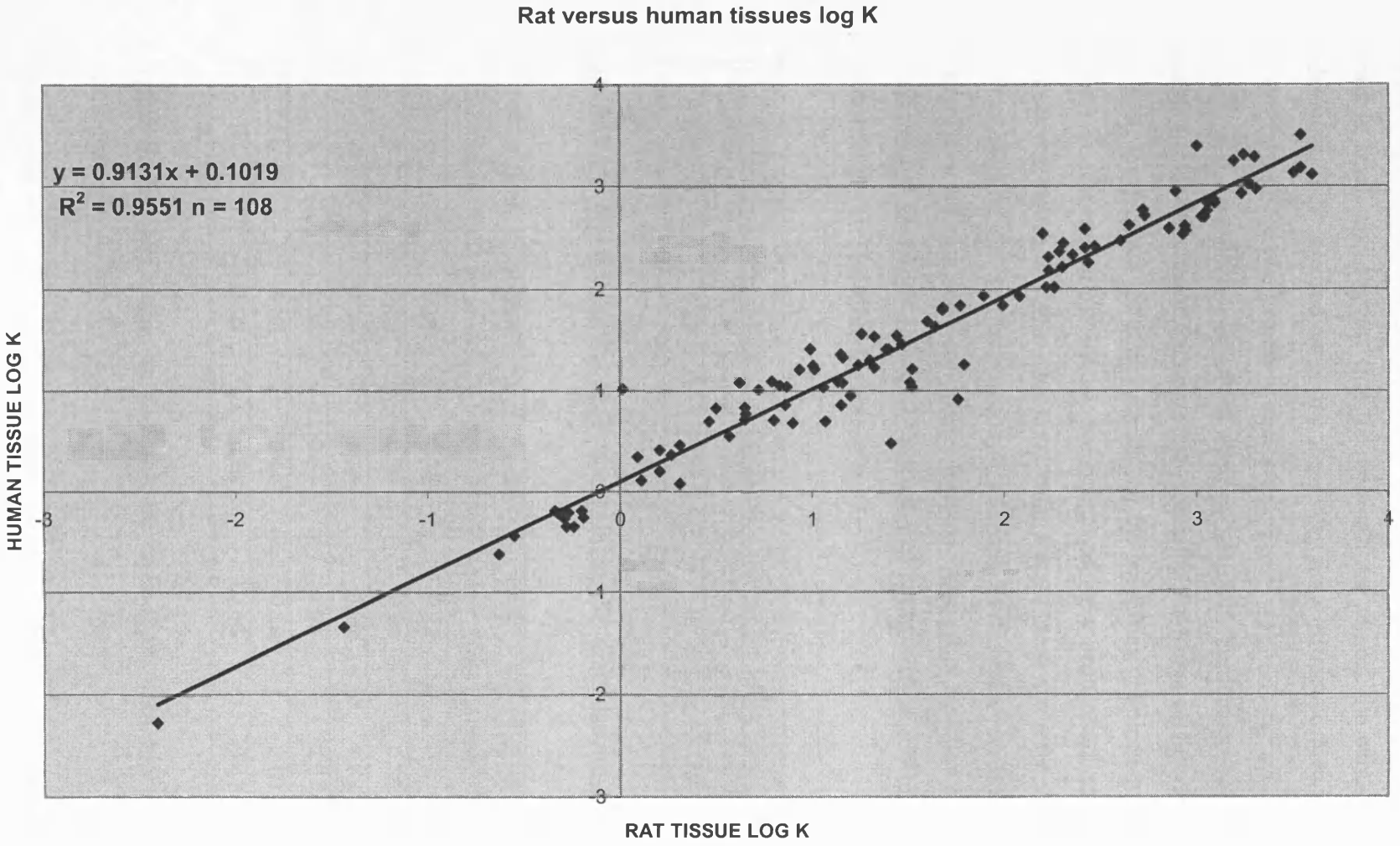

Rat versus human tissues log K
y = 0.9131x + 0.1019
R2 = 0.9551 n = 108
HUMAN TISSUE LOG K
RAT TISSUE LOG K
4
3
2
1
0
-1
-2
-3
-3
-2
-1
0
1
2
3
4

Table 9.11 the average absolute error of 0.19 for the two sets of data for human and rats is small and suggest that the two sets are almost the same. The standard deviation 0.27 also indicates that there very little difference between the errors for the means of the two groups of species.

It is also finally important again to verify if the data on both species are the same for air to tissue distribution by comparing the tissue equations. In this case fat has the largest data set of any tissues for the equations to be constructed for humans and rats. From table 9.12 the two equations have been obtained for fat by multi linear regression analysis. By comparing the standard error for the coefficients of the equations, the coefficients c, e, s, a, b and l are same within the standard error (see Table 9.13) at 95% confidence interval. The plots are shown in figure 9.0 for humans and figure 9.1 for rats. This suggests again that rat and human air to tissue distributions are the same for both species.

Species	c	e	S	a	b	l	n	r^2	SD (n-1)	F
HUMAN	-0.03	0.02	0.70	2.14	0.04	0.77	30	0.97	0.15	153.53
RAT	0.09	0.03	0.53	2.15	0.20	0.74	30	0.97	0.14	145.17

Table 9.12. Human and Rat log K tissue (fat log K verses descriptors)

Coefficients	HUMAN Lower 95%	HUMAN Upper 95%	RAT Lower 95%	RAT Upper 95%
c	-0.24	0.18	-0.12	0.30
e	-0.32	0.36	-0.30	0.36
s	0.33	1.06	0.18	0.89
a	1.60	2.67	1.63	2.67
b	-0.36	0.44	-0.19	0.60
l	0.69	0.85	0.66	0.81

Table 9.13. The range for the standard error (95% confidence interval) of the coefficient (for fat log K)

Conclusion

Both the equations and plots show that human and rat air to blood (and air to tissue) distributions are the same within the experimental error. Human and rat log K data for air to blood and air to tissue distribution can be combined and averages taken where necessary to develop new Abraham equations to predict distribution for VOCs.

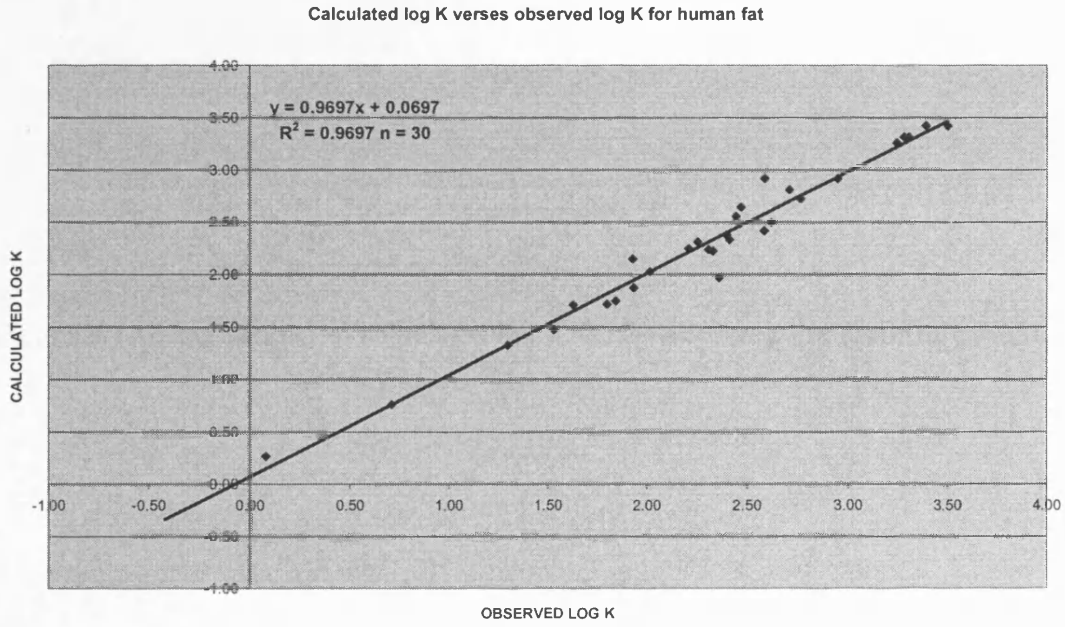


Figure 9.0. Graph showing calculated log K verses observed log K for human blood

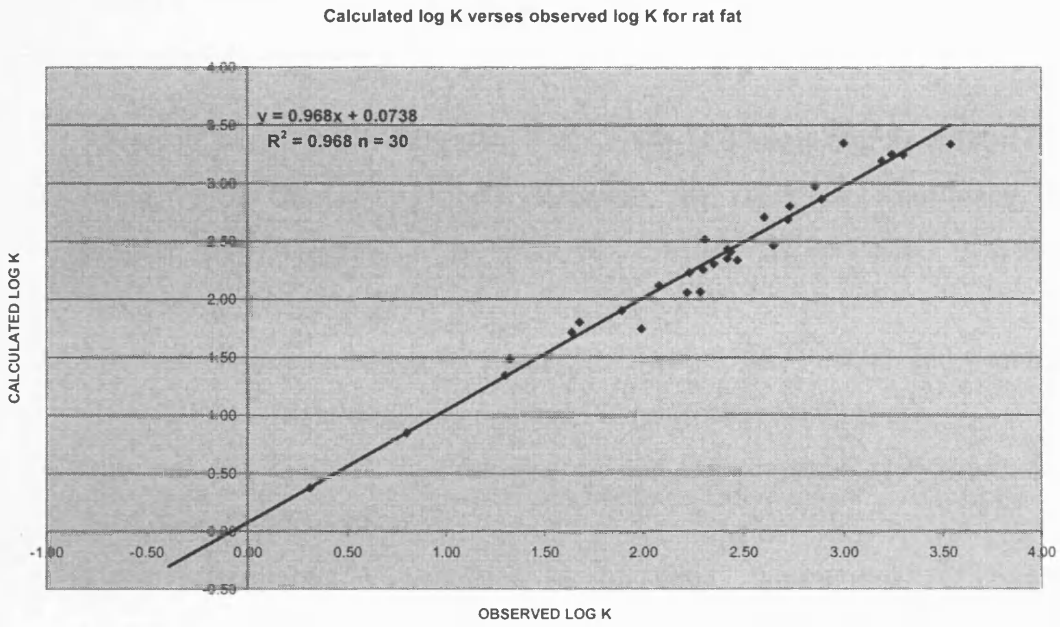


Figure 9.1. Graph showing calculated log K verses observed log K for rat blood

References

1. Abraham, M. H. and Weathersby, P. K (1994). J. Pharm. Sci, **83**, 1450-1455.
2. Filser, J. G., Csanady, GY. A., Hartmann, M., Denk, B., Kauffmann, A., Kessler, W., Kreuzer, P. E., Putz, C., Shen,J. H. and Stei, P (1996). Toxicology, **113**, 278-287.
3. Filser, J. G., Schmidbauer, R., Rampf, F., Baur, C. M., Putz, C. and Csanady, G. A (2000). Toxicology and Applied Pharmacology, **169**, 40-51.
4. Imbriani, M., Ghittori, S., Pezzagno, G. and Capodagio, E (1985). G. Ital. Med. Lav., **7**, 133-140.
5. Fiserova-Bergerova, V. and Diaz, M. L (1986). International Archives of Occupational and Environmental Health, **58**, 75-87.
6. Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H. and Anderson, M. E (1989). Toxicology and Applied Pharmacology, **98**, 87-99.
7. Meulenberg, C. J. W. and Vijverberg, H. P. M (2000). Toxicology and Applied Pharmacology, **165**, 206-216.
8. Kaneko, T., Kim, H. Y., Wang, P. Y. and Sato, A (1997). J. Occup. Health, **39**, 341-342.
9. Kedderis, G. L., Carfagna, M. A., Held, S. D., Batra, R., Murphy, J. E. and Gargas, M. L (1993). Toxicology and Applied Pharmacology, **123**, 274-282.
10. Knaak, J. B. and Smith, L. W (1998). Inhalation Toxicology, **10**, 65-85.
11. Borghoff, S. J., Murphy, J. E. and Medinsky, M. A (1996). Fundamental and Applied Toxicology, **30**, 264-275.
12. Loizou, G. D. Urban, G. Dekant, W. and Anders, M.W (1994). Drug Metabolism and Disposition, **22**, 511-517.
13. Loizou, G. D. and Anders, M.W (1993). Drug Metabolism and Disposition, **21**, 634-639.
14. Medinsky, M. A., Bechtold, W. E., Birnbaum, L. S., Chico, D. M., Gerlach, R. F. and Henderson, R. F (1988). Fundamental and Applied Toxicology, **11**, 250-260.
15. Nussbaum, E. and Hursh, J. B (1957). Science, **125**, 552-553.
16. Zahlsen, K. and Eide, I (1996). Arch. Toxicol, **70**, 397-404.

17. Zahlsen, K., Eide, I., Nilsen, A. M. and Nisen, O. G (1992). Pharmacology & Toxicology, **71**, 144-149.
18. Zahlsen, K., Eide, I., Nilsen, A. M. and Nisen, O. G (1993). Pharmacology & Toxicology, **73**, 163-168.
19. Kaneko, T., Wang, P-Y. and Sato, A (2000). J. Occup. Health, **42**, 86-87.
20. Johanson, G. and Dynesius, B (1988). British Journal of Industrial Medicine, **45**, 561-564.
21. Sweeney, L. M., Himmelstein, M. W. and Gargas, M. L (2001). Chemical-Biological Interactions, **135-136**, 303-322.
22. Morris, J. B., Hassett, D. N. and Blanchard, K. T (1993). Toxicolgy and Applied Pharmacology, **123**, 120-129.
23. Krishnan, K., Gargas, M. L., Fennell, T. R. and Anderson, M. E (1992). Toxicol. Indust. Health, **8**, 121-140.
24. Csanady, Gy. A., Kreuzer, P. E., Baur, C. and Filser, J. G (1996). Toxicology, **113**, 300-305.
25. Gargas, M. L., Anderson, M. E., Teo, S. K. O., Batra, R., Fennell, T. R. and Kedderis, G. L (1995). Toxicology and Applied Pharmacology, **134**, 185-194.
26. Guitart, R (1993). Revista Espanola de Fisiologia., **49**, 195-202.
27. Barton, H. A., Creech, J. R., Godin, C. S., Randall, G. M. and Seckel, C. S (1995). Toxicology and Applied Pharmacology, **130**, 237-247.
28. Simmons, J. E., Boyes, W. K., Bushnell, P. J., Raymer, J. H., Limsakun, T., McDonald, A., Sey, Y. M. and Evans, M. V (2002). Toxicological Sciences, **69**, 3-15.
29. Csanady, Gy. A., Denk, B., Putz, C., Kreuzer, P. E., Kessler, W., Baur, C., Gargas, M. L. and Filser, J. G (2000). Toxicology and Applied Pharmacology, **165**, 1-26.
30. Reitz, R. H., Gargas, M. L., Anderson, M. E., Provan, W. M. and Green, T. L (1996). Toxicology and Applied Pharmacology, **137**, 253-267.
31. Beliveau, M., Charest-Tardif, G and Krishnan, K (2001). Chemosphere, **44**, 377-381.
32. Batterman, S., Zhang, L., Wang, S. and Franzblau, A (2002). The Science of the Total Environment, **284**, 237-247.

33. Lilly, P. D., Anderson, M. E., Ross, T. M. and Pegram, R. A (1997). Toxicology, **124**, 141-152.
34. Sato, A., Nakajima, T (1979). British Journal of Industrial Medicine, **36**, 231-234.
35. Young, I. H. and Wagner, P. D (1979). J. Applied Physiology, **46**, 240-248.
36. Sherwood, R. J (1976). British Journal of Industrial Medicine, **33**, 106-107.
37. Sato, A., Nakajima, T (1979). Archives of Environmental Health, **34**, 69-75.
38. Wilson, P. F., Freeman, C. G., McEwan, M. J., Allardyce, R. A. and Shaw, G. M (2003). Applied Occupational and Environmental Hygiene, **18**, 759-763.
39. Thrall, K. D., Gies, R. A., Muniz, J., Woodstock, A. D. and Higgins, G (2002). J. Toxicol. Letters, **116**, 183-188.

Chapter

10

Air to blood distribution for VOCs (rat and human average data)

The solubility of VOCs in biological fluids like blood is of importance in anaesthesiology, toxicology and environmental health. There has been a large number of reported log K data for air to blood distribution and their references are cited in chapter 9 of this thesis. Several cited papers, where models have been used to predict air to blood distribution in humans, rats or for both species shall be discussed. The purpose of this work is to add new literature data to the existing data set of Abrahams and to make an attempt to construct a new equation to predict these processes. Correlative and predictive equations shall be constructed based on datasets that represent average human and rat data combined. The complete list of the VOC log K data for these averages, are shown in table 10.0 (see chapter 9 how this data is obtained). The reported literature models for the correlation of air to blood distribution for VOCs are found for the following authors: Gargas (et al)[1], Weathersby (et al)[2] and Meulenberg (et al).[3]

Gargas et al, determined log K values (air to blood distribution) for 55 low molecular weight, volatile compounds in humans. From this data set he successfully modelled human air to blood distribution in terms of contribution from air to oil and air to saline values using linear regression techniques as shown below.[1]

$$\text{Log}\,(K_{\text{Human blood}}) = 0.553(\pm 0.030)\log(K_{oil}) + 0.351(\pm 0.025)\log(K_{sal}) - 0.286(\pm 0.080) \qquad \text{Equation 10.0}$$

$r^2 = 0.957$ RMSE = 0.179 n = 55

Gargas also obtained 36 log K (air to blood distribution) values, common to rats and humans. From this data set he used the rat log K in the equations to predict human air to blood distribution (see below for equation). Unfortunately his model doesn't have a test set, thus any predictions cannot be assessed.[1]

$$\text{Log}(K_{\text{Human blood}}) = 1.014(\pm 0.037)\ \log(K_{\text{Rat blood}}) - 0.232(\pm 0.051) \qquad \text{Equation 10.1}$$

$r^2 = 0.957$ RMSE = 0.132 n=36

Weathersby compiled a dataset of 86 solutes in his model to correlate air to blood distribution for humans. Most of the data is compiled from previous literature reviews and new data have been added to update any current models in air to blood distribution. Weathersby used the Abraham solvation equation to correlate the log K data (Ostwald solubility coefficient at 310K) for air to blood with the five Abraham descriptors (**E**, **S**, **A**, **B** and **L**). The Abraham equation used for this model is shown in equation 10.2.[2]

$$\text{Log K}_{(\text{air:blood})} = -1.269 + 0.612\mathbf{E} + 0.916\mathbf{S} + 3.614\mathbf{A} + 3.381\mathbf{B} + 0.362\mathbf{L} \qquad \text{Equation 10.2}$$

$r^2 = 0.977$ S.D = 0.203 n = 82

Meulenberg obtained a reasonable amount of data in his training set to produce two correlative equations for air to blood distribution in rats and humans. He found that the partition coefficients of VOCs in human (or rat) blood are well described by a linear combination of $K_{\text{air:oil}}$ and $K_{\text{air:saline}}$ with specific regression coefficient for the biological sample i.e. blood. The partition (K) data for air to blood distribution are averaged values from literature sources and have been used for the bilinear regression. This method is very much similar to the method employed by Gargas et al.[1] Equations 10.3 and 10.4 shows how $K_{\text{air:blood}}$ is described as a bilinear function of $K_{\text{air:olive oil}}$ and $K_{\text{air:saline}}$. The coefficients α_{oil} and α_{saline} of both equations represents the tissue-specific contributions of the lipophilic and hydrophilic interactions to the solubility of the VOCs in blood. The c constant has no specific physical meaning but it was required by Meulenberg to compensate for systematic errors in blood, oil, or saline partition coefficients.[3]

Human equation

$K_{air:blood} = 0.0072K_{air:olive\ oil} + 0.898K_{air:saline} + 0.03$ Equation 10.3

$r^2 = 0.99$ SD = n/a n = 109

Rat equation

$K_{air:blood} = 0.0054K_{air:olive\ oil} + 0.931K_{air:saline} + 1.16$ Equation 10.3

$r^2 = 0.93$ SD = n/a n = 92

It appears that none of the three models have any test sets that includes statistical information that can be used to verify their predictions. It was very unfortunate that Meulenberg did not calculate the fits (i.e. standard deviation) for his training set used for his correlation.

An improvement can be made on air to blood distribution by using training sets and test sets containing a large number of compounds (including statistical information). It would be very interesting to see if the diversity of compounds helps to improve models that are currently available. Most of the models discussed before are outdated and many new measurements have been made for VOCs which will be included into the new training sets (see appendix table 10.0).

Discussion on air to blood distribution for VOCs in human and rat combined

The Abraham solvation equations can successfully be employed to correlate VOC air to blood distribution (see table 10.1). Considering that this model has now got 196 compounds in the dataset, the fits for this new equation appear to very good (S.D = 0.324 and $r^2 = 0.938$). This work is the most up to date model for VOC air to blood distribution, it appears this is the most general model yet reported for both species (humans and rats).

There were two outliers observed from this training set: CFH2CH2CFH2 and cyclopentadiene. The residual for these two compounds are approximately 1 log unit and may result from being an experimental error. Other possibilities can arise as these compounds may bind irreversibly to proteins contained in blood or interact and bind to hydrophobic pockets in haemoglobin.

All the log K equations for air to tissue distribution are based on the same model, given by:

$$\text{Log K} = c + e.\mathbf{E} + s.\mathbf{S} + a.\mathbf{A} + b.\mathbf{B} + l.\mathbf{L}$$

In the present case, if the two outliers are included in the model, this gives the following equation.

Log K (air:blood) = -1.049 + 0.450**E** + 1.088**S** + 3.720**A** + 2.565**B** + 0.373**L**

N = 198 r^2 = 0.932 SD = 0.338 F = 523.776 D = 5

The statistics of the equation with n = 198 are very close to those of the equal with n = 196 (as shown in table 10.1). The coefficients and statistics for both of the equations are very much the same. This model shown above is used throughout the thesis for VOCs.

EQUATIONS	N	r^2	SD	F	D
FULL SET-AVERAGED **Log K** (air:blood) = - 1.069 + 0.456**E** + 1.083**S** + 3.738**A** + 2.580**B** + 0.376**L**	196	0.938	0.324	572.751	5
FULL SET-COMBINED (NOT AVERAGED) **Log K** (air:blood) = - 1.062 + 0.460**E** + 1.067**S** + 3.777**A** + 2.556**B** + 0.375**L**	282	0.927	0.330	699.100	5
TRAINING SET **Log K** (air:blood) = - 0.978 + 0.596**E** + 1.000**S** + 3.494**A** + 2.914**B** + 0.329**L** **TEST SET** **N** = 98 **RMSE**= 0.326 **AAE** = 0.255 **AE** = 0.043 **SD** (n-1) = 0.327	98	0.933	0.338	257.339	5

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = the root mean square error

AAE = absolute average error

AE = average error

Table 10.1. Summary of the equations for VOC air to blood distributions and statistics for test sets of 98 solutes

The inter-correlations of descriptors used for VOC air to blood distribution are shown in table 10.2. The largest r^2 value for E versus L descriptor is 0.37. It appears that there is no strong inter-correlation between the set of descriptors used in this model.

The Abraham solvation equation has successfully been employed to predict air to blood distribution for VOCs provided the solute is within the given descriptor range of the training set to make the correlative equations. To test the predictability of the model it was necessary to split the full data into training and test set using a random data set generator. A new model is developed using only those data in the training set, which is then used to predict values for the test set. The full data set of 196 compounds in two sets of 98, using the first set to construct equation similar to the first equation from table 10.1 and predicting 96 log K values (air to blood distribution of VOCs) for the second set. The regression yielded similar training equation and statistics to the first equation in table 10.1. The test of 96 compounds shows the average absolute error to be 0.255 for the two sets of data (observed and predicted log K). The root mean square error of 0.326 also indicates that there very little difference between the observed value and predicted data for air to blood distribution for VOCs. The average error of 0.043 indicates that the equation of the model is not biased, as it is close 0.

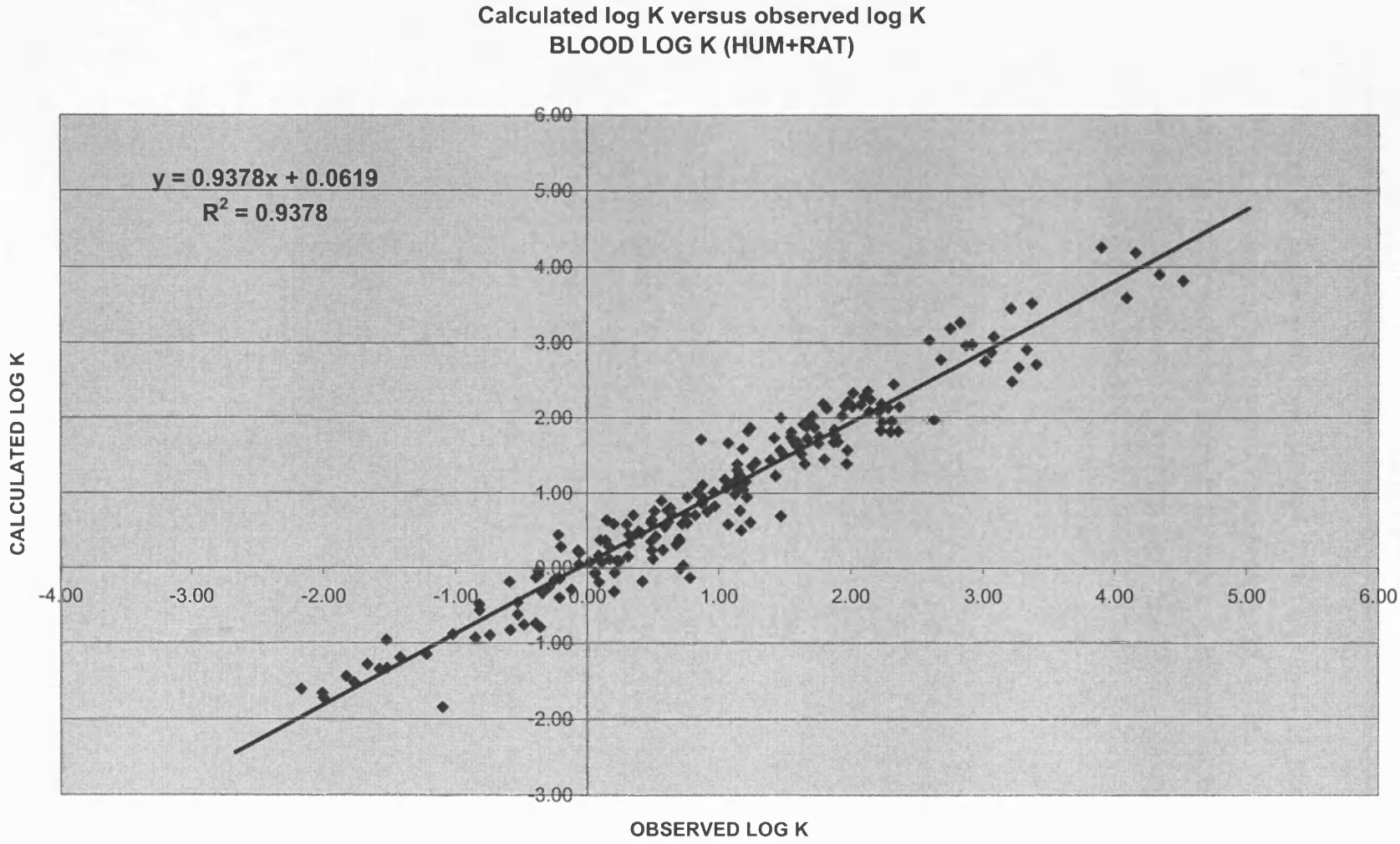
r	E	S	A	B	L
E	1.00				
S	0.54	1.00			
A	-0.09	0.08	1.00		
B	0.05	0.41	0.42	1.00	
L	0.61	0.44	-0.06	0.21	1.00

Table 10.2. The inter-correlation of descriptors for 196 solutes using descriptors **E**, **S**, **A**, **B** and **L**.

The positive **e**- and **l**-coefficients in the air to blood equation (10.1) indicate that increasing molecular size, cavity effects, London dispersion interaction and the presence of n- and Π-electrons pairs can push VOCS out of air and into the blood phase. The coefficients of the 'polar' descriptors **S**, **A**, and **B** are all positive, indicating that polarity and hydrogen-bond donor and acceptor ability is stronger in the blood phase than in gas phase. The solutes are pushed out of the air and into the blood phase. Finally, the Abraham equation can successfully be used to make predictions for air to blood distribution for a diverse range of solutes (VOCs) as shown before.

We can also use the data on human blood and rat blood without averaging the K_{blood}(human) and K_{blood}(rat) values. This leads to 282 data points and the corresponding equation is shown in blue in table 10.1. The coefficients in the combined equation and the average equation (table 10.1) are almost identical. The r^2 value in the combined equation is a little less than that in the average equation, but the F-statistic is much better, simply reflecting the larger number of data points.

The statistics for this model is very good, as indicated by the large spread of data shown from the plot. It is found to be possible to construct an equation capable of predicting further values of log K_{blood} to around 0.3 log units. This equation is one of the soundest equations for predicting log K values for air to blood distribution of VOCs.


Calculated log K versus observed log K
BLOOD LOG K (HUM+RAT)
y = 0.9378x + 0.0619
R2 = 0.9378
CALCULATED LOG K
OBSERVED LOG K
6.00
5.00
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-4.00
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00
5.00
6.00

References

1. Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H. and Anderson, M. E (1989). Toxicology and Applied Pharmacology, **98**, 87-99.
2. Abraham, M. H. and Weathersby, P. K (1994). J. Pharm. Sci, **83**, 1450-1455.
3. Meulenberg, C. J. W. and Vijverberg, H. P. M (2000). Toxicology and Applied Pharmacology, **165**, 206-216.

Chapter

11

Air to plasma, air to blood and plasma to blood distribution for VOCs

In this chapter the Abraham solvation equation shall be used for models that include: air to plasma distribution for VOCS and air to blood (blood and plasma data combined) distribution for VOCs. A comparison between plasma and blood log K data for VOCs and blood to plasma distribution (log P) for VOCs shall be discussed.

Air to plasma distribution for VOCs

The distribution of VOCs from air to plasma is of importance in clinical medicine, toxicology and environmental sciences. There has been only been one reported model for air to plasma distribution.[1] The purpose of this work is to add new literature data to the existing data set of previous authors and to make an attempt to construct a new equation to predict these processes. The correlative equations of Abraham shall be constructed based on datasets that represent average human and rat data combined. The complete list of the VOC log K data for these literature averages is shown in table 11.0. The reported literature models for the correlation of air to plasma distribution for VOCs are found only from Weathersby and Abraham.[1]

Weathersby and Abraham compiled a dataset of 32 VOCs in their model to predict air to plasma distribution for humans. Most of their data is compiled from previous literature reviews and new data have been added to update any current models in air to plasma distribution.[1] Weathersby used the Abraham solvation

equation to correlate the log K data (Ostwald solubility coefficient at 310K) for air to plasma with the five Abraham descriptors (**E**, **S**, **A**, **B** and **L**). The Abraham equation used for this model is shown in equation 11.0.[1]

$$\text{Log K}_{(\text{air:plasma})} = -1.480 + 0.490\mathbf{E} + 2.047\mathbf{S} + 3.507\mathbf{A} + 3.911\mathbf{B} + 0.157\mathbf{L} \qquad \text{Equation 11.0}$$

$r^2 = 0.984$ S.D = 0.232 n = 32

The equation above is the only literature model for air to plasma distribution of VOCs. A small improvement can be made by adding new data to that of Weathersby and Abraham for air to plasma distribution of VOCs to see if the new equation for the full set has better fits than the above model. Table 11.0 (see appendix) is the up to date literature data for air to plasma distribution for VOCs in general.

Discussion on air to plasma distribution for VOCs

The solvation equation of Abraham can successfully be employed to correlate VOC air to plasma distribution (see table 11.1). Considering that this model has now got 36 compounds in the dataset, the fits for this new equation appear to very good (S.D = 0.249 and $r^2 = 0.979$). This work is the most up to date model for VOC air to plasma distribution; it appears this is the most general model yet reported for both species (humans and rats). There were no outliers observed from this full set.

The inter-correlations of descriptors used for VOC air to plasma distribution are shown in table 11.2. The largest r^2 value for S versus L descriptor is 0.45. It appears that there is no strong inter-correlation between the set of descriptors used in this model.

EQUATIONS	N	r^2	SD	F	D
FULL SET **Log K** (air:plasma) = -1.435 + 0.543**E** + 1.677**S** + 3.518**A** + 3.982**B** + 0.192**L**	36	0.979	0.249	285.359	5

N = number of compounds used in the model

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

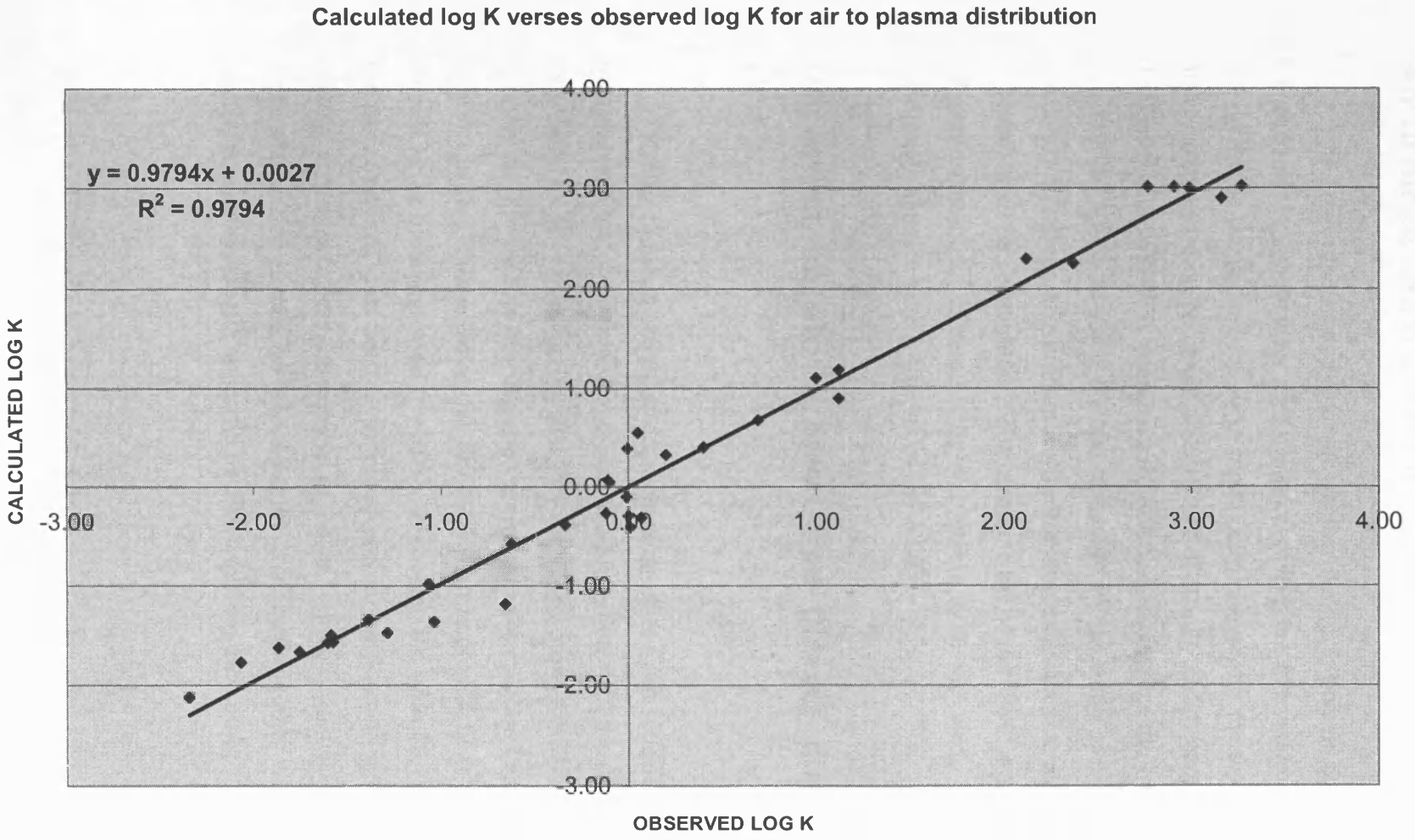
Table 11.1. Summary of the equations for VOC air to plasma distributions

r	E	S	A	B	L
E	1.00				
S	0.39	1.00			
A	0.20	0.37	1.00		
B	0.26	0.60	0.66	1.00	
L	0.43	0.67	0.34	0.54	1.00

Table 11.2. The inter-correlation of descriptors for 36 solutes using descriptors **E**, **S**, **A**, **B** and **L**.

The Abraham solvation equation has successfully been employed to correlate air to plasma distribution for VOCs. No training and test sets were used as the number of compounds was too small. The positive **e**- and **l**-coefficients in the air to plasma equation (11.1) indicate that increasing molecular size, London dispersion interaction and the presence of n- and Π-electrons pairs can push VOCS out of air and into the plasma tissue phase. The coefficients of the 'polar' descriptors **S**, **A**, and **B** are all positive, indicating that polarity and hydrogen-bond donor and acceptor ability is stronger in the plasma phase than in gas phase. The solutes are pushed out of the air and into the plasma phase.

The statistics for this correlative model (air-plasma distribution) is good, as indicated by the even spread of data shown from the plot. The air to plasma model is comparable to the air to blood model, as blood and plasma behave very similar to other especially when they interact with VOCs.


Calculated log K verses observed log K for air to plasma distribution
y = 0.9794x + 0.0027
R2 = 0.9794
CALCULATED LOG K
OBSERVED LOG K
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00

Air to blood and plasma distribution for VOCs

It is now important to find out whether the air to plasma equation for VOC is the same as the equation for air to blood distribution. One way of investigating this is to combine the dataset for air to blood distribution with the data set of air plasma distribution and then construct an equation of Abraham solvation based on the full data set combined. This equation can then be compared to the original air to blood equation for VOCs to see if the two equations are the same or not. Correlative and predictive equations of Abraham for the combined dataset (i.e. air to plasma and air to blood) shall be constructed. The complete list of the VOC log K data is obtained by combining the air to blood (chapter 10, table 10.0) with air to plasma (table 11.0) distribution data, to give a total set of 232 VOCs.

Discussion on air to blood and plasma distribution for VOCs

The air to blood and plasma solvation equations can successfully be employed to correlate VOC air to blood and plasma distribution with the Abraham descriptors (see table 11.3). Considering that this model has now got 232 compounds in the dataset, the fits for this new equation appear to be very good (S.D = 0.332 and r^2 = 0.943). In fact, this equation for full set has an improved fit (r^2 = 0.943) over the current air to blood equation (r^2=0.938, in chapter 10, table 10.1). The coefficients for the combined data set (air to blood and air to plasma) equation for the 232 solutes, are more or less the same as in the air to blood equation. This suggest that air to plasma log K data for VOCs can be combined together with the air to blood log K data, as they do not appear to be different from each other.

EQUATIONS	N	r^2	SD	F	D
FULL SET					
Log K (air:sample) = - 1.154 + 0.446**E** + 1.130**S** + 3.844**A** + 2.673**B** + 0.386**L**	232 (196+36)	0.943	0.332	741.611	5
Log K (air:blood) = - 1.069 + 0.456**E** + 1.083**S** + 3.738**A** + 2.580**B** + 0.376**L**	196	0.938	0.324	572.751	5
TRAINING SET					
Log K (air:sample) = - 0.997 + 0.629**E** + 1.015**S** + 3.517**A** + 2.769**B** + 0.337**L**	116	0.933	0.332	308.020	5
TEST SET					
N = 116 **RMSE** = 0.352 **AAE** = 0.288					
AE =0.082 **SD** (n-1) = 0.353					

Sample = blood and plasma data combined

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = the root mean square error

AAE = absolute average error

AE = average error

Table 11.3. Summary of the equations for VOC air to blood and plasma distributions and statistics for training sets and test sets of 116 solutes

The Abraham solvation equation has successfully been employed to predict air to blood and plasma distribution for VOCs provided the solute is within the given descriptor range of the training set to make the correlative equations. To test the predictivity of the model it was necessary to split the full data into training and test set using a random data set generator. A new model is developed using only those data in the training set, which is then used to predict values for the test set. The full data set of 232 compounds is divided into two sets of 116, using the first set to construct an equation similar to the first equation from table 11.3 and then 116 log K values are predicted for the second set. The regression yielded a similar training equation and statistics to the first equation in table 11.3. The test of 116 compounds

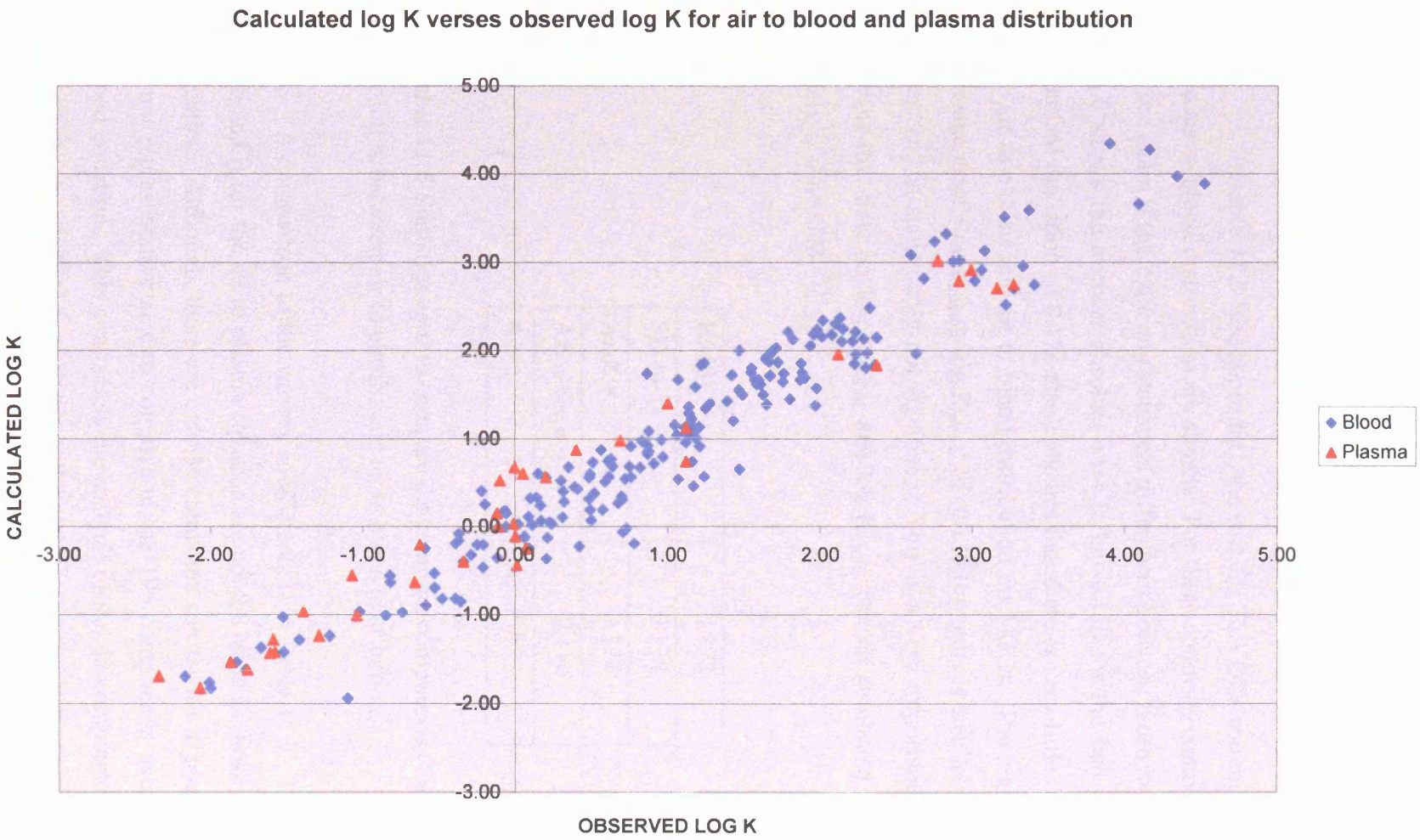
shows the average absolute error to be 0.288 for the two sets of data (observed and predicted log K). The root mean square error of 0.352 also indicates that there is a reasonable predictive capability for the equations shown in table 11.3. The average error of 0.082 indicates that the equation of the model is not biased, as it is just over zero. The plot on page 147, indicates visually that there is no difference between the blood and plasma observed data points.

The coefficients for the air to blood and plasma equation are the same as the coefficients as in the air to blood equation (shown in blue in table 11.3).

The inter-correlations of descriptors used for VOC air to blood and plasma distribution are shown in table 11.4. The largest r^2 for E versus L descriptor is 0.36. It appears that there is no strong inter-correlation between the set of descriptors used in this model.

r	E	S	A	B	L
E	1.00				
S	0.53	1.00			
A	-0.06	0.11	1.00		
B	0.07	0.43	0.46	1.00	
L	0.60	0.49	-0.02	0.24	1.00

Table 11.4. The inter-correlation of descriptors for 232 solutes using descriptors **E**, **S**, **A**, **B** and **L**.


Calculated log K verses observed log K for air to blood and plasma distribution
CALCULATED LOG K
OBSERVED LOG K
5.00
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00
5.00
Blood
Plasma

Comparison between air to plasma and air to blood distribution data for common VOCs

In table 11.5 (see appendix), the data for 35 VOCs common to air to plasma and air to blood distribution are shown. For these common compounds it important to see again if there are any statistical differences between these two data sets. Table 11.6 shows the average absolute error to be -0.116 for the two sets of data. The standard deviation of 0.174 also indicates that there is very little difference between the air to plasma and air to blood distribution for VOCs. The reason for this is that average error (log K_{plasma}-log K_{blood}) of -0.116 log units is smaller than experimental error and is statistically not significant. This is a very important result, because it means that data on log K_{plasma} and log K_{blood} can be combined in any correlative analysis if needed.

RMSE =	0.172
SD (n-1) =	0.174
AAE =	0.127
AE $_{(plasma\text{-}blood)}$ **=**	-0.116
n =	35

Table 11.6 Statistics used to compare the common compounds (found in table 11.5) for log K between air to plasma and air to blood distribution.

A comparison of the various coefficients is in Table 11.7. The coefficients (**c**, **s**, **b** and **l**) for the air to plasma equation are not the same as those in the air to blood equation. Although, there are only 36 compounds in the air to plasma equation, this is not representative size to compare to the 196 compounds used make the air to blood equation. This contradicts the analysis of the 35 compounds, but is probably due to the small number of VOCs in the plasma equation.

The combined data set for air to blood and plasma equation, has very much the same coefficients as the air to blood equation. This also suggests that air to blood

and air to plasma data for VOCs are the same, although the compound numbers for blood and plasma are not the same.

MODEL	N	c	e	s	a	b	l
Air to blood	196	-1.069	0.456	1.083	3.738	2.580	0.376
Air to plasma	36	-1.435	0.543	1.677	3.518	3.982	0.192
Air to blood and plasma	232	-1.154	0.446	1.130	3.844	2.673	0.386

Table 11.7 The comparison of the various coefficients

Now it was important to see if the air to blood equation for VOC from table 11.7 can be used to predict the 35 VOCs for air to plasma distribution as shown from table 11.8. If the average error of 0.04 for the test set of the air to blood equation is combined (from table 10.1) with the experimental error of 0.163, a combined error of 0.203 is found. The average error for this test set 0.18 (table 11.8) is found to be less than the combined error of 0.203. This concludes that blood and plasma log K data for VOCs are the within the given error.

RMSE =	0.390
SD (n-1) =	0.396
AAE =	0.334
AE =	0.179
n =	35

Table 11.8 Statistics used to compare the difference between observed and predicted values for plasma log K; the predicted values are from the air to blood equation.

Blood to plasma distribution for VOCs

It is also important to find out whether there is correlation between the blood to plasma distribution for VOCs and the Abraham descriptors. The log P data for blood to plasma distribution for VOCs can be obtained by combining the log K_{plasma} data with log K_{blood} VOC data.

Weathersby and Abraham compiled a dataset of 32 VOCs in his model to predict blood to plasma distribution for humans. Most of their data is compiled from previous literature reviews and new data have been added to update any current models in air to plasma distribution. Weathersby used the Abraham solvation equation to correlate the log P data for blood to plasma with the four Abraham descriptors (**E**, **S**, **B** and **V**). The Abraham equation obtained by Weathersby and Abraham is shown as 11.1.[1] Note the '**aA**' was left out as appeared to be not significant in the equation.

$$\text{Log P}_{(\text{blood to plasma})} = -0.049 - 0.298\mathbf{E} + 0.422\mathbf{S} + 0.340\mathbf{B} - 0.344\mathbf{V} \qquad \text{Equation 11.1}$$

$r^2 = 0.660$ S.D = 0.086 n = 32

There have been no other reports in the literature, where models have been constructed for blood to plasma distribution for humans or rats. The correlative equation shall be constructed based on datasets that represent average human and rat data combined. The complete list of the VOC log P data for these averages, are shown in table 11.5.

Discussion on blood to plasma distribution for VOCs

The correlation for blood to plasma distribution for VOCs can be successfully constructed by using the Abraham solvation equation (see table 11.9). This model has 35 compounds in the dataset, but the fits for this new equation appear to be very poor (S.D = 0.119 and r^2 = 0.270). The reason for this, as suggested before, is that there is a very little difference between the experimental values of blood log K and of plasma log K for VOCs. This means that the spread of log P values is very small, 0.6 log units and this is only 3 times the standard error. This explains why the correlation between the two biological samples is poor. There were no outliers observed in the correlation. The equation for log P blood to plasma, with a very poor r^2 value and very small coefficients suggests that there is little difference between the two sets of data. air and into the plasma phase. The plot for this blood to plasma model also indicates that the data has not got an even spread and most VOC data is too cluttered in the middle of the graph (see page 153). This model is poor when compared to other good models like blood to brain distribution for VOCs, where there is an even spread of data (as shown from the plots in chapter 12). Although Weathersby and Abrahams model has a better fit for humans, the present model takes into account average log K data for air to plasma and air to blood (humans and rats combined). The final conclusion is that the air to blood and air to plasma distribution data can be combined for VOCs.

EQUATIONS	N	r^2	SD	F	D
FULL SET					
Log P (blood:plasma) = -0.056 – 0.227**E** + 0.103**S** - 0.268**A** + 0.303**B** – 0.201**V**	35	0.270	0.119	2.145	5

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

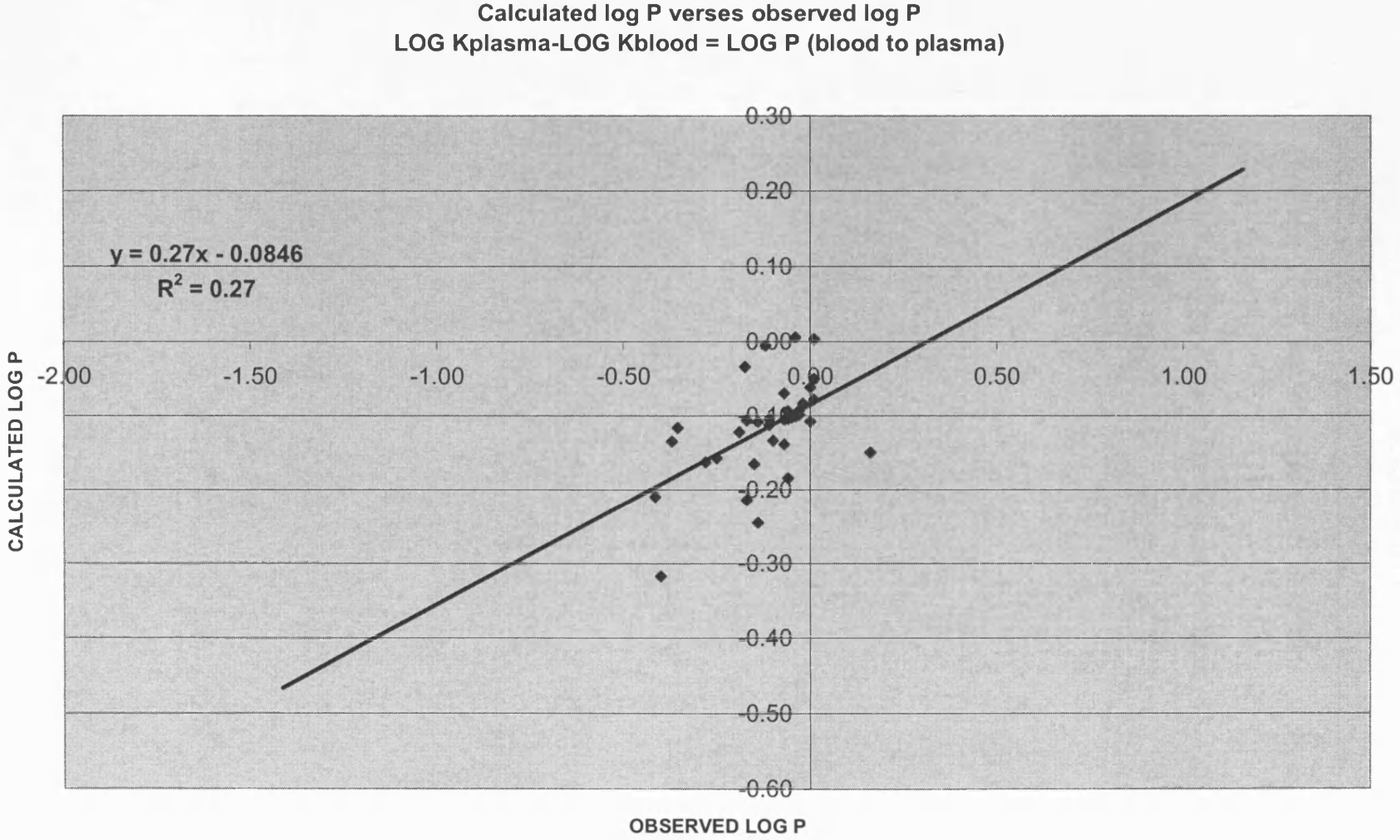
D = the number of descriptors used in the model

Table 11.9. Summary of the equations for VOC blood to plasma distributions

The inter-correlations of descriptors used for VOC blood to plasma distribution are shown in table 11.10. It appears that there is no strong inter-correlation between the set of descriptors used in this model.

r	E	S	A	B	V
E	1.00				
S	0.39	1.00			
A	0.20	0.38	1.00		
B	0.25	0.61	0.66	1.00	
V	0.05	0.57	0.20	0.39	1.00

Table 11.10. The inter-correlation of descriptors for 35 solutes using descriptors **E**, **S**, **A**, **B** and **V**.


Calculated log P verses observed log P
LOG Kplasma-LOG Kblood = LOG P (blood to plasma)
y = 0.27x - 0.0846
R2 = 0.27
CALCULATED LOG P
OBSERVED LOG P
0.30
0.20
0.10
0.00
-0.20
-0.30
-0.40
-0.50
-0.60
-2.00
-1.50
-1.00
-0.50
0.00
0.50
1.00
1.50

References

1. Abraham, M. H. and Weathersby, P. K (1994). J. Pharm. Sci, **83**, 1450-1455.
2. Guitart, R (1993). Revista Espanola de Fisiologia., **49**, 195-202.

Chapter

12

Air to brain, blood to brain and plasma to brain distribution (rat and human) for VOCs

In this chapter the Abraham solvation equation shall be used for models that include: air to brain distribution for VOCs, blood to brain distribution for VOCs and plasma to brain distribution for VOCs.

Air to brain distribution for VOCs

The distribution of VOCs from air to brain is of importance in medicine, toxicological sciences and environmental health. There has been a large amount of reported log K data for air to brain distribution.[1-15] Several cited papers, where models have been suggested to predict air to brain distribution in humans, rats or for both species shall be discussed. The purpose of this work is to add new literature data to the existing data set of Abraham and to make an attempt to construct new equations to predict these processes. Correlative and predictive equations shall be constructed based on datasets that represent average human and rat data combined. The complete list of the VOC log K data for these averages, are shown in table 12.0 (see appendix). The reported literature models for the correlation of air to brain distribution for VOCs are found from the following authors: Meulenberg and Vijverberg[1] and Abraham and Weathersby[3]

Meulenberg and Vijverberg[1] have obtained a reasonable amount of data used in a training set to produce two correlative equations for air to brain distribution in rats and humans. They found that the partition coefficients of VOCs in human (or

rat) brain are well described by a linear combination of $K_{air:oil}$ and $K_{air:saline}$ with specific regression coefficients for the biological sample i.e. brain. The partition data (K) for air to brain distribution were averaged values from literature sources and were used for the bilinear regression. Equations 12.0 and 12.1 show how $K_{air:brain}$ is described as a bilinear function of $K_{air:olive\ oil}$ and $K_{air:saline}$. The coefficients α_{oil} and α_{saline} of both equations represents the tissue-specific contributions of the lipophilic and hydrophilic interactions to the solubility of the VOCs in brain.

Human equation Equation 12.0

$K_{air:brain} = 0.020K_{air:olive\ oil} + 0.380K_{air:saline} + 0.94$

$r^2 = 0.98$ SD = n/a n = 35

Rat equation Equation 12.1

$K_{air:brain} = 0.054K_{air:olive\ oil} + 0.832K_{air:saline}$

$r^2 = 0.90$ SD = n/a n = 19

Note. The **c** coefficient for the rat equation was not determined by Meulenberg and Vijverberg.

Abraham and Weathersby[3] compiled a dataset of 41 solutes to predict air to brain distribution for humans. Most of the data was compiled from previous literature. Abraham and Weathersby used the Abraham solvation equation to correlate the log K data (Ostwald solubility coefficient at 310K) for air to brain with the five Abraham descriptors (**E, S, A, B** and **L**). The Abraham equation used for this model is shown in equation 12.2.[3]

$\text{Log K}_{(air:brain)} = -1.074 + 0.427\mathbf{E} + 0.286\mathbf{S} + 2.781\mathbf{A} + 2.787\mathbf{B} + 0.609\mathbf{L}$ Equation 12.2

$r^2 = 0.968$ S.D = 0.240 n = 41

Neither of the two models have any test sets that include statistical information that can be used to verify their predictions. It is very unfortunate that Meulenberg and Vijverberg did not calculate the fits (i.e. standard deviation) for the training set used in the correlation.

An improvement can be made on air to brain distribution by using training sets and test sets containing a large number of compounds (including statistical

information). It would be very interesting to see if the diversity of compounds helps to improve models that are currently available. The models discussed before are outdated and many new measurements have been made for VOCs which will be included in the new training sets (see appendix table 12.0).

Discussion on air to brain distribution for VOCs (humans and rats combined)

The Abraham solvation equation can successfully be employed to correlate VOC air to brain distribution (see table 12.1). Considering that this model has now 81 compounds in the dataset, the fits for this new equation appear to be quite good (S.D = 0.346 and $r^2 = 0.923$). This work is the most up to date model for VOC air to brain distribution; it appears that this is the most general model yet reported for both species (humans and rats).

There were two outliers observed from this training set: cyanoethylene oxide and methanol. The residual for these two compounds are approximately 1 log unit and may result from being an experimental error.

EQUATIONS	N	r^2	SD	F	D
FULL SET **Log K** (air:brain) = - 0.987 + 0.263**E** + 0.411**S** + 3.358**A** + 2.025**B** + 0.591**L**	81	0.923	0.346	178.980	5
TRAINING SET **Log K** (air:brain) = - 1.018 + 0.560**E** + 0.563**S** + 3.440**A** + 2.240**B** + 0.521**L** **TEST SET** **N** = 40 **RMSE** = 0.404 **AAE** = 0.347 **AE** = -0.136 **SD** (n-1) = 0.409	41	0.930	0.342	92.680	5

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = the root mean square error

AAE = absolute average error

AE = average error

SD = standard deviation

Table 12.1 Summary of the equations for VOC air to brain distributions and the statistics for test sets of 40 solutes

The inter-correlations of descriptors used for VOC air to brain distribution are shown in table 12.2. The largest r^2 value for S versus B descriptor is 0.32. It appears that there is no strong inter-correlation between the set of descriptors used in this model.

r	E	S	A	B	L
E	1.00				
S	0.36	1.00			
A	0.01	0.25	1.00		
B	0.04	0.57	0.45	1.00	
L	0.39	0.16	-0.08	0.10	1.00

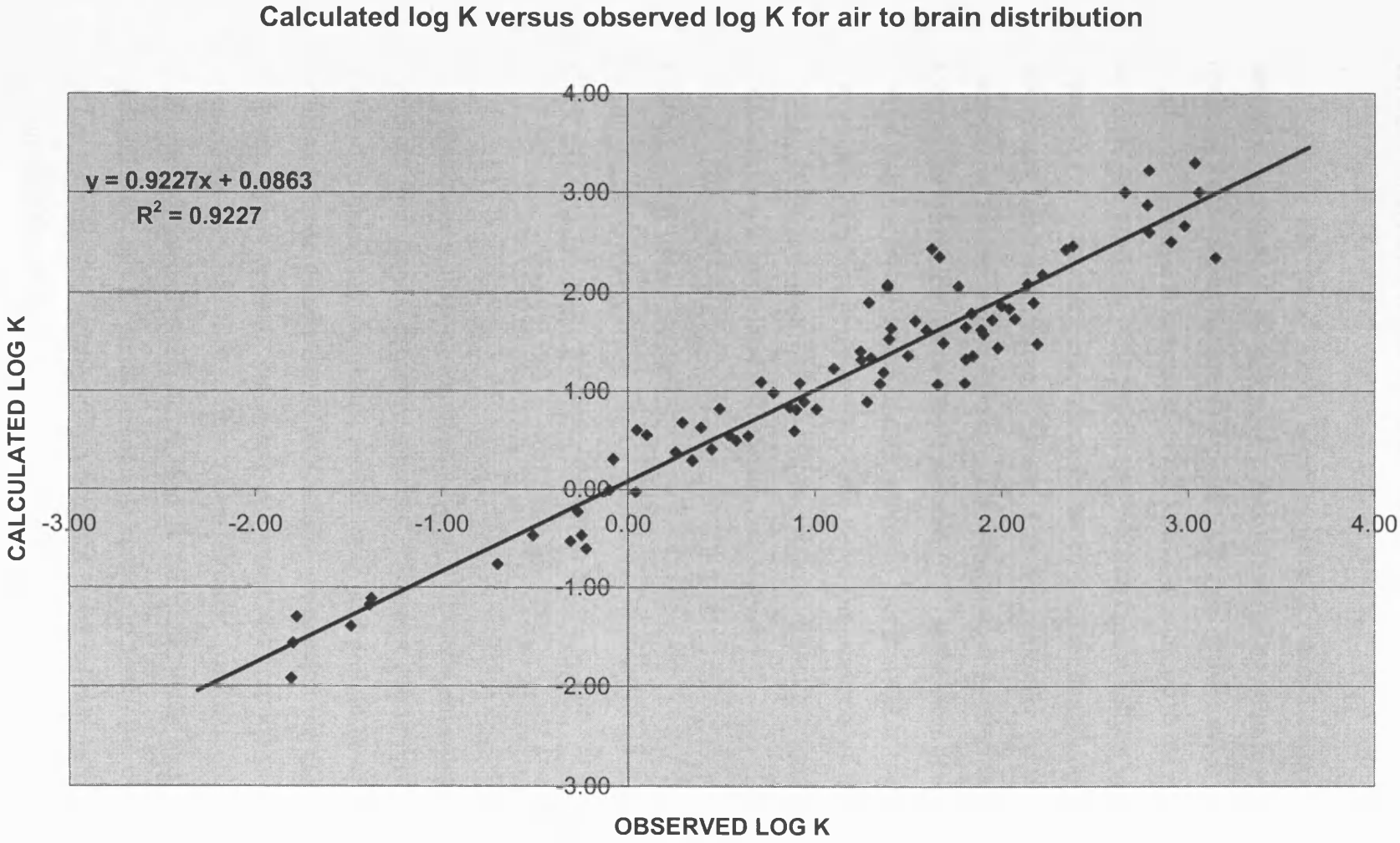
Table 12.2. The inter-correlation of descriptors for 81 solutes using descriptors **E**, **S**, **A**, **B** and **L**.

The Abraham solvation equation has successfully been employed to predict air to brain distribution for VOCs provided the solute is within the given descriptor range of the training set to make the correlative equations. To test the predictivity of the model it was necessary to split the full data into training and test set using a random data set generator. A new model was developed using only those data in the training set, which was then used to predict values for the test set. The full data set of 81 compounds was divided into two sets, using the first set to construct an equation similar to the first equation from table 12.1 and then 40 log K values (air to brain distribution of VOCs) are predicted for the second set. The regression yielded a similar training equation and statistics to the first equation in table 12.1. The test set of 40 compounds shows the average absolute error to be 0.347 for the two sets of data (observed and predicted log K). The root mean square error of 0.404 also indicates that there is a reasonable predictive capability for the equations shown in

table 12.1. The average error of -0.136 indicates that the equation of the model is not biased, as it is just below zero.

The **e**-coefficient (P = 0.15) and **s**-coefficient (P = 0.06) is not significant. The positive **l**-coefficient in the air to brain equation (12.1) indicate that increasing molecular size and London dispersion interaction can push VOCS out of air and into the brain tissue phase. The coefficients of the descriptors **A**, and **B** are all positive, indicating that the hydrogen-bond donor and acceptor ability is stronger in the brain phase than in gas phase. The solutes are pushed out of the air and into the brain phase. Finally, the Abraham equation can successfully be used to make predictions for air to brain distribution for a diverse range of solutes (VOCs).

The statistics for this model is good, as indicated by the large spread of data shown from the plot. It is found that to be possible to construct an equation capable of predicting further values of log K_{brain} to around 0.4 log units. This equation is not much better in terms of predictions when compared to the air to blood equation shown in chapter 10. This may be due to the fact there are fewer data points used in this model.


Calculated log K versus observed log K for air to brain distribution
y = 0.9227x + 0.0863
R2 = 0.9227
CALCULATED LOG K
OBSERVED LOG K
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00

Blood to brain distribution for VOCs

Zhang developed a non linear equation for blood to brain distribution of neutral compounds.[16] He compiled a list of 35 different volatile organic compounds obtained from various literatures sources. The first two equations in 12.3, takes into account the biological interaction of the VOCs in the brain tissue phase, with regards to lipids and proteins. The product of the first two equations is used in the third equation (12.3) to make predictions for blood to brain distribution for VOCs. The descriptors used in equation 12.3 are, α is the molecular polarizability (used to characterise the steric bulk effeck), $\sum Ca$ is the sum of H-bond factor values for all acceptor substructures in the molecule and $\sum Q^+$ is the sum of all positive partial atomic charges for all atoms in the molecule. His non-linear regression equations, was compared to the composition of the tissue (i.e. contents of lipids, proteins and water in brain and other tissues). Zhang concludes that his equations for blood to brain distribution are related to the tissue composition.

$$\text{Log } P_{(\text{lipid})} = 0.124\alpha - 0.714\sum Ca \qquad \text{Equation 12.3}$$

$$\text{Log } P_{(\text{protein})} = 0.630\sum Q^+$$

$$\text{Log } P_{(\text{blood to brain})} = \log(10^{\text{logP (lipid)}} + 10^{\text{logP(protein)}} + 1) - 0.642$$

$r^2 = 0.920$ SD = 0.120 n = 35

Balaz and Lukacova compiled a list of 35 experimental data of non-ionisable chemicals from several literature sources.[17] Their aim was to predict blood to brain distribution based on: membrane accumulation and protein binding. The equation for blood to brain distribution of VOCs takes into account the tissue composition of the brain (lipids, proteins and water) in humans and function of lipophilicity. The independent variable is P_{oct} (the 1-octanol/water partition coefficient) and the adjustable parameters were A_i, A_b and B for brain tissue and blood. β is the adjustable constants for series compounds binding to proteins in the bi-layer system) and I as the indicator variable. The fits for the blood to brain distribution model was good and this model had a better fit than other current published models.

$$\text{Log } P_{(\text{blood to brain})} = \beta \text{ x log } P_{oct} + \log\left(1+\sum_{i=1}^{2} Ai' \text{ x } I_i\right) - \log(P^{\beta} \text{ x } A_b + 1) + B' \qquad \text{Equation 12.4}$$

$r^2 = 0.956$ SD = 0.083 n = 35

The solvation equation obtained by Abraham and Weathersby was shown to be useful in order to predict blood to brain distribution for VOCs.[3] Their 39 blood to brain distribution dataset was compiled from previous literature reviews for human subjects. The solvation equation was used to correlate the log K data (Ostwald solubility coefficient at 310K) for blood to brain with the five Abraham descriptors (**E**, **S**, **A**, **B** and **V**). The Abraham equation used for this model is shown in equation 12.5.[3]

$$\text{Log } P_{(\text{blood:brain})} = -0.166 + 0.239\mathbf{E} - 0.626\mathbf{S} - 0.368\mathbf{A} - 0.615\mathbf{B} + 1.072\mathbf{V} \qquad \text{Equation 12.5}$$

$r^2 = 0.865$ S.D = 0.154 n = 39

Fiserova-Bergerova and Diaz used the head space chromatography method to determine experimentally air to brain distribution for seven VOCs.[2] They also compiled a literature survey of experimental values based on human subjects. Air to brain and air to blood partition coefficients were used to calculate blood to brain partition coefficients for VOCs. From the statistical analysis of Fiserova-Bergerova Diaz, it was found that there is a good inter-correlation between blood to brain partition coefficients and blood to fat partition coefficients as shown from a plot. The intercept and the gradient of the straight line (linear) plots were included in the model to make a correlative equation to predict blood to brain distribution of VOCs (see equation 12.6). It appears that there is no test set for this model which limits the information with regards to its models predictability of VOCs partition. The standard deviation of this model is very large and suggests that the equation is of little use in prediction.[2]

$$P_{(\text{blood to brain})} = 0.0372P_{\text{blood to fat}} + 0.5199 \qquad \text{Equation 12.6}$$

n = 35 $r^2 = 0.940$ SD = 0.779

P denotes to the appropriate partition of the VOCs

An improvement of the current models can be made on blood to brain distribution by using training sets and test sets containing a larger amount of data (including statistical information). It would be very interesting to see if the diversity of compounds helps to improve models that are currently available. Most of the models discussed before are outdated and many new measurements have been made recently and can be included in a new training set (see appendix table 12.3).

Discussion on blood to brain distribution for VOCs (humans and rats combined)

The solvation equation of Abraham can successfully be used to correlate VOC blood to brain distribution (see table 12.4). The model has 78 compounds in the dataset, and the fits for this new equation appear to be very good (S.D = 0.203 and r^2 = 0.725). This work is the most up to date model for VOC blood to brain distribution; it appears this is the most general model yet reported for both species (humans and rats).

There were two outliers observed from this training set: decane and cyclohexane. The residual for these two compounds are approximately between 0.5 to 1 log units and may result from experimental error.

EQUATIONS	N	r^2	SD	F	D
FULL SET **Log P** (blood:brain) = - 0.057 + 0.017**E** - 0.536**S** - 0.323**A** - 0.335**B** + 0.731**V**	78	0.725	0.203	37.908	5
TRAINING SET **Log P** (blood:brain) = 0.009 + 0.053**E** - 0.584**S** - 0.238**A** - 0.338**B** + 0.638**V** **TEST SET** **N** = 39 **RMSE** = 0.200 **AAE** = 0.155 **AE** = -0.027 **SD** (n-1) = 0.203	39	0.657	0.215	12.654	5

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = the root mean square error

AAE = absolute average error

AE = average error

Table 12.4 Summary of the equations for VOC blood to brain distributions and statistics for test sets of 39 solutes

The inter-correlations of descriptors used for VOC blood to brain distribution are shown in table 12.5. The largest r^2 value for S versus B descriptor is 0.31. It appears that there is no strong inter-correlation between the set of descriptors used in this model.

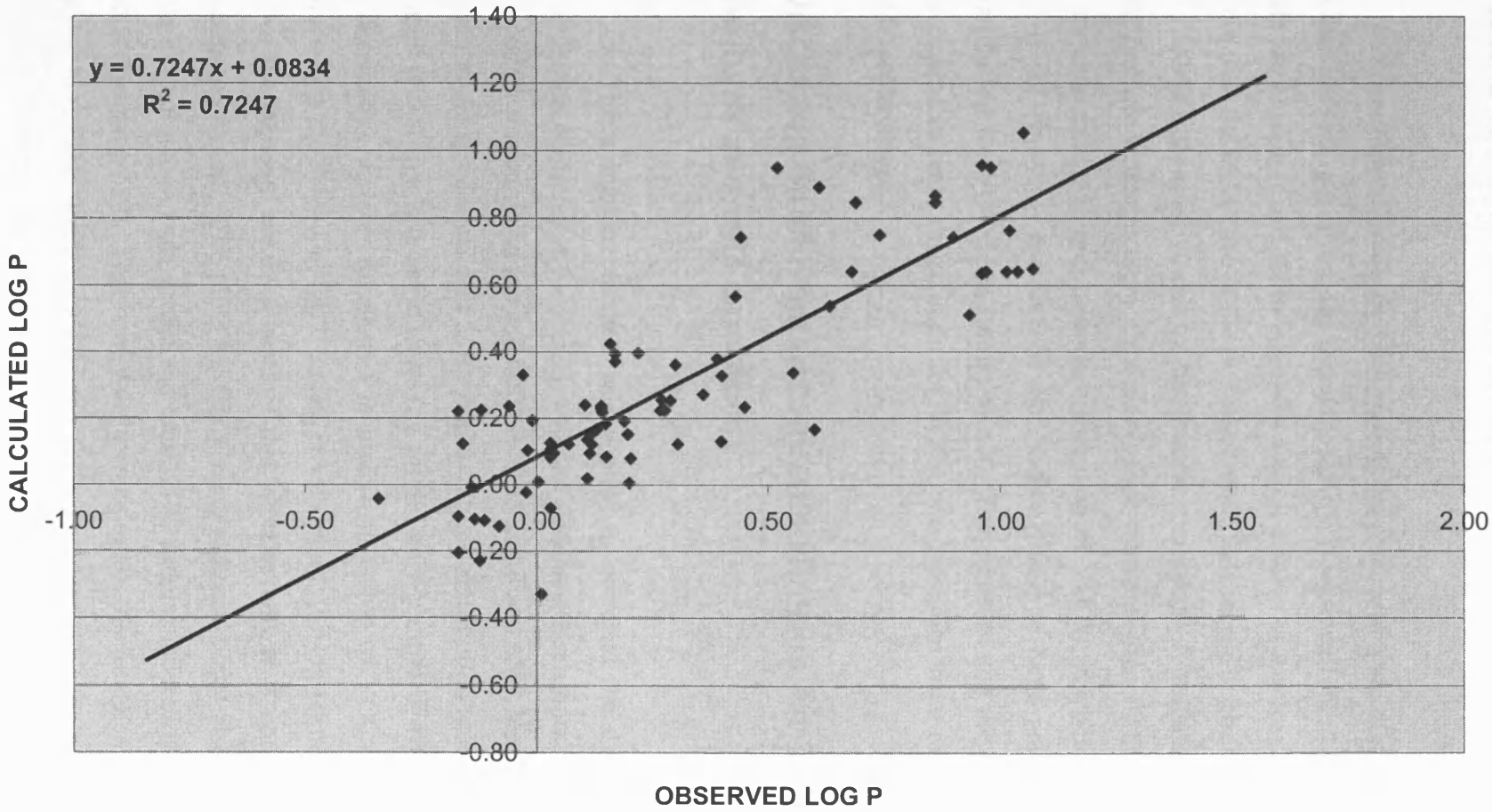
r	E	S	A	B	V
E	1.00				
S	0.36	1.00			
A	0.01	0.24	1.00		
B	0.03	0.56	0.45	1.00	
V	0.13	-0.04	-0.18	0.03	1.00

Table 12.5. The inter-correlation of descriptors for 78 solutes using descriptors **E**, **S**, **A**, **B** and **V**.

The Abraham solvation equation has successfully been employed to predict blood to brain distribution for VOCs, provided the solute is within the given descriptor range of the test set used to make the correlative equations. To test the predictivity of the model, it was necessary to split the full data into training and test set using a random data set generator. A new model is developed using only those data in the training set, which is then used to predict values for the test set. The full data set of 78 compounds is split into two sets of 39, using the first set to construct an equation similar to the first equation from table 12.4 and predicting 39 log K values (blood to brain distribution of VOCs) for the second set. The regression yielded similar training equation and statistics to the first equation in table 12.4. The test of 39 compounds shows the average absolute error to be 0.155 for the two sets of data (observed and predicted log K). The root mean square error of 0.200 also indicates that there is little difference between the observed value and predicted data

for blood to brain distribution for VOCs. The average error of -0.027 indicates that the equation of the model is not biased, as it is very close to zero.

The **e**-coefficient (P = 0.87) and **a**-coefficient (P = 0.16) is not significant. The positive **v**-coefficients in the blood to brain equation (12.4) indicate that increasing molecular size and London dispersion interaction can push VOCS out of blood and into the brain tissue phase. The coefficients of the 'polar' descriptors **S** and **B** are negative, indicating that polarity and hydrogen-bond acceptor ability is stronger in the blood phase than in brain phase. The solutes are pushed out of the brain and into the blood phase. From this analysis the Abraham equation can successfully be employed to make predictions for blood to brain distribution as shown from table 12.4, for a wide range of VOCs.


Calculated log P versus observed log P for blood to brain distribution for VOCs
y = 0.7247x + 0.0834
R2 = 0.7247
CALCULATED LOG P
OBSERVED LOG P
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
-0.80
-1.00
-0.50
0.00
0.50
1.00
1.50
2.00

Plasma to brain distribution for VOCs

There have been no reported models for plasma to brain distribution for VOCs in humans and rats. From the literature survey, it was possible to obtain distribution data (air to plasma and air to brain) for a small number of solutes (21 VOCs in total). From this, the difference in the two distribution data for air to plasma and air to brain (for VOCs), would give rise to distribution coefficients for solutes between plasma and brain. Although there are fewer solute numbers than in the previous model mentioned earlier (blood to brain distribution of VOCs), it would be interesting to see if the equations for both of the models (blood to brain and plasma to brain distribution for VOCs) are similar or not. The equation and the statistical analysis for plasma to brain distribution for VOCs, was compiled from data shown in table 12.6 (see appendix).

Discussion on plasma to brain distribution for VOCs (humans and rats combined)

The solvation equation was applied to correlate the 21 plasma to brain distribution coefficients for VOCS, with the Abraham descriptors. From table 12.7., the fits for this equation are good (SD = 0.153 and $r^2 = 0.619$), but coefficients of the equation are not comparable to the blood to brain distribution model of VOCs. The reason is that this data set is small (21 solutes only compared with 78) and is not a representative number for plasma to brain distribution of VOCs in general. This means that the spread of log P values is very small, 0.7 log units and this is only 3 times the standard error. Therefore this explains why it is not possible to get a reasonable model. The plots for this plasma to brain model also indicate that the data has not got an even spread and most VOC data is cluttered in the middle of the graph (see page 169). That is why the correlation for this model is poor and this may be due to the fact that the model has a limited amount of data points as shown from the plot. Whereas, the plot for blood to brain distribution has an even spread of data for a larger number of solutes and gives rise to a better correlative model than that of the plasma to brain model. There were no outliers observed from this training set.

No test sets have been used for this model as there are a limited number of compounds available.

EQUATIONS	N	r^2	SD	F	D
FULL SET **Log P** (plasma:brain) = 0.049 + 0.037**E** - 0.211**S** + 0.361**A** - 0.829**B** + 0.592**V**	21	0.619	0.153	4.881	5

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

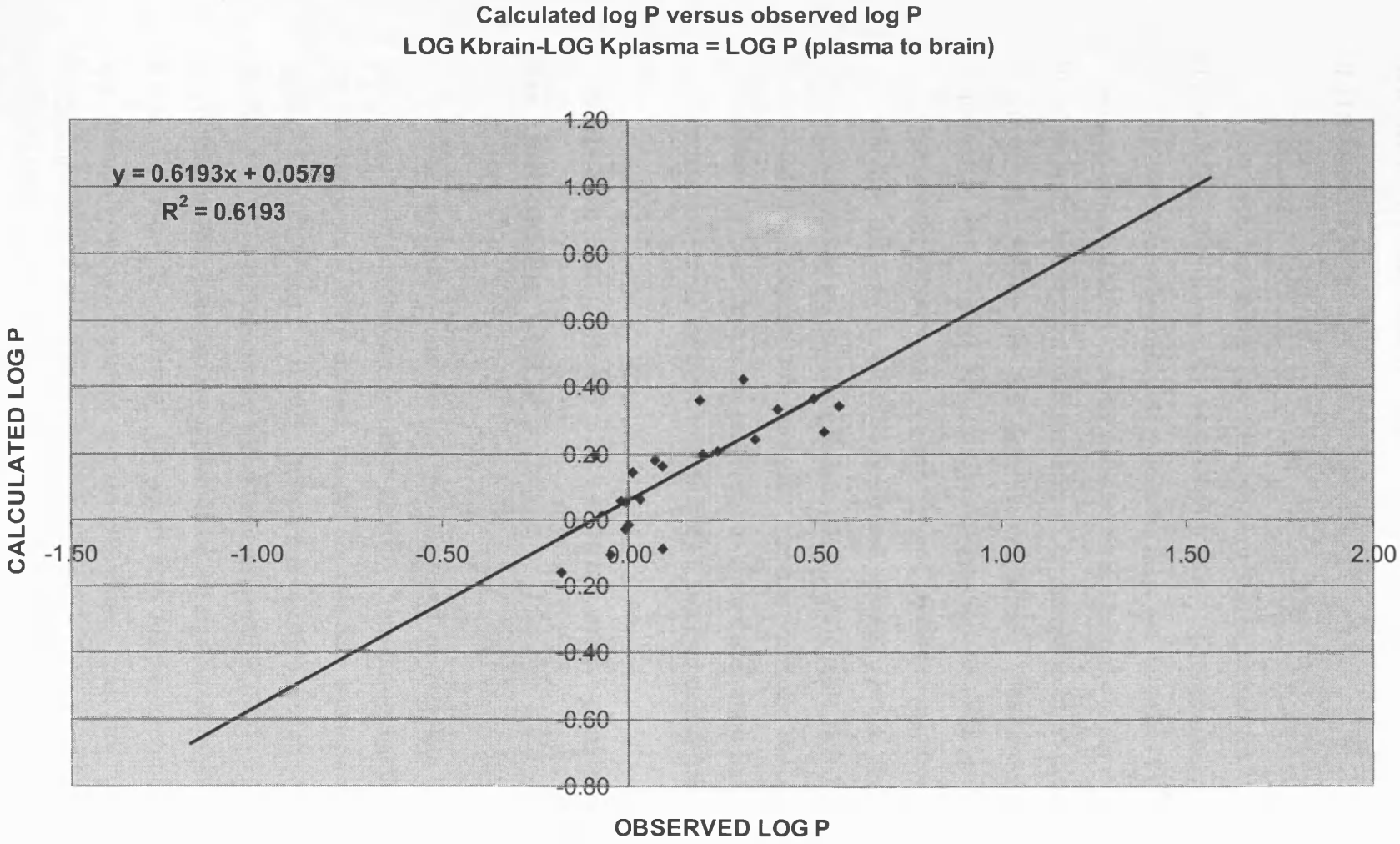
D = the number of descriptors used in the model

Table 12.7 Summary of the equations for VOC plasma to brain distributions

Finally the inter-correlations of descriptors used for VOC plasma to brain distribution are shown in table 12.8. The largest r^2 value for A versus B descriptor is 0.42. It appears that there is no strong inter-correlation between these sets of descriptors used in this current model.

r	E	S	A	B	V
E	1.00				
S	0.47	1.00			
A	0.40	0.36	1.00		
B	0.50	0.64	0.65	1.00	
V	-0.02	0.64	0.17	0.38	1.00

Table 12.8 The inter-correlations of descriptors used for VOC plasma to brain distribution for 21 solutes


Calculated log P versus observed log P
LOG Kbrain-LOG Kplasma = LOG P (plasma to brain)
y = 0.6193x + 0.0579
R2 = 0.6193
CALCULATED LOG P
OBSERVED LOG P
1.20
1.00
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
-0.80
-1.50
-1.00
-0.50
0.00
0.50
1.00
1.50
2.00

Blood to brain combined with plasma to brain distribution data for VOCs

It is important to find out if the blood to brain equation for VOC is the same as the equation for plasma to brain distribution. Unfortunately, examination of coefficients can be very misleading because there are only data 21 data points in the plasma to brain equation. Another way of investigation is to combine the dataset for blood to brain distribution with the data set of plasma to brain distribution and then construct an equation for full data set combined. This equation can then be compared to the original blood to brain equation for VOCs to see if the two equations are the same or not. Correlative and predictive equations of Abraham for the combined dataset (i.e. blood to brain and plasma to brain) shall be constructed. The complete list of the VOC log K data, are shown in table 12.9 (see appendix).

Finally, to support this analysis, one can compare the common partition coefficients for plasma to brain and blood to brain distribution for 19 VOCs and see if it is possible to combine these two dataset for these processes (table 12.12).

Discussion on blood to brain and plasma to brain distribution for VOCs

The solvation equation can be employed to correlate VOC (blood and plasma) to brain distribution (see table 12.10). Considering that this model has now got 99 compounds in the dataset, the fits for this new equation appear to be very good (S.D = 0.196 and $r^2 = 0.703$). In fact, this equation for full set has an improved fit (SD = 0.196) over the current blood to brain equation (SD = 0.203). Note the '**e**' coefficient for the combined set appears to be not significant in this model, although it was still used in the equation. The coefficients for the combined data set, (blood and plasma to brain) equation for the 99 solutes, has more or less the same coefficients (**c**, **e**, **s**, **b** and **v**) as the blood to brain equation. There is one exception where the '**a**' coefficient for the combined equation appeared to be smaller than in the blood to brain equation. This suggests that blood to brain log P data for VOCs can be combined together (if needed) with the plasma to brain log P data, as they do

not appear to be different from each other (as shown by the coefficients of the equations and the values of r^2 and SD).

EQUATIONS	N	r^2	SD	F	D
FULL SET					
Log P ((b+p):brain) = - 0.028 + 0.003**E** - 0.485**S** - 0.117**A** - 0.408**B** + 0.703**V**	99	0.703	0.197	44.062	5
Log P (blood:brain) = - 0.057 + 0.017**E** - 0.536**S** - 0.323**A** - 0.335**B** + 0.731**V**	78	0.725	0.203	37.908	5
TRAINING SET					
Log P ((b+p):brain) = 0.023 + 0.074**E** - 0.655**S** - 0.117**A** - 0.283**B** + 0.652**V**	50	0.700	0.206	20.495	5
TEST SET					
N = 49 **RMSE** = 0.195 **AAE** = 0.151					
AE = -0.003 **SD** (n-1) = 0.197					

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = the root mean square error

AAE = absolute average error

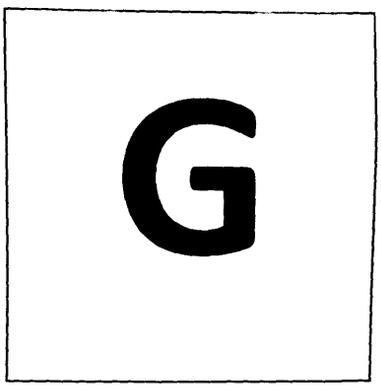
AE = average error

b = blood

p = plasma

Table 12.10 Summary of the equations for VOC blood (and plasma) to brain distributions and statistics for test sets of 99 solutes

The Abraham solvation equation has been employed to predict blood and plasma to brain distribution for VOCs provided the solute is within the given descriptor range of the test set to make the correlative equations. Note that no outliers were found in this combined data set (blood to brain and plasma to brain). The values of the coefficients for the blood and plasma to brain equation are nearly the same as the coefficients in the blood to brain equation (shown in blue in table 12.10).


G

The inter-correlations of descriptors used for VOC blood and plasma to brain distribution are shown in table 12.11. The largest r^2 value for S versus B descriptor is 0.34. It appears that there is no strong inter-correlation between the set of descriptors used in this model.

r	E	S	A	B	V
E	1.00				
S	0.37	1.00			
A	0.07	0.27	1.00		
B	0.11	0.58	0.50	1.00	
V	0.15	0.05	-0.16	0.05	1.00

Table 12.11. The inter-correlation of descriptors for 99 solutes using descriptors **E, S, A, B** and **V**.

SOLUTE NAME	Log P (plasma to brain)	
Argon	**0.09**	**0.03**
Krypton	**-0.09**	**-0.16**
Xenon	**0.34**	**0.15**
Nitrogen	**0.07**	**0.03**
Nitrous Oxide	**0.03**	**0.03**
Methane	**0.20**	**0.03**
Cyclopropane	**0.53**	**0.11**
CF3CBrClH (Halothane)	**0.31**	**0.14**
Diethyl ether	**-0.02**	**-0.01**
CHF2OCHClCF3 (Isoflurane)	**0.19**	**0.19**
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	**0.24**	**0.14**
Acetone (Propan-2-one)	**-0.18**	**-0.17**
Butan-2-one (Methyl ethyl ketone)	**-0.05**	**-0.17**
Ethanol	**-0.01**	**-0.12**
1-Propanol	**-0.01**	**-0.08**
2-Propanol (Isopropanol)	**0.00**	**-0.11**

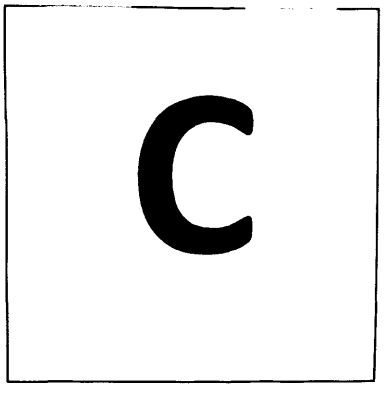
2-Methyl-1-propanol (Isobutanol)	**0.01**	**-0.14**
Sulphur hexafluride	**0.57**	**0.39**
CHF2OCF2CHFCl (Enflurane)	**0.49**	**0.14**

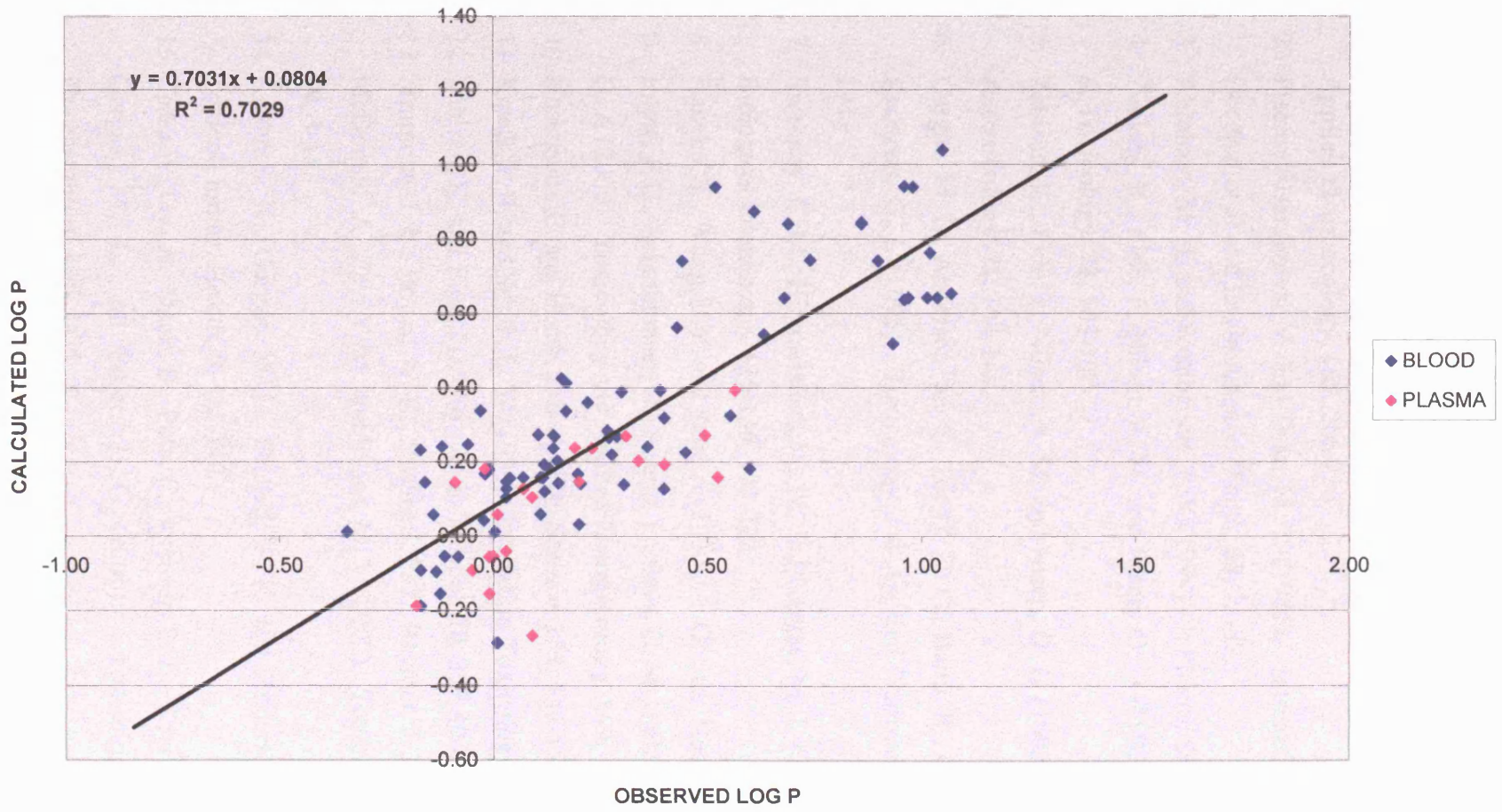
SD (n-1) =	**0.168**
RMSE =	**0.164**
AAE =	**0.122**
AE (plasma to brain - blood to brain) =	**0.120**
N =	**19**

Table 12.12. 19 common compounds, for plasma to brain and blood to brain distribution, which includes the statistics between the two columns of data.

From table 12.12, the 19 common compounds for plasma to brain and blood to brain distribution for VOCs appear to be slightly different. This is also supported by the plot shown page on page 174, where the plasma to brain distribution data appears to below the line of the regression for the whole plot.

The fact that the equation for blood to brain and plasma (plus blood) to brain are nearly the same is mostly due to the fact that they have 78 common compounds. The total data set used in the (plasma plus blood) to brain equations has the 78 blood to brain VOCs and only 21 plasma to brain VOCs. Therefore the best statistical parameter is the average error between the two partitions for the 19 data set. This is 0.12 log unit and definitely indicates that plasma to brain and blood to brain distribution are not quite the same. However the SD in the blood to brain equation is 0.20 log unit, so that it may be justified to combine the two data sets, even though they are not quite the same.


C


Calculated log P versus observed log P of blood or plasma to brain distribution for VOCs
y = 0.7031x + 0.0804
R2 = 0.7029
CALCULATED LOG P
OBSERVED LOG P
1.40
1.20
1.00
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
-1.00
-0.50
0.00
0.50
1.00
1.50
2.00
BLOOD
PLASMA

References

1. Meulenberg, C. J. W. and Vijverberg, H. P. M (2000). Toxicology and Applied Pharmacology, **165**, 206-216.
2. Fiserova-Bergerova, V. and Diaz, M. L (1986). International Archives of Occupational and Environmental Health, **58**, 75-87.
3. Abraham, M. H. and Weathersby, P. K (1994). J. Pharm. Sci, **83**, 1450-1455.
4. Zahlsen, K., Eide, I., Nilsen, A. M. and Nisen, O. G (1993). Pharmacology & Toxicology, **73**, 163-168.
5. Zahlsen, K., Eide, I., Nilsen, A. M. and Nisen, O. G (1992). Pharmacology & Toxicology, **71**, 144-149.
6. Gargas, M. L., Anderson, M. E., Teo, S. K. O., Batra, R., Fennell, T. R. and Kedderis, G. L (1995). Toxicology and Applied Pharmacology, **134**, 185-194.
7. Sweeney, L. M., Himmelstein, M. W. and Gargas, M. L (2001). Chemical-Biological Interactions, **135-136**, 303-322.
8. Kaneko, T., Wang, P-Y. and Sato, A (2000). J. Occup. Health, **42**, 86-87.
9. Filser, J. G., Schmidbauer, R., Rampf, F., Baur, C. M., Putz, C. and Csanady, G. A (2000). Toxicology and Applied Pharmacology, **169,** 40-51.
10. Nussbaum, E. and Hursh, J. B (1957). Science, **125**, 552-553.
11. Knaak, J. B. and Smith, L. W (1998). Inhalation Toxicology, **10**, 65-85.
12. Zahlsen, K. and Eide, I (1996). Arch. Toxicol, **70**, 397-404.
13. Simmons, J. E., Boyes, W. K., Bushnell, P. J., Raymer, J. H., Limsakun, T., McDonald, A., Sey, Y. M. and Evans, M. V (2002). Toxicological Sciences, **69**, 3-15.
14. Krishnan, K., Gargas, M. L., Fennell, T. R. and Anderson, M. E (1992). Toxicol. Indust. Health, **8**, 121-140.
15. Csanady, Gy. A., Denk, B., Putz, C., Kreuzer, P. E., Kessler, W., Baur, C., Gargas, M. L. and Filser, J. G (2000). Toxicology and Applied Pharmacology, **165**, 1-26.
16. Zhang, H (2004). J. Pharmaceutical Sciences, **93**, 1595-1604.
17. Balaz, S and Lukacova, V (1999). Quant. Struct.-Act. Relat., **18**, 361-368.

Chapter

13

Air to kidney and blood to kidney distribution for VOCs (human and rats)

In this chapter the Abraham solvation equation shall be used to construct models that include air to kidney and blood to kidney distribution for VOCs. This model will contain training and test sets for air to kidney distribution and blood to kidney distribution.

Air to kidney distribution for VOCs

The distribution of VOCs from air to kidney is of importance in environmental toxicology and health. There have been only a relatively small number of citations for air to kidney distribution data for VOCs.[1-15] The aim of this work is to gather a set of existing data from literatures sources and to attempt to construct a new equation to predict these distribution processes. The correlative equations of Abraham shall be constructed based on datasets that represent average human and rat data combined. The complete list of the VOC log K data for these literature averages is shown in table 13.0 (see appendix). The reported literature models for the correlation of air to kidney distribution for VOCs are from Meulenberg & Vijverberg[1] and Weathersby and Abraham.[4]

Meulenberg and Vijverberg have used a small amount of literature data in their training set to produce two correlative equations for air to kidney distribution in rats and humans.[1] They found that the partition coefficients of VOCs in human (or rat) kidney are well described by a linear combination of $K_{air:oil}$ and $K_{air:saline}$ with specific

regression coefficient for the biological sample i.e. kidney. The partition data (K) for air to kidney distribution are averaged values from literature sources and have been used for the bilinear regression. The two equations 13.0 and 13.1 show how $K_{air:kidney}$ is described as a bilinear function of $K_{air:olive\ oil}$ and $K_{air:saline}$. The coefficients α_{oil} and α_{saline} of both equations represents the tissue-specific contributions of the lipophilic and hydrophilic interactions to the solubility of the VOCs in kidney. The c constant has no specific physical meaning but it was required by Meulenberg and Vijverberg to compensate for systematic errors in kidney, oil, or saline partition coefficients. It appears that none of the two models have any test sets that includes statistical information that can be used to verify their predictions. It was very unfortunate that Meulenberg and Vijverberg did not calculate the fits (i.e. standard deviation) for the obtained correlation.[1]

Human equation Equation 13.0

$K_{(air\ to\ kidney)} = 0.011K_{air:olive\ oil} + 0.400K_{air:saline} + 0.69$

$R^2 = 0.98$ SD = n/a n = 34

Rat equation Equation 13.1

$K_{(air\ to\ kidney)} = 0.097K_{air:olive\ oil} + 0.826K_{air:saline}$

$R^2 = 0.91$ SD = n/a n = 16

Weathersby and Abraham compiled a dataset of 36 VOCs in order to correlate air to kidney distribution for humans.[4] They used the Abraham solvation equation to correlate the log K data (310K) for air to kidney with the five Abraham descriptors (**E**, **S**, **A**, **B** and **L**). The Abraham equation used for this model is shown in equation 13.2.[4]

$\text{Log } K_{(air\ to\ human\ kidney)} = -1.084 + 0.417\mathbf{E} + 0.226\mathbf{S} + 3.624\mathbf{A} + 2.926\mathbf{B} + 0.534\mathbf{L}$ Equation 13.2

$r^2 = 0.951$ S.D = 0.266 n = 36

It appears that both of the above models have not got any test sets that can be used to test predictions; this may be due to the limited number of compounds used in

the air to kidney distribution model. It was very unfortunate that Meulenberg & Vijverberg did not calculate the fits (i.e. standard deviation) for the training set used in the correlation.

Some improvement to these models can be made by compiling new VOC data to see if the new equation for the full set has better fits than the published models. With larger data sets for air to kidney distribution, it would be possible to construct training sets and test sets. Table 13.0 (see appendix) contains the up to date literature data for air to kidney distribution for VOCs.[1-15]

Discussion on air to kidney distribution for VOCs (humans and rats combined)

The Abraham solvation equation has been employed to correlate VOC air to kidney distribution as shown in table 13.1.

EQUATIONS	N	r^2	SD	F	D
FULL SET **Log K** (air to kidney) = -1.154 + 0.208**E** + 0.382**S** + 2.919**A** + 2.533**B** + 0.651**L**	76	0.923	0.327	167.846	5
TRAINING SET **Log K** (air to kidney) = -1.223 + 0.390**E** + 0.408**S** + 3.381**A** + 2.627**B** + 0.640**L** **TEST SET** **N** = 38 **RMSE** = 0.298 **SD** (n-1) = 0.302 **AAE** = 0.238 **AE** = -0.022	38	0.907	0.380	62.741	5

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = root mean square error

Table 13.1. Summary of the equations for VOC air to kidney distributions for humans and rats.

This model has now 76 compounds in the dataset, and the fits for the new equation appear to be good (S.D = 0.327 and r^2 = 0.923), considering now there is a diverse range of VOCs. Although, the standard deviation is little larger than the earlier model of Abraham and Weathersby, this work is the most up to date model for VOC air to kidney distribution model yet reported for both species (humans and rats). There were two outliers observed from this full set: methanol and methylcyclohexane; the residuals for the two outliers are approximately 1 log unit.

r	E	S	A	B	L
E	1.00				
S	0.32	1.00			
A	0.00	0.24	1.00		
B	-0.05	0.52	0.46	1.00	
L	0.31	0.02	-0.15	-0.01	1.00

Table 13.2. The inter-correlation of descriptors for 76 solutes using descriptors **E**, **S**, **A**, **B** and **L** for air to kidney distribution model.

The inter-correlations of the descriptors (**E**, **S**, **A**, **B** and **L**) used for VOC air to kidney distribution are shown in table 13.2. The largest r^2 value for S versus B descriptor is 0.27. It appears that there is no strong inter-correlation between the set of descriptors used in this air to kidney distribution model.

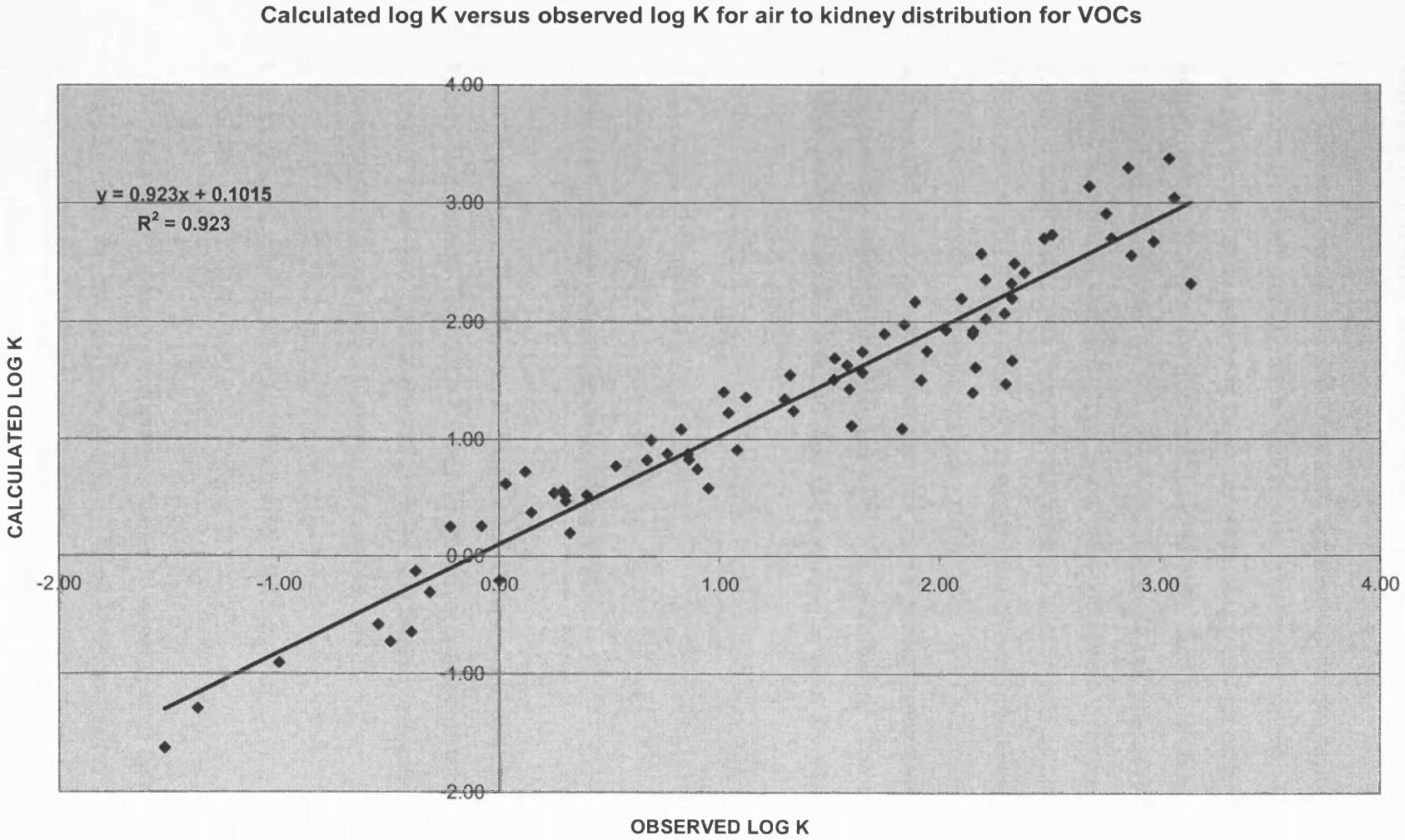
The Abraham solvation equation has been employed to predict air to kidney distribution for VOCs, provided the solute is within the given descriptor range of the training set to make the correlative equations. To examine the predictive ability of the model, it was necessary to split the full data into training and test set using a random data set generator (Kennard-Stone). A new model was developed using only those data in the training set, which was then used to predict values for the test set. The regression yielded similar training equations and statistics to the first equation in table 13.1, although, the fits of the training set appeared to be littler poorer than the full set.

The test set of 38 compounds shows the average absolute error to be 0.238 for the two sets of data (observed and predicted log K). The root mean square error of 0.298 also indicates that there is a small difference between the observed value and

predicted data for air to kidney distribution for VOCs. The average error of -0.022 indicates that the equation of the model is not biased, as it is very close to 0.

The statistics for this model is satisfactory, as this is indicated by the even spread of data shown from the plot. It is found that it is possible to construct an equation capable of predicting further values of log K_{kidney} to around 0.3 log units. This equation is not much better in terms of predictions when compared to the air to saline or air to olive oil equation shown in chapter 18 and 19. This maybe due to the fact there are other processes like metabolism (cytochrome-p450 enzymes) and protein binding taking place within the kidney.

The **e**-coefficient (P = 0.24) and **s**-coefficient (P = 0.07) is not significant. The positive **l**-coefficient in the air to kidney equation (13.1) indicates that increasing molecular size and London dispersion interaction can push VOCs out of air and into the kidney tissue phase. The coefficients of the descriptors **A** and **B** are all positive, indicating that the hydrogen-bond donor and acceptor ability is stronger in the kidney tissue phase than in the gas phase. The solutes are pushed out of the air and into the kidney phase. Finally, the Abraham equation can successfully be used to correlate and predict air to kidney distribution for humans and rats (see graph on page 181).


Calculated log K versus observed log K for air to kidney distribution for VOCs
y = 0.923x + 0.1015
R2 = 0.923
CALCULATED LOG K
OBSERVED LOG K
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00

Blood to kidney distribution for VOCs

Fiserova-Bergerova and Diaz used the head space chromatography method to determine experimentally air to kidney distribution for several VOCs.[2] They also compiled a literature survey of experimental values based on human subjects. Air to kidney and air to blood partition coefficients were used to calculate blood to kidney partition coefficients for VOCs. From the statistical analysis of Fiserova-Bergerova Diaz, it was found that there is a good inter-correlation between blood to kidney partition coefficients and blood to fat partition coefficients as shown from a plot. The intercept and the gradient of the straight line (linear) plots were used to make a correlative equation for blood to kidney distribution of VOCs (see equation 13.3). It appears that there is no test set for this model which limits the information with regards to the model's predictability of VOCs partitioning. The standard deviation of this model is very large, so that the equation is poor by comparison to other models mentioned latter.

$P_{(blood\ to\ kidney)} = 0.0211P_{blood\ to\ fat} + 0.6442$ Equation 13.3

$n = 35 \qquad r^2 = 0.945 \qquad SD = 0.421$

P denotes to the appropriate partition of the VOCs

A solvation equation was obtained by Abraham and Weathersby for blood to kidney distribution for VOCs.[4] Their 35 blood to kidney distribution dataset was compiled from previous literature reviews for human subjects. The Abraham equation used for this model is shown in equation 13.4.[4]

$\text{Log } P_{(blood\ to\ kidney)} = -0.188 + 0.226\mathbf{E} - 0.559\mathbf{S} - 0.433\mathbf{B} + 0.832\mathbf{V}$ Equation 13.4

$r^2 = 0.774 \qquad S.D = 0.168 \qquad n = 35$

Balaz and Lukacova compiled a list of 33 experimental data of non-ionisable chemicals from several literature sources.[16] Their aim was to predict blood to kidney distribution based on: membrane accumulation and protein binding. The equation for blood to kidney distribution of VOCs takes into account the tissue composition of the kidney (lipids, proteins and water) in humans and function of lipophilicity. The independent variable is P_{oct} (1-octanol/water partition coefficient) and the adjustable

parameters were A_i, A_b and B for tissue and blood. β is the adjustable constants for series compounds binding to proteins in the bi-layer system) and I as the indicator variable. Fits for the blood to kidney distribution model were very good and this model had a better fit than other published models.

$$\text{Log } P_{(\text{blood to kidney})} = \beta \text{ x log } P_{oct} + \log\left(1+\sum_{i=1}^{2} Ai' \text{ x } I_i\right) - \log(P^{\beta} \text{ x } A_b + 1) + B' \qquad \text{Equation 13.5}$$

$r^2 = 0.924$ SD = 0.089 n = 33

Zhang developed a non linear equation for blood to kidney distribution of neutral compounds.[17] He compiled a list of 34 different volatile organic compounds obtained from various literatures sources. The first two equations in 13.6, takes into account the biological interaction of the VOCs in the kidney tissue phase, with regard to lipids and proteins. The product of these two equations is used in the third equation (13.6), to make predictions for blood to kidney distribution for VOCs. The descriptors used in equation 13.6 are: α is the molecular polarizability (used to characterise the steric bulk effeck), Ca^{max} is the maximum H-bond acceptor descriptor in a molecule and $\sum Q^+$ is the sum of all positive partial atomic charges for all atoms in the molecule.

$$\text{Log } P_{(\text{lipid})} = 0.094\alpha - 3.661 Ca^{max} \qquad \text{Equation 13.6}$$

$$\text{Log } P_{(\text{protein})} = 0.637 \sum Q^+$$

$$\text{Log P } (_{\text{blood to kidney}}) = \log(10^{\text{logP (lipid)}} + 10^{\text{logP(protein)}} + 1) - 0.562$$

$r^2 = 0.897$ SD = 0.110 n = 34

The aim of the current work was to construct an equation as part of a general method for 'high-throughput' prediction of equilibrium blood-kidney distribution based on structure using the Abraham descriptors (**E**, **S**, **A**, **B** and **V**). Table 13.3 (see appendix) has the observed and calculated log P data for blood to kidney distribution of VOCs.

Discussion on blood to kidney distribution for VOCs (humans and rats combined)

The Abraham solvation equation has been employed to correlate VOC blood to kidney distribution as shown in table 13.4.

EQUATIONS	N	r^2	SD	F	D
FULL SET **Log P** (blood to kidney) = -0.361 + 0.134**E** - 0.515**S** - 0.823**A** + 0.273**B** + 0.981**V**	73	0.757	0.217	41.661	5
FULL SET TRAINING SET **Log P** (blood to kidney) = -0.305 + 0.151**E** - 0.451**S** - 0.797**A** + 0.195**B** + 0.881**V** **TEST SET** **N** = 36 **RMSE** = 0.189 **SD** (n-1) = 0.192 **AAE** = 0.153 **AE** = -0.036	37	0.678	0.253	13.003	5

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = root mean square error

Table 13.4. Summary of the equations for VOC blood to kidney distributions for humans and rats

The solvation equation of Abraham can be used to correlate VOC blood to kidney distribution (see table 13.4). This model has 73 compounds in total in the dataset and the fits for this new equation appear to be reasonably good (S.D = 0.217 and r^2 = 0.757). Although the fits for this model are poorer than the previous equations for the blood to kidney distribution, the present model is much more general. There were three outliers observed in the training set: 1,2-dimethylcyclo-hexane, decane and 1,2,4-trimethylcyclohexane. The residual for these three compounds are approximately 0.6-0.8 log units and may result from experimental error.

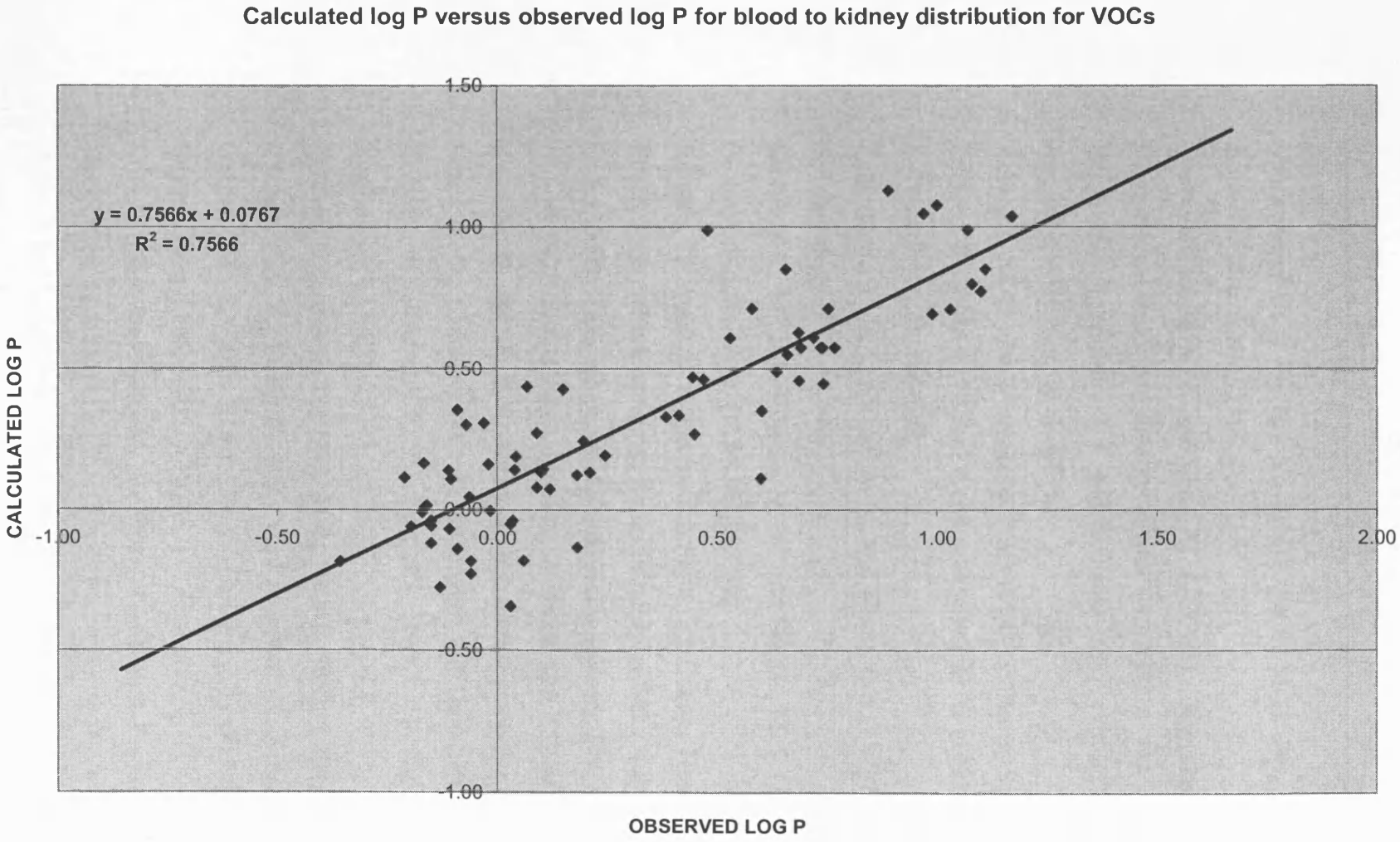
r	E	S	A	B	V
E	1.00				
S	0.32	1.00			
A	0.00	0.23	1.00		
B	-0.04	0.51	0.45	1.00	
V	0.00	-0.18	-0.22	-0.02	1.00

Table 13.5. The inter-correlation of descriptors for 73 solutes using descriptors **E**, **S**, **A**, **B** and **V**.

The inter-correlations of descriptors used for VOC blood to kidney distribution are shown in table 13.5. The largest r^2 value for S versus B descriptor is 0.26. It appears that there is no strong inter-correlation between the set of descriptors used in this model.

The Abraham solvation equation has been employed to correlate blood to kidney distribution for VOCs. To test the predictiveness of the model it was necessary to split the full data into training and test set using a random data set generator. A new model was developed using only those data in the training set, which was then used to predict values for the test set. The regression yielded a similar training equation and statistics to the first equation in table 13.4. The test set of 36 compounds shows the average absolute error to be 0.153 for the observed and predicted log K. The root mean square error of 0.189 also indicates that there is a small difference between the observed value and predicted data for air to kidney distribution for VOCs. The average error of -0.036 indicates that the equation of the model is not biased, as it is very close to 0.

The **e**-coefficient (P = 0.23) and **b**-coefficient (P = 0.08) is not significant. The positive **v**-coefficients in the blood to kidney equation (13.4) indicate that increasing molecular size and London dispersion interaction can push VOCs out of blood and into the kidney tissue phase. The coefficients of the 'polar' descriptors **S** and **A** are all negative, indicating that polarity and hydrogen-bond donor is stronger in the blood phase than in the kidney tissue phase. The solutes are pushed out of the kidney and into the blood phase. Finally the Abraham equation can successfully be used to correlate and predict blood to kidney distribution for humans and rats (see graph on page 186).


Calculated log P versus observed log P for blood to kidney distribution for VOCs
y = 0.7566x + 0.0767
R2 = 0.7566
CALCULATED LOG P
OBSERVED LOG P
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.00
-0.50
0.00
0.50
1.00
1.50
2.00

References

1. Meulenberg, C. J. W. and Vijverberg, H. P. M (2000). Toxicology and Applied Pharmacology, **165**, 206-216.
2. Fiserova-Bergerova, V. and Diaz, M. L (1986). International Archives of Occupational and Environmental Health, **58**, 75-87.
3. Borghoff, S. J., Murphy, J. E. and Medinsky, M. A (1996). Fundamental and Applied Toxicology, **30**, 264-275.
4. Abraham, M. H. and Weathersby, P. K (1994). J. Pharm. Sci, **83**, 1450-1455.
5. Zahlsen, K., Eide, I., Nilsen, A. M. and Nisen, O. G (1993). Pharmacology & Toxicology, **73**, 163-168.
6. Zahlsen, K., Eide, I., Nilsen, A. M. and Nisen, O. G (1992). Pharmacology & Toxicology, **71**, 144-149.
7. Sweeney, L. M., Himmelstein, M. W. and Gargas, M. L (2001). Chemical-Biological Interactions, (**135-136**), 303-322.
8. Kaneko, T., Wang, P-Y. and Sato, A (2000). J. Occup. Health, **42**, 86-87.
9. Filser, J. G., Schmidbauer, R., Rampf, F., Baur, C. M., Putz, C. and Csanady, G. A (2000). Toxicology and Applied Pharmacology, **169**, 40-51.
10. Nussbaum, E. and Hursh, J. B (1957). Science, **125**, 552-553.
11. Knaak, J. B. and Smith, L. W (1998). Inhalation Toxicology, **10**, 65-85.
12. Filser, J. G., Csanady, GY. A., Hartmann, M., Denk, B., Kauffmann, A., Kessler, W., Kreuzer, P. E., Putz, C., Shen,J. H. and Stei, P (1996). Toxicology, **113**, 278-287.
13. Zahlsen, K. and Eide, I (1996). Arch. Toxicol, **70**, 397-404.
14. Csanady, Gy. A., Denk, B., Putz, C., Kreuzer, P. E., Kessler, W., Baur, C., Gargas, M. L. and Filser, J. G (2000). Toxicology and Applied Pharmacology, **165**, 1-26.
15. Lilly, P. D., Anderson, M. E., Ross, T. M. and Pegram, R. A (1997). Toxicology, **124**, 141-152.
16. Balaz, S and Lukacova, V (1999). Quant. Struct.-Act. Relat., **18**, 361-368.
17. Zhang, H (2004). J. Pharmaceutical Sciences, **93**, 1595-1604.

Chapter

14

Air to fat and blood to fat distribution for VOCs (human and rats)

In this chapter the Abraham solvation equation shall be used to construct models that include air to fat and blood to fat distribution for VOCs.

Air to fat distribution for VOCs

The distribution of VOCs from air to fat is of importance in pharmaceuticals, and in environmental toxicology. There have been a number of citations for air to fat distribution data for VOCs.[1-25] The purpose of this work is to gather a set of existing data from the literature of previous authors and to attempt to construct a new equation to predict these distribution processes. The correlative equations of Abraham shall be constructed based on datasets that represent average human and rat data combined. The complete list of the VOC log K data for these literature averages is shown in table 14.0 (see appendix). The reported literature models for the correlation of air to fat distribution for VOCs were made by Meulenberg & Vijverberg,[1] Gargas et al [3] and Weathersby and Abraham.[4]

Meulenberg & Vijverberg have obtained a reasonable amount of data in their training set to produce two correlative equations for air to fat distribution in rats (equation 14.0) and humans (equation 14.1). They found that the partition coefficients of VOCs in human (or rat) fat are well described by a linear combination of $K_{air:oil}$ and $K_{air:saline}$ with specific regression coefficients for the biological sample i.e. fat. The partition data (K) for air to fat distribution are averaged values from literature sources and have been used for the bilinear regression (note K is used in

the equations, and not log K). This method is very similar to the method employed by Gargas et al.[3] Equations (14.0 and 14.1) show how $K_{air:fat}$ is described as a bilinear function of $K_{air:olive\ oil}$ and $K_{air:saline}$. The coefficients α_{oil} and α_{saline} of both equations represents the tissue-specific contributions of the lipophilic and hydrophilic interactions to the solubility of the VOCs in fat. The c constant has no specific physical meaning but was required by Meulenberg & Vijverberg to compensate for systematic errors in fat, oil, or saline partition coefficients.[1] Note that the standard deviation and test sets are absent in this model.

Human equation Equation 14.0

$K_{(air\ to\ fat)} = 0.447K_{air:olive\ oil} + 0.075K_{air:saline} + 6.59$

$r^2 = 0.92$ SD = n/a n = 41

Rat equation Equation 14.1

$K_{(air\ to\ fat)} = 0.594K_{air:olive\ oil} + 0.085K_{air:saline} + 9.40$

$r^2 = 0.86$ SD = n/a n = 76

Gargas et al, obtained 55 air to oil partition (log K) data; this was used to model VOC air to fat distribution for rats. For this data set they used the air to oil log K data in equations to estimate rat air to fat distribution for VOCs (see below for the equation). Unfortunately the model doesn't have a test set, thus the predictive capability cannot be assessed.[3]

$\text{Log } K_{(air\ to\ fat)} = 0.920(\pm0.030)\log(K_{air:oil}) + 0.136(\pm0.083)$ Equation 14.2

$r^2 = 0.946$ RMSE = 0.187 n= 55

Gargas also modelled VOC air to rat fat distribution in terms of contribution from air to oil and air to saline values using linear regression techniques as shown below. With this combination there was very little improvement on the fits.[3]

$\text{Log } K_{(air\ to\ fat)} = 0.927(\pm0.031)\log(K_{air:oil}) - 0.032(\pm0.026)\log(K_{air:sal}) + 0.120(\pm0.083)$ Equation 14.3

$r^2 = 0.947$ RMSE = 0.186 n = 55

Abraham and Weathersby[4] compiled a dataset of 36 VOCs to model air to fat distribution in humans. Most of their data was collected from previous literature reviews. They used the Abraham solvation equation to correlate the log K data (310K) for air to fat with the five Abraham descriptors (**E**, **S**, **A**, **B** and **L**). The Abraham equation used for this model is shown in equation 14.4.[4]

$$\text{Log K}_{(\text{air to human fat})} = -0.294 - 0.172\mathbf{E} + 0.729\mathbf{S} + 1.747\mathbf{A} + 0.219\mathbf{B} + 0.895\mathbf{L} \qquad \text{Equation 14.4}$$

$r^2 = 0.988$ S.D = 0.118 n = 36

It appears that none of the three models have any test sets that include statistical information that can be used to test predictions. It was very unfortunate that Meulenberg & Vijverberg did not calculate the fits (i.e. standard deviation) for the training set equation they constructed.

Some improvement to these models can be made by adding new VOC data to see if the new equation for the full set has better fits than the published models. Table 14.0 (see appendix) contains the up to date literature data for air to fat distribution for VOCs.[1-25]

Discussion on air to fat distribution for VOCs (humans and rats combined)

The Abraham solvation equation has successfully been employed to correlate VOC air to fat distribution as shown in table 14.1. This model has now got 129 compounds in the dataset, and the fits for this new equation appear to be quite good (S.D = 0.194 and r^2 = 0.958), considering now there is a diverse range of VOCs. This work is the most up to date model for VOC air to fat distribution; it appears that this is the most general model yet reported for both species (humans and rats). There were no outliers observed in this full set.

EQUATIONS	N	r^2	SD	F	D
FULL SET **Log K** (air to fat) = -0.052 + 0.051**E** + 0.728**S** + 1.783**A** + 0.332**B** + 0.743**L**	129	0.958	0.194	562.789	5
TRAINING SET **Log K** (air to fat) = -0.138 + 0.067**E** + 0.770**S** + 1.965**A** + 0.356**B** + 0.756**L** **TEST SET** **N** = 64 **RMSE** = 0.177 **SD** (n-1) = 0.178 **AAE** = 0.142 **AE** = -0.036	65	0.962	0.217	298.023	5

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = root mean square error

Table 14.1. Summary of the equations for VOC air to fat distributions for rats and humans.

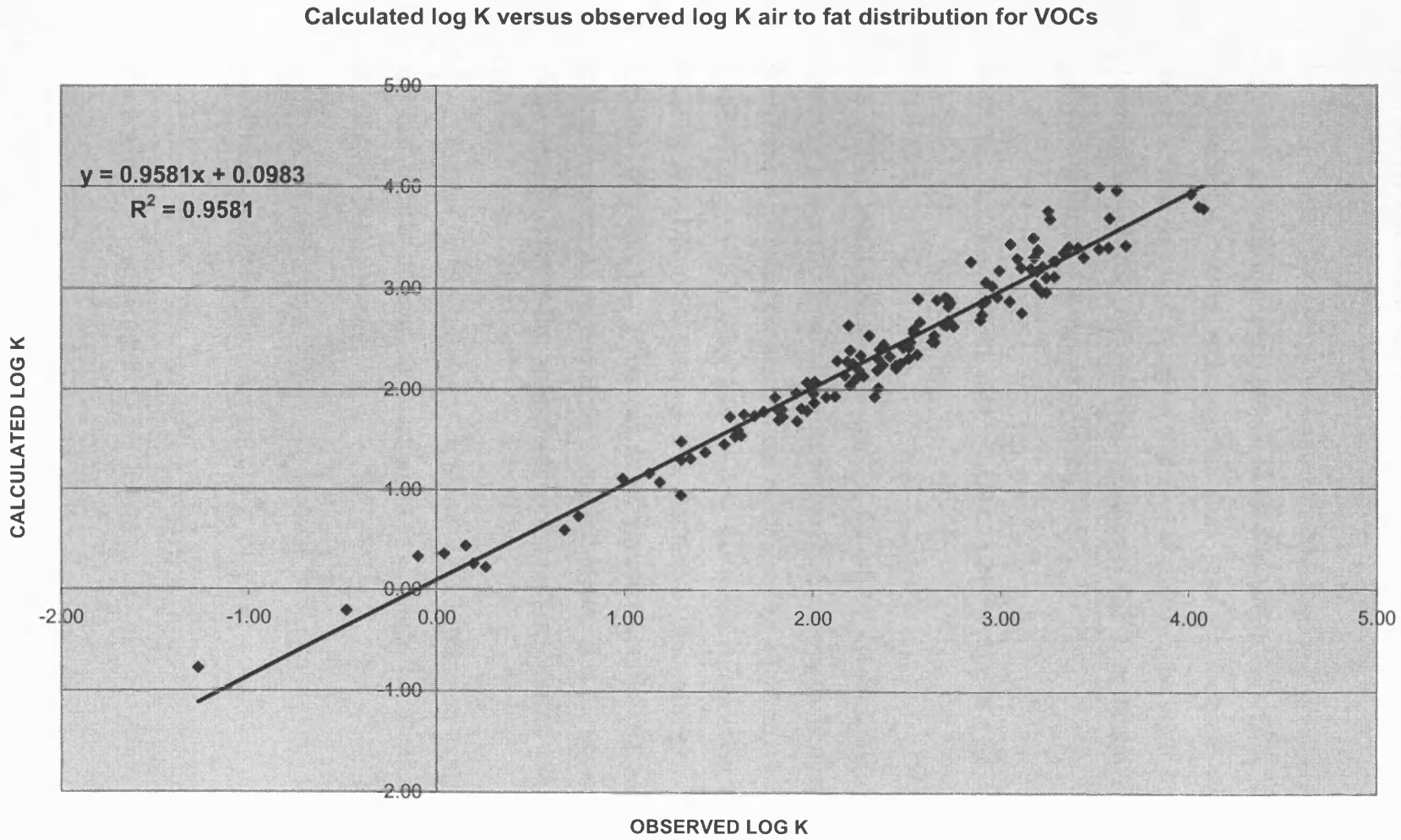
The inter-correlations of the descriptors (**E**, **S**, **A**, **B** and **L**) used for VOC air to fat distribution are shown in table 14.2. The largest r^2 value for E versus S is 0.25. It appears that there is no strong inter-correlation between the set of descriptors used in this current model.

r	E	S	A	B	L
E	1.00				
S	0.50	1.00			
A	-0.06	0.11	1.00		
B	-0.07	0.39	0.34	1.00	
L	0.48	0.16	-0.18	0.00	1.00

Table 14.2. The inter-correlation of descriptors for 129 solutes using descriptors **E**, **S**, **A**, **B** and **L** for air to fat distribution model.

The Abraham solvation equation has successfully been employed to correlate air to fat distribution for VOCs. To test the predictive ability of the model it was necessary to split the full data into training and test set using a random data set generator (Kennard-Stone). A new model was developed using only those data in the training set, which was then used to predict values for the test set. The regression yielded similar training equations and statistics to the first equation in table 14.1. The test set of 64 compounds leads to an average absolute error of 0.142 for the two sets of data (observed and predicted log K) and a root mean square error of 0.177, which indicates that the predictive capability of the training equation is very good. The average error of -0.036 indicates that the equation of the model is not biased, as it is very close to zero. The statistics for this model is the best out of all the air to tissue distribution models for VOCs. This is indicated by the large spread of data shown from the plot (page 194) giving rise to strong correlative fits.

The **e**-coefficient is not significant (P = 0.56). The positive **l**-coefficient in the air to fat equation (14.1) indicate that increasing molecular size and London dispersion interaction can push VOCS out of air and into the fat tissue phase. The coefficients of the 'polar' descriptors **S**, **A**, and **B** are all positive, indicating that polarity and hydrogen-bond donor and acceptor ability is stronger in the fat tissue phase than in gas phase. The solutes are pushed out of the air and into the fat phase. Finally the Abraham equation can successfully be used to correlate and predict air to fat distribution for humans and rats (see graph on page 194).


Calculated log K versus observed log K air to fat distribution for VOCs
y = 0.9581x + 0.0983
R2 = 0.9581
CALCULATED LOG K
OBSERVED LOG K
5.00
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00
5.00

Blood to fat distribution for VOCs

The solvation equation obtained by Abraham and Weathersby was shown to be useful in order to correlate blood to fat distribution for VOCs.[4] Their 35 blood to fat distribution dataset was compiled from previous literature reviews for human subjects. The solvation equation was used to correlate the log P data for blood to fat distribution with the five Abraham descriptors (**E**, **S**, **A**, **B** and **V**). The Abraham equation obtained is shown as equation 14.5.[4]

$$\text{Log P}_{(\text{blood to fat})} = 0.168 + 0.198\mathbf{E} + 0.130\mathbf{S} - 1.211\mathbf{A} - 3.267\mathbf{B} + 2.275\mathbf{V} \qquad \text{Equation 14.5}$$

$r^2 = 0.967$ S.D = 0.188 n = 35

Balaz and Lukacova compiled a list of 36 experimental data from several literature sources.[26] Their objective was to predict blood to fat distribution based on: membrane accumulation and protein binding. The equation 14.6 for blood to fat distribution of VOCs takes into account the tissue composition of the fat (lipids, proteins and water) in humans and function of lipophilicity. The independent variable was the P_{oct} (1-octanol/water partition coefficient) and the adjustable parameters were A_i, A_b and B for tissue and blood. β is the adjustable constant for series compounds binding to proteins in the bi-layer system) and I is an indicator variable.

The fits for the correlation of blood to fat distribution model were very good and had similar statistics to equation 14.5, although the Abraham & Weathersby model was slightly better.

$$\text{Log P}_{(\text{blood to fat})} = \beta \text{ x } \log P_{oct} + \log\left(1+\sum_{i=1}^{2} Ai' \text{ x } I_i\right) - \log(P^{\beta} \text{ x } A_b + 1) + B' \qquad \text{Equation 14.6}$$

$r^2 = 0.960$ SD = 0.20 n = 36

Zhang developed a non linear equation for blood to fat distribution of neutral compounds.[27] He compiled a list of 36 different volatile organic compounds obtained from various literatures sources. The first two equations in 14.7 takes into

account the biological interaction of the VOCs in the fat tissue phase, with regards to lipids and water. Note here, that protein contribution in this equation (14.7) was not taken into account for fat tissue, as it found to be the lowest in weight out of all the tissues used by Zhang. The product of these two equations is used in the third equation (14.7), to make predictions for blood to fat distribution for VOCs. The descriptors used in equation 14.7 are: α is the molecular polarizability (used to characterise the steric bulk effeck), $\sum Q^+$ is the sum of all positive partial atomic charges for all atoms in the molecule, Ca^{max} is the maximum H-bond acceptor descriptor in a molecule and $\sum Cd$ is the sum of the H-bond factor values for all the donor atoms in the molecule. His non-linear regression equations, was compared to the composition of the tissue (i.e. contents of lipids and water in fat and other tissues). Zhang concludes that his equations for blood to fat distribution are related to the tissue composition.

$$\text{Log } P_{(lipid)} = 0.123\alpha + 0.868\sum Q^+ - 0.849Ca^{max} + 0.621\sum Cd + 0.419 \qquad \text{Equation 14.7}$$

$$\text{Log } P_{(water)} = -0.318\alpha$$

$$\text{Log } P_{(blood\ to\ fat)} = \log(10^{\log P\ (lipid)} + 10^{\log P(water)} + 1)$$

$r^2 = 0.972$ SD = 0.178 n = 36

The aim of the current work was to construct an equation as part of a general method for prediction of equilibrium blood-fat distribution based on structure, using the Abraham descriptors (**E**, **S**, **A**, **B** and **V**). It would be interesting to see if the new literature for VOC air to fat distribution (humans and rats combined), leads to better statistics than the previous models. Table 14.3 (see appendix) lists the observed log P data for blood to fat distribution of VOCs.

Discussion on blood to fat distribution for VOCs (humans and rats combined)

The solvation equation of Abraham can be used to correlate VOC blood to fat distribution (see table 14.4). This model has 126 compounds in total in the dataset and the fits for this new equation appear to be reasonably good (S.D = 0.304 and r^2 = 0.847). This work is the most up to date model for VOC blood to fat distribution; it appears this is the most general model with the largest dataset yet reported for both species (humans and rats).

There were three outliers observed from this training set: nonane, decane and 2-nitropropane. The residual for these three compounds are approximately 0.82-1.12 log units and may result from experimental error.

EQUATIONS	N	r^2	SD	F	D
FULL SET **Log P** (blood to fat) = 0.473 + 0.017**E** - 0.003**S** - 1.577**A** - 2.246**B** + 1.560**V**	126	0.847	0.304	132.700	5
TRAINING SET **Log P** (blood to fat) = 0.419 - 0.006**E** - 0.016**S** - 1.245**A** - 2.240**B** + 1.587**V** **TEST SET** **N** = 63 **RMSE** = 0.303 **SD** (n-1) = 0.305 **AAE** = 0.249 **AE** = -0.047	63	0.835	0.311	57.739	5

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = root mean square error

Table 14.4. Summary of the equations for VOC blood to fat distributions

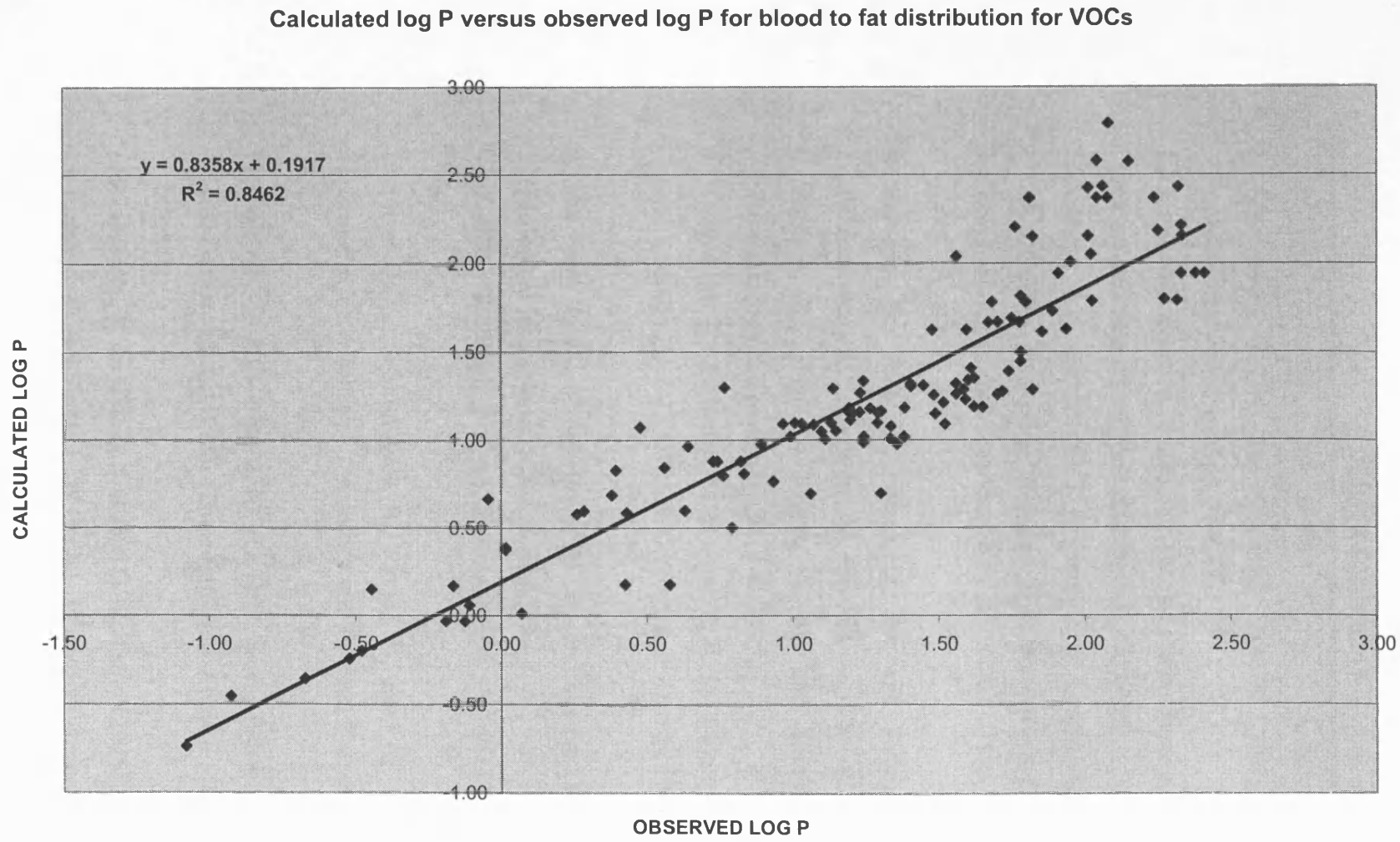
The inter-correlations of descriptors used for VOC blood to fat distribution are shown in table 14.5. The largest r^2 value for E versus S is 0.26. It appears that there is no strong inter-correlation between the set of descriptors used in this model.

r	E	S	A	B	V
E	1.00				
S	0.51	1.00			
A	-0.07	0.11	1.00		
B	-0.08	0.37	0.34	1.00	
V	0.12	-0.17	-0.24	0.00	1.00

Table 14.5. The inter-correlation of descriptors for 126 solutes using descriptors **E**, **S**, **A**, **B** and **V**

The Abraham solvation equation has been employed to correlate blood to fat distribution for VOCs. To test the predictiveness of the model it was necessary to split the full data into training and test set using a random data set generator. A new model was developed using only those data in the training set, which was then used to predict values for the test set. The regression yielded similar training equation and statistics to the first equation in table 14.4. The test of 63 compounds shows the average absolute error to be 0.249 for the two sets of data (observed and predicted log K). The root mean square error of 0.303 also indicates that there is reasonable predictive capability. The average error of -0.047 indicates that the equation of the model is not biased, as it is very close to zero. The blood to fat model is the best statistical model reported for blood to tissue distribution of VOCs in the thesis.

The **e**-coefficient (P = 0.90) and **s**-coefficient (P = 0.98) is not significant. The positive **v**-coefficient in the blood to fat equation (14.4) indicate that increasing molecular size and London dispersion interaction can push VOCS out of blood and into the fat tissue phase. The coefficients of the descriptors **A** and **B** are all negative, indicating that the hydrogen-bond donor and acceptor ability is stronger in the blood phase than in fat tissue phase. The solutes are pushed out of the fat and into the blood phase. The Abraham equation can therefore successfully be used to correlate and predict blood to fat distribution for humans and rats (see graph on page 198).


Calculated log P versus observed log P for blood to fat distribution for VOCs
y = 0.8358x + 0.1917
R2 = 0.8462
CALCULATED LOG P
OBSERVED LOG P
3.00
2.50
2.00
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-1.00
-0.50
0.00
0.50
1.00
1.50
2.00
2.50
3.00

References

1. Meulenberg, C. J. W. and Vijverberg, H. P. M (2000). Toxicology and Applied Pharmacology, **165**, 206-216.
2. Fiserova-Bergerova, V. and Diaz, M. L (1986). International Archives of Occupational and Environmental Health, **58**, 75-87.
3. Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H. and Anderson, M. E (1989). Toxicology and Applied Pharmacology, **98**, 87-99.
4. Abraham, M. H. and Weathersby, P. K (1994). J. Pharm. Sci, **83**, 1450-1455.
5. Zahlsen, K., Eide, I., Nilsen, A. M. and Nisen, O. G (1993). Pharmacology & Toxicology, **73**, 163-168.
6. Zahlsen, K., Eide, I., Nilsen, A. M. and Nisen, O. G (1992). Pharmacology & Toxicology, **71**, 144-149.
7. Gargas, M. L., Anderson, M. E., Teo, S. K. O., Batra, R., Fennell, T. R. and Kedderis, G. L (1995). Toxicology and Applied Pharmacology, **134**, 185-194.
8. Sweeney, L. M., Himmelstein, M. W. and Gargas, M. L. Chemical-Biological Interactions, **135-136**, 303-322.
9. Kaneko, T., Wang, P-Y. and Sato, A (2000). J. Occup. Health, **42**, 86-87.
10. Filser, J. G., Schmidbauer, R., Rampf, F., Baur, C. M., Putz, C. and Csanady, G. A (2000). Toxicology and Applied Pharmacology, **169**, 40-51.
11. Kedderis, G. L., Carfagna, M. A., Held, S. D., Batra, R., Murphy, J. E. and Gargas, M. L (1993). Toxicology and Applied Pharmacology, **123**, 274-282.
12. Nussbaum, E. and Hursh, J. B (1957). Science, **125**, 552-553.
13. Knaak, J. B. and Smith, L. W (1998). Inhalation Toxicology, **10**, 65-85.
14. Medinsky, M. A., Bechtold, W. E., Birnbaum, L. S., Chico, D. M., Gerlach, R. F. and Henderson, R. F (1988). Fundamental and Applied Toxicology, **11**, 250-260.

15. Loizou, G. D. and Anders, M.W (1993). Drug Metabolism and Disposition, **21**, 634-639.
16. Loizou, G. D. Urban, G. Dekant, W. and Anders, M.W (1994). Drug Metabolism and Disposition, **22**, 511-517.
17. Filser, J. G., Csanady, GY. A., Hartmann, M., Denk, B., Kauffmann, A., Kessler, W., Kreuzer, P. E., Putz, C., Shen,J. H. and Stei, P (1996). Toxicology, **113**, 278-287.
18. Zahlsen, K. and Eide, I (1996). Arch. Toxicol, **70**, 397-404.
19. Simmons, J. E., Boyes, W. K., Bushnell, P. J., Raymer, J. H., Limsakun, T., McDonald, A., Sey, Y. M. and Evans, M. V (2002). Toxicological Sciences, **69**, 3-15.
20. Krishnan, K., Gargas, M. L., Fennell, T. R. and Anderson, M. E (1992). Toxicol. Indust. Health, **8**, 121-140.
21. Csanady, Gy. A., Denk, B., Putz, C., Kreuzer, P. E., Kessler, W., Baur, C., Gargas, M. L. and Filser, J. G (2000). Toxicology and Applied Pharmacology, **165**, 1-26.
22. Reitz, R. H., Gargas, M. L., Anderson, M. E., Provan, W. M. and Green, T. L (1996). Toxicology and Applied Pharmacology, **137**, 253-267.
23. Lilly, P. D., Anderson, M. E., Ross, T. M. and Pegram, R. A (1997). Toxicology, **124**, 141-152.
24. Csanady, Gy. A., Kreuzer, P. E., Baur, C. and Filser, J. G (1996). Toxicology, **113**, 300-305
25. Borghoff, S. J., Murphy, J. E. and Medinsky, M. A (1996). Fundamental and Applied Toxicology, **30**, 264-275.
26. Balaz, S and Lukacova, V (1999). Quant. Struct.-Act. Relat., **18**, 361-368.
27. Zhang, H (2004). J. Pharmaceutical Sciences, **93**, 1595-1604.

Chapter

15

Air to liver and blood to liver distribution for VOCs

In this chapter, the Abraham solvation equation shall be used to construct models that include air to liver and blood to liver distribution for VOCs.

Air to liver distribution for VOCs

The partition of VOCs from air to liver is of importance in medical and environmental sciences. There have been a number of citations for air to liver distribution data for VOCs.[1-24] The purpose of this work is to gather a set of existing data from the literature of previous authors and to attempt to construct a new equation to predict these distribution processes. The correlative equations of Abraham shall be constructed based on datasets that represent average human and rat data combined. The complete list of the VOC log K data for these literature averages is shown in table 15.0 (see appendix). The reported literature models for the correlation of air to liver distribution for VOCs were made by Meulenberg and Vijverberg,[1] Gargas et al. [3] and Weathersby and Abraham.[4]

Meulenberg and Vijverberg have obtained a small number of literature data in his training set to produce two correlative equations for air to liver distribution in rats and humans.[1] They found that the partition coefficients of VOCs in human (or rat) liver are well described by a linear combination of $K_{air:oil}$ and $K_{air:saline}$ with a specific regression coefficient for the biological sample i.e. liver. The partition (K) data for air to liver distribution were averaged values from literature sources and were used

for the bilinear regression. The two equations 15.0 and 15.1 shows how $K_{air:liver}$ is described as a bilinear function of $K_{air:olive\ oil}$ and $K_{air:saline}$. The coefficients α_{oil} and α_{saline} of both equations represents the tissue-specific contributions of the lipophilic and hydrophilic interactions to the solubility of the VOCs in liver. The c constant has no specific physical meaning but it was required by Meulenberg and Vijverberg to compensate for systematic errors in liver, oil, or saline partition coefficients. It appears that neither of the two models have any test sets that includes statistical information that can be used to verify their predictions. It was very unfortunate that Meulenberg and Vijverberg did not calculate the fits (i.e. standard deviation) for the obtained correlations.[1]

Human equation Equation 15.0

$K_{(air\ to\ liver)} = 0.028K_{air:olive\ oil} + 0.79$

$r^2 = 0.88$ SD = n/a n = 28

Note K air to saline was not determined for the human air to liver equation. K is not log of K for air to liver partition.

Rat equation Equation 15.1

$K_{(air\ to\ liver)} = 0.026K_{air:olive\ oil} + 0.878K_{air:saline} + 2.36$

$r^2 = 0.92$ SD = n/a n = 77

Gargas et al.[3] also modelled VOC air to rat liver distribution in terms of contribution from air to oil and air to saline values using linear regression techniques as shown below. With this combination reasonable fits were obtained.[3] Unfortunately the model does not have a test set.

$$\text{Log } K_{(rat\ air\ to\ liver)} = 0.730(\pm 0.036)\log(K_{air:oil}) + 0.128(\pm 0.030)\log(K_{air:sal}) - 0.550(\pm 0.097) \quad \text{Equation 15.2}$$

$r^2 = 0.903$ RMSE = 0.217 n = 55

Weathersby and Abraham compiled a small dataset of 29 VOCs in order to model air to liver distribution for humans. Most of their data was compiled from previous literature reviews, and they used the Abraham solvation equation to correlate with the log K data (310K) for air to liver distribution with the five

Abraham descriptors (**E**, **S**, **A**, **B** and **L**). The Abraham equation used for this model is shown in equation 15.3.[4]

$$\text{Log K}_{(\text{air to human liver})} = -1.031 + 0.059\mathbf{E} + 0.774\mathbf{S} + 0.593\mathbf{A} + 1.049\mathbf{B} + 0.654\mathbf{L} \qquad \text{Equation 15.3}$$

$r^2 = 0.981$ S.D = 0.101 n = 29

None of the three above models have any test sets that include statistical information that can be used to test predictions. It was very unfortunate that Meulenberg and Vijverberg did not calculate the fits (i.e. standard deviation) for the training set equation they constructed. This also makes it very difficult to compare the training equations for the three authors based on the fits.

Some improvement to these models can possibly be made by adding new VOC data, to see if the new equation has better fits than the published models mentioned. Table 15.0 (see appendix) contains up to date literature data for air to liver distribution for VOCs.[1-25]

Discussion on air to liver distribution for VOCs (humans and rats combined)

The Abraham solvation equation has been employed to correlate VOC air to liver distribution as shown in table 15.1. This model has now 124 compounds in the data set, and the fits for this new equation appear to be good (S.D = 0.285 and r^2 = 0.909), considering that there is a diverse range of VOCs. This work is the most up to date model for VOC air to liver distribution; it appears that this is the most general model yet reported for both species (humans and rats). There were four outliers observed (0.75 to 1.04 log units difference in observed and calculated values) in this full set: methanol, ethanol, allyl chloride and 1, 2, 4-trimethylbenzene. There is no direct explanation to why these compounds are outliers, but it may be due experimental errors.

EQUATIONS	N	r^2	SD	F	D
FULL SET **Log K** (air to liver) = -0.919 + 0.083**E** + 0.768**S** + 2.796**A** + 2.090**B** + 0.560**L**	124	0.909	0.285	236.684	5
TRAINING SET **Log K** (air to liver) = -0.846 + 0.249**E** + 0.835**S** + 2.594**A** + 2.196**B** + 0.504**L** **TEST SET** **N** = 62 **RMSE** = 0.271 **SD** (n-1) = 0.273 **AAE** = 0.223 **AE** = 0.000	62	0.905	0.315	106.502	5

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = root mean square error

Table 15.1. Summary of the equations for VOC air to liver distributions

The inter-correlations of the descriptors (**E**, **S**, **A**, **B** and **L**) used for VOC air to liver distribution are shown in table 15.2. The largest r^2 value for E versus S descriptor is 0.23. It appears that there is no strong inter-correlation between the set of descriptors used in this current model.

r	E	S	A	B	L
E	1.00				
S	0.48	1.00			
A	-0.06	0.12	1.00		
B	-0.10	0.39	0.30	1.00	
L	0.46	0.13	-0.14	0.01	1.00

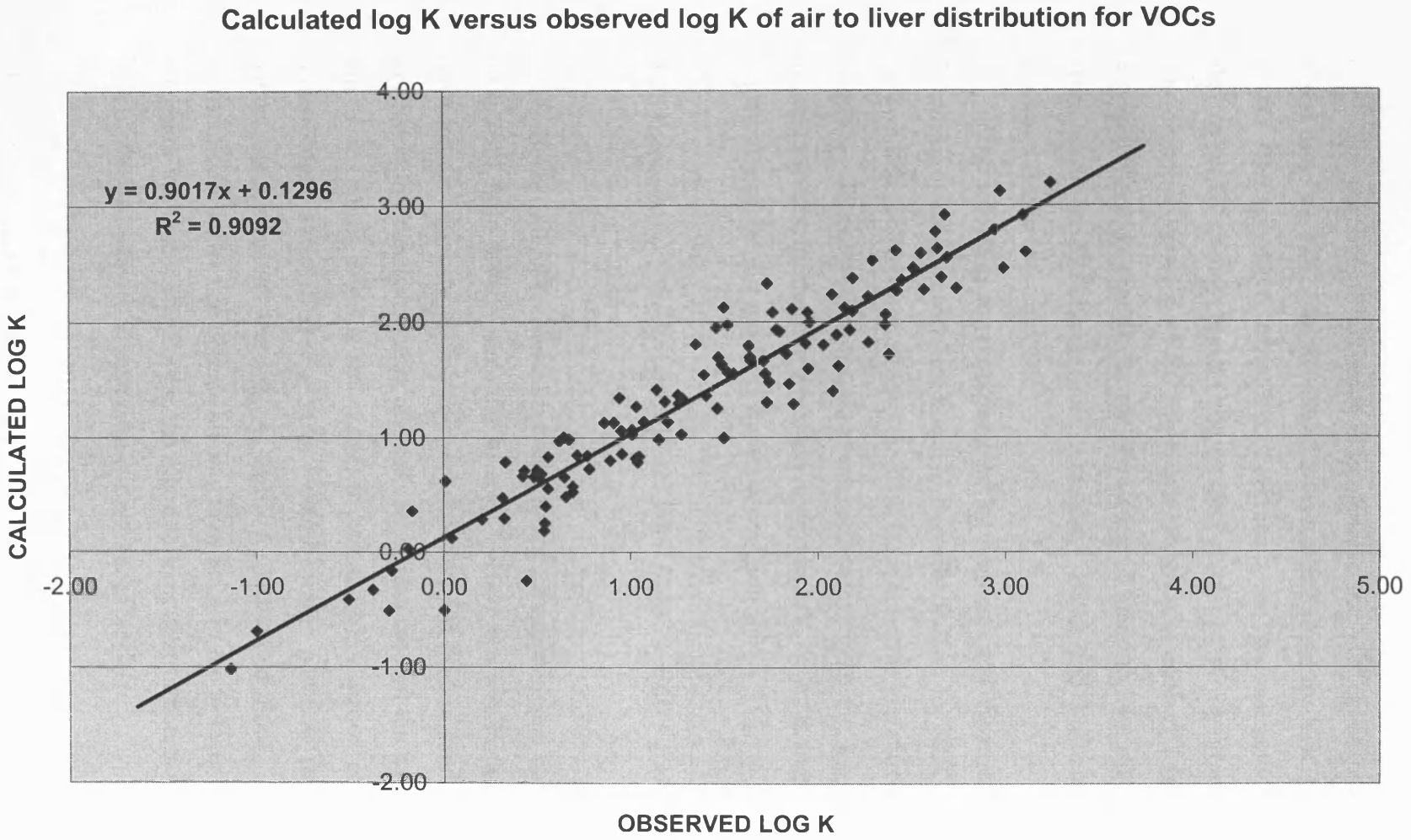
Table 15.2. The inter-correlation of descriptors for 124 solutes using descriptors **E**, **S**, **A**, **B** and **L** for air to liver distribution model.

The Abraham solvation equation has effectively been employed to correlate air to liver distribution for VOCs. To test the predictive ability of the model it was

necessary to split the full data into training and test set using a random data set generator (Kennard-Stone). A new model was developed using only those data in the training set, which was then used to predict values for the test set. The regression yielded similar training equations and statistics to the first equation in table 15.1, although the full set has a better fit than the training equation. Note also the **e** coefficient in the full set is slightly different from that in the training equation. The test set of 62 compounds leads to an average absolute error of 0.223 for the two sets of data (observed and predicted log K) and a root mean square error of 0.271, which indicates that the predictive capability of the training equation is good. The average error of 0.0 indicates that the equation of the model is not biased, as it is zero.

The **e**-coefficient is not significant (P = 0.51). The positive **l**-coefficients in the air to liver equation (15.1) indicate that increasing molecular size and London dispersion interaction can push VOCS out of air and into the liver tissue phase. The coefficients of the 'polar' descriptors **S**, **A**, and **B** are all positive, indicating that polarity and hydrogen-bond donor and acceptor ability is stronger in the liver tissue phase than in gas phase. The solutes are pushed out of the air and into the liver tissue phase. Finally the Abraham equation can successfully be used to correlate and predict air to liver distribution for humans and rats (see graph on page 206).

The performance of this model is satisfactory, as this is indicated by the even spread of data shown from the plot. It is found that it is possible to construct an equation capable of predicting further values of log K_{liver} to around 0.3 log units. This equation is not much better in terms of predictions when compared to the air to fat equation shown in chapter 14. This may be due to the fact there are other processes like metabolism (cytochrome-p450 enzymes) and protein binding taking part within the liver.


Calculated log K versus observed log K of air to liver distribution for VOCs
y = 0.9017x + 0.1296
R2 = 0.9092
CALCULATED LOG K
OBSERVED LOG K
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00
5.00

Blood to liver distribution for VOCs

The purpose of this work is to combine the VOC liver data (log K_{tissue}) with blood data (log K_{blood}) to give blood to liver distribution for VOCs (log $P_{blood\ to\ liver}$) at 310K. The correlative equations of Abraham can be constructed based on datasets that represent average human and rat data combined for blood to liver distribution. The complete list of the VOC log P data is shown in table 15.3 (see appendix). The reported literature models for the correlation of blood to liver distribution for VOCs were made by Weathersby and Abraham,[4] Balaz and Lukacova,[25] Zhang[26] and Fiserova-Bergerova and Diaz.[2]

The solvation equation obtained by Abraham and Weathersby was shown to be useful in order to correlate blood to liver distribution for VOCs.[4] Their 28 blood to liver distribution dataset was compiled from previous literature reviews on human subjects. The solvation equation was used to correlate the log P data (310K) for blood to liver with the five Abraham descriptors (**E**, **S**, **A**, **B** and **V**). The Abraham equation used for this model is shown in equation 15.4.[4]

$$\text{Log P}_{(\text{blood to liver})} = -0.270 + 0.233\mathbf{E} - 0.375\mathbf{S} - 1.004\mathbf{A} - 1.118\mathbf{B} + 1.241\mathbf{V} \qquad \text{Equation 15.4}$$

$r^2 = 0.865$ S.D = 0.155 n = 28

Balaz and Lukacova compiled a list of 28 experimental data of non-ionisable chemicals from several literature sources.[25] Their aim was to predict blood to liver distribution based on: membrane accumulation and protein binding. The equation for blood to liver distribution of VOCs takes into account the tissue composition of the liver (lipids, proteins and water) in humans and function of lipophilicity. The independent variable is P_{oct} (the 1-octanol/water partition coefficient) and the adjustable parameters were A_i, A_b and B for tissue and blood. β is the adjustable constants for series compounds binding to proteins in the bi-layer system and I as the indicator variable. The fits for the blood to liver distribution model were good and this model had a better fit than previous published models.

$$\text{Log P}_{(\text{blood to liver})} = \beta \times \log P_{oct} + \log\left(1 + \sum_{i=1}^{2} Ai' \times I_i\right) - \log(P^{\beta} \times A_b + 1) + B' \qquad \text{Equation 15.5}$$

$r^2 = 0.931$ SD = 0.095 n = 28

Zhang developed a non linear equation for blood to liver distribution of neutral compounds.[26] He compiled a list of 29 different volatile organic compounds obtained from various literatures sources. The first two equations in 15.6, takes into account the biological interaction of the VOCs in the liver tissue phase, with regards to lipids and proteins. The product of the first two equations is used in the third equation (15.6) to make predictions for blood to liver distribution for VOCs. The descriptors used in equation 15.6 are, α is the molecular polarizability (used to characterise the steric bulk effeck), $\sum Ca$ is the sum of H-bond factor values for all acceptor substructures in the molecule and $\sum Q^+$ is the sum of all positive partial atomic charges for all atoms in the molecule. Zhang concludes that his equations for blood to liver distribution are related to the tissue composition.

$$\text{Log } P_{(\text{lipid})} = 0.130\alpha - 0.668\sum Ca \qquad \text{Equation 15.6}$$

$$\text{Log } P_{(\text{protein})} = 0.843\sum Q^+$$

$$\text{Log } P_{(\text{blood to liver})} = \log(10^{\text{logP (lipid)}} + 10^{\text{logP(protein)}} + 1) - 0.658$$

$r^2 = 0.872$ SD = 0.134 n = 29

Fiserova-Bergerova and Diaz used the head space chromatography method to determine experimentally air to liver distribution for several VOCs.[2] They also compiled a literature survey of experimental values based on human subjects. Air to liver and air to blood partition coefficients were used to calculate blood to liver partition coefficients for VOCs. From the statistical analysis of Fiserova-Bergerova Diaz, it was found that there is a good inter-correlation between blood to liver partition coefficients and blood to fat partition coefficients as shown from a plot. The intercept and the gradient of the straight line (linear) plots were included in the model to make a correlative equation for blood to liver distribution of VOCs (see equation 15.7). The standard deviation of this model is very large compared to other models mentioned earlier and this suggests that the equation is poor in terms of correlation in contrast to other models mentioned earlier.[2]

$$P_{(\text{blood to liver})} = 0.0425P_{\text{blood to fat}} + 0.5770 \qquad \text{Equation 15.7}$$

n = 27 r^2 = 0.919 SD = 0.997

P denotes to the appropriate partition of the VOCs

The aim of the current work was to construct an equation for prediction of equilibrium blood-liver distribution based on structure using the Abraham descriptors (**E, S, A, B** and **V**). Now it would be interesting to see if the new data for VOC blood to liver distribution (humans and rats combined) makes a difference in the fits to the previous blood to liver distribution models as mentioned. Table 15.3 (see appendix) lists the observed log P ($_{\text{blood to liver}}$) data for blood to liver distribution of VOCs. Note all of the work is for blood and not blood or plasma.

Discussion on blood to liver distribution for VOCs (humans and rats combined)

The solvation equation of Abraham can be used to correlate VOC blood to liver distribution (see table 15.4). This model has 120 compounds in total in the dataset. The fits for this new equation appear to be not as good (S.D = 0.211 and r^2 = 0.648), compared to other models constructed previously by other authors. Although the standard deviation is good (SD= 0.211) and even better than the model by Fiserova-Bergerova and Diaz (SD = 0.997), the overall square of the correlation is poorer (r^2 = 0.648). This work is the most up to date model for VOC blood to liver distribution; it appears this is the most general model with the largest dataset yet reported for both species (humans and rats).

There were four outliers observed from this training set: nonane, decane, t-butylcyclohexane and 1,1-difluoroethylene. The residual for these four compounds are approximately 0.60-0.90 log units and may result from experimental error.

EQUATIONS	N	r^2	SD	F	D
FULL SET **Log P** (blood to liver) = -0.202 - 0.048**E** - 0.295**S** - 0.213**A** - 0.359**B** + 0.876**V**	120	0.648	0.211	42.012	5

TRAINING SET **Log P** (blood to liver) = -0.213 - 0.002**E** - 0.253**S** - 0.236**A** - 0.522**B** + 0.879**V** **TEST SET** **N** = 60 **RMSE** = 0.233 **SD** (n-1) = 0.235 **AAE** = 0.184 **AE** = -0.018	60	0.693	0.193	24.389	5

N = number of compounds used the in training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = root mean square error

Table 15.4. Summary of the equations for VOC blood to liver distributions

The inter-correlations of descriptors used for VOC blood to liver distribution are shown in table 15.5. The largest r^2 value for E versus S descriptor is 0.21. It appears that there is no strong inter-correlation between the set of descriptors used in this model.

r	E	S	A	B	V
E	1.00				
S	0.46	1.00			
A	-0.07	0.11	1.00		
B	-0.13	0.37	0.30	1.00	
V	0.05	-0.20	-0.18	0.02	1.00

Table 15.5. The inter-correlation of descriptors for 120 solutes using descriptors **E**, **S**, **A**, **B** and **V**.

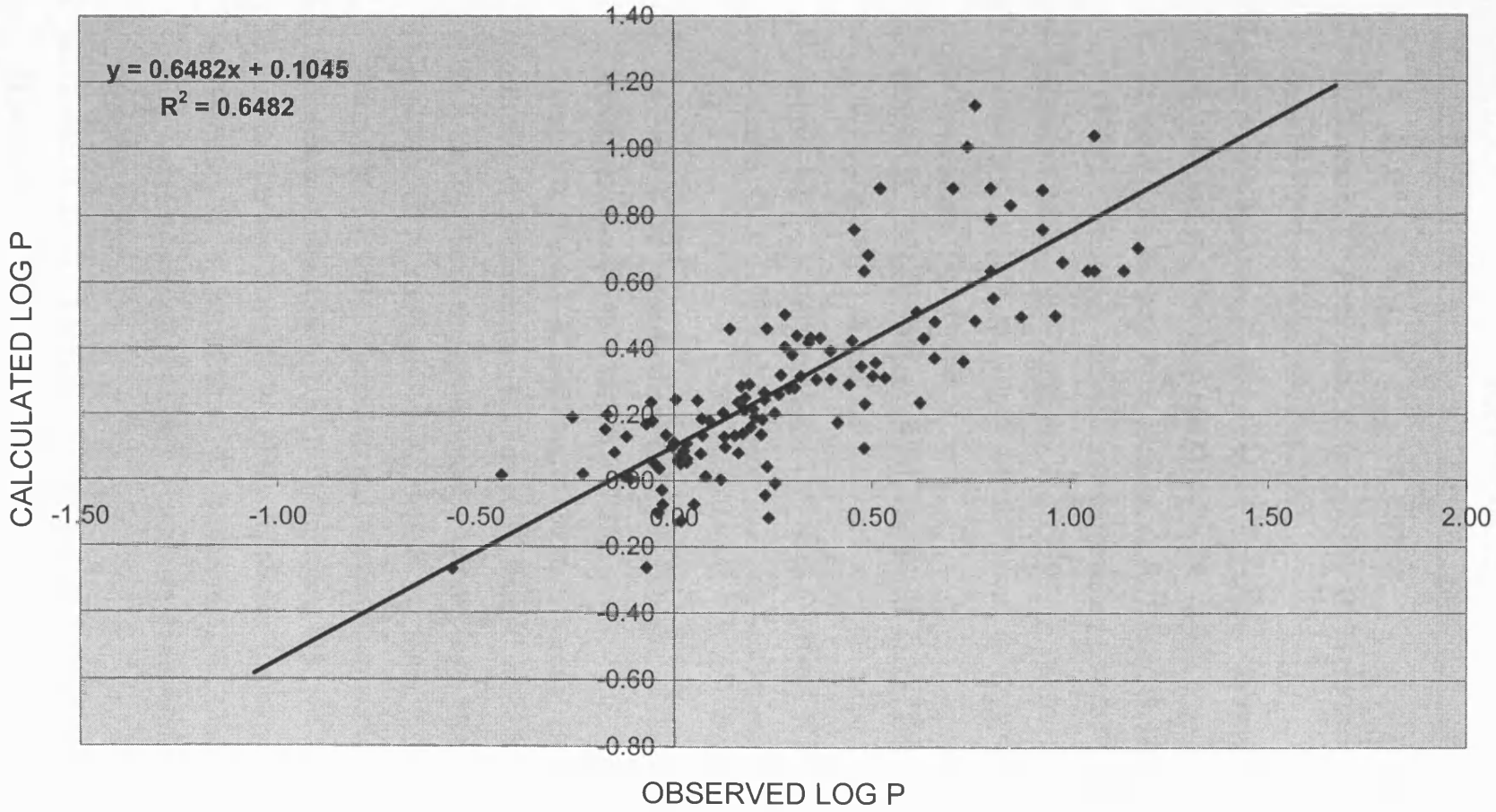
The Abraham solvation equation has been employed to correlate blood to liver distribution for VOCs. To test the predictiveness of the model it was necessary to split the full data into training and test set using a random data set generator. A new model was developed using only those data in the training set, which was then used to predict values for the test set. The regression yielded similar training equation, although the training set had a slightly better fit than the first full equation in table

15.4. The test of 60 compounds shows the average absolute error to be 0.184 for the two sets of data (observed and predicted log K). The root mean square error of 0.233 also indicates that there is reasonable predictive capability. The average error of -0.018 indicates that the equation of the model is not biased, as it is very close to zero.

The **e**-coefficient (P = 0.58) and **a**-coefficient (P = 0.33) is not significant. The positive **v**-coefficients in the blood to liver equation (15.4) indicate that increasing molecular size can push VOCS out of blood and into the liver tissue phase. The coefficients of the 'polar' descriptors **S** and **B** are all negative, indicating that polarity and hydrogen-bond acceptor ability is stronger in the blood phase than in liver tissue phase. The solutes are pushed out of the liver and into the blood phase. The Abraham equation can therefore successfully be used to correlate and predict blood to liver distribution for humans and rats (see graph on page 212).

It appears that this up to date model has a poorer fit than the current models already published by Weathersby and Abraham,[4] Balaz and Lukacova,[25] Zhang[26] and Fiserova-Bergerova and Diaz.[2] Although this model has a poorer fit, it represents a larger dataset of compounds for blood to liver distribution for VOCs. Previous published models have a limited number of compounds (27-28 compounds) in the data set and this may well not be a total representation of VOC blood to liver distribution. In addition, the smaller number of data points per descriptor, the better will the fit be in terms of the correlation coefficient and SD (but not in terms of F). The published models above have small data sets, and this is probably the reason for the good fits. The present model has a much larger and more varied data set and should represent a better coverage of chemical space.

Calculated log P versus observed log P of blood to liver distribution for VOCs



References

1. Meulenberg, C. J. W. and Vijverberg, H. P. M (2000). Toxicology and Applied Pharmacology, **165**, 206-216.
2. Fiserova-Bergerova, V. and Diaz, M. L (1986). International Archives of Occupational and Environmental Health, **58**, 75-87.
3. Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H. and Anderson, M. E (1989). Toxicology and Applied Pharmacology, **98**, 87-99.
4. Abraham, M. H. and Weathersby, P. K (1994). J. Pharm. Sci, **83**, 1450-1455.
5. Zahlsen, K., Eide, I., Nilsen, A. M. and Nisen, O. G (1993). Pharmacology & Toxicology, **73**, 163-168.
6. Zahlsen, K., Eide, I., Nilsen, A. M. and Nisen, O. G (1992). Pharmacology & Toxicology, **71**, 144-149.
7. Gargas, M. L., Anderson, M. E., Teo, S. K. O., Batra, R., Fennell, T. R. and Kedderis, G. L (1995). Toxicology and Applied Pharmacology, **134**, 185-194.
8. Sweeney, L. M., Himmelstein, M. W. and Gargas, M. L (2001). Chemical-Biological Interactions, **135-136**, 303-322.
9. Kaneko, T., Wang, P-Y. and Sato, A (2000). J. Occup. Health, **42**, 86-87.
10. Filser, J. G., Schmidbauer, R., Rampf, F., Baur, C. M., Putz, C. and Csanady, G. A (2000). Toxicology and Applied Pharmacology, **169**, 40-51.
11. Kedderis, G. L., Carfagna, M. A., Held, S. D., Batra, R., Murphy, J. E. and Gargas, M. L (1993). Toxicology and Applied Pharmacology, **123**, 274-282.
12. Nussbaum, E. and Hursh, J. B (1957). Science, **125**, 552-553.
13. Knaak, J. B. and Smith, L. W (1998). Inhalation Toxicology, **10**, 65-85.
14. Medinsky, M. A., Bechtold, W. E., Birnbaum, L. S., Chico, D. M., Gerlach, R. F. and Henderson, R. F (1988). Fundamental and Applied Toxicology, **11**, 250-260.
15. Loizou, G. D. and Anders, M.W (1993). Drug Metabolism and Disposition, **21**, 634-639.
16. Loizou, G. D. Urban, G. Dekant, W. and Anders, M.W (1994). Drug Metabolism and Disposition, **22**, 511-517.

17. Filser, J. G., Csanady, GY. A., Hartmann, M., Denk, B., Kauffmann, A., Kessler, W., Kreuzer, P. E., Putz, C., Shen, J. H. and Stei, P (1996). Toxicology, **113**, 278-287.
18. Zahlsen, K. and Eide, I (1996). Arch. Toxicol, **70**, 397-404.
19. Simmons, J. E., Boyes, W. K., Bushnell, P. J., Raymer, J. H., Limsakun, T., McDonald, A., Sey, Y. M. and Evans, M. V (2002). Toxicological Sciences, **69**, 3-15.
20. Csanady, Gy. A., Denk, B., Putz, C., Kreuzer, P.E., Kessler, W., Baur, C., Gargas, M. L. and Filser, J. G (2000). Toxicology and Applied Pharmacology, **165**, 1-26.
21. Reitz, R. H., Gargas, M. L., Anderson, M. E., Provan, W. M. and Green, T. L (1996). Toxicology and Applied Pharmacology, **137**, 253-267.
22. Lilly, P. D., Anderson, M. E., Ross, T. M. and Pegram, R. A (1997). Toxicology, **124**, 141-152.
23. Csanady, Gy. A., Kreuzer, P. E., Baur, C. and Filser, J. G (1996). Toxicology, **113**, 300-305.
24. Borghoff, S. J., Murphy, J. E. and Medinsky, M. A (1996). Fundamental and Applied Toxicology, **30**, 264-275.
25. Balaz, S. and Luckacova (1999). Quant. Struct.-Act. Relat., **18**, 361-368.
26. Zhang, H (2004). J. Pharmaceutical Sciences, **93**, 1595-1604.

Chapter

16

Air to muscle and blood to muscle distribution for VOCs

In this chapter the Abraham solvation equation shall be used to construct predictive equations that include air to muscle and blood to muscle distribution for VOCs.

Air to muscle distribution for VOCs

The distribution of VOCs from air to muscle is of importance in environmental and toxicological sciences. There have been a few citations for air to muscle distribution data for VOCs.[1-17] The purpose of this work is to gather a set of existing data from the literature of previous authors and to attempt to construct a new equation to predict these distribution processes. The Abraham equation used in the correlation shall be constructed based on datasets that represent average human and rat data combined. The complete list of the VOC log K data for these literature averages is shown in table 16.0 (see appendix). The reported literature models for the correlation of air to muscle distribution for VOCs are found from Meulenberg & Vijverberg,[1] Gargas et al [3] and Weathersby and Abraham.[4]

Meulenberg and Vijverberg have used a small amount of literature data in training sets to produce two correlative equations for air to muscle distribution in rats and humans.[1] They found that the partition coefficients of VOCs in human (or rat) kidney are well described by a linear combination of $K_{air:oil}$ and $K_{air:saline}$ with specific

regression coefficient for the biological sample i.e. muscle. The partition (K) data for air to muscle distribution are averaged values from literature sources and have been used for the bilinear regression. The two equations 16.0 and 16.1 shows how $K_{air:muscle}$ is described as a bilinear function of $K_{air:olive\ oil}$ and $K_{air:saline}$. The coefficients α_{oil} and α_{saline} of both equations represents the tissue-specific contributions of the lipophilic and hydrophilic interactions to the solubility of the VOCs in muscle. The c constant has no specific physical meaning but it was required by Meulenberg and Vijverberg to compensate for systematic errors in muscle, oil, or saline partition coefficients. Neither of the two models have any test sets that can be used to assess predictions. It is very unfortunate that Meulenberg and Vijverberg did not calculate the fits (i.e. standard deviation) for the obtained correlations.[1]

Human equation Equation 16.0

$$K_{(air\ to\ muscle)} = 0.014K_{air:olive\ oil} + 0.384K_{air:saline} + 0.94$$

$r^2 = 0.99$ SD = n/a n = 35

Note K is not the log of K for air to muscle distribution

Rat equation Equation 16.1

$$K_{(air\ to\ muscle)} = 0.010K_{air:olive\ oil} + 0.772K_{air:saline} + 0.29$$

$r^2 = 0.82$ SD = n/a n = 76

Gargas et al. also modelled VOC air to rat muscle distribution in terms of contribution from air to oil and air to saline values using linear regression techniques as shown below. With this combination they achieved reasonable fits.[3] Unfortunately the model again does not have a test set.[3]

$$\text{Log } K_{(air\ to\ muscle)} = 0.644(\pm0.038)\log(K_{air:oil}) + 0.180(\pm0.032)\log(K_{air:sal}) - 0.725(\pm0.104)$$ Equation 16.2

$r^2 = 0.879$ RMSE = 0.233 n = 55

Weathersby and Abraham compiled a dataset of 41 VOCs in order to correlate air to muscle distribution for humans. Their data is compiled from previous literature reviews.[4] They used the Abraham solvation equation to correlate the log K

data (310K) for air to muscle with the five Abraham descriptors (**E**, **S**, **A**, **B** and **L**). The Abraham equation used for this model is shown in equation 16.3.[4]

$$\text{Log } K_{(\text{air to human muscle})} = -1.140 + 0.544\mathbf{E} + 0.216\mathbf{S} + 3.471\mathbf{A} + 2.924\mathbf{B} + 0.578\mathbf{L} \qquad \text{Equation 16.3}$$

$r^2 = 0.966$ S.D = 0.262 n = 41

None of the three above models have any test sets that include statistical information that can be used to assess predictions. It is disappointing that Meulenberg and Vijverberg did not calculate the fits (i.e. standard deviation) for the training set equation they have constructed.

Some improvement to these models can be made by adding new VOC data to see if the new equation for the full set has better fits than the published models. Table 16.0 (see appendix) contains up to date literature data for air to muscle distribution for VOCs.[1-17]

Discussion on air to muscle distribution for VOCs (humans and rats combined)

The solvation equation of Abraham is used to correlate and predict VOC air to muscle distribution as shown in table 16.1. This model has now got 114 compounds in the dataset, and the fits for this new equation appear to be quite good (S.D = 0.267 and $r^2 = 0.944$), for a diverse range of VOCs. This work is the most up to date model for VOC air to muscle distribution (humans and rats). Although this model has the largest data set ever reported; the fits for the full set (table 16.1) appear to be not too far off compared to other models mentioned earlier. There were three outliers observed from this training set: tricyclo[5.2.1.0]-decane (JP10), methanol and 4-chlorobenzotrifluoride The residual for these three compounds are approximately 0.78-0.97 log units and may result from experimental error.

EQUATIONS	N	r^2	SD	F	D
FULL SET **Log K** (air to muscle) = -1.039 + 0.207**E** + 0.723**S** + 3.242**A** + 2.469**B** + 0.463**L**	114	0.944	0.267	362.966	5
TRAINING SET **Log K** (air to muscle) = -1.061 + 0.427**E** + 0.982**S** + 3.328**A** + 2.557**B** + 0.387**L** **TEST SET** **N** = 57 **RMSE** = 0.304 **SD** (n-1) = 0.306 **AAE** = 0.236 **AE** = -0.043	57	0.947	0.262	180.787	5

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = root mean square error

Table 16.1. Summary of the equations for VOC air to muscle distributions for humans and rats combined.

The inter-correlations of the descriptors (**E**, **S**, **A**, **B** and **L**) used for VOC air to muscle distribution are shown in table 16.2. The largest r^2 value for E versus L descriptor is 0.34. It appears that there is no strong inter-correlation between the set of descriptors used in this current model.

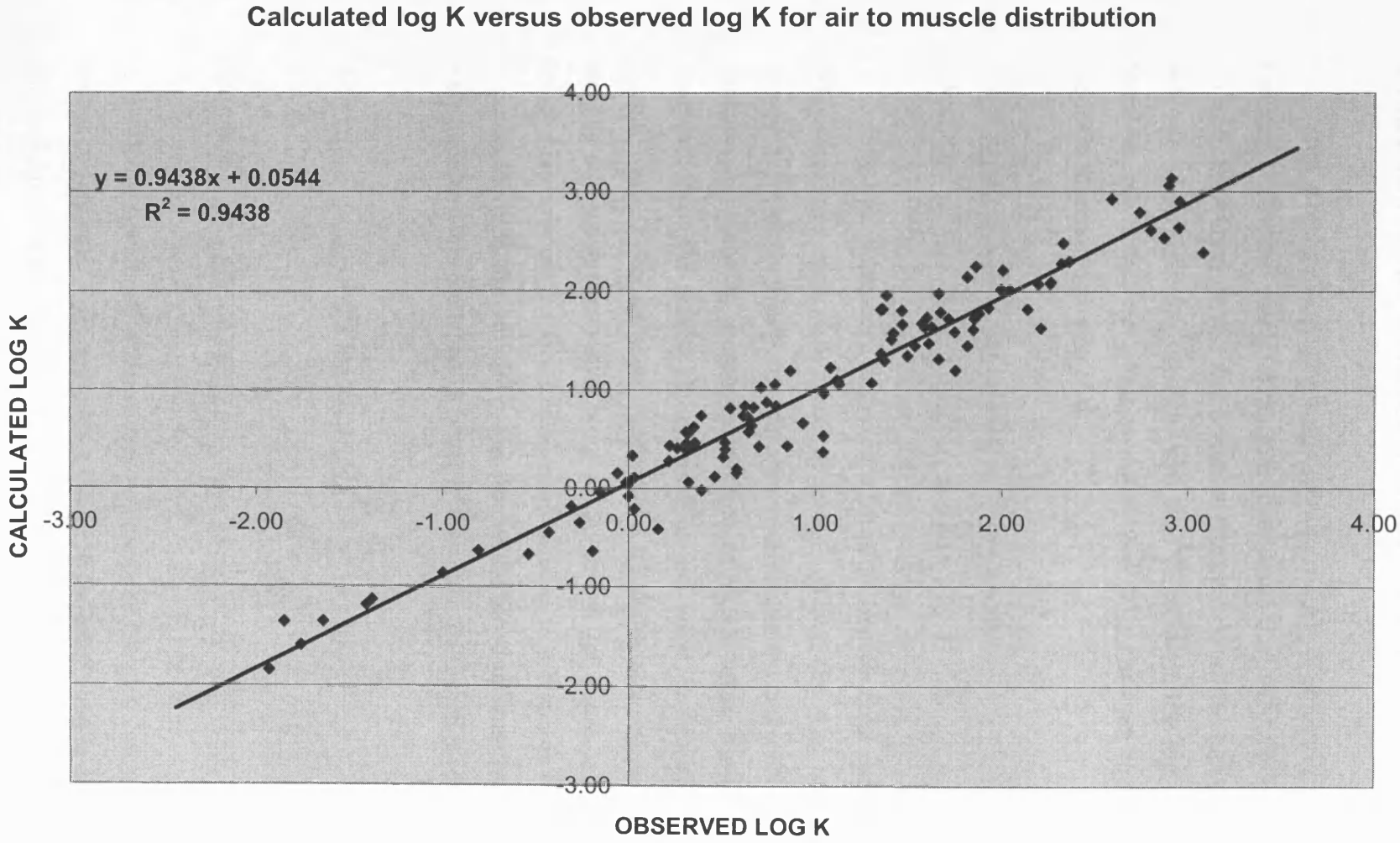
r	E	S	A	B	L
E	1.00				
S	0.57	1.00			
A	-0.03	0.10	1.00		
B	-0.05	0.35	0.31	1.00	
L	0.58	0.49	-0.01	0.24	1.00

Table 16.2. The inter-correlation of descriptors for 114 solutes using descriptors **E**, **S**, **A**, **B** and **L** for air to muscle distribution model.

The Abraham solvation equation has successfully been employed to correlate air to muscle distribution for VOCs provided the solute is within the given descriptor

range of the training set to make the correlative equations. To test the predictive ability of the model it was necessary to split the full data into training and test set using a random data set generator (Kennard-Stone). A new model was developed using only those data in the training set, which was then used to predict values for the test set. The regression yielded similar training equations and statistics to the first equation in table 16.1, except for the **e** coefficient, which appeared to be slightly different. The test set of 57 compounds leads to an average absolute error of 0.236 for the two sets of data (observed and predicted log K) and a root mean square error of 0.304, which indicates that the predictive capability of the training equation is good. The average error of -0.043 indicates that the equation of the model is not biased, as it is very close to zero. The statistics for this model is very good, as indicated by the even spread of data shown from the plot. It is found that it is possible to construct an equation capable of predicting further values of log K_{muscle} to around 0.3 log units. The fits for this equation are little better than that of the air to blood model mentioned earlier in chapter 10.

The **e**-coefficient is not significant (P = 0.11). The positive **l**-coefficients in the air to muscle equation (16.1) indicate that increasing molecular size and London dispersion interaction can push VOCS out of air and into the muscle tissue phase. The coefficients of the 'polar' descriptors **S**, **A**, and **B** are all positive, indicating that polarity and hydrogen-bond donor and acceptor ability is stronger in the muscle tissue phase than in gas phase. The solutes are pushed out of the air and into the muscle phase. Finally the Abraham equation can successfully be used to correlate and predict air to muscle distribution for humans and rats (see graph on page 220).


Calculated log K versus observed log K for air to muscle distribution
y = 0.9438x + 0.0544
R2 = 0.9438
CALCULATED LOG K
OBSERVED LOG K
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-3.00
-3.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00

Blood to muscle distribution for VOCs

The Abraham solvation equation obtained by Weathersby and Abraham was shown to correlate blood to muscle distribution for VOCs.[4] Their 40 blood to muscle distribution dataset was compiled from previous literature reviews for human subjects. The solvation equation was used to correlate the log P data (310K) for blood to muscle with the Abraham descriptors. The Abraham equation used for this model is shown in equation 16.4.[4]

$$\text{Log P}_{(\text{blood to muscle})} = -0.222 - 0.479\mathbf{S} - 0.517\mathbf{B} + 0.999\mathbf{V} \qquad \text{Equation 16.4}$$

$r^2 = 0.790$ S.D = 0.176 n = 40

Balaz and Lukacova compiled a list of 34 experimental data of non-ionisable chemicals from several literature sources.[18] Their aim was to predict blood to muscle distribution based on: membrane accumulation and protein binding. The equation for blood to muscle distribution of VOCs takes into account the tissue composition of the muscle (lipids, proteins and water) in humans and function of lipophilicity. The independent variable is P_{oct} (1-octanol/water partition coefficient) and the adjustable parameters were A_i, A_b and B for tissue and blood. β is the adjustable constants for series compounds binding to proteins in the bi-layer system and I_i as the indicator variable. The fits for the blood to muscle distribution model were good and this model had a better fit than other published models.

$$\text{Log P}_{(\text{blood to muscle})} = \beta \times \log P_{oct} + \log\left(1+\sum_{i=1}^{2} Ai' \times I_i\right) - \log(P_{oct}^{\beta} \times A_b + 1) + B' \qquad \text{Equation 16.5}$$

$r^2 = 0.908$ SD = 0.108 n = 34

Zhang developed a non linear equation for blood to muscle distribution of neutral compounds.[19] He compiled a list of 36 different volatile organic compounds obtained from various literatures sources. The first two equations in 16.6, take into account the biological interaction of the VOCs in the muscle tissue phase, with regards to lipids and proteins. The product of the first two equations is used in the third equation (16.6) to make predictions for blood to muscle distribution for VOCs.

The descriptors used in equation 16.6 are, α is the molecular polarizability (used to characterise the steric bulk effeck), $\sum Ca$ is the sum of H-bond factor values for all acceptor substructures in the molecule and $\sum Q^+$ is the sum of all positive partial atomic charges for all atoms in the molecule. His non-linear regression equations, was compared to the composition of the tissue (i.e. contents of lipids, proteins and water in muscle and other tissues). Zhang concludes that his equations for blood to muscle distribution are related to the tissue composition.

$$\text{Log } P_{(\text{lipid})} = 0.091\alpha - 0.760\sum Ca + 0.600\sum Q^+ \qquad \text{Equation 16.6}$$

$$\text{Log } P_{(\text{protein})} = 0.553\sum Q^+$$

$$\text{Log } P_{(\text{blood to muscle})} = \log(10^{\text{logP (lipid)}} + 10^{\text{logP(protein)}} + 1) - 0.617$$

$r^2 = 0.884$ SD = 0.133 n = 36

Fiserova-Bergerova and Diaz used the head space chromatography method to determine experimentally air to muscle distribution for several VOCs.[2] They also compiled a literature survey of experimental values based on human subjects. Air to muscle and air to blood partition coefficients were used to calculate blood to muscle partition coefficients for VOCs. From the statistical analysis of Fiserova-Bergerova Diaz, it was found that there is a good inter-correlation between blood to muscle partition coefficients and blood to fat partition coefficients as shown from a plot. The intercept and the gradient of the straight line (linear) plots were included in the model to make a correlative equation to predict blood to muscle distribution of VOCs (see equation 16.6). There is no test set for this model which limits the information with regards to its predictability of VOCs partition. The standard deviation of this model is very large compared to other models mentioned earlier, and so the equation is very poor in terms of correlation, in contrast to other models mentioned earlier.[2]

$$P_{(\text{blood to muscle})} = 0.0326P_{\text{blood to fat}} + 0.4504 \qquad \text{Equation 16.7}$$

n = 34 $r^2 = 0.928$ SD = 0.7022

P denotes to the appropriate partition of the VOCs

The aim of the current work was to construct an equation as part of a general method for prediction of equilibrium blood-muscle distribution based on structure using the Abraham descriptors (**E**, **S**, **A**, **B** and **V**). Now it would be interesting to see if the new data for VOC blood to muscle distribution (humans and rats combined) makes a difference in the fits to the previous blood to muscle distribution models as mentioned. Table 16.3 (see appendix) lists the observed log P ($_{\text{blood to muscle}}$) data for blood to muscle distribution of VOCs. Note that all of the work is for blood and not blood or plasma.

Discussion on blood to muscle distribution for VOCs (humans and rats combined)

The Abraham solvation equation was used to correlate VOC blood to muscle distribution (see table 16.4). This model has 110 compounds in total in the dataset and the fits for this new equation appear to be poor (S.D = 0.207 and r^2 = 0.537), compared to other models by other authors. Although the standard deviation is good (SD= 0.207) and even better than the model by Fiserova-Bergerova and Diaz (SD = 0. 702), the overall square of the correlation is poorer (r^2 = 0.537). From the plot on page 226 most of the data is cluttered to the middle of the graph and that might explain why this model has a poor correlation for the full set and training set. This work is the most up to date model for VOC blood to muscle distribution reported for both species (humans and rats).

There were three outliers observed in this training set: 2,3,4-trimethylpentane, octane and 2,2,4-trimethylpentane (isooctane). The residual for these three compounds are approximately 0.35-0.56 log units and may result from experimental error.

EQUATIONS	N	r^2	SD	F	D
FULL SET **Log P** (blood to muscle) = -0.185 - 0.209**E** - 0.593**S** - 0.081**A** - 0.168**B** + 0.741**V**	110	0.537	0.207	24.125	5

TRAINING SET **Log P** (blood to muscle) = -0.057 - 0.131**E** - 0.606**S** - 0.230**A** - 0.218**B** + 0.594**V** **TEST SET** **N** = 55 **RMSE** = 0.213 **SD** (n-1) = 0.215 **AAE** = 0.169 **AE** = 0.037	55	0.481	0.214	9.087	5

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = root mean square error

Table 16.4. Summary of the equations for VOC blood to muscle distributions for humans and rats combined.

The inter-correlations of descriptors used for VOC blood to muscle distribution are shown in table 16.5. The largest r^2 value for E versus S descriptor is 0.31. It appears that there is no strong inter-correlation between the set of descriptors used in this model.

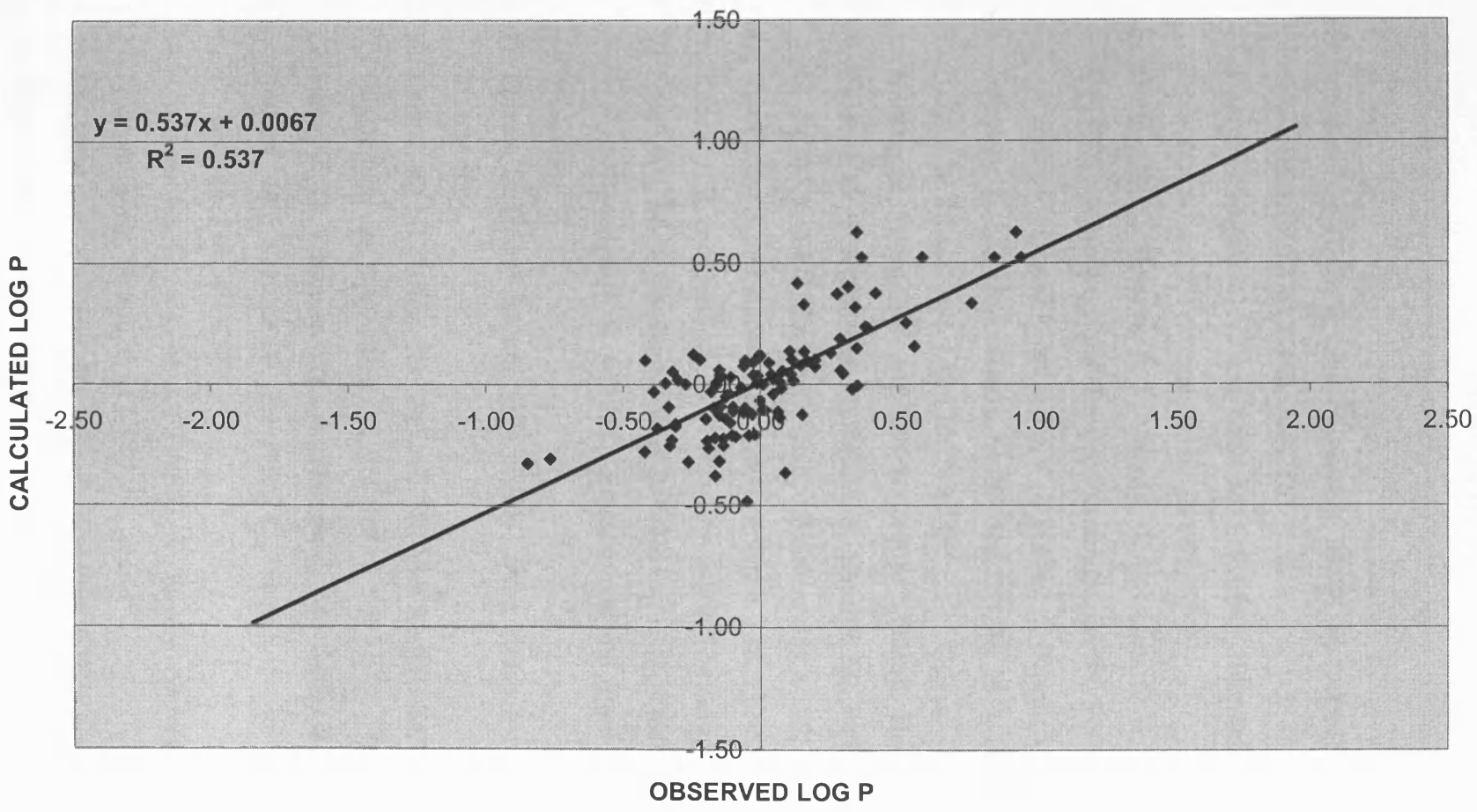
r	E	S	A	B	V
E	1.00				
S	0.56	1.00			
A	-0.03	0.08	1.00		
B	-0.07	0.32	0.32	1.00	
V	0.25	0.27	-0.04	0.33	1.00

Table 16.5. The inter-correlation of descriptors for 110 solutes using descriptors **E**, **S**, **A**, **B** and **V**

The Abraham solvation equation has been employed to correlate blood to muscle distribution for VOCs. To test the predictiveness of the model it was necessary to split the full data into training and test set using a random data set generator. A new model was developed using only those data in the training set, which was then used to predict values for the test set. The regression yielded

similar training equation and statistics to the first equation in table 16.4., except for the '**a**' coefficient. The test set of 55 compounds shows the average absolute error to be 0.169 for the two sets of data (observed and predicted log P). The root mean square error of 0.213 also indicates that there is reasonable predictive capability. The average error of 0.037 indicates that the equation of the model is not biased, as it is very close to zero. The blood to muscle model is the poorest statistical model reported for blood to tissue distribution of VOCs in the thesis and is ranked the lowest. The data points are cluttered in the middle of the graph and give rise to poor correlations as shown from the plot on page 226. The reason is that the coefficients in the equations for blood to muscle distribution are harder to interpret than those in the air to muscle distributions. This is because the coefficients in the blood to muscle equations reflect the difference in interactions between VOCs and blood and VOCs and muscle. This will also give rise to larger percentage error in the blood to muscle equation than the air to muscle equation due to these cancellation effects.

The **a**-coefficient (P = 0.70) and **b**-coefficient (P = 0.19) is not significant. The positive **v**-coefficients in the blood to muscle equation (16.4) indicate that increasing molecular size, can push VOCS out of blood and into the muscle tissue phase. The negative **e**-coefficients in the blood to muscle equation (16.4) indicate that London dispersion interaction and the presence of n- and Π-electrons pairs can push VOCS out of muscle and into the blood tissue phase. The coefficient of the 'polar' descriptor **S**, indicating that polarity is stronger in the blood phase than in muscle tissue phase. The solutes are pushed out of the muscle and into the blood phase. The Abraham equation can therefore successfully be used to correlate and predict blood to muscle distribution for humans and rats (see graph on page 226).


Observed log P versus claculated log P for blood to muscle distribution for VOCs
y = 0.537x + 0.0067
R2 = 0.537
CALCULATED LOG P
OBSERVED LOG P
1.50
1.00
0.50
-0.50
-1.00
-1.50
-2.50
-2.00
-1.50
-1.00
-0.50
0.50
1.00
1.50
2.00
2.50

References

1. Meulenberg, C. J. W. and Vijverberg, H. P. M (2000). Toxicology and Applied Pharmacology, **165**, 206-216.
2. Fiserova-Bergerova, V. and Diaz, M. L (1986). International Archives of Occupational and Environmental Health, **58**, 75-87.
3. Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H. and Anderson, M. E (1989). Toxicology and Applied Pharmacology, **98**, 87-99.
4. Abraham, M. H. and Weathersby, P. K (1994). J. Pharm. Sci, **83**, 1450-1455.
5. Sweeney, L. M., Himmelstein, M. W. and Gargas, M. L (2001). Chemical-Biological Interactions, **135-136**, 303-322.
6. Kaneko, T., Wang, P-Y. and Sato, A (2000). J. Occup. Health, **42**, 86-87.
7. Filser, J. G., Schmidbauer, R., Rampf, F., Baur, C. M., Putz, C. and Csanady, G. A (2000). Toxicology and Applied Pharmacology, **169**, 40-51.
8. Nussbaum, E. and Hursh, J. B (1957). Science, **125**, 552-553.
9. Knaak, J. B. and Smith, L. W (1998). Inhalation Toxicology, **10**, 65-85.
10. Medinsky, M. A., Bechtold, W. E., Birnbaum, L. S., Chico, D. M., Gerlach, R. F. and Henderson, R. F (1988). Fundamental and Applied Toxicology, **11**, 250-260.
11. Loizou, G. D. and Anders, M.W (1993). Drug Metabolism and Disposition, **21**, 634-639.
12. Loizou, G. D. Urban, G. Dekant, W. and Anders, M.W (1994). Drug Metabolism and Disposition, **22**, 511-517.
13. Filser, J. G., Csanady, GY. A., Hartmann, M., Denk, B., Kauffmann, A., Kessler, W., Kreuzer, P. E., Putz, C., Shen,J. H. and Stei, P (1996). Toxicology, **113**, 278-287.
14. Krishnan, K., Gargas, M. L., Fennell, T. R. and Anderson, M. E (1992). Toxicol. Indust. Health, **8**, 121-140.
15. Csanady, Gy. A., Denk, B., Putz, C., Kreuzer, P.E., Kessler, W., Baur, C., Gargas, M. L. and Filser, J. G (2000). Toxicology and Applied Pharmacology, **165**, 1-26.

16. Csanady, Gy. A., Kreuzer, P. E., Baur, C. and Filser, J. G (1996). A Toxicology, **113**, 300-305.
17. Borghoff, S. J., Murphy, J. E. and Medinsky, M. A (1996). Fundamental and Applied Toxicology, **30**, 264-275.
18. Balaz, S. and Luckacova (1999). Quant. Struct.-Act. Relat., **18**, 361-368.
19. Zhang, H (2004). J. Pharmaceutical Sciences, **93**, 1595-1604.

Chapter

17

Air to urine distribution for VOCs (human only)

The solubility of volatile organic compounds in biological fluids and tissues is of importance in areas such as anaesthesiology and toxicology. Indeed, a quantitative knowledge of such solubility is needed in pharmacokinetic modelling for VOCs in human subjects. Air to urine distributions for VOCs (or the ability of a molecule to enter urine from the air), have not frequently been obtained. The normal healthy urine sample for humans contains many chemical components of which water is the most abundant substance (95%), see figure 17.0.

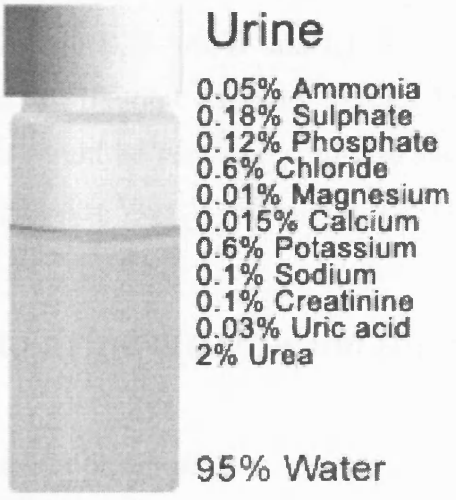


Figure 17.0. The chemical composition for normal healthy human urine

Abnormal urine has additional substances that include: glucose, ketones and various types of protein. Most of the reported published data for VOC distributions are for normal human urine samples (that does not affect binding of VOCs to

proteins). The reported distribution data from the literature for VOCs is collected in table 17.0 (see appendix).

A few reviews are available in which partition data for VOCs into urine are available.[1-4] The values for log K that were obtained from the literatures are displayed in Table 17.0 and they refer to the VOC distribution (air:urine) in humans. The solubility in urine for VOCs was first reported by Pezzagno (et al).[1] They reported partitions of a number of VOCs in human urine at 310 K. Each of these data was expressed in terms of partition coefficient, K. Abraham then used this data to construct an equation for air to urine distribution for 34 solutes (see table 17.1).[4] The E descriptor in the equation appeared not to be significant in the correlation, so this was removed. Besides this review, there has not been any other reported model for air to urine distribution in humans (with any other data sets).

EQUATION	N	r^2	SD	F	D
Log K (air:urine) = -0.314 + 0.854**S** + 3.445**A** + 3.720**B** + 0.056**L**	34	0.951	0.263	140.10	4

Table 17.1. Equation for air to urine distribution for VOCS in humans.[4]

The aim of this work is attempt to correlate 50 log K (solutes in urine) values with the Abraham descriptors (**E**, **S**, **A**, **B** and **L**), to obtain equations to enable prediction of air to urine distribution (humans only). The human urine sample contains 95% water, so it would be very useful to also see if the second Abraham equation can be constructed using V as the fifth descriptor rather than L.

Discussion on air to urine distribution for VOCs in humans

The two Abraham solvation equations were used to correlate VOC air to urine distribution (see table 17.2). The fits for the first equation (s.d = 0.322) appear to be poorer than for the second equation (s.d = 0.306). This shows that **V** descriptor is a rather better descriptor than **L** in this distribution process (thus giving rise to correlation r^2 = 0.936). There were six outliers observed for this set of 50 compounds and they are: enflurane, sulphur hexafluoride, cyclopentadiene, 2,3-

dimethylbutane, decane and nonane with residuals of 0.80-1.07 log units. The second model with the **V** descriptor has the same compounds as outliers as the first model, plus an additional compound that is ethane. The residuals for this second model for the seven outliers are 0.71-1.28 log units. Although there is no direct explanation to why these compounds are outliers, but it may be due experimental errors.

EQUATIONS	N	r^2	SD	F	D
Log K (air:urine) = - 0.973 - 0.689**E** + 1.285**S** + 4.451**A** + 3.288**B** + 0.300**L**	44	0.936	0.322	111.825	5
Log K (air:urine) = - 0.716 - 0.322**E** + 1.320**S** + 4.062**A** + 3.257**B** + 0.643**V**	43	0.936	0.306	107.407	5

N = number of compounds used in the equation

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

Table 17.2. Summary of the two equations for VOC air to urine distributions

The inter-correlations of descriptors are shown in table 17.3. The largest r^2 value for E versus S and E versus L descriptor is 0.30. It appears that there is no strong inter-correlation between the sets of descriptors used in this model.

The Abraham solvation equation has been employed to correlate air to urine distribution for VOCs.. No test sets have been used for this model as there is a limited number of compounds available. The descriptors used in the equation reflect solute/solvent interactions. Therefore, one can interpret these coefficients in terms of effect that particular interactions have on the process under consideration.

The statistics for this model is good, as indicated by the even spread of data shown from the plot (see graph on page 233). This model is not much better in describing the VOC air to urine distribution process compared to the air to muscle model mentioned earlier in chapter 16. The reason could be due to the limited amount of data in the dataset and to the presence of other components in urine.

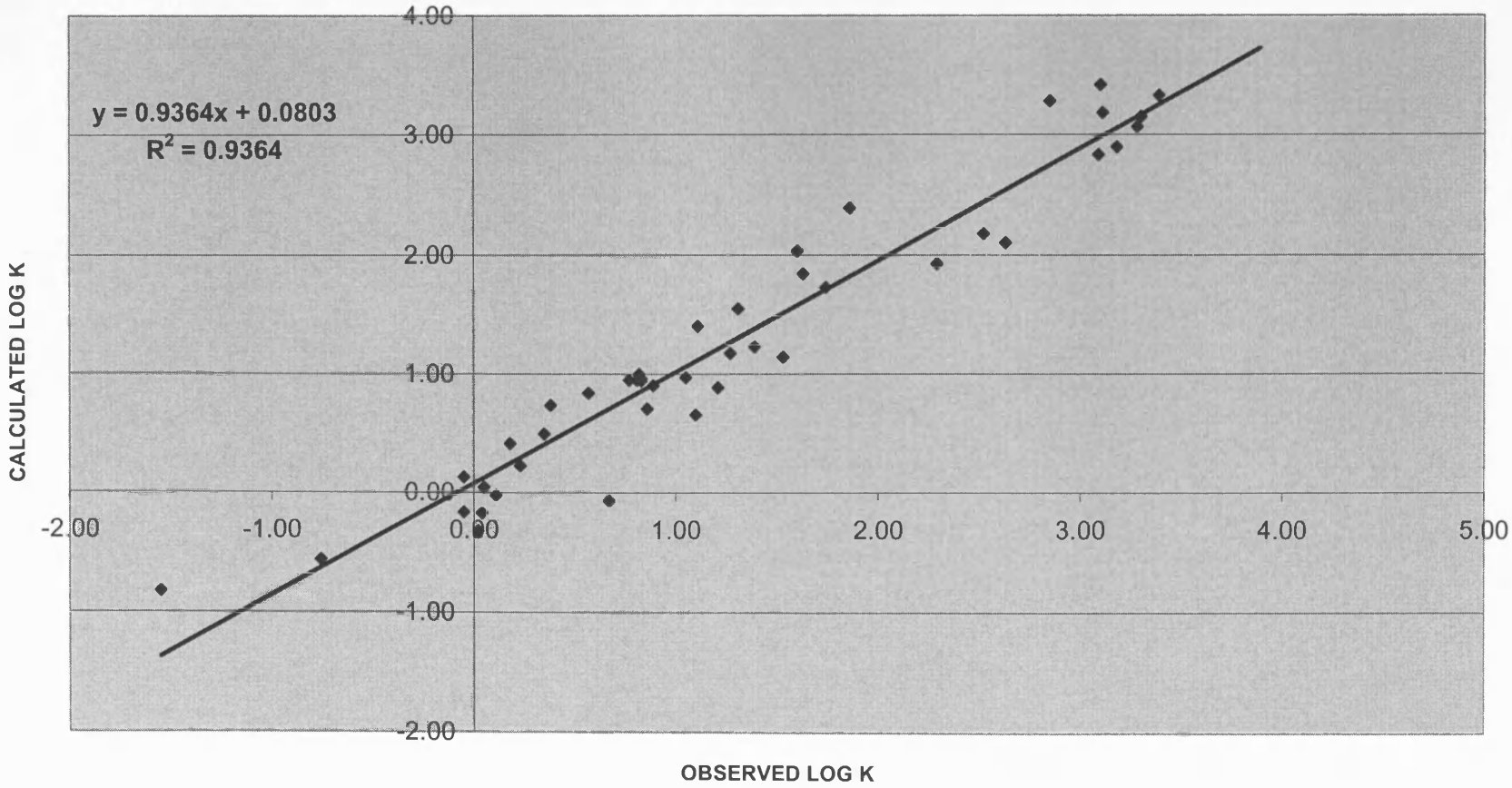
r	E	S	A	B	L
E	1.00				
S	0.55	1.00			
A	-0.04	0.01	1.00		
B	-0.25	0.46	0.35	1.00	
L	0.55	0.39	-0.38	-0.07	1.00

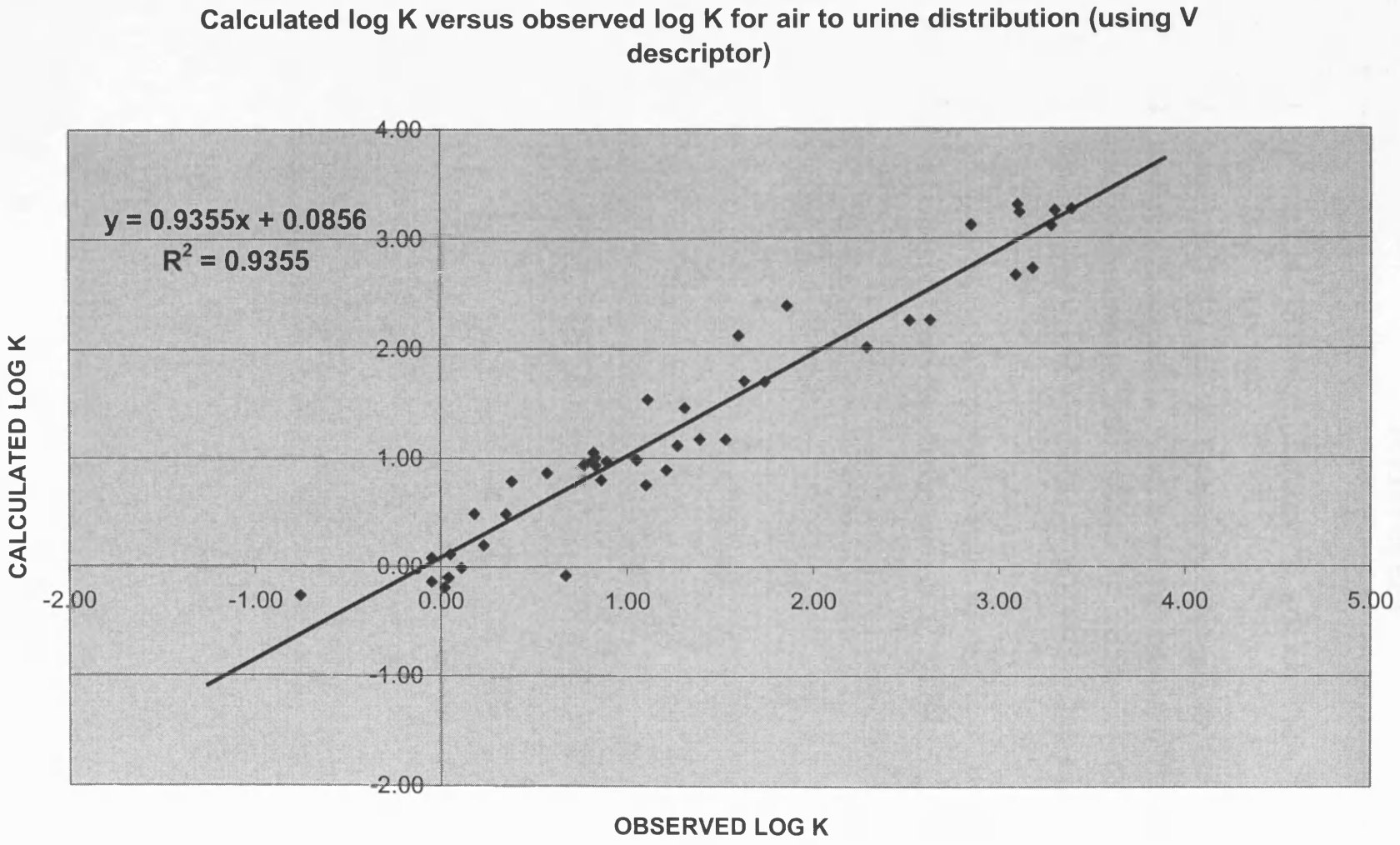
Table 17.3. The inter-correlation of descriptors for 44 solutes using descriptors **E**, **S**, **A**, **B** and **L**

EQUATIONS	N	r^2	SD	F	D
Log K (air:water) = -1.271 + 0.822**E** + 2.743**S** + 3.904**A** + 4.814**B** -0.213**L**	392	0.992	0.185	10229	5
Log K (air:water) = -0.994 + 0.577**E** + 2.549**S** + 3.813**A** + 4.841**B** -0.869**V**	408	0.995	0.151	16810	5

Table 17.4. VOC air to water distributions for 408 solutes[5]

The positive **l**-coefficient in the air to urine equation from table 17.2 indicates that increasing molecular size and London dispersion interaction can push VOCs out of air and into urine. The **E**-coefficient is not significant (P = 0.08). The coefficients of the 'polar' descriptors **S**, **A**, and **B** are all positive, indicates that polarity and hydrogen-bond donor and acceptor ability is stronger in the urine phase than in gas phase. The solutes are pushed out the air and into the urine phase. The positive **S**, **A** and **B** coefficients in this equation are also similar to the air to water equation (see table 17.4), although this equation has a larger training set of 408 solutes.[5] It is interesting that the **v**-coefficient for log K (air:urine) is positive (Table 17.2), but for log K (air:water) it is negative (table 17.4). The 5% components other than water in urine must have a considerable effect in making urine less hydrophilic than water.


Calculated log K versus observed log K for air to urine distribution
(using L descriptor)
y = 0.9364x + 0.0803
R2 = 0.9364
CALCULATED LOG K
OBSERVED LOG K
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00
5.00


Calculated log K versus observed log K for air to urine distribution (using V descriptor)
y = 0.9355x + 0.0856
R2 = 0.9355
CALCULATED LOG K
OBSERVED LOG K
4.00
3.00
2.00
1.00
0.00
-1.00
-2.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00
5.00

References

1. Imbriani, M., Ghittori, S., Pezzagno, G. and Capodagio, E (1985). G. Ital. Med. Lav., **7**, 133-140.
2. Guitart, R (1993). Revista Espanola de Fisiologia., **49**, 195-202.
3. Batterman, S., Zhang, L., Wang, S. and Franzblau, A (2002). The Science of the Total Environment, **284**, 237-247.
4. Abraham, M. H. and Weathersby, P. K (1994). J. Pharmaceutical Sciences, **83**, 1450-1455.
5. Abraham, M. H., Andonian-Haftvan, J., Whiting, G. S., Leo, A. and Taft, R. SA (1994). J. Chem. Soc. Perkin Trans 2., 1777-1791.

Chapter

18

Air to olive oil distribution for VOCs

In this chapter, the Abraham solvation equation shall be used to construct models that include air to olive oil distribution for VOCs.

Air to olive oil distribution for VOCs

The partition of VOCs from air to olive is of importance in the pharmaceutical industry, biochemical and environmental sciences. Air-saline (as discussed in chapter 19) and air-olive oil partition coefficients are used to predict various air-solvent or air-tissue partitions for VOCs. The differences in the air-tissue and air to blood partition can used to predict blood to tissue distributions for a large number of VOCs. There have been a number of citations for air to olive oil distribution data for VOCs.[1-11]

Scheme 18.0. Bottles of pure olive oil

The purpose of this work is to gather a set of data from the literature and to attempt to construct a new equation to predict the distribution of solutes from air to olive oil. The correlative equations of Abraham shall be constructed based on the obtained dataset at a temperature of 310K. The complete list of the VOC log K data is shown in table 18.0 (appendix). Various models for the correlation of air to olive oil distribution for VOCs have been made by Weathersby and Abraham[1], Klopman, et al.,[4] and Fuchs and Abraham.[5]

Weathersby and Abraham compiled a dataset of 88 VOCs in their model to predict air to olive oil distribution at 310 K. Most of their data was compiled from previous literature reviews. They used the Abraham solvation equation to correlate the log K data for air to olive oil with the five Abraham descriptors (**E**, **S**, **A**, **B** and **L**). The Abraham equation used for this model is shown in equation 18.0.[1]

$$\text{Log K}_{(\text{air to olive oil})} = -0.240 - 0.018\mathbf{E} + 0.806\mathbf{S} + 1.469\mathbf{A} + 0.891\mathbf{L} \qquad \text{Equation 18.0}$$

$r^2 = 0.997$ S.D = 0.082 n = 88

Klopman et al modelled the air to olive oil distribution for VOCs based on a number of properties of solutes.[4] The purpose of the work was to derive coefficients for the equation to predict air to olive oil distributions as surrogates for air to blood distribution, used for pharmacokinetic profiling for humans or rats. The Log K $_{(\text{air to olive oil})}$ models were also derived from multivariate regression analysis, based on the basic sets of functional groups ($L_{1\text{-}94}$) and correction factors obtained from the CASE computer program which calculates the intramolecular group interaction of VOCs. The basic parameters for the functional groups ($L_{1\text{-}94}$) are: OH, CHO, COOH, COO, $CONH_2$, CONH, CON, CON=, CO, CS, NO, NO_2, PO, SO, SO_2, NH_2, NH, CN, SH, etc). The predictions for the first equation appear to be good.

TRAINING SET Equation 18.1

$$\text{Log } K^{a}_{(\text{air to olive oil})} = 0.9608 - 0.0001L_1 + 0.4706L_2 + 0.7265L_3 + 0.9577L_4 - 0.0101L_5 + 0.4368L_6 + 0.3896L_{11} + 0.4284L_{14} + 0.9096L_{15} - 0.4270L_{17} + 0.3436L_{19} + 0.7891L_{21} + 0.5198L_{24} + 0.5431L_{25} + 0.0911L_{30} + 0.8844L_{34} + 0.9121L_{40} + 0.0914L_{55} - 0.3485L_{68} + 0.1128L_{73} - 0.0883L_{89} - 0.0249L_{93} - 0.9110L_{94}$$

n = 159 $r^2 = 0.9377$ SD = 0.2954

TEST SET

n = 36 SD = 0.271

TOTAL SET Equation 18.2

$$\text{Log } K^{b}_{(\text{air to olive oil})} = 0.2541 + 0.0057L_1 + 0.2399L_2 + 0.3233L_3 + 0.2153L_4 - 0.0678L_5 + 0.2688L_6 + 0.0847L_7 + 0.2357L_{11} + 0.2598L_{14} + 0.4826L_{15} - 0.4539L_{17} - 0.0329L_{19} - 0.2673L_{21} + 0.6331L_{24} + 0.3297L_{55} + 0.9196L_{57} - 0.1014L_{68} + 0.0478L_{73} - 0.0678L_{89} - 0.0198L_{93} - 0.4828L_{94} + 0.1675L_{95}$$

n = 192 $r^2 = 0.9678$ SD = 0.2056

NO TEST SET

[a]Equation was derived from 119 experimental values and 40 calculated values for log K air to olive oil distribution. [b]Equation was derived from 192 calculated values for log K air to olive oil distribution

Klopman et al. found that the most important descriptor in the model for determining Log K air to olive oil distribution was the molecular weight (L_{95}). The fits for the second equation appears to be a little better because of this additional descriptor taking into account of the molecular weight of the VOC. Note that both equations include a large amount of calculated data as opposed to experimental data,[4] and this may well lead to much better statistics than for the usual 'observed' values.

Abraham and Fuchs also modelled VOC air to olive distribution using a linear regression technique (olive oil at 310K).[5] They achieved reasonable fits for the correlation of air to olive oil distribution for 52 VOCs (see equation 18.3).

$$\text{Log } K = 0.6468 - 0.0384V + 0.2233(MR) + 0.0869(\mu^2)$$ Equation 18.3

n = 52 SD = 0.233 $r^2 = 0.947$

The endoergic work of creating a cavity in the solvent is given by vV, where V is a solute volume and the exoergic solute-solvent interactions are given by c(MR) and $d\mu^2$, being dispersion and dipole-induced dipole (or dipole-dipole) effects,

respectively; MR and μ are the solute molar refraction and dipole moment. The air to olive oil distribution is denoted by log K at 310 K.[5]

It is found that two out of the three models do not have any test sets that include statistical information that can be used to test predictions. It is disappointing that Klopman et al. did not base their models fully on experimental data for the training set equations they have constructed. Some improvement to these models can be made by adding new VOC experimental data, to see if the new equation for the full set has better fits than the published models. Table 18.0 (see appendix) contains up to date literature data for air to olive distribution for VOCs.[1-11]

Discussion on air to olive oil distribution for VOCs

Abraham's solvation equation has successfully been employed to correlate and predict VOC air to olive oil distribution as shown in table 18.1. This model contains 213 compounds in the dataset, and the fits for this new equation appear to be quite good (S.D = 0.197 and $r^2 = 0.981$), for a diverse range of VOCs. This work is the most up to date model for VOC air to olive oil distribution. There were three outliers observed in the training set: CF2HCFHCH2CF2H, N-Methylimidazole and 2-Methyl-1,3-butadiene (isoprene). The residuals for these three compounds are approximately 0.80-2.62 log units and may result from experimental error.

EQUATIONS	N	r^2	SD	F	D
FULL SET					
Log K (air to olive oil) = -0.159 - 0.277**E** + 0.903**S** + 1.698**A** - 0.091**B** + 0.876**L**	213	0.981	0.197	2090.954	5
TRAINING SET					
Log K (air to olive oil) = -0.124 - 0.266**E** + 0.830**S** + 1.620**A** - 0.018**B** + 0.868**L**	107	0.973	0.232	739.167	5
TEST SET					
N = 106 **RMSE** = 0.160 **SD** (n-1) = 0.161					
AAE = 0.105 **AE** = -0.004					

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = root mean square error

Table 18.1. Summary of the equations for VOC air to olive oil distributions

The inter-correlations of the descriptors (**E**, **S**, **A**, **B** and **L**) used for VOC air to olive oil distribution are shown in table 18.2. The largest r^2 value for E versus S descriptor is 0.35. It appears that there is no strong inter-correlation between the set of descriptors used in this current model.

r	E	S	A	B	L
E	1.00				
S	0.59	1.00			
A	-0.01	0.11	1.00		
B	0.12	0.52	0.30	1.00	
L	0.52	0.30	-0.07	0.18	1.00

Table 18.2. The inter-correlation of descriptors for 213 solutes using descriptors **E**, **S**, **A**, **B** and **L** for air to olive oil distribution model

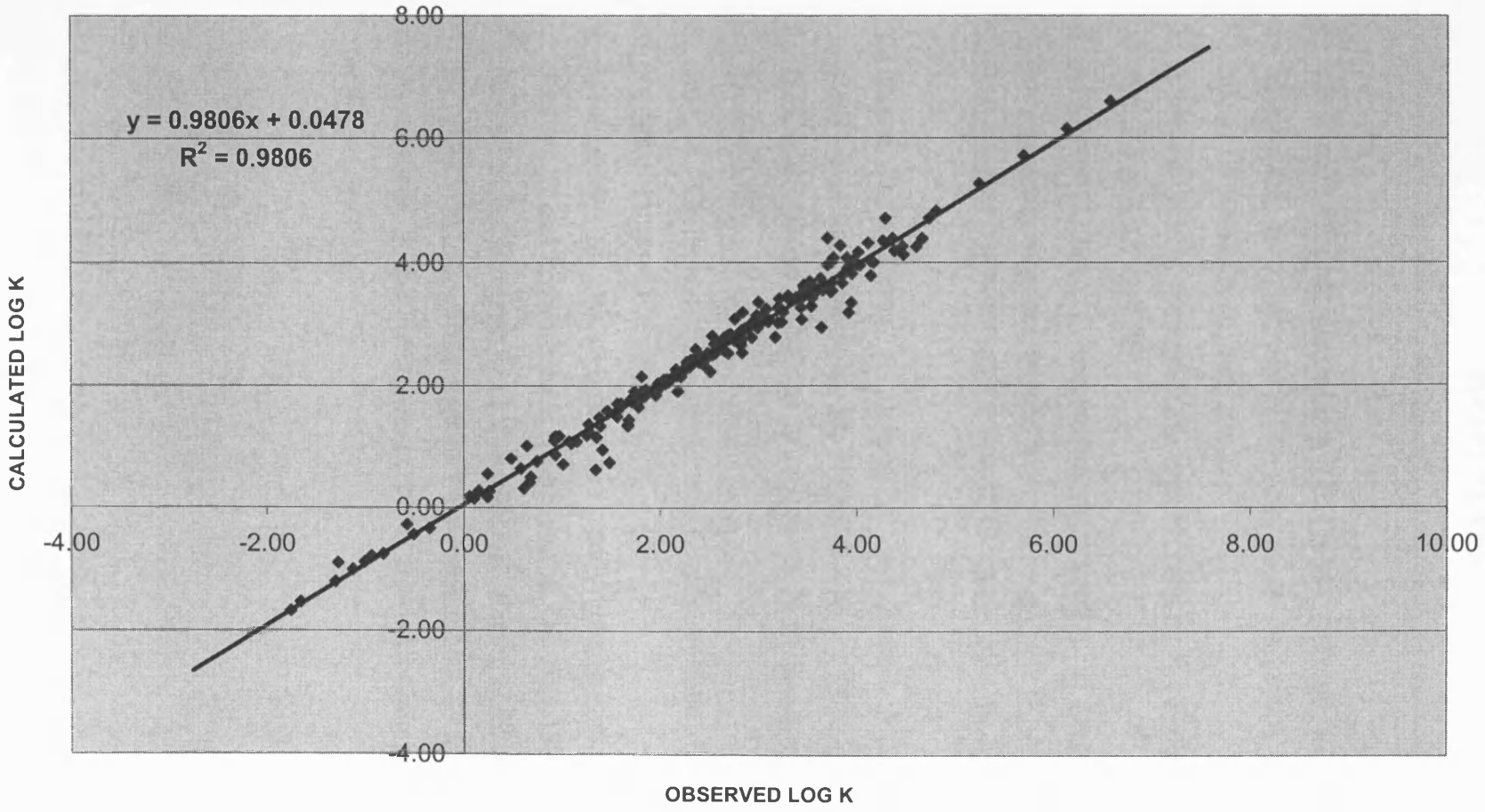
The Abraham solvation equation has successfully been employed to correlate air to olive oil distribution for VOCs. To test the predictive ability of the model it was necessary to split the full data into training and test set using a random data set generator (Kennard-Stone). A new model was developed using only those data in the training set, which was then used to predict values for the test set. The regression yielded similar training equations and statistics to the first equation in table 18.1., except for the **b** coefficient, which appeared to be slightly different. The test set of 106 compounds leads to an average absolute error of 0.105 for the two sets of data (observed and predicted log K) and a root mean square error of 0.160, which indicates that the predictive capability of the training equation is good. The average error of -0.004 indicates that the equation of the model is not biased, as it is very close to zero.

The **b**-coefficient is not significant (P = 0.22). The positive **l**-coefficients in the air to olive oil equation (18.1) indicate that increasing molecular size can push

VOCs out of air and into the olive oil phase. The negative **e**-coefficients in the air to olive oil equation (18.1) indicate that London dispersion interaction and the presence of n- and Π-electrons pairs can push VOCs out of the olive oil phase and into the air. The coefficients of the 'polar' descriptors **S** and **A** are positive, indicating that polarity and hydrogen-bond acceptor ability is stronger in the olive oil tissue phase than in gas phase. The solutes are pushed out of the air and into the olive oil phase. Finally, the Abraham equation can successfully be used to correlate and predict air to olive oil distribution (see graph on page 242)

The statistics for this model is very good, as indicated by the large spread of data shown from the plot. It is found that it is possible to construct an equation capable of predicting further values of log $K_{olive\ oil}$ to around 0.2 log units. This equation is the best model for predicting log K values for air to olive oil distribution for VOCs and is far better in terms of fits than other models (air to tissue distribution of VOCs) reported earlier in this thesis

Although, now this model has the largest data set ever reported; the fits for the full set (table 18.1) appear to better than the models of Klopman et al., and of Fuchs and Abraham, but not as good as the fit found by Abraham and Weathersby. However, there is now a larger diversity of compounds in the data set than ever before, and on this basis the present equations are more soundly based than any of those previously published.


Calculated log K versus observed log K for air to olive oil distribution for VOCs
y = 0.9806x + 0.0478
R2 = 0.9806
8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
10.00
CALCULATED LOG K
OBSERVED LOG K

References

1. Abraham, M. H. and Weathersby, P. K (1994). J. Pharm. Sci, **83**, 1450-1455.
2. Meulenberg, C. J. W. and Vijverberg, H. P. M (2000). Toxicology and Applied Pharmacology, **165**, 206-216.
3. Abraham, M. H., Grellier, P. L. and McGill, R. A (1987). J. Chem. Soc. Perkin Trans. **II**., 797-803.
4. Klopman, G., Ding, C. and Macina, O. T (1997). J. Chemical Information and Computer Sciences, **37**, 569-575.
5. Abraham, M. H. and Fuchs, R (1988). J. Chem. Soc. Perkin Trans. **II**, 523-527.
6. Loizou, G. D. Urban, G. Dekant, W. and Anders, M.W (1994). Drug Metabolism and Disposition, **22**, 511-517.
7. Loizou, G. D. and Anders, M.W (1993). Drug Metabolism and Disposition, **21**, 634-639.
8. Falk, A., Gullstrand, E., Lof, A. and Wigaeous-Hjelm, E (1990). British Journal of Industrial Medicine, **47**, 62-64.
9. Meulenberg, C. J. W., Wijnker, A. G. and Vijverberg, H. P. M (2003). J. Toxicology and Environmental Health, Part A, **66**, 1985-1998.
10. Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H. and Anderson, M. E (1989). Toxicology and Applied Pharmacology, **98**, 87-99.
11. Olivier-Droz, P. and Fernandez, J (1977). Helvetica Chemica Acta., **60**, 454-458.

Chapter

19

Air to saline and saline to olive oil distribution for VOCs

In this chapter, the Abraham solvation equation shall be used to construct models that include air to saline and saline to olive oil distribution for VOCs.

Air to saline distribution for VOCs

The air to saline distribution of VOCs, or the ability of a volatile organic molecule to enter the saline phase has become a subject of interest to the pharmaceutical industry, biochemical and environmental sciences. Air-saline and air-olive oil partition coefficients are used to predict various air-solvent or air-tissue partitions for VOCs. The differences in the air-tissue and air to blood partition can used to predict blood to tissue distributions for a large number of VOCs.

Most of the literature data compiled for saline is for 0.9% NaCl solution and this is indicated clearly in table 19.0 (see appendix). There have been a number of reported published data for air to saline distribution for VOCs.[1-5] There have been no specific reviews reported yet that attempt to make models to predict air to saline distribution for VOCs. The present work consists of a large data set of 134 compounds, which is all that has been published.[1-5] The VOC data for air to saline distribution is at 310 K. The aim is to attempt to construct a model based on the Abraham equations to predict the distribution process. This equilibrium distribution for air to saline distribution will be the first reported model in this area. Since saline

is so close to water, it would be useful to see if a second Abraham equation can be constructed using V as the fifth descriptor, rather than L, in correlating the distribution process.

The aim of this work is to construct an equation as part of a general method for 'high-throughput' prediction of equilibrium air to saline distribution based on structure using the Abraham descriptors (E, S, A, B and L or V).[6]

Discussion on air to saline distribution for VOCs

The Abraham solvation equation has been employed to correlate and predict VOC air to saline distribution, as shown in table 19.1.

EQUATIONS	N	r^2	SD	F	D
FULL SET **Log K** (air to saline) = -1.254 + 0.653**E** + 2.447**S** + 4.144**A** + 4.247**B** - 0.211**L**	134	0.961	0.302	625.836	5
TRAINING SET **Log K** (air to saline) = -1.214 + 0.804**E** + 2.206**S** + 3.978**A** + 4.274**B** - 0.200**L** **TEST SET** **N** = 67 **RMSE** = 0.259 **SD** (n-1) = 0.261 **AAE** = 0.206 **AE** = -0.009	67	0.954	0.353	252.645	5
FULL SET (with 'v' coefficient) **Log K** (air to saline) = -1.034 + 0.300**E** + 2.310**S** + 4.079**A** + 4.211**B** - 0.763**V**	133	0.966	0.284	714.153	5
TRAINING SET (with 'v' coefficient) **Log K** (air to saline) = -1.083 + 0.345**E** + 2.277**S** + 3.791**A** + 4.348**B** - 0.728**V** **TEST SET** **N** = 66 **RMSE** = 0.241 **SD** (n-1) = 0.243 **AAE** = 0.178 **AE** = -0.033	67	0.960	0.329	291.539	5

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = root mean square error

Table 19.1. Summary of the equations for VOC air to saline distributions

This model has 134 compounds in the dataset, and the fits for the first full equation appear to be good (S.D = 0.302 and r^2 = 0.961), considering this data set contains a diverse range of VOCs. The fits for the second full equation (s.d = 0.284) appear to be slightly better than for the first equation (s.d = 0.302), although there were only 133 compounds in this data set. The reason for this is that, a new set of outliers was observed in this regression when using the **V** descriptor. These outliers are: 1,3-difluoropropane, heptane, 2,2,4-trimethylpentane, 3-methylpentane and CF2HCFHCH2CF2H (residuals of 0.91-1.55 log units). This justifies the fact that **V** descriptor is a more important factor than the **L** descriptor in this distribution process (thus giving rise to better correlation r^2 = 0.966). There were five outliers observed from this full set of the equation using the **L**-descriptor: 1,3-difluoropropane, heptane, 2,2,4-trimethylpentane, 3-methylpentane and CF2HCFHCH2CF2H; the residuals for the five outliers are approximately 0.91-1.55 log units.

The inter-correlations of the descriptors (**E**, **S**, **A**, **B** and **L**) used for VOC air to saline distribution are shown in table 19.2. The largest r^2 value for E versus L descriptor is 0.50. It appears that there is no strong inter-correlation between the set of descriptors used in this current model.

r	E	S	A	B	L
E	1.00				
S	0.61	1.00			
A	-0.14	-0.04	1.00		
B	-0.04	0.28	0.46	1.00	
L	0.71	0.47	-0.19	0.13	1.00

Table 19.2. The inter-correlation of descriptors for 134 solutes using descriptors **E**, **S**, **A**, **B** and **L** for air to saline distribution model

The Abraham solvation equation has been employed to predict air to saline distribution for VOCs. To examine the predictive ability of the model, it was

necessary to split the full data into training and test set using a random data set generator (Kennard-Stone). A new model was developed using only those data in the training set, which was then used to predict values for the test set. The regression yielded similar training equations and statistics to the first equation in table 19.1., except for the **e**-coefficient, which appeared to be slightly different.

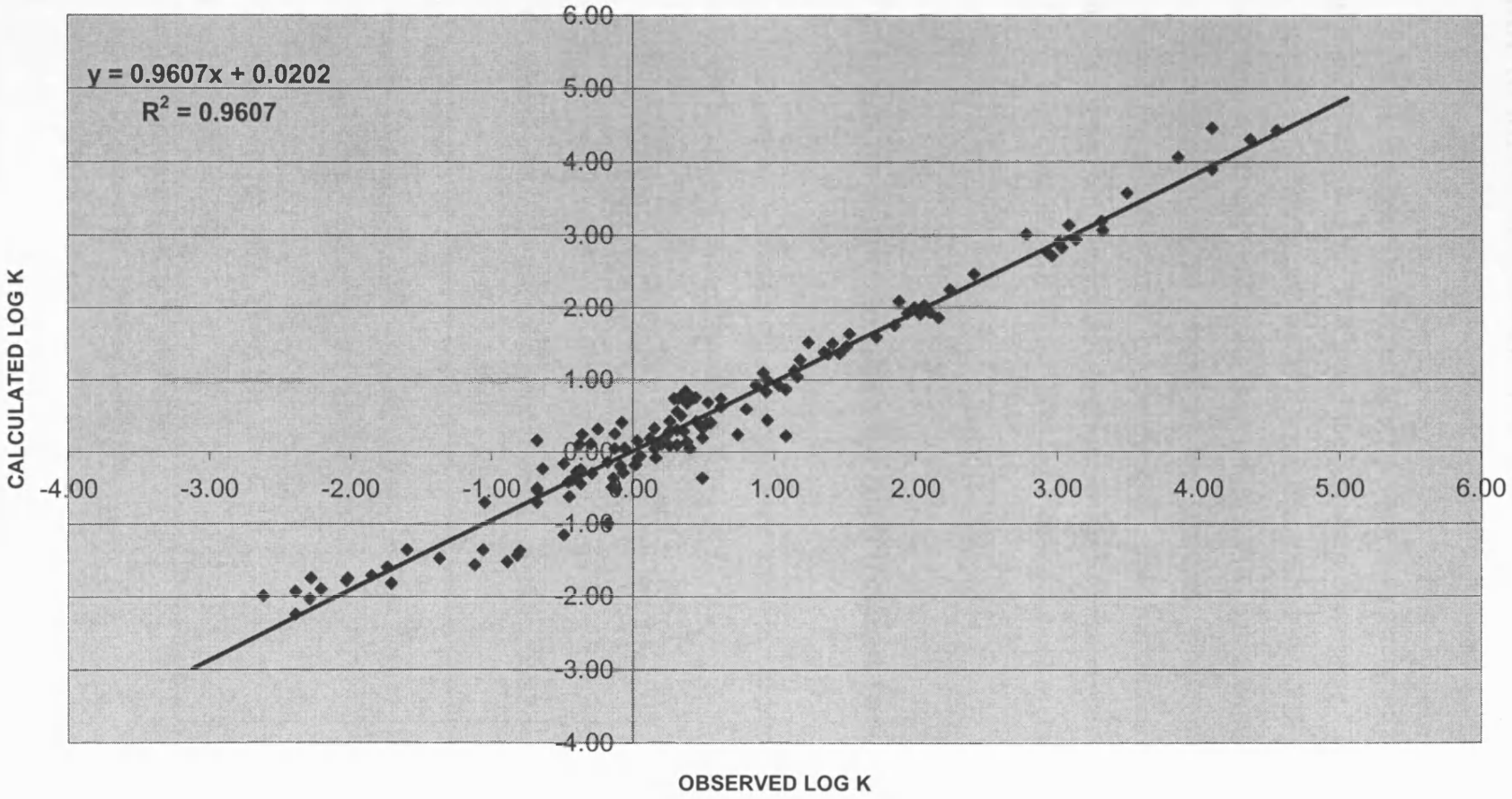
The test set of 67 compounds shows the average absolute error to be 0.206 for the two sets of data (observed and predicted log K). The root mean square error of 0.259 also indicates that there is a small difference between the observed value and predicted data for air to saline distribution for VOCs. The average error of -0.009 indicates that the equation of the model is not biased, as it is very close to 0.

The statistics for this model is good, as is indicated by the even spread of data shown from the plot. It is found that it is possible to construct an equation capable of predicting further values of log K_{saline} to around 0.2 log units. The reason why this model is better than other air to tissue models (VOCs) at steady-state is that the measurements for solvents (saline or olive oil) have less background noise than tissue samples.

The positive **e**-coefficients in the air to saline equation (19.1) indicates that increasing the presence of n- and Π-electrons pairs can push VOCs out of air phase and into the saline phase. The negative **l**-coefficients in the air to saline equation indicates that increasing molecular size and the London dispersion interaction this can push VOCs out of saline and into the air phase. The coefficients of the 'polar' descriptors **S**, **A**, and **B** are all positive, indicating that polarity and hydrogen-bond donor and acceptor ability is stronger in the saline phase than in the gas phase. The solutes are pushed out of the air and into the saline phase. Finally, the Abraham equation can successfully be used to correlate and predict air to saline distribution (see graph on page 248).

The general chemistry associated with the second equation, (containing the **V** coefficient in table 19.1), is similar to that of the first equation.

Calculated log K versus observed log K for air to saline distribution for VOCs (using the L-descriptor)



Saline to olive oil distribution for VOCs

It is also important to find out whether there is a correlation between the saline to olive oil distribution for VOCs and the Abraham descriptors. The log P data for saline to olive oil distribution for VOCs can be obtained by combining the log $K_{olive\ oil}$ data (chapter 18, table 18.0) with the log K_{saline} VOC data (chapter 19, table 19.0).

There have been no reports in the literature, where models have been constructed for saline to olive oil distribution. The correlative equation shall be constructed based on datasets that represent average literature values for distribution at 310 K. The complete list of the VOC log P data for these averages, are shown in table 19.3 (see appendix).

Discussion on saline to olive oil distribution for VOCs

The Abraham solvation equation has been used to correlate VOC saline to olive oil distribution as shown in table 19.4.

EQUATIONS	N	r^2	SD	F	D
FULL SET					
Log P (saline to olive oil)= 0.066 + 0.700**E** - 0.961**S** -2.152**A** - 4.504**B** + 4.068**V**	131	0.982	0.207	1370.563	5
TRAINING SET					
Log P (saline to olive oil)= 0.006 + 0.695**E** - 0.959**S** -2.035**A** - 4.512**B** + 4.172**V**	66	0.977	0.253	500.655	5
TEST SET					
N = 65 **RMSE** = 0.162 **SD** (n-1) = 0.164					
AAE = 0.129 **AE** = 0.057					

N = number of compounds used in the training set

r^2 = is the overall square of the correlation coefficient

SD = the standard deviation

F = F-statistic

D = the number of descriptors used in the model

RMSE = root mean square error

Table 19.4 Summary of the equations for VOC saline to olive oil distributions

This saline to olive oil model has a total of 131 compounds in the dataset, and the fits for the new equation appear to be good (SD = 0.207 and r^2 = 0.982), considering now there is a diverse range of VOCs. There were two outliers observed from this full set: ethyl t-pentyl ether and hexaflurobenzene; the residuals for the two outliers are approximately 1.22-1.36 log unit. The inter-correlations of descriptors used for VOC saline to olive oil distribution are shown in table 19.5. The largest r^2 value for E versus S descriptor is 0.37. It appears that there is no strong inter-correlation between the set of descriptors used in this model.

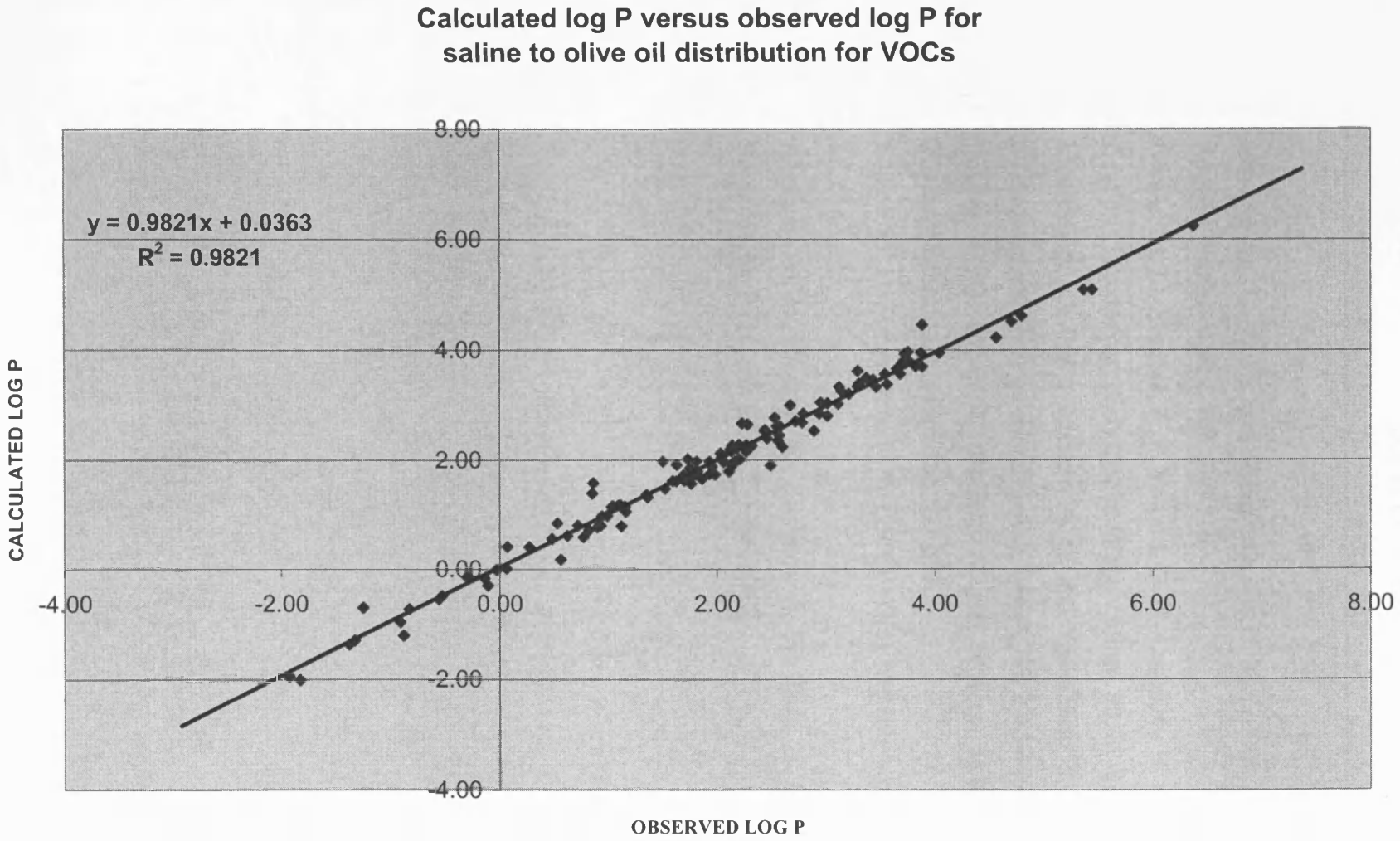
r	E	S	A	B	V
E	1.00				
S	0.61	1.00			
A	-0.15	-0.05	1.00		
B	-0.04	0.29	0.46	1.00	
V	0.24	0.08	-0.23	0.13	1.00

Table 19.5. The inter-correlation of descriptors for 131 solutes using descriptors **E**, **S**, **A**, **B** and **V**

The solvation equation of Abraham has been employed to correlate saline to olive oil distribution for VOCs. To test the predictiveness of the model it was necessary to split the full data into a training set and a test set using a random data set generator. A new model was developed using only those data in the training set, which was then used to predict values for the test set. The regression yielded similar training equation and statistics to the first equation in table 19.4. The root mean square error of 0.162 also indicates that there is reasonable predictive capability. The average error of 0.057 indicates that the equation of the model is not biased, as it is very close to zero.

The positive **e**- and **v**-coefficients in the saline to olive oil equation (19.4) indicate that increasing molecular size, London dispersion interaction and the presence of n- and Π-electrons pairs can push VOCS out of saline and into the olive oil phase. The coefficients of the 'polar' descriptors **S**, **A**, and **B** are all negative, indicating that polarity and hydrogen-bond donor and acceptor ability is stronger in the saline phase than in olive oil phase. The solutes are pushed out of the olive oil

phase and into the saline phase. The Abraham equation can therefore successfully be used to correlate and predict saline to olive oil distribution (see graph on page 252).


Calculated log P versus observed log P for saline to olive oil distribution for VOCs
y = 0.9821x + 0.0363
R2 = 0.9821
8.00
6.00
4.00
2.00
0.00
-2.00
-4.00
-4.00
-2.00
0.00
2.00
4.00
6.00
8.00
CALCULATED LOG P
OBSERVED LOG P

References

1. Meulenberg, C. J. W. and Vijverberg, H. P. M (2000). Toxicology and Applied Pharmacology, **165**, 206-216.
2. Gargas, M. L., Burgess, R. J., Voisard, D. E., Cason, G. H. and Anderson, M. E (1989). Toxicology and Applied Pharmacology, **98**, 87-99.
3. Meulenberg, C. J. W., Wijnker, A. G. and Vijverberg, H. P. M (2003). J. Toxicology and Environmental Health, Part A, **66**, 1985-1998
4. Loizou, G. D. Urban, G. Dekant, W. and Anders, M.W (1994). Drug Metabolism and Disposition, **22**, 511-517.
5. Loizou, G. D. and Anders, M.W (1993). Drug Metabolism and Disposition, **21**, 634-639.
6. Abraham, M. H., Ibrahim, A., Zissimos, A. M., Zhao, Y. H., Comer, J. and Reynolds, D. P (2002). DDT, **7**, 1056-1063.

Chapter

20

Conclusion on air to tissue and blood to tissue distribution for VOCs

The equilibrium distribution models for: air to tissue, air to solvent, blood to tissue and blood to biophase distribution were successfully employed using the Abraham equations. The following tissues, biophases and solvents thus examined are blood, plasma, brain, kidney, fat, heart, liver, lung, muscle, urine, olive oil and saline. The coefficients for the equations are summarised in tables 20.0 and 20.1.

The partition coefficients, $\mathbf{K}_{air:tissues}$ & $\mathbf{P}_{blood:tissues}$ for volatile organic compounds, VOCs, from air to tissues and blood to tissues have been collected for a large number of compounds, measured in two different species, namely humans and rats. The general solvation equations from table 20.0 and 20.1 can be used to correlate the partitions of these VOCs as log **K** or log **P** values at 310K. The equations can be used to screen large numbers of compounds very quickly and to identify out-of-line log **K** or log **P** values that may be a result of experimental error or some specific effect.

It has been found that human and rat partition data are the same within a given experimental error of 0.16 log units (see chapter 9) and that these two sets of data can be combined for all the tissue models discussed earlier in this thesis.

Coefficients e, s, a, b and l versus the tissue composition for air to tissue distribution of VOCs

An attempt was made to compare the coefficients (**e**, **s**, **a**, **b**, **l** and **v**) of the equations (table 20.0 and 20.1) with the composition of the tissues (weight fraction of lipid and water). It was not possible to construct mathematical connections, and so a useful visual comparison of plots (20.0-20.15) between the tissue equation coefficients and the tissue composition was made using the software program, MINIITAB.[1]

AIR: TISSUE	e	s	a	b	l	TISSUE COMPOSITION (WEIGHT FRACTIONS)[2]		
						LIPID	PROTEIN	WATER
OIL[a]	-0.277	0.903	1.698	-0.091	0.876	1.000	0.000	0.000
FAT	0.051	0.728	1.783	0.332	0.743	0.800	0.050	0.150
BRAIN	0.263	0.411	3.358	2.025	0.591	0.107	0.079	0.790
HEART	0.139	0.941	0.778	1.870	0.595	0.100	0.167	0.727
LIVER	0.083	0.768	2.796	2.090	0.560	0.070	0.180	0.720
KIDNEY	0.208	0.382	2.919	2.533	0.651	0.050	0.170	0.770
MUSCLE	0.207	0.723	3.242	2.469	0.463	0.020	0.170	0.790
LUNG	0.534	1.052	3.679	3.078	0.421	0.010	0.177	0.780
BLOOD	0.456	1.083	3.738	2.580	0.376	0.007	0.180	0.800
URINE	-0.689	1.285	4.451	3.288	0.300	-	-	-
PLASMA	0.543	1.677	3.518	3.982	0.192	-	-	-
SALINE	0.653	2.447	4.144	4.247	-0.211	-	-	-
WATER[b]	0.822	2.743	3.904	4.814	-0.213	0.000	0.000	1.000

[a]Olive oil

[b]Water equation at 298K

Table 20.0. Summary of equations for air to tissue (or phase) distribution, and the tissue composition (lipid, protein and water) in humans.[2]

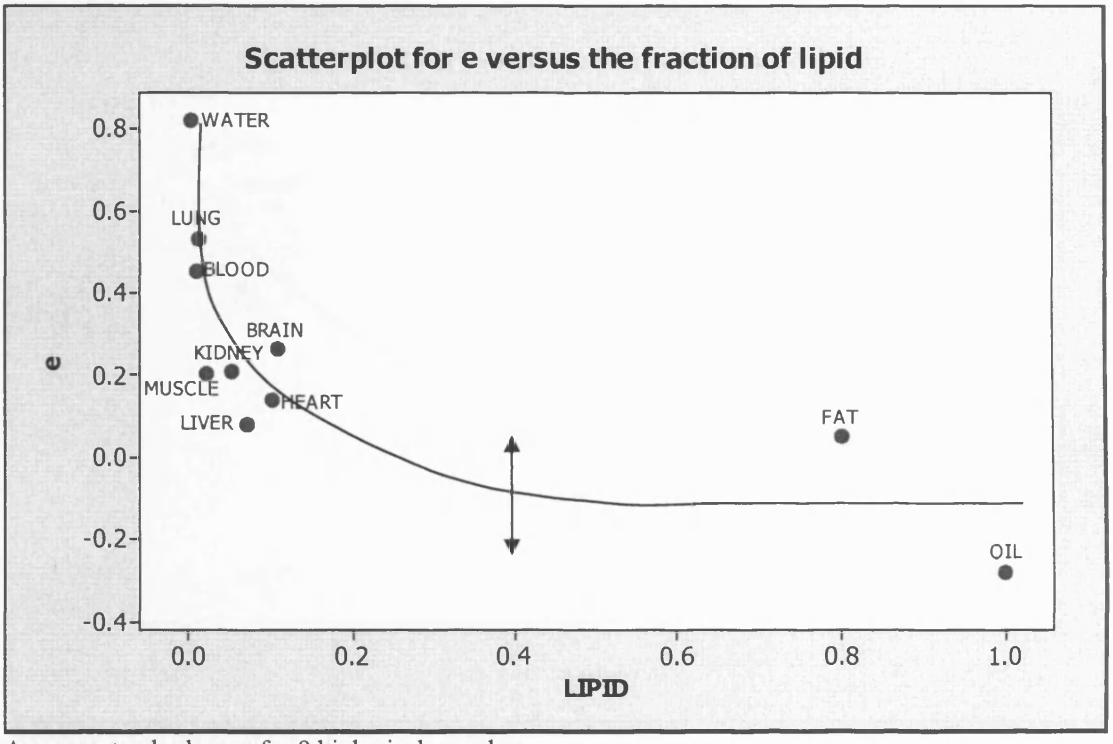
Due to small number of data found in the literature, the new equations have not been mentioned for air-lung and air-heart distribution. The new equations are not different to the current published produced by Weathersby and Abraham.[3] The amount of data in the literature for air-lung and air-heart distribution was very small, being only for 43 and 26 compounds respectively. Hence these systems were not reported earlier in this work. The equations are given here for the completeness (see table 20.0).

When the lipid fraction is increased in tissues (lung, blood, muscle, kidney, brain, heart, liver and fat) or solvents (water and olive oil), the '**e**' coefficient for the equations appears to get smaller. This indicates that the London dispersion interaction is smaller for partition into lipids than into the more polar water and protein constituents. This is shown by the graph in scheme 20.0.

The **s** coefficient also becomes smaller when the lipid fraction is increased. Then dipole/dipole and dipole/induced dipole interactions are also smaller for partition into lipid than into the more polar water and protein constituents. This is displayed by the graph shown in scheme 20.1.

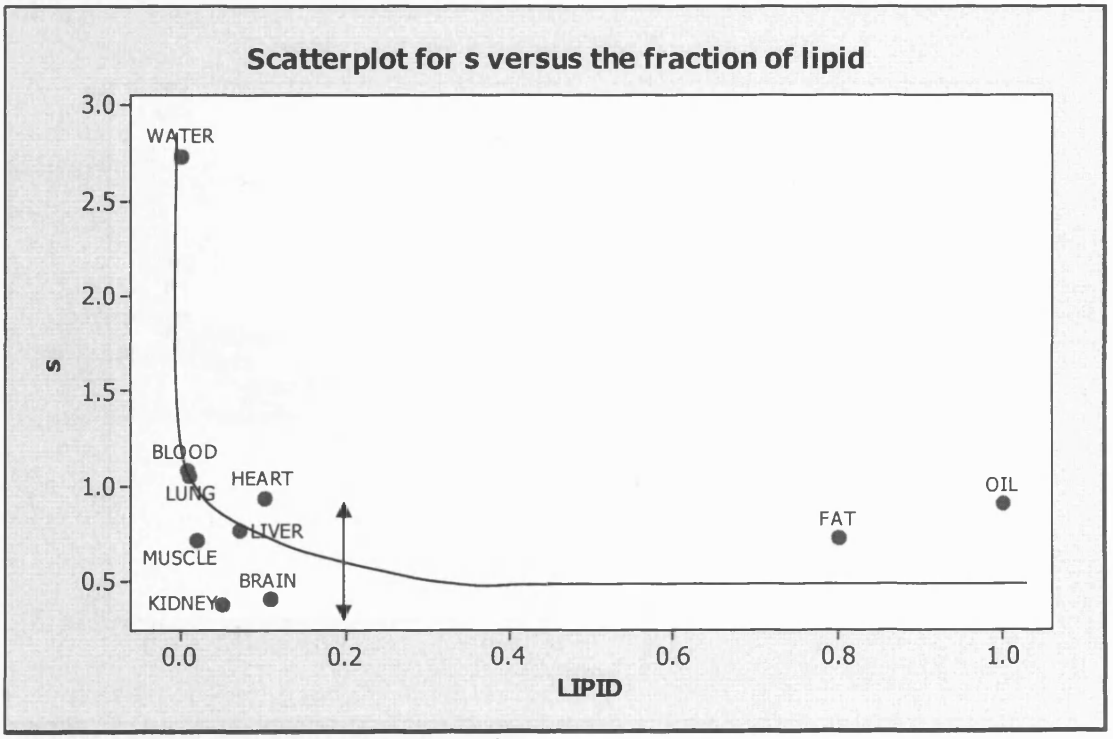
When the lipid fraction is increased in tissues or solvents, the '**a** and **b**' coefficients for the equations appear to get smaller. Again, the more polar constituents of the tissues, hydrogen bond better to VOCs that are themselves hydrogen bond acids or bases, see scheme 20.2 and 20.3. Note that the coefficients for the heart equation in scheme 20.2. do not follow this trend. This could be because the dataset does not represent the true chemical space for that particular tissue. This problem is discussed again later, with the blood to heart distribution equation. Note also that the **a**-coefficient in an equation will reflect an interaction between a hydrogen bond acidic VOC and a hydrogen bond basic tissue, so that a large **a**-coefficient means that the tissue is a strong hydrogen bond base.

Finally, as the lipid fraction is increased in tissues or solvents, the **l**-coefficient for the equation increases and seems to reach a plateau. The **l**-coefficient is a measure of the 'lipophilicity' of a phase. Thus for partition from the gas phase into hexadecane, **l** = 1, for partition into many organic solvents **l** is around 0.7 to 0.9, and for partition into saline, **l** = -0.21 (the only solvent phase for which **l** is negative). The trend shown in scheme 20.4 is exactly as expected, although the numerical value for **l** in water (at 298K) is also negative (**l** = -0.21), this data is the same for that of saline at 310K.



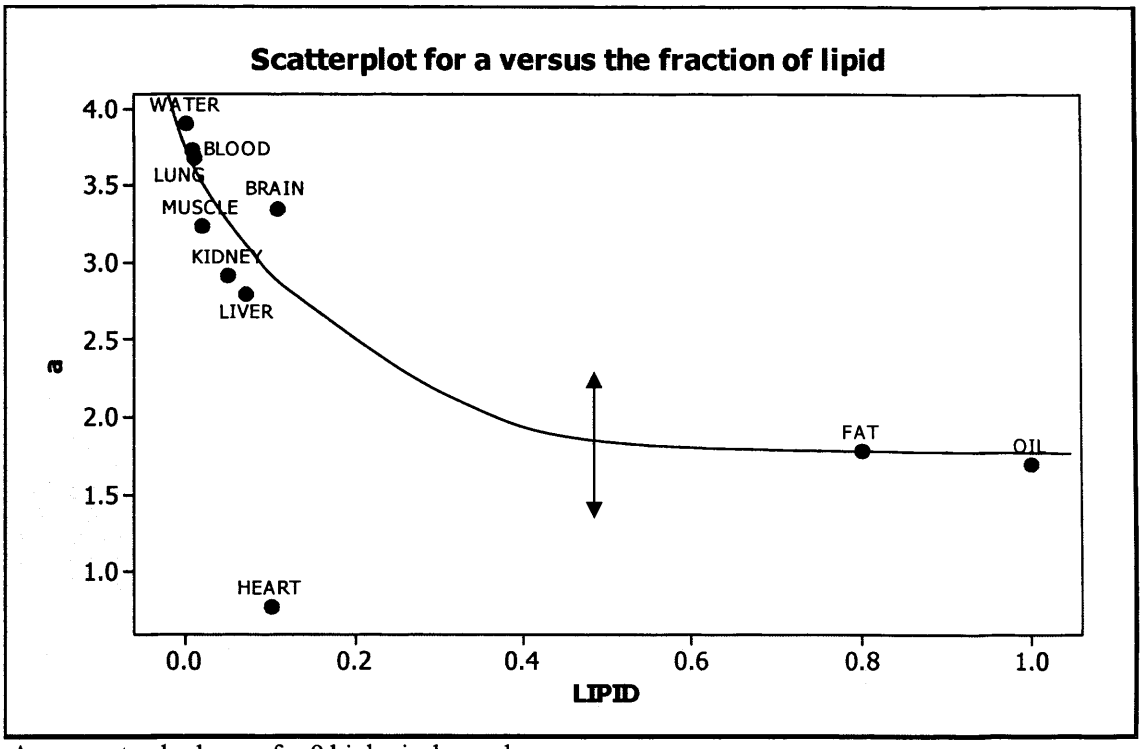
Average standard error for 9 biological samples

Scheme 20.0. Scatter plot for **e** coefficient versus the fraction of lipid for VOCs air to tissue distribution



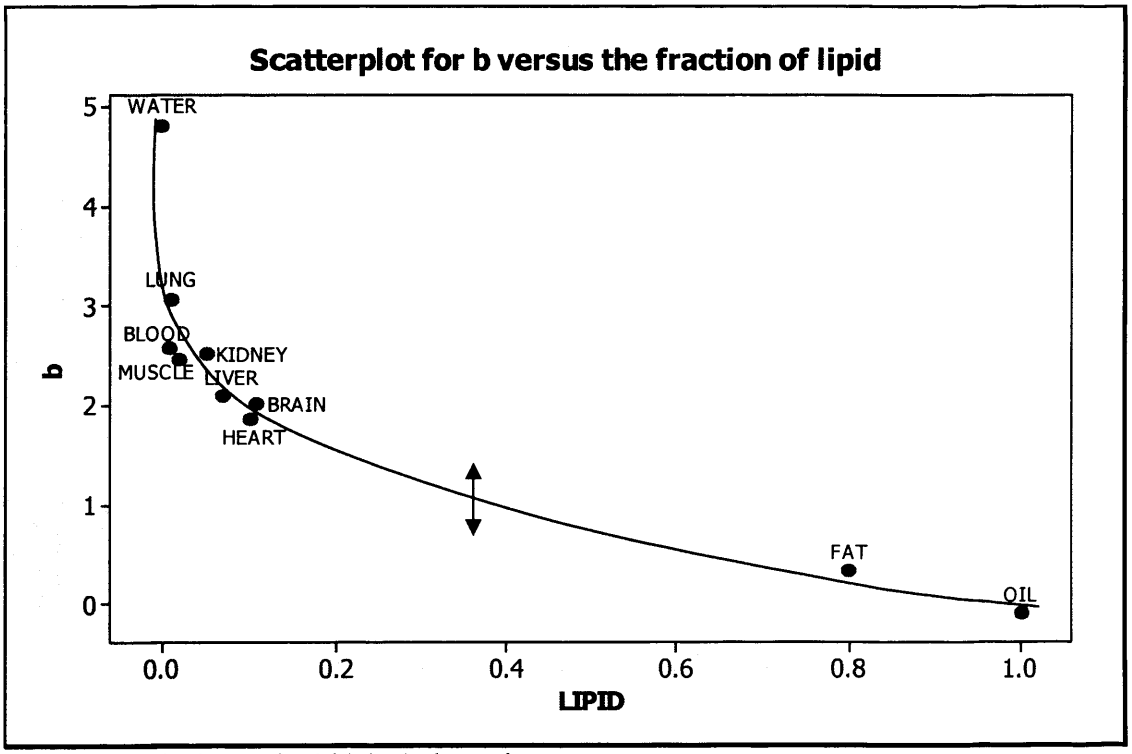
Average standard error for 9 biological samples

Scheme 20.1. Scatter plot for **s** coefficient versus the fraction of lipid for VOCs air to tissue distribution



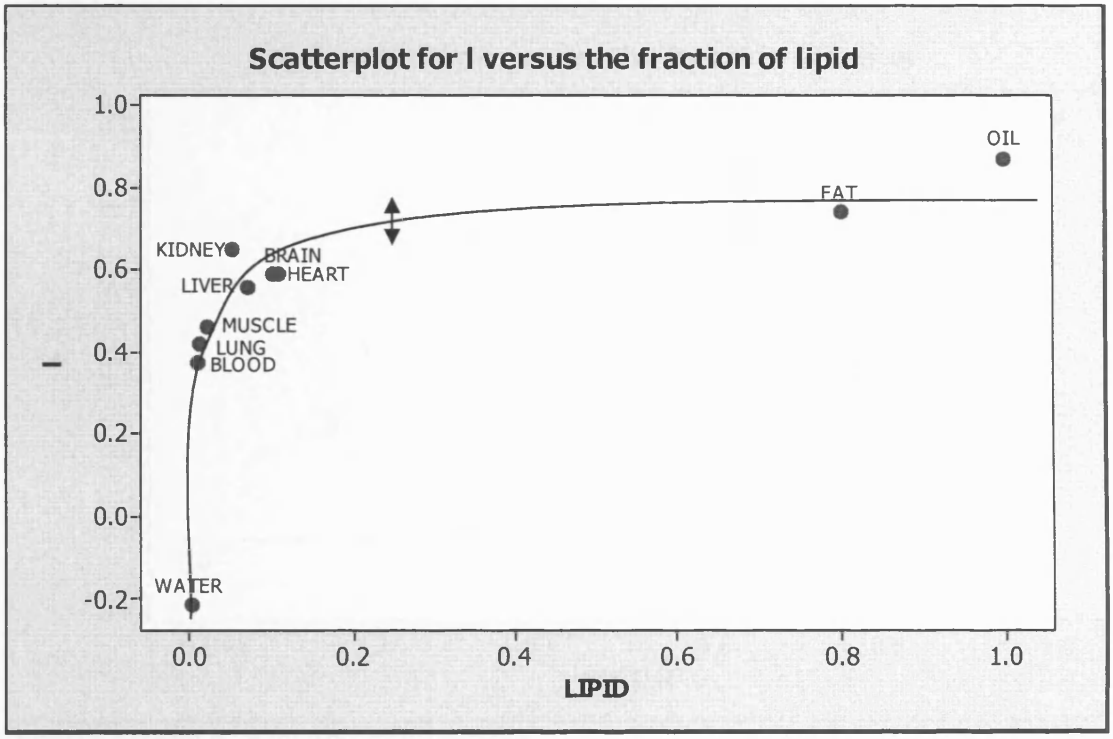
Average standard error for 9 biological samples

Scheme 20.2. Scatter plot for **a** coefficient versus the fraction of lipid for VOCs air to tissue distribution



Average standard error for 9 biological samples

Scheme 20.3. Scatter plot for **b** coefficient versus the fraction of lipid for VOCs air to tissue distribution

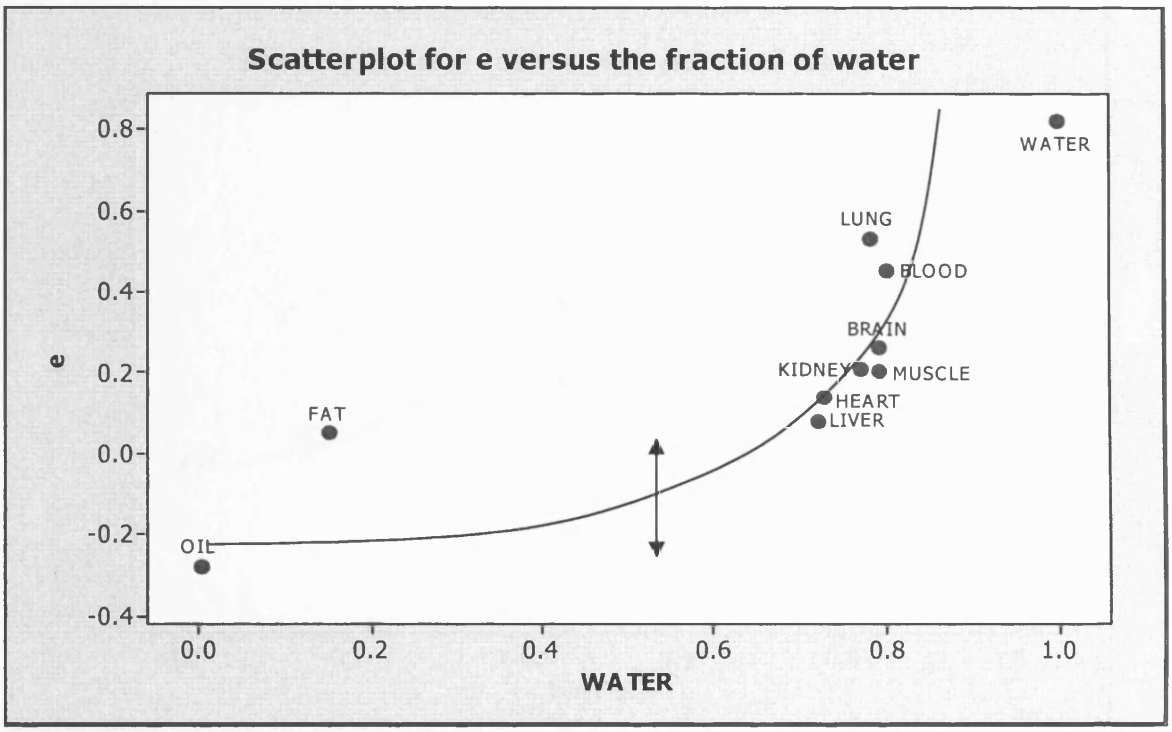


Average standard error for 9 biological samples

Scheme 20.4. Scatter plot for **L** coefficient versus the fraction of lipid for VOCs air to tissue distribution

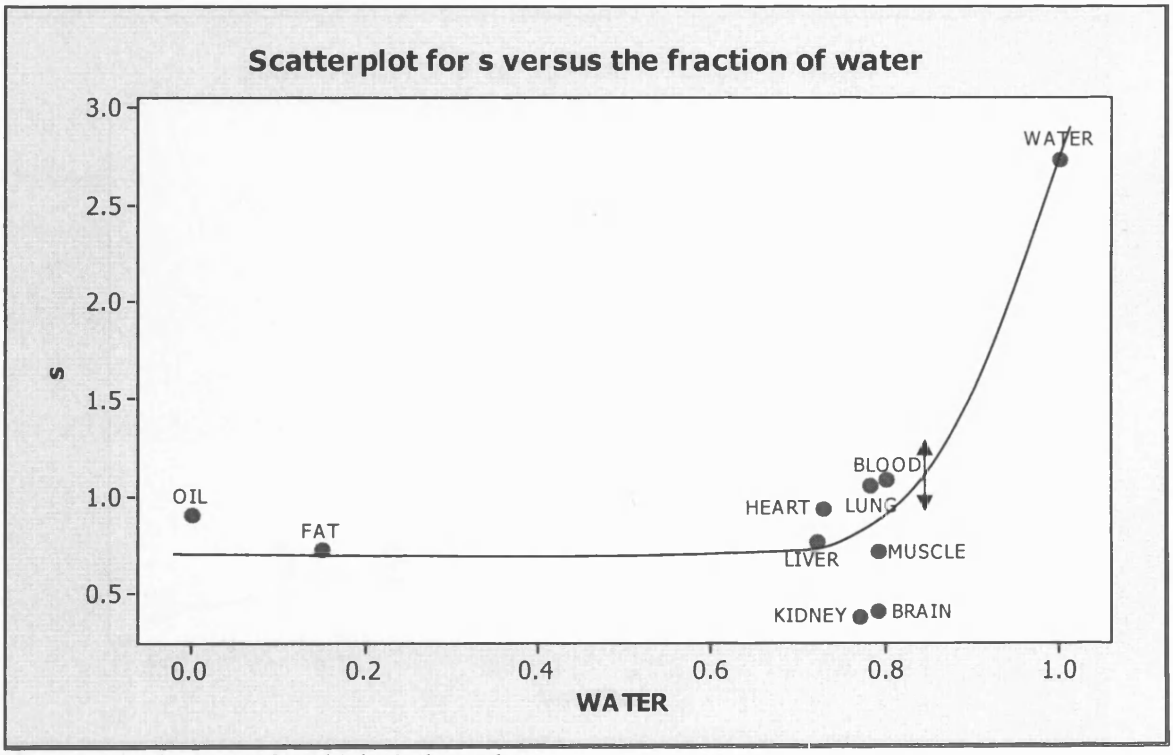
Plots of the coefficients **e**, **s**, **a**, **b** and **l** versus the fraction of water are almost mirror images of the plots versus the fraction of lipid. The explanation is the same as that given above. The **e**, **s**, **a** and **b**-coefficients increase with water fraction because the coefficients refer to polar-type interactions. The **l**-coefficient decreases with water content because it refers to essentially a non-polar type interaction. This can all be seen in the plots shown in schemes 20.5 to 20.9. The **a**-coefficient for the air to heart distribution is anomalous, just as it is for the plot of **a** versus the lipid content, but this is the only very large anomaly.

The conclusion is that the coefficients in the equations for air to tissue distribution are chemically reasonable. They reflect the chemical composition of the tissues, and are not just fitting coefficients in the multiple linear regression equations.



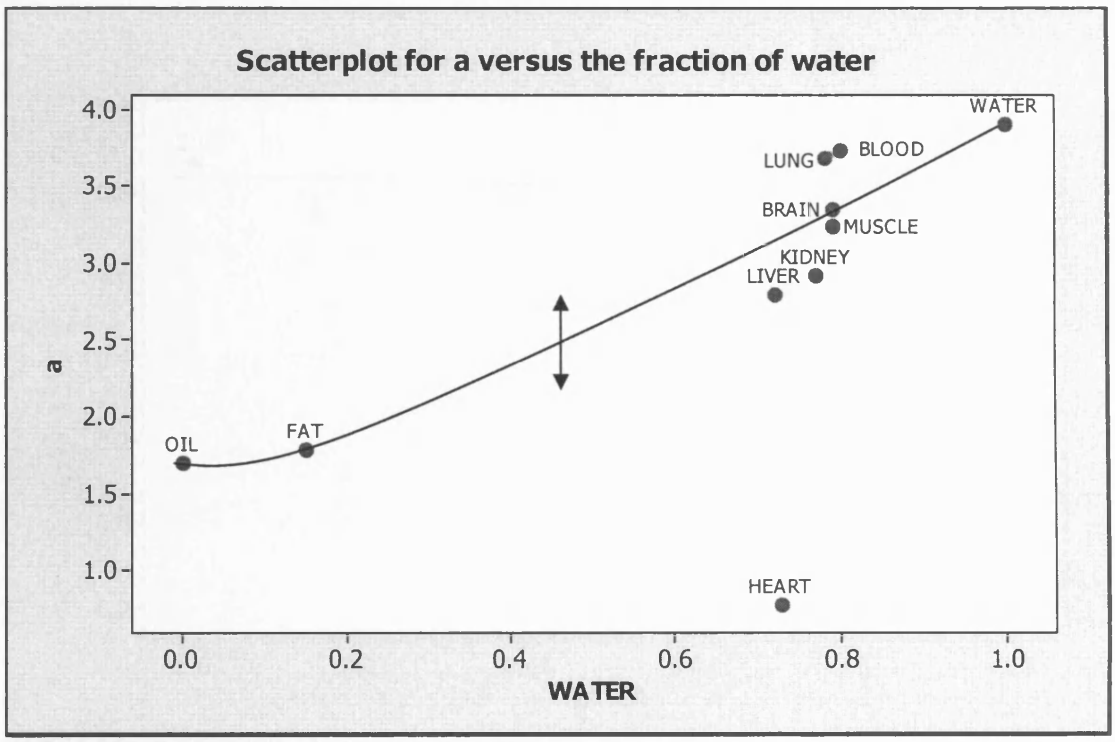
Average standard error for 9 biological samples

Scheme 20.5. Scatter plot for **e** coefficient versus the fraction of water for VOCs air to tissue distribution



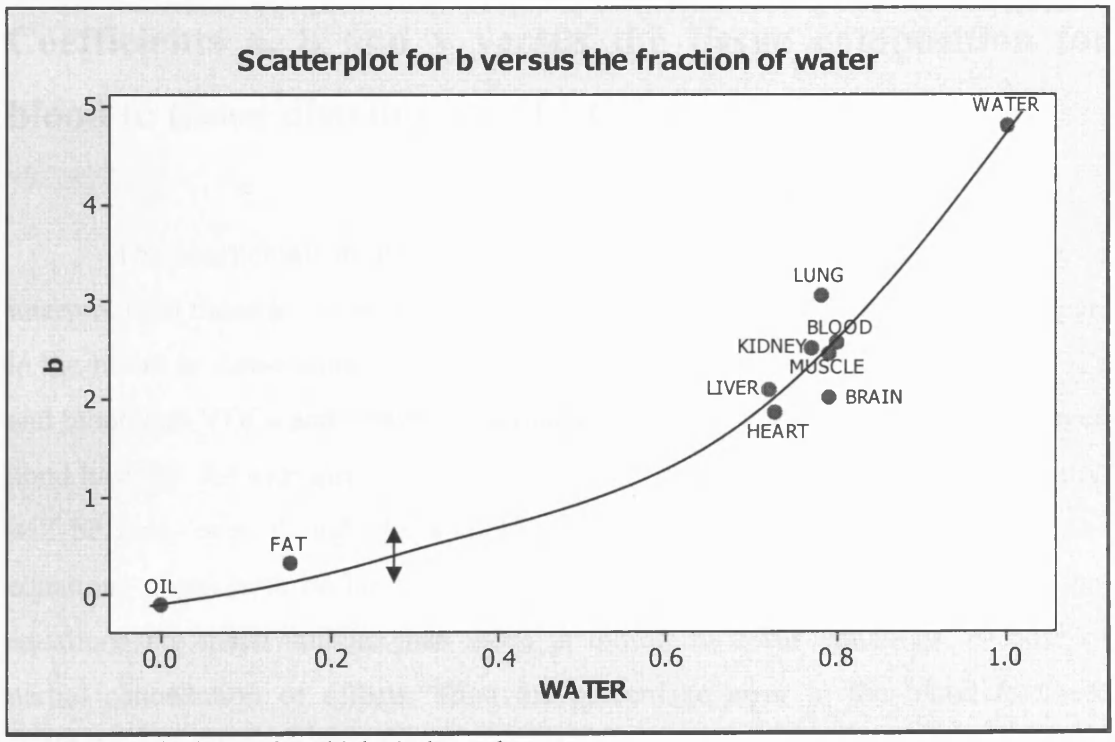
Average standard error for 9 biological samples

Scheme 20.6. Scatter plot for **s** coefficient versus the fraction of water for VOCs air to tissue distribution



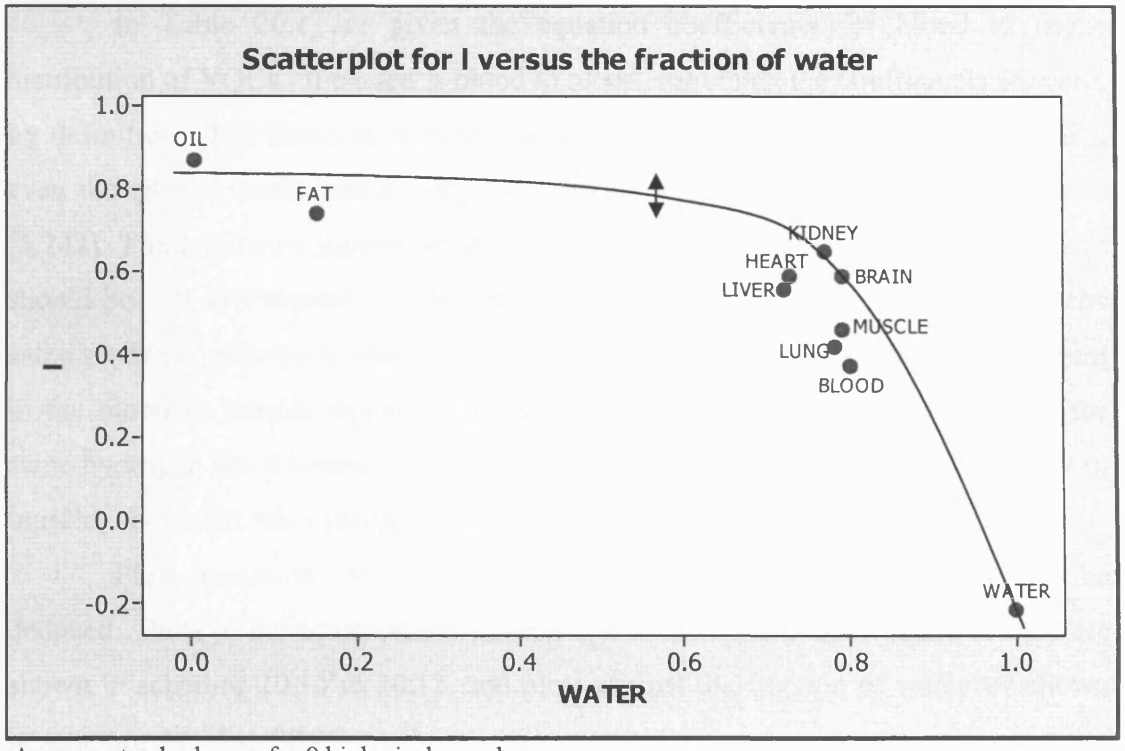
Average standard error for 9 biological samples

Scheme 20.7. Scatter plot for **a** coefficient versus the fraction of water for VOCs air to tissue distribution



Average standard error for 9 biological samples

Scheme 20.8. Scatter plot for **b** coefficient versus the fraction of water for VOCs air to tissue distribution



Average standard error for 9 biological samples

Scheme 20.9. Scatter plot for **L** coefficient versus the fraction of water for VOCs air to tissue distribution

Coefficients a, b and v versus the tissue composition for blood to tissue distribution of VOCs

The coefficients in the equations for blood to tissue distribution are harder to interpret than those in the air to tissue distributions. This is because the coefficients in the blood to tissue equations reflect the difference in interactions between VOCs and blood and VOCs and tissue. If blood and a given tissue have the same hydrogen bond basicity, for example, then the **a**-coefficient for the blood to tissue distribution will be zero, even though the **a**-coefficients in the air to blood and air to tissue equations might both be large. In addition, the coefficients in the blood to tissue equations are much smaller than those in the air to tissue equations, because of partial cancellation of effects. Then the percentage error in the blood to tissue coefficients will be much larger than the percentage error in the air to tissue coefficients. This is another factor that makes interpretation of the blood to tissue equation coefficients difficult.

In Table 20.1 are given the equation coefficients for blood to tissue distribution of VOCs. Included is blood to blood, for which the coefficients are zero, by definition. For blood to muscle, the **a**-coefficient is very small at 0.081 units, even though the coefficient is very large for air to blood (3.738) and air to muscle (3.242). The latter two figures would suggest that the blood to muscle **a**-coefficient should be - 0.496 instead of 0.081, but this will only be the case if the exactly the same set of compounds is used in all three equations. As it is, the small **a**-coefficient in the blood to muscle equation shows that blood and muscle are have about the same hydrogen bond basicity – the coefficients 3.738 (air to blood) and 3.242 (air to muscle) show that this hydrogen bond basicity is very large.

Plots involving the **e**- and **s**-coefficients are so scattered that little can be deduced. Plots of the equation coefficients **a**, **b** and **v** against the fraction of lipid are shown in schemes 20.10 to 20.12, and plots against the fraction of water are shown in schemes 20.13 to 20.15

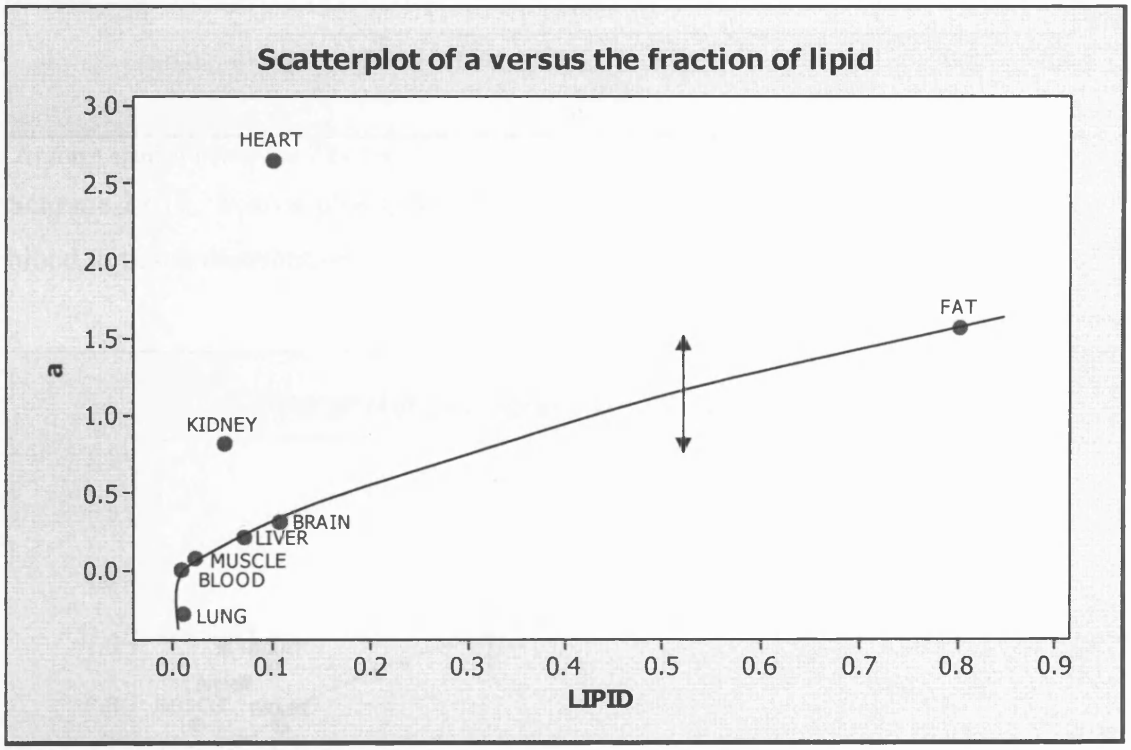
BLOOD: TISSUE	e	s	a	b	v	TISSUE COMPOSITION (WEIGHT FRACTIONS)[2]		
						LIPID	PROTEIN	WATER
FAT	0.017	-0.003	1.577	-2.246	1.560	0.800	0.050	0.150
BRAIN	0.017	-0.536	0.323	-0.335	0.731	0.107	0.079	0.790
HEART	0.033	0.177	2.664	-0.749	0.923	0.100	0.167	0.727
LIVER	-0.048	-0.295	0.213	-0.359	0.876	0.070	0.180	0.720
KIDNEY	0.134	-0.515	0.823	0.273	0.981	0.050	0.170	0.770
MUSCLE	-0.209	-0.593	0.081	-0.168	0.741	0.020	0.170	0.790
LUNG	-0.193	0.004	-0.279	-0.206	0.256	0.010	0.177	0.780
BLOOD	0.000	0.000	0.000	0.000	0.000	0.007	0.180	0.800
PLASMA	-0.227	0.103	-0.268	0.303	-0.201	-	-	-

Table 20.1. Summary of equations for blood to tissue distribution, and the tissue composition (lipid, protein and water) for humans.[2]

The plot of the **a**-coefficients against the lipid fraction is shown in scheme 20.10, and appears to be quite anomalous. It would be expected that as the tissue gets less polar, so the **a**-coefficient should get less (the value for the blood to heart

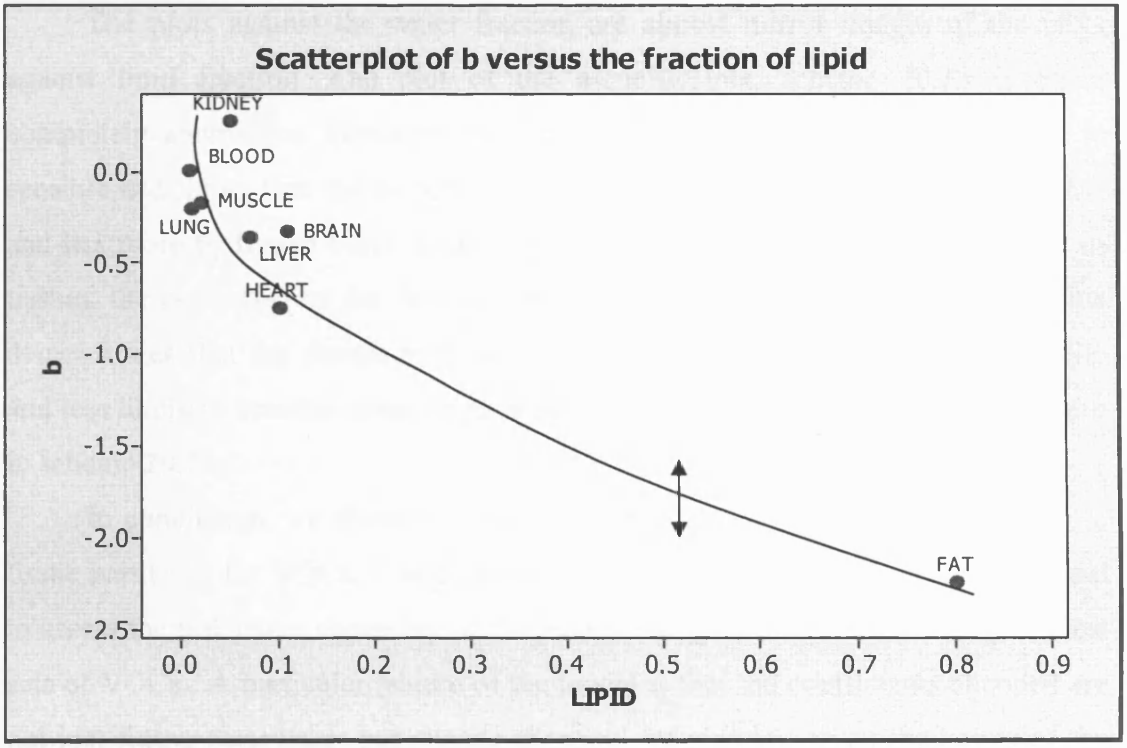
equation is again out of line). However, the **a**-coefficients get larger, for which there is no explanation. The plot of the **a**-coefficients against the fraction of lipid for the air to tissue equations shows that as the tissue gets less polar, the **a**-coefficient gets less, as expected, so the peculiar trend in scheme 20.10 might be due to errors in the coefficients.

If the lipid fraction is increased in tissues, the **b**-coefficient for the equations gets smaller, see scheme 20.11. This is reasonable, because the tissue becomes less polar. and less positive. Finally, as the lipid fraction is increased in tissues, the **v**-coefficient for the equations increases, scheme 20.12. This is also reasonable because the **v**-coefficient is also an indication of lipophilicity of a phase.



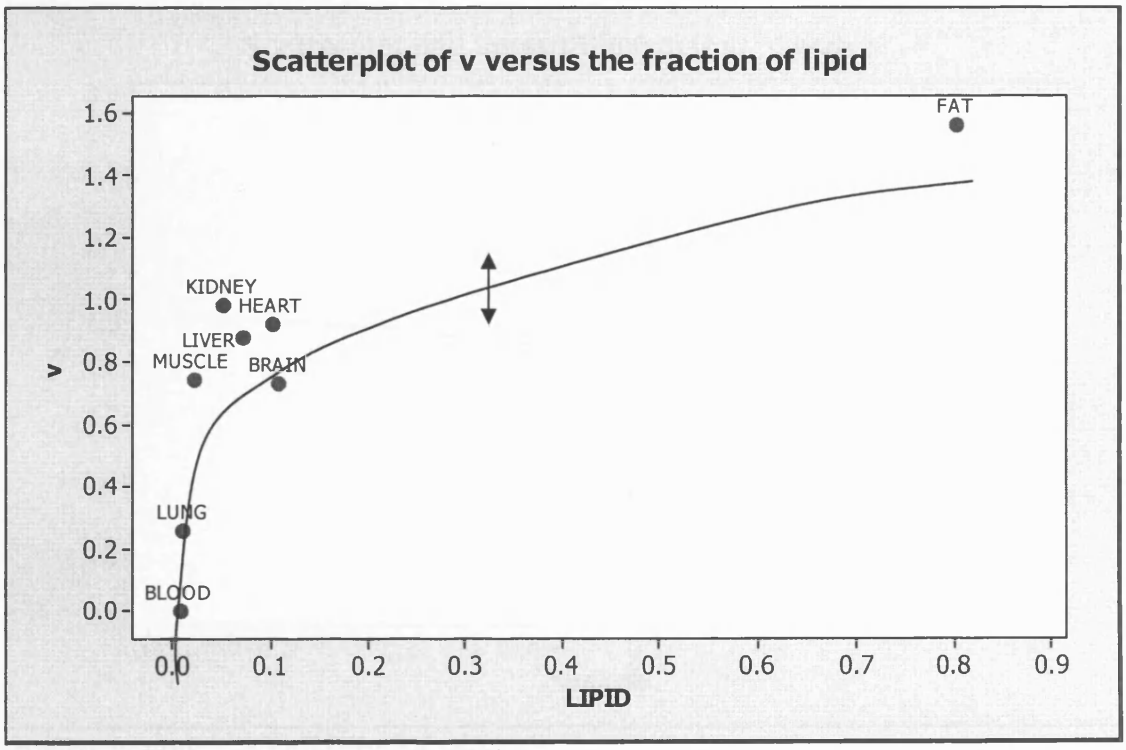
Average standard error for 7 biological samples

Scheme 20.10. Scatter plot for **a** coefficient versus the fraction of lipid for VOCs blood to tissue distribution



Average standard error for 7 biological samples

Scheme 20.11. Scatter plot for **b** coefficient versus the fraction of lipid for VOCs blood to tissue distribution

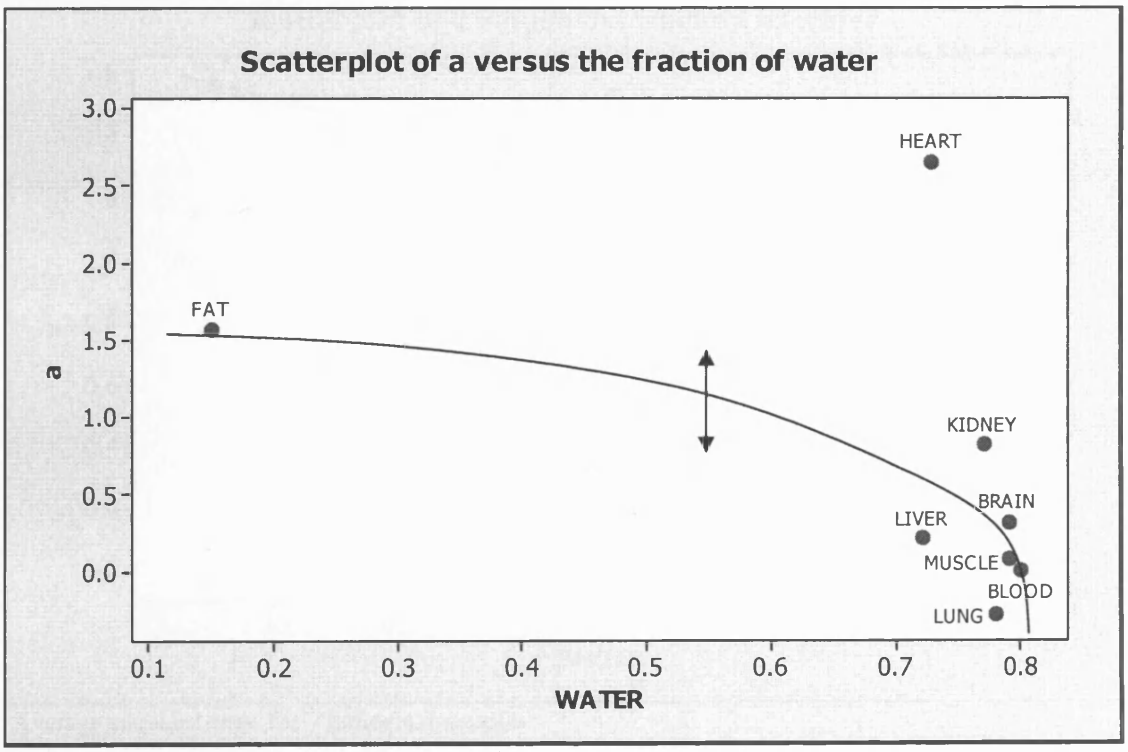


Average standard error for 7 biological samples

Scheme 20.12. Scatter plot for **v** coefficient versus the fraction of lipid for VOCs blood to tissue distribution

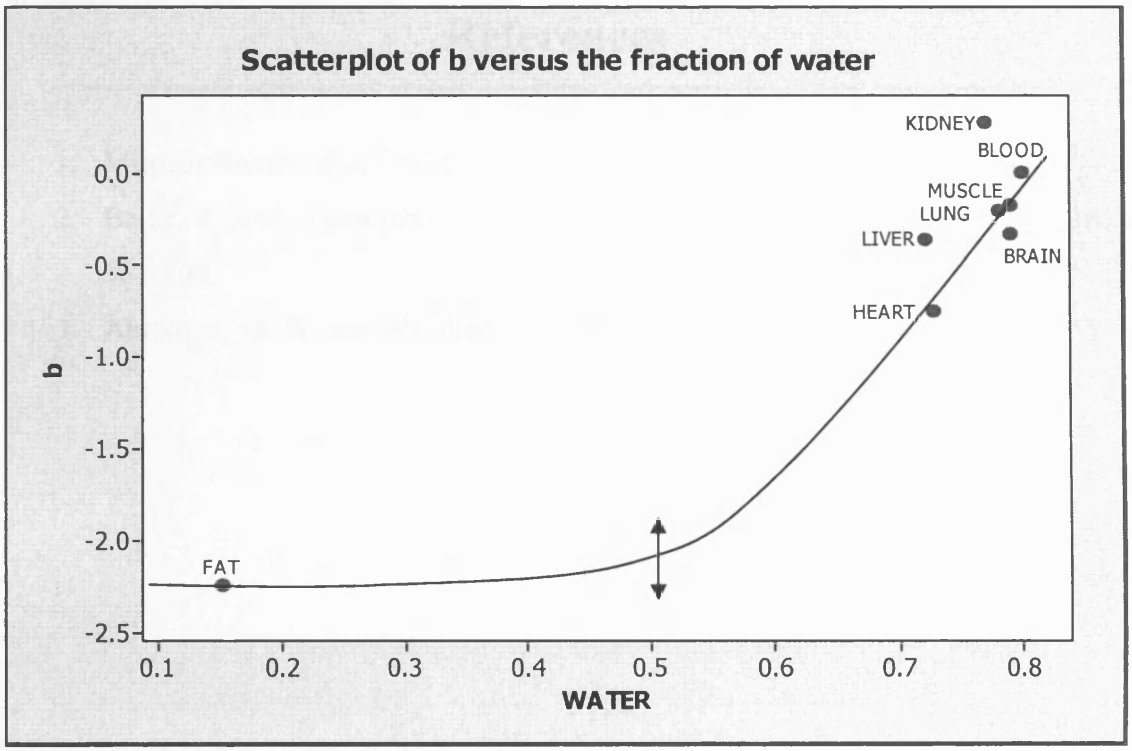
The plots against the water fraction are almost mirror images of the plots against lipid fraction. The plot of the **a**-coefficients, scheme 20.13 is again completely anomalous. However, the plot for the **b**-coefficient, scheme 20.14, is sensible and shows that the **b**-coefficient increases as the tissue becomes more polar, and has more hydrogen bond acidity. Finally, as the water fraction is increased in tissues, the **v-** coefficient for the equations, appears to also gradually decrease. This demonstrates that the tissues with larger amounts water content are less lipophilic and less likely to accommodate larger solutes. This relationship is shown by a graph in scheme 20.15.

In conclusion, we show that for a large data set of air to tissue and blood to tissue partitions for VOCs, it is possible to construct statistically sound models and to assess the predictive capability of the model through selection of training and test sets of VOCs. A particular feature of the model is that the coefficients obtained are not just fitting parameters but encode chemical information about the nature of the process. This enables the air to tissue or blood to tissue distribution process to be compared in terms of the chemical interactions between the VOCs and blood and tissues.



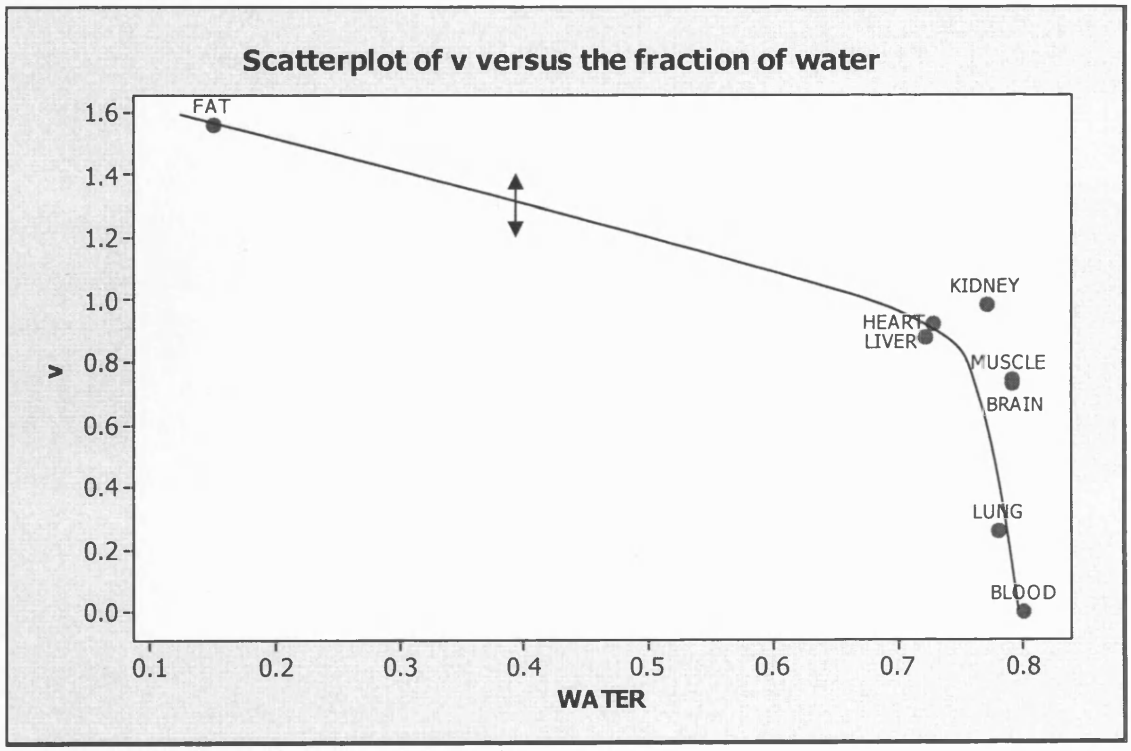
Average standard error for 7 biological samples

Scheme 20.13. Scatter plot for **a** coefficient versus the fraction of water for VOCs blood to tissue distribution



Average standard error for 7 biological samples

Scheme 20.14. Scatter plot for **b** coefficient versus the fraction of water for VOCs blood to tissue distribution



Average standard error for 7 biological samples

Scheme 20.15. Scatter plot for **v** coefficient versus the fraction of water for VOCs blood to tissue distribution

References

1. Minitab statistical software (version 14), www.minitab.com
2. Balaz, S. and Lukacova, V (1999). Quant. Struct.-Act. Relationships, **18**, 361-368.
3. Abraham, M. H. and Weathersby, P. K (1994). J. Pharm. Sci, **83**, 1450-1455.

Chapter

21

Blood/plasma/serum to brain distribution for drugs and VOCs combined

In this chapter the Abraham solvation equation shall be used for models that include: blood/plasma/serum to brain distribution for drugs and blood/plasma/serum to brain distribution for VOCs and drugs combined

Blood/plasma/serum to brain distribution for drugs

Brain uptake, or the ability of a molecule to enter the brain tissue, has been a subject of great interest to the pharmaceutical industry for over 35 years. There are two quite different methods used to obtain blood to brain distribution coefficients, denoted as BB values. In the first method, gas to blood and gas to brain partition coefficients are determined separately, using blood and brain samples from either humans or rats. Then the two partition coefficients can be combined to yield *in vitro* BB values. This method has invariably been used to obtain BB values for volatile organic compounds and some volatile inorganic compounds; the term VOC is used to cover both of these and has already been discussed in chapter 12 of this thesis.

The second method of obtaining blood/plasma/serum to brain distribution data for drugs in rats is by having the drug administered to the rat and the latter then being sacrificed after a given period of time. The drug concentration in blood/plasma/serum and in brain gives the blood/plasma/serum to brain ratio. The blood to brain distribution data (*in vivo*) for steady state is obtained by the following equation (21.0).

log BB = (conc. of drug in the brain tissue)/(conc. of drug in the blood/serum/plasma) Equation 21.0

Literature values of the *in vivo* distribution of drugs from blood/plasma/serum to rat brain have been assembled for 120 compounds (152 data points), by far the largest data set ever reported.[1-34] Several cited papers, where models have been suggested to predict blood to brain distribution in rats will be discussed. The purpose of this work is to add new literature data to the existing data sets and to make an attempt to construct new equations to predict blood to brain distribution. The complete list of drug log BB data is shown in table 21.0 (see appendix); where two or more results are reported for the same compound in the same system (blood/plasma/serum) they have been averaged.

The usual method of obtaining blood to brain distribution data is that of Young and Mitchell.[11] A number of rats are sacrificed at given times and the blood/plasma/serum to brain ratios are examined to see if they reach a constant value. Quite often, data at only one or two time points are obtained, and the assumption is made that equilibrium has been reached. The distribution of drugs from the blood to the brain is of importance in pharmaceuticals and as well as toxicological sciences.

Unfortunately, the experimental determination of the blood/plasma/serum to brain distribution is time consuming, expensive, requiring animal experiments, synthesis of compounds (often in a radiolabeled form) and is a difficult technique to carry out. It would therefore be of very considerable value to devise a methodology whereby, log BB values could be predicted from other experimental results (such as other partition coefficients) that are easier to obtain than log BB values themselves, or even whereby log BB could be predicted from molecular structure.

Literature models for the correlation of blood to brain distribution for drugs have been reported by the following authors: Chadha *et al.*,[35] Platts *et al.*,[36-37] Norinder *et al.*,[38] Lombardo *et al.*,[39] Clark,[40] Kaliszan and Markuszewski,[41] Luco,[42] Keseru and Molnar,[43] Hou and Xu,[44-45] Ooms *et al.*,[46] Liu *et al.*,[47] Cabrera *et al.*,[21] Kaznessis *et al.*,[48] Feher *et al.*,[49] Kelder *et al.*,[22] Ertl *et al.*,[50] Salminen *et al.*,[13] Hutter,[51] Rose *et al.*[52] and Lobell *et al.*[53]

Chadha *et al.*[35] modelled the passive distribution of drugs from the blood and into the brain by using the Abraham equation. Their data set of 57 compounds was

used to correlate these distribution processes (equation 21.1). Here and elsewhere, N is the number of compounds, r^2 is the square of the correlation coefficient, SD is the standard deviation and F is the Fisher statistics.

FULL SET

log BB = -0.038 + 0.198**E** - 0.687**S** - 0.715**A** - 0.698**B** + 0.995**V** Equation 21.1

n = 57 $\mathbf{r^2}$ = 0.92 **SD** = 0.20 **F** = 99

NO TEST SETS

Although the model contained eight outliers (SKB3, SKB4, SKB6, SKB9, SKB13, SKB24, SKB25 and SKB45) the fits for the correlative equation were very good. It was very unfortunate to find that there were no test sets for this model, since it would have been very useful to see how good the predictions are for this blood to brain model.[35]

The method used to model the equilibrium distribution of drugs between blood and brain was developed by Platts *et al.*[36] They used the Abraham equation to predict and correlate these passive distribution processes for 148 drug compounds. The five descriptors (**E**, **S**, **A**, **B** and **V**) plus an indicator variable (**Ia**) were used in a multiple linear regression analysis. Platts *et al.*[36] also used a training set of 74 compounds to predict the remaining log BB values in the test set with an SD value of 0.38 log units. They used an indicator variable, **Ia**, for carboxylic acids such that **Ia** = 1 for carboxylic acids, and **Ia** = 0 for all other compounds. The equation for the total set of 148 compounds found by Platts *et al.* is given as equation 21.2.

FULL SET

log BB = 0.062 + 0.469**E** - 0.864**S** - 0.586**A** - 0.713**B** + 0.895**V** Equation 21.2

- 0.564**Ia**

N = 148 r^2 = 0.74 SD = 0.34 F = 83

TRAINING SET

N = 74 r^2 = 0.76 SD = 0.34

TEST SET

N = 74 SD = 0.38

The indicator variable used by Platts *et al.*, shows that carboxylic acids are less likely to distribute into the brain than expected from the usual descriptors. Nine outliers (SKB4, SKB9, SKB20, SKB21, YG-19, YG-20, thioridazine, org12962 and fluphenazine) were removed from model, as they appeared to have large residuals. Platt suggested that the large residuals resulted from active transport, metabolism or error in the experimental measurements. The advantage of using this model, is that the descriptors for the unknown compounds can be calculated quickly, by using a commercially available program based on various fragment schemes.[37]

Norinder *et al.*[38] used a quantitative structure-activity property analysis based on MolSurf parameterisation with the Partial Least Squares (PLS) method as a statistical engine. The computed properties of the solutes were calculated by using MolSurf (computer program) and this was used to create the training equation for blood to brain distribution for drugs. Hence all 14 descriptors can be calculated with available software. Norinder *et al.* divided the data set of Lombardo *et al.*[39] into two parts: a training set (56 compounds) and a test set (6 compounds). The statistics for the PLS analysis, are shown below. The equations are not shown, as they absent from the author's paper:

TRAINING SET

N = 56 $r^2 = 0.83$ SD = 0.31 F = 87

TEST SET

N = 6 RMSE = 0.51

The statistics were good, although the data set was not as large as that of Platts *et al.*,[36] and the test set is too small to afford a reliable predictive assessment. It would be useful to see if the training equation can be used to predict larger test sets for future work. This method is complicated and time consuming, as computational methods are required for the conformational analysis of the solutes, via the intensive semi-empirical and ab-intio calculations which can slow down the screening process. Norinder *et al.*[38] also found that the most important factors that affected the models were associated with polarity and base strength. They also found that the properties associated with hydrogen bonding were important and should be kept to a minimum to facilitate high blood to brain distribution. High lipophilicity (log P_{oct}) was also identified to be favourable for high partitioning. Furthermore, the presence of

polarisable electrons, e.g. conjugated and aromatic substructures as well as the larger halogens, is also beneficial for the compound to penetrate the brain effectively.[38]

Clark used a simple QSAR (quantitative structure-activity relationship) model for the prediction of log BB from a set of 55 diverse organic compounds.[40] His data set was also obtained from Lombardo *et al.*[39] Although, this data set is not as large as the Platts *et al* BB model,[36] it can be used effectively to carry out rapid predictions of blood to brain distribution. The model contains two important variables: polar surface area (PSA) and calculated log P_{oct}, both of which can be rapidly computed. The two descriptors (PSA and CLOG P) were used in a multiple linear regression analysis and the equations (21.3) are shown below. Note that two outliers were observed by Clark; they are: YG-19 and YG-20.

TRAINING SET

LogBB=-0.0148(±0.001)PSA + 0.152(±0.036)ClogP + 0.139(±0.073) Equation 21.3

N = 55 r^2 = 0.79 SD = 0.35 F = 96

TEST SET 1 AND 2

1. N = 5 and MAE = 0.14 2. N = 5 and MAE = 0.24

The two test sets shows that there are good predictions, although, the compound numbers are still quite small. It would interesting to see again if larger test sets give more or less the same predictions.[40]

Kaliszan and Markuszewski[41] initially used 20 compounds in their blood to brain data set. They used two parameters for the equation and this appeared to be one of the simplest forms of models to date, for blood to brain distribution for drugs. In the first correlative equation, they used square root of the molecular mass (M_m) and log P_{oct} (octanol-water) as the main parameters, this is shown below as equation 21.4. The fits for the second equation (21.5) were improved by replacing log P_{oct} with log $P_{cyclohexane}$ (cyclohexane-water). The third equation (21.6) lead to better fits with the replacement of molecular mass (M_m), with the new volume parameter for the solute (V). This V parameter was used to represent the bulkiness of the solute. Again, the fits are found to be little better with this new parameter than found from equations 21.4 to 21.5. Finally, the parameter for equation 21.7 was reverted back to the parameters of equation 21.4. This was done to see if the fits of the equation would improve with larger data sets for the 33 blood to brain drug compounds. The

fits for this equation improved with these larger data sets. It is very unfortunate that the data set for equation 21.4-21.7 is small, as it would be interesting to see how the correlations and predictions would be like with larger number of compounds (or data points), such as used in Platt's model.[36]

FULL SET [1]

$LogBB = -0.476(\pm 0.454) + 0.541(\pm 0.106)logP_{oct} - 0.008(\pm 0.002)M_m$ Equation 21.4

$N = 20 \quad r^2 = 0.64 \quad SE = 0.49 \quad F = 15$

FULL SET [2]

$LogBB = 1.296(\pm 0.313) + 0.309(\pm 0.034)logP_{cyh} - 0.006(\pm 0.001)M_m$ Equation 21.5

$N = 20 \quad r^2 = 0.84 \quad SE = 0.32 \quad F = 46$

FULL SET [3]

$LogBB = 1.979(\pm 0.339) + 0.373(\pm 0.032)logP_{cyh} - 0.003(\pm 0.001)V$ Equation 21.6

$N = 20 \quad r^2 = 0.89 \quad SE = 0.27 \quad F = 68$

FULL SET [4]

$LogBB = -0.088(\pm 0.051) + 0.272(\pm 0.017)logP_{oct} - 0.001(\pm 0.000)M_m$ Equation 21.7

$N = 33 \quad r^2 = 0.90 \quad SE = 0.13 \quad F = 131$

Luco used the partial least squares method to model the blood to brain distribution for 61 structurally diverse compounds.[42] The PLS model was based on molecular descriptors that can be calculated for any compound simply from knowledge of its molecular structure. This provides information to polarity (hydrogen bond acceptors and donors), molecular size and the molecular shape. The statistics for the training equation were good, after removing three outliers form the model. The fits for the training set are shown below:

TRAINING SET

$N = 58 \quad r^2 = 0.85 \quad S = 0.32 \quad F = 102$

TEST SET [1]

$N = 12 \quad RMSE = 0.24$

TEST SET [2]

$N = 25 \quad RMSE = 0.54$

The test set (1) containing 12 other simple drug compounds, that were not used in the training set was predicted quite well. Whereas, the larger test set (2) of

25 large drug compounds was rather poorly predicted. The reason is that these compounds are much larger and more complicated in structure, so that the training model may not predict compounds in descriptor space very well.

Lombardo and co-workers took a rather different approach in order to analyse the 57 drug compounds of the Abraham data set.[35] The free energy of solvation was computed with the AMSOL 5.0 computer program using the AM1-SM2.1 solvation model. Lombardo *et al.*[39] also identified two outliers and obtained the following model as shown by equation 21.8. The correlation established by this model for blood to brain distribution for drugs is not so impressive and the subsequent predictions of some additional six compounds as an external test set confirm this (SD =0.62).

TRAINING SET

$LogBB = 0.054(\pm 0.001)\Delta G_w^{o} + 0.43(\pm 0.070)$ Equation 21.8

$N = 55 \quad r^2 = 0.67 \quad SE = 0.41 \quad F = 108$

TEST SET

$N = 6 \quad SD = 0.62$

They also computed the free energy of solvation in hexadecane (ΔG_{HEX}^{o}) in the hope that this term might be more representative of the environment that the compounds would encounter while crossing the membrane of the brain. However, the inclusion of this term did not lead to a significant improvement in the correlation. Finally, although there is a correlation between the experimental and calculated values for the blood to brain distribution for drugs, the semi-empirical calculation required for each molecule made this approach very-time consuming.

Keseru and Molnar used the thermodynamic approach of the solvation of free energies.[43] They used only one descriptor in their model – the standard free energy of hydration of the gaseous compound, ΔG_{hyd}, in kJ mol^{-1} Their training set involved a structurally diverse set of 55 compounds and the fits for the model can be found from equation 21.9.[43] The prediction for the test sets appeared to be better than the equation produced by Lombardo.[39]

TRAINING SET

$LogBB = 0.035\Delta G_{hyd}^{o} + 0.2592$ Equation 21.9

N = 55 r^2 = 0.72 SE = 0.37

TEST SET

N = 5, 25 SD = 0.14, 0.37

Hou and Xu[44-45] developed a blood to brain distribution model for drugs that is based on three descriptors, which are: octanol/water partition coefficient calculated using the SLOG approach; the high charged polar surface area (HCPSA) and a molecular weight term that takes the value (MW–360) for compounds with molecular weight greater than 360 and zero for other compounds. A combination of these properties gives rise to the following equation, 21.10, below:

TRAINING SET

Log BB = 0.00845 + 0.197SLOGP – 0.0135HCPSA – 0.0140 (MW –360) Equation 21.10

N = 72 r^2 = 0.77 SD = 0.36 F = 82

TEST SET

N = 35 SD = 0.41

The prediction of the model was assessed through two external test sets of a total of 35 drug compounds. Note the two test sets were predicted through equation 21.10 as shown above.[44-45]

Following the success of equation 21.10, Hou and Xu[44-45] re-parameterised the calculation of SLOGP, and replaced HCPSA with the more easily calculated high-charged topological polar surface area, HCTPSA, and derived the equation shown below, 21.11. The fits for this training and test set appeared little worse than that of equation 21.10., but equation, 21.10 can be used for high throughput screening, as the descriptors can be calculated quicker.

TRAINING SET

Log BB = 0.1256 + 0.160SLOGP – 0.0133HCTPSA – 0.0148(MW – 360) Equation 21.11

N = 78 r^2 = 0.74 SD = 0.38 F = 67

TEST SET

N = 35 SD = 0.44

Ooms *et al.*[46] developed a model to correlate blood to brain distribution for 79 drug compounds. Their method involved the minimisation of the 3D structure of the drugs. Their descriptors were automatically calculated using a computer program called VolSurf. They obtained 31 useful descriptors and these refer to: molecular size, hydrogen bonding and the shape of hydrophilic and hydrophobic regions within the solute. The statistical analysis was performed with principle component analysis (PCA). Note that no test sets were used. The statistics for the model are shown below but no SD value was given.

FULL SET
N = 79 $r^2 = 0.76$ SD = - F = -
NO TEST SET

Liu *et al.*[47] constructed equation (21.12) for blood to brain distribution for drugs using multiple linear regression. This was based on two important descriptors that are: lipoaffinity (LA) and molecular weight (MW). Their model was based on 55 compounds (that included VOCs and drugs) and this was used to predict a test set of 11 solutes. They found that the lipoaffinity descriptor is the most significant descriptor for molecular transport of drugs across the blood to brain barrier. Finally, four outliers (comp-12, chlorambucil, trifluoropromazine and org-12962) were removed from model, as they appeared to have large residuals.

TRAINING SET
Log BB = 0.138(0.104) − 0.0112(0.0008)MW + 0.364(0.033)LA Equation 21.12
N = 55 $r^2 = 0.79$ SD = 0.35 F = 98
TEST SET
N = 11 SD = 0.30

Note that these statistical parameters are nearly identical to Clark's model (equation 21.3) based on polar surface area and CLOGP. Finally, four outliers were removed from Liu's model and they are: nitrogen, compound 12, YG19 and YG20.

Cabrera *et al.*[21] used only three descriptors in their equation for the correlation and the prediction of blood to brain distribution for drugs. Five outliers

(chlorambucil, temelastine, tiotidine, YG-20 and indomethacin) were omitted from the model and large residuals possibly resulted from active transport or metabolism.

TRAINING SET

Log BB = -0.032 − 0.046(exp-3)PS * AM + 0.227H Equation 21.13

N = 114 r^2 = 0.70 SD = 0.42 F = 128

TEST SET

N = 28 MAE = 0.33

They use a simple notation for the descriptors: PS is related to the polar surface, AM is the atomic mass (which appears to be the mass of atoms in a molecule excluding hydrogen atoms), and H is related to hydrophobicity. The descriptors were calculated by an in-house program, called TOPS-MODE (topological sub-structural molecular design approach). They have found that hydrophobicity of the drug compound has a positive effect with blood to brain distribution, whilst the polar surface and atomic mass has a negative effect on this distribution. Although the equation, 21.13, is very good and so is the predictive power, with MAE = 0.33 for a test set, the method will not be very easy to carry out.

The results of Kaznessis *et al.*[48] are also noteworthy. The best model, with SD = 0.17 and r^2 = 0.94 for a training set seems to be over fitted, because the SD value is considerably smaller than the general error of the measurements. This is indicated by a much larger SD value of 0.48 log units for a very small test set (4 compounds). Kaznessis *et al.*[48] also used another model with only three descriptors that yielded SD = 0.26 and r^2 = 0.87; however, both of these models require very heavy computing time. Nine outliers (icotidine, temlastine, tiotidine, Org12962, YG-19, YG-20 and compound 2, 4 and 13) were omitted from the model. The author suggested that these large residuals resulted from active transport, metabolism or error in the experimental measurements.

Of calculations mentioned earlier by other authors, the Feher *et al.*[49] method for predicting blood to brain distribution for drugs is very simple indeed. Only three descriptors are used in this model, which are: a calculated log P_{oct} value, the polar surface area and the number of hydrogen bond acceptors in an aqueous medium, n_a. For a training set of 61 compounds, Feher *et al* [49] obtained the following equation (21.14) shown below:

TRAINING SET

$\text{Log BB} = 0.4275 - 0.3873n_a + 0.1092 \log P_{oct}(\text{calc}) - 0.0017\text{PSA}$ Equation 21.14

$N = 61 \quad r^2 = 0.73 \quad \text{RMSE} = 0.42 \quad F = 51$

TEST SET

$N = 14 \quad SD = 0.76$

$N = 25 \quad SD = 0.81$

The results of the predictions through the test sets are disappointing, with SD values of 0.76 and 0.805 for test sets of 14 and 25 compounds. These SD values have been calculated in this work. Feher *et al.*[49] give RMSE values of 0.63 and 0.79 but these seem to be calculated from plots of predicted vs. experimental values.

Of the models that had either no test set or a very limited test set, that of Kelder *et al.*[22] is the simplest; with the sole descriptor being the 3D dynamic polar surface area, as calculated by available software. The data set was restricted to only 45 compounds and no validation of any sort was carried out. Nevertheless an r^2 value of 0.84 is impressive for a one-descriptor model. They also found that brain penetration decreases with increasing polar surface area.

Ertl *et al.*[50] used polar surface as a descriptor in their model to correlate with the partition data of 45 drugs for blood to brain distribution. Although, there was no solid statistics on the fits of the correlation (i.e. standard deviation being absent), they did provide some information to the square of the correlation coefficients (r^2) to be 0.84. Finally, no test sets have been provided for the model and this makes it difficult to determine whether or not the equation can be used to make any prediction for blood to brain distribution of drugs.[50]

Salminen et al[13] examined two set of structurally dissimilar drugs compounds (dataset of 21 and 23 compounds) using retention factors (K_{IAM}) measured by immobilised artificial membrane (packing material for HPLC column) using chromatography technique. From the two training sets, Salminen *et al* [13] obtained the following equation shown below, 21.15 and 21.16. They concluded that the molecular size (V_m), the ionization of acidic ($I_3 = 1$) or basic ($I_2 = -1$) groups and the lipophilicity, reflected either by log P_{oct} or log KIAM are the principle factors contributing to the brain uptake of the compounds examined. It is very unfortunate that they did not have a test set to verify their predictions, and, in any case, there are

a very limited number of compounds used in their training set. Finally, six outliers were found in the Salminen model and they are: diphenhydramine, cimetidine, indomethacin, ranitidine, salicylic acid and thioridazine.

TRAINING SET

$\text{Log BB} = 0.58\log K_{IAM} + 0.89I_2 - 0.01V_m + 1.28$ Equation 21.15

$N = 21 \quad r^2 = 0.848 \quad SD = 0.27 \quad F = 32$

TRAINING SET

$\text{Log BB} = 0.35\log K_{oct} + 0.99I_3 - 0.01V_m + 1.25$ Equation 21.16

$N = 23 \quad r^2 = 0.848 \quad SD = 0.32 \quad F = 35$

NO TEST SET

Hutter[51] has a model for the correlation of the blood to brain distribution (log BB) of drugs that is obtained by using multiple linear regression analysis of molecular descriptors, for a training set of 90 compounds. The descriptors for the drug compounds were computed using a computer program called VAMP. The twelve descriptors are grouped into 5 categories, each representing specific properties: polar surface area, hydrogen bond donor, hydrogen bond acceptor, the presence of chemical groups, physical quantities, shape of drug compound and entropy descriptors. This software can compute 1 compound per minute which includes time for structural optimization. Although this method is very slow, the statistics for the correlation (training set only) appear to be good.

TRAINING SET

$N = 90 \quad r^2 = 0.87 \quad SE = 0.31 \quad F = 41$

NO TEST SETS

Rose and Hall [52] developed a QSAR model for correlation and the prediction of blood to the brain distribution at steady-steady for drugs. Their model is based on five structural properties which are: the hydrogen E-state index for hydrogen bond donors (HBD), the hydrogen E-state index for aromatic CHs (AROM), and the second order difference valence molecular connectivity index (VMC). The model indicates that molecules that penetrate the blood to brain barrier have large AROM values (presence of aromatic groups) but small values of HBD (fewer or weaker H-

Bond donors) and smaller VMC values (less branched molecules with electronegative atoms). Finally, the presence of fluorine or chlorine atoms also leads to higher calculated log BB values. The correlation for this model is poor, compared to earlier models like that of Platts that have larger data sets.[36] It is a shame again that no statistics for the predictions were provided for this large data set of 102 drug compounds.

TRAINING SET

Log BB = -0.202(HBD) + 0.006(AROM) — 0.105(VMC) — 0.425 Equation 21.17

N = 102 r^2 = 0.66 SD = 0.45 F = 62

NO TEST SETS

Lobell *et al.*[53] used 65 drug like structures with experimental log BB values from literature. Lobell *et al.*, split the 65 compounds into 48 compounds in training set and 17 compounds in a test set, with equally homogenous log BB distributions. The 5 descriptors were calculated and obtained by using two computer programs that are: Cerius 2 and the ACD/log D suite. The stepwise linear regression analysis method was employed to obtain the training equation for the blood to brain distribution model for rats. Although, the training equation has very good correlation (n = 48, r^2 = 0.84) and predictions (n = 17, predictive power, with MAE = 0.33 for a test set), it would be interesting to see if this statistics can also be represented for larger data sets.[53]

Ultimately, there are several models with good statistics for training sets that could be further tested using additional compounds for test sets. The general conclusion is that there are a number of models available with good statistics for both training sets and test sets. Some are very simple and can give rapid estimates of log BB. Others are more complicated but might provide better estimates. In addition, there are models that have good statistics for training sets. Most of these models have been constructed with much smaller data sets than are currently available and that is why the aim of this work is to construct a model with a larger data set than has been published before. Also, it is important to show if blood/plasma/serum to brain data for drugs in rats can combined or not and to ascertain if the *in vivo* data of the drugs can be combined with the *in vitro* data of the VOCs and see if a general equation can be made for this combined data set.

Blood to brain/tissues distribution combined with plasma or serum to brain/tissues distribution data for drugs only

It was first necessary to examine the database of *in vivo* log BB (and log blood to tissue data for common compounds in serum, plasma or blood) values, Table 21.0 (see appendix), and to distinguish those that were actually derived from blood to rat brain or tissues (skin, fat, heart, kidney, liver, lung and muscle), $P_{b\text{-brain or tissues}}$, plasma to rat brain or tissues (skin, fat, heart, kidney, liver, lung and muscle), $P_{p\text{-brain or tissues}}$, and serum to rat brain, $P_{s\text{-brain}}$, distribution coefficients. The values of log $P_{b\text{-brain}}$ and log $P_{p\text{-brain}}$ are known for 11 common compounds, and the following statistics were calculated: AE (the average error), AAE (the absolute average error), RMSE (the root mean square error) and SD (the standard deviation). The values of log $P_{b\text{-tissues}}$ and log $P_{p\text{-tissues}}$ are also found for 12 common compounds and the following statistics were also calculated (table 21.1). Results are found in table 21.1.

SAMPLE	N	SD(n-1)	RMSE	AAE	AE
BLOOD-PLASMA[a]	11	0.449	0.428	0.301	0.104
BLOOD-SERUM[a]	8	0.221	0.207	0.110	-0.052
PLASMA-SERUM[a]	10	0.274	0.260	0.156	-0.058
BLOOD-PLASMA[b]	12	0.184	0.176	0.118	0.008

[a] Brain data only

[b]Includes data from the following tissues: skin, fat, heart, kidney, liver, lung and muscle (brain data excluded)

Table 21.1. Correspondence between blood to brain, plasma to brain and serum to brain distribution coefficients, as log P values

The most useful statistic to assess whether there is any systematic difference between the two sets of brain data is the average error. For the 11 common compounds for blood-brain and plasma-brain, this is only 0.1 log units, statistically insignificant by comparison to the experimental error. The AAE and SD statistics indicate random differences; both of these are large, at 0.301 and 0.449 log units

respectively. Since the AE is very small, these large values must either be due to experimental error or to some specific effects for certain compounds. There are two unusual observations in the blood/plasma set, due to codeine, and valproic acid, and some specific effect may be present in these cases.

For log $P_{b\text{-}brain}$ and log $P_{s\text{-}brain}$ there are only 8 common compounds for which AE is very small, at -0.05 log units. Values of AAE (0.11) and SD (0.22) are higher, as expected, but within the range of general experimental error. There is only one unusual observation in this set and that is promazine; again it is not possible to decide if this is the result of experimental error or specific effects. Since promazine is not unusual in comparisons of log $P_{b\text{-}brain}$ and log $P_{p\text{-}brain}$ there may be experimental error in the log $P_{s\text{-}brain}$ value. For the plasma and serum data sets, there are 10 common compounds for which there is a very small value of AE, -0.06 log units, and reasonable values of AAE (0.16) and SD (0.27).

Finally, it was important to assess the systematic difference in statistics between the two sets of tissue data, blood to tissue and plasma to tissue. For the 12 common compounds, this is only 0.01 log units, statistically insignificant by comparison to the experimental error. The AAE and SD statistics indicate random differences; both of these are not too large, at 0.118 and 0.184 log units respectively. Since the AE is very small, these values must either be due to experimental error or to some specific effects for certain compounds. There is one unusual observation in the blood/plasma set, due to toliprolol. This may be experimental error or may result from some specific effect.

The conclusion is straightforward; the systematic differences between log $P_{b\text{-}brain\ or\ tisues}$, log $P_{p\text{-}brain\ or\ tissues}$, and log $P_{s\text{-}brain}$ are so small that all three measures of distribution into brain or tissues (skin, fat, heart, kidney, liver, lung and muscle) can be combined into a global term, log BB or log BT. The random differences, as indicated by AAE and SD, are quite large and suggest that the experimental error in log BB (or log BT) must be around 0.2 log units (AAE) or 0.3 log units (SD).

The three *in vivo* distribution sets (log $P_{b\text{-}brain\ or\ tissues}$, log $P_{p\text{-}brain\ or\ tissues}$, and log $P_{s\text{-}brain\ or\ tissues}$) can now be combined together for brain and all tissues mentioned in the thesis (skin, fat, heart, kidney, liver, lung and muscle), but these ratios will not

be averaged; thus the data points will be greater than the number of compounds. The data set for brain tissue only is shown in table 21.0.

Discussion on blood/plasma/serum to brain distribution for drugs only

The 120 compounds (all of which are drugs) were used to construct the blood brain distribution model, see table 21.2. The structures for compounds with trivial names can be found in the appendix of this thesis.

Salminen also found in their dataset that some drugs have basic groups and it might be useful to investigate this further. They found that this indicator variable (**Ib**) in the correlation for the BB model was not significant and tried to use an alternative new indicator variable for drug acids (**Ia**). They then found that this acid indicator variable for drugs was significant and it was found to be useful in their equation to describe the blood to brain transport process for various drugs.[13] A first attempt to construct a model was to use the indicator variable for carboxylic acids (**Ia**). Platts[36] and Salminen[13] found that acidic drugs bind more strongly to albumin then basic compounds, hence reducing their brain penetration. This descriptor has been used before by Platts and Salminen, and this is why it is important to use this for all drug blood to tissue distribution models. The independent variable **Ia** for a value of '1' means the presence of carboxylic acid as a functional group to the drug compound. The independent variable **Ia** for a value of '0' means that the carboxylic acid is absent drug compound. This descriptor was used in this model to correlate and predict blood to brain distribution for drugs and VOCs as shown in table 21.4 (see later). Note no other descriptors have been in this model (or other models in the thesis) for neutral compounds that may have high plasma protein binding. This has not been investigated by anyone known in the literature for any blood to tissue distribution models for drug compounds.

A number of compounds had previously been identified as outliers,[13, 21, 35, 36, 39, 40, 47, 48] possibly because of active transport or metabolism. There are also five additional compounds (AI-1, alfentanil, inaperisone, methylphenidate and tolbutamide) that were outliers, not only as regards blood to brain distribution, but also in distribution from blood to several other biological tissues. These were

removed completely from subsequent models, as the residuals appeared to follow the same direction in terms of error. For the present model, aspirin, 2,4-dichlorophenoxyacetic acid, terbinafine, promazine and praxanthine were marked as new outliers. It is also not very surprising that from this model and other reported models that the remaining nine compounds (SKB3, SKB4, SKB9, SKB13, SKB19, Fluphenazine, Y-G 20, Org12962 and Temelastine) have found to be outliers. [13, 21, 35, 36, 39, 40, 47, 48]

The original method[54, 55] for the determination of *in vivo* log BB values analysed material by a radioactive method. All radioactive species, including metabolites, were included in the count. As pointed out, [35] this would lead to observed log BB values for compounds that were metabolized to be much more negative than the (true) log BB value for the parent compound. In addition, *in silico* methods for log BB all refer to passive transport. Active transport and efflux mechanisms will result in compounds appearing to be outliers to a passive transport algorithm.

EQUATIONS (rat model)	N(dp)	N(nc)	r^2	SD	F	OUTLIERS	D
FULL SET (*IN VIVO*)							
Log BB = 0.607 + 0.183**E** – 0.778**S** – 0.697**A** – 0.462**B** + 0.679**V** – 1.199**Ia**	152	120	0.708	0.383	58.567	14(nc)	6
TRAINING SET							
Log BB = 0.616 + 0.115**E** – 0.781**S** – 0.766**A** – 0.357**B** + 0.675**V** – 1.195**Ia**	76	76	0.722	0.415	29.832	-	6
TEST SET							
N(nc) = 65 **N**(dp) = 76							
SD (n-1) = 0.359 **RMSE** = 0.356							
AAE = 0.287 **AE** = -0.010							

N (dp) = number of data points used in the training set
N (nc) = number of compounds used in the training set
r^2 = is the square of the overall correlation coefficient
SD = the standard deviation
F = F-statistic
D = the number of descriptors used in the model

Table 21.2. Summary of the *in vivo* equations for blood/serum/plasma to brain distributions for drugs only

The statistics of the equation for the full set in table 21.2. are satisfactory and the predictive capability of the equation can be assessed by the usual method of constructing a training set and an independent test set. The selection of the training and test sets was carried out using the usual computing method called Kennard-Stone and that the two sets were made up of the same number of data points. These two criteria ensure that the chemical space covered by the training and test sets is the same. The training equation is shown table 21.2.

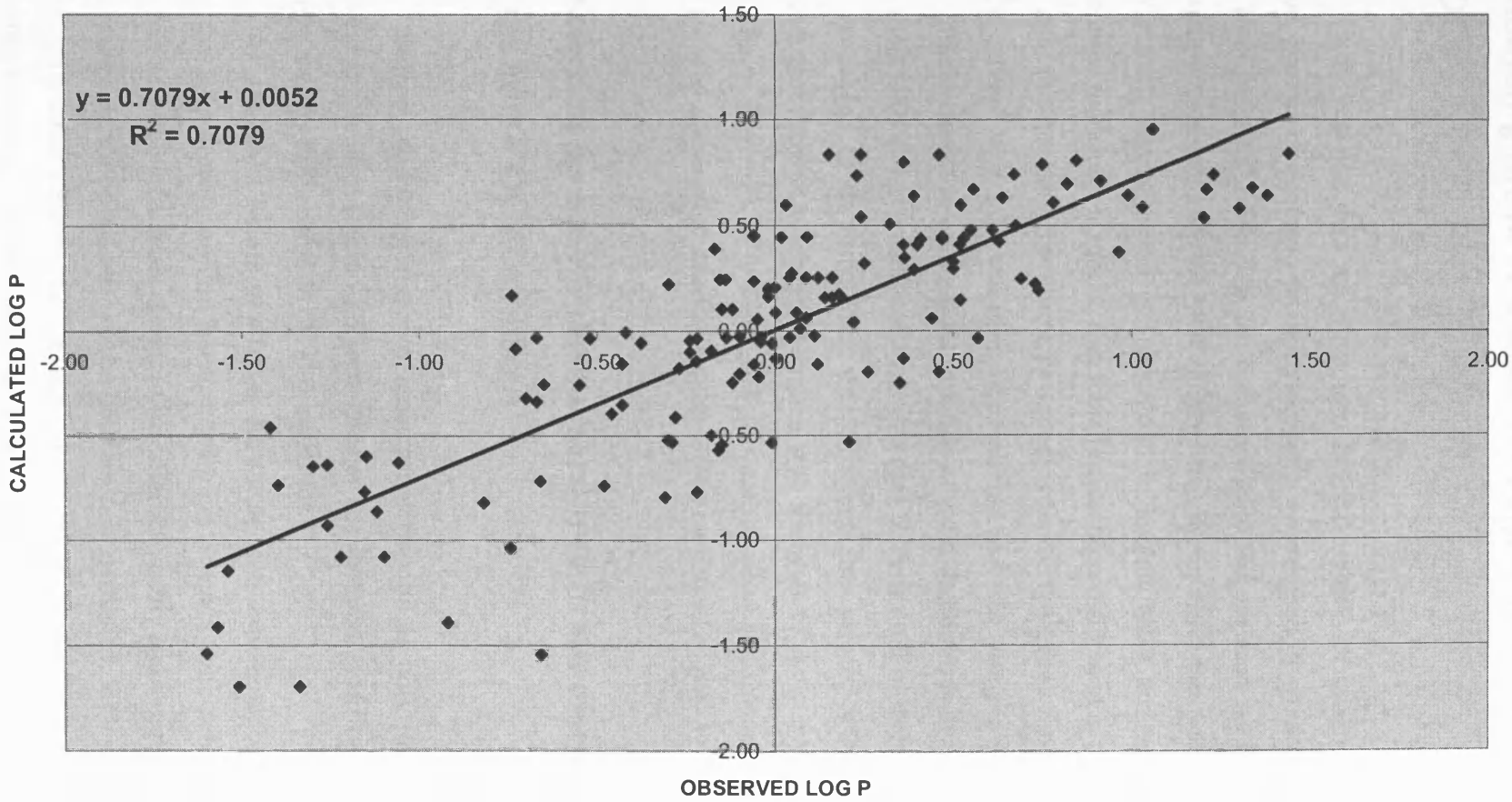
The statistics of the full set (table 21.2) and the training set (table 21.2), are similar and so are the coefficients. If the training set equation is used to predict the remainder of the data in the (unused) test set, it is found that AE = -0.01, AAE = 0.29 RMSE = 0.36 and SD = 0.36, that is there is very little bias, and the predictive capability of the training equation (table 21.2) and by implication the full set equation (table 21.2) is therefore assessed. Thus the full (*in vivo*) equation and the training set can be used to predict further values of log BB close to the suggested experimental error of SD = 0.3 and AAE =0.2 log units.

The inter-correlations of descriptors used for drug blood/plasma/serum to brain distribution are shown in table 21.3. The largest r^2 value for E versus S descriptor is 0.49. It appears that there is no strong inter-correlation between the set of descriptors used in this model.

r	E	S	A	B	V	Ia
E	1.00					
S	0.70	1.00				
A	-0.11	-0.07	1.00			
B	0.54	0.62	0.08	1.00		
V	0.66	0.67	-0.13	0.67	1.00	
Ia	-0.23	-0.23	0.34	-0.28	-0.13	1.00

Table 21.3. The inter-correlation of descriptors for 120 compounds

Calculated log P (BB) versus observed log P (BB) for blood, serum, plasma to brain distribution for drugs only



Combining the *in vivo* data of drugs with the *in vitro* data of VOCs

Now that the data set of *in vivo* log BB values for passive transport from blood, plasma or serum to rat brain has been investigated, it would be interesting to see what happens if this data are combined with the previous set of *in vitro* values (chapter 12, table 12.3), for VOCs. Note that the *in vitro* data set is for transfer of VOCs from human blood and rat blood to human brain and rat brain. It was concluded in chapter 9 and 12, that there is very little difference in data from humans and rats, and that data from these two species can be combined for VOCs. Note that in the present work, the use of *in vitro* plasma to brain data for VOCs will not be used, mainly because of the small number of extra compounds involved. It appears that most authors have used some *in vitro* data in their drug models. Although, they have managed to combine these two different data sets of VOCs and drugs, it appears that there is no single indication with statistics to show how this is possible. The aim of this work is to show how this can be done and to see if these two data sets (*in vivo* and *in vitro* measurements) can be combined or not.

Discussion on blood/plasma/serum to brain distribution for drugs and VOCs combined

Before combining the two models together of *in vivo* (table 21.0 of this chapter) and *in vitro* (table 12.3 from chapter 12) data sets, it is first important to see if there are any differences in the two equations.

As described before in chapter 12, the *in vitro* equation for VOCs does not have the Ia-indicator variable, because there are no carboxylic acids in the data set. The two sets of coefficients appear to be different visibly and suggest that the *in vivo* and *in vitro* distributions may not be combined. In particular, the *c*-coefficients differ appreciably 0.607 (SD = 0.127) as against – 0.057 (SD = 0.070), which suggests that there is a systematic difference between the *in vivo* and *in vitro* distributions.

One way to check for systematic deviations is to use the coefficients in the *in vivo* equation to predict log BB values for the 78 VOCs used to construct the *in vitro*

equation. It is found that for the predicted and observed *in vitro* log BB values, AE = 0.53, AAE = 0.53, RMSE = 0.58 and SD = 0.58 log units. The numerical values here show that the *in vivo* equation can only predict the *in vitro* dataset within 0.58 log units (SD). The large value of AE, and the exact equivalent of AAE, suggests that there is a systematic difference in the two distribution models. It is possible that this could be due to the two series of compounds inhabiting different areas of chemical space and resulting in poor predictions.

However, especially in view of the marked difference in the *c*-coefficient for the *in vivo* and the *in vitro* equation, it might be possible to apply another indicator variable, **Iv**, defined as **Iv** = 1 for the *in vitro* data and **Iv** = 0 for the *in vivo* data. The resulting equation with this indicator variable is shown in blue in table 21.4 (*in vivo* + *in vitro* combined). The **Iv** term is statistically very significant (T-test = -5.04, p-value = 9.78E-07), and the coefficient of the indicator variable, -0.487, is quite close to the difference in the *c*-coefficients of the *in vivo* and *in vitro* (-0.66) and to the AE for predictions of the VOC data from the drug data (0.53 log units). The coefficients of drug (*in vivo*) and the drug and VOC (*in vivo* + *in vitro*) combined equation are the same within experimental error (SD values of the coefficients are in the drug equation: c=0.127, e=0.071, s=0.078, a=0.113, b=0.108, v=0.082, ia=0.130 and in the combined equation (*in vivo* + *in vitro*): c=0.106, e=0.057, s=0.063, a=0.093, b=0.084, v=0.058, ia=0.109, iv=0.097). It is concluded that there is a systematic difference between the *in vivo* and *in vitro* distributions (as **Iv** is -0.5 log units), and that the two sets of distribution data can only be combined by taking into account this difference (see equation 21.4).

The full data set (*in vivo* and *in vitro* measurements) is given in table 12.3 and table 21.0, and consists of 206 compounds and 229 data points, the difference arising because blood/plasma/serum to brain distributions are counted separately. The correlation of this set of data points with the Abraham descriptors is shown in table 21.4. The fits for this very large data set appear to be good (S.D = 0.328 and r^2 = 0.721). This represents a good model for a combined data set of VOCs and drugs for blood/plasma/serum to brain distribution, and is the largest general model yet reported for both types of data sets (*in vivo* rats + *in vitro* rats and humans combined).

There was only one additional outlier observed from this training set and this is cimetidine (SKB1). The residual for this compound is 1 log unit and may result from experimental error.

EQUATIONS	N(dp)	N(nc)	r^2	SD	F	OUTLIERS	D
FULL SET (*IN VIVO*) **Log BB** = 0.607 + 0.183**E** — 0.778**S** — 0.697**A** – 0.462**B** + 0.679**V** – 1.199**Ia**	152	120	0.708	0.383	58.567	14(nc)	6
FULL SET (*IN VITRO*) **Log BB** = - 0.057 + 0.017**E** — 0.536**S** — 0.323**A** — 0.335**B** + 0.731**V**	78	78	0.725	0.203	37.908	2(nc)	5
FULL SET (*IN VIVO + IN VITRO*) **Log BB** = 0.565 + 0.152**E** – 0.756**S** – 0.664**A** – 0.382**B** + 0.652**V** – 1.196**Ia** – 0.487**Iv**	229	206	0.721	0.328	81.519	1	7
TRAINING SET **Log BB** = 0.602 + 0.158**E** – 0.739**S** – 0.758**A** – 0.398**B** + 0.632**V** – 1.179**Ia** – 0.477**Iv**	115	115	0.729	0.369	41.048	-	7
TEST SET **N**(nc) = 104 **N**(dp) = 114 **SD** (n-1) = 0.290 **RMSE** = 0.289 **AAE** = 0.221 **AE** = -0.001							

N (dp) = number of data points used in the training set
N (nc) = number of compounds used in the training set
r^2 = is the square of the overall correlation coefficient
SD = the standard deviation
F = F-statistic
D = the number of descriptors used in the model

Table 21.4. Summary of the equations for blood/serum/plasma to brain distributions for VOCs and drugs combined.

In order to assess the predictive capability of the combined equation (*in vivo + in vitro*) a training set and a test set was constructed. The equation for the training set was used to predict log BB values for the test set, as shown in table 21.4.

The coefficients and statistics of the full set of the combined equation (*in vivo + in vitro*) and training equation (*in vivo + in vitro*) are reasonably consistent in terms of coefficients. The training equation of the combined set can then be used to

predict log BB values for the (unused) test set. The statistic and predictions are for N(nc) = 104 and N(dp) = 114, that AE = 0.00, AAE = 0.22, RMSE = 0.29 and the SD = 0.29 log units. Thus the combined (*in vivo* + *in vitro*) training equation and the full combined equation (in blue) can be used to predict further values of log BB to close to our suggested experimental error of around 0.2 log units (AAE) or 0.3 log units (SD).

r	E	S	A	B	V	Ia	Iv
E	1.00						
S	0.89	1.00					
A	0.31	0.33	1.00				
B	0.84	0.87	0.42	1.00			
V	0.85	0.84	0.26	0.85	1.00		
Ia	0.02	0.01	0.37	0.00	0.06	1.00	
Iv	-0.82	-0.81	-0.44	-0.83	-0.76	-0.18	1.00

Table 21.5. The inter-correlation of descriptors for 120 compounds

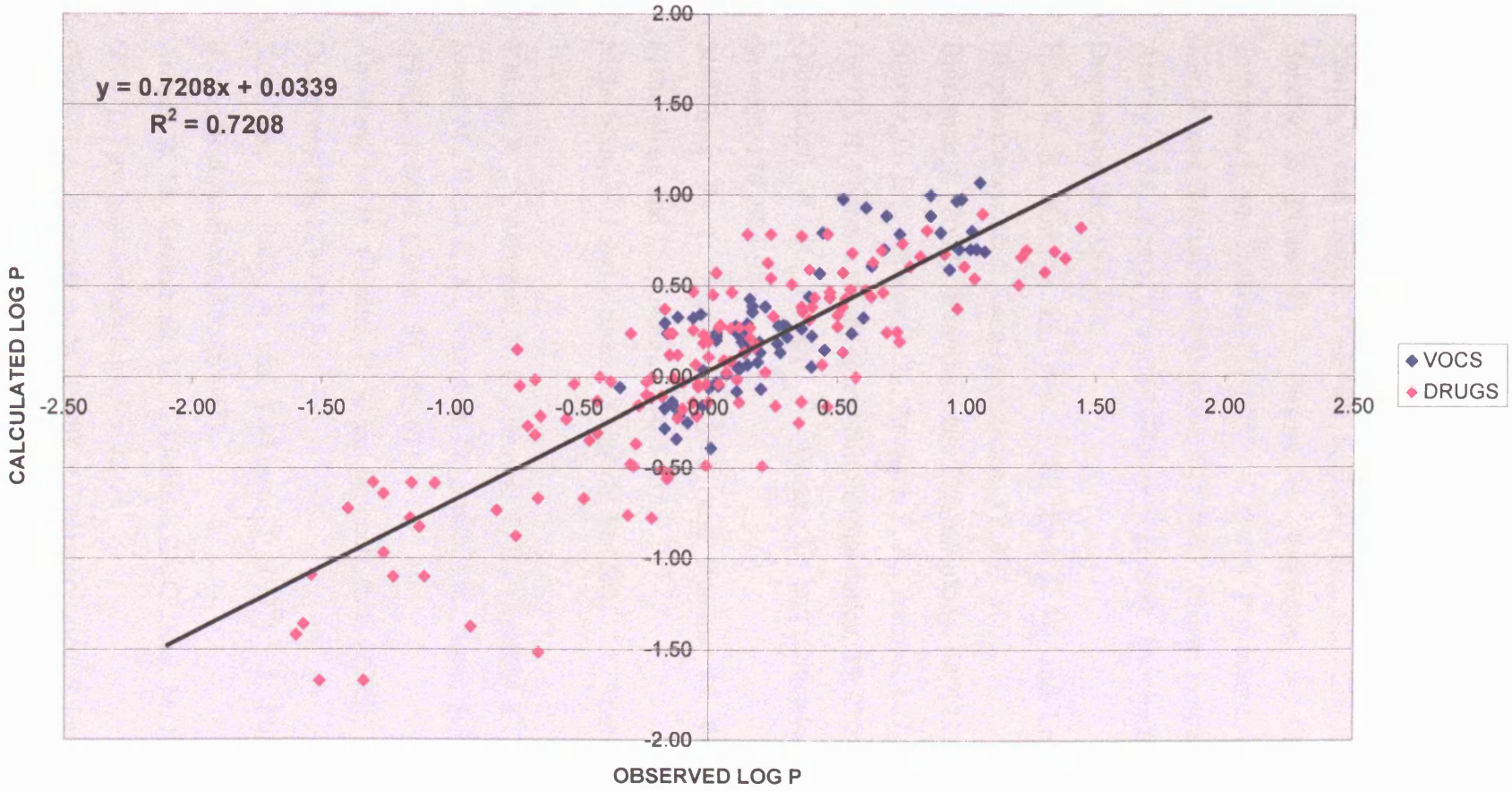
The inter-correlations of descriptors are obtained for the combined *in vivo* and *in vitro* equation (from table 21.4) is shown by table 21.5. The largest r^2 value for E versus S descriptor is 0.79. It appears that there is no other strong inter-correlation between the set of descriptors used in this model.

The descriptors used in the equation reflect solute/solvent interactions. One can, therefore, interpret the coefficients in terms of the effect that particular interactions have on the process under consideration. The positive **e**- and **v**-coefficients in the blood/plasma/serum to brain (*in vivo* + *in vitro*) equation from table 21.4, indicate that increasing molecular size and the presence of n- and Π-electrons pairs can push drugs out of blood/plasma/serum and into brain. The coefficient of the 'polar' descriptor **S** is negative, indicating that polarity is stronger in the blood/plasma/serum phase than in the brain tissue phase. Therefore the drug compounds will move from the brain phase and into the blood/plasma/serum phase. The coefficients of the descriptors **A** and **B** are negative, indicating that the hydrogen-bond acidity and basicity act to keep drug compounds in

blood/plasma/serum and out of the brain. In turn, this indicates that the brain has less hydrogen bond acidity and basicity than is blood/plasma/serum, but is more able to accommodate bulky solutes and interact with them via dispersive forces than blood/plasma/serum. The large negative coefficient of the indicator variable (***Ia***), in the combined equation from table 21.4., indicates a hitherto unknown factor affecting the transfer of solutes between blood/plasma/serum and the brain. It is apparent that the presence of $-CO_2H$ group acts to hinder brain penetration further. This could be simply due to intrinsic hydrogen bonding and the polarity properties of neutral acids. Finally, it is known that some carboxylic acids drugs may bind to albumin present in plasma/serum and blood, which may also account for the negative contribution of the coefficient ***Ia***.

Finally, this new model based on 206 compounds suggest that this can now be reliably and be used to correlate and predict blood (plasma or serum) to brain distribution for *in vitro* and *in vivo* drug compounds. Of course, any prediction using the full *in vivo* equation (drugs only) or the full combined equation (*in vivo* plus *in vitro*) can, or should, only be made within the chemical space of the test set used to assess the predictions. Note also that this model can also be used to predict a test set (104 compounds) within a standard deviation value 0.29 log units. This up to date model has now superseded current BB models like Platts *et al.*[36] in terms of compound numbers and data points. This model for blood (plasma or serum) to brain distribution is very simple to calculate, and the method employed is robust and requires very little computing time.[37] Whereas the Norinder *et al*[38] method is complicated and time consuming, as computational methods are required for the conformational analysis of the solutes via the use of intensive computer calculations (which can slow down the screening process).

Calculated log P versus observed log P for the distribution of compounds from blood, serum and plasma to brain



References

1. Poulin, P. and Theil, F. P (2002). J. Pharm. Sci., **91**, 1358-1367.
2. Baudry, S., Pham, Y. T., Baune, B., Vidrequin, S., Crevoisier, C. H., Gimenez, F. and Farinotti, R (1997). J. Pharm. Pharmacol., **49**, 1086-1090.
3. Xie, R. and Hammarlund-Udenaes, M (1998). Pharm Research, **15**, 570-575.
4. Aravagiri, M., Teper, Y. and Marder, S. R (1999). Biopharmaceutics & Drug Disposition, **20**, 369-377.
5. Han, K. S., Kim, Y. G., Yoo, J. K., Lee, J. W. and Lee, M. G (1998). Biopharmaceutics & Drug Disposition, **19**, 493-500.
6. Bjorkman, S (2002). Pharmacy and Pharmacology, **54**, 1237-1245.
7. Blakey, G. E., Nestorov, I. A., Arundel, P. A., Aarons, L. J. and Rowland, M (1997). J. Pharmacokinetics and Biopharmaceutics, **25**, 277-312.
8. Oldendorf, W. H., Hyman, S., Braun, L. and Oldendorf, S. Z (1972). Science, **178**, 984-986.
9. Schillings, R. T., Sisenwine, S. F. and Ruelius, H. W (1977). Drug Metabolism and Disposition, **5**, 425-435.
10. Mclachlan, A. J and Hosseini-Yegabeh, M (2001). J. Pharm. Sci., **90**, 1817-1828.
11. Young, R. C., Mitchell, R. C., Brown, T. H., Ganellin, C. R., Griffiths, R., Jones, M., Rana, K. K., Saunders, D., Smith, I. R., Sore, N. E. and Wilks, T. J (1988). J. Med. Chem., **31**, 656-671
12. Abraham, M. H., Chadha, H. S. and Mitchell, R. C (1995). Drug Design and Discovery, **13**, 123-131.
13. Salminen, T., Pulli, A. and Taskinen, J (1997). J. Pharmaceutical and Biomedical Analysis, **15**, 469-477.
14. Arendt, R. M., Greenblatt, D. J., Liebisch, D. C., Luu, M. D. and Paul, S. M (1987). Psychopharmacology, **93**, 72-76.
15. Glynn, S. L., Yazdanian, M (1998). J. Pharmaceutical Sciences, **87**, 306-310.
16. Borlander, H. G., Wahistrom, G. and Norberg, L (1984). Acta Pharmacol. et Toxicol., **54**, 33-40.

17. Wilkinson, J. M. and Pollard, I (1993). Development Brain Research, 75, 193-199.
18. Bickel, M. H. and Gerny, R (1980). J. Pharm. Pharmacology, 32, 669-674.
19. Hu, O. Y-P. and Curry, S. H (1989). Biopharmaceutics & Drug Disposition, **10**, 537-548.
20. Okuyama, S. and Aihara, H (1984). Japan. J. Pharmacology, **35**, 95-103.
21. Cabrera, M. A., Bermejo, M., Perez, Maykel. and Ramos, R (2004). J. Pharm. Sci, **93**, 1701-1717.
22. Kelder, J., Grootenhuis, P. D. J., Bayada, D. M., Delbressine, L. P. C. and Ploemen, J. P (1999). Pharmaceutical Research, **16**, 1514-1519.
23. Tsuneizumi, T., Babb, M. S. and Cohen, B. M (1992). Biol. Psychiatry, **32**, 817-824.
24. Yamamoto, F., Oka, H., Antoku, S., Ichiya, Yu-ichi., Masuda, K. and Maeda, M (1999). Biol. Pharm. Bull., **22**, 590-597.
25. Sako, K., Diksic, M., Kato, A., Yamamoto, Y. L. and Feindel, W (1984). J. Cerebral Blood Flow and Metabolism, **4**, 259-263.
26. Haddad, S., Withey, J., Lapare, S., Law, F. and Krishnan, K (1998). Environmental Toxicology and Pharmacology, **5**, 245-255.
27. Yazdanian, M (1999). J. Pharm. Sci., **88**, 950-954.
28. Csanady, G. A., Oberste-Frielinghaus, H. R., Semder, B., Baur, C., Schneider, K. T. and Filser, J. G (2002). Arch. Toxicology, 76, 299-305.
29. Moerlein, S. M., Laufer, P. and Stocklin, G (1985). Int. J. Nucl. Med. Bio., **12**, 353-356.
30. Doze, P., Elsinga, P. H., Maas, B., Van Waarde, A., Wegman, T. and Vaalburg, W (2002). Neurochemistry International, **40**, 145-155.
31. Wojcikowski, J. and Daniel, W. A (2002). Pol. J. Pharmacol., **54**, 647-654
32. Bonate, P. L., Swann, A. and Silverman, P. B (1996). J. Pharm. Sci., **85**, 878-883.
33. Poulin, P. and Theil, F. P (2000). J. Pharm. Sci., **89**, 16-35.
34. Javaid, J. I. and Davis, J. M (1993). Biopharmaceutics & Drug Disposition, **14**, 357-364.
35. Abraham, M. H., Chadha, H. S. and Mitchell, R. C (1994). J. Pharm Sci., **83**, 1257-1268.

36. Platts, J. A., Abraham, M.H., Zhao, Y.H., Hersey, A., Ijaz, L. and Butina, D (2001). Eur. J. Med. Chem., **36**, 719-730.

37. PharmaAlgorithms, ADME Boxes, Version 2.2, PharmaAlgorithms Inc., 591 Indian Road, Toronto, ON M6P 2C4, Canada.

38. Norinder, U., Sjoberg, P. and Osterberg, T (1998). J. Pharm. Sci., **87**, 952-959.

39. Lombardo, F., Blake, J. F. and Curatolo, J. W (1996). J. Med. Chem., **39**, 4750-4755.

40. Clark, D (1999). J. Pharm. Sci., **88**, 815-821.

41. Kaliszan, R. and Markuszewski, M (1996). International Journal of Pharmaceutics, **145**, 9-16.

42. Luco, J. M (1999). J. Chem. Inf. Comput. Sci., **39**, 396-404.

43. Keseru, G. M., Molnar, L (2001). J. Chem., Inf. Comput. Sci., **41**, 120-128.

44. Hou, T. J. and Xu, X. J (2003). J. Chem. Inf. Comput. Sci., **43**, 2137-2152.

45. Hou, T. J. and Xu, X. J (2004). J. Chem. Inf. Comput. Sci., **44**, 766-770.

46. Ooms, F., Weber, P., Carrupt, P-A. and Testa, B (2002). Biochimica et Biophysica Acta, **1587**, 118-125.

47. Liu, R., Sun, H. and So, S-S (2001). J. Chem. Inf. Comput. Sci., **41**, 1623-1632.

48. Kaznessis, Y. N., Snow, M. E. and Blankley, C. J (2001). J. Computer-Aided Molecular Design, **15**, 697-708.

49. Fehur, M., Sourial, E. and Schimidt, J. H (2000). Int. J. Pharm., **201**, 239-247.

50. Ertl, P., Rohde, B. and Selzer, P (2000). J. Med. Chem., **43**, 3714-3717.

51. Hutter, M. C (2003). J. Computer-Aided Molecular Design., **17**, 415-433.

52. Rose, K., Hall, L. H. and Kier, L (2002). J. Chem. Inf. Comput. Sci., 42, 651-666.

53. Lobell, M., Molnar, L. and Keseru, G. M (2003). J. Pharm Sci., **92**, 360-370.

54. Young, R. C., Mitchell, R. C., Brown, T.H., Ganellin, C.R., Griffiths, R., Jone, M., Rana, K K., Saunders, D., Smith, I. R., Sore, N. E. and Wilks, T. J (1988). J Med Chem., **31**, 656-671.

55. Young, R. C., Ganellin, C. R., Griffiths, R., Mitchell, R. C., Parsons, M. E., Saunders, D. and Sore, NE (1993). Eur. J. Med. Chem., **28**, 201-211.

Chapter

22

Blood/plasma to fat distribution for drugs and VOCs combined

The distribution drugs between blood (or plasma) and fat tissue has become a matter of interest for the pharmaceutical industry and toxicological sciences. Stored body fat is a kind of body tissue that serves as a source of energy and it also fulfils its purpose to cushion and insulate vital organs in human body. The fat tissue does not contain a rich supply of blood capillaries, unlike other tissues such as heart which has a huge supply of blood capillaries to the tissue.

There have been a small number of reports of data for drug blood (or plasma) to fat distribution in rats (*in vivo*).[1-5] There have been no specific reviews reported yet that attempt to make models to predict blood to fat distribution either for rats or humans. The present work is restricted to a small data set of 17 drug compounds (18 data points), which is all that has been published. The drug data (*in vivo*) for blood (or plasma) to fat distribution of rats, is at steady state concentration and the distribution from blood to fat, BF, is defined by the equation (22.0).

BF = (conc. of drug in fat)/(conc. of drug in blood or plasma) Equation 22.0

The aim is to attempt to construct a model based on the Abraham equations to predict the distribution process. Note that no attempts have been made in this work to construct a correlative equation to predict blood (or plasma) to fat distribution based on drugs only, as there are only a limited number of data points that can be used in this model. This small data set of drugs (*in vivo* data) can be combined with

the VOC data (*in vitro* data), to construct an overall equation to predict blood (or plasma) to fat distribution for VOCs and drugs in general. The statistics will be obtained for the equations and it will be possible to see if VOC (*in vitro*) and the drugs (*in vivo*) can be combined. This equilibrium distribution for blood (or plasma) to fat distribution will be based on rats, and will be the first reported model in this area for blood (or plasma) to fat tissue transport. It should be possible to build bigger models in the near future providing further literature data are published based on drug blood to fat distribution.

Finally, the aim of this work is to construct an equation as part of a general method for 'high-throughput' prediction of equilibrium blood (or plasma)-fat distribution, based on structure using the Abraham descriptors (E, S, A, B, and V).[6] The drug data set used in this model to obtain model to predict blood (or plasma) to fat distributions in rats (see appendix) is found in table 22.0 (see appendix).

Discussion on blood/plasma to fat distribution for drugs

As discussed before (chapter 21), the distribution sets for the tissues that include fat, can be combined together for plasma, blood and serum. In this case, the two distribution sets (log $P_{p\text{-}fat}$ and log $P_{b\text{-}fat}$) from table 22.0 (see appendix) can now be combined together for fat, but these ratios will not be averaged; thus the data points will be greater than the number of compounds.

Most of the data on blood to fat distribution is actually on VOCs, for which indirect (*in vitro*) values are available for 126 compounds (table 14.3, from chapter 14). The total set of compounds (*in vitro* and *in vivo*) numbers were 141 and this combined set, resulted in the equation shown in table 22.1

The structures for compounds with trivial names can be found in the appendix of this thesis. A first attempt to construct a model, included the indicator variable for carboxylic acids (**Ia**). This descriptor has been used before for blood to brain distribution of drugs by Platt[7] and Salminen.[8] The independent variable **Ia** for a value of '1' means the presence of carboxylic acid as a functional group to the drug compound. The independent variable **Ia** for a value of '0' means that the carboxylic acid is absent. The second attempt is to see if there is any difference in the *in vitro* and *in vivo* data set for VOCs and drugs. In this case, one can apply another

indicator variable, **Iv**, defined as **Iv** = 1 for the *in vitro* data and **Iv** = 0 for the *in vivo* data. This indicator descriptor was also used in this model to correlate and predict blood to fat distribution (drugs and VOCs combined) resulting in the equation as shown in table 22.1. The resulting equation is shown in blue in table 22.1 (*in vivo* + *in vitro* combined).

EQUATIONS	N(dp)	N(nc)	r^2	SD	F	OUTLIERS	D
FULL SET (*IN VIVO+IN VITRO*)							
Log BF = 0.432 + 0.117**E** – 0.099**S** – 1.543**A** – 2.106**B** + 1.596**V** – 0.946**Ia**	142	141	0.873	0.312	154.982	2	6
TRAINING SET							
Log BF = 0.362 + 0.105**E** – 0.137**S** – 1.541**A** – 2.113**B** + 1.688**V** – 1.037**Ia**	71	70	0.872	0.343	72.858	-	6
TEST SET							
N(nc) = 70 **N**(dp) = 71 **SD** (n-1) = 0.288							
RMSE = 0.286 **AAE** = 0.234							
AE = -0.023							
FULL SET (Iv not significant)							
Log BF = 0.484 + 0.111**E** – 0.101**S** – 1.553**A** – 2.117**B** + 1.589**V** – 0.956**Ia** – 0.043**Iv**	142	141	0.873	0.313	131.891	2	7

N (dp) = number of data points used in the training set
N (nc) = number of compounds used in the training set
r^2 = is the square of the overall correlation coefficient
SD = the standard deviation
F = F-statistic
D = the number of descriptors used in the model

Table 22.1. Summary of the two equations for blood/plasma to fat distributions for VOCs and drugs combined.

The full data set (*in vitro* and *in vivo* measurements) is given in table 14.3 (chapter 14) and table 22.0, and consists of 141 compounds and 142 data points, the difference arising because blood/plasma to fat distributions was counted separately. The correlation of this set of data points with the Abraham descriptors is shown in table 22.1. The fits for this large data set appear to be very good (S.D = 0.312 and r^2 = 0.873). This represents the most up to date (and only) model for a combined data set of VOCs and drugs for blood to fat distribution.

There were only two outliers observed from this training set: pyrene and p-phenylbenzoic acid. The residual for both of these compounds is 0.9 and 1.1 log units. This may of resulted from experimental error.

The **Iv** term is statistically **not significant** (T-test = -0.17, p-value = 0.86), and the coefficient of the indicator variable is 0.043. This suggests that there is a very small systematic difference between the *in vivo* and *in vitro* distributions, and that the distribution data can be combined, without taking into account the indicator variable **(Iv)**. The *in vivo* and *in vitro* largely overlap in terms of their log BB values, see the graph on page 303, although the chemical space they occupy must be considerably different.

In order to assess the predictive capability of the combined equation (*in vivo* + *in vitro*) by the stratagem of constructing a training set and using the equation for the training set to predict log BF values for the test set, as shown in table 22.1.

The coefficients and statistics of the full set of the combined equation (*in vivo* + *in vitro*) and training equation (*in vivo* + *in vitro*) are reasonably consistent in terms of coefficients. The training equation of the combined set can then be used to predict log BF values for the (unused) test set. The statistic and predictions are for N(nc) = 70 and N(dp) = 71, that AE = -0.02, AAE = 0.23, RMSE = 0.29 and the SD = 0.29 log units. Thus the combined (*in vivo* + *in vitro*) training equation and the full combined equation (in blue) can be used to predict further values of log BF within to close to the suggested experimental error (SD = 0.3 log units, chapter 21).

The inter-correlations of descriptors are obtained for the combined *in vivo* and *in vitro* equation is shown by table 22.2. The largest r^2 value for E versus S descriptor is 0.74. It appears that there is no strong inter-correlation between the other set of descriptors used in this model.

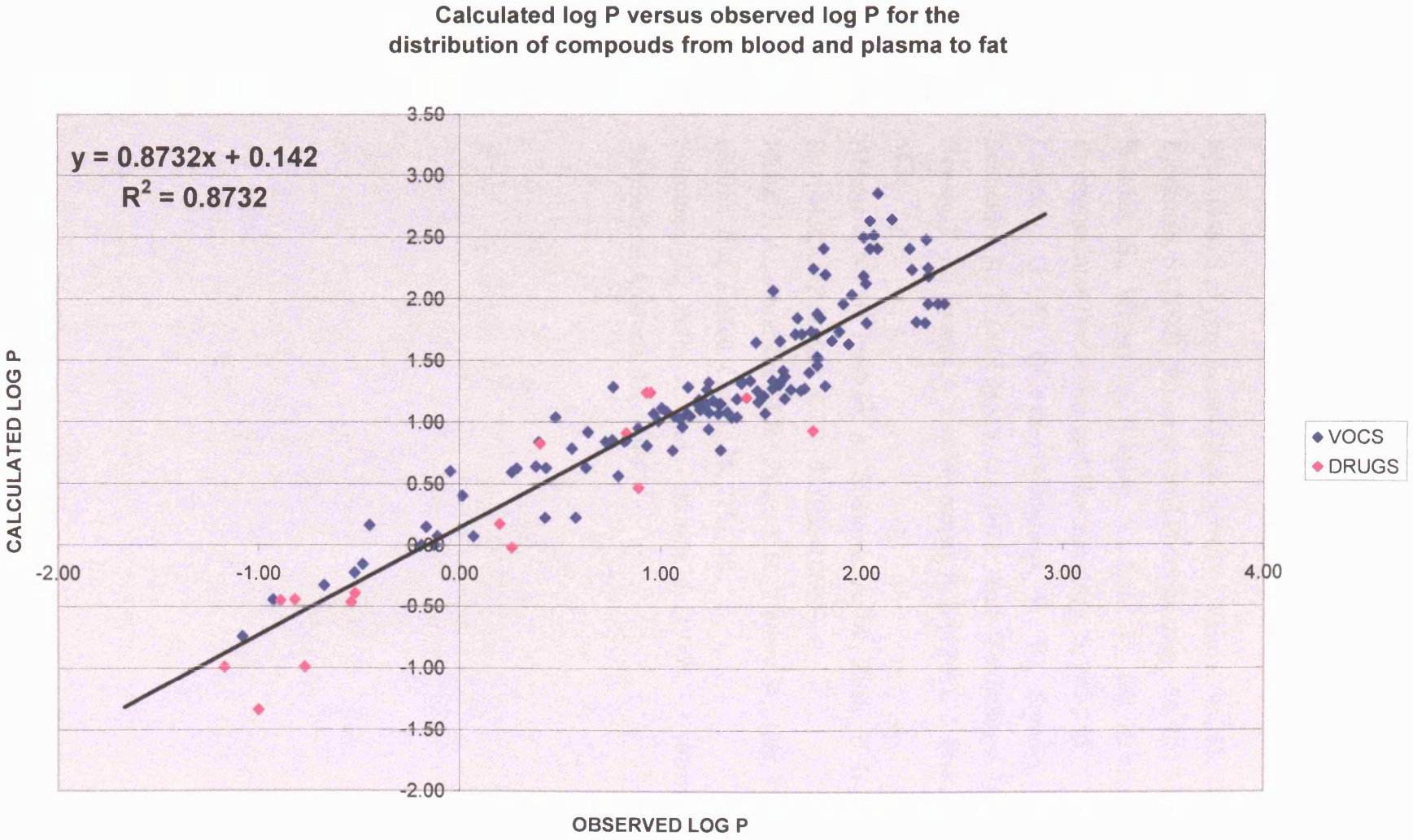
r	E	S	A	B	V	Ia
E	1.00					
S	0.86	1.00				
A	0.49	0.61	1.00			
B	0.72	0.82	0.56	1.00		
V	0.73	0.66	0.42	0.75	1.00	
Ia	0.28	0.38	0.58	0.31	0.32	1.00

Table 22.2. The inter-correlation of descriptors for 141 (142 data points) compounds

The descriptors used in the equation reflect solute/solvent interactions. One can, therefore, interpret the coefficients in terms of the effect that particular interactions have on the process under consideration. The **e**-coefficient (P = 0.27) and **s**-coefficient (P = 0.46) is not significant. The positive **v**-coefficient in the blood/plasma to fat (*in vivo* + *in vitro*) equation from table 22.1, indicate that increasing molecular size can push drugs out of blood/plasma and into the fat phase. The coefficients of the descriptors **A**, and **B** are all negative, indicating that the hydrogen-bond acidity and basicity act to keep drug compounds in blood/plasma and out of the fat phase. In turn, this indicates that the fat is less acidic and basic than is blood/plasma, but is more able to accommodate bulky solutes and interact with them via dispersive forces than blood/plasma. The large negative coefficient of the indicator variable (***Ia***), in the combined equation from table 22.1., indicates a hitherto unknown factor affecting the transfer of solutes between blood/plasma and the fat. It is apparent that the presence of $-CO_2H$ group acts to hinder fat penetration further. This could be simply due to intrinsic hydrogen bonding and the polarity properties of neutral acids. Finally, it is known that some acidic drugs may bind to albumin present in plasma and blood, which may also account for this negative contribution of the coefficient ***Ia***.

The statistics for this model is the best out of all the blood (plasma or serum) to tissue distribution models for VOCs and drugs combined. This is indicated by the large spread of data shown from the plot (page 303) giving rise to strong correlative fits.

This new model based on 141 compounds suggest that this can now be reliably and be used to correlate and predict blood/plasma to fat distribution for *in vitro* and *in vivo* drug compounds. Of course, any prediction using the full combined equation (*in vivo* plus *in vitro*) can, or should, only be made within the chemical space of the test set used to assess the predictions.


Calculated log P versus observed log P for the distribution of compouds from blood and plasma to fat
y = 0.8732x + 0.142
R2 = 0.8732
CALCULATED LOG P
OBSERVED LOG P
3.50
3.00
2.50
2.00
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-2.00
-2.00
-1.00
0.00
1.00
2.00
3.00
4.00
VOCS
DRUGS

References

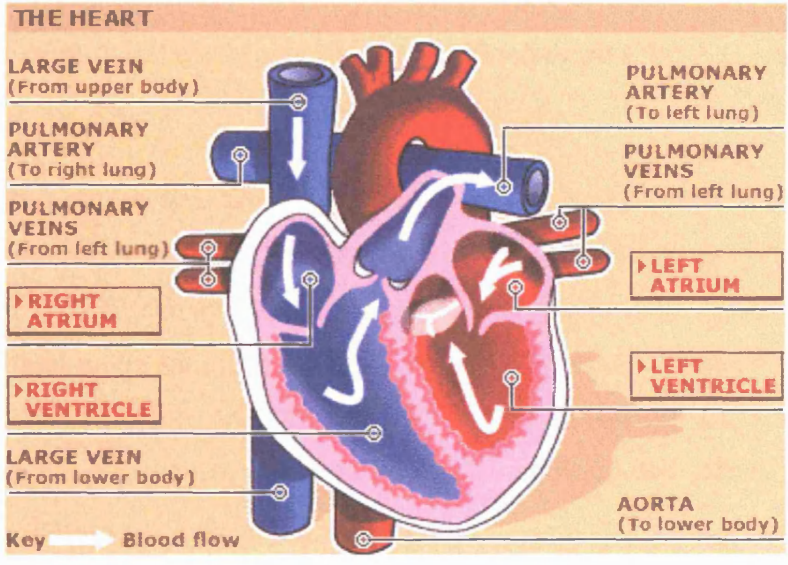
1. Bjorkman, S., Fyge, A. and Qi, Z (1996). J. Pharm. Sci, **85**, 887-889.
2. Bjorkman, S (2002). Pharmacy and Pharmacology, **54**, 1237-1245.
3. Haddad, S., Withey, J., Lapare, S., Law, F. and Krishnan, K (1998). Environmental Toxicology and Pharmacology, **5**, 245-255.
4. Csanady, G. A., Oberste-Frielinghaus, H. R., Semder, B., Baur, C., Schneider, K. T. and Filser, J. G (2002). Arch. Toxicology, **76**, 299-305.
5. Bonate, P. L., Swann, A. and Silverman, P. B (1996). J. Pharm. Sci, **85**, 878-883.
6. Abraham, M. H., Ibrahim, A., Zissimos, A. M., Zhao, Y. H., Comer, J. and Reynolds, D. P (2002). DDT, **7**, 1056-1063.
7. Platts, J. A., Abraham, M.H., Zhao, Y.H., Hersey, A., Ijaz, L. and Butina, D (2001). Eur. J. Med. Chem., **36**, 719-730.
8. Salminen, T., Pulli, A. and Taskinen, J (1997). J. Pharmaceutical and Biomedical Analysis, **15**, 469-477.

Chapter

23

Blood/plasma to heart distribution for drugs and VOCs combined

The distribution drugs between blood (or plasma) and heart tissue has become a matter of interest for the pharmaceutical industry, toxicological sciences and environmental health.



Scheme 23.0. Diagram showing the large muscular heart found in humans and rats.

The circulatory system is the body's main transport system, carrying food and oxygen to the cells and taking waste products away. Humans have a *double circulatory system* because it has two loops - one from the heart to the lungs and back, and another from the heart to the rest of the body and back. The heart is a

four-chambered muscular pump which pumps blood round the circulatory system. The right side of the heart pumps de-oxygenated blood to the lungs to pick up oxygen. The left side of the heart pumps the oxygenated blood from the lungs around the rest of the body. More pumping force is required for this much longer journey, which is why the left side of the heart has more muscular walls. Note that the heart has huge supply of blood from capillaries, arteries and veins. The muscular wall of the heart is where solutes can distribute between the two different phases at steady state concentration.

There have been a number of reports of data for drug blood (or plasma) to heart distribution in rats (*in vivo*).[1-15] Whereas, there have been no specific reviews reported yet that attempt to make models to predict blood (or plasma) to heart distribution either for rats or humans. The present work is restricted to a data set of 51 different drug compounds (54 data points), which is all that has been published. The *in vivo* drug data for blood (or plasma) to heart distribution of rats is at steady state concentration and the distribution from blood (or plasma) to heart, BH, is defined by the equation (23.0).

BH = (conc. of drug in heart)/(conc. of drug in blood or plasma) Equation 23.0

The purpose of this work is to make an attempt to construct equations to predict blood/plasma to heart distribution. The complete list of drug log BH data is shown in table 23.0 (see appendix); where two or more results are reported for the same compound in the same system (blood/plasma) they have been averaged.

As discussed before (chapter 21), the distribution sets for the tissues that also include heart can be combined together for plasma, blood and serum. In this case, the two distribution sets blood (log $P_{b\text{-}heart}$) and plasma (log $P_{p\text{-}heart}$) into heart from table 23.0 (see appendix) can now be combined together, but these ratios will not be averaged; thus the data points will be greater than the number of compounds.

Discussion on blood/plasma to heart distribution for drugs only

The 51 compounds (all of which are drugs) were used to construct the blood (or plasma) heart distribution model, see table 23.1. The structures for compounds with trivial names can be found in the appendix of this thesis. A first attempt to construct a model was to use the indicator variable for carboxylic acids (**Ia**). This descriptor has been used before for blood to brain distribution of drugs by Platt's[16] and Salminen.[17] The independent variable **Ia** for a value of '1' means the presence of carboxylic acid as a functional group to the drug compound. The independent variable **Ia** for a value of '0' means that the carboxylic acid is absent from the drug compound. This descriptor was used in this model to correlate and predict blood (or plasma) to heart distribution for drugs as shown in table 23.1.

EQUATIONS (rat model)	N(dp)	N(nc)	r^2	SD	F	OUTLIERS	D
FULL SET (*IN VIVO*)							
Log BH = 0.141 + 0.280**E** – 0.399**S** + 0.189**A** + 0.066**B** + 0.093**V** – 1.003**Ia**	54	51	0.728	0.279	20.975	0	6
TRAINING SET							
Log BH = 0.258 + 0.235**E** – 0.454**S** + 0.154**A** + 0.188**B** + 0.052**V** – 1.018**Ia**	40	40	0.769	0.287	18.334	-	6
TEST SET							
N(nc) = 12 **N**(dp) = 14							
SD (n-1) = 0.297 **RMSE** = 0.287							
AAE = 0.235 **AE** = 0.026							

N (dp) = number of data points used in the training set
N (nc) = number of compounds used in the training set
r^2 = is the square of the overall correlation coefficient
SD = the standard deviation
F = F-statistic
D = the number of descriptors used in the model

Table 23.1. Summary of the two equations for blood/plasma to heart distributions for drugs only.

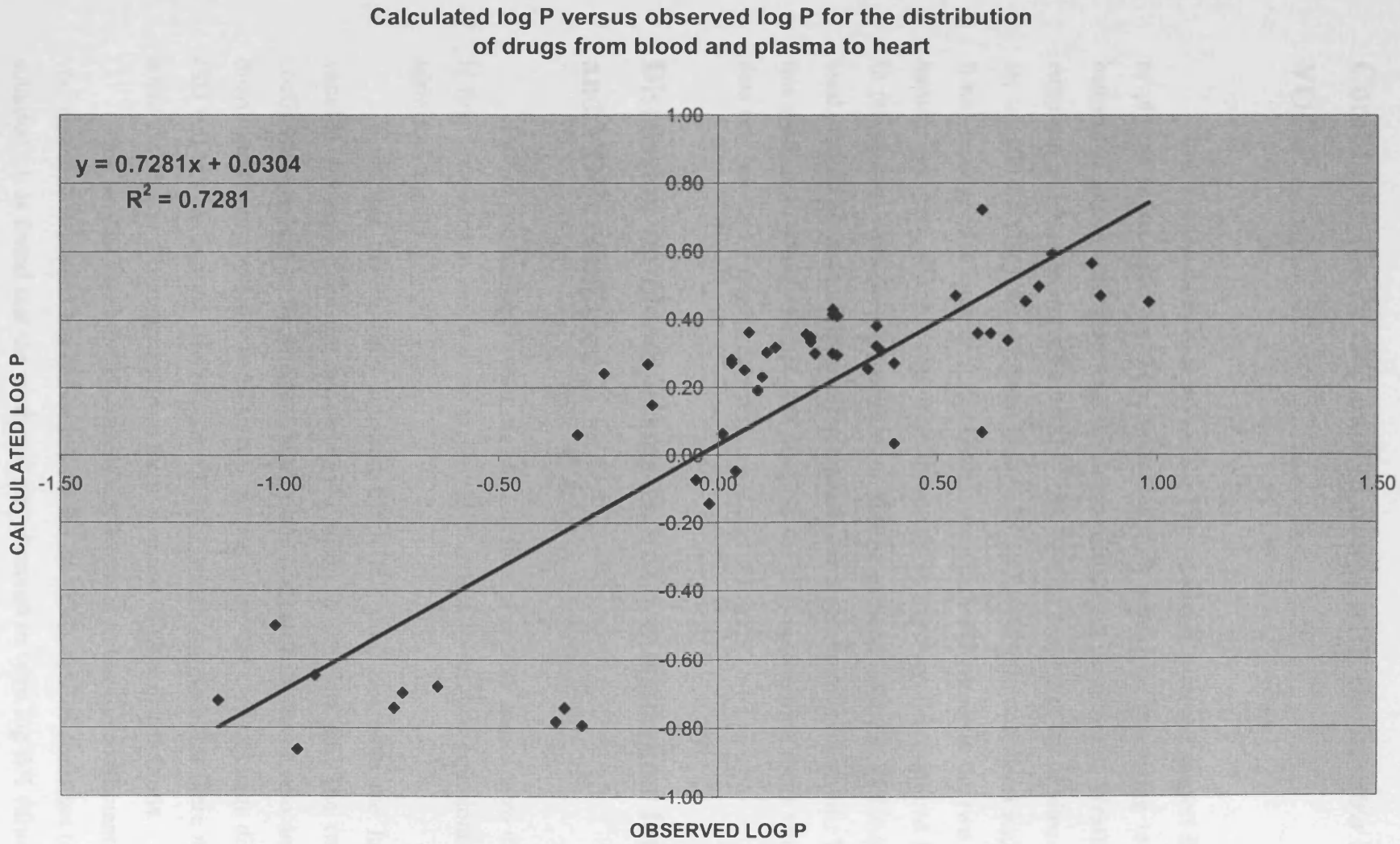
For the present model for blood/plasma to heart distribution, there were no marked outliers. The statistics of the equation for the full set in table 23.1. are satisfactory and the predictive capability of the equation can be assessed by the usual method of constructing a training set and an independent test set. The selection of the training (40 data points) and test sets (14 data points) was carried out using the usual computing method called Kennard-Stone. The training equation is also shown in table 23.1.

The statistics of the full set (table 23.1) and the training set (table 23.1), are similar and so are some of the coefficients. If the training set equation is used to predict the remainder of the data in the (unused) test set, it is found that AE = 0.03, AAE = 0.24 RMSE = 0.29 and SD = 0.30, there is very little bias, and the predictive capability of the training equation (table 23.1) and by implication the full set equation (table 23.1) is therefore assessed. Thus the full (*in vivo*) equation and the training equation can be used to predict further values of log BH close to the suggested experimental error of SD = 0.3 and AAE = 0.2 log units (from chapter 21).

The inter-correlations of descriptors used for drug blood/plasma to heart distribution are shown in table 23.2. The largest r^2 value for S versus B descriptor is 0.71. It appears that there is no strong inter-correlation between the other set of descriptors used in this model.

r	E	S	A	B	V	I
E	1.00					
S	0.83	1.00				
A	0.41	0.59	1.00			
B	0.70	0.84	0.58	1.00		
V	0.73	0.80	0.51	0.83	1.00	
I	0.05	0.13	0.40	-0.06	0.03	1.00

Table 23.2. The inter-correlation of descriptors for 51 compounds


Calculated log P versus observed log P for the distribution of drugs from blood and plasma to heart
y = 0.7281x + 0.0304
R2 = 0.7281
CALCULATED LOG P
OBSERVED LOG P
1.00
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
-0.80
-1.00
-1.50
-1.00
-0.50
0.00
0.50
1.00
1.50

Combining the *in vivo* data of drugs with the *in vitro* data of VOCs

Now that the data set of *in vivo* log BH values for passive transport from blood or plasma to rat heart has been investigated, it would be interesting to see what happens if this is combined with the previous data set (*in vitro)* of Weathersby and Abraham and Nussbaum and Hursh,[18-19] for VOCs. Note that the *in vitro* data set is for transfer of VOCs from human blood and rat blood to human heart and rat heart. It was concluded in chapter 9, that there is very little difference in the two species of humans and rats, and that data from these two species can be combined for VOCs. In the present work, the use of *in vitro* plasma to heart data for VOCs will not be used, mainly because of the small number of extra compounds involved. The aim of this work is to show whether it is possible (or not) to combine these two different data sets (*in vitro* + *in vivo* data set).

Discussion on blood/plasma to heart distribution for drugs and VOCs combined

Before combining the two models together of *in vivo* and *in vitro* data sets, it is first important to see if there are any differences in the two equations shown in table 23.1 and 23.3.

Note that the *in vitro* equation for VOCs does not have the Ia-indicator variable, because there are no carboxylic acids in the data set. The two sets of coefficients appear to be different visibly and suggest that the *in vivo* and *in vitro* distributions may possibly be different. In particular, the *c*-coefficients differ 0.141 (SD = 0.110) as against –0.410 (SD = 0.101), which suggests that there might be a small systematic difference between the *in vivo* and *in vitro* distributions.

One way to check for systematic deviations is to use the coefficients in the *in vivo* equation to predict log BH values for the 24 VOCs used to construct the *in vitro* equation. It is found that the predicted and observed *in vitro* log BH values, AE = -0.01, AAE = 0.18, RMSE = 0.23 and SD = 0.23 log units. The numerical values here show that the *in vivo* equation can predict the *in vitro* dataset within 0.23 log

units (SD). This suggests that the *in vivo* model can successfully predict the *in vitro* dataset as covers a larger descriptor space than the *in vitro* model. Any systematic difference between the *in vivo* (drugs) *in vitro* (VOCs) distributions could be due to the two series of compounds inhabiting different areas of chemical space.

Another indicator variable was then used to test if there is any difference in the *in vitro* and *in vivo* data set for VOCs and drugs. In this case, one can apply the indicator variable, **Iv**, defined as **Iv** = 1 for the *in vitro* data and **Iv** = 0 for the *in vivo* data. An equation with the **Iv** term was constructed but is not shown because the **Iv** term was not statistically significant (T-test = 0.355, p-value = 0.724) and the value of the **Iv** coefficient was quite small (0.034). The coefficients of drug (*in vivo*) and the drug and VOC (*in vivo* + *in vitro*) combined equation are the same within experimental error (SD values of the coefficients are in the drug equation: c=0.110, e=0.084, s=0.114, a=0.149, b=0.119, v=0.089, ia=0.120 and in the combined equation (*in vivo* + *in vitro*): c=0.071, e=0.075, s=0.100, a=0.141, b=0.105, v=0.079, ia=0.109). It is concluded that there is no systematic difference between the *in vivo* and *in vitro* distributions (as **Iv** is 0.0 log units), and that the two sets of distribution data can be combined without taking into account this difference (see equation 23.3).

The full data set (*in vivo* and *in vitro* measurements) consists of 74 compounds and 77 data points, the difference arising because blood/plasma to heart distributions are counted separately. The correlation equation of this set of data points with the Abraham descriptors is shown in table 23.3. The fits for this full equation (in blue) appear to be reasonable (S.D = 0.258 and r^2 = 0.689). This represents a good model for a combined data set of VOCs and drugs for blood/plasma to heart distribution, and is the only model yet reported for both types of data sets (*in vivo* rats + *in vitro* rats and humans combined).

There were one additional outlier observed from this full set and that is AI-6 (structure can be found in the appendix of this thesis). The residual for this compound is 0.7 log unit and may result from experimental error.

EQUATIONS	N(dp)	N(nc)	r^2	SD	F	OUTLIERS	D
FULL SET (*IN VIVO*) **Log BH** = 0.141 + 0.280**E** – 0.399**S** + 0.189**A** + 0.066**B** + 0.093**V** – 1.003**Ia**	54	51	0.728	0.279	20.975	0	6
FULL SET (*IN VITRO*) **Log BH** = -0.410 + 0.033**E** + 0.177**S** – 2.664**A** – 0.749**B** + 0.923**V**	24	24	0.822	0.145	16.619	2(nc)	5
FULL SET (*IN VIVO* + *IN VITRO*) **Log BH** = 0.041 + 0.209**E** – 0.393**S** + 0.076**A** – 0.046**B** + 0.279**V** – 0.965**Ia**	77	74	0.689	0.258	25.839	1(nc)	6
TRAINING SET **Log BH** = 0.201 + 0.179**E** – 0.397**S** + 0.080**A** + 0.138**B** + 0.115**V** – 0.992**Ia**	39	39	0.774	0.274	18.297	-	6
TEST SET **N**(nc) = 36 **N**(dp) = 38 **SD** (n-1) = 0.273 **RMSE** = 0.269 **AAE** = 0.219 **AE** = 0.024							

N (dp) = number of data points used in the training set
N (nc) = number of compounds used in the training set
r^2 = is the square of the overall correlation coefficient
SD = the standard deviation
F = F-statistic
D = the number of descriptors used in the model

Table 23.3. Summary of the two equations for blood/plasma to heart distributions for VOCs and drugs combined.

In order to assess the predictive capability of the combined equation (*in vivo* + *in vitro*), a training set was constructed and the equation for the training set used to predict log BH values for the test set, as shown in table 23.3. The coefficients and statistics of the full set of the combined equation (*in vivo* + *in vitro*) and training equation (*in vivo* + *in vitro*) are similar in terms of coefficients. The training equation predicted log BH values for the (unused) test set as follows. N(nc) = 36 and N(dp) = 38, that AE = 0.024, AAE = 0.219, RMSE = 0.269 and the SD = 0.273 log units. Thus the combined (*in vivo* + *in vitro*) training equation and the full combined equation (in blue) can be used to predict further values of log BH to close to our suggested experimental error of around 0.2 log units (AAE) or 0.3 log units (SD).

r	E	S	A	B	V	Ia
E	1.00					
S	0.91	1.00				
A	0.62	0.73	1.00			
B	0.84	0.91	0.72	1.00		
V	0.85	0.88	0.69	0.91	1.00	
Ia	0.22	0.27	0.48	0.15	0.19	1.00

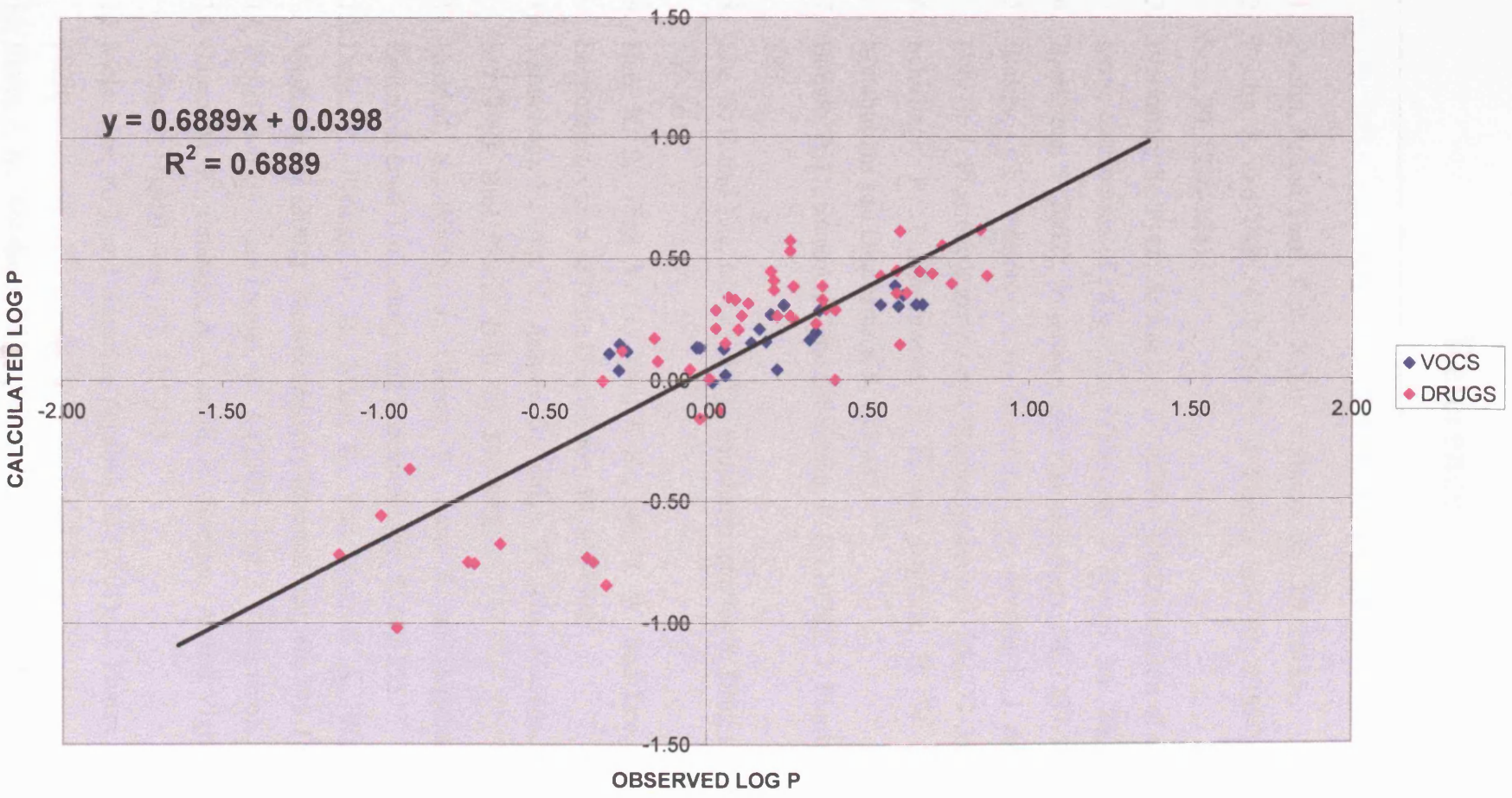
Table 23.4. The inter-correlation of descriptors for 74 compounds

The inter-correlations of descriptors for the combined *in vivo* and *in vitro* equation (from table 23.3) is shown by table 23.4. The largest r^2 value for E versus S descriptor, S versus B and B versus V is 0.83. It appears that there is no strong inter-correlation between the other set of descriptors used in this model.

The descriptors used in the equation reflect solute/solvent interactions. One can, therefore, interpret the coefficients in terms of the effect that particular interactions have on the process under consideration. The positive **e**- and **v**-coefficients in the blood/plasma to heart (*in vivo* + *in vitro*) equation from table 23.3, indicate that increasing molecular size and the presence of n- and Π-electrons pairs can push drugs out of blood/plasma and into the heart phase. The coefficient of the 'polar' descriptors **S** is negative, indicating that polarity is stronger in the blood/plasma phase than in the heart tissue phase. Therefore the drug compounds will move from the heart phase and into the blood/plasma phase. The **a**-coefficients (P = 0.59) and **b**-coefficient (P = 0.66) is not significant. The large negative coefficient of the indicator variable (***Ia***), in the combined equation from table 23.3., indicates a hitherto unknown factor affecting the transfer of solutes between blood/plasma and the heart. It is apparent that the presence of $-CO_2H$ group acts to hinder heart penetration further. This could be simply due to intrinsic hydrogen bonding and the polarity properties of neutral acids. Finally, it is known that some acidic drugs may bind to albumin present in plasma and blood, which may also account for this negative contribution of the coefficient ***Ia***.

This new model based on 74 compounds suggests that it can now be used to correlate and predict blood/plasma to heart distribution for *in vitro* and *in vivo* drug compounds. Of course, any prediction using the full combined equation (*in vivo* plus

in vitro) can, or should, only be made within the chemical space of the test set used to assess the predictions.


Calculated log P versus observed log P for the
distribution of compouds (VOCs+ drugs) from blood and plasma to heart
y = 0.6889x + 0.0398
R2 = 0.6889
CALCULATED LOG P
OBSERVED LOG P
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-2.00
-1.50
-1.00
-0.50
0.00
0.50
1.00
1.50
2.00
VOCS
DRUGS

References

1. Poulin, P. and Theil, F. P (2000). J. Pharm. Sci, **89**, 16-35.
2. Poulin, P. and Theil, F. P (2002). J. Pharm. Sci., **91**, 1358-1367. J. Pharm. Sci., **90**, 1226-1241.
3. Bjorkman, S., Fyge, A. and Qi, Z (1996). Determination of the steady state tissue distribution of midazolam in the rats. J. Pharm. Sci., **85**, 887-889.
4. Bjorkman, S (2002). Pharmacy and Pharmacology, **54**, 1237-1245.
5. Blakey, G. E., Nestorov, I. A., Arundel, P. A., Aarons, L. J. and Rowland, M (1997). J. Pharmacokinetics and Biopharmaceutics, **25**, 277-312.
6. Schillings, R. T., Sisenwine, S. F. and Ruelius, H. W (1977). Drug Metabolism and Disposition, **5**, 425-435.
7. Bonate, P. L., Swann, A. and Silverman, P. B (1996). J. Pharm. Sci, **85**, 878-883.
8. Lee, S. H and Lee, M. G (1995). Biopharmaceutics & Drug Disposition, **16**, 547-561.
9. Han, K. S., Kim, Y. G., Yoo, J. K., Lee, J. W. and Lee, M. G (1998). Biopharmaceutics & Drug Disposition, **19**, 493-500.
10. Yamamoto, F., Oka, H., Antoku, S., Ichiya, Yu-ichi., Masuda, K. and Maeda, M (1999). Biol. Pharm. Bull., **22**, 590-597.
11. Haddad, S., Withey, J., Lapare, S., Law, F. and Krishnan, K (1998). Environmental Toxicology and Pharmacology, **5**, 245-255.
12. Doze, P., Elsinga, P. H., Maas, B., Van Waarde, A., Wegman, T. and Vaalburg, W (2002). Neurochemistry International, **40**, 145-155.
13. Wojcikowski, J. and Daniel, W. A (2002). Pol. J. Pharmacol., **54**, 647-654.
14. Gasco, M. R., Fundaro, A., Cavalli, R., Bargoni, A. and Vighetto, D (2000). Pharmacological. Res., **42**, 337-343.
15. Mclachlan, A. J. and Hosseini-Yegabeh, M (2001). J. Pharm. Sci., **90**, 1817-1828.
16. Platts, J. A., Abraham, M.H., Zhao, Y.H., Hersey, A., Ijaz, L. and Butina, D (2001). Eur. J. Med. Chem., **36**, 719-730.

17. Salminen, T., Pulli, A. and Taskinen, J (1997). J. Pharmaceutical and Biomedical Analysis, **15**, 469-477.
18. Abraham, M. H. and Weathersby, P. K (1994). J. Pharm. Sci, **83**, 1450-1455.
19. Nussbaum, E. and Hursh, J. B (1957). Science, **125**, 552-553.

Chapter

24

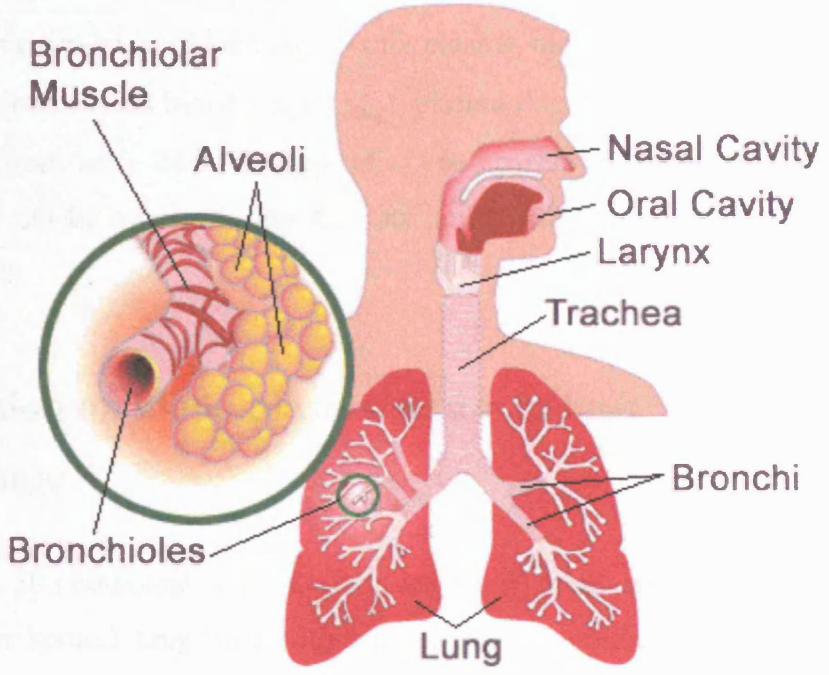
Blood/plasma/serum to lung distribution for drugs and VOCs combined

The distribution of drugs between blood (plasma or serum) and lung tissue has become a matter of interest for the pharmaceutical industry and environmental health.

Lungs are paired organs in the chest that carry on respiration. In the adult human, each lung is 25 to 30 cm (10 to 12 in) long and roughly conical. The two lungs are separated by a structure called the mediastinum, which contains the heart, trachea, oesophagus, and blood vessels. They are covered by a protective membrane called the pulmonary pleura, which is separated from the parietal pleura a similar membrane on the chest wall by a lubricating fluid. Inhaled air passes through the trachea, which divides into two tubes called bronchi; each bronchus leads to one lung. Within the lungs the bronchi subdivide into bronchioles, which give rise to alveolar ducts; these end in sacs called alveoli.

In this life-supporting process, oxygen from incoming air enters the blood and carbon dioxide, a waste gas from the metabolism of food, is exhaled into the atmosphere. The exchange of gases takes place when air reaches the alveoli. These small sacs are only one cell thick, and they are surrounded by a rich supply of blood capillaries that are also only one cell thick. Air diffuses through these cells into the capillary blood, which carries the oxygen-rich air or drugs to the heart to be distributed throughout the body. In the alveoli, at the same time, gaseous carbon dioxide diffuses from the blood into the lung and is expired. From the capillary

network, there are also distribution of drugs or VOCS between blood and the lung tissue phase at steady-state.



Scheme 24.0. Diagram showing the lung tissue present in humans

There have been a number of reports of data for drug blood (plasma or serum) to lung distribution in rats (*in vivo*).[1-21] There have been no specific reviews reported yet that attempt to make models to predict blood (plasma or serum) to lung distribution either for rats or humans. The present work is restricted to a data set of 50 different drug compounds (53 data points), which is all that has been published. The drug data (*in vivo*) for blood (plasma or serum) to lung distribution of rats is at steady state concentration and the distribution from blood to lung, BL, is defined by the following equation (24.0).

BL = (conc. of drug in lung)/(conc. of drug in blood/plasma/serum) Equation 24.0

The purpose of this work is to collect as many existing literature data on distribution (as log BL values) and to make an attempt to construct equations to predict *in vivo* blood/plasma/serum to lung distribution. The complete list of *in vivo*

drug log BL data is shown in table 24.0 (see appendix); where two or more results are reported for the same compound in the same system (blood/plasma/serum) they have been averaged.

As discussed before (chapter 21), the distribution sets for the tissues that also include lung can be combined together for plasma, blood and serum. In this case, the three distribution sets blood (log $P_{b\text{-}lung}$), plasma (log $P_{p\text{-}lung}$) and serum (log $P_{s\text{-}lung}$) into lung from table 24.0 (see appendix) can now be combined together, but these ratios will not be averaged; thus the data points will be greater than the number of compounds.

Discussion on blood/plasma/serum to lung distribution for drugs only

The 50 compounds (all of which are drugs) were used to construct the blood (plasma or serum) lung distribution model are in table 24.1. The structures for compounds with trivial names can be found in the appendix of this thesis. A first attempt to construct a model was to use the indicator variable for carboxylic acids (**Ia**). This descriptor has been used before for blood to brain distribution of drugs by Platt's[22] and Salminen.[23] The independent variable **Ia** for a value of '1' means the presence of carboxylic acid as a functional group to the drug compound. The independent variable **Ia** for a value of '0' means that the carboxylic acid is absent from the drug compound. This descriptor was used in a model to correlate and predict blood (plasma or serum) to lung distribution for drugs as shown in table 24.1.

EQUATIONS (rat model)	N(dp)	N(nc)	r^2	SD	F	OUTLIERS	D
FULL SET							
Log BL = – 0.012 + 0.031**E** – 0.174**S** – 0.666**A** + 0.248**B** + 0.321**V** – 0.833**Ia**	53	50	0.749	0.351	22.846	3(nc)	6
TRAINING SET							
Log BL = 0.093 + 0.026**E** – 0.217**S** – 0.603**A** + 0.243**B** + 0.321**V** – 0.882**Ia**	40	40	0.770	0.381	18.439	-	6
TEST SET							
N(nc) = 13 **N**(dp) = 13							
SD (n-1) = 0.291 **RMSE** = 0.280							
AAE = 0.238 **AE** = 0.145							

N (dp) = number of data points used in the training set
N (nc) = number of compounds used in the training set
r^2 = is the square of the overall correlation coefficient
SD = the standard deviation
F = F-statistic
D = the number of descriptors used in the model

Table 24.1 Summary of the two equations for blood/plasma/serum to lung distributions for drugs only.

For the present model for blood (plasma or serum) to lung distribution, there were three compounds that were marked outliers: biperiden, pentazocine and propranolol. The residual for these compounds is 0.8, 0.8 and 1.0 log unit, and may result from experimental error. The statistics of the equation for the full set is shown in table 24.1., are good and the predictive capability of the equation can be assessed by the usual method of constructing a training set (40 data points) and an independent test set (13 data points). The selection of the training and test sets was carried out using the usual computing method called Kennard-Stone. The training equation constructed is also shown in table 24.1.

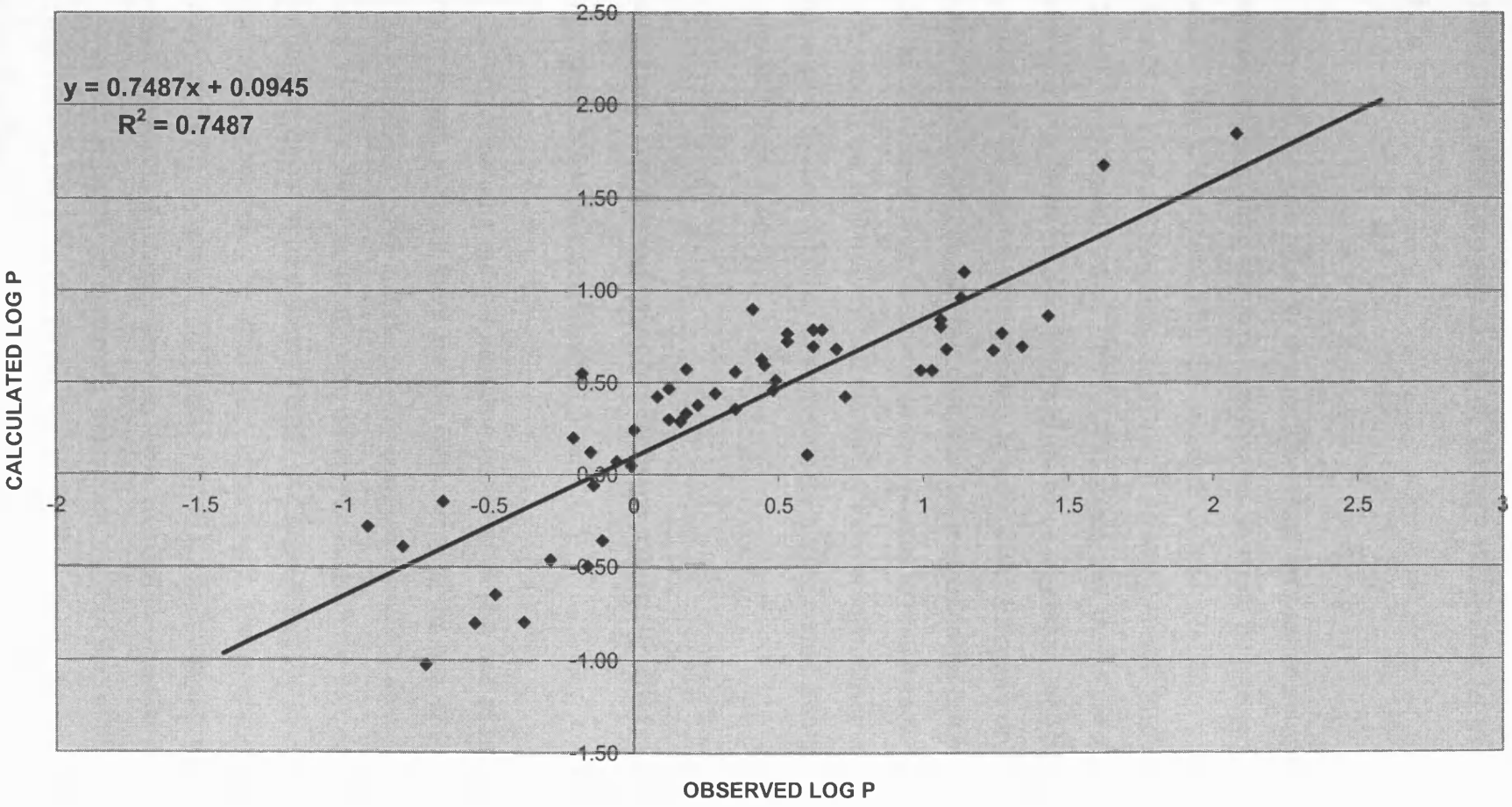
The statistics of the full set (table 24.1) and the training set (table 24.1), are similar and so are some of the coefficients. If the training set equation is used to predict the remainder of the data in the (unused) test set, it is found that AE = 0.15, AAE = 0.24 RMSE = 0.28 and SD = 0.29, there is very little bias, and the predictive capability of the training equation (table 24.1) and by implication of the full set equation (table 24.1) is therefore assessed. Thus the full (*in vivo*) equation and the training equation can be used to predict further values of log BL close to the suggested experimental error of SD = 0.3 and AAE = 0.2 log units (from chapter 21).

The inter-correlations of descriptors used for drug blood (plasma or serum) to lung distribution are shown in table 24.2. The largest r^2 value for B versus V descriptor is 0.76. It appears that there is no strong inter-correlation between the other set of descriptors used in this model.

r	E	S	A	B	V	I
E	1.00					
S	0.82	1.00				
A	0.13	0.44	1.00			
B	0.50	0.76	0.51	1.00		
V	0.53	0.73	0.37	0.87	1.00	
I	0.06	0.12	0.28	-0.08	-0.02	1.00

Table 24.2. The inter-correlation of descriptors for 50 compounds

Calculated log P versus observed log P for the distribution of drugs from blood, plasma and serum to lung



Combining the *in vivo* data of drugs with the *in vitro* data of VOCs

Now that the data set of *in vivo* log BL values for passive transport from blood /plasma/serum to rat lung has been investigated, it would be interesting to see what happens if this is combined with the previous data set *(in vitro)* of Abraham and Weathersby,[24] Fiserova-Bergerova and Diaz,[25] Sweeney *et al*,[26] Filser et al,[27] Knaak and Smith[28] and Csandy *et al.*[29] for VOCs. Note that the *in vitro* data set is for transfer of VOCs from human blood and rat blood to human lung and rat lung. It was concluded in chapter 9, that there is very little difference in data from the species of humans and rats, and that data from these two species can be combined for VOCs. In the present work, the use of *in vitro* plasma to lung data for VOCs will not be used, mainly because of the small number of extra compounds involved. The aim of this work is to show whether it is possible (or not) to combine these two different data sets (*in vitro* + *in vivo*).

Discussion on blood/plasma/serum to lung distribution for drugs and VOCs combined

Before combining the two models together of *in vivo* and *in vitro* data sets, it is first important to see if there are any differences in the two equations shown in table 24.1 and 24.3. Note that the *in vitro* equation for VOCs does not have the Ia-indicator variable, because there are no carboxylic acids in the data set. The two sets of coefficients appear to be different visibly and suggest that the *in vivo* and *in vitro* distributions may possibly be different. In particular, the *c*-coefficients differ – 0.012 (SD = 0.143) as against – 0.114 (SD = 0.091), which suggests that there might be a small systematic difference between the *in vivo* and *in vitro* distributions.

One way to check for systematic deviations is to use the coefficients in the *in vivo* equation to predict log BL values for the 43 VOCs used to construct the *in vitro* equation. It is found that for the predicted and observed *in vitro* log BL values, AE = 0.153, AAE = 0.190, RMSE = 0.249 and SD = 0.252 log units. The numerical values here show that the *in vivo* equation can predict the *in vitro* dataset within 0.26

log units (SD). This suggests that the *in vivo* model can successfully predict the *in vitro* dataset as it covers a larger descriptor space than the *in vitro* model. Any systematic difference between the *in vivo* (drugs) and the *in vitro* (VOCs) distributions could be due to the two series of compounds inhabiting different areas of chemical space.

Another indicator variable was then used to test if there is any difference in the *in vitro* and *in vivo* data set for VOCs and drugs. In this case one can apply the indicator variable, **Iv**, defined as **Iv** = 1 for the *in vitro* data and **Iv** = 0 for the *in vivo* data. An equation with the **Iv** term was constructed but is not shown, because the **Iv** term was not statistically significant (T-test = -2.271, p-value = 0.026) and the value of the **Iv** coefficient was quite small (-0.23). The coefficients of drug (*in vivo*) and the drug and VOC (*in vivo* + *in vitro*) combined equation are the same within experimental error (SD values of the coefficients are in the drug equation: c=0.143, e=0.122, s=0.179, a=0.189 b=0.137, v=0.107, ia=0.145, and in the combined equation (*in vivo* + *in vitro*): c=0.058, e=0.078, s=0.119, a=0.141, b=0.097, v=0.072, ia=0.115). It is concluded that there is a very small systematic difference between the *in vivo* and *in vitro* distributions (as **Iv** is -0.2 log units), and that the two sets of distribution data can be combined without taking into account this small difference (see equation 24.3).

The full data set (*in vivo* and *in vitro* measurements) consists of 91 compounds and 94 data points, the difference arising because blood/plasma/serum to lung distributions are counted separately. The correlation of this set of data points with the Abraham descriptors is shown in table 24.3. The fits for this full equation (in blue) appear to be good (S.D = 0.272 and $r^2 = 0.760$). This represents a good model for a combined data set of VOCs and drugs for blood/plasma/serum to lung distribution, and is the only model yet reported for both types of data sets (*in vivo* rats + *in vitro* rats and humans combined).

There were two additional outliers observed in the full set; cotinine and dicloxacillin. The residual for these compounds is 0.7 and 0.7 log unit and may result from experimental error.

EQUATIONS	N(dp)	N(nc)	r^2	SD	F	OUTLIERS	D
FULL SET (*IN VIVO*) **Log BL** = – 0.012 + 0.031**E** – 0.174**S** – 0.666**A** + 0.248**B** + 0.321**V** – 0.833**Ia**	53	50	0.749	0.351	22.846	3(nc)	6
FULL SET (*IN VITRO*) **Log BL** = – 0.114 – 0.193**E** + 0.004**S** – 0.279**A** – 0.206**B** + 0.256**V**	43	43	0.275	0.194	2.802	0	5
FULL SET (*IN VIVO + IN VITRO*) **Log BL** = -0.157 – 0.004**E** – 0.036**S** – 0.742**A** + 0.252**B** + 0.312**V** – 0.734**Ia**	94	91	0.760	0.272	45.925	2(nc)	6
TRAINING SET **Log BL** = -0.067 – 0.010**E** – 0.033**S** – 0.683**A** + 0.279**B** + 0.253**V** – 0.774**Ia**	47	47	0.801	0.314	26.812	-	6
TEST SET **N**(nc) = 46 **N**(dp) = 47 **SD** (n-1) = 0.247 **RMSE** = 0.244 **AAE** = 0.201 **AE** = 0.080							

N (dp) = number of data points used in the training set
N (nc) = number of compounds used in the training set
r^2 = is the square of the overall correlation coefficient
SD = the standard deviation
F = F-statistic
D = the number of descriptors used in the model

Table 24.3. Summary of the two equations for blood/plasma/serum to lung distributions for VOCs and drugs combined.

In order to assess the predictive capability of the combined equation (*in vivo + in vitro*) a training set was constructed and the equation for the training set used to predict log BL values for the test set, as shown in table 24.3. The coefficients and statistics of the full set of the combined equation (*in vivo + in vitro*) and training equation (*in vivo + in vitro*) are similar in terms of coefficients. The training equation predicted log BL values for the (unused) test set as follows. N(nc) = 46 and N(dp) = 47, AE = 0.080, AAE = 0.201, RMSE = 0.244 and SD = 0.247 log units. Thus the combined (*in vivo + in vitro*) training equation and the full combined equation (in blue) can be used to predict further values of log BL to close to our suggested experimental error of around 0.2 log units (AAE) or 0.3 log units (SD).

r	E	S	A	B	V	I
E	1.00					
S	0.92	1.00				
A	0.51	0.67	1.00			
B	0.77	0.88	0.71	1.00		
V	0.78	0.86	0.60	0.92	1.00	
I	0.21	0.25	0.35	0.13	0.16	1.00

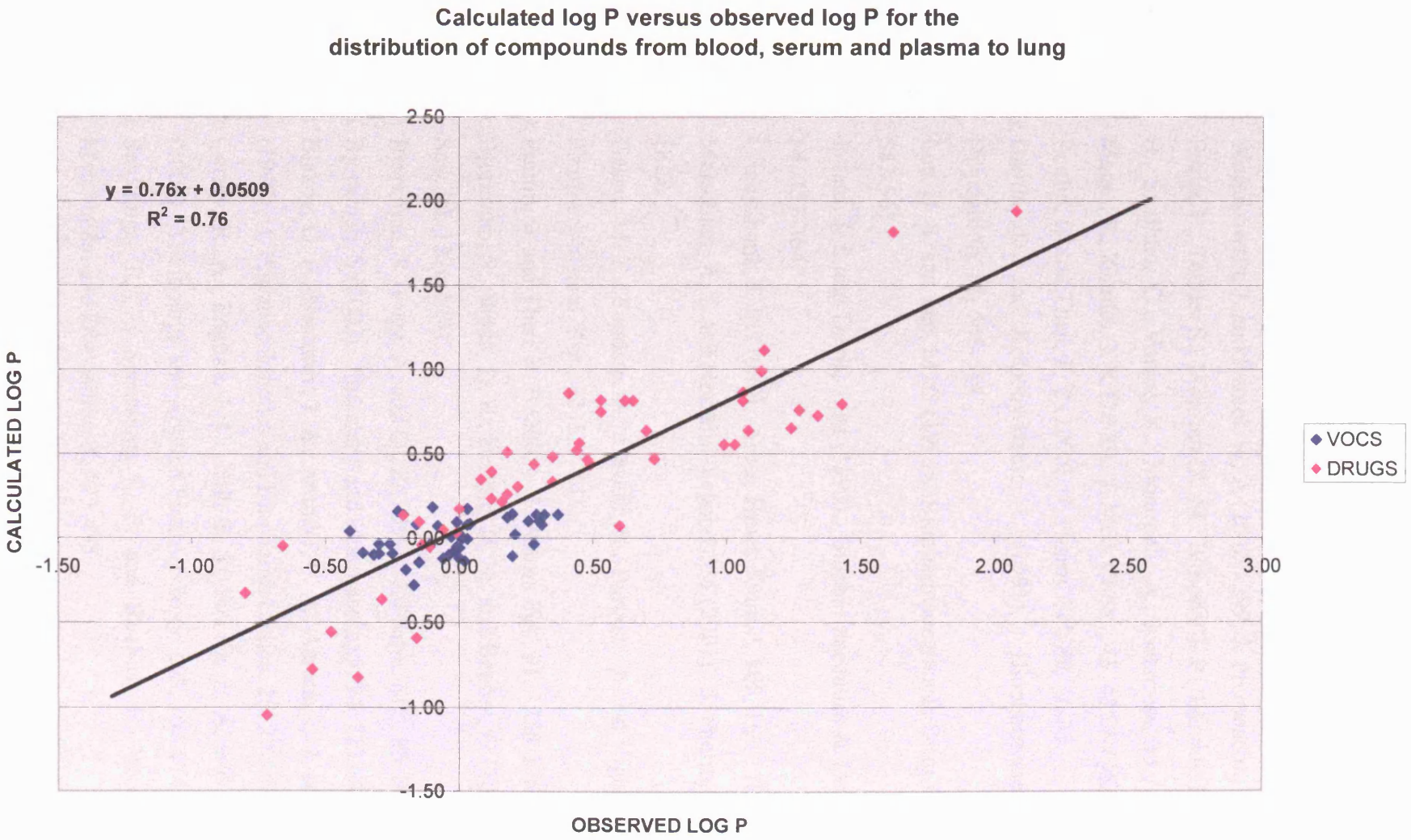
Table 24.4. The inter-correlation of descriptors for 91 compounds

The inter-correlations of descriptors obtained for the combined *in vivo* and *in vitro* equation (from table 24.3) as shown by table 24.4. The largest r^2 value for E versus S and B versus V descriptor is 0.85. It appears that there is no strong inter-correlation between the other set of descriptors used in this model.

The descriptors used in the equation reflect solute/solvent interactions. One can, therefore, interpret the coefficients in terms of the effect that particular interactions have on the process under consideration. The **e**-coefficients (P = 0.96) and **s**-coefficient (P = 0.76) is not significant. The positive **v**-coefficient in the blood/plasma/serum to lung (*in vivo* + *in vitro*) equation, indicate that increasing molecular size can push drugs out of blood/plasma/serum and into the lung phase. The coefficient of the **A** descriptors are negative, indicating that the hydrogen-bond acidity acts to keep drug compounds in the blood/plasma/serum phase and out of the lung tissue phase. In turn, this indicates that the lung tissue is less basic than is the blood/plasma/serum phase, so the solute will move from the lung tissue into the blood/plasma/serum phase. The coefficients of the descriptor **B** are positive, indicating that the hydrogen-bond basicity acts to keep drug compounds in the lung and out of the blood/plasma/serum phase. In turn, this indicates that the blood/plasma/serum is less acidic than is the lung phase, but is more able to accommodate bulky solutes and interact with them via dispersive forces than in the lung phase. The large negative coefficient of the indicator variable (***Ia***), in the combined equation from table 24.3., indicates a hitherto unknown factor affecting the transfer of solutes between blood/plasma/serum and the lung. It is apparent that the presence of $-CO_2H$ group acts to hinder lung penetration further. This could be simply due to intrinsic hydrogen bonding and the polarity properties of neutral acids.

Finally, it is known that some acidic drugs may bind to albumin present in plasma, blood and serum, which may also account for this negative contribution of the coefficient ***Ia***.

This new model based on 91 compounds suggest that it can now be used to correlate and predict blood (plasma or serum) to lung distribution for *in vitro* and *in vivo* drug compounds. Of course, any prediction using the full combined equation (*in vivo* plus *in vitro*) can, or should, only be made within the chemical space of the test set used to assess the predictions.


Calculated log P versus observed log P for the distribution of compounds from blood, serum and plasma to lung
y = 0.76x + 0.0509
R2 = 0.76
CALCULATED LOG P
OBSERVED LOG P
2.50
2.00
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-1.50
-1.00
-0.50
0.00
0.50
1.00
1.50
2.00
2.50
3.00
VOCS
DRUGS

References

1. Wojcikowski, J. and Daniel, W. A (2002). Pol. J. Pharmacol., **54**, 647-654.
2. Fazio, F., Todde, S., Moresco, R. M., Simonelli, P., Baraldi, P. G., Cacciari, B., Spalluto, G., Varani, K., Monopoli, A., Matarrese, M., Carpinelli, A., Magni, F., Kienle, G. K (2000). J. Med. Chem., **43**, 4359-4362.
3. Poulin, P. and Theil, F. P (2000). J. Pharm. Sci, **89**, 16-35.
4. Davila, D. and Kolacny-Babic, L (1991). Biopharmaceutics & Drug Disposition, **12**, 505-514.
5. Lee, H. S. and Lee, M. G (1995). Biopharmaceutics & Drug Disposition, **16**, 547-561
6. Javaid, J. I. and Davis, J. M (1993). Biopharmaceutics & Drug Disposition, **14**, 357-364.
7. Cruickshank, J. M (1980). Amer. Heart. Journal, **100**, 160-178.
8. Mclachlan, A. J. and Hosseini-Yegabeh, M (2001). J. Pharm. Sci., **90**, 1817-1828.
9. Gasco, M. R., Fundaro, A., Cavalli, R., Bargoni, A.and Vighetto, D (2000). Pharmacological. Res., **42**, 337-343.
10. Poulin, P. and Theil, F. P (2002). J. Pharm. Sci., **91**, 1358-1367.
11. Bjorkman, S., Wada, D. R., Berling, B. M. and Benoni, G (2001). J. Pharm. Sci., **90**, 1226-1241.
12. Bjorkman, S., Fyge, A. and Qi, Z (1996). J. Pharm. Sci., **85**, 887-889.
13. Bjorkman, S (2002). Pharmacy and Pharmacology, **54**, 1237-1245.
14. Blakey, G. E., Nestorov, I. A., Arundel, P. A., Aarons, L. J. and Rowland, M (1997). J. Pharmacokinetics and Biopharmaceutics, **25**, 277-312.
15. Corley, R. A., English, J. C., Hill, T. S., Fiorcia, L. A. and Morgott, D. A (2000). Toxicology and Applied Pharmacology, **165**, 163-174.
16. Schillings, R. T., Sisenwine, S. F. and Ruelius, H. W (1997). Drug Metabolism and Disposition, **5**, 425-435.
17. Aravagiri, M., Teper, Y. and Marder, S. R (1999). Biopharmaceutics & Drug Disposition, **20**, 369-377.
18. Han, K. S., Kim, Y. G., Yoo, J. K., Lee, J. W. and Lee, M. G (1998). Biopharmaceutics & Drug Disposition, **19**, 493-500.

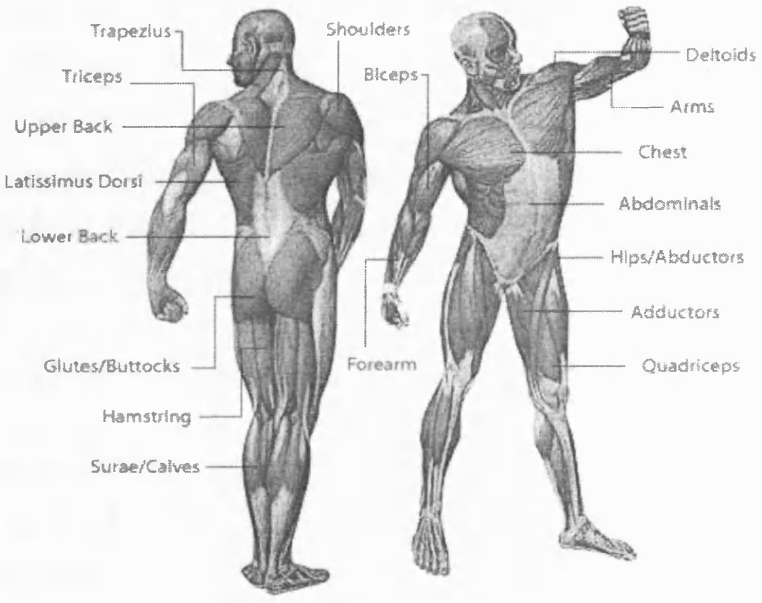
19. Haddad, S., Withey, J., Lapare, S., Law, F. and Krishnan, K (1998). Environmental Toxicology and Pharmacology, **5**, 245-255.
20. Bonate, P. L., Swann, A. and Silverman, P. B (1996). J. Pharm. Sci., **85**, 878-883.
21. Doze, P., Elsinga, P. H., Maas, B., Van Waarde, A., Wegman, T. and Vaalburg, W (2002). Neurochemistry International, **40**, 145-155.
22. Platts, J. A., Abraham, M.H., Zhao, Y.H., Hersey, A., Ijaz, L. and Butina, D (2001). Eur. J. Med. Chem., **36**, 719-730.
23. Salminen, T., Pulli, A. and Taskinen, J (1997). J. Pharmaceutical and Biomedical Analysis, **15**, 469-477.
24. Abraham, M. H. and Weathersby, P. K (1994). J. Pharm. Sci, **83**, 1450-1455.
25. Fiserova-Bergerova, V. and Diaz, M. L (1986). International Archives of Occupational and Environmental Health, **58**, 75-87.
26. Sweeney, L. M., Himmelstein, M. W. and Gargas, M. L (2001). Chemical-Biological Interactions, **135-136**, 303-322.
27. Filser, J. G., Schmidbauer, R., Rampf, F., Baur, C. M., Putz, C. and Csanady, G. A (2000). Toxicology and Applied Pharmacology, **169**, 40-51.
28. Knaak, J. B. and Smith, L. W (1998). Inhalation Toxicology, **10**, 65-85.
29. Csanady, Gy. A., Denk, B., Putz, C., Kreuzer, P. E., Kessler, W., Baur, C., Gargas, M. L. and Filser, J. G (2000). Toxicology and Applied Pharmacology, **165**, 1-26.

Chapter

25

Blood/plasma to muscle distribution for drugs and VOCs combined

The distribution of drugs between blood (or plasma) and muscle tissue has become a matter of interest for the pharmaceutical industry, toxicological sciences, environmental health, sports and forensic sciences.



Scheme 25.0. Diagram showing the muscle tissue of the human body

The skeletal muscles of humans and rats are characterized by the ability to contract, usually in response to a stimulus from the nervous system. Muscles are made up of long cells called fibres that contract, or shorten, to move parts of the body. Most of the muscular system of up to half of the body weight of humans

consists of skeletal muscle. More than 650 skeletal muscles cover the human skeleton in layers, give the body its shape, and are attached to, and pull on to the bones. By contracting they make the human body run, jump, or perform any one of the thousands of movements. Two other muscle types work unseen inside the body: cardiac muscle powers the heart beat, while the smooth muscle moves food and other materials around the body.

Skeletal muscle is supplied with nerves from the central nervous system and because it is partly under conscious control, it is also called voluntary muscle. Most skeletal muscle is attached to portions of the skeleton by connective-tissue attachments called tendons. Contractions of skeletal muscle serve to move the various bones and cartilages of the skeleton. Skeletal muscle forms most of the underlying flesh of humans and rats. The skeletal muscle is where solutes can distribute between the blood capillaries and muscle phase at steady state concentration.

There have been a number of reports of data for drug blood (or plasma) to muscle distribution in rats (*in vivo*).[1-17] However, there have been no specific reviews reported that attempt to obtain models to predict blood (or plasma) to muscle distribution either for rats or humans. The present work is restricted to a data set of 59 different drug compounds (60 data points), which is all that has been published. The *in vivo* drug data for blood (or plasma) to muscle distribution of rats is at steady state concentration and the distribution from blood (or plasma) to muscle, BM, is defined by the equation (25.0).

BM = (conc. of drug in muscle)/(conc. of drug in blood or plasma) Equation 25.0

The purpose of this work is to collect as much existing literature data on distribution (as log BM values) and to make an attempt to construct equations to predict blood/plasma to muscle distribution. The complete list of *in vivo* drug log BM data is shown in table 25.0 (see appendix); where two or more results are reported for the same compound in the same system (blood/plasma) they have been averaged.

As discussed before (chapter 21), the distribution sets for the tissues that also include muscle can be combined together for blood, plasma and serum. In this case, the two distribution sets blood (log $P_{b\text{-}muscle}$) and plasma (log $P_{p\text{-}muscle}$) into muscle

from table 25.0 (see appendix) can now be combined together, but these ratios will not be averaged; thus the data points will be greater than the number of compounds.

Discussion on blood/plasma to muscle distribution for drugs only

The 59 compounds (all of which are drugs) were used to construct the blood to muscle distribution model, see table 25.1. The structures for compounds with trivial names can be found in the appendix of this thesis. A first attempt to construct a model was to use the indicator variable for carboxylic acids (**Ia**). This descriptor has been used before for blood to brain distribution of drugs by Platt's[18] and Salminen.[19] The independent variable **Ia** for a value of '1' means the presence of carboxylic acid as a functional group to the drug compound. The independent variable **Ia** for a value of '0' means that the carboxylic acid is absent from the drug compound. This descriptor was used in this model to correlate and predict blood (or plasma) to muscle distribution for drugs as shown in table 25.1.

EQUATIONS	N(dp)	N(nc)	r^2	SD	F	OUTLIERS	D
FULL SET (*IN VIVO*)							
Log BM = 0.082 – 0.059**E** + 0.010**S** – 0.248**A** + 0.028**B** + 0.110**V** – 1.022**Ia**	60	59	0.745	0.253	25.850	0	6
TRAINING SET							
Log BM = 0.179 – 0.093**E** – 0.022**S** – 0.206**A** + 0.086**B** + 0.086**V** – 1.054**Ia**	40	40	0.777	0.285	19.139	-	6
TEST SET							
N(nc) = 20 **N**(dp) = 20							
SD (n-1) = 0.213 **RMSE** = 0.208							
AAE = 0.169 **AE** = 0.051							

N (dp) = number of data points used in the training set
N (nc) = number of compounds used in the training set
r^2 = is the square of the overall correlation coefficient
SD = the standard deviation
F = F-statistic
D = the number of descriptors used in the model

Table 25.1. Summary of the two equations for blood/plasma to muscle distributions for drugs only.

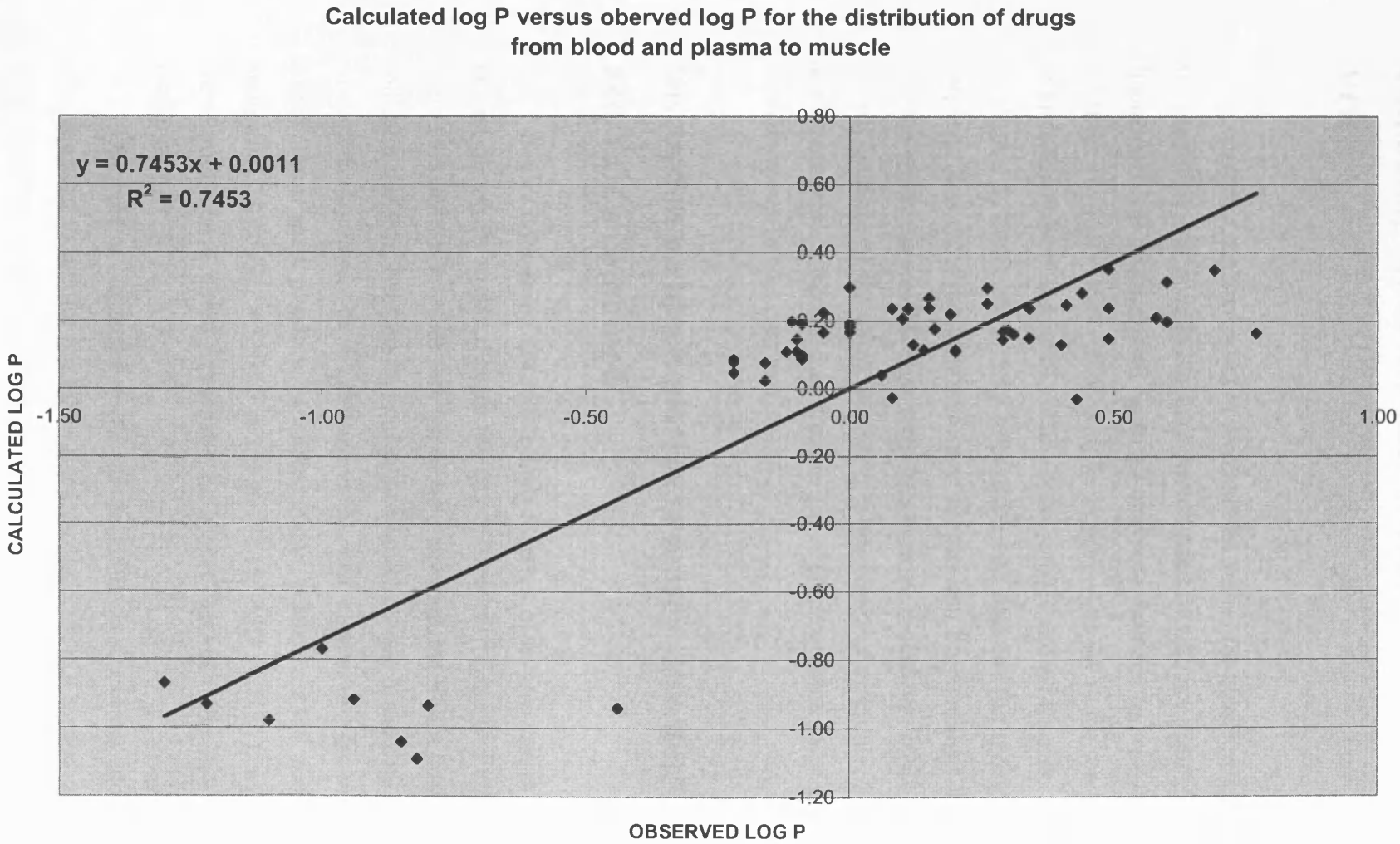
For the present model for blood or plasma to muscle distribution, there were no marked outliers. The statistics of the equation for the full set in table 25.1. are good and the predictive capability of the equation can be assessed by the usual method of constructing a training set and an independent test set. The selection of the training (40 data points) and test sets (20 data points) was carried out using the usual computing method called Kennard-Stone. The training equation is also shown in table 25.1.

The statistics of the full set (table 25.1) and the training set (table 25.1), are similar and so are some of the coefficients. If the training set equation is used to predict the remainder of the data in the (unused) test set, it is found that AE = 0.05, AAE = 0.17, RMSE = 0.21 and SD = 0.21, that there is very little bias, and the predictive capability of the training equation (table 25.1) and by implication of the full set equation (table 25.1) is therefore assessed. Thus the full (*in vivo*) equation and the training equation can be used to predict further values of log BM, within the suggested experimental error of SD = 0.3 and AAE = 0.2 log units (from chapter 21).

The inter-correlations of descriptors used for drug blood (or plasma) to muscle distribution are shown in table 25.2. The largest r^2 value for E versus S descriptor is 0.76. It appears that there is no strong inter-correlation between the other set of descriptors used in this model.

r	E	S	A	B	V	I
E	1.00					
S	0.87	1.00				
A	0.29	0.51	1.00			
B	0.63	0.80	0.64	1.00		
V	0.68	0.78	0.43	0.77	1.00	
I	0.08	0.16	0.42	0.12	0.09	1.00

Table 25.2. The inter-correlation of descriptors for 59 compounds


Calculated log P versus oberved log P for the distribution of drugs from blood and plasma to muscle
y = 0.7453x + 0.0011
R2 = 0.7453
CALCULATED LOG P
OBSERVED LOG P
0.80
0.60
0.40
0.20
0.00
-0.20
-0.40
-0.60
-0.80
-1.00
-1.20
-1.50
-1.00
-0.50
0.00
0.50
1.00

Combining the *in vivo* data of drugs with the *in vitro* data of VOCs

Now that the data set of *in vivo* log BM values for passive transport from blood (or plasma) to rat muscle has been investigated, it would be interesting to see what happens if this is combined with the previous data set (*in vitro)* from table 16.3 (chapter 16) for VOCs. Note that the *in vitro* data set is for the transfer of VOCs from human blood and rat blood to human muscle and rat muscle. It was concluded in chapter 9, that there is very little difference data for the species of humans and rats, and that the data from these two species can be combined for VOCs. In the present work, the use of *in vitro* plasma to muscle data for VOCs will not be used, mainly because of the small number of extra compounds involved. The aim of this work is to show whether it is possible (or not) to combine these two different data sets (*in vitro* + *in vivo* data set).

Discussion on blood/plasma to muscle distribution for drugs and VOCs combined

Before combining the two models together of *in vivo* and *in vitro* data sets, it is first important to see if there are any differences in the two equations shown in table 25.1 and 25.3. Note that the *in vitro* equation for VOCs does not have the Ia-indicator variable, because there are no carboxylic acids in the data set. The two sets of coefficients appear to be different visibly and suggest that the *in vivo* and *in vitro* distributions may possibly be different. In particular, the *c*-coefficients differ somewhat 0.082 (SD = 0.100) as against -0.185 (SD = 0.067), which suggests that there might be a small systematic difference between the *in vivo* and *in vitro* distributions.

One way to check for systematic deviations is to use the coefficients in the *in vivo* equation to predict log BM values for the 110 VOCs used to construct the *in vitro* equation. It is found that the predicted and observed *in vitro* log BM values, AE = 0.13, AAE = 0.25, RMSE = 0.31 and SD = 0.31 log units. The numerical values here show that the *in vivo* equation can predict the *in vitro* dataset within 0.31

log units (SD). This is probably because the *in vivo* data set covers a larger descriptor space than the *in vitro* model. Any systematic difference in the equations for *in vivo* (drugs) and *in vitro* (VOCs) distributions could be due to the two series of compounds inhabiting different areas of chemical space.

Another indicator variable was then used to test if there is any systematic difference in the *in vitro* and *in vivo* data set for VOCs and drugs. In this case one can apply the indicator variable, **Iv**, defined as **Iv** = 1 for the *in vitro* data and **Iv** = 0 for the *in vivo* data. An equation with the **Iv** term was constructed shown in table 25.3, because the **Iv** term was statistically significant (T-test = -2.124, p-value = 0.035) and the value of the **Iv** coefficient was -0.140. The coefficients of drug (*in vivo*) and the drug and VOC (*in vivo* + *in vitro*) combined equation are the same within experimental error (SD values of the coefficients are in the drug equation: c=0.100, e=0.084, s=0.110, a=0.128, b=0.084, v=0.073, ia= 0.104 and in the combined equation (*in vivo* + *in vitro*): c=0.080, e=0.057, s=0.076, a=0.097, b=0.063, v=0.054, ia=0.088, iv=0.066). It is concluded that there is a systematic difference between the *in vivo* and *in vitro* distributions (as **Iv** is 0.1 log units), and that the two sets of distribution data can be combined by taking into account this difference (see equation 25.3).

The full data set (*in vivo* and *in vitro* measurements) consists of 163 compounds and 164 data points, the difference arising because blood (or plasma) to muscle distributions are counted separately. The correlation equation of this set of data points with the Abraham descriptors is shown in table 25.3. The fits for this full equation (in blue) appear to be satisfactory (S.D = 0.220 and r^2 = 0.595). This represents a relatively good model for a combined data set of VOCs and drugs for blood/plasma to muscle distribution, and is the only one yet reported for both types of data sets (*in vivo* rats + *in vitro* rats and humans combined).

There were six additional outliers observed in the full set: 1-nitropropane, 2-nitropropane, 2-methylpentane (isohexane), methylcyclopentane, 3-methylhexane and 3-methylpentane. The residual for these compounds is: -0.8, -0.7, 0.7, 0.7, 0.7 and 0.7 log units and may result from experimental error.

EQUATIONS	N(dp)	N(nc)	r^2	SD	F	OUTLIERS	D
FULL SET (*IN VIVO*) **Log BM** = 0.082 – 0.059**E** + 0.010**S** – 0.248**A** + 0.028**B** + 0.110**V** – 1.022**Ia**	60	59	0.745	0.253	25.850	0	6
FULL SET (*IN VITRO*) **Log BM** = -0.185 – 0.209**E** – 0.593**S** – 0.081**A** – 0.168**B** + 0.741**V**	110	110	0.537	0.207	24.125	3	5
FULL SET (*IN VIVO+IN VITRO*) **Log BM** = 0.039 – 0.100**E** – 0.080**S** – 0.254**A** + 0.041**B** + 0.233**V** – 1.005**Ia** – 0.140**Iv**	164	163	0.595	0.220	32.699	6	7
TRAINING SET **Log BM** = 0.114 – 0.102**E** – 0.021**S** – 0.246**A** + 0.072**B** + 0.125**V** – 1.009**Ia** – 0.183**Iv**	82	82	0.697	0.235	24.365	-	7
TEST SET **N**(nc) = 82 **N**(dp) = 82 **SD** (n-1) = 0.213 **RMSE** = 0.212 **AAE** = 0.166 **AE** = -0.023							

N (dp) = number of data points used in the training set
N (nc) = number of compounds used in the training set
r^2 = is the square of the overall correlation coefficient
SD = the standard deviation
F = F-statistic
D = the number of descriptors used in the model

Table 25.3. Summary of the two equations for blood/plasma to muscle distributions for VOCs and drugs combined.

In order to assess the predictive capability of the combined equation (*in vivo* + *in vitro*) a training set was constructed and the equation for the training set used to predict log BM values for the test set, as shown in table 25.3. The coefficients and statistics of the full set of the combined equation (*in vivo* + *in vitro*) and training equation (*in vivo* + *in vitro*) are similar in terms of coefficients. The training equation predicted log BM values for the (unused) test set as follows. N(nc) = 82 and N(dp) = 82, that AE = -0.023, AAE = 0.166, RMSE = 0.212 and SD = 0.213 log units. Thus the combined (*in vivo* + *in vitro*) training equation and the full combined equation (in blue) can be used to predict further values of log BM to close to our suggested experimental error of around 0.2 log units (AAE) or 0.3 log units (SD).

r	E	S	A	B	V	Ia	Iv
E	1.00						
S	0.94	1.00					
A	0.58	0.68	1.00				
B	0.80	0.89	0.76	1.00			
V	0.87	0.90	0.64	0.88	1.00		
Ia	0.30	0.34	0.49	0.31	0.30	1.00	
Iv	-0.81	-0.77	-0.57	-0.73	-0.80	-0.32	1.00

Table 25.4. The inter-correlation of descriptors for 163 compounds

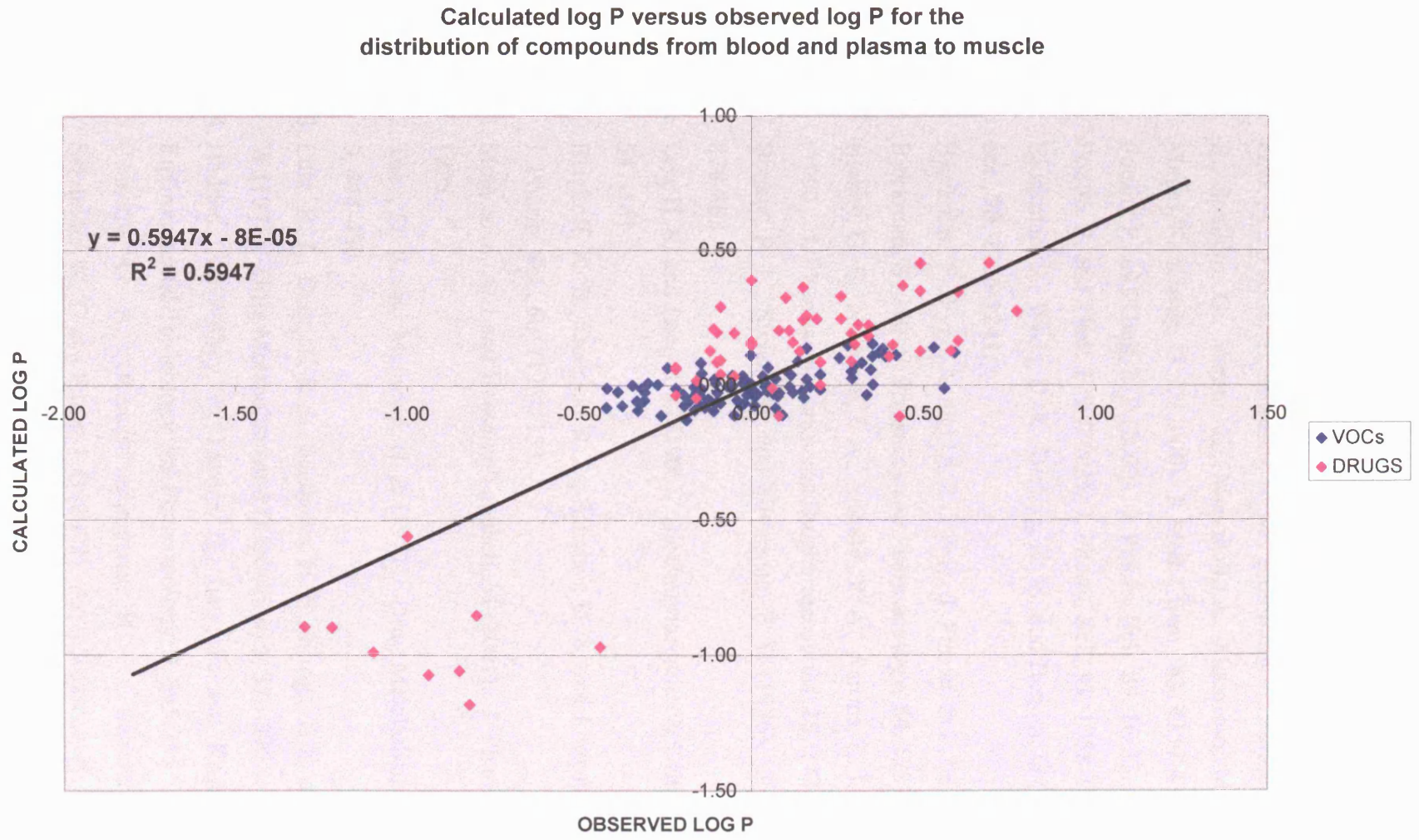
The inter-correlations of descriptors obtained for the combined *in vivo* and *in vitro* equation (from table 25.3) as shown by table 25.4. The largest r^2 value for E versus S descriptor is 0.88. It appears that there is no strong inter-correlation between the other set of descriptors used in this model

The descriptors used in the equation reflect solute/solvent interactions. One can, therefore, interpret the coefficients in terms of the effect that particular interactions have on the process under consideration. The **e**-coefficients (P = 0.08), **s**-coefficient (P = 0.29) and **b**-coefficient (P = 0.52) is not significant. The positive **v**-coefficient in the blood (or plasma) to muscle (*in vivo* + *in vitro*) equation, indicate that increasing molecular size can push drugs out of blood (or plasma) and into the muscle phase. The coefficient of the **A** descriptors are negative, indicating that the hydrogen-bond basicity acts to keep drug compounds in the blood (or plasma) phase and out of the muscle tissue phase. The large negative coefficient of the indicator variable (***Ia***), in the combined equation from table 25.3, indicates a hitherto unknown factor affecting the transfer of solutes between blood (or plasma) and the muscle. It is apparent that the presence of $-CO_2H$ group acts to hinder muscle penetration. This could be simply due to intrinsic hydrogen bonding and the polarity properties of neutral acids. Finally, it is known that some acidic drugs may bind to albumin present in plasma and blood, which may also account for this negative contribution of the coefficient ***Ia***.

This new model based on 163 compounds can now be used to correlate and predict blood (or plasma) to muscle distribution for *in vitro* and *in vivo* drug

compounds. Of course, any prediction using the full combined equation (*in vivo* plus *in vitro*) can, or should, only be made within the chemical space of the test set used to assess the predictions.

The blood/plasma to muscle model is the poorest statistical model reported for blood to tissue distribution for VOCs and drugs combined The plots for this blood/ plasma to muscle distribution model also indicates that the data has not got an even spread data. The VOC and drug are too cluttered in specific regions of the graph (see page 342); this indicates that both datasets (in vitro and in vivo) inhabit different chemical space (see page 153). This model is poor when compared to other good models like blood/plasma to fat distribution (VOCs and drugs combined) that has an even spread of data as shown from the plots (see chapter 22) and give rise to good correlative equations


Calculated log P versus observed log P for the distribution of compounds from blood and plasma to muscle
y = 0.5947x - 8E-05
R2 = 0.5947
CALCULATED LOG P
OBSERVED LOG P
1.00
0.50
0.00
-0.50
-1.00
-1.50
-2.00
-1.50
-1.00
-0.50
0.00
0.50
1.00
1.50
VOCs
DRUGS

References

1. Fazio, F., Todde, S., Moresco, R. M., Simonelli, P., Baraldi, P. G., Cacciari, B., Spalluto, G., Varani, K., Monopoli, A., Matarrese, M., Carpinelli, A., Magni, F., Kienle, G. K (2000). J. Med. Chem., **43**, 4359-4362.
2. Poulin, P. and Theil, F. P (2000). J. Pharm. Sci., **89**, 16-35.
3. Poulin, P. and Theil, F. P (2002). J. Pharm. Sci., **91**, 1358-1367.
4. Bjorkman, S., Wada, D. R., Berling, B. M. and Benoni, G (2001). J. Pharm. Sci., **90**, 1226-1241.
5. Bjorkman, S., Fyge, A. and Qi, Z (1996). J. Pharm. Sci., **85**, 887-889.
6. Bjorkman, S (2001). Pharmacy and Pharmacology, **54**, 1237-1245.
7. Blakey, G. E., Nestorov, I. A., Arundel, P. A., Aarons, L. J. and Rowland, M (1997). J. Pharmacokinetics and Biopharmaceutics, **25**, 277-312.
8. Bonate, P. L., Swann, A. and Silverman, P. B (1996). J. Pharm. Sci., **85**, 878-883.
9. Lee, H. S. and Lee, M. G (1995). Biopharmaceutics & Drug Disposition, **16**, 547-561.
10. Bischoff, K. B., Dedrick, R. L., Zaharko, D. S. and Longstreth, J. A (1971). J. Pharm. Sci., **60**, 1128-1133.
11. Mclachlan, A. J and Hosseini-Yegabeh, M (2001). J. Pharm. Sci., **90**, 1817-1828.
12. Tuey, D. B. and Mathews, H. B (1977). Drug Metabolism and Disposition, **5**, 444-450.
13. Lutz, R. J., Dedrick, R. L., Mathews, H. B., Eling., T. E. and Anderson, M. W (1977). Drug Metabolism and Disposition, **5**, 386-396.
14. Haddad, S., Withey, J., Lapare, S., Law, F. and Krishnan, K (1998). Environmental Toxicology and Pharmacology, **5**, 245-255.
15. Csanady, G. A., Oberste-Frielinghaus, H. R., Semder, B., Baur, C., Schneider, K. T. and Filser, J. G (2002). Arch. Toxicology, **76**, 299-305.
16. Parham, F. M., Mathews, H. B. and Portier, C. J (2002). Toxicology and Applied Pharmacology, **181**, 153-163.
17. Wojcikowski, J. and Daniel, W. A (2002). Pol. J. Pharmacol., **54**, 647-654.

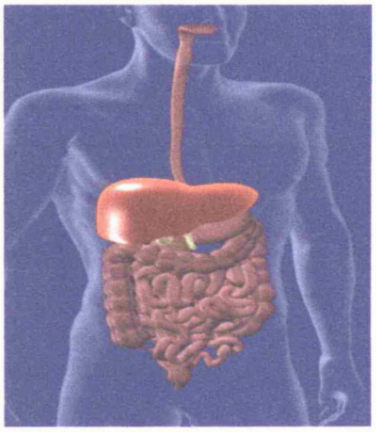
18. Platts, J. A., Abraham, M.H., Zhao, Y.H., Hersey, A., Ijaz, L. and Butina, D (2001). Eur. J. Med. Chem., **36**, 719-730.
19. Salminen, T., Pulli, A. and Taskinen, J (1997). J. Pharmaceutical and Biomedical Analysis, **15**, 469-477.

Chapter

26

Blood/plasma/serum to liver distribution for drugs and VOCs combined

The distribution of drugs between blood (plasma or serum) and liver tissue has become a matter of interest for the pharmaceutical industry, toxicological sciences and environmental health.



Scheme 26.0. Diagram showing the liver, the largest internal organ of the body in humans and rats.

The liver is the largest internal organ of the human body. This organ, which is part of the digestive system, performs more than 500 different functions, all of which are essential to life. Its essential functions include helping the body to digest fats, storing reserves of nutrients, filtering poisons and wastes from the blood, synthesizing a variety of proteins, and regulating the levels of many chemicals found in the bloodstream. The liver is unique among the body's vital organs in that it can

regenerate, or grow back, cells that have been destroyed by some short-term injury or disease. But if the liver is damaged repeatedly over a long period of time, it may undergo irreversible changes that permanently interfere with function.

The human liver is a dark red-brown organ with a soft, spongy texture. It is located at the top of the abdomen, on the right side of the body just below the diaphragm (a sheet of muscle tissue that separates the lungs from the abdominal organs). The lower part of the rib cage covers the liver, protecting it from injury. In a healthy adult, the liver weighs about 1.5 kg (3 lb) and is about 15 cm (6 in) thick.

The liver has not one, but two blood supplies of oxygen rich blood, through the hepatic artery and nutrient rich blood from the portal vein. Inside the liver, fine branches of these blood vessels empty their blood into common vessels. Mixed blood flows past and is processed by liver cells. Cleaned blood leaves the liver through the hepatic veins. The liver is where solutes can distribute between the blood capillaries and liver phase at steady state concentration.

There have been a number of reports of data for drug blood (plasma or serum) to liver distribution in rats (*in vivo*).[1-23] However, there have been no reviews reported yet that attempt to make models to predict blood (plasma or serum) to liver distribution either for rats or humans. The present work is restricted to a data set of 52 different drug compounds (53 data points), which is all that has been published. The *in vivo* drug data for blood (plasma or serum) to liver distribution of rats is at steady state concentration and the distribution from blood (plasma or serum) to liver, BL, is defined by the equation (26.0).

BH = (conc. of drug in liver)/(conc. of drug in blood, plasma or serum) Equation 26.0

The purpose of this work is to collect as many existing literature data on distribution (as log BL values) and to make an attempt to construct equations to predict *in vivo* blood/plasma/serum to liver distribution. The complete list of drug log BL data is shown in table 26.0 (see appendix); where two or more results are reported for the same compound, in the same system (blood/plasma/serum) they have been averaged.

As discussed before (chapter 21), the distribution sets for the tissues that also include liver can be combined together for plasma, blood and serum. In this case, the three distribution sets blood (log $P_{b\text{-}liver}$) plasma (log $P_{p\text{-}liver}$) and serum (log $P_{s\text{-}}$

$_{liver}$) into liver from table 26.0 (see appendix) can now be combined together, but these ratios will not be averaged; thus the data points will be greater than the number of compounds.

Discussion on blood/plasma/serum to liver distribution for drugs only

The 52 compounds (all of which are drugs) were used to construct the blood (plasma or serum) liver distribution model, see table 26.1. The structures for compounds with trivial names can be found in the appendix of this thesis. A first attempt to construct a model was to use the indicator variable for carboxylic acids (**Ia**). This descriptor has been used before for blood to brain distribution of drugs by Platt's[24] and Salminen.[25] The independent variable **Ia** for a value of '1' means the presence of carboxylic acid as a functional group to the drug compound. The independent variable **Ia** for a value of '0' means that the carboxylic acid is absent from the drug compound. This descriptor was used in this model to correlate and predict blood to liver distribution for drugs as shown in table 26.1.

EQUATIONS	N(dp)	N(nc)	r^2	SD	F	OUTLIERS	D
FULL SET (*IN VIVO*)							
Log BL = -0.008 + 0.119**E** – 0.077**S** – 0.276**A** + 0.149**B** + 0.194**V** – 0.871**Ia**	53	52	0.708	0.320	18.574	2	6
TRAINING SET							
Log BL = 0.073 + 0.113**E** – 0.137**S** – 0.351**A** + 0.293**B** + 0.121**V** – 0.828**Ia**	40	39	0.733	0.342	15.126	-	6
TEST SET							
N(nc) = 13 **N**(dp) = 13							
SD (n-1) = 0.305 **RMSE** = 0.293							
AAE = 0.227							
AE = -0.128							

N (dp) = number of data points used in the training set
N (nc) = number of compounds used in the training set
r^2 = is the square of the overall correlation coefficient
SD = the standard deviation
F = F-statistic
D = the number of descriptors used in the model

Table 26.1. Summary of the two equations for blood/plasma/serum to liver distributions for drugs only

For the present model for blood (plasma or serum) to liver distribution, there were two compounds that were marked outliers: p,p'-dichlorodiphenylsulfone and valproic acid. The residuals for these two compounds are both 0.8 units. These two outliers may have resulted from experimental errors. The statistics of the equation for the full set in table 26.1 are satisfactory and the predictive capability of the equation can be assessed by the usual method of constructing a training set and an independent test set. The selection of the training (40 data points) and test sets (13 data points) was carried out using the usual computing method called Kennard-Stone. The training equation is also shown in table 26.1.

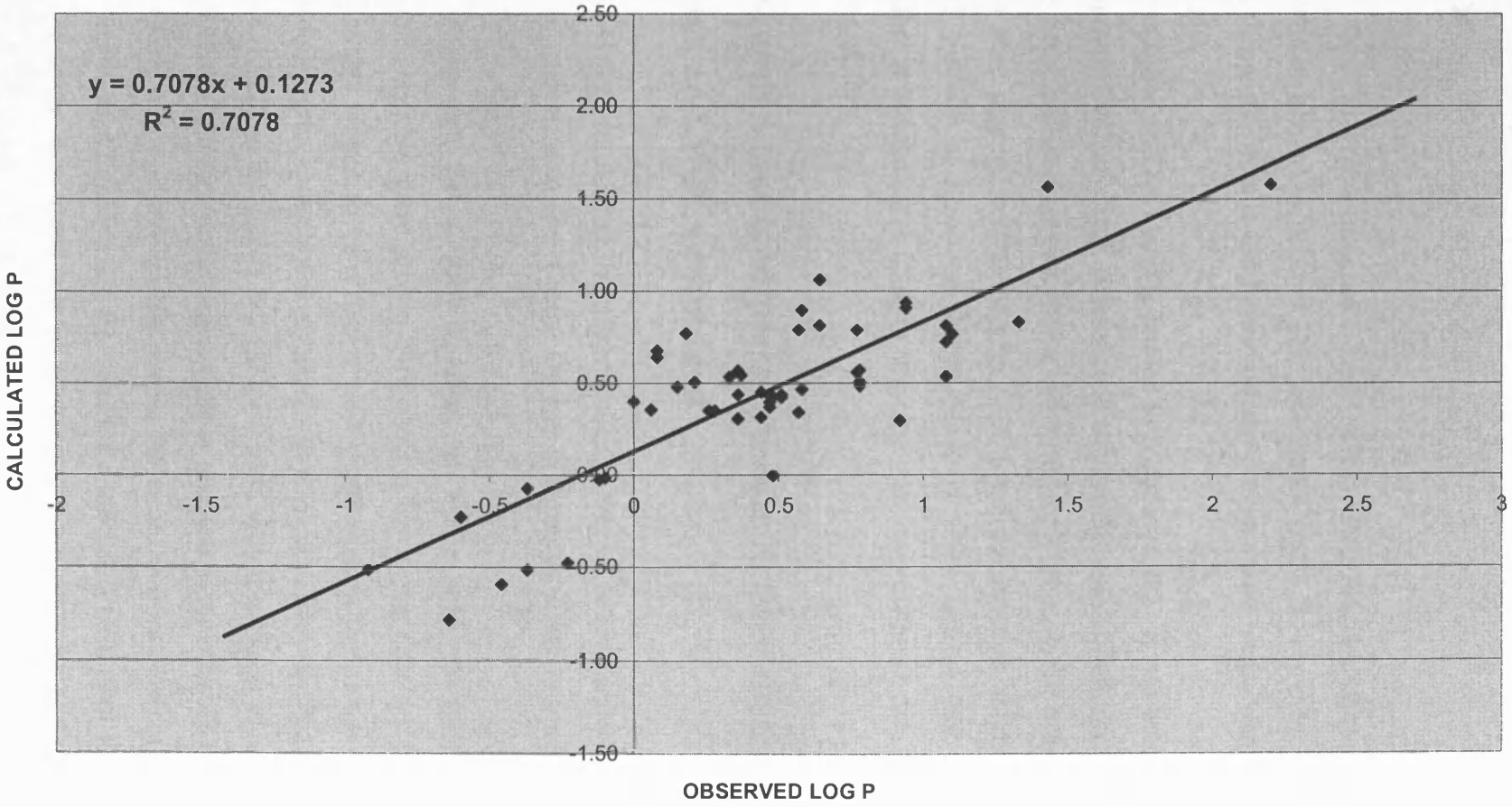
The statistics of the full set (table 26.1) and the training set (table 26.1) are similar and so are some of the coefficients. If the training set equation is used to predict the remainder of the data in the (unused) test set, it is found that AE = -0.13, AAE = 0.23 RMSE = 0.29 and SD = 0.31, that there is very little bias, and the predictive capability of the training equation (table 26.1) and by implication of the full set equation (table 26.1) is therefore assessed. Thus the full (*in vivo*) equation and the training equation can be used to predict further values of log BL, close to the suggested experimental error of SD = 0.3 and AAE = 0.2 log units (from chapter 21).

r	E	S	A	B	V	I
E	1.00					
S	0.81	1.00				
A	0.30	0.52	1.00			
B	0.53	0.75	0.63	1.00		
V	0.57	0.72	0.40	0.86	1.00	
I	0.18	0.27	0.42	0.09	0.00	1.00

Table 26.2. The inter-correlation of descriptors for 52 compounds

The inter-correlations of descriptors used for drug blood (plasma or serum) to liver distribution are shown in table 26.2. The largest r^2 value for B versus V descriptor is 0.74. It appears that there is no strong inter-correlation between the other set of descriptors used in this model.

Calculated log P versus observed log P for the distribution of drugs from blood, plasma and serum to liver



Combining the *in vivo* data of drugs with the *in vitro* data of VOCs

Now that the data set of *in vivo* log BL values for passive transport from blood (plasma or serum) to rat liver has been investigated, it would be interesting to see what happens if this is combined with the previous data set (*in vitro)* from table 15.3 (chapter 15) for VOCs. Note that the *in vitro* data set is transfer of VOCs from human blood and rat blood to human liver and rat liver. It was concluded in chapter 9, that there is very little difference in the two species of humans and rats, and that data from these two species can be combined for VOCs. In the present work, the use of *in vitro* plasma to liver data for VOCs will not be used, mainly because of the small number of extra compounds involved. The aim of this work is to show whether it is possible (or not) to combine these two different data sets (*in vitro* + *in vivo* data set).

Discussion on blood/plasma/serum to liver distribution for drugs and VOCs combined

Before combining the two models together of *in vivo* and *in vitro* data sets, it is first important to see if there are any differences in the two equations shown in table 26.1 and 26.3.

Note that the *in vitro* equation for VOCs does not have the Ia-indicator variable, because there are no carboxylic acids in the data set. The two sets of coefficients appear to be different visibly and suggest that the *in vivo* and *in vitro* distributions may possibly be different. In particular, the *c*-coefficients differ -0.008 (SD = 0.152) as against -0.202 (SD = 0.078), which suggests that there might be a small systematic difference between the *in vivo* and *in vitro* distributions.

One way to check for systematic deviations is to use the coefficients in the *in vivo* equation, to predict log BL values for the 120 VOCs used to construct the *in vitro* equation. It is found that for the predicted and observed *in vitro* log BL values, AE = -0.14, AAE = 0.25, RMSE = 0.34 and SD = 0.34 log units. The numerical values here show that the *in vivo* equation can predict the *in vitro* dataset within 0.34 log units (SD). This suggests that the *in vivo* model can successfully predict the *in*

vitro data set as it covers a larger descriptor space than the *in vitro* model. Any systematic difference between the *in vivo* (drugs) and the *in vitro* (VOCs) distributions could be due to the two series of compounds inhabiting different areas of chemical space.

Another indicator variable was then used to test if there is any difference in the *in vitro* and *in vivo* data set for VOCs and drugs. In this case one can apply the indicator variable, **Iv**, defined as **Iv** = 1 for the *in vitro* data and **Iv** = 0 for the *in vivo* data. An equation with the **Iv** term was constructed but is not shown, because the **Iv** term was not statistically significant (T-test = 1.612, p-value = 0.109) and the value of the **Iv** coefficient was quite small (0.137). The coefficients of drug (*in vivo*) and the drug and VOC (*in vivo* + *in vitro*) combined equation are the same within experimental error (SD values of the coefficients are in the drug equation: c= 0.152, e= 0.093, s= 0.137, a= 0.167, b= 0.118, v= 0.105, ia= 0.140 and in the combined equation (*in vivo* + *in vitro*): c= 0.048, e= 0.058, s= 0.082, a= 0.114, b= 0.073, v= 0.057, ia= 0.109). It is concluded that there is a very small systematic difference between the *in vivo* and *in vitro* distributions (as **Iv** is 0.1 log units), and that the two sets of distribution data can be combined without taking into account this difference (see equation 26.3).

The full data set (*in vivo* and *in vitro* measurements) consists of 169 compounds and 170 data points, the difference arising because blood (plasma or serum) to liver distributions are counted separately. The correlation equation of this set of data points with the Abraham descriptors is shown in table 26.3. The fits for this full equation (in blue) appear to be reasonable (S.D = 0.254 and r^2 = 0.660). This represents a satisfactory model for a combined data set of VOCs and drugs for blood/plasma/serum to liver distribution, and is the only model yet reported for both types of data sets (*in vivo* rats + *in vitro* rats and humans combined).

There were three additional outliers observed in the full set: terbinafine, 1-(3-fluoropropyl)-2-nitroimidazole (FPN) and methotrexate. The residuals for these compounds are: -0.9, 0.8 and 0.8 log units and may result from experimental error.

EQUATIONS	N(dp)	N(nc)	r^2	SD	F	OUTLIERS	D
FULL SET (*IN VIVO*) **Log BL** = -0.008 + 0.119**E** – 0.077**S** – 0.276**A** + 0.149**B** + 0.194**V** – 0.871**Ia**	53	52	0.708	0.320	18.574	2	6
FULL SET (*IN VITRO*) **Log BL** = -0.202 – 0.048**E** – 0.295**S** – 0.213**A** – 0.359**B** + 0.876**V**	120	120	0.648	0.211	42.012	4	5
FULL SET (*IN VIVO* + *IN VITRO*) **Log BL** = – 0.062 + 0.104**E** – 0.351**S** – 0.202**A** – 0.127**B** + 0.617**V** – 0.831**Ia**	170	169	0.660	0.254	52.762	3	6
TRAINING SET **Log BL** = 0.022 + 0.148**E** – 0.377**S** – 0.220**A** – 0.007**B** + 0.491**V** – 0.851**I**	85	85	0.711	0.273	32.006	-	6
TEST SET **N**(nc) = 85 **N**(dp) = 85 **SD** (n-1) = 0.245 **RMSE** = 0.243 **AAE** = 0.182 **AE** = 0.001							

N (dp) = number of data points used in the training set
N (nc) = number of compounds used in the training set
r^2 = is the square of the overall correlation coefficient
SD = the standard deviation
F = F-statistic
D = the number of descriptors used in the model

Table 26.3. Summary of the two equations for blood/plasma/serum to liver distributions for VOCs and drugs combined.

In order to assess the predictive capability of the combined equation (*in vivo* + *in vitro*) a training set was constructed and the equation for the training set used to predict log BL values for the test set, as shown in table 26.3. The coefficients and statistics of the full set of the combined equation (*in vivo* + *in vitro*) and training equation (*in vivo* + *in vitro*) are similar in terms of coefficients. The training equation predicted log BL values for the (unused) test sets as follows. N(nc) = 85 and N(dp) = 85, that AE = 0.001, AAE = 0.182, RMSE = 0.243 and the SD = 0.245 log units. Thus the combined (*in vivo* + *in vitro*) training equation and the full combined equation (in blue) can be used to predict further values of log BL to close to our suggested experimental error of around 0.2 log units (AAE) or 0.3 log units (SD).

r	E	S	A	B	V	Ia
E	1.00					
S	0.91	1.00				
A	0.61	0.72	1.00			
B	0.74	0.85	0.76	1.00		
V	0.80	0.83	0.65	0.89	1.00	
Ia	0.32	0.37	0.45	0.24	0.23	1.00

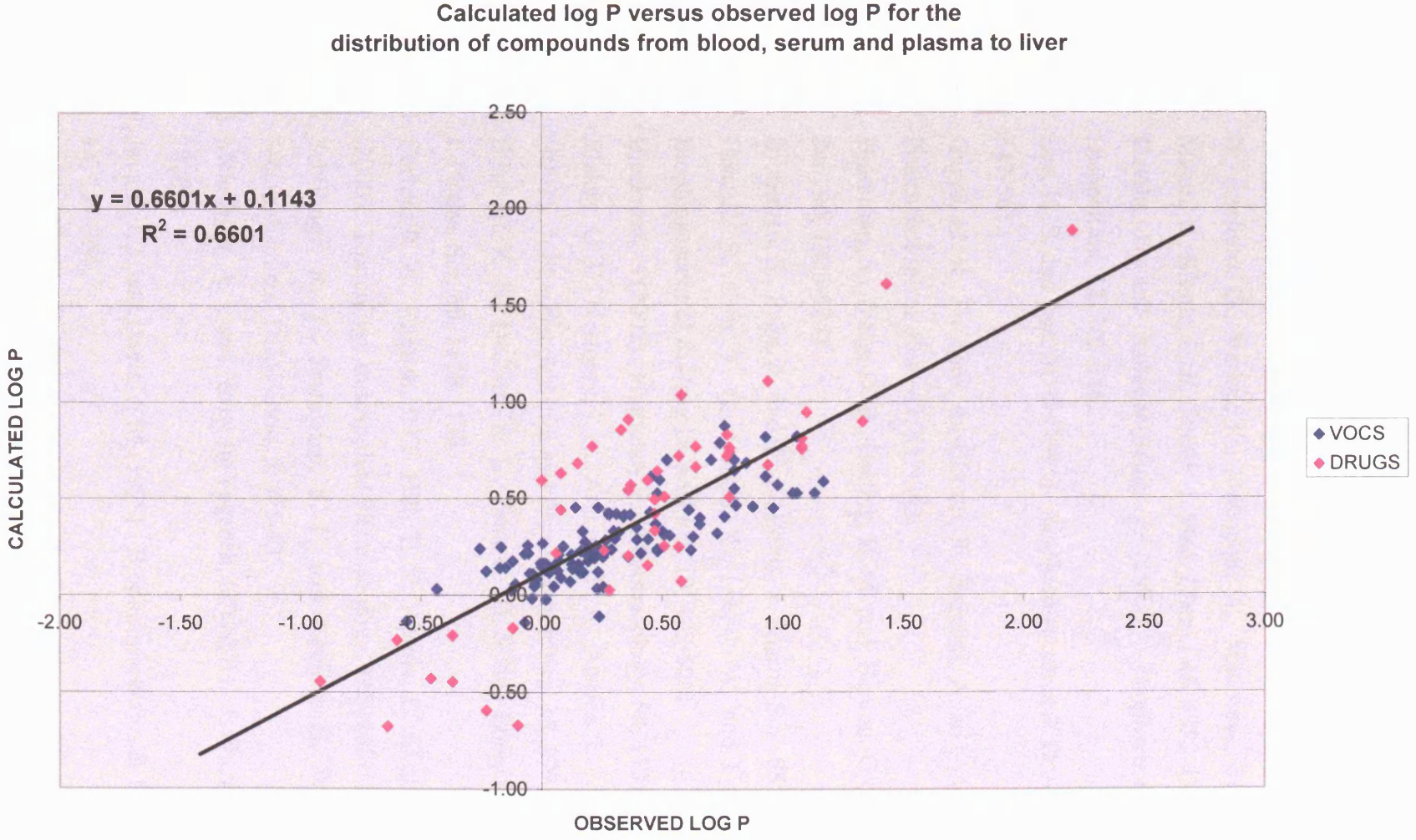
Table 26.4. The inter-correlation of descriptors for 169 compounds

The inter-correlations of descriptors obtained for the combined *in vivo* and *in vitro* equations (from table 26.3) as shown by table 26.4. The largest r^2 value for E versus S descriptor is 0.83. It appears that there is no strong inter-correlation between the other set of descriptors used in this model.

The descriptors used in the equation reflect solute/solvent interactions. One can, therefore, interpret the coefficients in terms of the effect that particular interactions have on the process under consideration. The **e**-coefficients (P = 0.08), **a**-coefficient (P = 0.08) and **b**-coefficient (P = 0.09) is not significant. The positive **v**-coefficient in the blood/plasma/serum to liver (*in vivo* + *in vitro*) equation from table 26.3, indicate that increasing molecular size can push drugs out of blood/plasma/serum and into liver. The coefficient of the 'polar' descriptor **S** is negative, indicating that polarity is stronger in the blood (plasma or serum) phase than in the liver tissue phase. Therefore the drug compounds will move from the liver phase and into the blood/plasma/serum phase. The negative coefficient of the indicator variable (***Ia***), in the combined equation from table 26.3., indicates a hitherto unknown factor affecting the transfer of solutes between blood/plasma/serum and the liver. It is apparent that the presence of $-CO_2H$ group acts to hinder liver penetration further. This could be simply due to intrinsic hydrogen bonding and the polarity properties of neutral acids. Finally, it is known that some acidic drugs may bind to albumin present in blood, plasma and serum, which may also account for this negative contribution of the coefficient ***Ia***.

This new model based on 169 compounds can now be used to correlate and predict blood (plasma or serum) to liver distribution for *in vitro* and *in vivo* drug compounds. Of course, any prediction using the full combined equation (*in vivo* plus

in vitro) can, or should, only be made within the chemical space of the test set used to assess the predictions.


Calculated log P versus observed log P for the distribution of compounds from blood, serum and plasma to liver
y = 0.6601x + 0.1143
R2 = 0.6601
CALCULATED LOG P
OBSERVED LOG P
2.50
2.00
1.50
1.00
0.50
0.00
-0.50
-1.00
-2.00
-1.50
-1.00
-0.50
0.00
0.50
1.00
1.50
2.00
2.50
3.00
VOCS
DRUGS

References

1. Fazio, F., Todde, S., Moresco, R. M., Simonelli, P., Baraldi, P. G., Cacciari, B., Spalluto, G., Varani, K., Monopoli, A., Matarrese, M., Carpinelli, A., Magni, F., Kienle, G. K (2000). J. Med. Chem., **43**, 4359-4362.
2. Davila, D. and Kolacny-Babic, L (1991). Biopharmaceutics & Drug Disposition, **12**, 505-514.
3. Lee, H. S. and Lee, M. G (1995). Biopharmaceutics & Drug Disposition, **16**, 547-561.
4. Gasco, M. R., Fundaro, A., Cavalli, R., Bargoni, A. and Vighetto, D (2000). Pharmacological. Res., **42**, 337-343.
5. Bjorkman, S., Wada, D. R., Berling, B. M. and Benoni, G (2001). J. Pharm. Sci., **90**, 1226-1241.
6. Bjorkman, S., Fyge, A. and Qi, Z (1996). J. Pharm. Sci., **85**, 887-889.
7. Han, K. S., Kim, Y. G., Yoo, J. K., Lee, J. W. and Lee, M. G (1998). Biopharmaceutics & Drug Disposition, **19**, 493-500.
8. Bjorkman, S (2002). Pharmacy and Pharmacology, **54**, 1237-1245.
9. Blakey, G. E., Nestorov, I. A., Arundel, P. A., Aarons, L. J. and Rowland, M (1997). J. Pharmacokinetics and Biopharmaceutics, **25**, 277-312.
10. Bischoff, K. B., Dedrick, R. L., Zaharko, D. S. and Longstreth, J. A (1971). J. Pharm. Sci, **60**, 1128-1133.
11. Corley, R. A., English, J. C., Hill, T. S., Fiorcia, L. A. and Morgott, D. A (2000). Toxicology and Applied Pharmacology, **165**, 163-174.
12. Schillings, R. T., Sisenwine, S. F. and Ruelius, H. W (1977). Drug Metabolism and Disposition, **5**, 425-435.
13. Mclachlan, A. J. and Hosseini-Yegabeh, M (2001). J. Pharm. Sci., **90**, 1817-1828.
14. Javaid, J. I. and Davis, J. M (1993). Biopharmaceutics & Drug Disposition, **14**, 357-364.
15. Yamamoto, F., Oka, H., Antoku, S., Ichiya, Yu-ichi., Masuda, K. and Maeda, M (1999). Biol. Pharm. Bull., **22**, 590-597.
16. Tuey, D. B. and Mathews, H. B (1977). Drug Metabolism and Disposition, **5**, 444-450.

17. Lutz, R. J., Dedrick, R. L., Mathews, H. B., Eling., T. E. and Anderson, M. W (1977). Drug Metabolism and Disposition, **5**, 386-396.
18. Haddad, S., Withey, J., Lapare, S., Law, F. and Krishnan, K (1998). Environmental Toxicology and Pharmacology, **5**, 245-255.
19. Bonate, P. L., Swann, A. and Silverman, P. B (1996). J. Pharm. Sci., **85**, 878-883.
20. Csanady, G. A., Oberste-Frielinghaus, H. R., Semder, B., Baur, C., Schneider, K. T. and Filser, J. G (2002). Arch. Toxicology, **76**, 299-305.
21. Parham, F. M., Mathews, H. B. and Portier, C. J (2002). Toxicology and Applied Pharmacology, **181**, 153-163.
22. Wojcikowski, J. and Daniel, W. A (2002). Pol. J. Pharmacol., **54**, 647-654.
23. Aravagiri, M., Teper, Y. and Marder, S. R (1999). Biopharmaceutics & Drug Disposition, **20**, 369-377.
24. Platts, J. A., Abraham, M.H., Zhao, Y.H., Hersey, A., Ijaz, L. and Butina, D (2001). Eur. J. Med. Chem., **36**, 719-730.
25. Salminen, T., Pulli, A. and Taskinen, J (1997). J. Pharmaceutical and Biomedical Analysis, **15**, 469-477.

Chapter

27

Blood/plasma to skin distribution (rat only)

Blood/plasma to skin distribution, or the ability of a molecule to enter the skin tissue, has become a subject of interest of interest to pharmaceutical industry. The skin acts as a physical barrier to the external environment and protects the body from environmental pollutants. The skin is made up of several layers also containing a rich supply of capillary blood network (see figure 27.0). Here the distribution of drugs can occur at steady state concentration between the blood and the skin (and vice versa).

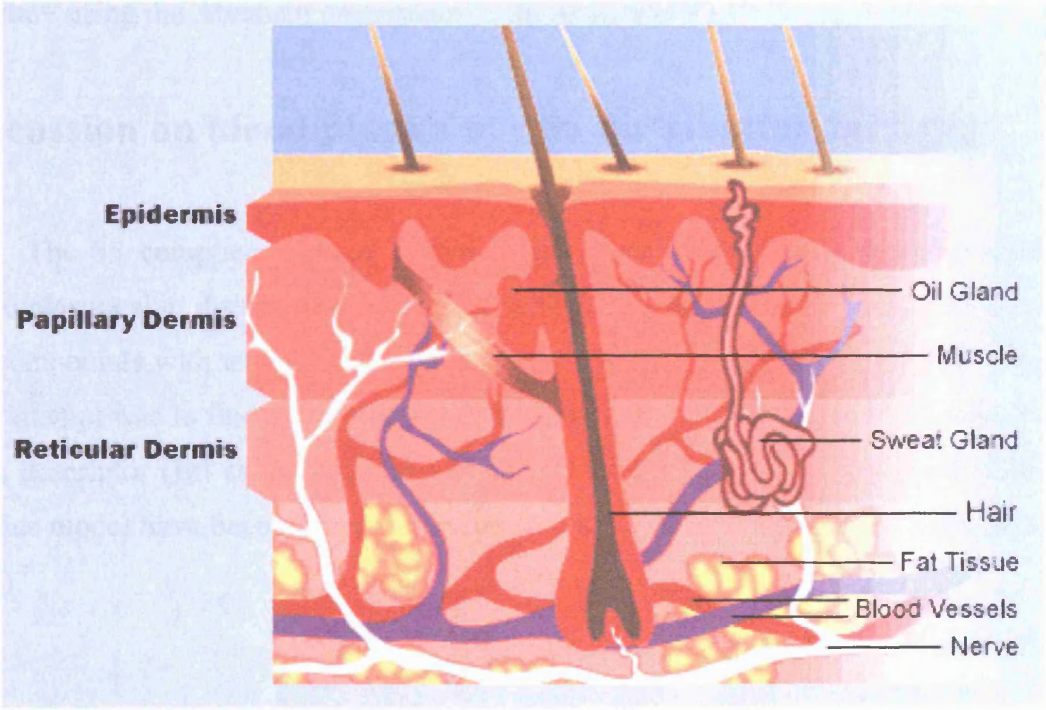


Figure 27.0. The structure of skin.

There have been a number of reports of data for drug blood/plasma to skin distribution in rats.[1-12] There have been no specific reviews reported yet that to attempt to make models to predict blood/plasma to skin distribution either for rats or humans. The present work is restricted to a small data set of 35 compounds, which is all that has been published. The drug data for blood/plasma to skin distribution of rats is at steady state concentration. The aim is to attempt to construct a model based on the Abraham equations to predict the distribution process. This equilibrium distribution for blood/plasma to skin distribution will be based on rats, and will be the first reported model in this area for blood/plasma to skin transport. It should be possible to build bigger models in the near future providing further literature data are published. Blood/plasma to skin distribution, BS, is expressed by the equation below.

BS = (conc. of drug in skin)/(conc. of drug in blood/plasma) Equation 27.0

The aim of this work is to construct an equation as part of a general method for 'high-throughput' prediction of equilibrium blood/plasma-skin distribution based on structure using the Abraham descriptors (E, S, A, B, and V).[13]

Discussion on blood/plasma to skin distribution for drugs

The 35 compounds (most of which are drugs) were used to construct the blood/plasma skin distribution model, see table 27.0 (see appendix). The structures for compounds with trivial names can be found in the appendix of this thesis. The first attempt was to find out whether it is was necessary for this model to include the sixth descriptor (**Ia**) as an indicator variable for carboxylic acids. The correlations for this model have been carried out in order to test this indicator variable (see table 27.1).

EQUATIONS (*IN VIVO* DATA ONLY)	N	r^2	SD	F	OUTLIERS	D
Log BS = 0.557 + 0.246**E** – 0.155**S** – 0.750**A** – 0.453**B** + 0.144**V**	34	0.658	0.326	10.803	p-Phenylbenzoic acid Terbinafine	5
Log BS = 0.527 + 0.071**E** + 0.085**S** – 0.468**A** – 0.490**B** + 0.134**V** – 0.674**Ia**	35	0.762	0.288	14.979	Terbinafine	6

N = number of compounds used in the training set
r^2 = is the square of the overall correlation coefficient
SD = the standard deviation
F = F-statistic
D = the number of descriptors used in the model

Table 27.1. Summary of the two equations for drug blood/plasma skin distributions

The first equation in Table 27.1 does not include the **Ia** descriptor. The fit for this equation appears to be poorer than the second equation. However, several discrepancies were observed for carboxylic acid containing compounds, such as p-phenylbenzoic acid, glycyrrhetinic acid, nalidixic acid, salicylic acid, valproic acid and acrylic acid. From this set of carboxylic acid containing compounds p-phenylbenzoic acid appeared to be the biggest outlier (residual of 0.93 log units). An improvement over the first equation can be achieved by the inclusion of the indicator variable (**Ia**) that is set to 1 for a compound containing a carboxylic acid fragment and 0 otherwise. Carboxylic acid containing compounds are known to ionize in blood and therefore prevent passive distribution into tissues.[14] From the above equations it seems to make sense to include this indicator variable to take into account the transport problems that are known to occur. The second equation has the best fit of the two equations (smaller standard deviation and the larger square of the correlation coefficient) and can be viewed graphically on page 363.

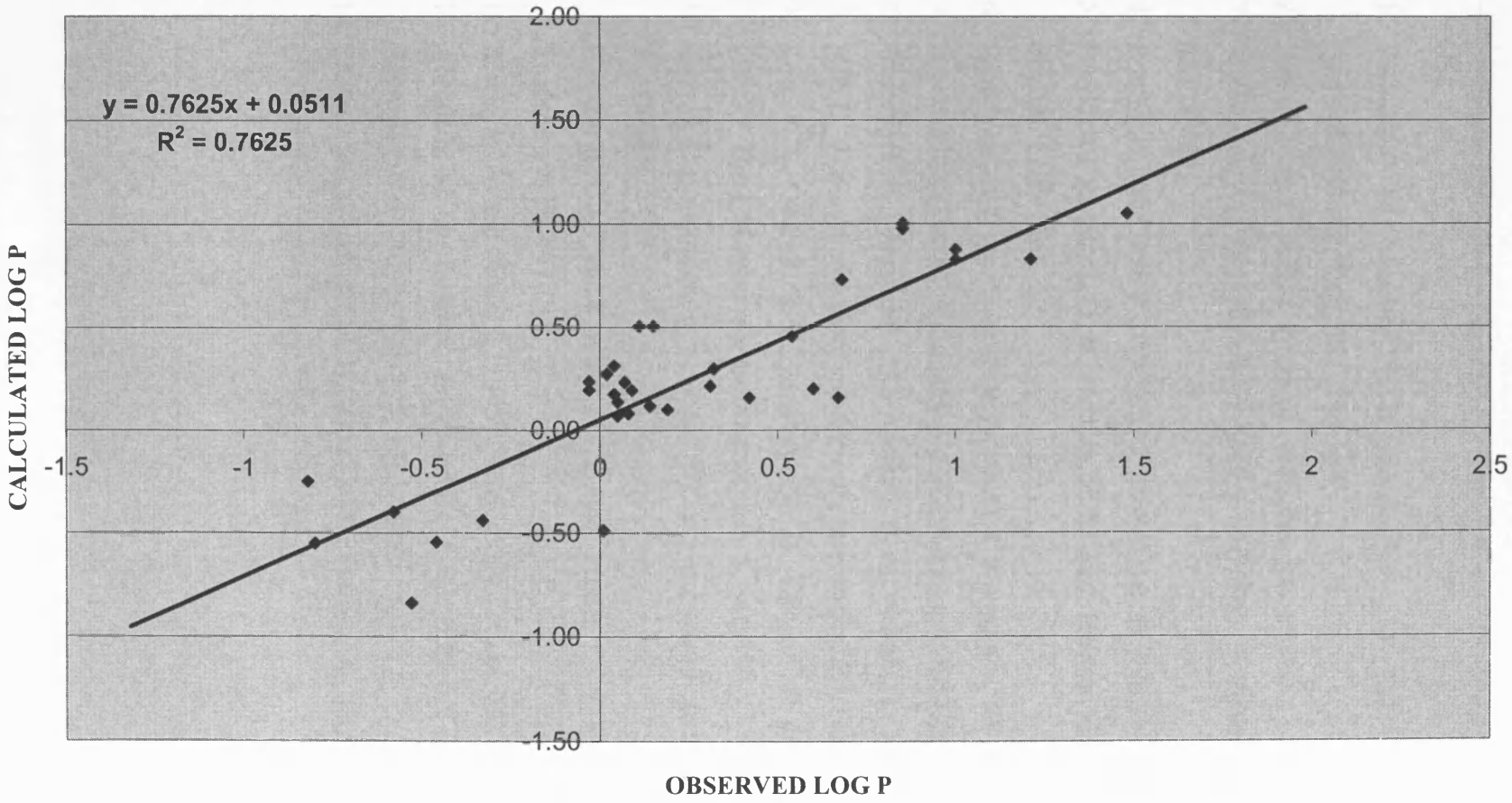
The molecule omitted from this model is terbinafine (residual of 0.88 log units). There could be several reasons why terbinafine appears as an outlier from this model. Possible reasons include experimental error (difficulties in determining log P values *in vivo*), metabolism of this compound (in skin, blood or plasma) and efflux that may affect the equilibrium processes. The inter-correlations of descriptors are obtained for the second equation and are shown by table 27.2. The largest r^2 value for E versus S descriptor is 0.76. It appears that there is no strong inter-correlation between the other set of descriptors used in this model.

r	E	S	A	B	V	Ia
E	1.00					
S	0.87	1.00				
A	-0.21	0.14	1.00			
B	0.42	0.57	0.35	1.00		
V	0.55	0.57	0.06	0.58	1.00	
Ia	-0.05	0.15	0.54	0.06	-0.04	1.00

Table 27.2. The inter-correlation of descriptors for 35 compounds

The Abraham solvation equation has successfully been employed to correlate blood/plasma to skin distribution for drugs. The equation could be used to predict further values provided the solute is within the descriptor space of the training set. No test sets have been used for this model, as there are a limited number of compounds available. The descriptors used in the equation reflect solute/solvent interactions. Therefore, one can interpret the coefficients in terms of effect that particular interactions have on the process under consideration. The **e**-coefficients (P = 0.73), **s**-coefficient (P = 0.70), **a**-coefficient (P = 0.09) and **v**-coefficient (P = 0.24) is not significant. The coefficient of the 'polar' descriptors **S** is positive, indicating that polarity is stronger in the skin phase than in blood/plasma. Therefore the drug compounds will move from the blood/plasma phase and into the skin phase. The large negative coefficient of the indicator variable (***Ia***), in equation 27.1 and its low correlation with other descriptors (27.2), indicates a hitherto unknown factor affecting the transfer of solutes between blood/plasma and the skin. It is apparent that the presence of $-CO_2H$ group acts to hinder skin penetration further. This could be simply due to intrinsic hydrogen bonding and the polarity properties of neutral acids. Finally, it is known that some acidic drugs may bind to albumin present in plasma and blood, which may also account for this negative contribution of the coefficient **Ia**.

Calculated log P versus observed log P for the distribution of drugs from blood and plasma to skin



References

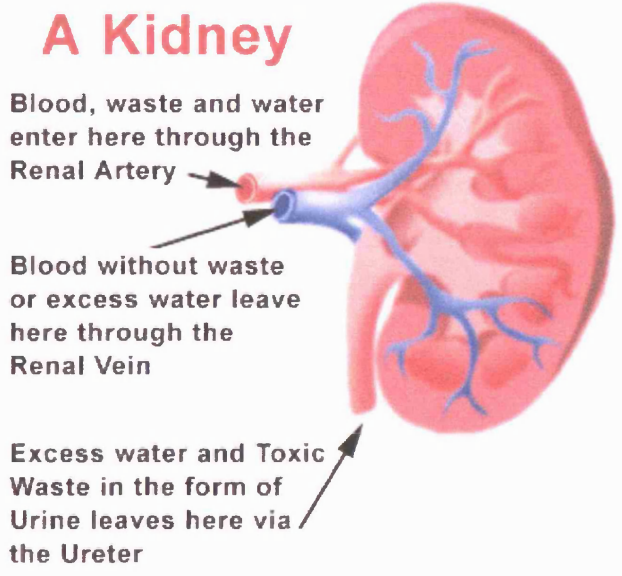
1. Parham, F. M., Mathews, H. B. and Portier, C. J (2002). Toxicology and Applied Pharmacology, **181**, 153-163.
2. Poulin, P. and Theil, F. P (2000). J. Pharm. Sci, **89** (No 1), 16-35.
3. Mclachlan, A. J and Hosseini-Yegabeh, M (2001). J. Pharm. Sci, **90** (No 11), 1817-1828.
4. Poulin, P. and Theil, F. P (2002). J. Pharm. Sci, **91** (No 5), 1358-1367.
5. Bjorkman, S., Wada, D. R., Berling, B. M. and Benoni, G (2001). J. Pharm. Sci, **90**, 1226-1241.
6. Bjorkman, S., Fyge, A. and Qi, Z (1996). J. Pharm. Sci, **85** (No 8), 887-889.
7. Bjorkman, S (2002). Pharmacy and Pharmacology, **54**, 1237-1245.
8. Blakey, G. E., Nestorov, I. A., Arundel, P. A., Aarons, L. J. and Rowland, M (1997). J. Pharmacokinetics and Biopharmaceutics, **25** (No 3), 277-312.
9. Lutz, R. J., Dedrick, R. L., Matthews, H. B., Eling, T. E. and Aderson, M. W (1977). Drug Metabolism and Disposition, **5**, 386-396.
10. Tuey, D. B. and Matthews, H. B (1977). Drug Metabolism and Disposition, **5**, 444-450.
11. Perleberg, U. R. and Keys, D. A. and Fisler, J. W (2004). Inhalation Toxicology, **16**, 771-783.
12. Black, K. A. and Finch, L (1995). Toxicology Letters, **78**, 73-78.
13. Abraham, M. H., Ibrahim, A., Zissimos, A. M., Zhao, Y. H., Comer, J. and Reynolds, D. P (2002). DDT, **7**, 1056-1063.
14. Platts, J. A., Abraham, M. H., Zhao, Y. H., Hersey, A., Ijaz, L. and Butina, D (2001). Eur. J. Med. Chem., **36**, 719-730.

Chapter

28

Blood/plasma/serum to kidney distribution for drugs and VOCs combined

The distribution of drugs between blood (plasma or serum) and kidney tissue has become a matter of interest for the pharmaceutical industry, toxicological sciences and environmental health.



Scheme 28.0. Diagram showing the kidney found in humans and rats.

In humans, kidneys are situated one on each side of the spine and are embedded in fatty tissue. They have a bean-shaped structure, possessing a convex outer border and a concave inner border. The inner border presents an indentation,

the hilum, at which the blood vessels enter and leave. In front is the renal vein carrying blood from the kidney; behind it lies the renal artery carrying blood to the kidney. The ureter is a tube that carries urine to the bladder from the kidney. The kidney also embodies glomeruli, aggregations or loops of capillaries enclosed within thin envelopes of endothelial lining called Bowman's capsules, located at the blind ends of the renal tubules.

The kidneys are important in maintaining a balance of fluid, salt and a normal degree of acidity. When disorders upset these delicate balances, the kidneys act to restore them by excreting more or less water, salt, and hydrogen ions. The kidneys help maintain normal blood pressure by secreting the hormone renin and elaborate a hormone that stimulates the production of red blood cells.

The kidney has a rich supply of blood from the renal artery and waste products are removed from the kidney via the renal vein. The distribution of drugs takes place between the blood capillaries and the kidney tissue at steady-state concentration.

There have been a number of reports of data for drug blood (plasma or serum) to kidney distribution in rats (*in vivo*).[1-20] There have been no specific reviews reported yet that attempt to make models to predict blood (plasma or serum) to kidney distribution either for rats or humans. The present work is restricted to a data set of 50 different drug compounds (51 data points), which is all that has been published. The *in vivo* drug data for blood (plasma or serum) to kidney distribution of rats is at steady state concentration and the distribution from blood (plasma or serum) to kidney, BK, is defined by the equation (28.0).

BK = (conc. of drug in kidney)/(conc. of drug in blood, plasma or serum) Equation 28.0

The purpose of this work is to collect as many existing literature data on distribution (as log BK values) and to make attempt to construct equations to predict *in vivo* blood/plasma/serum to kidney distribution. The complete list of drug log BK data is shown in table 28.0 (see appendix); where two or more results are reported for the same compound in the same system (blood/plasma/serum) they have been averaged.

As discussed before (chapter 21), the distribution sets for the tissues that also include kidney can be combined together for plasma, blood and serum. In this case, the three distribution sets: blood (log $P_{b\text{-}kidney}$), plasma (log $P_{p\text{-}kidney}$) and serum (log

$P_{s\text{-}kidney}$) into kidney from table 28.0 (see appendix) can now be combined together, but these ratios will not be averaged; thus the data points will be greater than the number of compounds.

Discussion on blood/plasma/serum to kidney distribution for drugs only

The 50 compounds (all of which are drugs) were used to construct the blood (plasma or serum) to kidney distribution model, see table 28.1. The structures for compounds with trivial names can be found in the appendix of this thesis. A first attempt to construct a model was to use the indicator variable for carboxylic acids (**Ia**). This descriptor has been used before for blood to brain distribution of drugs by Platt's[21] and Salminen.[22] The independent variable **Ia** for a value of '1' means the presence of carboxylic acid as a functional group to the drug compound. The independent variable **Ia** for a value of '0' means that the carboxylic acid is absent from the drug compound. This descriptor was used in a model to correlate and predict blood (plasma or serum) to kidney distribution for drugs as shown in table 28.1.

EQUATIONS (rat model)	N(dp)	N(nc)	r^2	SD	F	OUTLIERS	D
FULL SET (*IN VIVO*)							
Log BK = 0.048 – 0.022**E** – 0.118**S** – 0.486**A** + 0.215**B** + 0.317**V** – 0.411**Ia**	51	50	0.726	0.299	19.465	1	6
TRAINING SET							
Log BK = 0.073 – 0.026**E** – 0.122**S** – 0.520**A** + 0.236**B** + 0.304**V** – 0.372**Ia**	40	40	0.726	0.327	14.550	-	6
TEST SET							
N(nc) = 11 **N**(dp) = 11							
SD (n-1) = 0.202 **RMSE** = 0.192							
AAE = 0.180 **AE** = 0.004							

N (dp) = number of data points used in the training set
N (nc) = number of compounds used in the training set
r^2 = is the square of the overall correlation coefficient
SD = the standard deviation, F = F-statistic
D = the number of descriptors used in the model

Table 28.1. Summary of the two equations for blood/plasma/serum to kidney distributions for drugs only

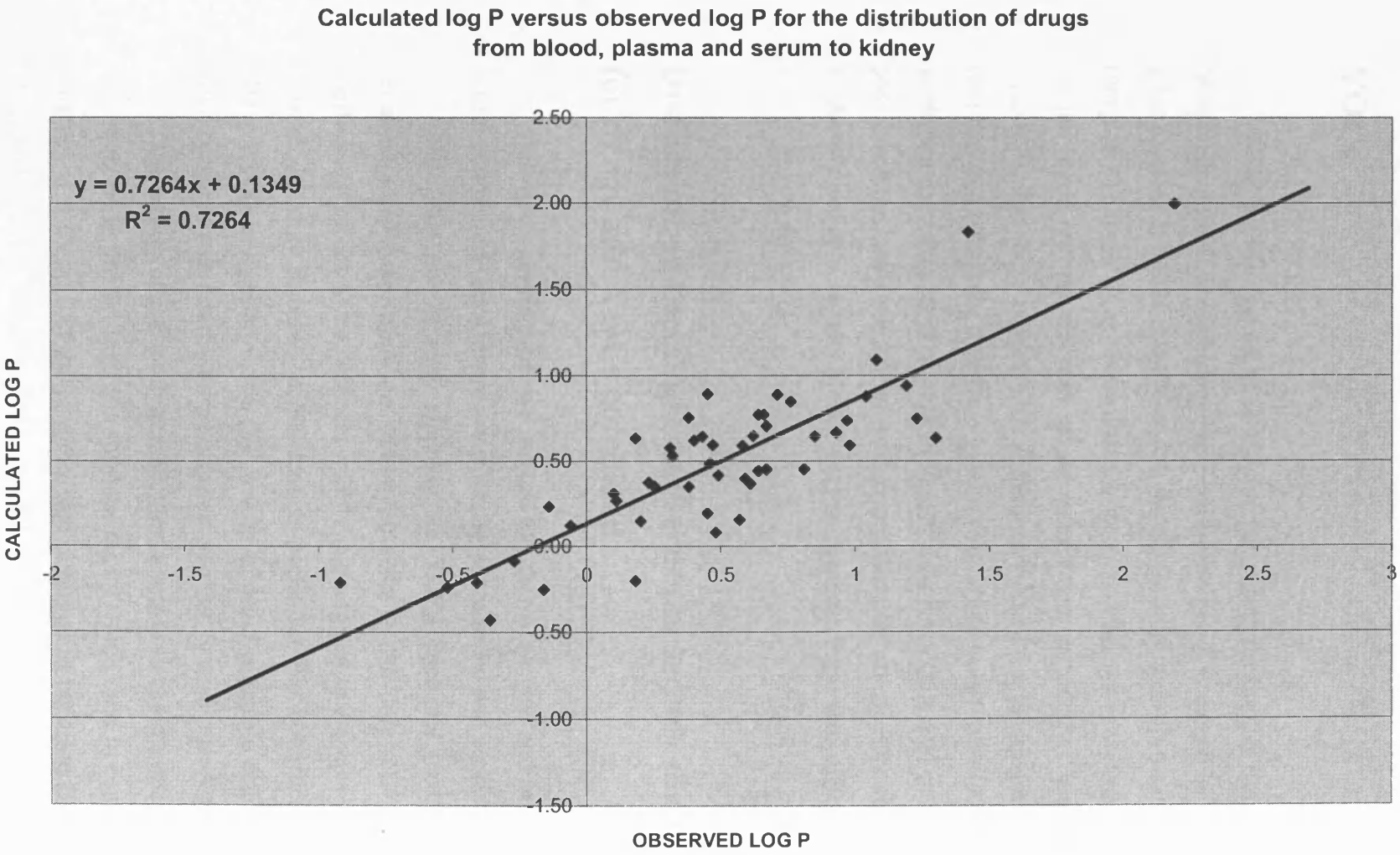
For the present model for blood (plasma or serum) to kidney distribution, there was one outlier observed. This outlier known as AI-6, has a residual of 1.2 log units which may possibly result from experimental error. The statistics of the equation for the full set (table 28.1) are good and the predictive capability of the equation can be assessed by the usual method of constructing a training set and an independent test set. The selection of the training (40 data points) and test sets (11 data points) was carried out using the usual computing method called Kennard-Stone. The training equation is also shown in table 28.1.

The statistics of the full set (table 28.1) and the training set (table 28.1), are similar and so are some of the coefficients. If the training set equation is used to predict the remainder of the data in the (unused) test set, it is found that AE = 0.00, AAE = 0.18, RMSE = 0.19 and SD = 0.20, that there is very little bias, and the predictive capability of the training equation (table 28.1) and by implication the full set equation (table 28.1) is therefore assessed. Thus, the full (*in vivo*) equation and the training equation can be used to predict further values of log BK, close to the suggested experimental error of SD = 0.3 and AAE = 0.2 log units (from chapter 21).

The inter-correlations of descriptors used for drug blood (plasma or serum) to kidney distribution are shown in table 28.2. The largest r^2 value for B versus V descriptor is 0.81. It appears that there is no strong inter-correlation between the other set of descriptors used in this model.

r	E	S	A	B	V	Ia
E	1.00					
S	0.83	1.00				
A	0.27	0.48	1.00			
B	0.58	0.74	0.56	1.00		
V	0.59	0.71	0.39	0.90	1.00	
Ia	0.09	0.18	0.42	0.00	-0.06	1.00

Table 28.2. The inter-correlation of descriptors for 50 compounds


Calculated log P versus observed log P for the distribution of drugs from blood, plasma and serum to kidney
y = 0.7264x + 0.1349
R2 = 0.7264
CALCULATED LOG P
OBSERVED LOG P
2.50
2.00
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-2
-1.5
-1
-0.5
0
0.5
1
1.5
2
2.5
3

Combining the *in vivo* data of drugs with the *in vitro* data of VOCs

Now that the data set of *in vivo* log BK values for passive transport from blood (plasma or serum) to rat kidney has been investigated, it would be interesting to see what happens if this is combined with the previous data set (*in vitro)* from table 13.3 (chapter 13) for VOCs. Note that the *in vitro* data set is transfer of VOCs from human blood and rat blood to human kidney and rat kidney. It was concluded in chapter 9, that there is very little difference in data for the species of humans and rats, and that data from these two species can be combined for VOCs. In the present work, the use of *in vitro* plasma to kidney data for VOCs will not be used, mainly because of the small number of extra compounds involved. The aim of this work is to show whether it is possible (or not) to combine these two different data sets (*in vitro* + *in vivo* data set).

Discussion on blood/plasma/serum to kidney distribution for drugs and VOCs combined

Before combining the two models together of *in vivo* and *in vitro* data sets, it is first important to see if there are any differences in the two equations.

Note that the *in vitro* equation for VOCs does not have the Ia-indicator variable, because there are no carboxylic acids in the data set. The two sets of coefficients appear to be different visibly and suggest that the *in vivo* and *in vitro* distributions may possibly be different. In particular, the *c*-coefficients differ 0.048 (SD = 0.144) as against -0.361 (SD = 0.089), which suggests that there might be a small systematic difference between the *in vivo* and *in vitro* distributions.

One way to check for systematic deviations is to use the coefficients in the *in vivo* equation to predict log BK values for the 73 VOCs used to construct the *in vitro* equation. It is found that the predicted and observed *in vitro* log BK values, AE = -0.03, AAE = 0.28, RMSE = 0.33 and SD = 0.33 log units. The numerical values here show that the *in vivo* equation can predict the *in vitro* dataset within 0.33 log units (SD). This suggests that the *in vivo* model can successfully predict the *in vitro*

dataset as it covers a larger descriptor space than the *in vitro* model. Any systematic difference between the *in vivo* (drugs) and the *in vitro* (VOCs) distributions could be due to the two series of compounds inhabiting different areas of chemical space.

Another indicator variable was then used to test if there is any difference in the *in vitro* and *in vivo* data set for VOCs and drugs. In this case one can apply the indicator variable, **Iv**, defined as **Iv** = 1 for the *in vitro* data and **Iv** = 0 for the *in vivo* data. An equation with the **Iv** term was constructed but is not shown because the **Iv** term was not statistically significant (T-test = 0.709, p-value = 0.480) and the value of the **Iv** coefficient was quite small (0.065). The coefficients of drug (*in vivo*) and the drug and VOC (*in vivo* + *in vitro*) combined equation are the same within experimental error (SD values of the coefficients are in the drug equation: c= 0.144, e= 0.086, s= 0.120, a= 0.158, b= 0.127, v= 0.110, ia= 0.125, and in the combined equation (*in vivo* + *in vitro*): c= 0.055, e= 0.062, s= 0.086, a= 0.123, b= 0.086, v= 0.065, ia= 0.109). It is concluded that there is very small systematic difference between the *in vivo* and *in vitro* distributions (as **Iv** is 0.1 log units), and that the two sets of distribution data can be combined without taking into account this difference (see equation 28.3).

The full data set (*in vivo* and *in vitro* measurements) consists of 120 compounds and 121 data points, the difference arising because blood (plasma or serum) to kidney distributions are counted separately. The correlation equation for this set of data points with the Abraham descriptors is shown in table 28.3. The fits for this full equation (in blue) appear to be reasonable (S.D = 0.259 and $r^2 = 0.711$). This represents a good model for a combined data set of VOCs and drugs for blood/plasma/serum to kidney distribution, and is the only model yet reported for both types of data sets (*in vivo* rats + *in vitro* rats and humans combined).

There were three additional outliers observed in the full set: methotrexate, terbinafine and r-etodolac (structures can be found in the appendix of this thesis). The residual for these compounds are: 0.7, -0.7 and -0.8 log units and may result from experimental error.

EQUATIONS	N(dp)	N(nc)	r^2	SD	F	OUTLIERS	D
FULL SET (*IN VIVO*) **Log BK** = 0.048 – 0.022**E** – 0.118**S** – 0.486**A** + 0.215**B** + 0.317**V** – 0.411**Ia**	51	50	0.726	0.299	19.465	1	6
FULL SET (*IN VITRO*) **Log BK** = -0.361 + 0.134**E** – 0.515**S** – 0.823**A** + 0.273**B** + 0.981**V**	73	73	0.757	0.217	41.661	3	5
FULL SET (*IN VIVO+IN VITRO*) **Log BK** = – 0.148 – 0.025**E** – 0.243**S** – 0.493**A** – 0.109**B** + 0.722**V** – 0.312**Ia**	121	120	0.711	0.259	46.815	3	6
TRAINING SET **Log BK** = – 0.014 – 0.013**E** – 0.221**S** – 0.491**A** + 0.014**B** + 0.559**V** – 0.345**Ia**	61	61	0.763	0.263	28.948	-	6
TEST SET **N**(nc) = 60 **N**(dp) = 60 **SD** (n-1) = 0.274 **RMSE** = 0.271 **AAE** = 0.235 **AE** = 0.046							

N (dp) = number of data points used in the training set
N (nc) = number of compounds used in the training set
r^2 = is the square of the overall correlation coefficient
SD = the standard deviation
F = F-statistic
D = the number of descriptors used in the model

Table 28.3. Summary of the two equations for blood/plasma/serum to kidney distributions for VOCs and drugs combined

In order to assess the predictive capability of the combined equation (*in vivo* + *in vitro*) a training set and was constructed and the equation for the training set used to predict log BK values for the test set, as shown in table 28.3. The coefficients and statistics of the full set of the combined equation (*in vivo* + *in vitro*) and training equation (*in vivo* + *in vitro*) are similar in terms of coefficients. The training equation predicted log BK values for the (unused) test sets as follows. N(nc) = 60 and N(dp) = 60, that AE = 0.05, AAE = 0.24, RMSE = 0.27 and the SD = 0.27 log units. Thus the combined (*in vivo* + *in vitro*) training equation and the full combined equation (in blue) can be used to predict further values of log BK to close to our suggested experimental error of around 0.2 log units (AAE) or 0.3 log units (SD).

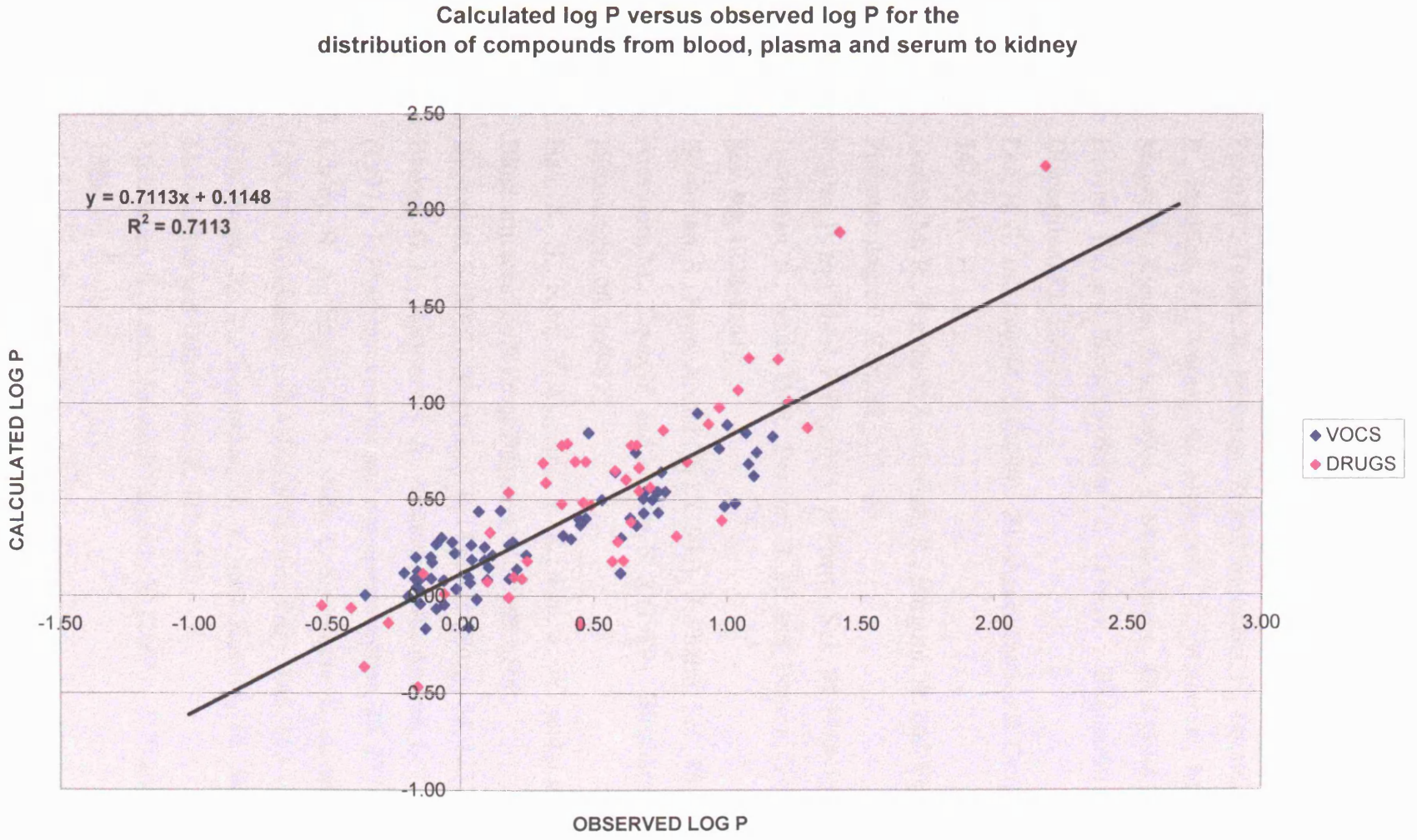
r	E	S	A	B	V	Ia
E	1.00					
S	0.92	1.00				
A	0.60	0.70	1.00			
B	0.78	0.87	0.74	1.00		
V	0.80	0.83	0.64	0.91	1.00	
Ia	0.26	0.31	0.41	0.18	0.17	1.00

Table 28.4. The inter-correlation of descriptors for 120 compounds

The inter-correlations of descriptors obtained for the combined *in vivo* and *in vitro* equations (from table 28.3) are shown by table 28.4. The largest r^2 value for E versus S ($r^2 = 0.85$) and B versus V ($r^2 = 0.83$) descriptor. It appears that there is no strong inter-correlation between the other set of descriptors used in this model.

The descriptors used in the equation reflect solute/solvent interactions. One can, therefore, interpret the coefficients in terms of the effect that particular interactions have on the process under consideration. The **e**-coefficient (P = 0.68) and **b**-coefficient (P = 0.21) is not significant. The positive **v**-coefficient in the blood/plasma/serum to kidney (*in vivo* + *in vitro*) equation, indicates that increasing molecular size can push drugs out of blood/plasma/serum and into the kidney phase. The coefficient of the 'polar' descriptors **S** is negative, indicating that polarity is stronger in the blood/plasma/serum phase than in the kidney tissue phase. Therefore the drug compounds will move from the kidney phase and into the blood/plasma/serum phase. The coefficient of the **A** descriptor is negative, indicating that the hydrogen-bond basicity ability act to keep drug compounds in the blood/plasma/serum phase and out of the kidney tissue. The negative coefficient of the indicator variable (***Ia***), in the combined equation from table 28.3., indicates a hitherto unknown factor affecting the transfer of solutes between blood/plasma/serum and the kidney. It is apparent that the presence of $-CO_2H$ group acts to hinder kidney penetration further. This could be simply due to intrinsic hydrogen bonding and the polarity properties of neutral acids. Finally, it is known that some acidic drugs may bind to albumin present in blood, plasma and serum, which may also account for this negative contribution of the coefficient ***Ia***.

This new model based on 120 compounds suggests that it can now be used to correlate and predict blood (plasma or serum) to kidney distribution for *in vitro* and *in vivo* drug compounds. Of course, any prediction using the full combined equation (*in vivo* plus *in vitro*) can, or should, only be made within the chemical space of the test set used to assess the predictions.


Calculated log P versus observed log P for the distribution of compounds from blood, plasma and serum to kidney
y = 0.7113x + 0.1148
R2 = 0.7113
CALCULATED LOG P
OBSERVED LOG P
2.50
2.00
1.50
1.00
0.50
0.00
-0.50
-1.00
-1.50
-1.00
-0.50
0.00
0.50
1.00
1.50
2.00
2.50
3.00
VOCS
DRUGS

References

1. Wojcikowski, J. and Daniel, W. A (2002). Pol. J. Pharmacol., **54**, 647-654.
2. Fazio, F., Todde, S., Moresco, R. M., Simonelli, P., Baraldi, P. G., Cacciari, B., Spalluto, G., Varani, K., Monopoli, A., Matarrese, M., Carpinelli, A., Magni, F., Kienle, G. K (2000). J. Med. Chem., **43**, 4359-4362.
3. Davila, D. and Kolacny-Babic, L (1991). Biopharmaceutics & Drug Disposition, **12**, 505-514.
4. Lee, H. S. and Lee, M. G (2000). Biopharmaceutics & Drug Disposition, **16**, 547-561.
5. Gasco, M. R., Fundaro, A., Cavalli, R., Bargoni, A. and Vighetto, D (2000). Pharmacological. Res., **42**, 337-343.
6. Poulin, P. and Theil, F. P (2002). J. Pharm. Sci., **91**,1358-1367.
7. Bjorkman, S., Wada, D. R., Berling, B. M. and Benoni, G (2001). J. Pharm. Sci., **90**, 1226-1241.
8. Bjorkman, S., Fyge, A. and Qi, Z (1996). J. Pharm. Sci., **85**, 887-889.
9. Aravagiri, M., Teper, Y. and Marder, S. R (1999). Biopharmaceutics & Drug Disposition, **20**, 369-377.
10. Han, K. S., Kim, Y. G., Yoo, J. K., Lee, J. W. and Lee, M. G (1998). Biopharmaceutics & Drug Disposition, **19**, 493-500.
11. Bjorkman, S (2002). Pharmacy and Pharmacology, **54**, 1237-1245.
12. Blakey, G. E., Nestorov, I. A., Arundel, P. A., Aarons, L. J. and Rowland, M (1997). J. Pharmacokinetics and Biopharmaceutics, **25**, 277-312.
13. Corley, R. A., English, J. C., Hill, T. S., Fiorcia, L. A. and Morgott, D. A (2000). Toxicology and Applied Pharmacology, **165**, 163-174.
14. Schillings, R. T., Sisenwine, S. F. and Ruelius, H. W (1997). Drug Metabolism and Disposition, **5**, 425-435.
15. Mclachlan, A. J and Hosseini-Yegabeh, M (2001). J. Pharm. Sci., **90**, 1817-1828.
16. Yamamoto, F., Oka, H., Antoku, S., Ichiya, Yu-ichi., Masuda, K. and Maeda, M (1999). Biol. Pharm. Bull., **22**, 590-597.
17. Haddad, S., Withey, J., Lapare, S., Law, F. and Krishnan, K (1998). Environmental Toxicology and Pharmacology, **5**, 245-255.

18. Bischoff, K. B., Dedrick, R. L., Zaharko, D. S. and Longstreth, J. A (1971). J. Pharm. Sci., **60**, 1128-1133.
19. Csanady, G. A., Oberste-Frielinghaus, H. R., Semder, B., Baur, C., Schneider, K. T. and Filser, J. G (2002). Arch. Toxicology, **76**, 299-305.
20. Parham, F. M., Mathews, H. B. and Portier, C. J (2002). Toxicology and Applied Pharmacology, **181**, 153-163.
21. Platts, J. A., Abraham, M.H., Zhao, Y.H., Hersey, A., Ijaz, L. and Butina, D (2001). Eur. J. Med. Chem., **36**, 719-730.
22. Salminen, T., Pulli, A. and Taskinen, J (1997). J. Pharmaceutical and Biomedical Analysis, **15**, 469-477.

Chapter

29

Conclusion on blood to tissue distribution for VOCs and drugs combined

The equilibrium distribution models for blood to tissue for drugs (VOCs and drugs combined) were set up using the Abraham equations (note that blood also represents plasma and serum as shown in earlier chapters). The following blood to tissue phase distributions that were examined were for the following tissues: brain, heart, liver, kidney, muscle, lung, skin and fat. The coefficients for the equations are summarised in tables 29.0 and 29.1.

The partition coefficients, $\mathbf{P}_{\text{blood:tissues}}$ for blood to tissues (VOCs and drugs combined) have been collected for a large number of compounds and measured in two different species, namely for humans and rats for VOCs, and rats for drugs. The general solvation equations from table 29.0 and 29.1 can be used to correlate the partition coefficients of these VOCs (*in vitro*) and drugs (*in vivo*) as log **P** values. The equations from table 29.0 and 29.1 can be used for compound design. For example if one has a set of compounds with low partition into a particular tissue and it is required to synthesise related compounds with increased partition coefficients. Predictions can be made on compounds before they are synthesised and then it is only necessary to synthesise and test those most likely to have the right properties.

It has been found that human and rat partition data are the same within a given experimental error, as shown in chapter 9, for VOCs. This data set can also be combined with the dataset of drugs in rats to make new models to correlate and predict blood to tissue distribution as shown earlier in this thesis (chapter 21-28).

Table 29.0 is a summary of the equations constructed for blood to tissue distribution in rats for drugs only. These set of equations, were necessary in order to

see if there were similarities to the VOC models. Finally, it was useful to examine how equations depend on tissue composition for the combined data sets (VOCs and drugs).

BLOOD: TISSUE	e	s	a	b	v	Ia	TISSUE COMPOSITION (WEIGHT FRACTIONS)		
							LIPID	PROTEIN	WATER
BRAIN	0.183	-0.778	-0.697	-0.462	0.679	-1.199	0.107	0.079	0.790
HEART	0.280	-0.399	0.189	0.066	0.093	-1.003	0.100	0.167	0.727
LIVER	0.119	-0.077	-0.276	0.149	0.194	-0.871	0.070	0.180	0.720
KIDNEY	-0.022	-0.118	-0.486	0.215	0.317	-0.411	0.050	0.170	0.770
MUSCLE	-0.059	0.010	-0.248	0.028	0.110	-1.022	0.020	0.170	0.790
LUNG	0.031	-0.174	-0.666	0.248	0.321	-0.833	0.010	0.177	0.780
FAT	-	-	-	-	-	-	0.800	0.050	0.150
SKIN	0.071	0.085	-0.468	-0.490	0.134	-0.674	-	-	-
BLOOD	0.000	0.000	0.000	0.000	0.000	0.000	0.007	0.180	0.800

Table 29.0. Summary of equations for blood to tissue distribution (drugs only), and the tissue composition (lipid, protein and water) in humans.[1]

Coefficients e, s, a, b, v and Ia versus the tissue composition for blood to tissue distribution of VOCs and drugs combined

It was found that plots of the coefficients **e**, **s**, **a**, **b**, **v** and **Ia**-coefficients versus weight fraction of proteins were so scattered that little could be deduced. An attempt was made to compare the coefficients (**e**, **s**, **a**, **b**, **v** and **Ia**) of the equations (table 29.1) with the composition of the tissues in terms of weight fractions of lipid and water. The compositions are given for humans (and not rats), as this is the only composition that is available from the literature.[1] Visual comparisons are shown in schemes 29.0-29.11 using the program MINIITAB.[2]

The coefficients in the equations for blood to tissue distribution are harder to interpret than those for air to tissue distribution. The coefficients in the blood to tissue equations (VOCs and drugs combined) are much smaller than those in the air to tissue equations, because of partial cancellation of effects. This is because the

coefficients in the blood to tissue equations reflect the difference in interactions between compounds and blood and compounds and tissue, whereas the coefficients in the air to tissue equations reflect only the interactions with the tissue. In addition, the percentage error in the blood to tissue coefficients will be much larger than the percentage error in the air to tissue coefficients. This is another factor that makes interpretation of the blood to tissue equation coefficients much more difficult, especially when the points in the plots are all clustered together, for example in schemes 29.2 to 29.4. From these plots one can establish that there is not great deal of connection between the equation coefficients and tissue composition.

In Table 29.1., are given the final equation coefficients for blood to tissue distribution of VOCs and drugs combined. Included in this table, is the blood to blood equation for which the coefficients are zero, by definition.

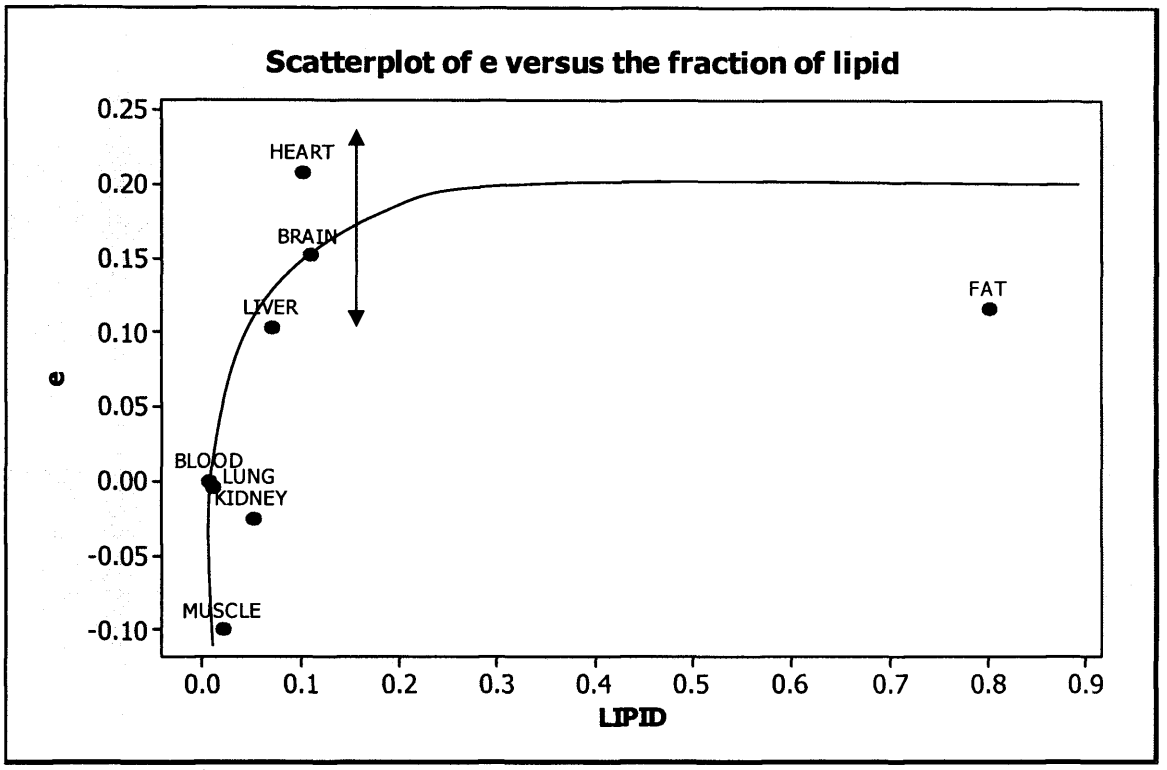
BLOOD: TISSUE	e	s	a	b	v	la	lv	TISSUE COMPOSITION (WEIGHT FRACTIONS)		
								LIPID	PROTEIN	WATER
FAT	0.117	-0.099	-1.543	-2.106	1.596	-0.946	0	0.800	0.050	0.150
BRAIN	0.152	-0.756	-0.664	-0.382	0.652	-1.196	-0.487	0.107	0.079	0.790
HEART	0.209	-0.393	0.076	-0.046	0.279	-0.965	0	0.100	0.167	0.727
LIVER	0.104	-0.351	-0.202	-0.127	0.617	-0.831	0	0.070	0.180	0.720
KIDNEY	-0.025	-0.243	-0.493	-0.109	0.722	-0.312	0	0.050	0.170	0.770
MUSCLE	-0.100	-0.080	-0.254	0.041	0.233	-1.005	-0.140	0.020	0.170	0.790
LUNG	-0.004	-0.036	-0.742	0.252	0.312	-0.734	0	0.010	0.177	0.780
BLOOD	0.000	0.000	0.000	0.000	0.000	0.000	0	0.007	0.180	0.800

Table 29.1. Summary of equations for blood to tissue distribution (VOCs and drugs combined), and the tissue composition (lipid, protein and water) for humans.[1]

When the lipid fraction is increased in the blood to tissue equations, the **e**-coefficients for the equations appear to get larger and reach a plateau. This indicates that the London dispersion interaction is larger for partitions into lipids than into the polar water and protein constituents. This is shown in scheme 29.0. For blood to lung equation, the **e**-coefficient is very small at -0.004 units. These figures would suggest that the blood and lung have about the same London dispersion interactions, and this is why this point appears near the top left of the plot in scheme 29.1.

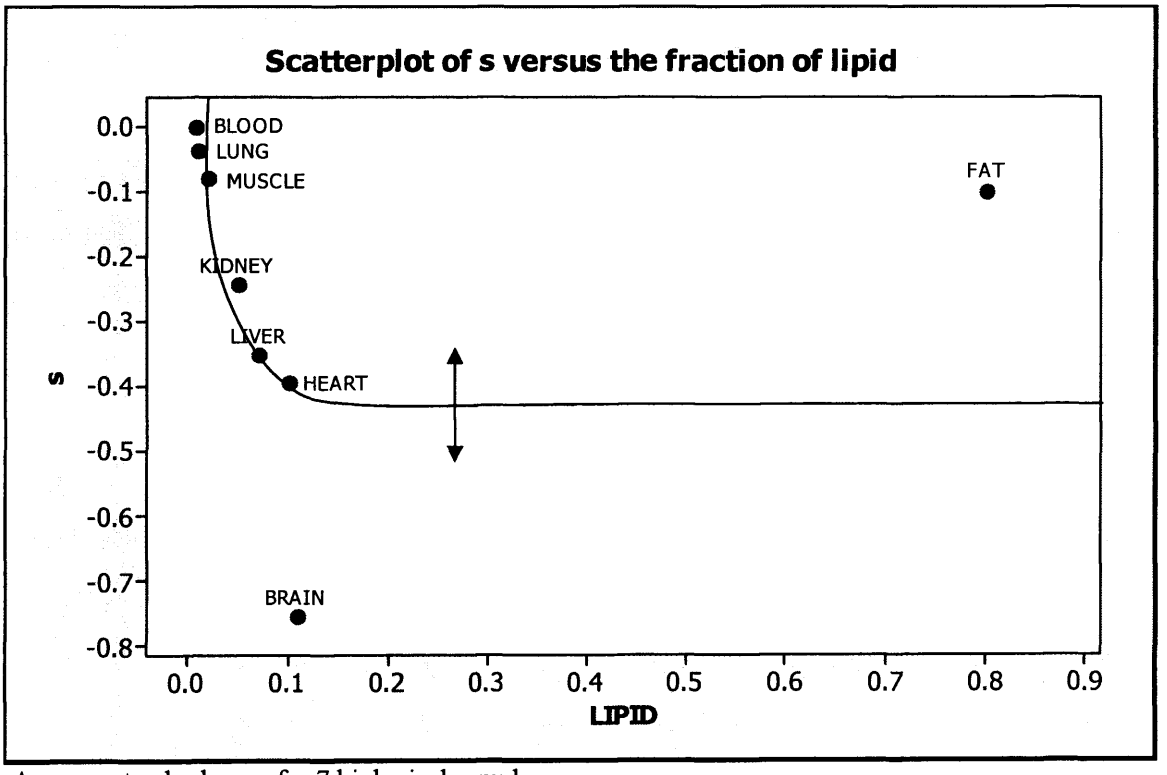
If the lipid fraction is increased in the blood to tissue model, the **s**-coefficient for the equations gets smaller, see scheme 29.1. This is reasonable, because the tissue becomes less polar and less positive. Note that the coefficients for brain and heart do not follow this trend very well. This could be because the data set does not represent the true chemical space for that particular tissue. This problem is also observed later with fat tissue as shown in schemes 29.2-29.11, where the data point for fat is far away from the other data points.

The plot of the **a**-coefficients against the lipid fraction is shown in scheme 29.2. It is found that as the lipid fraction increases in the tissue (thus the latter becomes less polar), so the **a**-coefficient gets smaller. If the lipid fraction is increased in tissues, the **b**-coefficient for the equations also gets smaller, see scheme 29.3. This is reasonable, because as the tissue becomes less polar, so it becomes less basic and less acidic. As the lipid fraction is increased in tissues, the **v**-coefficient for the equations increases, as shown in scheme 29.4. This is also reasonable because the **v**-coefficient is an indication of the lipophilicity of a phase. In this case fat has the highest lipid fraction of all the tissues and this data point is observed at the top right of the plot. Finally, as the lipid fractions were increased in tissues, the **Ia**-coefficient for the equations decreases and reaches a plateau, see scheme 29.5. It is apparent that the presence of $-CO_2H$ group in drugs acts to hinder tissue penetration further, simply due to intrinsic hydrogen bonding and the polarity properties of neutral acids. This suggests that the fat tissue is least polar out of all tissues due high lipid content (as shown by graph) and is less likely to interact with drugs via hydrogen bonding than the blood or kidney phase (both of which are more polar and able to accommodate carboxylic acid containing drug compounds).



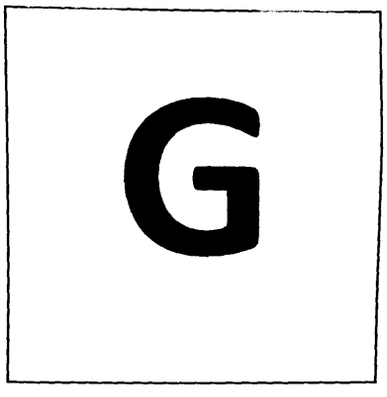
Average standard error for 7 biological samples

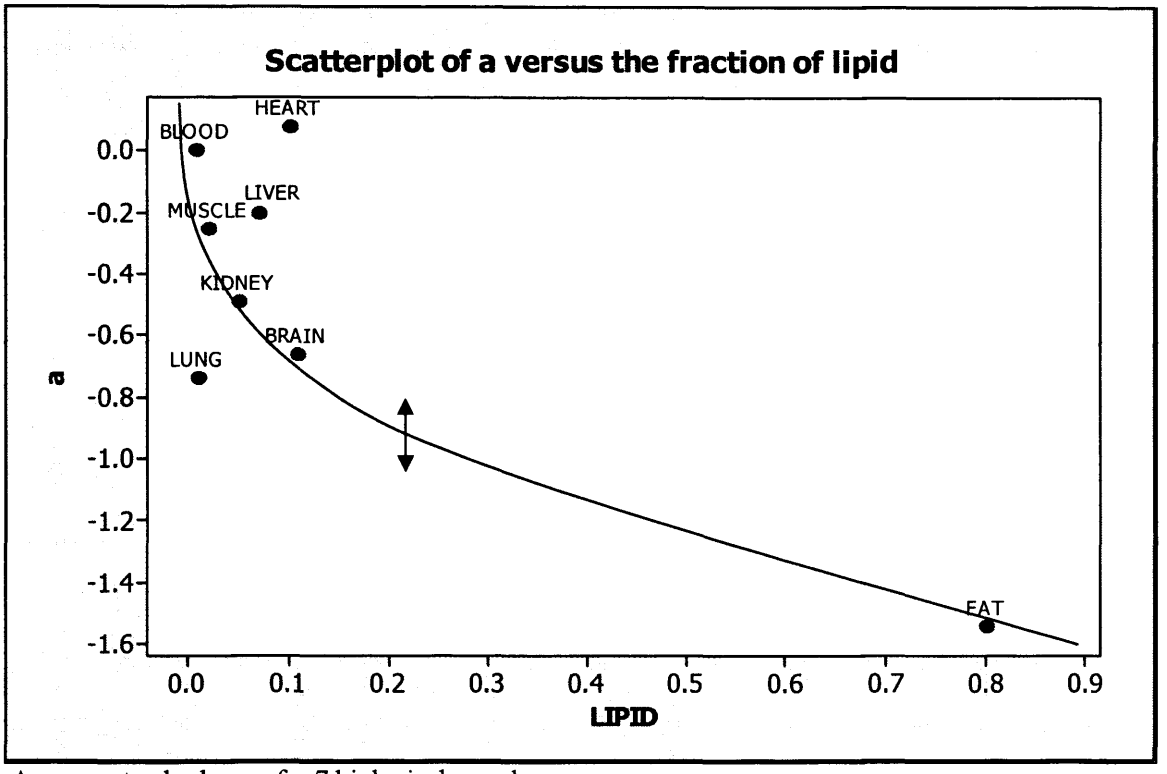
Scheme 29.0. Scatter plot for **e** coefficient versus the fraction of lipid for blood to tissue distribution (VOCs and drugs combined)



Average standard error for 7 biological samples

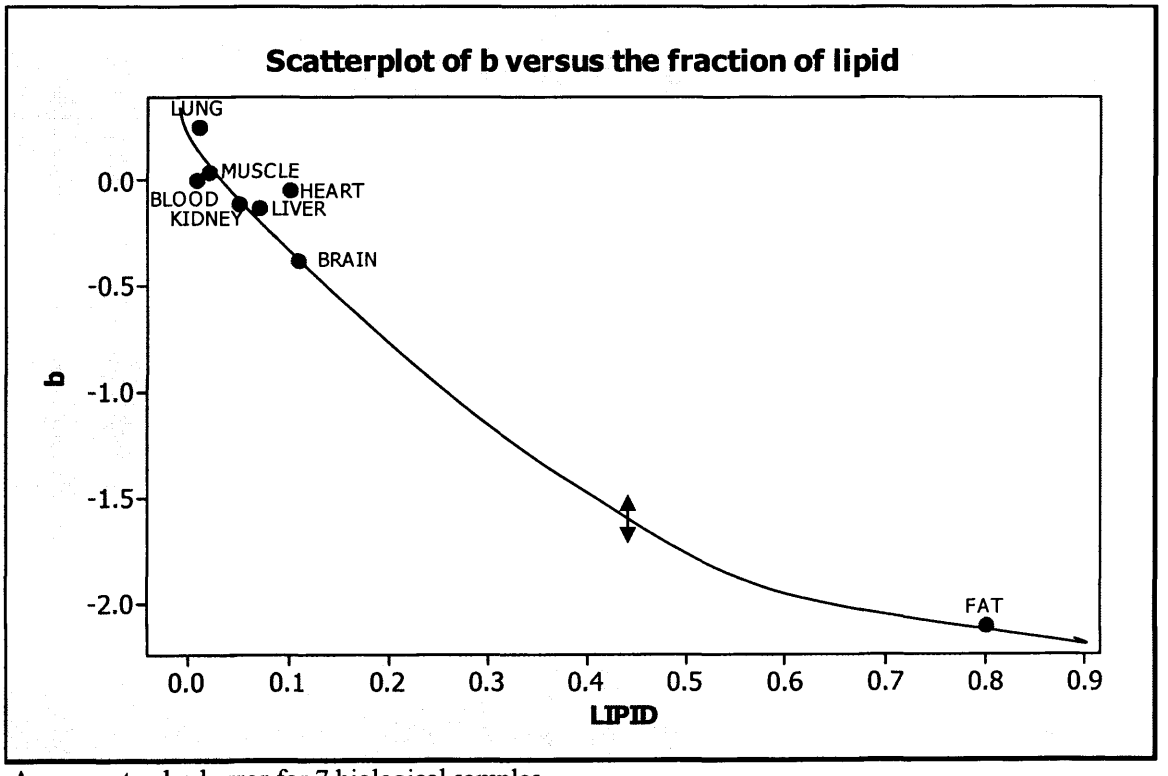
Scheme 29.1. Scatter plot for **s** coefficient versus the fraction of lipid for blood to tissue distribution (VOCs and drugs combined)


G



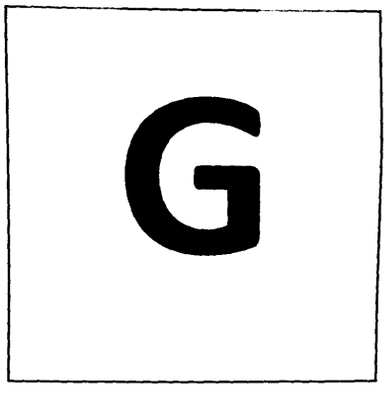
Average standard error for 7 biological samples

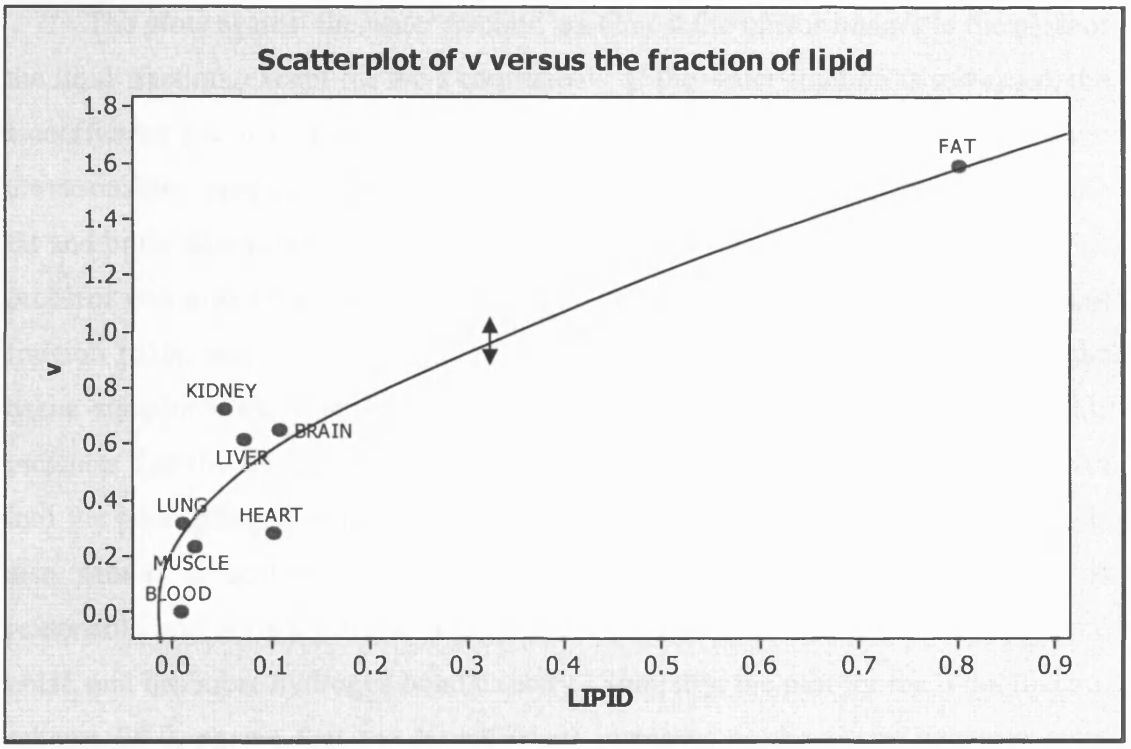
Scheme 29.2. Scatter plot for **a** coefficient versus the fraction of lipid for blood to tissue distribution (VOCs and drugs combined)



Average standard error for 7 biological samples

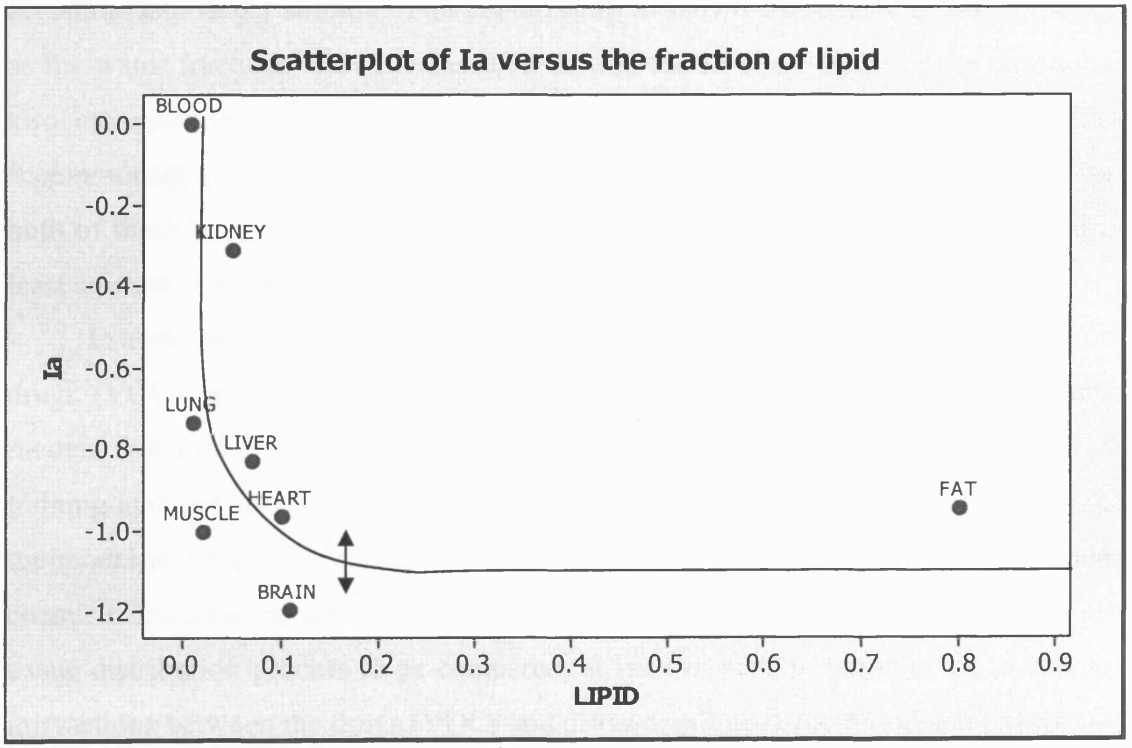
Scheme 29.3. Scatter plot for **b** coefficient versus the fraction of lipid for blood to tissue distribution (VOCs and drugs combined)


G



Average standard error for 7 biological samples

Scheme 29.4. Scatter plot for **v** coefficient versus the fraction of lipid for blood to tissue distribution (VOCs and drugs combined)



Average standard error for 7 biological samples

Scheme 29.5. Scatter plot for **Ia** coefficient versus the fraction of lipid for blood to tissue distribution (VOCs and drugs combined)

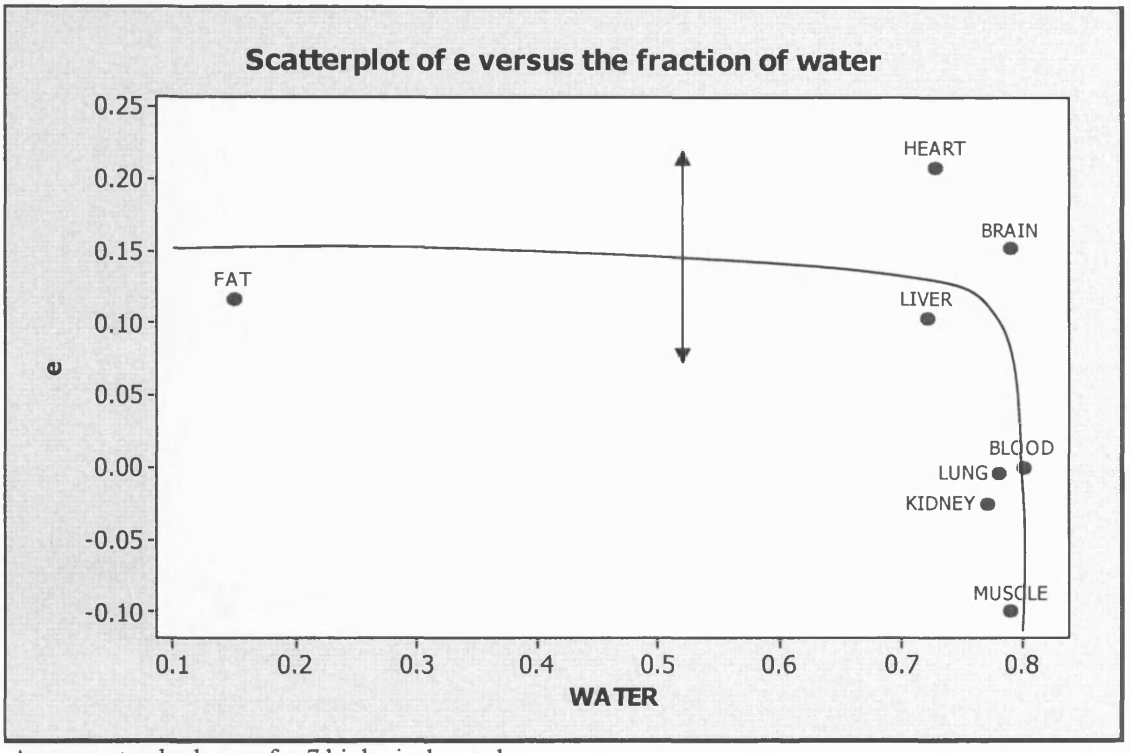
The plots against the water fraction are almost the mirror images to the plots of the lipid fraction, except for the **s**-coefficient. If the water fraction is increased, the **s**-coefficient for the equations gets smaller, see scheme 29.7. This seems quite unreasonable. However, this graph is difficult to interpret as one can not tell if the fat and brain data point are reliable data points or if they are out of line again. This problem was also observed for the data points of fat and brain in some of the lipid fraction plots, as mentioned earlier. When the water fraction is increased in the tissue equations, the **e** -coefficients for the equations appear to get smaller. This indicates that the London dispersion interaction is larger for partitions into fats than into the polar phase like kidney and muscle that have high water fractions. This is also shown in scheme 29.6. The plot for the **a**-coefficient, scheme 29.8, is reasonable and shows that the **a**-coefficient increases as the tissue becomes more polar, and has more hydrogen bond basicity. Similarly, the plot for the **b**-coefficient, scheme 29.9, shows that the **b**-coefficient increases as the tissue becomes more polar, and has more hydrogen bond acidity. Also, as the water fraction is increased in tissues, the **v**-coefficient for the equations gradually decreases. This demonstrates that the tissues with larger water content are less lipophilic and less likely to accommodate larger solutes. This relationship is shown by scheme 29.10. Finally, as the water fractions were increased in tissues, the **Ia**-coefficient for the equations also increased (see scheme 29.11). This suggest that blood and kidney are able accommodate carboxylic acid containing compounds (via hydrogen bonding), as both of these tissue phases are more polar than the fat phase (which in turn has the least amount of water, hence less polar).

In conclusion, we show that for a large data set of blood to tissue partitions for drugs (VOCs and drugs combined), it is possible to construct statistically sound models and to assess the predictive capability of the models through selection of training and test sets of drugs (VOCs and drugs combined). A particular feature of the model is that the coefficients obtained are not just fitting parameters but encode chemical information about the nature of the process. This enables the blood to tissue distribution process to be compared, at least in part, in terms of the chemical interactions between the drugs (VOCs and drugs combined) and blood and tissues.

However, the comparisons of equation coefficients with tissue composition are not very good, and there seem to be many anomalies. It is unlikely that this is due to the compounds in the various equations occupying different areas of chemical space,

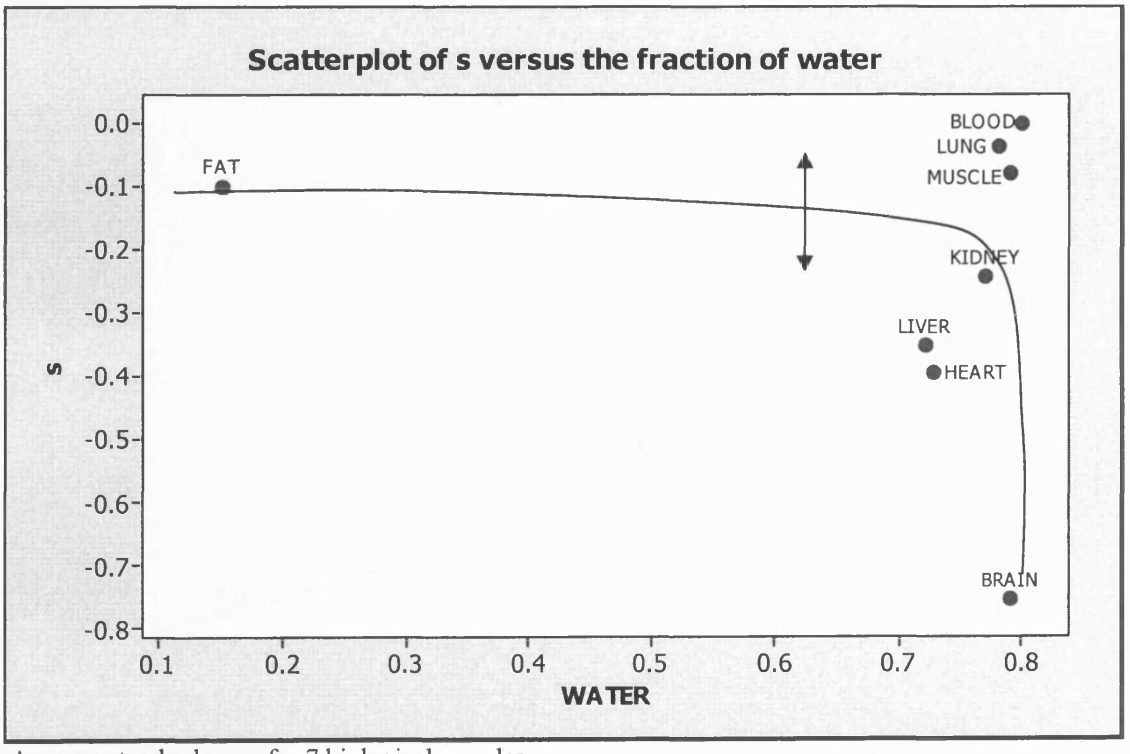
because similar sets of compound are used in the various equations. Another possibility for the lack of connection between the equation coefficients and tissue composition is more fundamental. When a tissue is notionally divided up into separate components, water, protein and lipid, it is implied that, for example, the water in a tissue is somehow the same as water in bulk water, and that water in one tissue is somehow the same as water in another tissue. This cannot be completely correct. The properties of water in a tissue are unlikely to be the same as the properties of water in another tissue or as bulk water.

For example, Park et al.[3] have measured the hydrogen bond acidity, a(m), of solvent mixtures. An interesting example is the water-tetrahydrofuran (THF) mixture, because the acidity is due only to the water component and not THF. A plot of α_m versus the solvent composition, as volume % is shown in Scheme 29.12. If the properties of water were the same at all compositions, the plot should be a straight line, but it is clear that the acidity of water in the mixed solvent is not the same as the acidity of bulk water. The conclusion would be the same no matter what composition scale was chosen. There have been several models for gas to tissue and blood or plasma to tissue partition that include tissue compositions,[1,4-6] but these models are not satisfactory if they rely on properties of tissue components to be the same as properties of the bulk components.



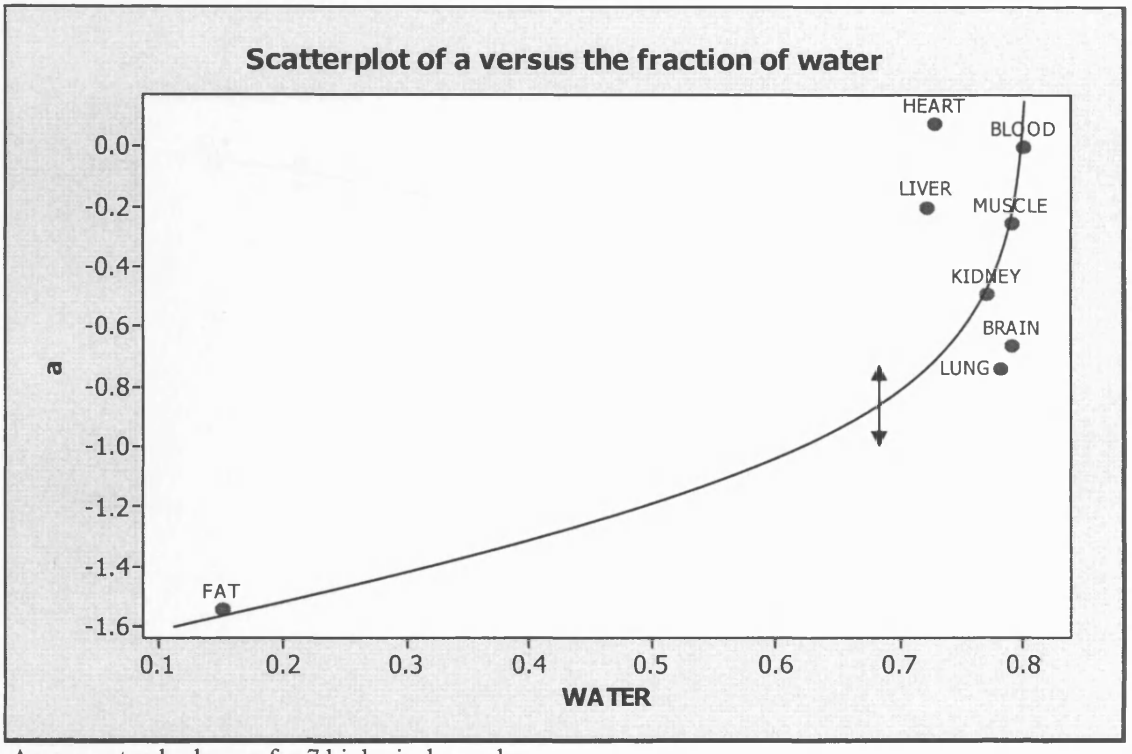
Average standard error for 7 biological samples

Scheme 29.6. Scatter plot for **e** coefficient versus the fraction of water for blood to tissue distribution (VOCs and drugs combined)



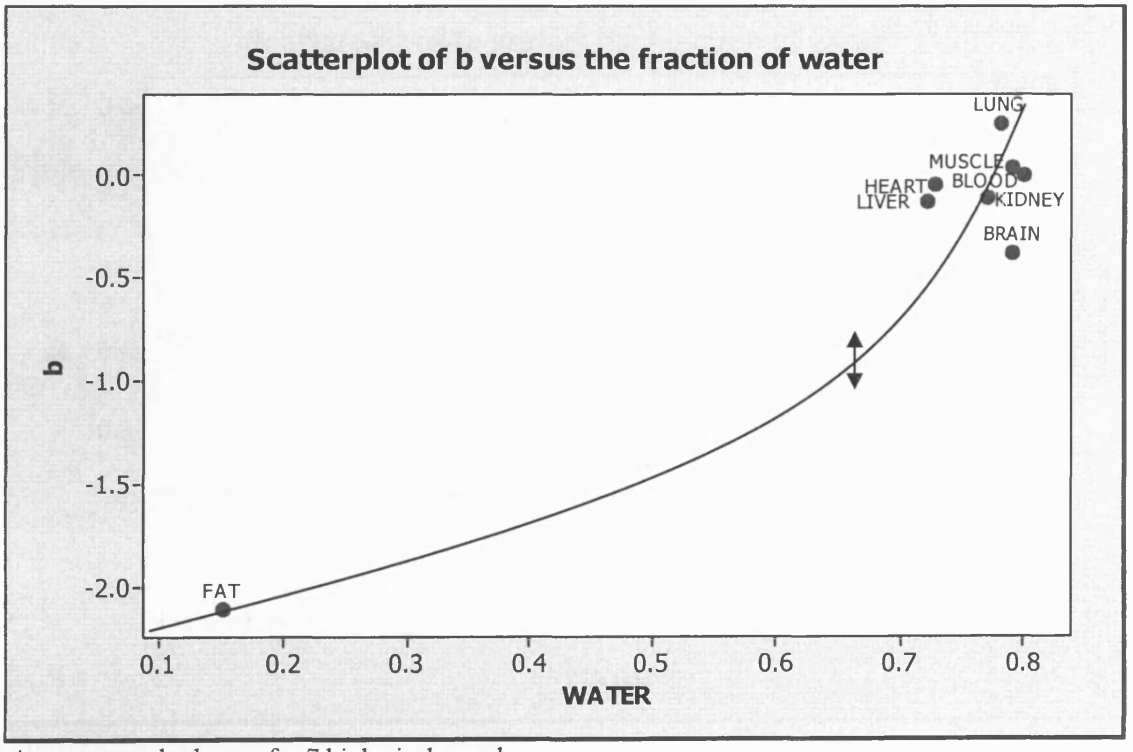
Average standard error for 7 biological samples

Scheme 29.7. Scatter plot for **e** coefficient versus the fraction of water for blood to tissue distribution (VOCs and drugs combined)



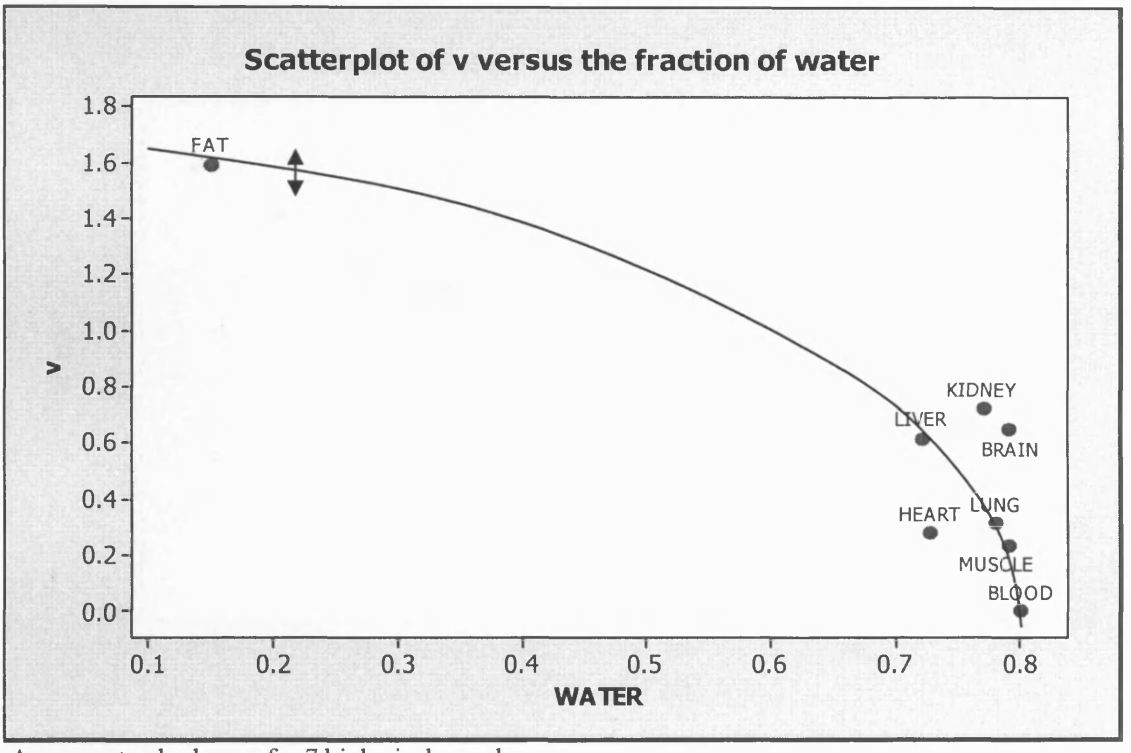
Average standard error for 7 biological samples

Scheme 29.8. Scatter plot for **a** coefficient versus the fraction of water for blood to tissue distribution (VOCs and drugs combined)



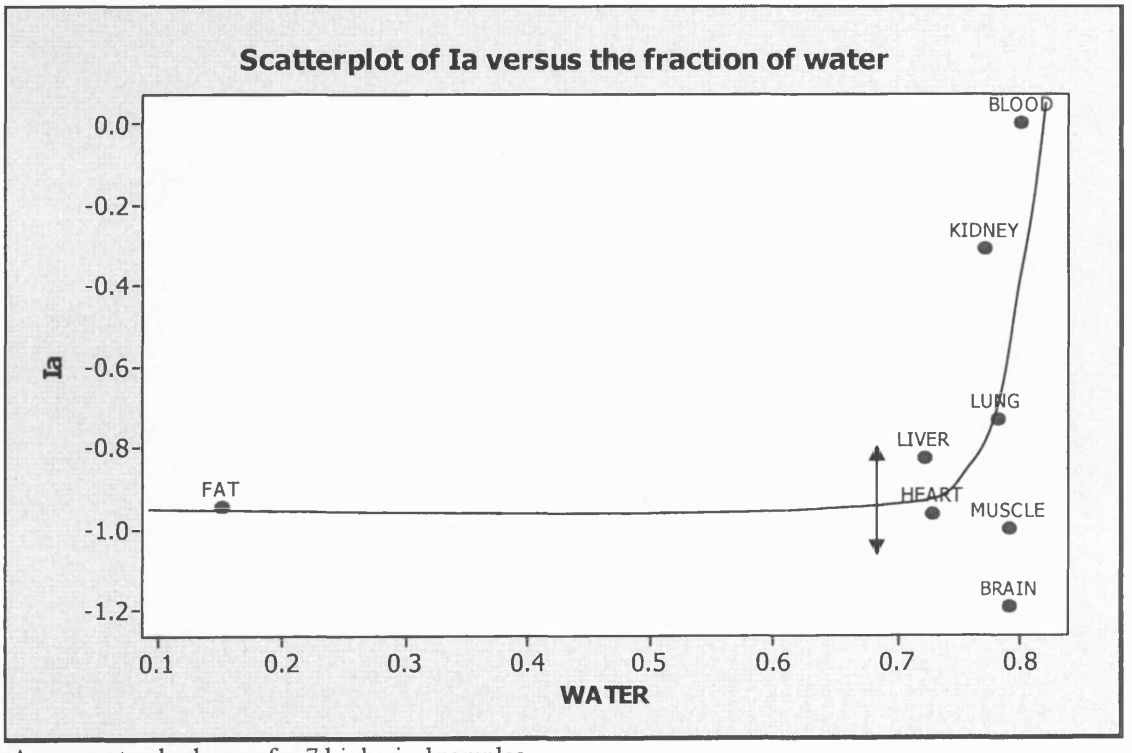
Average standard error for 7 biological samples

Scheme 29.9. Scatter plot for **b** coefficient versus the fraction of water for blood to tissue distribution (VOCs and drugs combined)



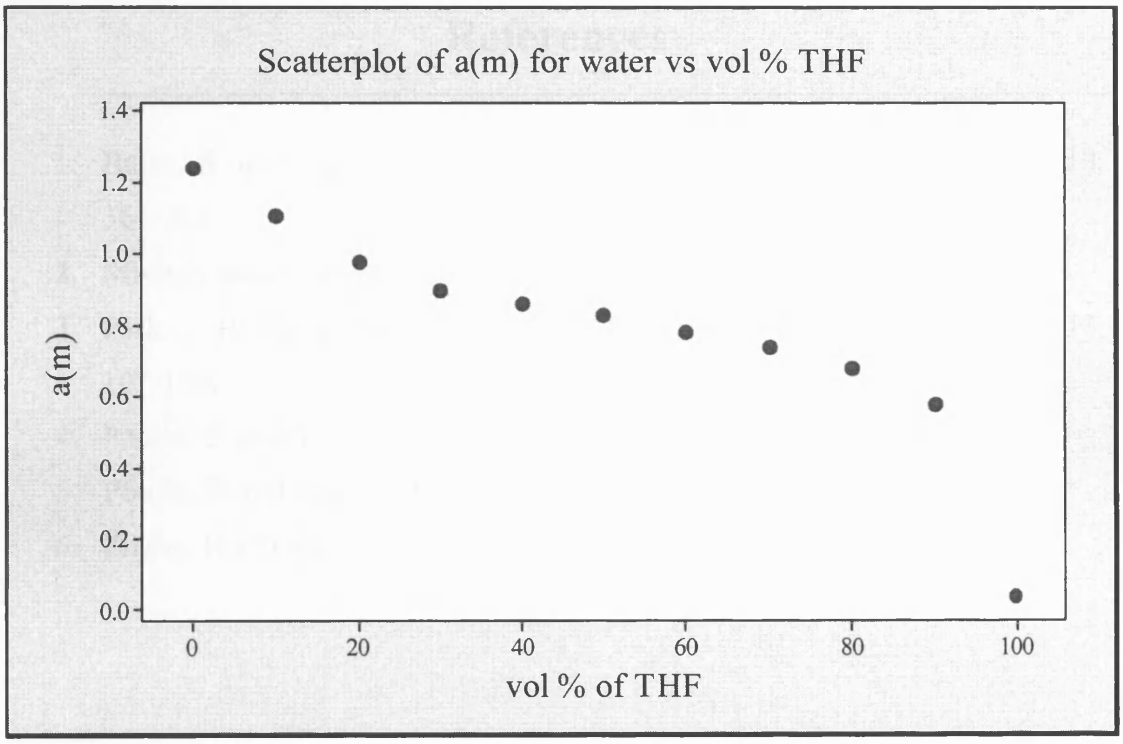
Average standard error for 7 biological samples

Scheme 29.10. Scatter plot for **v** coefficient versus the fraction of water for blood to tissue distribution (VOCs and drugs combined)



Average standard error for 7 biological samples

Scheme 29.11. Scatter plot for **Ia** coefficient versus the fraction of water for blood to tissue distribution (VOCs and drugs combined)



Scheme 29.12. A plot of a(m) versus the solvent composition of THF, as volume %

References

1. Balaz, S. and Lukacova, V (1999). Quant. Struct.-Act. Relationships, **18**, 361-368.
2. Minitab statistical software (version 14), www.minitab.com
3. Park, J. H., Jang, M. D. and Kim, D. S (1990). J. Chromatography., **513**, 107-116.
4. Poulin, P. and Theil, F. P (2000). J. Pharm. Sci., **89**, 16-35
5. Poulin, P. and Theil, F. P (2002). J. Pharm. Sci., **91**, 1358-1370
6. Zhang, H (2004). J. Pharmaceutical Sciences, **93**, 1595-1604.

Chapter

30

Conclusion and future work

Conclusion

This work contains the largest sets of literature values for *in vitro* and *in vivo* distribution of VOCs and drugs from air to blood (or air to tissue) and blood to tissue in rats and humans. Models based on the Abraham equations have been constructed to predict air to blood, air to tissue and blood to tissue distributions for VOCs in rats and humans, as shown in chapters 9-20. The Abraham equations have also been constructed for blood to tissue distribution for drugs in rats only (see chapters 21-29). This data on drugs was later combined with VOC data to enable the construction of new, extended equations, to correlate and predict drug and VOC blood to tissue distribution.

It was shown in this work that for a large data set of air to tissue and blood to tissue partitions for VOCs and drugs, it is possible to construct statistically sound models and to assess the predictive capability of the models, through selection of training and test sets. Although there have been previous attempts to model air to tissue and blood to tissue distributions, these have all used much smaller data sets and, except for blood to brain distributions, have not assessed at all the predictive capability of the obtained models.

The determination of blood to tissue distributions for drugs is difficult (due to compound instability and confounding impurities) and time-consuming, requiring animal experiments and synthesis (sometimes in radiolabelled form) of the compounds to be tested. That is why it is desirable that equations can be constructed

to predict these distributions computationally with reasonable accuracy. This will therefore allow rejection of unsuitable candidates early on in drug discovery and will have a significant impact on drug research and development.

Future work

There are new data being published on blood to brain distribution for drugs, with respect to different compartments of the brain. These include: pituitary gland, hypothalamus, hippocampus, striatum, thalamus, cerebellum, amygdyla, caudate putamen, frontal cortex, globus pallidus, hippocampus, nucleus accumbens and septum. There are possibilities for future work to make models to predict distribution of drugs from blood to the different regions of the brain at steady-state concentrations.

The databases of compounds studied in the blood to tissue distribution model are mainly VOCs and drug compounds. Currently, there are new datasets being published for dioxin blood to tissue distribution in rats. It would be interesting to see how the various equations may look, if this new dataset is combined with the current model used in this thesis. Unfortunately, due to limited time, this was not possible during this project and would be a very suitable subject for future work.

Above a third of drugs are zwitterions, thus it would be useful to find a way in the future to obtain descriptors for these compounds, in order to predict blood to tissue distributions for zwitterionic drugs.[1] Recently, there have been published a small amount of data for blood to tissue distribution, for the following tissues: adipose, bone, testes, ovary, stomach, pancreas, bile, adrenal gland, placenta, cerebrospinal fluid, trachea, ileum and spleen. It would be very interesting to find out if it is possible to obtain correlative and predictive models for these tissues. If so, how would the new equations compare in terms of coefficients, with the current models obtained in the present work?

In addition, small amounts of drug data (*in vivo*) have recently been published for blood to tissue distribution in the following species: humans, pregnant rats, mouse, rabbits, monkeys, guinea pigs, and dogs. It would be interesting to find out whether it is possible to make correlative and predictive models for blood to tissue distribution in these species too.

References

1. Yamazaki, K. and Kanaoka (2004). Computational predication of the plasma protein-binding percent of diverse pharmaceutical compounds. J. Pharm. Sci., 93, 1480-1494.

Appendix

Tables

Tables for distribution data

Table 9.0. Air to blood distribution for VOCs

SOLUTE NAME	E	S	A	B	L	HUMAN REF	HUMAN LOG K	RAT REF	RAT LOG K	COMBINED REF	AVERAGE LOG K
1,1,1,2-Tetrachloroethane	0.542	0.630	0.100	0.080	3.641	1, 6, 7	1.48	6, 7	1.62	1, 6, 7	1.55
1,1,1-Trichloroethane (Methyl chloroform)	0.369	0.410	0.000	0.090	2.733	1, 4, 5, 6, 7	0.50	6, 7	0.76	1, 4, 5, 6, 7	0.63
1,1,2,2-Tetrachloroethane	0.595	0.760	0.160	0.120	3.803	1, 4, 6, 7,	2.11	6, 7,	2.15	1, 4, 6, 7	2.13
1,1,2-Trichloroethane	0.499	0.680	0.130	0.130	3.290	1, 4, 6, 7	1.58	6, 7	1.76	1, 4, 6, 7	1.67
1,1-Dichloroethane	0.322	0.490	0.100	0.100	2.316	1, 4, 6, 7	0.70	6, 7	1.05	1, 4, 6, 7	0.88
1,2,4-Trimethylbenzene (Pseudocumene)	0.677	0.560	0.000	0.190	4.441	7,	1.77	16, 17	1.16	7, 16, 17	1.47
1,2-Dichloroethane	0.416	0.640	0.100	0.110	2.573	1, 4, 6, 7	1.30	6, 7	1.48	1, 4, 6, 7	1.39
1,2-Dichloropropane	0.371	0.680	0.000	0.150	2.866	1, 4, 6, 7	1.01	6, 7	1.27	1, 4, 6, 7	1.14
1-Bromo-2-chloroethane (1-Chloro-2-Bromoethane)	0.572	0.700	0.100	0.090	2.982	6, 7	1.47	6, 7	1.72	6, 7	1.60
1-Butanol	0.224	0.420	0.370	0.480	2.601	4, 7,	2.97	7,	3.19	4, 7	3.08
1-Chloropropane (Propyl chloride)	0.216	0.400	0.000	0.100	2.202	1, 6, 7,	0.46	6, 7,	0.72	1, 6, 7	0.59
1-Nitropropane	0.240	0.950	0.000	0.310	2.894	6, 7	2.27	6, 7	2.35	6, 7	2.31
1-Pentanol	0.219	0.420	0.370	0.480	3.106	7,	2.73	7,	2.92	7,	2.83
1-Propanol	0.236	0.420	0.370	0.480	2.031	1, 5, 7	2.99	7,	3.13	1, 5, 7	3.06
2,2,4-Trimethylpentane (Isooctane)	0.000	0.000	0.000	0.000	3.106	6, 7	0.20	6, 7	0.25	6, 7	0.23
2-Chloropropane	0.177	0.350	0.000	0.120	1.970	6, 7	0.14	6, 7,	0.49	6, 7	0.32

2-Heptanone (Methyl pentyl ketone)	0.123	0.680	0.000	0.510	3.760	1, 7	2.30	7,	2.35	1, 7	2.33
2-Methyl-1,3-butadiene (Isoprene)	0.313	0.230	0.000	0.100	2.101	2,	-0.12	2, 6, 7	0.32	2, 6, 7	0.10
2-Methyl-1-propanol (Isobutanol)	0.217	0.390	0.370	0.480	2.413	1, 4, 5, 7	2.89	7,	2.94	1, 4, 5, 7	2.92
2-Nitropropane	0.216	0.920	0.000	0.330	2.550	6, 7	2.19	6, 7	2.26	6, 7	2.23
2-Pentanone (Methyl propyl ketone)	0.143	0.680	0.000	0.510	2.755	1,	2.18	7,	2.10	1, 7	2.14
2-Propanol (Isopropanol)	0.212	0.360	0.330	0.560	1.764	1, 4, 5, 7	2.92	7,	3.11	1, 4, 5, 7	3.02
3-Methyl-1-butanol (Isopentanol) (Isoamyl alcohol)	0.192	0.390	0.370	0.480	3.011	7,	2.58	7, 22	2.92	7, 22	2.75
4-Chlorobenzotrifluoride	0.530	0.580	0.000	0.010	3.730	10,	1.22	10,	1.64	10,	1.43
4-Methyl-2-pentanone (Methyl isobutyl ketone)	0.111	0.650	0.000	0.510	3.089	1, 4, 7,	2.01	7,	1.90	1, 4, 7	1.96
Acetone (Propan-2-one)	0.179	0.700	0.040	0.490	1.696	1, 4, 5, 7	2.35	7, 22	2.37	1, 4, 5, 7, 22	2.36
Benzene	0.610	0.520	0.000	0.140	2.786	1, 5, 6, 7	0.87	6, 7, 17	1.22	1, 5, 6, 7, 17	1.05
Carbon tetrachloride (Tetrachloromethane)	0.458	0.380	0.000	0.000	2.823	1, 4, 6, 7,	0.57	6, 7,	0.66	1, 4, 6, 7	0.62
Bromochloromethane (Chlorobromomethane)	0.541	0.800	0.010	0.060	2.445	4,	0.79	6, 7,	1.62	4, 6, 7	1.21
Bromodichloromethane	0.593	0.690	0.100	0.040	2.891	32,	1.42	31, 33	1.56	31, 32, 33	1.49
Butyl acetate	0.071	0.600	0.000	0.450	3.353	7,	1.92	7,	1.95	7,	1.94
Butan-2-one (Methyl ethyl ketone)	0.166	0.700	0.000	0.510	2.287	1, 4, 5, 7	2.19	7,	2.28	1, 4, 5, 7	2.24
CF3CBrClH (Halothane)	0.102	0.380	0.150	0.050	2.177	1, 5, 7	0.40	7, 12	0.73	1, 5, 7, 12	0.57

CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	0.010	0.400	0.150	0.000	1.168	1, 5, 7	0.18	6, 7,	0.10	1, 5, 6, 7	0.14
CHF2OCF2CHFCl (Enflurane)	-0.230	0.400	0.120	0.130	1.750	1, 5, 7, 26	0.25	26,	0.45	1, 5, 7, 26	0.35
CHF2OCHClCF3 (Isoflurane)	-0.240	0.500	0.100	0.100	1.576	1, 5, 7	0.15	6, 7,	0.25	1, 5, 6, 7	0.20
Chlorobenzene	0.718	0.650	0.000	0.070	3.657	1, 6, 7	1.48	6, 7	1.77	1, 6, 7	1.63
Chlorodibromomethane	0.775	0.710	0.070	0.080	3.304	6, 7, 32	1.71	6, 7, 31	2.04	6, 7, 31, 32	1.88
Chloroethane	0.227	0.400	0.000	0.100	1.678	1, 6, 7	0.36	6, 7	0.61	1, 6, 7	0.49
Chloroform (Trichloromethane)	0.425	0.490	0.150	0.020	2.480	1, 4, 5, 6, 7, 32	0.98	7, 31	1.32	1, 4, 5, 6, 7, 31, 32	1.15
cis-1,2-Dichloroethylene	0.436	0.610	0.110	0.050	2.4390	1, 6, 7,	0.98	6, 7,	1.33	1, 6, 7	1.16
Cyclohexane	0.305	0.100	0.000	0.000	2.964	1, 4, 5, 6	0.19	6, 7	0.14	1, 4, 5, 6, 7	0.17
Cyclopropane	0.408	0.230	0.000	0.000	1.314	1, 5, 7, 26	-0.29	26,	-0.12	1, 5, 7, 26	-0.21
Decane	0.000	0.000	0.000	0.000	4.686	4,	1.92	7, 17	1.02	4, 7, 17	1.47
Dichloromethane (Methylene chloride)	0.387	0.570	0.100	0.050	2.019	1, 4, 5, 6, 7	0.95	6, 7	1.29	1, 4, 5, 6, 7	1.12
Diethyl ether	0.041	0.250	0.000	0.450	2.015	1, 5, 7, 26	1.09	7, 26	1.12	1, 5, 7, 26	1.11
Ethane	0.000	0.000	0.000	0.000	0.492	1, 26	-1.07	7, 26	-0.97	1, 7, 26	-1.02
Ethanol	0.246	0.420	0.370	0.480	1.485	1, 4, 5, 7	3.17	7,	3.37	1, 4, 5, 7	3.27
Ethene	0.107	0.100	0.000	0.070	0.289	1, 7, 29	-0.75	29,	-0.31	1, 7, 29	-0.53
Ethyl acetate	0.106	0.620	0.000	0.450	2.314	4, 7	1.91	7, 22	1.89	4, 7, 22	1.90
Ethylbenzene	0.613	0.510	0.000	0.150	3.778	1, 4, 7	1.45	7,	1.48	1, 4, 7	1.47

Ethylene Oxide	0.250	0.740	0.070	0.320	1.371	29,	1.79	23,	1.81	23, 29	1.80
Heptane	0.000	0.000	0.000	0.000	3.173	1, 4, 5, 6, 7	0.42	6, 7, 17	0.58	1, 4, 5, 6, 7, 17	0.50
Hexachloroethane (Perchloroethane)	0.680	0.680	0.000	0.000	4.718	6, 7	1.72	6, 7,	1.80	6, 7	1.76
Hexane	0.000	0.000	0.000	0.000	2.668	1, 4, 5, 7	0.07	6, 7, 17	0.35	1, 4, 5, 6, 7, 17	0.21
Isobutyl acetate	0.052	0.570	0.000	0.470	3.161	7,	1.65	7,	1.72	7,	1.69
Isopentylacetate (Aceticacidpentylester)	0.051	0.570	0.000	0.470	3.740	7,	1.77	7,	1.81	7,	1.79
Isopropyl acetate	0.055	0.570	0.000	0.470	2.546	4, 7	1.54	7,	1.55	4, 7	1.55
Isopropyl bromide (2-Brompropane)	0.332	0.350	0.000	0.140	2.390	6, 7,	0.41	6, 7, 8	0.86	6, 7, 8	0.64
JP-10	0.590	0.450	0.000	0.060	4.840	6,	1.72	6,	1.79	6,	1.76
MeOCF2CHCl2 (Methoxyflurane)	0.109	0.670	0.070	0.140	2.864	1, 5, 7	1.16	7,	1.40	1, 5, 7	1.28
Methanol	0.278	0.440	0.430	0.470	0.970	1, 4, 5, 7	3.29	7,	3.52	1, 4, 5, 7	3.41
Methyl acetate	0.142	0.640	0.000	0.450	1.911	7,	1.95	7,	2.00	7,	1.98
Methyl tertiary-butyl ether (MTBE)	0.024	0.210	0.000	0.590	2.380	7, 19	1.25	11, 19	1.11	7, 11, 19	1.18
Methylchloride (Chloromethane)	0.249	0.430	0.000	0.080	1.163	1, 6, 7	0.23	6, 7	0.39	1, 6, 7	0.31
Methylcyclohexane	0.244	0.060	0.000	0.000	3.319	4,	0.61	17,	0.79	4, 17	0.70
m-Xylene (1,3-Dimethylbenzene)	0.623	0.520	0.000	0.160	3.839	1, 4, 6, 7	1.51	6, 7	1.66	1, 4, 6, 7	1.59
Nonane	0.000	0.000	0.000	0.000	4.182	4,	1.70	16, 17	0.63	4, 16, 17	1.17
Octane	0.000	0.000	0.000	0.000	3.677	4,	0.61	7, 17	0.74	4, 7, 17	0.68
o-Xylene	0.663	0.560	0.000	0.160	3.939	1, 7	1.53	7, 17	1.30	1, 7, 17	1.42
Pentyl acetate	0.067	0.600	0.000	0.450	3.844	7,	1.97	7,	1.99	7,	1.98

Propene (Propylene)	0.103	0.080	0.000	0.070	0.946	3,	-0.36	3,	-0.06	3,	-0.21
Propyl acetate	0.092	0.600	0.000	0.450	2.819	7,	1.87	7,	1.88	7,	1.88
Propyl bromide (1-Bromopropane)	0.366	0.400	0.000	0.120	2.620	6, 7	0.85	6, 7, 8	1.09	6, 7, 8	0.97
p-Xylene	0.613	0.520	0.000	0.160	3.839	1, 4, 6, 7	1.60	6, 7	1.62	1, 4, 6, 7	1.61
Styrene (Vinyl benzene)	0.849	0.650	0.000	0.160	3.856	1, 4, 7	1.73	6, 7	1.60	1, 4, 6, 7	1.67
Sulphur Hexafluoride	-0.600	-0.200	0.000	0.000	-0.120	1, 26	-2.22	26,	-2.12	1, 26	-2.17
t-butanol (tert-butanol)(TBA) (2-Methyl-2-propanol)	0.180	0.300	0.310	0.600	1.963	7, 19	2.66	11, 19	2.70	7, 11, 19	2.68
tertiary-Amyl methyl ether (TAME)	0.050	0.210	0.000	0.600	2.916	19	1.25	19,	1.19	19,	1.22
Tetrachloroethene (Perchloroethylene)	0.639	0.440	0.000	0.000	3.584	1, 4, 6, 7	1.09	6, 7	1.28	1, 4, 6, 7	1.19
Toluene	0.601	0.520	0.000	0.140	3.325	1, 4, 5	1.12	6, 7, 17	1.16	1, 4, 5, 6, 7, 17	1.14
trans-1,2-Dichloroethylene	0.425	0.410	0.090	0.050	2.278	1, 6, 7	0.77	6, 7	0.98	1, 6, 7	0.88
Tribromomethane (Bromoform)	0.974	0.680	0.150	0.060	3.784	4, 32	2.02	31,	2.27	4, 31, 32	2.15
Trichloroethene (Trichloroethylene)	0.524	0.370	0.080	0.030	2.997	1, 4, 5, 6, 7	0.94	6, 7, 27, 28	1.33	1, 4, 5, 6, 7, 27, 28	1.14
Vinyl bromide (Bromoethene)	0.564	0.500	0.000	0.070	1.846	6, 7	0.36	6, 7	0.61	6, 7	0.49
Vinyl chloride (Chloroethene)	0.258	0.380	0.000	0.050	1.404	6, 7, 30	0.06	6, 7, 27, 30	0.27	6, 7, 27, 30	0.17
Propylbenzene	0.604	0.500	0.000	0.150	4.230	1, 7	1.67			1, 7	1.67
(CF3)2CHOCH2F (Sevoflurane)	-0.465	0.232	0.080	0.147	1.688	1, 5, 7	-0.20			1, 5, 7	-0.20
1,2,3-Trichloropropane	0.547	0.650	0.020	0.330	3.566	4,	2.01			4,	2.01

1,2,3-Trimethylbenzene (Hemimellitene)	0.728	0.610	0.000	0.190	4.565	7,	1.82			7,	1.82
1,2-Dichlorobenzene (o-Dichlorobenzene)	0.872	0.780	0.000	0.040	4.518	1, 7	2.63			1, 7	2.63
1,3,5-Trimethylbenzene (Mesitylene)	0.649	0.520	0.000	0.190	4.344	4, 7	1.64			4, 7	1.64
1,3-Butadiene	0.320	0.230	0.000	0.100	1.543	21,	0.09			21,	0.09
1,3-Dichlorobenzene (m-Dichlorobenzene)	0.847	0.730	0.000	0.020	4.410	1, 7	2.30			1, 7	2.30
1-Chlorobutane	0.210	0.400	0.000	0.100	2.722	1, 7	0.63			1, 7	0.63
1-Chloropentane	0.208	0.380	0.000	0.090	3.223	1, 7	0.87			1, 7	0.87
1-Fluropropane	0.034	0.350	0.000	0.130	1.103	1,	0.02			1,	0.02
1-Methoxy-2-propanol	0.218	0.610	0.350	0.620	2.655	7, 20	4.09			7, 20	4.09
2,2-Dimethylbutane	0.000	0.000	0.000	0.000	2.352	1, 5, 7	-0.59			1, 5, 7	-0.59
2,3-Dimethylbutane	0.000	0.000	0.000	0.000	2.495	4	0.78			4	0.78
2-Butoxyethanol	0.201	0.500	0.300	0.830	3.806	7, 20	3.90			7, 20	3.90
2-Ethoxyethanol	0.237	0.520	0.310	0.810	2.792	7, 20	4.34			7, 20	4.34
2-Fluropropane	0.004	0.320	0.000	0.100	1.070	1,	0.06			1,	0.06
2-Hexanone	0.136	0.680	0.000	0.510	3.286	1,	2.10			1,	2.10
2-Isopropoxyethanol	0.196	0.470	0.300	0.910	3.170	7, 20	4.16			7, 20	4.16
2-Methoxyethanol	0.269	0.500	0.300	0.840	2.490	7, 20	4.52			7, 20	4.52
2-Methylcyclohexanone (o-Methylcyclohexanone)	0.372	0.830	0.000	0.560	4.055	4,	2.87			4,	2.87

2-Methylpentane (Isohexane)	0.000	0.000	0.000	0.000	2.503	1, 5, 7	-0.39			1, 5, 7	-0.39
3-Methylhexane	0.000	0.000	0.000	0.000	3.044	1, 5, 7	0.11			1, 5, 7	0.11
3-Methylpentane	0.000	0.000	0.000	0.000	2.581	1, 5, 7	-0.37			1, 5, 7	-0.37
3-Pentanone (Diethyl ketone)	0.154	0.660	0.000	0.510	2.811	1, 7	2.21			1, 7	2.21
Acetylene	0.190	0.600	0.060	0.040	0.140	1,	-0.06			1,	-0.06
Allylbenzene	0.717	0.600	0.000	0.220	4.136	1, 7	1.71			1, 7	1.71
Argon	0.000	0.000	0.000	0.000	-0.688	1,	-1.52			1,	-1.52
Carbon Disulfide	0.876	0.260	0.000	0.030	2.370	1,	0.30			1,	0.30
Carbon Monoxide	0.000	0.000	0.000	0.040	-0.836	1,	-1.67			1,	-1.67
CF2=CHCl (1-Chloro-2,2-difluroethylene) (1,1-Difluro-2-chloroethene)	-0.340	0.290	0.150	0.000	0.723	1, 5	0.06			1, 5	0.06
CF2ClCF2Cl (1,2-Dichlorotetrafluoroethane)	-0.190	0.050	0.000	0.000	1.427	1,	-0.82			1,	-0.82
CF2HCF2CF2CF2H	-0.710	0.040	0.090	0.000	0.590	7,	-0.36			7,	-0.36
CF2HCF2CFH2	-0.450	0.170	0.000	0.030	0.680	7,	-0.48			7,	-0.48
CF2HCF2H (1,1,2,2-Tetrafluoroethane)	-0.280	-0.300	0.300	0.000	0.289	7,	-0.12			7,	-0.12
CF2HCFHCH2CF2H	-0.500	1.250	0.120	0.130	2.324	7,	0.87			7,	0.87
CF2HCH3 (1,1-Difluroethane)	-0.250	0.490	0.040	0.050	0.517	7,	0.42			7,	0.42
CF3CBrFH (Teflurane)	-0.070	0.210	0.200	0.020	1.370	1, 7	-0.22			1, 7	-0.22
CF3CF2CF2CF2H	-0.780	-0.300	0.100	0.100	0.420	7,	-1.52			7,	-1.52

CF3CFH2 (Norflurane) (1,1,1,2-Tetrafluoroethane)	-0.640	0.200	0.240	0.000	0.226	7,	-0.25			7,	-0.25
CF3CFHCFHCF3	-0.710	-0.090	0.090	0.040	0.590	7,	-0.59			7,	-0.59
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	0.183	0.300	0.000	0.270	1.600	1, 5, 7	0.15			1, 5, 7	0.15
CF4 (Perfluoromethane) (Carbon Tetrafluoride)	-0.550	-0.200	0.000	0.000	-0.819	7,	-1.10			7,	-1.10
CHF2CF2CH2Br (Halopropane)	-0.070	0.280	0.200	0.000	2.030	1,	0.75			1,	0.75
CHF2OCHFCF3 (Desflurane)	-0.540	0.270	0.070	0.170	0.740	1,	-0.37			1,	-0.37
Cyclohexanone	0.403	0.860	0.000	0.560	3.792	4,	3.33			4,	3.33
Diflurodichloromethane	0.037	0.130	0.000	0.000	1.124	1,	-0.82			1,	-0.82
Dimethyl ether	0.000	0.270	0.000	0.410	1.285	1,	1.16			1,	1.16
Divinyl ether	0.259	0.390	0.000	0.130	1.760	1, 5, 7	0.41			1, 5, 7	0.41
Ethyl formate	0.146	0.660	0.000	0.380	1.845	1,	1.65			1,	1.65
Ethyl t-butyl ether (ETBE)	-0.020	0.160	0.000	0.600	2.720	7, 19	1.07			7, 19	1.07
Ethyl t-pentyl ether	0.030	0.230	0.000	0.370	3.200	7,	1.25			7,	1.25
Fluoroethane	0.052	0.350	0.000	0.100	0.576	1,	0.09			1,	0.09
Fluorotrichloromethane (Trichlorofluromethane)	0.207	0.240	0.000	0.070	1.950	1,	-0.06			1,	-0.06
Helium	0.000	0.000	0.000	0.000	-1.741	1,	-2.00			1,	-2.00
Hydrogen	0.000	0.000	0.000	0.000	-1.200	1,	-1.77			1,	-1.77

Iodoethane	0.640	0.400	0.000	0.150	2.573	1,	0.83			1,	0.83
Isophorone (3,5,5-trimethylcyclohex-2-en-1-one)	0.511	1.120	0.000	0.530	4.740	4,	3.37			4,	3.37
Isopropylbenzene (Cumene)	0.602	0.490	0.000	0.160	4.084	1, 7	1.57			1, 7	1.57
Krypton	0.000	0.000	0.000	0.000	-0.211	1,	-1.22			1,	-1.22
Methane	0.000	0.000	0.000	0.000	-0.323	1,	-1.42			1,	-1.42
Methyl butyl ketone (3-Methylpentan-2-one)	0.110	0.650	0.000	0.510	3.163	7,	2.23			7,	2.23
Methylcyclopentane	0.225	0.100	0.000	0.000	2.907	1, 5, 7	-0.07			1, 5, 7	-0.07
Neon	0.000	0.000	0.000	0.000	-1.575	1,	-2.01			1,	-2.01
Nitrogen	0.000	0.000	0.000	0.000	-0.978	1,	-1.83			1,	-1.83
Nitrous Oxide	0.068	0.350	0.000	0.100	0.164	1, 5	-0.34			1, 5	-0.34
Oxygen	0.000	0.000	0.000	0.000	-0.723	1,	-1.58			1,	-1.58
Pentane	0.000	0.000	0.000	0.000	2.162	1, 4, 5, 7	-0.29			1, 4, 5, 7	-0.29
Xenon	0.000	0.000	0.000	0.000	0.378	1, 5	-0.85			1, 5	-0.85
1,1-Dichloro-1-Fluoroethane (HCFC-141b)	0.084	0.430	0.010	0.050	1.920			13,	0.32	13,	0.32
1,1-Dichloroethylene	0.362	0.340	0.000	0.050	2.110			6, 7	0.70	6, 7	0.70
1,2,4-Trifluorobenzene	0.410	0.650	0.000	0.020	2.850			7,	0.76	7,	0.76
1,2,4-Trimethylcyclohexane	0.360	0.210	0.000	0.000	4.100			16, 17	0.87	16, 17	0.87
1,2-Dibromoethane	0.747	0.760	0.100	0.170	3.382			6, 7,	2.08	6, 7,	2.08

1,2-Diflurobenzene (o-Diflurobenzene)	0.390	0.630	0.000	0.060	2.843			7,	0.96	7,	0.96
1,2-Dimethylcyclohexane	0.320	0.230	0.000	0.000	3.800			17,	0.91	17,	0.91
1,2-Epoxy-3-butene (BMO)	0.370	0.470	0.000	0.360	2.257			24,	1.97	24,	1.97
1,3,5-Triflurobenzene	0.390	0.490	0.000	0.000	2.660			7,	0.49	7,	0.49
1,4-Diflurobenzene (p-Diflurobenzene)	0.384	0.600	0.000	0.060	2.766			7,	0.87	7,	0.87
1-Decene	0.093	0.080	0.000	0.070	4.533			18,	1.21	18,	1.21
1-Hexanol	0.210	0.420	0.370	0.480	3.610			7,	3.21	7,	3.21
1-Nonene	0.900	0.080	0.000	0.070	4.073			18,	1.18	18,	1.18
1-Octene	0.094	0.080	0.000	0.070	3.568			18,	1.07	18,	1.07
2,2-Dichloro-1,1,1-Trifluroethane (HCFC-123)	-0.160	0.400	0.220	0.000	1.746			12,	0.61	12,	0.61
2,3,4-Trimethylpentane	0.000	0.000	0.000	0.000	3.481			6, 7	0.57	6, 7	0.57
2-Methyl heptane	0.000	0.000	0.000	0.000	3.480			18,	0.49	18,	0.49
2-Methyl nonane	0.000	0.000	0.000	0.000	4.453			18,	0.76	18,	0.76
2-Methyl octane	0.000	0.000	0.000	0.000	3.966			18,	0.52	18,	0.52
Allyl chloride (3-Chloro-1-propylene)	0.327	0.560	0.000	0.050	2.109			6, 7	1.24	6, 7	1.24
Bromobenzene	0.882	0.730	0.000	0.090	4.041			22,	1.72	22,	1.72
Butane	0.000	0.000	0.000	0.000	1.615			7,	-0.53	7,	-0.53
Cyanoethylene Oxide	0.390	1.000	0.000	0.520	2.543			25,	3.22	25,	3.22
Cycloheptane	0.350	0.100	0.000	0.000	3.704			7,	0.72	7,	0.72
Cyclopentane	0.263	0.100	0.000	0.000	2.477			7,	0.24	7,	0.24

Dibromomethane	0.714	0.690	0.110	0.070	2.886			6, 7	1.87	6, 7	1.87
Difluromethane (CF2H2)	-0.320	0.490	0.060	0.050	0.040			6, 7	0.20	6, 7	0.20
Fluorobenzene	0.477	0.570	0.000	0.100	2.788			7,	1.06	7,	1.06
Flurochloromethane	-0.080	0.270	0.090	0.030	1.030			6, 7	0.71	6, 7	0.71
Furan	0.369	0.510	0.000	0.130	1.913			9,	0.82	9,	0.82
Hexaflurobenzene	0.088	0.560	0.000	0.010	2.345			7,	0.39	7,	0.39
Methylpentaflurobenzene (2,3,4,5,6-Pentylfluorotoluene)	0.240	0.450	0.040	0.000	0.946			7,	0.73	7,	0.73
m-Methylstyrene (1-vinyl-3-methylbenzene)	0.866	0.650	0.000	0.180	4.375			6, 7	2.28	6, 7	2.28
Pentachloroethane	0.648	0.660	0.170	0.060	4.267			6, 7	2.02	6, 7	2.02
Pentaflurobenzene	0.154	0.680	0.000	0.020	2.578			7,	0.51	7,	0.51
p-Methylstyrene (1-vinyl-4-methylbenzene)	0.871	0.650	0.000	0.180	4.399			6, 7	2.37	6, 7	2.37
Radon	0.000	0.000	0.000	0.000	0.877			15,	-0.39	15,	-0.39
t-Butylbenzene	0.619	0.490	0.000	0.180	4.413			17,	1.24	17,	1.24
t-Butylcyclohexane	0.300	0.100	0.000	0.100	4.603			17,	1.16	17,	1.16
tertiary-amyl alcohol (TAA)	0.194	0.300	0.310	0.600	2.630			19,	2.59	19,	2.59

Vinylidene fluoride (1,1-Difluoroethylene)	-0.100	0.000	0.000	0.050	0.240			14,	-0.74	14,	-0.74

Table 9.1. Air to tissue (brain, fat, kidney, liver, lung, muscle and plasma) distribution for VOCs

TISSUE	SOLUTE NAME	E	S	A	B	L	TISSUE REF	HUMAN LOG K	TISSUE REF	RAT LOG K
LUNG	(CF3)2CHOCH2F (Sevoflurane)	-0.465	0.232	0.080	0.147	1.688	1, 5	0.11	1, 5,	0.11
LUNG	Propene (Propylene)	0.103	0.080	0.000	0.070	0.946	3,	-0.21	3,	-0.27
PLASMA	Acetone (Propan-2-one)	0.179	0.700	0.040	0.490	1.696	1,	2.54	26,	2.20
PLASMA	Cyclopropane	0.408	0.230	0.000	0.000	1.314	1,	-0.62	26,	-0.63
PLASMA	Diethyl ether	0.041	0.250	0.000	0.450	2.015	1,	1.08	26,	1.16
PLASMA	Ethane	0.000	0.000	0.000	0.000	0.492	1,	-1.34	26,	-1.44
PLASMA	Sulphur Hexafluoride	-0.600	-0.200	0.000	0.000	-0.120	1,	-2.28	26,	-2.41
FAT	1,1,1-Trichloroethane (Methyl chloroform)	0.369	0.410	0.000	0.090	2.733	1, 5	2.40	6, 7	2.42
FAT	1-Propanol	0.236	0.420	0.370	0.480	2.031	1, 5, 7	2.47	7,	2.60
FAT	2-Methyl-1-propanol (Isobutanol)	0.217	0.390	0.370	0.480	2.413	1, 5, 7	2.59	7,	2.86
FAT	2-Propanol (Isopropanol)	0.212	0.360	0.330	0.560	1.764	1, 5, 7	2.26	7,	2.44
FAT	Acetone (Propan-2-one)	0.179	0.700	0.040	0.490	1.696	1, 5, 7	1.93	7,	1.89

FAT	Benzene	0.610	0.520	0.000	0.140	2.786	1, 5, 7	2.62	6, 7, 17	2.65
FAT	Butan-2-one (Methyl ethyl ketone)	0.166	0.700	0.000	0.510	2.287	1, 5, 7	2.21	7,	2.30
FAT	CF3CBrClH (Halothane)	0.102	0.380	0.150	0.050	2.177	1, 5, 7	2.31	6, 7, 12	2.23
FAT	CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	0.010	0.400	0.150	0.000	1.168	1, 5, 7	1.53	6, 7	1.33
FAT	CHF2OCHClCF3 (Isoflurane)	-0.240	0.500	0.100	0.100	1.576	1, 5, 7	1.84	6, 7	1.99
FAT	Chloroform (Trichloromethane)	0.425	0.490	0.150	0.020	2.480	1, 5, 7	2.45	6, 7	2.31
FAT	Cyclohexane	0.305	0.100	0.000	0.000	2.964	1, 5, 7	2.41	6, 7, 17	2.47
FAT	Dichloromethane (Methylene chloride)	0.387	0.570	0.100	0.050	2.019	1, 5, 7	1.93	6, 7	2.08
FAT	Diethyl ether	0.041	0.250	0.000	0.450	2.015	1, 5, 7	1.80	6, 7	1.68
FAT	Ethanol	0.246	0.420	0.370	0.480	1.485	1, 5, 7	2.33	7	2.35
FAT	Ethylbenzene	0.613	0.510	0.000	0.150	3.778	7,	3.25	7	3.19
FAT	Ethylene Oxide	0.250	0.740	0.070	0.320	1.371	29	1.63	23	1.64
FAT	Ethylene (Ethene)	0.107	0.100	0.000	0.070	0.289	1, 7, 29	0.08	29	0.31
FAT	Heptane	0.000	0.000	0.000	0.000	3.173	1, 5, 7	2.59	6, 7, 17	2.42
FAT	Hexane	0.000	0.000	0.000	0.000	2.668	1, 5, 7	2.02	6, 7, 17	2.22
FAT	Methanol	0.278	0.440	0.430	0.470	0.970	5, 7	2.36	7,	2.29
FAT	m-Xylene (1,3-Dimethylbenzene)	0.623	0.520	0.000	0.160	3.839	7,	3.28	6, 7	3.30
FAT	Octane	0.000	0.000	0.000	0.000	3.677	1, 7	2.71	17,	2.73
FAT	o-Xylene (1,2-dimethylbenzene)	0.663	0.560	0.000	0.160	3.939	7,	3.39	6, 7, 17	3.00
FAT	Propene (Propylene)	0.103	0.080	0.000	0.070	0.946	3,	0.71	3,	0.80
FAT	p-Xylene (1,4-dimethylbenzene)	0.613	0.520	0.000	0.160	3.839	7,	3.31	6, 7	3.24

FAT	Styrene (Vinyl benzene)	0.849	0.650	0.000	0.160	3.856	7,	3.50	6, 7	3.54
FAT	Toluene (methylbenzene)	0.601	0.520	0.000	0.140	3.325	1, 5, 7	2.95	6, 7, 17	2.89
FAT	Trichloroethene (Trichloroethylene)	0.524	0.370	0.080	0.030	2.997	1, 5, 7	2.77	6, 7, 28	2.72
FAT	Vinyl chloride (Chloroethene)	0.258	0.380	0.000	0.050	1.404	30,	1.30	6, 7, 30	1.30
LIVER	1,1,1-Trichloroethane (Methyl chloroform)	0.369	0.410	0.000	0.090	2.733	1, 5, 7	1.21	6, 7	0.93
LIVER	Benzene	0.610	0.520	0.000	0.140	2.786	1, 5, 7	1.36	6, 7, 17	1.15
LIVER	CF3CBrClH (Halothane)	0.102	0.380	0.150	0.050	2.177	1, 5, 7	0.86	6, 7, 12	0.86
LIVER	CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	0.010	0.400	0.150	0.000	1.168	1, 5, 7	0.36	6, 7	0.26
LIVER	CHF2OCHClCF3 (Isoflurane)	-0.240	0.500	0.100	0.100	1.576	1, 5, 7	0.55	6, 7	0.57
LIVER	Chloroform (Trichloromethane)	0.425	0.490	0.150	0.020	2.480	1, 5, 7	1.23	6, 7	1.32
LIVER	Cyclohexane	0.305	0.100	0.000	0.000	2.964	1, 5, 7	1.03	6, 7, 17	1.06
LIVER	Dichloromethane (Methylene chloride)	0.387	0.570	0.100	0.050	2.019	1, 5, 7	0.86	6, 7	1.15
LIVER	Diethyl ether	0.041	0.250	0.000	0.450	2.015	1, 5, 7	1.05	6, 7	0.83
LIVER	Ethylene (Ethene)	0.107	0.100	0.000	0.070	0.289	29,	-0.35	29,	-0.24
LIVER	Heptane	0.000	0.000	0.000	0.000	3.173	1, 5, 7	1.04	6, 7, 17	0.87
LIVER	Hexane	0.000	0.000	0.000	0.000	2.668	1, 5, 7	0.72	6, 7, 17	0.65
LIVER	MeOCF2CHCl2 (Methoxyflurane)	0.109	0.670	0.070	0.140	2.864	1, 5, 7	1.46	7,	1.47
LIVER	Octane	0.000	0.000	0.000	0.000	3.677	1, 7	1.41	17,	0.99
LIVER	Propene (Propylene)	0.103	0.080	0.000	0.070	0.946	3,	-0.27	3,	-0.29
LIVER	Toluene (Methylbenzene)	0.601	0.520	0.000	0.140	3.325	1, 5, 7	1.68	6, 7, 17	1.60
LIVER	Trichloroethene (Trichloroethylene)	0.524	0.370	0.080	0.030	2.997	1, 5, 7	1.41	6, 7, 28	1.40
LIVER	Vinyl chloride (Chloroethene)	0.258	0.380	0.000	0.050	1.404	30,	0.20	6, 7, 30	0.20

MUSCLE	1,1,1-Trichloroethane (Methyl chloroform)	0.369	0.410	0.000	0.090	2.733	1, 5, 7	0.83	6, 7	0.50
MUSCLE	1-Propanol	0.236	0.420	0.370	0.480	2.031	1, 5, 7	2.83	7,	3.06
MUSCLE	2-Methyl-1-propanol (Isobutanol)	0.217	0.390	0.370	0.480	2.413	1, 5, 7	2.54	7,	2.93
MUSCLE	2-Propanol (Isopropanol)	0.212	0.360	0.330	0.560	1.764	1, 5, 7	2.70	7,	3.04
MUSCLE	Acetone (Propan-2-one)	0.179	0.700	0.040	0.490	1.696	1, 5, 7	2.18	7,	2.23
MUSCLE	Benzene	0.610	0.520	0.000	0.140	2.786	1, 5, 7	1.21	6, 7	1.01
MUSCLE	Butan-2-one (Methyl ethyl ketone)	0.166	0.700	0.000	0.510	2.287	1, 5, 7	2.01	7,	2.26
MUSCLE	CF3CBrClH (Halothane)	0.102	0.380	0.150	0.050	2.177	1, 5, 7	0.83	6, 7	0.65
MUSCLE	CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	0.010	0.400	0.150	0.000	1.168	1, 5, 7	0.34	6, 7	0.09
MUSCLE	CHF2OCHClCF3 (Isoflurane)	-0.240	0.500	0.100	0.100	1.576	1, 5, 7	0.41	6, 7	0.20
MUSCLE	Chloroform (Trichloromethane)	0.425	0.490	0.150	0.020	2.480	1, 5, 7	1.09	6, 7	1.14
MUSCLE	Cyclohexane	0.305	0.100	0.000	0.000	2.964	1, 5, 7	1.02	6, 7	0.01
MUSCLE	Dichloromethane (Methylene chloride)	0.387	0.570	0.100	0.050	2.019	1, 5, 7	0.68	6, 7	0.90
MUSCLE	Diethyl ether	0.041	0.250	0.000	0.450	2.015	1, 5, 7	1.01	6, 7	0.72
MUSCLE	Ethanol	0.246	0.420	0.370	0.480	1.485	1, 5, 7	2.93	7,	3.23
MUSCLE	Ethylene Oxide	0.250	0.740	0.070	0.320	1.371	29,	1.82	23,	1.68
MUSCLE	Ethylene (Ethene)	0.107	0.100	0.000	0.070	0.289	29,	-0.19	29,	-0.20
MUSCLE	Heptane	0.000	0.000	0.000	0.000	3.173	1, 5, 7	1.08	6, 7	0.62
MUSCLE	Hexane	0.000	0.000	0.000	0.000	2.668	1, 5, 7	0.70	6, 7	0.46
MUSCLE	Methanol	0.278	0.440	0.430	0.470	0.970	1, 5, 7	3.12	7,	3.60
MUSCLE	Propene (Propylene)	0.103	0.080	0.000	0.070	0.946	3,	-0.19	3,	-0.34

MUSCLE	Toluene (methylbenzene)	0.601	0.520	0.000	0.140	3.325	1, 5, 7	1.54	6, 7	1.44
MUSCLE	Trichloroethene (Trichloroethylene)	0.524	0.370	0.080	0.030	2.997	1, 5, 7	1.25	6, 7	1.00
KIDNEY	1-Propanol	0.236	0.420	0.370	0.480	2.031	1, 5, 7	2.84	7,	3.09
KIDNEY	2-Methyl-1-propanol (Isobutanol)	0.217	0.390	0.370	0.480	2.413	1, 5, 7	2.57	7,	2.94
KIDNEY	2-Propanol (Isopropanol)	0.212	0.360	0.330	0.560	1.764	1, 5, 7	2.70	7,	3.03
KIDNEY	Benzene	0.610	0.520	0.000	0.140	2.786	1, 5, 7	1.08	17	1.51
KIDNEY	Ethanol	0.246	0.420	0.370	0.480	1.485	1, 5, 7	2.97	7,	3.31
KIDNEY	Ethylene (Ethene)	0.107	0.100	0.000	0.070	0.289	29,	-0.44	29,	-0.55
KIDNEY	Heptane	0.000	0.000	0.000	0.000	3.173	1, 5, 7	0.95	17,	1.20
KIDNEY	Hexane	0.000	0.000	0.000	0.000	2.668	1, 5, 7	0.48	17	1.41
KIDNEY	Methanol	0.278	0.440	0.430	0.470	0.970	1, 5, 7	3.13	7,	3.50
KIDNEY	Octane	0.000	0.000	0.000	0.000	3.677	1, 7	0.91	17,	1.76
KIDNEY	Propene (Propylene)	0.103	0.080	0.000	0.070	0.946	3,	-0.35	3,	-0.28
KIDNEY	Toluene (methylbenzene)	0.601	0.520	0.000	0.140	3.325	1, 5, 7	1.26	17,	1.79
BRAIN	1-Propanol	0.236	0.420	0.370	0.480	2.031	1, 5, 7	2.87	7,	3.09
BRAIN	2-Methyl-1-propanol (Isobutanol)	0.217	0.390	0.370	0.480	2.413	1, 5, 7	2.61	7,	2.94
BRAIN	2-Propanol (Isopropanol)	0.212	0.360	0.330	0.560	1.764	1, 5, 7	2.76	7,	3.05
BRAIN	Benzene	0.610	0.520	0.000	0.140	2.786	1, 5, 7	1.26	17,	1.24
BRAIN	CF3CBrClH (Halothane)	0.102	0.380	0.150	0.050	2.177	1, 5, 7	0.77	7,	0.65
BRAIN	CHF2OCHClCF3 (Isoflurane)	-0.240	0.500	0.100	0.100	1.576	1, 5, 7	0.46	7,	0.31
BRAIN	Cyclohexane	0.305	0.100	0.000	0.000	2.964	1, 5, 7	1.04	17	1.52
BRAIN	Ethanol	0.246	0.420	0.370	0.480	1.485	1, 5, 7	3.02	7,	3.27
BRAIN	Ethylene Oxide	0.250	0.740	0.070	0.320	1.371	29,	1.84	23,	1.77

BRAIN	Ethylene (Ethene)	0.107	0.100	0.000	0.070	0.289	29,	-0.26	29,	-0.19
BRAIN	Heptane	0.000	0.000	0.000	0.000	3.173	1, 5, 7	1.09	17,	0.79
BRAIN	Hexane	0.000	0.000	0.000	0.000	2.668	1, 5, 7	0.70	17,	1.07
BRAIN	MeOCF2CHCl2 (Methoxyflurane)	0.109	0.670	0.070	0.140	2.864	1, 5, 7	1.41	7,	1.39
BRAIN	Methanol	0.278	0.440	0.430	0.470	0.970	5, 7	3.18	7,	3.54
BRAIN	Octane	0.000	0.000	0.000	0.000	3.677	1, 7	1.22	17,	1.52
BRAIN	Propene (Propylene)	0.103	0.080	0.000	0.070	0.946	3,	-0.22	3,	-0.32
BRAIN	Toluene (Methylbenzene)	0.601	0.520	0.000	0.140	3.325	1, 5, 7	1.56	17	1.26
BRAIN	Trichloroethene (Trichloroethylene)	0.524	0.370	0.080	0.030	2.997	1, 5, 7	1.33	28,	1.16

Table 10.0. Air to blood distribution for VOCs

SOLUTE NAME	AVERAGE LOG K	CALC LOG K
(CF3)2CHOCH2F (Sevoflurane)	-0.20	0.28
1,1,1,2-Tetrachloroethane	1.55	1.81
1,1,1-Trichloroethane (Methyl chloroform)	0.63	0.80
1,1,2,2-Tetrachloroethane	2.13	2.36
1,1,2-Trichloroethane	1.67	1.95
1,1-Dichloro-1-Fluoroethane (HCFC-141b)	0.32	0.32
1,1-Dichloroethane	0.88	1.11
1,1-Dichloroethylene	0.70	0.39
1,2,3-Trichloropropane	2.01	2.15
1,2,3-Trimethylbenzene (Hemimellitene)	1.82	2.13
1,2,4-Trifluorobenzene	0.76	0.94
1,2,4-Trimethylbenzene (Pseudocumene)	1.47	2.01
1,2,4-Trimethylcyclohexane	0.87	0.86
1,2-Dibromoethane	2.08	2.18
1,2-Dichlorobenzene (o-Dichlorobenzene)	2.63	1.97
1,2-Dichloroethane	1.39	1.44
1,2-Dichloropropane	1.14	1.30
1,2-Diflurobenzene (o-Diflurobenzene)	0.96	1.01
1,2-Dimethylcyclohexane	0.91	0.75
1,2-Epoxy-3-butene (BMO)	1.97	1.39
1,3,5-Triflurobenzene	0.49	0.64
1,3,5-Trimethylbenzene (Mesitylene)	1.64	1.91
1,3-Butadiene	0.09	0.16
1,3-Dichlorobenzene (m-Dichlorobenzene)	2.30	1.82
1,4-Diflurobenzene (p-Diflurobenzene)	0.87	0.95
1-Bromo-2-chloroethane (1-Chloro-2-Bromoethane)	1.60	1.68
1-Butanol	3.08	3.09
1-Chlorobutane	0.63	0.74
1-Chloropentane	0.87	0.88
1-Chloropropane (Propyl chloride)	0.59	0.55
1-Decene	1.21	0.94
1-Fluropropane	0.02	0.08
1-Hexanol	3.21	3.46
1-Methoxy-2-propanol	4.09	3.60
1-Nitropropane	2.31	1.96

1-Nonene	1.18	1.14
1-Octene	1.07	0.58
1-Pentanol	2.83	3.27
1-Propanol	3.06	2.88
2,2,4-Trimethylpentane (Isooctane)	0.23	0.10
2,2-Dichloro-1,1,1-Trifluroethane (HCFC-123)	0.61	0.77
2,2-Dimethylbutane	-0.59	-0.18
2,3,4-Trimethylpentane	0.57	0.24
2,3-Dimethylbutane	0.78	-0.13
2-Butoxyethanol	3.90	4.26
2-Chloropropane	0.32	0.44
2-Ethoxyethanol	4.34	3.90
2-Fluropropane	0.06	-0.06
2-Heptanone (Methyl pentyl ketone)	2.33	2.45
2-Hexanone	2.10	2.28
2-Isopropoxyethanol	4.16	4.19
2-Methoxyethanol	4.52	3.82
2-Methyl heptane	0.49	0.24
2-Methyl nonane	0.76	0.60
2-Methyl octane	0.52	0.42
2-Methyl-1,3-butadiene (Isoprene)	0.10	0.37
2-Methyl-1-propanol (Isobutanol)	2.92	2.98
2-Methylcyclohexanone (o-Methylcyclohexanone)	2.87	2.97
2-Methylpentane (Isohexane)	-0.39	-0.13
2-Nitropropane	2.23	1.84
2-Pentanone (Methyl propyl ketone)	2.14	2.08
2-Propanol (Isopropanol)	3.02	2.76
3-Methyl-1-butanol (Isopentanol) (Isoamyl alcohol)	2.75	3.19
3-Methylhexane	0.11	0.08
3-Methylpentane	-0.37	-0.10
3-Pentanone (Diethyl ketone)	2.21	2.09
4-Chlorobenzotrifluoride	1.43	1.23
4-Methyl-2-pentanone (Methyl isobutyl ketone)	1.96	2.16
Acetone (Propan-2-one)	2.36	1.82
Acetylene	-0.06	0.05
Allyl chloride (3-Chloro-1-propylene)	1.24	0.61

Allylbenzene	1.71	2.03
Argon	-1.52	-1.33
Benzene	1.05	1.18
Bromobenzene	1.72	1.87
Bromochloromethane (Chlorobromomethane)	1.21	1.16
Bromodichloromethane	1.49	1.51
Butan-2-one (Methyl ethyl ketone)	2.24	1.94
Butane	-0.53	-0.46
Butyl acetate	1.94	2.03
Carbon Disulfide	0.30	0.58
Carbon Monoxide	-1.67	-1.28
Carbon tetrachloride (Tetrachloromethane)	0.62	0.61
CF2=CHCl (1-Chloro-2,2-difluroethylene) (1,1-Difluro-2-chloroethene)	0.06	-0.08
CF2ClCF2Cl (1,2-Dichlorotetrafluoroethane)	-0.82	-0.56
CF2HCF2CF2CF2H	-0.36	-0.79
CF2HCF2CFH2	-0.48	-0.76
CF2HCF2H (1,1,2,2-Tetrafluoroethane)	-0.12	-0.29
CF2HCFHCH2CF2H	0.87	1.71
CF2HCH3 (1,1-Difluroethane)	0.42	-0.18
CF3CBrClH (Halothane)	0.57	0.90
CF3CBrFH (Teflurane)	-0.22	0.44
CF3CF2CF2CF2H	-1.52	-0.96
CF3CFH2 (Norflurane) (1,1,1,2-Tetrafluoroethane)	-0.25	-0.16
CF3CFHCFHCF3	-0.59	-0.83
CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	0.14	0.37
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	0.15	0.64
CF4 (Perfluoromethane)(Carbon Tetrafluoride)	-1.10	-1.84
CHF2CF2CH2Br (Halopropane)	0.75	0.71
CHF2OCF2CHFCl (Enflurane)	0.35	0.70
CHF2OCHClCF3 (Isoflurane)	0.20	0.59
CHF2OCHFCF3 (Desflurane)	-0.37	-0.04
Chlorobenzene	1.63	1.52
Chlorodibromomethane	1.88	1.76
Chloroethane	0.49	0.36
Chloroform (Trichloromethane)	1.15	1.20
cis-1,2-Dichloroethylene	1.16	1.25

Cyanoethylene Oxide	3.22	2.49
Cycloheptane	0.72	0.59
Cyclohexane	0.17	0.29
Cyclohexanone	3.33	2.92
Cyclopentane	0.24	0.09
Cyclopropane	-0.21	-0.14
Decane	1.47	0.69
Dibromomethane	1.87	1.68
Dichloromethane (Methylene chloride)	1.12	0.99
Diethyl ether	1.11	1.14
Diflurodichloromethane	-0.82	-0.49
Difluromethane (CF2H2)	0.20	-0.32
Dimethyl ether	1.16	0.76
Divinyl ether	0.41	0.47
Ethane	-1.02	-0.88
Ethanol	3.27	2.68
Ethene	-0.53	-0.62
Ethyl acetate	1.90	1.68
Ethyl formate	1.65	1.39
Ethyl t-butyl ether (ETBE)	1.07	1.67
Ethyl t-pentyl ether	1.25	1.35
Ethylbenzene	1.47	1.57
Ethylene Oxide	1.80	1.45
Fluorobenzene	1.06	1.07
Fluoroethane	0.09	-0.19
Fluorotrichloromethane (Trichlorofluromethane)	-0.06	0.20
Flurochloromethane	0.71	-0.01
Furan	0.82	0.71
Helium	-2.00	-1.72
Heptane	0.50	0.12
Hexachloroethane (Perchloroethane)	1.76	1.75
Hexaflurobenzene	0.39	0.48
Hexane	0.21	-0.07
Hydrogen	-1.77	-1.52
Iodoethane	0.83	1.01
Isobutyl acetate	1.69	1.97
Isopentylacetate (Aceticacidpentylester)	1.79	2.19
Isophorone (3,5,5-trimethylcyclohex-2-en-1-one)	3.37	3.53
Isopropyl acetate	1.55	1.74

Isopropyl bromide (2-Brompropane)	0.64	0.72
Isopropylbenzene (Cumene)	1.57	1.68
JP-10	1.76	1.66
Krypton	-1.22	-1.15
MeOCF2CHCl2 (Methoxyflurane)	1.28	1.41
Methane	-1.42	-1.19
Methanol	3.41	2.72
Methyl acetate	1.98	1.57
Methyl butyl ketone (3-Methylpentan-2-one)	2.23	2.19
Methyl tertiary-butyl ether (MTBE)	1.18	1.59
Methylchloride (Chloromethane)	0.31	0.15
Methylcyclohexane	0.70	0.35
Methylcyclopentane	-0.07	0.23
Methylpentaflurobenzene (2,3,4,5,6-Pentylfluorotoluene)	0.73	0.03
m-Methylstyrene (1-vinyl-3-methylbenzene)	2.28	2.14
m-Xylene (1,3-Dimethylbenzene)	1.59	1.63
Neon	-2.01	-1.66
Nitrogen	-1.83	-1.44
Nitrous Oxide	-0.34	-0.34
Nonane	1.17	0.50
Octane	0.68	0.31
Oxygen	-1.58	-1.34
o-Xylene	1.42	1.73
Pentachloroethane	2.02	2.34
Pentaflurobenzene	0.51	0.76
Pentane	-0.29	-0.26
Pentyl acetate	1.98	2.22
p-Methylstyrene (1-vinyl-4-methylbenzene)	2.37	2.15
Propene (Propylene)	-0.21	-0.40
Propyl acetate	1.88	1.84
Propyl bromide (1-Bromopropane)	0.97	0.83
Propylbenzene	1.67	1.72
p-Xylene	1.61	1.63
Radon	-0.39	-0.74
Styrene (Vinyl benzene)	1.67	1.88
Sulphur Hexafluoride	-2.17	-1.60
t-butanol (tert-butanol)(TBA) (2-Methyl-2-propanol)	2.68	2.78

t-Butylbenzene	1.24	1.87
t-Butylcyclohexane	1.16	1.16
tertiary-amyl alcohol (TAA)	2.59	3.04
tertiary-Amyl methyl ether (TAME)	1.22	1.83
Tetrachloroethene (Perchloroethylene)	1.19	1.05
Toluene	1.14	1.38
trans-1,2-Dichloroethylene	0.88	0.89
Tribromomethane (Bromoform)	2.15	2.25
Trichloroethene (Trichloroethylene)	1.14	1.07
Vinyl bromide (Bromoethene)	0.49	0.60
Vinyl chloride (Chloroethene)	0.17	0.12
Vinylidene fluoride (1,1-Difluoroethylene)	-0.74	-0.90
Xenon	-0.85	-0.93
OUTLIERS		
CFH2CH2CFH2	0.97	-0.05
Cyclopentadiene	1.61	0.65

Descriptors E, S, A, B and L are found from chapter 9, table 9.0.

Ref = Literature reference, Comp No = Abraham's compound number

Table 11.0. Data set used to obtain a model to correlate air to plasma distributions for VOCs in humans and rats combined

SOLUTE NAME	HUM REF[b]	HUM PLASMA LOG K	RAT REF[b]	RAT PLASMA LOG K	REF[b]	AVERAGE OF H&R LOG K	CALC[a] LOG K
Helium	1,	-2.07			1,	-2.07	-1.77
Argon	1,	-1.58			1,	-1.58	-1.57
Krypton	1,	-1.29			1,	-1.29	-1.48
Xenon	1,	-1.04			1,	-1.04	-1.36
Hydrogen	1,	-1.76			1,	-1.76	-1.67
Oxygen	1,	-1.61			1,	-1.61	-1.57
Nitrogen	1,	-1.87			1,	-1.87	-1.62
Nitrous Oxide	1,	-0.34			1,	-0.34	-0.38
Methane	1,	-1.59			1,	-1.59	-1.50
Ethane	1,	-1.34	2,	-1.44	1, 2	-1.39	-1.34
Cyclopropane	1,	-0.62	2,	-0.63	1, 2	-0.63	-0.58
Acetylene	1,	-0.11			1,	-0.11	0.07
Fluoroethane	1,	0.07			1,	0.07	-0.31
Fluoromethane	1,	0.01			1,	0.01	-0.40
1-Fluropropane	1,	-0.01			1,	-0.01	-0.10
2-Fluropropane	1,	0.00			1,	0.00	-0.29
Iodoethane	1,	0.69			1,	0.69	0.68
Fluorotrichloromethane (Trichloroflurometh ane)	1,	-0.12			1,	-0.12	-0.27
Diflurodichloromethane	1,	-1.07			1,	-1.07	-0.98
CF3CBrClH (Halothane)	1,	0.40			1,	0.40	0.40
CF2ClCF2Cl (1,2-Dichlorotetrafluoroethane)	1,	-0.66			1,	-0.66	-1.18
Dimethyl ether	1,	1.12			1,	1.12	0.90
Diethyl ether	1,	1.08	2,	1.16	1, 2	1.12	1.19
MeOCF2CHCl2 (Methoxyflurane)	1,	1.00			1,	1.00	1.10
CHF2OCHClCF3 (Isoflurane)	1,	0.20			1,	0.20	0.33
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	1,	0.05			1,	0.05	0.55
Acetone (Propan-2-one)	1,	2.54	2,	2.20	1, 2	2.37	2.25
Butan-2-one	1,	2.12			1,	2.12	2.30

(Methyl ethyl ketone)							
Methanol	1,	3.27			1,	3.27	3.02
Ethanol	1,	3.16			1,	3.16	2.90
1-Propanol	1,	2.99			1,	2.99	3.00
2-Propanol (Isopropanol)	1,	2.91			1,	2.91	3.01
2-Methyl-1-propanol (Isobutanol)	1,	2.77			1,	2.77	3.01
Sulphur hexafluride	1,	-2.28	2,	-2.41	1, 2	-2.35	-2.12
Carbon Disulfide	1,	-0.10			1,	-0.10	0.05
CHF2OCF2CHFCl (Enflurane)			2,	0.00	2,	0.00	0.39

[a]Calculated values for log K, obtained when correlating with the Abraham descriptors (E, S, A, B and L)

Descriptors E, S, A, B and L for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 18.

[b]REF= Literature reference

Table 11.5. Data set used in this model to obtain model to predict blood to plasma distributions for VOCs

SOLUTE NAME	V	LOG K PLASMA[b] (H&R)	LOG K BLOOD[b] (H&R)	LOG P (B to P)	CALCULATED LOG K[a]
Helium	0.0680	-2.07	-2.00	-0.07	-0.07
Argon	0.1900	-1.58	-1.52	-0.06	-0.09
Krypton	0.2460	-1.29	-1.22	-0.07	-0.11
Xenon	0.3290	-1.04	-0.85	-0.19	-0.12
Hydrogen	0.1086	-1.76	-1.77	0.01	-0.08
Oxygen	0.1830	-1.61	-1.58	-0.03	-0.09
Nitrogen	0.2222	-1.87	-1.83	-0.04	-0.10
Nitrous Oxide	0.2810	-0.34	-0.34	0.00	-0.06
Methane	0.2495	-1.59	-1.42	-0.17	-0.11
Ethane	0.3904	-1.39	-1.02	-0.37	-0.13
Cyclopropane	0.4227	-0.63	-0.21	-0.42	-0.21
Acetylene	0.3044	-0.11	-0.06	-0.05	-0.10
Fluoroethane	0.4081	0.07	0.09	-0.02	-0.08
1-Fluropropane	0.5490	-0.01	0.02	-0.03	-0.10
2-Fluropropane	0.5490	0.00	0.06	-0.06	-0.10
Iodoethane	0.6486	0.69	0.83	-0.14	-0.24
Fluorotrichloromethane (Trichlorofluromethane)	0.6344	-0.12	-0.06	-0.06	-0.18
Diflurodichloromethane	0.5297	-1.07	-0.82	-0.25	-0.16
CF3CBrClH (Halothane)	0.7410	0.40	0.57	-0.17	-0.21
CF2ClCF2Cl (1,2-Dichlorotetrafluoroethane)	0.7060	-0.66	-0.82	0.16	-0.15
Dimethyl ether	0.4491	1.12	1.16	-0.04	0.01
Diethyl ether	0.7309	1.12	1.11	0.01	-0.05
MeOCF2CHCl2 (Methoxyflurane)	0.8700	1.00	1.28	-0.28	-0.16
CHF2OCHClCF3 (Isoflurane)	0.8010	0.20	0.20	0.00	-0.11
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	0.7410	0.05	0.15	-0.10	-0.13
Acetone (Propan-2-one)	0.5470	2.37	2.36	0.01	0.00
Butan-2-one (Methyl ethyl ketone)	0.6879	2.12	2.24	-0.12	-0.01
Methanol	0.3082	3.27	3.41	-0.14	-0.11

Ethanol	0.4491	3.16	3.27	-0.11	-0.11
1-Propanol	0.5900	2.99	3.06	-0.07	-0.14
2-Propanol (Isopropanol)	0.5900	2.91	3.02	-0.11	-0.10
2-Methyl-1-propanol (Isobutanol)	0.7309	2.77	2.92	-0.15	-0.17
Sulphur hexafluoride	0.4643	-2.35	-2.17	-0.18	-0.03
Carbon Disulfide	0.4905	-0.10	0.30	-0.40	-0.32
CHF2OCF2CHFCl (Enflurane)	0.8009	0.00	0.35	-0.35	-0.12

[a]Calculated values for log K, obtained when correlating with the Abraham descriptors (E, S, A, B and V)

Descriptors E, S, A and B can be found in table 9.0 from chapter 9.

[b]literature reference for air to blood is found in chapter 10 (table 10.0). The reference for air to plasma data is found in table 11.0 of this chapter

Table 12.0. Air to brain distribution for VOCs

SOLUTE NAME	BRAIN REF	BRAIN HUM LOG K	BRAIN REF	BRAIN RAT LOG K	BRAIN REF (H&R)	BRAIN AVERAGE LOG K (H&R)	CALC LOG K
(CF3)2CHOCH2F (Sevoflurane)	1, 2, 3	0.10			1, 2, 3	0.10	0.55
1,1,1-Trichloroethane (Methyl chloroform)	1, 2, 3	0.92			1, 2, 3	0.92	1.08
1,2,4-Trimethylbenzene (Pseudocumene)			5, 12	1.63	5, 12	1.63	2.43
1,2,4-Trimethylcyclohexane			5, 12	1.89	5, 12	1.89	1.62
1,2-Dimethylcyclohexane			5,	1.98	5,	1.98	1.44
1,3-Butadiene	7,	-0.08			7,	-0.08	0.31
1-Butanol			1,	3.06	1,	3.06	3.00
1-Pentanol			1,	3.03	1,	3.03	3.29
1-Propanol	1, 2, 3	2.87	1,	3.09	1, 2, 3,	2.98	2.66
1-Octene			4,	1.81	4,	1.81	1.32
1-Nonene			4,	2.04	4,	2.04	1.83
1-Decene			4,	2.17	4,	2.17	1.89
2,2-Dimethylbutane	1, 2, 3	0.45			1, 2, 3	0.45	0.40
2-Methyl-1-propanol (Isobutanol)	1, 2, 3	2.61	1,	2.94	1, 2, 3	2.78	2.87
2-Methylpentane (Isohexane)	1, 2, 3	0.58			1, 2, 3	0.58	0.49
2-Methyl heptane			4,	1.35	4,	1.35	1.07
2-Methyl octane			4,	1.50	4,	1.50	1.36
2-Methyl nonane			4,	1.81	4,	1.81	1.65
2-Propanol (Isopropanol)	1, 2, 3	2.76	1,	3.05	1, 2, 3	2.91	2.50
3-Methyl-1-butanol (Isopentanol)			1,	2.79	1,	2.79	3.22
3-Methylhexane	1, 2, 3	1.01			1, 2, 3	1.01	0.81
3-Methylpentane	1, 2, 3	0.64			1, 2, 3	0.64	0.54
4-Chlorobenzotrifluoride			11,	1.60	11,	1.60	1.62
Acetone (Propan-2-one)	1, 2, 3	2.19			1, 2, 3	2.19	1.48
Argon	3,	-1.49			3,	-1.49	-1.39
Benzene	1, 2, 3	1.26	5,	1.24	1, 2, 3, 5	1.25	1.32
Butan-2-one (Methyl ethyl ketone)	1, 2, 3	2.07			1, 2, 3	2.07	1.73
Butyl acetate			1,	2.22	1,	2.22	2.17
Carbon Disulfide	3	0.90			3	0.90	0.81
CF2=CHCl (1-Chloro-2,2-difluroethylene)	2, 3	0.04			2, 3	0.04	-0.03

(1,1-Difluro-2-chloroethene)							
CF3CBrClH (Halothane)	1, 2, 3	0.77	1,	0.65	1, 2, 3	0.71	1.09
CF3CBrFH (Teflurane)	1, 3	0.05			1, 3	0.05	0.60
CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	1, 2, 3	0.26			1, 2, 3	0.26	0.37
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	1, 2, 3	0.29			1, 2, 3	0.29	0.68
CHF2OCF2CHFCl (Enflurane)	1, 2, 3	0.49			1, 2, 3	0.49	0.82
CHF2OCHClCF3 (Isoflurane)	1, 2, 3	0.46	1,	0.31	1, 2, 3	0.39	0.63
Chloroform (Trichloromethane)	1, 2, 3	1.30			1, 2, 3	1.30	1.34
Cyclohexane	1, 2, 3	1.04	5,	1.52	1, 2, 3, 5	1.28	0.89
Cyclopropane	1, 2	-0.10			1, 2	-0.10	-0.01
Decane			5,	1.84	5,	1.84	1.78
Dichloromethane (Methylene chloride)	1, 2, 3	0.78			1, 2, 3	0.78	0.98
Diethyl ether	1, 2, 3	1.10			1, 2, 3	1.10	1.23
Divinyl ether	1, 2, 3	0.54			1, 2, 3	0.54	0.55
Ethanol	1, 2, 3	3.02	1,	3.27	1, 2, 3	3.15	2.34
Ethyl acetate			1,	1.90	1,	1.90	1.58
Ethyl t-butyl ether (ETBE)			8,	1.29	8,	1.29	1.90
t-butanol (tert-butanol)(TBA) (2-Methyl-2-propanol)			8,	2.79	8,	2.79	2.60
Methyl tertiary-butyl ether (MTBE)			8,	1.54	8,	1.54	1.71
tertiary-Amyl methyl ether (TAME)			8,	1.39	8,	1.39	2.05
tertiary-amyl alcohol (TAA)			8,	2.66	8,	2.66	3.00
Ethylene Oxide	15,	1.84	14,	1.77	14, 15	1.81	1.08
Ethylene (Ethene)	15,	-0.26	15,	-0.19	15,	-0.23	-0.60
Heptane	1, 2, 3	1.09	5,	0.79	1, 2, 3, 5	0.94	0.89
Hexane	1, 2, 3	0.70	5,	1.07	1, 2, 3, 5	0.89	0.59
Isobutyl acetate			1,	2.14	1,	2.14	2.08
Isopentylacetate (Aceticacidpentylester)			1,	2.34	1,	2.34	2.42
Isopropyl acetate			1,	1.95	1,	1.95	1.72

Krypton	3,	-1.38			3,	-1.38	-1.11
MeOCF2CHCl2 (Methoxyflurane)	1, 2, 3	1.41	1,	1.39	1, 2, 3	1.40	1.53
Methane	3,	-1.39			3,	-1.39	-1.18
Methyl acetate			1,	1.85	1,	1.85	1.36
Methylcyclopentane	1, 2, 3	0.86			1, 2, 3	0.86	0.83
Methylcyclohexane			5,	1.66	5,	1.66	1.06
Neon	3,	-1.81			3,	-1.81	-1.92
Nitrogen	3,	-1.80			3,	-1.80	-1.56
Nitrous Oxide	2, 3	-0.31			2, 3	-0.31	-0.53
Nonane			5, 12	1.69	5, 12	1.69	1.49
Octane	1, 3	1.22	5,	1.52	1, 3, 5	1.37	1.19
o-Xylene (1,2-dimethylbenzene)			5,	1.39	5,	1.39	2.07
Pentane	1, 2, 3	0.34			1, 2, 3	0.34	0.29
Pentyl acetate			1,	2.38	1,	2.38	2.46
Propyl acetate			1,	2.00	1,	2.00	1.86
Propene (Propylene)	9,	-0.22	9,	-0.32	9,	-0.27	-0.23
Radon			10,	-0.51	10,	-0.51	-0.47
Sulphur Hexafluoride	3,	-1.78			3,	-1.78	-1.30
t-Butylcyclohexane			5,	1.77	5,	1.77	2.06
t-Butylbenzene			5,	1.67	5,	1.67	2.35
Toluene (Methylbenzene)	1, 2, 3	1.56	5,	1.26	1, 2, 3, 5	1.41	1.64
Trichloroethene (Trichloroethylene)	1, 2, 3	1.33	13,	1.16	1, 2, 3, 13	1.25	1.41
Xenon	2,	-0.70			2,	-0.70	-0.76
Carbon Dioxide	3,	-0.25			3,	-0.25	-0.47
OUTLIERS							
Cyanoethylene Oxide			6,	3.37	6,	3.37	2.25
Methanol	1, 2	3.18	1,	3.54	1, 2	3.36	2.39

Ref = Literature reference, H&R = humans and rats

Descriptors E, S, A, B and L for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 18.

Table 12.3. Blood to brain distribution for VOCs

SOLUTE NAME	V	BRAIN AVERAGE LOG K (H&R)	BLOOD AVERAGE LOG K (H&R)	BLOOD TO BRAIN LOG P (H&R)	CALC LOG P
(CF3)2CHOCH2F (Sevoflurane)	0.8548	0.10	-0.20	0.30	0.36
1,1,1-Trichloroethane (Methyl chloroform)	0.7576	0.92	0.63	0.29	0.25
1,2,4-Trimethylbenzene (Pseudocumene)	1.1391	1.63	1.47	0.16	0.42
1,2,4-Trimethylcyclohexane	1.2681	1.89	0.87	1.02	0.76
1,2-Dimethylcyclohexane	1.1270	1.98	0.91	1.07	0.65
1,3-Butadiene	0.5862	-0.08	0.09	-0.17	0.22
1-Butanol	0.7309	3.06	3.08	-0.02	-0.02
1-Pentanol	0.8718	3.03	2.83	0.20	0.08
1-Propanol	0.5900	2.98	3.06	-0.08	-0.13
1-Octene	1.1928	1.81	1.07	0.74	0.75
1-Nonene	1.3337	2.04	1.18	0.86	0.87
1-Decene	1.4746	2.17	1.21	0.96	0.96
2,2-Dimethylbutane	0.9540	0.45	-0.59	1.04	0.64
2-Methyl-1-propanol (Isobutanol)	0.7309	2.78	2.92	-0.14	-0.01
2-Methylpentane (Isohexane)	0.9540	0.58	-0.39	0.97	0.64
2-Methyl heptane	1.2358	1.35	0.49	0.86	0.85
2-Methyl octane	1.3767	1.50	0.52	0.98	0.95
2-Methyl nonane	1.5176	1.81	0.76	1.05	1.05
2-Propanol (Isopropanol)	0.5900	2.91	3.02	-0.11	-0.11
3-Methyl-1-butanol (Isopentanol)	0.8718	2.79	2.75	0.04	0.09
3-Methylhexane	1.0949	1.01	0.11	0.90	0.74
3-Methylpentane	0.9540	0.64	-0.37	1.01	0.64
4-Chlorobenzotrifluoride	1.0328	1.60	1.43	0.17	0.39
Acetone (Propan-2-one)	0.5470	2.19	2.36	-0.17	-0.21
Argon	0.1900	-1.49	-1.52	0.03	0.08
Benzene	0.7164	1.25	1.05	0.20	0.15
Butan-2-one (Methyl ethyl ketone)	0.6879	2.07	2.24	-0.17	-0.10
Butyl acetate	1.0284	2.22	1.94	0.28	0.22
Carbon Disulfide	0.4905	0.90	0.30	0.60	0.17
CF2=CHCl (1-Chloro-2,2-difluroethylene) (1,1-Difluro-2-chloroethene)	0.5052	0.04	0.06	-0.02	0.10
CF3CBrClH (Halothane)	0.7410	0.71	0.57	0.14	0.22
CF3CBrFH (Teflurane)	0.6360	0.05	-0.22	0.27	0.22
CF3CH2Cl (2chloro111trifluoroethane)	0.5659	0.26	0.14	0.12	0.09

CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	0.7410	0.29	0.15	0.14	0.24
CHF2OCF2CHFCl (Enflurane)	0.8009	0.49	0.35	0.14	0.23
CHF2OCHClCF3 (Isoflurane)	0.8010	0.39	0.20	0.19	0.19
Chloroform (Trichloromethane)	0.6167	1.30	1.15	0.15	0.08
Cyclopropane	0.4227	-0.10	-0.21	0.11	0.14
Dichloromethane (Methylene chloride)	0.4943	0.78	1.12	-0.34	-0.04
Diethyl ether	0.7309	1.10	1.11	-0.01	0.19
Divinyl ether	0.6449	0.54	0.41	0.13	0.17
Ethanol	0.4491	3.15	3.27	-0.12	-0.23
Ethyl acetate	0.7466	1.90	1.90	0.00	0.01
Ethyl t-butyl ether (ETBE)	1.0127	1.29	1.07	0.22	0.40
t-butanol (tert-butanol)(TBA) (2-Methyl-2-propanol)	0.7309	2.79	2.68	0.11	0.02
Methyl tertiary-butyl ether (MTBE)	0.8718	1.54	1.18	0.36	0.27
tertiary-Amyl methyl ether (TAME)	1.0127	1.39	1.22	0.17	0.37
tertiary-amyl alcohol (TAA)	0.8718	2.66	2.59	0.07	0.12
Ethylene Oxide	0.3405	1.81	1.80	0.01	-0.33
Ethylene (Ethene)	0.3474	-0.23	-0.53	0.31	0.12
Heptane	1.0949	0.94	0.50	0.44	0.74
Hexane	0.9540	0.89	0.21	0.68	0.64
Isobutyl acetate	1.0284	2.14	1.69	0.45	0.23
Isopentylacetate (Aceticacidpentylester)	1.1693	2.34	1.79	0.55	0.34
Isopropyl acetate	0.8875	1.95	1.55	0.40	0.13
Krypton	0.2460	-1.38	-1.22	-0.16	0.12
MeOCF2CHCl2 (Methoxyflurane)	0.8700	1.40	1.28	0.12	0.15
Methane	0.2495	-1.39	-1.42	0.03	0.13
Methyl acetate	0.6057	1.85	1.98	-0.13	-0.10
Methylcyclopentane	0.8454	0.86	-0.07	0.93	0.51
Methylcyclohexane	0.9863	1.66	0.70	0.96	0.64
Neon	0.0850	-1.81	-2.01	0.20	0.01
Nitrogen	0.2222	-1.80	-1.83	0.03	0.11
Nitrous Oxide	0.2810	-0.31	-0.34	0.03	-0.07
Nonane	1.3767	1.69	1.17	0.52	0.95
Octane	1.2358	1.37	0.68	0.69	0.85
o-Xylene (1,2-dimethylbenzene)	0.9982	1.39	1.42	-0.03	0.33
Pentane	0.8131	0.34	-0.29	0.63	0.54
Pentyl acetate	1.1693	2.38	1.98	0.40	0.33
Propyl acetate	0.8875	2.00	1.88	0.12	0.12
Propene (Propylene)	0.4883	-0.27	-0.21	-0.06	0.24

Radon	0.3840	-0.51	-0.39	-0.12	0.22
Sulphur Hexafluoride	0.4643	-1.78	-2.17	0.39	0.38
t-Butylcyclohexane	1.4090	1.77	1.16	0.61	0.89
t-Butylbenzene	1.2800	1.67	1.24	0.43	0.57
Toluene (methylbenzene)	0.8573	1.41	1.14	0.27	0.25
Trichloroethene (Trichloroethylene)	0.7146	1.25	1.14	0.11	0.24
Xenon	0.3290	-0.70	-0.85	0.15	0.18
OUTLIERS					
Decane	1.5176	1.84	1.47	0.37	1.01
Cyclohexane	0.8454	1.28	0.17	1.11	0.53

Descriptors E, S, A and B for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 18.

Note no blood values carbon dioxide

H&R = human and rat

Ref = Literature reference

Table 12.6. Plasma to brain distribution for VOCs

SOLUTE NAME	V	LOG K BRAIN (H&R)	LOG K PLASMA[a] (H&R)	LOG P (PLASMA TO BRAIN)	CALC LOG P
Argon	0.1900	-1.49	-1.58	0.09	0.16
Krypton	0.2460	-1.38	-1.29	-0.09	0.19
Xenon	0.3290	-0.70	-1.04	0.34	0.24
Nitrogen	0.2222	-1.80	-1.87	0.07	0.18
Nitrous Oxide	0.2810	-0.31	-0.34	0.03	0.06
Methane	0.2495	-1.39	-1.59	0.20	0.20
Cyclopropane	0.4227	-0.10	-0.63	0.53	0.27
CF3CBrClH (Halothane)	0.7410	0.71	0.40	0.31	0.42
Diethyl ether	0.7309	1.10	1.12	-0.02	0.06
MeOCF2CHCl2 (Methoxyflurane)	0.8700	1.40	1.00	0.40	0.34
CHF2OCHClCF3 (Isoflurane)	0.8010	0.39	0.20	0.19	0.36
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	0.7410	0.29	0.05	0.24	0.21
Acetone (Propan-2-one)	0.5470	2.19	2.37	-0.18	-0.16
Butan-2-one (Methyl ethyl ketone)	0.6879	2.07	2.12	-0.05	-0.11
Methanol	0.3082	3.36	3.27	0.09	-0.09
Ethanol	0.4491	3.15	3.16	-0.01	-0.03
1-Propanol	0.5900	2.98	2.99	-0.01	0.05
2-Propanol (Isopropanol)	0.5900	2.91	2.91	0.00	-0.02
2-Methyl-1-propanol (Isobutanol)	0.7309	2.78	2.77	0.01	0.14
Sulphur hexafluride	0.4643	-1.78	-2.35	0.57	0.34
CHF2OCF2CHFCl (Enflurane)	0.8009	0.49	0.00	0.49	0.37

[a]The literature data for air to plasma VOC distribution was obtained from chapter 11 (table 11.0)

Note references for plasma data can be found in the plasma chapter of the thesis

H&R = human and rat

Table 12.9. Plasma to brain and blood to brain distribution combined for VOCs

SOLUTE NAME	BLOOD OR PLASMA	BLOOD OR PLASMA TO BRAIN LOG P (H&R)	CALC LOG P
(CF3)2CHOCH2F (Sevoflurane)	BLOOD	0.30	0.39
1,1,1-Trichloroethane (Methyl chloroform)	BLOOD	0.29	0.27
1,2,4-Trimethylbenzene (Pseudocumene)	BLOOD	0.16	0.42
1,2,4-Trimethylcyclohexane	BLOOD	1.02	0.76
1,2-Dimethylcyclohexane	BLOOD	1.07	0.65
1,3-Butadiene	BLOOD	-0.17	0.23
1-Butanol	BLOOD	-0.02	0.04
1-Pentanol	BLOOD	0.20	0.14
1-Propanol	BLOOD	-0.08	-0.06
1-Octene	BLOOD	0.74	0.74
1-Nonene	BLOOD	0.86	0.84
1-Decene	BLOOD	0.96	0.94
2,2-Dimethylbutane	BLOOD	1.04	0.64
2-Methyl-1-propanol (Isobutanol)	BLOOD	-0.14	0.06
2-Methylpentane (Isohexane)	BLOOD	0.97	0.64
2-Methyl heptane	BLOOD	0.86	0.84
2-Methyl octane	BLOOD	0.98	0.94
2-Methyl nonane	BLOOD	1.05	1.04
2-Propanol (Isopropanol)	BLOOD	-0.11	-0.05
3-Methyl-1-butanol (Isopentanol)	BLOOD	0.04	0.16
3-Methylhexane	BLOOD	0.90	0.74
3-Methylpentane	BLOOD	1.01	0.64
4-Chlorobenzotrifluoride	BLOOD	0.17	0.41
Acetone (Propan-2-one)	BLOOD	-0.17	-0.19
Argon	BLOOD	0.03	0.11
Benzene	BLOOD	0.20	0.17
Butan-2-one (Methyl ethyl ketone)	BLOOD	-0.17	-0.09
Butyl acetate	BLOOD	0.28	0.22
Carbon Disulfide	BLOOD	0.60	0.18
CF2=CHCl	BLOOD	-0.02	0.17

CF3CBrClH (Halothane)	BLOOD	0.14	0.27
CF3CBrFH (Teflurane)	BLOOD	0.27	0.29
CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	BLOOD	0.12	0.16
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	BLOOD	0.14	0.24
CHF2OCF2CHFCl (Enflurane)	BLOOD	0.14	0.27
CHF2OCHClCF3 (Isoflurane)	BLOOD	0.19	0.24
Chloroform (Trichloromethane)	BLOOD	0.15	0.14
Cyclopropane	BLOOD	0.11	0.16
Dichloromethane (Methylene chloride)	BLOOD	-0.34	0.01
Diethyl ether	BLOOD	-0.01	0.18
Divinyl ether	BLOOD	0.13	0.18
Ethanol	BLOOD	-0.12	-0.15
Ethyl acetate	BLOOD	0.00	0.01
Ethyl t-butyl ether (ETBE)	BLOOD	0.22	0.36
t-butanol (tert-butanol)(TBA) (2-Methyl-2-propanol)	BLOOD	0.11	0.06
Methyl tertiary-butyl ether (MTBE)	BLOOD	0.36	0.24
tertiary-Amyl methyl ether (TAME)	BLOOD	0.17	0.34
tertiary-amyl alcohol (TAA)	BLOOD	0.07	0.16
Ethylene Oxide	BLOOD	0.01	-0.29
Ethylene (Ethene)	BLOOD	0.31	0.14
Heptane	BLOOD	0.44	0.74
Hexane	BLOOD	0.68	0.64
Isobutyl acetate	BLOOD	0.45	0.23
Isopentylacetate (Aceticacidpentylester)	BLOOD	0.55	0.33
Isopropyl acetate	BLOOD	0.40	0.13
Krypton	BLOOD	-0.16	0.15
MeOCF2CHCl2 (Methoxyflurane)	BLOOD	0.12	0.19
Methane	BLOOD	0.03	0.15
Methyl acetate	BLOOD	-0.13	-0.10
Methylcyclopentane	BLOOD	0.93	0.52

Methylcyclohexane	BLOOD	0.96	0.64
Neon	BLOOD	0.20	0.03
Nitrogen	BLOOD	0.03	0.13
Nitrous Oxide	BLOOD	0.03	-0.04
Nonane	BLOOD	0.52	0.94
Octane	BLOOD	0.69	0.84
o-Xylene (1,2-dimethylbenzene)	BLOOD	-0.03	0.34
Pentane	BLOOD	0.63	0.54
Pentyl acetate	BLOOD	0.40	0.32
Propyl acetate	BLOOD	0.12	0.12
Propene (Propylene)	BLOOD	-0.06	0.25
Radon	BLOOD	-0.12	0.24
Sulphur Hexafluoride	BLOOD	0.39	0.40
t-Butylcyclohexane	BLOOD	0.61	0.87
t-Butylbenzene	BLOOD	0.43	0.56
Toluene (methylbenzene)	BLOOD	0.27	0.27
Trichloroethene (Trichloroethylene)	BLOOD	0.11	0.27
Xenon	BLOOD	0.15	0.20
Argon	PLASMA	0.09	0.11
Krypton	PLASMA	-0.09	0.15
Xenon	PLASMA	0.34	0.20
Nitrogen	PLASMA	0.07	0.13
Nitrous Oxide	PLASMA	0.03	-0.04
Methane	PLASMA	0.20	0.15
Cyclopropane	PLASMA	0.53	0.16
CF3CBrClH (Halothane)	PLASMA	0.31	0.27
Diethyl ether	PLASMA	-0.02	0.18
MeOCF2CHCl2 (Methoxyflurane)	PLASMA	0.40	0.19
CHF2OCHClCF3 (Isoflurane)	PLASMA	0.19	0.24
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	PLASMA	0.23	0.24
Acetone (Propan-2-one)	PLASMA	-0.18	-0.19
Butan-2-one (Methyl ethyl ketone)	PLASMA	-0.05	-0.09
Methanol	PLASMA	0.09	-0.27
Ethanol	PLASMA	-0.01	-0.15
1-Propanol	PLASMA	-0.01	-0.06

2-Propanol (Isopropanol)	PLASMA	0.00	-0.05
2-Methyl-1-propanol (Isobutanol)	PLASMA	0.01	0.06
Sulphur hexafluride	PLASMA	0.57	0.40
CHF2OCF2CHFCl (Enflurane)	PLASMA	0.49	0.27

H&R = human and rat

Descriptor V, are found from tables 12.3 and 12.6.

Descriptors E, S, A and B for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 18.

Table 13.0. Air to kidney distribution for VOCs

SOLUTE NAME	KID REF	KIDNEY HUM LOG K	KID REF	KIDNEY RAT LOG K	KID REF (H&R)	KIDNEY AVERAGE LOG K (H&R)	CALC LOG K
(CF3)2CHOCH2F (Sevoflurane)	1, 2, 4	0.25			1, 2, 4	0.25	0.54
1,1,1-Trichloroethane (Methyl chloroform)	1, 2, 4	0.83			1, 2, 4	0.83	1.09
1,2,4-Trimethylbenzene (Pseudocumene)			6, 13	2.19	6, 13	2.19	2.57
1,2,4-Trimethylcyclohexane			6, 13	2.33	6, 13	2.33	1.67
1,2-Dimethylcyclohexane			6,	2.30	6,	2.30	1.47
1,3-Butadiene	7,	-0.08			7,	-0.08	0.26
1-Butanol			1	3.06	1	3.06	3.04
1-Pentanol			1	3.04	1	3.04	3.37
1-Propanol	1, 2, 4	2.84	1	3.09	1, 2, 4	2.97	2.67
1-Octene			5,	2.15	5,	2.15	1.39
1-Nonene			5,	2.15	5,	2.15	1.89
1-Decene			5,	2.21	5,	2.21	2.02
2,2-Dimethylbutane	1, 2, 4	0.15			1, 2, 4	0.15	0.38
2-Methyl-1,3-butadiene (Isoprene)			12,	0.03	12,	0.03	0.62
2-Methyl-1-propanol (Isobutanol)	1, 2, 4	2.57	1	2.94	1, 2, 4	2.76	2.91
2-Methylpentane (Isohexane)	1, 2, 4	0.30			1, 2, 4	0.30	0.47
2-Methyl heptane			5,	1.60	5,	1.60	1.11
2-Methyl octane			5,	1.59	5,	1.59	1.43
2-Methyl nonane			5,	1.65	5,	1.65	1.74
2-Propanol (Isopropanol)	1, 2, 4	2.70	1	3.03	1, 2, 4	2.87	2.56
3-Methyl-1-butanol (Isopentanol)			1	2.86	1	2.86	3.29
3-Methylhexane	1, 2, 4	0.86			1, 2, 4	0.86	0.83
3-Methylpentane	1, 2, 4	0.40			1, 2, 4	0.40	0.53
4-Chlorobenzotrifluoride			11,	1.58	11,	1.58	1.63
Acetone (Propan-2-one)	1, 2, 4	2.16			1, 2, 4	2.16	1.61
Benzene	1, 2, 4	1.08	6,	1.51	1, 2, 4, 6	1.30	1.34
Bromodichloromethane			15,	1.52	15,	1.52	1.51
Butan-2-one (Methyl ethyl ketone)	1, 2, 4	2.03			1, 2, 4	2.03	1.93
Butyl acetate			1	2.39	1	2.39	2.41

Carbon Disulfide	4,	0.90			4,	0.90	0.74
CF2=CHCl (1-Chloro-2,2-difluroethylene) (1,1-Difluro-2-chloroethene)	2, 4	0.00			2, 4	0.00	-0.21
CF3CBrClH (Halothane)	1, 2, 4	0.69			1, 2, 4	0.69	0.99
CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	1, 2, 4	0.32			1, 2, 4	0.32	0.20
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	1, 2, 4	0.12			1, 2, 4	0.12	0.72
CHF2OCF2CHFCl (Enflurane)	1, 2, 4	0.53			1, 2, 4	0.53	0.77
CHF2OCHClCF3 (Isoflurane)	1, 2, 4	0.29			1, 2, 4	0.29	0.56
Chloroform (Trichloromethane)	1, 2, 4	1.04			1, 2, 4	1.04	1.22
Cyclohexane	1, 2, 4	0.86			1, 2, 4	0.86	0.88
Cyclopropane	1, 2, 4	-0.38			1, 2, 4	-0.38	-0.13
Decane			6,	1.75	6,	1.75	1.90
Dichloromethane (Methylene chloride)	1, 2, 4	0.76			1, 2, 4	0.76	0.88
Diethyl ether	1, 2, 4	1.02			1, 2, 4	1.02	1.40
Divinyl ether	1, 2, 4	0.30			1, 2, 4	0.30	0.52
Ethanol	1, 2, 4	2.97	1	3.31	1, 2, 4	3.14	2.32
Ethyl acetate			1	1.94	1	1.94	1.75
Ethyl t-butyl ether (ETBE)			8,	2.10	8,	2.10	2.19
t-butanol (tert-butanol) (TBA) (2-Methyl-2-propanol)			3, 8	2.78	3, 8	2.78	2.70
Methyl tertiary-butyl ether (MTBE)			3, 8	1.84	3, 8	1.84	1.97
tertiary-Amyl methyl ether (TAME)			8,	2.21	8,	2.21	2.35
tertiary-amyl alcohol (TAA)			8,	2.68	8,	2.68	3.14
Ethylene Oxide	14,	1.83			14,	1.83	1.09
Ethylene (Ethene)	14,	-0.44	14,	-0.55	14,	-0.50	-0.73
Heptane	1, 2, 4	0.95	6,	1.20	1, 2, 4, 6	1.08	0.91
Hexane	1, 2, 4	0.48	6,	1.41	1, 2, 4, 6	0.95	0.58
Isobutyl acetate			1	2.33	1	2.33	2.32
Isopentylacetate (Aceticacidpentylester)			1	2.48	1	2.48	2.70

Isopropyl acetate			1	2.15	1	2.15	1.92
Krypton	4	-1.37			4	-1.37	-1.29
MeOCF2CHCl2 (Methoxyflurane)	1, 2, 4	1.32			1, 2, 4	1.32	1.55
Methyl acetate			1	1.92	1	1.92	1.50
Methylcyclopentane	1, 2, 4	0.67			1, 2, 4	0.67	0.82
Nitrous Oxide	2, 4	-0.40			2, 4	-0.40	-0.65
Nonane			6, 13	1.65	6, 13	1.65	1.57
Octane	1, 4	0.91	6,	1.76	1, 4, 6	1.34	1.24
Oxygen	4	-1.52			4	-1.52	-1.62
o-Xylene (1,2-dimethylbenzene)			6,	1.89	6,	1.89	2.17
Pentane	1, 2, 4	-0.22			1, 2, 4	-0.22	0.25
Pentyl acetate			1	2.51	1	2.51	2.73
Propyl acetate			1	2.29	1	2.29	2.07
Propene (Propylene)	9,	-0.35	9,	-0.28	9,	-0.32	-0.31
Radon			10,	-0.55	10,	-0.55	-0.58
t-Butylcyclohexane			6,	2.33	6,	2.33	2.19
t-Butylbenzene			6,	2.34	6,	2.34	2.49
Toluene (Methylbenzene)	1, 2, 4	1.26	6,	1.79	1, 2, 4, 6	1.53	1.69
Trichloroethene (Trichloroethylene)	1, 2, 4	1.12			1, 2, 4	1.12	1.36
Xenon	2	-1.00			2	-1.00	-0.91
OUTLIER							
Methanol	1, 2, 4	3.13	1	3.50	1, 2, 4	3.32	2.31
Methylcyclohexane			6,	2.05	6,	2.05	1.10

Ref = Literature reference

Descriptors E, S, A, B and L for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 18.

Table 13.3. Blood to kidney distribution for VOCs

SOLUTE NAME	V	KIDNEY AVERAGE LOG K (H&R)	BLOOD LOG K	KIDNEY LOG P (H&R)	CALC LOG P
(CF3)2CHOCH2F (Sevoflurane)	0.8548	0.25	-0.20	0.45	0.27
1,1,1-Trichloroethane (Methyl chloroform)	0.7576	0.83	0.63	0.20	0.25
1,2,4-Trimethylbenzene (Pseudocumene)	1.1391	2.19	1.47	0.72	0.61
1,3-Butadiene	0.5862	-0.08	0.09	-0.17	0.17
1-Butanol	0.7309	3.06	3.08	-0.02	0.00
1-Pentanol	0.8718	3.04	2.83	0.21	0.13
1-Propanol	0.5900	2.97	3.06	-0.09	-0.14
1-Octene	1.1928	2.15	1.07	1.08	0.80
1-Nonene	1.3337	2.15	1.18	0.97	1.05
1-Decene	1.4746	2.21	1.21	1.00	1.08
2,2-Dimethylbutane	0.9540	0.15	-0.59	0.74	0.57
2-Methyl-1,3-butadiene (Isoprene)	0.7271	0.03	0.10	-0.07	0.30
2-Methyl-1-propanol (Isobutanol)	0.7309	2.76	2.92	-0.16	0.01
2-Methylpentane (Isohexane)	0.9540	0.30	-0.39	0.69	0.57
2-Methyl heptane	1.2358	1.60	0.49	1.11	0.85
2-Methyl octane	1.3767	1.59	0.52	1.07	0.99
2-Methyl nonane	1.5176	1.65	0.76	0.89	1.13
2-Propanol (Isopropanol)	0.5900	2.87	3.02	-0.15	-0.06
3-Methyl-1-butanol (Isopentanol)	0.8718	2.86	2.75	0.11	0.15
3-Methylhexane	1.0949	0.86	0.11	0.75	0.71
3-Methylpentane	0.9540	0.40	-0.37	0.77	0.57
4-Chlorobenzotrifluoride	1.0328	1.58	1.43	0.15	0.43
Acetone (Propan-2-one)	0.5470	2.16	2.36	-0.20	-0.06
Benzene	0.7164	1.30	1.05	0.25	0.19
Bromodichloromethane	0.6693	1.52	1.49	0.03	-0.05
Butan-2-one (Methyl ethyl ketone)	0.6879	2.03	2.24	-0.21	0.11
Butyl acetate	1.0284	2.39	1.94	0.45	0.47
Carbon Disulfide	0.4905	0.90	0.30	0.60	0.11

CF2=CHCl (1-Chloro-2,2-difluroethylene) (1,1-Difluro-2-chloroethene)	0.5052	0.00	0.06	-0.06	-0.18
CF3CBrClH (Halothane)	0.7410	0.69	0.57	0.12	0.07
CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	0.5659	0.32	0.14	0.18	-0.13
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	0.7410	0.12	0.15	-0.03	0.31
CHF2OCF2CHFCl (Enflurane)	0.8009	0.53	0.35	0.18	0.12
CHF2OCHClCF3 (Isoflurane)	0.8010	0.29	0.20	0.09	0.08
Chloroform (Trichloromethane)	0.6167	1.04	1.15	-0.11	-0.07
Cyclohexane	0.8454	0.86	0.17	0.69	0.46
Cyclopropane	0.4227	-0.38	-0.21	-0.17	-0.01
Dichloromethane (Methylene chloride)	0.4943	0.76	1.12	-0.36	-0.19
Diethyl ether	0.7309	1.02	1.11	-0.09	0.36
Divinyl ether	0.6449	0.30	0.41	-0.11	0.14
Ethanol	0.4491	3.14	3.27	-0.13	-0.28
Ethyl acetate	0.7466	1.94	1.90	0.04	0.19
Ethyl t-butyl ether (ETBE)	1.0127	2.10	1.07	1.03	0.71
t-butanol (tert-butanol)(TBA) (2-Methyl-2-propanol)	0.7309	2.78	2.68	0.10	0.13
Methyl tertiary-butyl ether (MTBE)	0.8718	1.84	1.18	0.66	0.55
tertiary-Amyl methyl ether (TAME)	1.0127	2.21	1.22	0.99	0.69
tertiary-amyl alcohol (TAA)	0.8718	2.68	2.59	0.09	0.27
Ethylene Oxide	0.3405	1.83	1.80	0.03	-0.34
Ethylene (Ethene)	0.3474	-0.50	-0.53	0.04	-0.04
Heptane	1.0949	1.08	0.50	0.58	0.71
Hexane	0.9540	0.95	0.21	0.74	0.57
Isobutyl acetate	1.0284	2.33	1.69	0.64	0.49
Isopentylacetate (Aceticacidpentylester)	1.1693	2.48	1.79	0.69	0.63
Isopropyl acetate	0.8875	2.15	1.55	0.60	0.35
Krypton	0.2460	-1.37	-1.22	-0.15	-0.12

MeOCF2CHCl2 (Methoxyflurane)	0.8700	1.32	1.28	0.04	0.14
Methyl acetate	0.6057	1.92	1.98	-0.06	0.05
Methylcyclopentane	0.8454	0.67	-0.07	0.74	0.45
Nitrous Oxide	0.2810	-0.40	-0.34	-0.06	-0.23
Nonane	1.3767	1.65	1.17	0.48	0.99
Octane	1.2358	1.34	0.68	0.66	0.85
Oxygen	0.1830	-1.52	-1.58	0.06	-0.18
o-Xylene (1,2-dimethylbenzene)	0.9982	1.89	1.42	0.47	0.46
Pentane	0.8131	-0.22	-0.29	0.07	0.44
Pentyl acetate	1.1693	2.51	1.98	0.53	0.61
Propyl acetate	0.8875	2.29	1.88	0.41	0.34
Propene (Propylene)	0.4883	-0.32	-0.21	-0.11	0.11
Radon	0.3840	-0.55	-0.39	-0.16	0.02
t-Butylcyclohexane	1.4090	2.33	1.16	1.17	1.04
t-Butylbenzene	1.2800	2.34	1.24	1.10	0.77
Toluene (Methylbenzene)	0.8573	1.53	1.14	0.39	0.33
Trichloroethene (Trichloroethylene)	0.7146	1.12	1.14	-0.02	0.16
Xenon	0.3290	-1.00	-0.85	-0.15	-0.04
OUTLIER					
1,2-Dimethylcyclohexane	1.1270	2.30	0.91	1.39	0.69
Decane	1.5176	1.75	1.47	0.28	1.05
1,2,4-Trimethylcyclohexane	1.2681	2.33	0.87	1.46	0.86

Ref = Literature reference

Descriptors E, S, A and B for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 18.

Table 14.0. Air to fat distribution for VOCs

SOLUTE NAME	FAT REF	FAT HUM LOG K	FAT REF	FAT RAT LOG K	FAT REF	FAT AVERAGE LOG K (H&R)	CALC LOG K
(CF3)2CHOCH2F (Sevoflurane)	1, 2, 4	1.62			1, 2, 4	1.62	1.54
1,1,1,2-Tetrachloroethane			1, 3	3.33	1, 3	3.33	3.35
1,1,1-Trichloroethane (Methyl chloroform)	2, 4	2.40	1, 3	2.42	1, 2, 3, 4	2.41	2.33
1,1,2,2-Tetrachloroethane			1, 3	3.58	1, 3	3.58	3.68
1,1,2-Trichloroethane			1, 3	3.16	1, 3	3.16	3.19
1,1-Dichloroethane			1, 3	2.21	1, 3	2.21	2.25
1,1-Dichloroethylene			1, 3	1.84	1, 3	1.84	1.80
1,1-Dichloro-1-Fluoroethane (HCFC-141b)			15,	1.56	15,	1.56	1.73
1,2,4-Trimethylbenzene (Pseudocumene)			6, 18	3.25	6, 18	3.25	3.75
1,2,4-Trimethylcyclohexane			6, 18	3.19	6, 18	3.19	3.17
1,2-Dibromoethane			1, 3	3.09	1, 3	3.09	3.29
1,2-Dichloroethane			1, 3	2.54	1, 3	2.54	2.56
1,2-Dichloropropane			1, 3	2.70	1, 3	2.70	2.64
1,2-Dimethylcyclohexane			6,	3.24	6,	3.24	2.96
1,3-Butadiene	8,	1.35			8,	1.35	1.31
1,2-Epoxy-3-butene (BMO)			24,	2.23	24,	2.23	2.11
1-Bromo-2-chloroethane (1-Chloro-2-Bromoethane)			1, 3	2.98	1, 3	2.98	2.91
1-Butanol			1	2.95	1	2.95	3.02
1-Chloropropane (Propyl chloride)			1, 3	2.07	1, 3	2.07	1.92
1-Nitropropane			1, 3	2.70	1, 3	2.70	2.91
1-Pentanol			1	3.41	1	3.41	3.39
1-Propanol	1, 2, 4	2.47	1	2.60	1, 2, 4	2.54	2.59
1-Octene			5,	2.89	5,	2.89	2.69
1-Nonene			5,	3.24	5,	3.24	3.10

1-Decene			5,	3.36	5,	3.36	3.40
2,2,4-Trimethylpentane (Isooctane)			1, 3	2.47	1, 3	2.47	2.26
2,2-Dimethylbutane	1, 2, 4	1.82			1, 2, 4	1.82	1.70
2,2-Dichloro-1,1,1-Trifluroethane (HCFC-123)			16,	1.80	16,	1.80	1.92
2,3,4-Trimethylpentane			1, 3	2.65	1, 3	2.65	2.54
2-Chloropropane			1, 3	1.84	1, 3	1.84	1.72
2-Heptanone (Methyl pentyl ketone)			1	3.66	1	3.66	3.41
2-Methyl-1,3-butadiene (Isoprene)			1, 3, 17	1.84	1, 3, 17	1.84	1.73
2-Methyl-1-propanol (Isobutanol)	1, 2, 4	2.59	1	2.86	1, 2, 4	2.73	2.86
2-Methylpentane (Isohexane)	1, 2, 4	1.94			1, 2, 4	1.94	1.81
2-Methyl heptane			5,	2.30	5,	2.30	2.53
2-Methyl octane			5,	2.56	5,	2.56	2.90
2-Methyl nonane			5,	2.84	5,	2.84	3.26
2-Nitropropane			1, 3	2.19	1, 3	2.19	2.63
2-Pentanone (Methyl propyl ketone)			1	2.57	1	2.57	2.67
2-Propanol (Isopropanol)	1, 2, 4	2.26	1	2.44	1, 2, 4	2.35	2.31
3-Methyl-1-butanol (Isopentanol)			1	3.18	1	3.18	3.30
3-Methylhexane	1, 2, 4	2.44			1, 2, 4	2.44	2.21
3-Methylpentane	1, 2, 4	2.01			1, 2, 4	2.01	1.87
4-Methyl-2-pentanone (Methyl isobutyl ketone)			1	2.72	1	2.72	2.89
4-Chlorobenzotrifluoride			13	2.99	13	2.99	3.17
Acetone (Propan-2-one)	1, 2, 4	1.93	1	1.89	1, 2, 4	1.91	1.96
Allyl chloride (3-Chloro-1-propylene)			1, 3	2.00	1, 3	2.00	1.96
Benzene	1, 2, 4	2.62	1, 3, 6	2.65	1, 2, 3, 4, 6	2.64	2.47
Bromochloromethane (Chlorobromomethane)			1, 3	2.51	1, 3	2.51	2.41

Bromodichloromethane			23,	2.72	23,	2.72	2.82
Butan-2-one (Methyl ethyl ketone)	1, 2, 4	2.21	1	2.30	1, 2, 4	2.26	2.34
Butyl acetate			1	3.18	1	3.18	3.03
Carbon tetrachloride (Tetrachloromethane)			1, 3	2.56	1, 3	2.56	2.35
CF2=CHCl (1-Chloro-2,2-difluroethylene) (1,1-Difluro-2-chloroethene)	2, 4	1.30			2, 4	1.30	0.95
CF3CBrClH (Halothane)	1, 2, 4	2.31	1, 3, 16	2.23	1, 2, 3, 4, 16	2.27	2.13
CF3CBrFH (Teflurane)	1, 4	1.30			1, 4	1.30	1.48
CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	1, 2, 4	1.53	1, 3	1.33	1, 2, 3, 4	1.43	1.38
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	1, 2, 4	1.53			1, 2, 4	1.53	1.45
CHF2CF2CH2Br (Halopropane)	4	2.35			4	2.35	2.01
CHF2OCF2CHFCl (Enflurane)	1, 4	1.97			1, 4	1.97	1.79
CHF2OCHClCF3 (Isoflurane)	1, 2, 4	1.84	1, 3	1.99	1, 2, 3, 4	1.92	1.68
Chlorobenzene			1, 3	3.11	1, 3	3.11	3.20
Chlorodibromomethane			1, 3	3.28	1, 3	3.28	3.11
Chloroethane			1, 3	1.59	1, 3	1.59	1.53
Chloroform (Trichloromethane)	1, 2, 4	2.45	1, 3	2.31	1, 2, 3, 4	2.38	2.44
cis-1,2-Dichloroethylene			1, 3	2.36	1, 3	2.36	2.40
Cyclohexane	1, 2, 4	2.41	1, 3, 6	2.47	1, 2, 3, 4, 6,	2.44	2.24
Cyclopropane	1, 2, 4	0.99			1, 2, 4	0.99	1.11
Cyanoethylene Oxide			7,	3.11	7,	3.11	2.76
Decane			6,	3.05	6,	3.05	3.43
Dibromomethane			1, 3	2.90	1, 3	2.90	2.85
Dichloromethane (Methylene chloride)	1, 2, 4	1.93	1, 3	2.08	1, 2, 3, 4	2.01	2.08
Diethyl ether	1, 2, 4	1.80	1, 3	1.68	1, 2, 3, 4	1.74	1.78

Difluromethane (CF2H2)			1, 3	0.16	1, 3	0.16	0.44
Divinyl ether	1, 2, 4	1.60			1, 2, 4	1.60	1.60
Ethanol	1, 2, 4	2.33	1	2.35	1, 2, 4	2.34	2.19
Ethyl acetate			1	2.18	1	2.18	2.27
Ethyl t-butyl ether (ETBE)			9,	2.13	9,	2.13	2.28
t-butanol (tert-butanol)(TBA) (2-Methyl-2-propanol)			9, 25	2.20	9, 25	2.20	2.39
Methyl tertiary-butyl ether (MTBE)			9, 25	1.97	9, 25	1.97	2.07
tertiary-Amyl methyl ether (TAME)			9,	2.52	9,	2.52	2.47
tertiary-amyl alcohol (TAA)			9,	2.66	9,	2.66	2.88
Ethylbenzene	1	3.25	1	3.19	1,	3.22	3.21
Ethylene Oxide	21,	1.63	20,	1.64	20, 21	1.64	1.75
Ethylene (Ethene)	1, 4, 21	0.08	21,	0.31	1, 4, 21,	0.20	0.26
Flurochloromethane			1, 3	1.19	1, 3	1.19	1.08
Furan			11,	1.81	11,	1.81	1.80
Heptane	1, 2, 4	2.59	1, 3, 6	2.42	1, 2, 3, 4, 6	2.51	2.31
Hexachloroethane (Perchloroethane)			1, 3	3.52	1, 3	3.52	3.98
Hexane	1, 2, 4	2.02	1, 3, 6	2.22	1, 2, 3, 4, 6	2.12	1.93
Isobutyl acetate			1	3.05	1	3.05	2.87
Isopentylacetate (Aceticacidpentylester)			1	3.44	1	3.44	3.30
Isopropyl acetate			1	2.48	1	2.48	2.41
Isopropyl bromide (2-Brompropane)			1, 3	2.20	1, 3	2.20	2.04
JP-10			3,	4.01	3,	4.01	3.92
Krypton	4	-0.48			4	-0.48	-0.21
MeOCF2CHCl2 (Methoxyflurane)	1, 2, 4	2.90			1, 2, 4	2.90	2.74
Methanol	1, 2	2.36	1	2.29	1, 2	2.33	1.93
Methyl acetate			1	2.00	1	2.00	1.99
Methylchloride (Chloromethane)			1, 3	1.13	1, 3	1.13	1.16
Methylcyclopentane	1, 2, 4	2.25			1, 2, 4	2.25	2.19

Methylcyclohexane			6,	2.65	6,	2.65	2.47
m-Methylstyrene (1-vinyl-3-methylbenzene)			1, 3	4.08	1, 3	4.08	3.78
m-Xylene (1,3-Dimethylbenzene)	1	3.28	1, 3	3.30	1, 3	3.29	3.26
Nitrogen	4	-1.27			4	-1.27	-0.78
Nitrous Oxide	2, 4	0.04			2, 4	0.04	0.36
Nonane			6, 18	2.92	6, 18	2.92	3.06
Octane	1, 4	2.71	6,	2.73	1, 4, 6	2.72	2.68
o-Xylene (1,2-dimethylbenzene)	1	3.39	1, 3, 6	3.00	1, 3, 6	3.20	3.37
Pentachloroethane			1, 3	3.61	1, 3	3.61	3.96
Pentane	1, 2, 4	1.60			1, 2, 4	1.60	1.55
Pentyl acetate			1	3.57	1	3.57	3.39
p-Methylstyrene (1-vinyl-4-methylbenzene)			1, 3	4.05	1, 3	4.05	3.79
Propyl acetate			1	2.71	1	2.71	2.63
Propyl bromide (1-Bromopropane)			1, 3	2.37	1, 3	2.37	2.24
Propene (Propylene)	10,	0.71	10,	0.80	10,	0.76	0.74
p-Xylene (1,4-dimethylbenzene)	1	3.31	1, 3	3.24	1, 3	3.28	3.26
Radon			12,	0.68	12,	0.68	0.60
Styrene (Vinyl benzene)	1	3.50	1, 3	3.54	1, 3	3.52	3.38
t-Butylcyclohexane			6,	3.17	6,	3.17	3.49
t-Butylbenzene			6,	3.26	6,	3.26	3.68
Tetrachloroethene (Perchloroethylene)			1, 3	3.21	1, 3	3.21	2.96
Toluene (methylbenzene)	1, 2, 4	2.95	1, 3, 6	2.89	1, 2, 3, 4, 6	2.92	2.87
trans-1,2-Dichloroethylene			1, 3	2.17	1, 3	2.17	2.14
Trichloroethene (Trichloroethylene)	1, 2, 4	2.77	1, 3, 19	2.72	1, 2, 3, 4, 19	2.75	2.62
Vinyl bromide (Bromoethene)			1, 3	1.69	1, 3	1.69	1.74
Vinyl chloride (Chloroethene)	22,	1.30	1, 3, 22	1.30	1, 3, 22	1.30	1.30
Vinylidene fluoride (1,1-Difluoroethylene)			14,	-0.10	14,	-0.10	0.33

Xenon	2	0.26			2	0.26	0.23

Ref = Literature reference

Descriptors E, S, A, B and L for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 18.

Table 14.3. Blood to fat distribution for VOCs

SOLUTE NAME	V	FAT AVERAGE LOG K (H&R)	BLOOD AVERAGE LOG K (H&R)	LOG P (BLOOD TO FAT)	CALC LOG P
(CF3)2CHOCH2F (Sevoflurane)	0.8548	1.62	-0.20	1.82	1.29
1,1,1,2-Tetrachloroethane	0.8800	3.33	1.55	1.78	1.50
1,1,1-Trichloroethane (Methyl chloroform)	0.7576	2.41	0.63	1.78	1.45
1,1,2,2-Tetrachloroethane	0.8800	3.58	2.13	1.45	1.31
1,1,2-Trichloroethane	0.7576	3.16	1.67	1.49	1.15
1,1-Dichloroethane	0.6352	2.21	0.88	1.33	1.08
1,1-Dichloroethylene	0.5922	1.84	0.70	1.14	1.29
1,1-Dichloro-1-Fluoroethane (HCFC-141b)	0.6529	1.56	0.32	1.24	1.34
1,2,4-Trimethylbenzene (Pseudocumene)	1.1391	3.25	1.47	1.78	1.82
1,2,4-Trimethylcyclohexane	1.2681	3.19	0.87	2.32	2.43
1,2-Dibromoethane	0.7404	3.09	2.08	1.01	1.10
1,2-Dichloroethane	0.6352	2.54	1.39	1.15	1.05
1,2-Dichloropropane	0.7761	2.70	1.14	1.56	1.32
1,2-Dimethylcyclohexane	1.1270	3.24	0.91	2.33	2.21
1,3-Butadiene	0.5862	1.35	0.09	1.26	1.18
1,2-Epoxy-3-butene (BMO)	0.5793	2.23	1.97	0.26	0.58
1-Bromo-2-chloroethane (1-Chloro-2-Bromoethane)	0.6878	2.98	1.60	1.38	1.18
1-Butanol	0.7309	2.95	3.08	-0.13	-0.03
1-Chloropropane (Propyl chloride)	0.6537	2.07	0.59	1.48	1.26
1-Nitropropane	0.7055	2.70	2.31	0.39	0.83
1-Pentanol	0.8718	3.41	2.83	0.58	0.18
1-Propanol	0.5900	2.54	3.06	-0.52	-0.25
1-Octene	1.1928	2.89	1.07	1.82	2.15
1-Nonene	1.3337	3.24	1.18	2.06	2.43
1-Decene	1.4746	3.36	1.21	2.15	2.57
2,2,4-Trimethylpentane (Isooctane)	1.2358	2.47	0.23	2.24	2.37
2,2-Dimethylbutane	0.9540	1.82	-0.59	2.41	1.94
2,2-Dichloro-1,1,1-Trifluroethane (HCFC-123)	0.6883	1.80	0.61	1.19	1.16
2,3,4-Trimethylpentane	1.2358	2.65	0.57	2.08	2.37

2-Chloropropane	0.6537	1.84	0.32	1.52	1.21
2-Heptanone (Methyl pentyl ketone)	1.1106	3.66	2.33	1.33	1.01
2-Methyl-1,3-butadiene (Isoprene)	0.7271	1.84	0.10	1.74	1.39
2-Methyl-1-propanol (Isobutanol)	0.7309	2.73	2.92	-0.19	-0.03
2-Methylpentane (Isohexane)	0.9540	1.94	-0.39	2.33	1.94
2-Methyl heptane	1.2358	2.30	0.49	1.81	2.37
2-Methyl octane	1.3767	2.56	0.52	2.04	2.58
2-Methyl nonane	1.5176	2.84	0.76	2.08	2.79
2-Pentanone (Methyl propyl ketone)	0.8288	2.57	2.14	0.43	0.59
2-Propanol (Isopropanol)	0.5900	2.35	3.02	-0.67	-0.36
3-Methyl-1-butanol (Isopentanol)	0.8718	3.18	2.75	0.43	0.18
3-Methylhexane	1.0949	2.44	0.11	2.33	2.16
3-Methylpentane	0.9540	2.01	-0.37	2.38	1.94
4-Methyl-2-pentanone (Methyl isobutyl ketone)	0.9697	2.72	1.96	0.76	0.80
4-Chlorobenzotrifluoride	1.0328	2.99	1.43	1.56	2.04
Acetone (Propan-2-one)	0.5470	1.91	2.36	-0.45	0.15
Allyl chloride (3-Chloro-1-propylene)	0.6106	2.00	1.24	0.76	1.30
Benzene	0.7164	2.64	1.05	1.59	1.29
Bromochloromethane (Chlorobromomethane)	0.5469	2.51	1.21	1.30	1.17
Bromodichloromethane	0.6693	2.72	1.49	1.23	1.27
Butan-2-one (Methyl ethyl ketone)	0.6879	2.26	2.24	0.02	0.37
Butyl acetate	1.0284	3.18	1.94	1.24	1.02
Carbon tetrachloride (Tetrachloromethane)	0.7391	2.56	0.62	1.94	1.63
CF2=CHCl (1-Chloro-2,2-difluroethylene) (1,1-Difluro-2-chloroethene)	0.5052	1.30	0.06	1.24	0.98
CF3CBrClH (Halothane)	0.7410	2.27	0.57	1.70	1.26
CF3CBrFH (Teflurane)	0.6360	1.30	-0.22	1.52	1.09
CF3CH2Cl (2chloro1 1 1trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	0.5659	1.43	0.14	1.29	1.10
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	0.7410	1.53	0.15	1.38	1.02
CHF2CF2CH2Br (Halopropane)	0.7771	2.35	0.75	1.60	1.34
CHF2OCF2CHFCl (Enflurane)	0.8009	1.97	0.35	1.62	1.19
CHF2OCHClCF3 (Isoflurane)	0.8010	1.92	0.20	1.72	1.28
Chlorobenzene	0.8388	3.11	1.63	1.48	1.63

Chlorodibromomethane	0.7219	3.28	1.88	1.40	1.32
Chloroethane	0.5128	1.59	0.49	1.10	1.05
Chloroform (Trichloromethane)	0.6167	2.38	1.15	1.23	1.16
cis-1,2-Dichloroethylene	0.5922	2.36	1.16	1.20	1.11
Cyclohexane	0.8454	2.44	0.17	2.27	1.80
Cyclopropane	0.4227	0.99	-0.21	1.20	1.16
Cyanoethylene Oxide	0.4952	3.11	3.22	-0.11	0.06
Dibromomethane	0.5995	2.90	1.87	1.03	1.09
Dichloromethane (Methylene chloride)	0.4943	2.01	1.12	0.89	0.98
Diethyl ether	0.7309	1.74	1.11	0.63	0.60
Difluromethane (CF2H2)	0.2849	0.16	0.20	-0.04	0.67
Divinyl ether	0.6449	1.60	0.41	1.19	1.18
Ethanol	0.4491	2.34	3.27	-0.93	-0.46
Ethyl acetate	0.7466	2.18	1.90	0.28	0.60
Ethyl t-butyl ether (ETBE)	1.0127	2.13	1.07	1.06	0.69
t-butanol (tert-butanol)(TBA) (2-Methyl-2-propanol)	0.7309	2.20	2.68	-0.48	-0.20
Methyl tertiary-butyl ether (MTBE)	0.8718	1.97	1.18	0.79	0.50
tertiary-Amyl methyl ether (TAME)	1.0127	2.52	1.22	1.30	0.69
tertiary-amyl alcohol (TAA)	0.8718	2.66	2.59	0.07	0.01
Ethylbenzene	0.9982	3.22	1.47	1.75	1.69
Ethylene Oxide	0.3405	1.64	1.80	-0.17	0.17
Ethylene (Ethene)	0.3474	0.20	-0.53	0.73	0.88
Flurochloromethane	0.5300	1.19	0.71	0.48	1.07
Furan	0.5363	1.81	0.82	0.99	1.02
Heptane	1.0949	2.51	0.50	2.01	2.16
Hexachloroethane (Perchloroethane)	1.1248	3.52	1.76	1.76	2.20
Hexane	0.9540	2.12	0.21	1.91	1.94
Isobutyl acetate	1.0284	3.05	1.69	1.36	0.98
Isopentylacetate (Aceticacidpentylester)	1.1693	3.44	1.79	1.65	1.19
Isopropyl acetate	0.8875	2.48	1.55	0.93	0.76
Isopropyl bromide (2-Brompropane)	0.7063	2.20	0.64	1.56	1.26
JP-10	1.1918	4.01	1.76	2.25	2.18
Krypton	0.2460	-0.48	-1.22	0.74	0.88
MeOCF2CHCl2 (Methoxyflurane)	0.8700	2.90	1.28	1.62	1.35
Methanol	0.3082	2.33	3.41	-1.08	-0.74

Methyl acetate	0.6057	2.00	1.98	0.02	0.39
Methylchloride (Chloromethane)	0.3719	1.13	0.31	0.82	0.88
Methylcyclopentane	0.8454	2.25	-0.07	2.32	1.79
Methylcyclohexane	0.9863	2.65	0.70	1.95	2.01
m-Methylstyrene (1-vinyl-3-methylbenzene)	1.0961	4.08	2.28	1.80	1.78
m-Xylene (1,3-Dimethylbenzene)	0.9982	3.29	1.59	1.70	1.67
Nitrogen	0.2222	-1.27	-1.83	0.56	0.84
Nitrous Oxide	0.2810	0.04	-0.34	0.38	0.69
Octane	1.2358	2.72	0.68	2.04	2.37
o-Xylene (1,2-dimethylbenzene)	0.9982	3.20	1.42	1.78	1.67
Pentachloroethane	1.0024	3.61	2.02	1.59	1.62
Pentane	0.8131	1.60	-0.29	1.89	1.73
Pentyl acetate	1.1693	3.57	1.98	1.59	1.23
p-Methylstyrene (1-vinyl-4-methylbenzene)	1.0961	4.05	2.37	1.68	1.78
Propyl acetate	0.8875	2.71	1.88	0.83	0.81
Propyl bromide (1-Bromopropane)	0.7063	2.37	0.97	1.40	1.31
Propene (Propylene)	0.4883	0.76	-0.21	0.97	1.09
p-Xylene (1,4-dimethylbenzene)	0.9982	3.28	1.61	1.67	1.67
Radon	0.3840	0.68	-0.39	1.07	1.09
Styrene (Vinyl benzene)	0.9552	3.52	1.67	1.85	1.61
t-Butylcyclohexane	1.4090	3.17	1.16	2.01	2.42
t-Butylbenzene	1.2800	3.26	1.24	2.02	2.05
Tetrachloroethene (Perchloroethylene)	0.8370	3.21	1.19	2.02	1.79
Toluene (methylbenzene)	0.8573	2.92	1.14	1.78	1.50
trans-1,2-Dichloroethylene	0.5922	2.17	0.88	1.29	1.15
Trichloroethene (Trichloroethylene)	0.7146	2.75	1.14	1.61	1.41
Vinyl bromide (Bromoethene)	0.5224	1.69	0.49	1.20	1.15
Vinyl chloride (Chloroethene)	0.4698	1.30	0.17	1.13	1.10
Vinylidene fluoride (1,1-Difluoroethylene)	0.3828	-0.10	-0.74	0.64	0.96
Xenon	0.3290	0.26	-0.85	1.11	1.00
OUTLIERS					
Decane	1.5176	3.05	1.47	1.58	2.70
2-Nitropropane	0.7055	2.19	2.23	-0.04	0.78
Nonane	1.3767	2.92	1.17	1.75	2.58

Ref = Literature reference

Descriptors E, S, A and B for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 18.

Table 15.0. Data set used to obtain a model to correlate air to liver distributions for VOCs in humans and rats combined

SOLUTE NAME	LIVER REF	LIVER HUM LOG K	LIVER REF	LIVER RAT LOG K	LIVER REF (H&R)	LIVER AVERAGE LOG K (H&R)	CALC LOG K
(CF3)2CHOCH2F (Sevoflurane)	1, 2, 4	0.43			1, 2, 4	0.43	0.71
1,1,1,2-Tetrachloroethane			1, 3	1.95	1, 3	1.95	2.08
1,1,1-Trichloroethane (Methyl chloroform)	1, 2, 4	1.21	1, 3	0.93	1, 2, 3, 4	1.07	1.14
1,1,2,2-Tetrachloroethane			1, 3	2.29	1, 3	2.29	2.53
1,1,2-Trichloroethane			1, 3	1.86	1, 3	1.86	2.11
1,1-Dichloroethane			1, 3	1.03	1, 3	1.03	1.27
1,1-Dichloroethylene			1, 3	0.65	1, 3	0.65	0.65
1,1-Dichloro-1-Fluoroethane (HCFC-141b)			15,	0.52	15,	0.52	0.63
1,2,4-Trimethylcyclohexane			6, 18	1.72	6, 18	1.72	1.55
1,2-Dibromoethane			1, 3	2.08	1, 3	2.08	2.24
1,2-Dichloroethane			1, 3	1.55	1, 3	1.55	1.55
1,2-Dichloropropane			1, 3	1.39	1, 3	1.39	1.54
1,2-Dimethylcyclohexane			6,	2.08	6,	2.08	1.40
1,3-Butadiene	8,	-0.17			8,	-0.17	0.36
1,2-Epoxy-3-butene (BMO)			23,	1.74	23,	1.74	1.48
1-Bromo-2-chloroethane (1-Chloro-2-Bromoethane)			1, 3	1.63	1, 3	1.63	1.79
1-Butanol			1	3.10	1	3.10	2.92

1-Chloropropane (Propyl chloride)			1, 3	0.71	1, 3	0.71	0.85
1-Nitropropane			1, 3	2.18	1, 3	2.18	2.09
1-Pentanol			1	3.24	1	3.24	3.20
1-Propanol			1	3.11	1	3.11	2.60
1-Octene			5,	1.87	5,	1.87	1.29
1-Nonene			5,	2.11	5,	2.11	1.62
1-Decene			5,	2.27	5,	2.27	1.82
2,2,4-Trimethylpentane (Isooctane)			1, 3	1.03	1, 3	1.03	0.82
2,2-Dimethylbutane	1, 2, 4	0.54			1, 2, 4	0.54	0.40
2,2-Dichloro-1,1,1-Trifluroethane (HCFC-123)			16,	0.67	16,	0.67	0.98
2,3,4-Trimethylpentane			1, 3	1.27	1, 3	1.27	1.03
2-Chloropropane			1, 3	0.50	1, 3	0.50	0.72
2-Heptanone (Methyl pentyl ketone)			1	2.63	1	2.63	2.77
2-Methyl-1,3-butadiene (Isoprene)			1, 3, 17	0.42	1, 3, 17	0.42	0.67
2-Methyl-1-propanol (Isobutanol)			1	2.94	1	2.94	2.79
2-Methylpentane (Isohexane)	1, 2, 4	0.65			1, 2, 4	0.65	0.48
2-Methyl heptane			5,	1.01	5,	1.01	1.03
2-Methyl octane			5,	1.26	5,	1.26	1.29
2-Methyl nonane			5,	1.52	5,	1.52	1.56
2-Nitropropane			1, 3	1.80	1, 3	1.80	1.92
2-Pentanone (Methyl propyl ketone)			1	2.27	1	2.27	2.22
2-Propanol (Isopropanol)			1	2.99	1	2.99	2.46
3-Methyl-1-butanol (Isopentanol)			1	2.97	1	2.97	3.12

3-Methylhexane	1, 2, 4	1.04			1, 2, 4	1.04	0.78
3-Methylpentane	1, 2, 4	0.69			1, 2, 4	0.69	0.53
4-Methyl-2-pentanone (Methyl isobutyl ketone)			1	2.19	1	2.19	2.38
4-Chlorobenzotrifluoride			13,	1.71	13,	1.71	1.66
Acetone (Propan-2-one)			1	2.38	1	2.38	1.72
Benzene	1, 2, 4	1.36	1, 3, 6	1.15	1, 2, 3, 4, 6	1.26	1.37
Bromochloromethane (Chlorobromomethane)			1, 3	1.47	1, 3	1.47	1.25
Bromodichloromethane			22,	1.49	22,	1.49	1.63
Butan-2-one (Methyl ethyl ketone)			1	2.36	1	2.36	1.97
Butyl acetate			1	2.45	1	2.45	2.36
Carbon Disulfide	4	0.78			4	0.78	0.73
Carbon tetrachloride (Tetrachloromethane)			1, 3	1.15	1, 3	1.15	0.98
CF2=CHCl (1-Chloro-2,2-difluroethylene) (1,1-Difluro-2-chloroethene)	2, 4	0.04			2, 4	0.04	0.12
CF3CBrClH (Halothane)	1, 2, 4	0.86	1, 3, 16	0.86	1, 2, 3, 4, 16	0.86	1.13
CF3CBrFH (Teflurane)	1, 4	0.01			1, 4	0.01	0.62
CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	1, 2, 4	0.36	1, 3	0.26	1, 2, 3, 4	0.31	0.48
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	1, 2, 4	0.33			1, 2, 4	0.33	0.79
CHF2OCF2CHFCl (Enflurane)	1, 2, 4	0.62			1, 2, 4	0.62	0.97

CHF2OCHClCF3 (Isoflurane)	1, 2, 4	0.55	1, 3	0.57	1, 2, 3, 4	0.56	0.83
Chlorobenzene			1, 3	1.94	1, 3	1.94	1.81
Chlorodibromomethane			1, 3	2.10	1, 3	2.10	1.88
Chloroethane			1, 3	0.56	1, 3	0.56	0.56
Chloroform (Trichloromethane)	1, 2, 4	1.23	1, 3	1.32	1, 2, 3, 4	1.28	1.34
cis-1,2-Dichloroethylene			1, 3	1.18	1, 3	1.18	1.31
Cyclohexane	1, 2, 4	1.03	1, 3, 6	1.06	1, 2, 3, 4, 6	1.05	0.84
Cyclopropane	1, 2, 4	-0.19			1, 2, 4	-0.19	0.03
Cyanoethylene Oxide			7,	2.66	7,	2.66	2.38
Decane			6,	1.64	6,	1.64	1.69
Dibromomethane			1, 3	1.83	1, 3	1.83	1.73
Dichloromethane (Methylene chloride)	1, 2, 4	0.86	1, 3	1.15	1, 2, 3, 4	1.01	1.06
Diethyl ether	1, 2, 4	1.05	1, 3	0.83	1, 2, 3, 4	0.94	1.34
Difluromethane (CF2H2)			1, 3	0.44	1, 3	0.44	-0.25
Divinyl ether	1, 2, 4	0.48			1, 2, 4	0.48	0.66
Ethyl acetate			1	2.03	1	2.03	1.80
Ethyl t-butyl ether (ETBE)			9,	1.52	9,	1.52	1.97
t-butanol (tert-butanol)(TBA) (2-Methyl-2-propanol)			9, 24	2.69	9, 24	2.69	2.55
Methyl tertiary-butyl ether (MTBE)			9, 24	1.35	9, 24	1.35	1.81
tertiary-Amyl methyl ether (TAME)			9,	1.50	9,	1.50	2.12
tertiary-amyl alcohol (TAA)			9,	2.68	9,	2.68	2.92
Ethylbenzene			1	1.78	1	1.78	1.93

Ethylene Oxide	20,	1.73			20,	1.73	1.30
Ethylene (Ethene)	20,	-0.35	20,	-0.24	20,	-0.30	-0.51
Flurochloromethane			1, 3	0.54	1, 3	0.54	0.19
Furan			11,	0.77	11,	0.77	0.84
Heptane	1, 2, 4	1.04	1, 3, 6	0.87	1, 2, 3, 4, 6	0.96	0.86
Hexachloroethane (Perchloroethane)			1, 3	2.57	1, 3	2.57	2.28
Hexane	1, 2, 4	0.72	1, 3, 6	0.65	1, 2, 3, 4, 6	0.69	0.58
Isobutyl acetate			1	2.42	1	2.42	2.27
Isopentylacetate (Aceticacidpentylester)			1	2.55	1	2.55	2.59
Isopropyl acetate			1	2.17	1	2.17	1.93
Isopropyl bromide (2-Brompropane)			1, 3	0.64	1, 3	0.64	1.00
JP-10			3,	2.74	3,	2.74	2.29
Krypton	4	-1.14			4	-1.14	-1.02
MeOCF2CHCl2 (Methoxyflurane)	1, 2, 4	1.46	1	1.47	1, 2, 4	1.47	1.69
Methyl acetate			1	1.95	1	1.95	1.59
Methylchloride (Chloromethane)			1, 3	0.54	1, 3	0.54	0.25
Methylcyclopentane	1, 2, 4	0.89			1, 2, 4	0.89	0.80
Methylcyclohexane			6,	1.50	6,	1.50	1.00
m-Methylstyrene (1-vinyl-3-methylbenzene)			1, 3	2.51	1, 3	2.51	2.45
m-Xylene (1,3-Dimethylbenzene)			1, 3	1.96	1, 3	1.96	1.99

Nitrous Oxide	2, 4	-0.38			2, 4	-0.38	-0.33
Nonane			6, 18	1.14	6, 18	1.14	1.41
Octane	1, 4	1.41	6,	0.99	1, 4, 6	1.20	1.13
o-Xylene (1,2-dimethylbenzene)			1, 3, 6	1.76	1, 3, 6	1.76	2.08
Pentachloroethane			1, 3	2.41	1, 3	2.41	2.61
Pentane	1, 2, 4	0.32			1, 2, 4	0.32	0.30
Pentyl acetate			1	2.64	1	2.64	2.63
p-Methylstyrene (1-vinyl-4-methylbenzene)			1, 3	2.51	1, 3	2.51	2.46
Propyl acetate			1	2.36	1	2.36	2.06
Propyl bromide (1-Bromopropane)			1, 3	0.91	1, 3	0.91	1.13
Propene (Propylene)	10,	-0.27	10,	-0.29	10,	-0.28	-0.17
p-Xylene (1,4-dimethylbenzene)			1, 3	1.95	1, 3	1.95	1.99
Radon			12,	-0.51	12,	-0.51	-0.42
Styrene (Vinyl benzene)			1, 3	2.14	1, 3	2.14	2.12
t-Butylcyclohexane			6,	1.46	6,	1.46	1.95
t-Butylbenzene			6,	1.73	6,	1.73	2.33
Tetrachloroethene (Perchloroethylene)			1, 3	1.85	1, 3	1.85	1.46
Toluene (methylbenzene)	1, 2, 4	1.68	1, 3, 6	1.60	1, 2, 3, 4, 6	1.64	1.67
trans-1,2-Dichloroethylene			1, 3	0.95	1, 3	0.95	1.06
Trichloroethene (Trichloroethylene)	1, 2, 4	1.41	1, 3, 19	1.40	1, 2, 3, 4, 19	1.41	1.36
Vinyl bromide (Bromoethene)			1, 3	0.52	1, 3	0.52	0.68
Vinyl chloride (Chloroethene)	21,	0.20	1, 3, 21	0.20	1, 3, 21	0.20	0.29

Vinylidene fluoride (1,1-Difluoroethylene)			14,	0.00	14,	0.00	-0.51
Xenon	2	-1.00			2	-1.00	-0.69
OUTLIERS							
Methanol			1	3.49	1	3.49	2.45
Ethanol			1	3.24	1	3.24	2.39
Allyl chloride (3-Chloro-1-propylene)			1, 3	1.59	1, 3	1.59	0.82
1,2,4-Trimethylbenzene (Pseudocumene)			6, 18	1.67	6, 18	1.67	2.42

[a]Calculated values for log K, obtained when correlating with the Abraham descriptors (**E**, **S**, **A**, **B** and **L**)

[b]REF= Literature reference

Descriptors E, S, A, B and L for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 18.

Table 15.3. Data set used to obtain a model to correlate blood to liver distributions for VOCs in humans and rats combined

SOLUTE NAME	V	LIVER (AVE) LOG K (H&R)	BLOOD LOG K	LIVER LOG P (H&R)	CALC LOG P
(CF3)2CHOCH2F (Sevoflurane)	0.8548	0.43	-0.20	0.63	0.43
1,1,1,2-Tetrachloroethane	0.8800	1.95	1.55	0.40	0.31
1,1,1-Trichloroethane (Methyl chloroform)	0.7576	1.07	0.63	0.44	0.29
1,1,2,2-Tetrachloroethane	0.8800	2.29	2.13	0.16	0.24
1,1,2-Trichloroethane	0.7576	1.86	1.67	0.19	0.16
1,1-Dichloroethane	0.6352	1.03	0.88	0.15	0.14
1,1-Dichloroethylene	0.5922	0.65	0.70	-0.05	0.18
1,1-Dichloro-1-Fluoroethane (HCFC-141b)	0.6529	0.52	0.32	0.20	0.22
1,2,4-Trimethylcyclohexane	1.2681	1.72	0.87	0.85	0.83
1,2-Dibromoethane	0.7404	2.08	2.08	0.00	0.10
1,2-Dichloroethane	0.6352	1.55	1.39	0.16	0.09
1,2-Dichloropropane	0.7761	1.39	1.14	0.25	0.21
1,2-Dimethylcyclohexane	1.1270	2.08	0.91	1.17	0.70
1,3-Butadiene	0.5862	-0.17	0.09	-0.26	0.19
1,2-Epoxy-3-butene (BMO)	0.5793	1.74	1.97	-0.23	0.02
1-Bromo-2-chloroethane (1-Chloro-2-Bromoethane)	0.6878	1.63	1.60	0.03	0.11
1-Butanol	0.7309	3.10	3.08	0.02	0.05
1-Chloropropane (Propyl chloride)	0.6537	0.71	0.59	0.12	0.21
1-Nitropropane	0.7055	2.18	2.31	-0.13	0.01
1-Pentanol	0.8718	3.24	2.83	0.41	0.18
1-Propanol	0.5900	3.11	3.06	0.05	-0.07
1-Octene	1.1928	1.87	1.07	0.80	0.79
1-Nonene	1.3337	2.11	1.18	0.93	0.87
1-Decene	1.4746	2.27	1.21	1.06	1.04
2,2,4-Trimethylpentane (Isooctane)	1.2358	1.03	0.23	0.80	0.88
2,2-Dimethylbutane	0.9540	0.54	-0.59	1.13	0.63

2,2-Dichloro-1,1,1-Trifluroethane (HCFC-123)	0.6883	0.67	0.61	0.06	0.24
2,3,4-Trimethylpentane	1.2358	1.27	0.57	0.70	0.88
2-Chloropropane	0.6537	0.50	0.32	0.18	0.22
2-Heptanone (Methyl pentyl ketone)	1.1106	2.63	2.33	0.30	0.38
2-Methyl-1,3-butadiene (Isoprene)	0.7271	0.42	0.10	0.32	0.32
2-Methyl-1-propanol (Isobutanol)	0.7309	2.94	2.92	0.02	0.06
2-Methylpentane (Isohexane)	0.9540	0.65	-0.39	1.04	0.63
2-Methyl heptane	1.2358	1.01	0.49	0.52	0.88
2-Methyl octane	1.3767	1.26	0.52	0.74	1.00
2-Methyl nonane	1.5176	1.52	0.76	0.76	1.13
2-Nitropropane	0.7055	1.80	2.23	-0.43	0.02
2-Pentanone (Methyl propyl ketone)	0.8288	2.27	2.14	0.13	0.13
2-Propanol (Isopropanol)	0.5900	2.99	3.02	-0.03	-0.07
3-Methyl-1-butanol (Isopentanol)	0.8718	2.97	2.75	0.22	0.19
3-Methylhexane	1.0949	1.04	0.11	0.93	0.76
3-Methylpentane	0.9540	0.69	-0.37	1.06	0.63
4-Methyl-2-pentanone (Methyl isobutyl ketone)	0.9697	2.19	1.96	0.23	0.27
4-Chlorobenzotrifluoride	1.0328	1.71	1.43	0.28	0.50
Acetone (Propan-2-one)	0.5470	2.38	2.36	0.02	-0.12
Benzene	0.7164	1.26	1.05	0.21	0.19
Bromochloromethane (Chlorobromomethane)	0.5469	1.47	1.21	0.26	-0.01
Bromodichloromethane	0.6693	1.49	1.49	0.00	0.12
Butan-2-one (Methyl ethyl ketone)	0.6879	2.36	2.24	0.12	0.00
Butyl acetate	1.0284	2.45	1.94	0.51	0.36
Carbon Disulfide	0.4905	0.78	0.30	0.48	0.10
Carbon tetrachloride (Tetrachloromethane)	0.7391	1.15	0.62	0.53	0.31

CF2=CHCl (1-Chloro-2,2-difluroethylene) (1,1-Difluro-2-chloroethene)	0.5052	0.04	0.06	-0.02	0.14
CF3CBrClH (Halothane)	0.7410	0.86	0.57	0.29	0.28
CF3CBrFH (Teflurane)	0.6360	0.01	-0.22	0.23	0.25
CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	0.5659	0.31	0.14	0.17	0.14
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	0.7410	0.33	0.15	0.18	0.25
CHF2OCF2CHFCl (Enflurane)	0.8009	0.62	0.35	0.27	0.32
CHF2OCHClCF3 (Isoflurane)	0.8010	0.56	0.20	0.36	0.31
Chlorobenzene	0.8388	1.94	1.63	0.31	0.28
Chlorodibromomethane	0.7219	2.10	1.88	0.22	0.14
Chloroethane	0.5128	0.56	0.49	0.07	0.08
Chloroform (Trichloromethane)	0.6167	1.28	1.15	0.13	0.13
cis-1,2-Dichloroethylene	0.5922	1.18	1.16	0.02	0.09
Cyclohexane	0.8454	1.05	0.17	0.88	0.49
Cyclopropane	0.4227	-0.19	-0.21	0.02	0.08
Cyanoethylene Oxide	0.4952	2.66	3.22	-0.56	-0.27
Dibromomethane	0.5995	1.83	1.87	-0.04	0.04
Dichloromethane (Methylene chloride)	0.4943	1.01	1.12	-0.11	0.01
Diethyl ether	0.7309	0.94	1.11	-0.17	0.20
Difluromethane (CF2H2)	0.2849	0.44	0.20	0.24	-0.11
Divinyl ether	0.6449	0.48	0.41	0.07	0.19
Ethyl acetate	0.7466	2.03	1.90	0.13	0.10
Ethyl t-butyl ether (ETBE)	1.0127	1.52	1.07	0.45	0.42
t-butanol (tert-butanol)(TBA) (2-Methyl-2-propanol)	0.7309	2.69	2.68	0.01	0.06
Methyl tertiary-butyl ether (MTBE)	0.8718	1.35	1.18	0.17	0.29
tertiary-Amyl methyl ether (TAME)	1.0127	1.50	1.22	0.28	0.41
tertiary-amyl alcohol (TAA)	0.8718	2.68	2.59	0.09	0.18

Ethylbenzene	0.9982	1.78	1.47	0.31	0.44
Ethylene Oxide	0.3405	1.73	1.80	-0.07	-0.26
Ethylene (Ethene)	0.3474	-0.30	-0.53	0.24	0.04
Flurochloromethane	0.5300	0.54	0.71	-0.17	0.16
Furan	0.5363	0.77	0.82	-0.05	0.05
Heptane	1.0949	0.96	0.50	0.46	0.76
Hexachloroethane (Perchloroethane)	1.1248	2.57	1.76	0.81	0.55
Hexane	0.9540	0.69	0.21	0.48	0.63
Isobutyl acetate	1.0284	2.42	1.69	0.73	0.36
Isopentylacetate (Aceticacidpentylester)	1.1693	2.55	1.79	0.76	0.48
Isopropyl acetate	0.8875	2.17	1.55	0.62	0.24
Isopropyl bromide (2-Brompropane)	0.7063	0.64	0.64	0.00	0.25
JP-10	1.1918	2.74	1.76	0.98	0.66
Krypton	0.2460	-1.14	-1.22	0.08	0.01
MeOCF2CHCl2 (Methoxyflurane)	0.8700	1.47	1.28	0.19	0.29
Methyl acetate	0.6057	1.95	1.98	-0.03	-0.03
Methylchloride (Chloromethane)	0.3719	0.54	0.31	0.23	-0.04
Methylcyclopentane	0.8454	0.89	-0.07	0.96	0.50
Methylcyclohexane	0.9863	1.50	0.70	0.80	0.63
m-Methylstyrene (1-vinyl-3-methylbenzene)	1.0961	2.51	2.28	0.23	0.46
m-Xylene (1,3-Dimethylbenzene)	0.9982	1.96	1.59	0.37	0.43
Nitrous Oxide	0.2810	-0.38	-0.34	-0.04	-0.10
Octane	1.2358	1.20	0.68	0.52	0.88
o-Xylene (1,2-dimethylbenzene)	0.9982	1.76	1.42	0.34	0.42
Pentachloroethane	1.0024	2.41	2.02	0.39	0.39
Pentane	0.8131	0.32	-0.29	0.61	0.51
Pentyl acetate	1.1693	2.64	1.98	0.66	0.48
p-Methylstyrene (1-vinyl-4-methylbenzene)	1.0961	2.51	2.37	0.14	0.46
Propyl acetate	0.8875	2.36	1.88	0.48	0.23

Propyl bromide (1-Bromopropane)	0.7063	0.91	0.97	-0.06	0.24
Propene (Propylene)	0.4883	-0.28	-0.21	-0.07	0.17
p-Xylene (1,4-dimethylbenzene)	0.9982	1.95	1.61	0.34	0.43
Radon	0.3840	-0.51	-0.39	-0.12	0.13
Styrene (Vinyl benzene)	0.9552	2.14	1.67	0.47	0.35
t-Butylbenzene	1.2800	1.73	1.24	0.49	0.68
Tetrachloroethene (Perchloroethylene)	0.8370	1.85	1.19	0.66	0.37
Toluene (methylbenzene)	0.8573	1.64	1.14	0.50	0.32
trans-1,2-Dichloroethylene	0.5922	0.95	0.88	0.07	0.14
Trichloroethene (Trichloroethylene)	0.7146	1.41	1.14	0.27	0.26
Vinyl bromide (Bromoethene)	0.5224	0.52	0.49	0.03	0.06
Vinyl chloride (Chloroethene)	0.4698	0.20	0.17	0.03	0.07
Xenon	0.3290	-1.00	-0.85	-0.15	0.09
OUTLIERS					
Nonane	1.3767	1.14	1.17	-0.03	0.86
Decane	1.5176	1.64	1.47	0.17	0.95
t-Butylcyclohexane	1.4090	1.46	1.16	0.30	0.91
Vinylidene fluoride (1,1-Difluoroethylene)	0.3828	0.00	-0.74	0.74	0.14

[a]Calculated values for log P, obtained when correlating with the Abraham descriptors (**E**, **S**, **A**, **B** and **V**)

Descriptors E, S, A and B for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 18.

[b]REF= Literature reference

Table 16.0. Data set used to obtain a model to correlate air to muscle distributions for VOCs in humans and rats combined.

SOLUTE NAME	MUS REF	MUSCLE HUM LOG K	MUS REF	MUSCLE RAT LOG K	MUS REF (H&R)	MUSCLE AVERAGE LOG K (H&R)	CALC[a] LOG K
(CF3)2CHOCH2F (Sevoflurane)	1, 2, 4	0.22			1, 2, 4	0.22	0.44
1,1,1,2-Tetrachloroethane			1, 3	1.60	1, 3	1.60	1.74
1,1,1-Trichloroethane (Methyl chloroform)	1, 2, 4	0.83	1, 3	0.50	1, 2, 3, 4	0.67	0.82
1,1,2,2-Tetrachloroethane			1, 3	2.00	1, 3	2.00	2.21
1,1,2-Trichloroethane			1, 3	1.36	1, 3	1.36	1.82
1,1-Dichloroethane			1, 3	0.71	1, 3	0.71	1.03
1,1-Dichloroethylene			1, 3	0.31	1, 3	0.31	0.38
1,1-Dichloro-1-Fluoroethane (HCFC-141b)			11,	0.02	11,	0.02	0.33
1,2-Dibromoethane			1, 3	1.66	1, 3	1.66	1.97
1,2-Dichloroethane			1, 3	1.37	1, 3	1.37	1.30
1,2-Dichloropropane			1, 3	1.08	1, 3	1.08	1.23
1,3-Butadiene	5,	-0.06			5,	-0.06	0.16
1,2-Epoxy-3-butene (BMO)			16,	1.66	16,	1.66	1.31
1-Bromo-2-chloroethane (1-Chloro-2-Bromoethane)			1, 3	1.40	1, 3	1.40	1.51
1-Butanol			1	2.95	1	2.95	2.90
1-Chloropropane (Propyl chloride)			1, 3	0.32	1, 3	0.32	0.56
1-Nitropropane			1, 3	1.46	1, 3	1.46	1.80
1-Pentanol			1	2.91	1	2.91	3.13
1-Propanol	1, 2, 4	2.83	1	3.06	1, 2, 4	2.95	2.64

2,2,4-Trimethylpentane (Isooctane)			1, 3	0.52	1, 3	0.52	0.40
2,2-Dimethylbutane	1, 2, 4	0.00			1, 2, 4	0.00	0.05
2,2-Dichloro-1,1,1-Trifluroethane (HCFC-123)			12,	0.39	12,	0.39	0.74
2,3,4-Trimethylpentane			1, 3	0.64	1, 3	0.64	0.57
2-Chloropropane			1, 3	0.31	1, 3	0.31	0.46
2-Heptanone (Methyl pentyl ketone)			1	2.33	1	2.33	2.48
2-Methyl-1,3-butadiene (Isoprene)			1, 3, 13	0.26	1, 3, 13	0.26	0.41
2-Methyl-1-propanol (Isobutanol)	1, 2, 4	2.54	1	2.93	1, 2, 4	2.74	2.79
2-Methylpentane (Isohexane)	1, 2, 4	0.46			1, 2, 4	0.46	0.12
2-Nitropropane			1, 3	1.46	1, 3	1.46	1.67
2-Pentanone (Methyl propyl ketone)			1	2.00	1	2.00	2.02
2-Propanol (Isopropanol)	1, 2, 4	2.70	1	3.04	1, 2, 4	2.87	2.53
3-Methyl-1-butanol (Isopentanol)			1	2.90	1	2.90	3.06
3-Methylhexane	1, 2, 4	1.04			1, 2, 4	1.04	0.37
3-Methylpentane	1, 2, 4	0.58			1, 2, 4	0.58	0.16
4-Methyl-2-pentanone (Methyl isobutyl ketone)			1	1.81	1	1.81	2.14
Acetone (Propan-2-one)	1, 2, 4	2.18	1	2.23	1, 2, 4	2.21	1.63
Acetylene	4	0.03			4	0.03	-0.21
Allyl chloride (3-Chloro-1-propylene)			1, 3	1.04	1, 3	1.04	0.53
Argon	4	-1.64			4	-1.64	-1.36
Benzene	1, 2, 4	1.21	1, 3	1.01	1, 2, 3, 4	1.11	1.10
Bromochloromethane (Chlorobromomethane)			1, 3	1.05	1, 3	1.05	0.96

Butan-2-one (Methyl ethyl ketone)	1, 2, 4	2.01	1	2.26	1, 2, 4	2.14	1.82
Butyl acetate			1	2.20	1	2.20	2.07
Carbon tetrachloride (Tetrachloromethane)			1, 3	0.66	1, 3	0.66	0.64
CF2=CHCl (1-Chloro-2,2-difluroethylene) (1,1-Difluro-2-chloroethene)	2, 4	0.00			2, 4	0.00	-0.08
CF3CBrClH (Halothane)	1, 2, 4	0.83	1, 3	0.65	1, 2, 3, 4	0.74	0.87
CF3CBrFH (Teflurane)	2, 4	0.34			2, 4	0.34	0.43
CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	1, 2, 4	0.34	1, 3	0.09	1, 2, 3, 4	0.22	0.28
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	1, 2, 4	0.35			1, 2, 4	0.35	0.62
CHF2OCF2CHFCl (Enflurane)	1, 2, 4	0.64			1, 2, 4	0.64	0.72
CHF2OCHClCF3 (Isoflurane)	1, 2, 4	0.41	1, 3	0.20	1, 2, 3, 4	0.31	0.57
Chlorobenzene			1, 3	1.53	1, 3	1.53	1.45
Chlorodibromomethane			1, 3	1.75	1, 3	1.75	1.59
Chloroethane			1, 3	0.51	1, 3	0.51	0.32
Chloroform (Trichloromethane)	1, 2, 4	1.09	1, 3	1.14	1, 2, 3, 4	1.12	1.09
cis-1,2-Dichloroethylene			1, 3	0.78	1, 3	0.78	1.06
Cyclohexane	1, 2, 4	1.02	1, 3	0.01	1, 2, 3, 4	0.52	0.47
Cyclopropane	1, 2, 4	-0.31			1, 2, 4	-0.31	-0.18
Dibromomethane			1, 3	1.61	1, 3	1.61	1.47
Dichloromethane (Methylene chloride)	1, 2, 4	0.68	1, 3	0.90	1, 2, 3, 4	0.79	0.84
Diethyl ether	1, 2, 4	1.01	1, 3	0.72	1, 2, 3, 4	0.87	1.19
Difluromethane (CF2H2)			1, 3	0.16	1, 3	0.16	-0.41

Divinyl ether	1, 2, 4	0.34			1, 2, 4	0.34	0.43
Ethanol	1, 2, 4	2.93	1	3.23	1, 2, 4	3.08	2.39
Ethyl acetate			1	1.84	1	1.84	1.61
Ethyl t-butyl ether (ETBE)			6,	1.35	6,	1.35	1.81
t-butanol (tert-butanol)(TBA) (2-Methyl-2-propanol)			6, 17	2.80	6, 17	2.80	2.61
Methyl tertiary-butyl ether (MTBE)			6,	1.57	6,	1.57	1.68
tertiary-Amyl methyl ether (TAME)			6,	1.38	6,	1.38	1.95
tertiary-amyl alcohol (TAA)			6,	2.59	6,	2.59	2.92
Ethylbenzene			1	1.41	1	1.41	1.58
Ethylene Oxide	15,	1.82	14,	1.68	14, 15	1.75	1.20
Ethylene (Ethene)	15,	-0.19	15,	-0.20	15,	-0.20	-0.64
Flurochloromethane			1, 3	0.39	1, 3	0.39	-0.02
Helium	4	-1.93			4	-1.93	-1.84
Heptane	1, 2, 4	1.08	1, 3	0.62	1, 2, 3, 4	0.85	0.43
Hexachloroethane (Perchloroethane)			1, 3	1.88	1, 3	1.88	1.78
Hexane	1, 2, 4	0.70	1, 3	0.46	1, 2, 3, 4	0.58	0.20
Hydrogen	4	-1.76			4	-1.76	-1.59
Isobutyl acetate			1	2.04	1	2.04	2.01
Isopentylacetate (Aceticacidpentylester)			1	2.32	1	2.32	2.28
Isopropyl acetate			1	1.85	1	1.85	1.72
Isopropyl bromide (2-Brompropane)			1, 3	0.61	1, 3	0.61	0.74
Krypton	4	-1.38			4	-1.38	-1.14
MeOCF2CHCl2 (Methoxyflurane)	1, 2, 4	1.35			1, 2, 4	1.35	1.37

Methane	4	-1.41			4	-1.41	-1.19
Methyl acetate			1	1.81	1	1.81	1.45
Methylchloride (Chloromethane)			1, 3	-0.01	1, 3	-0.01	0.06
Methylcyclopentane	1, 2, 4	0.70			1, 2, 4	0.70	0.43
m-Methylstyrene (1-vinyl-3-methylbenzene)			1, 3	2.26	1, 3	2.26	2.08
m-Xylene (1,3-Dimethylbenzene)			1, 3	1.62	1, 3	1.62	1.64
Nitrous Oxide	2, 4	-0.43			2, 4	-0.43	-0.45
Octane	1, 4	0.93			1, 4	0.93	0.66
o-Xylene (1,2-dimethylbenzene)			1, 3	1.71	1, 3	1.71	1.72
Pentachloroethane			1, 3	1.86	1, 3	1.86	2.25
Pentane	1, 2, 4	-0.15			1, 2, 4	-0.15	-0.04
Pentyl acetate			1	2.36	1	2.36	2.30
p-Methylstyrene (1-vinyl-4-methylbenzene)			1, 3	2.26	1, 3	2.26	2.09
Propyl acetate			1	1.93	1	1.93	1.83
Propyl bromide (1-Bromopropane)			1, 3	0.62	1, 3	0.62	0.84
Propene (Propylene)	7,	-0.19	7,	-0.34	7,	-0.27	-0.35
p-Xylene (1,4-dimethylbenzene)			1, 3	1.58	1, 3	1.58	1.64
Radon			8,	-0.81	8,	-0.81	-0.63
Styrene (Vinyl benzene)			1, 3	1.67	1, 3	1.67	1.79
Sulphur Hexafluoride	4	-1.85			4	-1.85	-1.36
Tetrachloroethene (Perchloroethylene)			1, 3	1.30	1, 3	1.30	1.07
Toluene (Methylbenzene)	1, 2, 4	1.54	1, 3	1.44	1, 2, 3, 4	1.49	1.35

trans-1,2-Dichloroethylene			1, 3	0.55	1, 3	0.55	0.82
Trichloroethene (Trichloroethylene)	1, 2, 4	1.25	1, 3	1.00	1, 2, 3, 4	1.13	1.06
Vinyl bromide (Bromoethene)			1, 3	0.35	1, 3	0.35	0.47
Vinyl chloride (Chloroethene)			1, 3	0.32	1, 3	0.32	0.06
Vinylidene fluoride (1,1-Difluoroethylene)			10,	-0.54	10,	-0.54	-0.67
Xenon	2	-1.00			2	-1.00	-0.86
Difluorochloromethane (CHF2Cl)	4	0.03			4	0.03	0.11
OUTLIERS							
Tricyclo[5.2.1.0]-decane (JP10)			3,	2.83	3,	2.83	1.86
Methanol	1, 2, 4	3.12	1	3.60	1, 2, 4	3.36	2.46
4-Chlorobenzotrifluoride			9,	2.03	9,	2.03	1.25

[a]Calculated values for log K, obtained when correlating with the Abraham descriptors (**E**, **S**, **A**, **B** and **L**) using equation from table 16.0

REF= Literature reference

Descriptors E, S, A, B and L for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 1

Table 16.3. Data set used to obtain a model to correlate air to liver distributions for VOCs in humans and rats combined.

SOLUTE NAME	V	MUSCLE AVERAGE LOG	BLOOD LOG K	BLOOD TO MUSCLE LOG P	CALC LOG P
(CF3)2CHOCH2F (Sevoflurane)	0.8548	0.22	-0.20	0.42	0.38
1,1,1,2-Tetrachloroethane	0.8800	1.60	1.55	0.05	-0.04
1,1,1-Trichloroethane (Methyl chloroform)	0.7576	0.67	0.63	0.04	0.04
1,1,2,2-Tetrachloroethane	0.8800	2.00	2.13	-0.13	-0.14
1,1,2-Trichloroethane	0.7576	1.36	1.67	-0.31	-0.16
1,1-Dichloroethane	0.6352	0.71	0.88	-0.17	-0.10
1,1-Dichloroethylene	0.5922	0.31	0.70	-0.39	-0.03
1,1-Dichloro-1-Fluoroethane (HCFC-141b)	0.6529	0.02	0.32	-0.30	0.02
1,2-Dibromoethane	0.7404	1.66	2.08	-0.42	-0.28
1,2-Dichloroethane	0.6352	1.37	1.39	-0.02	-0.21
1,2-Dichloropropane	0.7761	1.08	1.14	-0.06	-0.12
1,3-Butadiene	0.5862	-0.06	0.09	-0.15	0.03
1,2-Epoxy-3-butene (BMO)	0.5793	1.66	1.97	-0.31	-0.17
1-Bromo-2-chloroethane (1-Chloro-2-Bromoethane)	0.6878	1.40	1.60	-0.20	-0.23
1-Butanol	0.7309	2.95	3.08	-0.13	-0.05
1-Chloropropane (Propyl chloride)	0.6537	0.32	0.59	-0.27	0.00
1-Nitropropane	0.7055	1.46	2.31	-0.85	-0.33
2-Nitropropane	0.7055	1.46	2.23	-0.77	-0.31
3-Methylpentane	0.9540	0.58	-0.37	0.95	0.52
1-Pentanol	0.8718	2.91	2.83	0.08	0.06
1-Propanol	0.5900	2.95	3.06	-0.11	-0.16
2,2-Dimethylbutane	0.9540	0.00	-0.59	0.59	0.52
2,2-Dichloro-1,1,1-Trifluroethane (HCFC-123)	0.6883	0.39	0.61	-0.22	0.10
2-Chloropropane	0.6537	0.31	0.32	-0.01	0.04
2-Heptanone (Methyl pentyl ketone)	1.1106	2.33	2.33	0.00	0.12

2-Methyl-1,3-butadiene (Isoprene)	0.7271	0.26	0.10	0.16	0.14
2-Methyl-1-propanol (Isobutanol)	0.7309	2.74	2.92	-0.18	-0.03
2-Methylpentane (Isohexane)	0.9540	0.46	-0.39	0.85	0.52
2-Pentanone (Methyl propyl ketone)	0.8288	2.00	2.14	-0.14	-0.09
2-Propanol (Isopropanol)	0.5900	2.87	3.02	-0.15	-0.13
3-Methyl-1-butanol (Isopentanol)	0.8718	2.90	2.75	0.15	0.08
3-Methylhexane	1.0949	1.04	0.11	0.93	0.63
4-Methyl-2-pentanone (Methyl isobutyl ketone)	0.9697	1.81	1.96	-0.15	0.04
Acetone (Propan-2-one)	0.5470	2.21	2.36	-0.15	-0.32
Acetylene	0.3044	0.03	-0.06	0.09	-0.37
Allyl chloride (3-Chloro-1-propylene)	0.6106	1.04	1.24	-0.20	-0.14
Argon	0.1900	-1.64	-1.52	-0.12	-0.04
Benzene	0.7164	1.11	1.05	0.06	-0.11
Bromochloromethane (Chlorobromomethane)	0.5469	1.05	1.21	-0.16	-0.38
Butan-2-one (Methyl ethyl ketone)	0.6879	2.14	2.24	-0.10	-0.21
Butyl acetate	1.0284	2.20	1.94	0.26	0.13
Carbon tetrachloride (Tetrachloromethane)	0.7391	0.66	0.62	0.04	0.04
CF2=CHCl (1-Chloro-2,2-difluroethylene) (1,1-Difluro-2-chloroethene)	0.5052	0.00	0.06	-0.06	0.08
CF3CBrClH (Halothane)	0.7410	0.74	0.57	0.17	0.10
CF3CBrFH (Teflurane)	0.6360	0.34	-0.22	0.56	0.16
CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	0.5659	0.22	0.14	0.08	-0.02
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	0.7410	0.35	0.15	0.20	0.10
CHF2OCF2CHFCl (Enflurane)	0.8009	0.64	0.35	0.29	0.19
CHF2OCHClCF3 (Isoflurane)	0.8010	0.31	0.20	0.11	0.14
Chlorobenzene	0.8388	1.53	1.63	-0.10	-0.11

Chlorodibromomethane	0.7219	1.75	1.88	-0.13	-0.25
Chloroethane	0.5128	0.51	0.49	0.02	-0.11
Chloroform (Trichloromethane)	0.6167	1.12	1.15	-0.03	-0.12
cis-1,2-Dichloroethylene	0.5922	0.78	1.16	-0.38	-0.18
Cyclohexane	0.8454	0.52	0.17	0.35	0.32
Cyclopropane	0.4227	-0.31	-0.21	-0.10	-0.09
Dibromomethane	0.5995	1.61	1.87	-0.26	-0.32
Dichloromethane (Methylene chloride)	0.4943	0.79	1.12	-0.33	-0.25
Diethyl ether	0.7309	0.87	1.11	-0.24	0.12
Difluromethane (CF2H2)	0.2849	0.16	0.20	-0.04	-0.21
Divinyl ether	0.6449	0.34	0.41	-0.07	-0.01
Ethanol	0.4491	3.08	3.27	-0.19	-0.26
Ethyl acetate	0.7466	1.84	1.90	-0.06	-0.10
Ethyl t-butyl ether (ETBE)	1.0127	1.35	1.07	0.28	0.37
t-butanol (tert-butanol)(TBA (2-Methyl-2-propanol)	0.7309	2.80	2.68	0.12	0.02
Methyl tertiary-butyl ether (MTBE)	0.8718	1.57	1.18	0.39	0.23
tertiary-Amyl methyl ether (TAME)	1.0127	1.38	1.22	0.16	0.33
tertiary-amyl alcohol (TAA)	0.8718	2.59	2.59	0.00	0.12
Ethylbenzene	0.9982	1.41	1.47	-0.06	0.10
Ethylene Oxide	0.3405	1.75	1.80	-0.05	-0.48
Ethylene (Ethene)	0.3474	-0.20	-0.53	0.34	-0.02
Flurochloromethane	0.5300	0.39	0.71	-0.32	0.05
Helium	0.0680	-1.93	-2.00	0.07	-0.13
Heptane	1.0949	0.85	0.50	0.35	0.63
Hexachloroethane (Perchloroethane)	1.1248	1.88	1.76	0.12	0.10
Hexane	0.9540	0.58	0.21	0.37	0.52
Hydrogen	0.1086	-1.76	-1.77	0.01	-0.10
Isobutyl acetate	1.0284	2.04	1.69	0.35	0.15
Isopentylacetate (Aceticacidpentylester)	1.1693	2.32	1.79	0.53	0.25
Isopropyl acetate	0.8875	1.85	1.55	0.30	0.04
Isopropyl bromide (2-Brompropane)	0.7063	0.61	0.64	-0.03	0.04

Krypton	0.2460	-1.38	-1.22	-0.16	0.00
MeOCF2CHCl2 (Methoxyflurane)	0.8700	1.35	1.28	0.07	0.01
Methane	0.2495	-1.41	-1.42	0.01	0.00
Methyl acetate	0.6057	1.81	1.98	-0.17	-0.22
Methylchloride (Chloromethane)	0.3719	-0.01	0.31	-0.32	-0.23
Methylcyclopentane	0.8454	0.70	-0.07	0.77	0.34
m-Methylstyrene (1-vinyl-3-methylbenzene)	1.0961	2.26	2.28	-0.02	0.03
m-Xylene (1,3-Dimethylbenzene)	0.9982	1.62	1.59	0.03	0.09
Nitrous Oxide	0.2810	-0.43	-0.34	-0.09	-0.21
o-Xylene (1,2-dimethylbenzene)	0.9982	1.71	1.42	0.29	0.06
Pentachloroethane	1.0024	1.86	2.02	-0.16	0.01
Pentane	0.8131	-0.15	-0.29	0.14	0.42
Pentyl acetate	1.1693	2.36	1.98	0.38	0.24
p-Methylstyrene (1-vinyl-4-methylbenzene)	1.0961	2.26	2.37	-0.11	0.03
Propyl acetate	0.8875	1.93	1.88	0.05	0.02
Propyl bromide (1-Bromopropane)	0.7063	0.62	0.97	-0.35	0.00
Propene (Propylene)	0.4883	-0.27	-0.21	-0.06	0.10
p-Xylene (1,4-dimethylbenzene)	0.9982	1.58	1.61	-0.03	0.09
Radon	0.3840	-0.81	-0.39	-0.42	0.10
Styrene (Vinyl benzene)	0.9552	1.67	1.67	0.00	-0.07
Sulphur Hexafluoride	0.4643	-1.85	-2.17	0.32	0.40
Tetrachloroethene (Perchloroethylene)	0.8370	1.30	1.19	0.11	0.04
Toluene (methylbenzene)	0.8573	1.49	1.14	0.35	-0.01
trans-1,2-Dichloroethylene	0.5922	0.55	0.88	-0.33	-0.09
Trichloroethene (Trichloroethylene)	0.7146	1.13	1.14	-0.01	0.00
Vinyl bromide (Bromoethene)	0.5224	0.35	0.49	-0.14	-0.22
Vinyl chloride (Chloroethene)	0.4698	0.32	0.17	0.15	-0.12
Vinylidene fluoride (1,1-Difluoroethylene)	0.3828	-0.54	-0.74	0.20	0.07
Xenon	0.3290	-1.00	-0.85	-0.15	0.06

OUTLIERS					
2,3,4-Trimethylpentane	1.2358	0.64	0.57	0.07	0.63
Octane	1.2358	0.93	0.68	0.25	0.63
2,2,4-Trimethylpentane (Isooctane)	1.2358	0.52	0.23	0.29	0.63

[a]Calculated values for log P, obtained when correlating with the Abraham descriptors (**E**, **S**, **A**, **B** and **V**)

Descriptors E, S, A and B for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 18.

[b]REF= Literature reference

Table 17.0. Data set used in this model to obtain model to correlate air to urine distributions in humans for VOCs

SOLUTE NAME	V	REF[b]	SPECIES	URINE LOG K	CALC LOG K[a]
1,1,1-Trichloroethane (Methyl chloroform)	0.7576	1,	HUMAN	0.18	0.42
1,1,2,2-Tetrachloroethane	0.8800	1,	HUMAN	1.63	1.84
1,1,2-Trichloroethane	0.7576	1,	HUMAN	1.31	1.55
1,1-Dichloroethane	0.6352	1,	HUMAN	0.89	0.90
1,2-Dichloroethane	0.6352	1,	HUMAN	1.53	1.14
1,2-Dichloropropane	0.7761	1,	HUMAN	0.82	1.00
1,3,5-Trimethylbenzene (Mesitylene)	1.1391	1,	HUMAN	1.27	1.18
1,2,3-Trichloropropane	0.8985	1,	HUMAN	1.74	1.73
2-methylcyclohexanone (o-Methylcyclohexanone)	1.0020	1,	HUMAN	3.18	2.90
2-Methylpropan-1-ol (isobutanol)	0.7309	1,	HUMAN	3.39	3.33
Methylcyclohexane	0.9863	1,	HUMAN	0.67	-0.07
4-Methylpentan-2-one	0.9697	1,	HUMAN	1.86	2.39
Butan-1-ol	0.7309	1,	HUMAN	3.10	3.42
Butanone (Methyl ethyl ketone)	0.6879	1,	HUMAN	2.52	2.18
Bromodichloromethane	0.6693	3,	HUMAN	0.81	0.95
Chlorobromomethane (Bromochloromethane)	0.5469	1,	HUMAN	1.10	0.66
Chlorodibromomethane	0.7219	3,	HUMAN	1.05	0.97
Cyclohexane	0.8454	1,	HUMAN	-0.05	-0.17
Cyclohexanone	0.8611	1,	HUMAN	3.09	2.83
Cyclopropane	0.4227	2,	HUMAN	-0.76	-0.56
Dichloromenthane (Methylene Chloride)	0.4943	1,	HUMAN	0.86	0.71
Diethyl ether	0.7309	2,	HUMAN	1.11	1.40
Ethanol	0.4491	1,	HUMAN	3.28	3.07
Ethane	0.3904	2,	HUMAN	-1.55	-0.83
Ethyl acetate	0.7466	1,	HUMAN	2.29	1.92
Ethylbenzene	0.9982	1,	HUMAN	1.21	0.89
Isopropyl acetate	0.8875	1,	HUMAN	1.60	2.03
Methanol	0.3082	1,	HUMAN	3.30	3.15
m-Xylene (1,3-Dimethylbenzene)	0.9982	1,	HUMAN	0.83	0.94
Heptane	1.0949	1,	HUMAN	0.11	-0.02
Hexane	0.9540	1,	HUMAN	0.04	-0.17

Octane	1.2358	1,	HUMAN	-0.05	0.13
Pentane	0.8131	1,	HUMAN	0.02	-0.32
Propan-2-ol	0.5900	1,	HUMAN	3.11	3.18
Propan-2-one (Acetone)	0.5470	1,	HUMAN	2.63	2.10
p-Xylene (1,4-Dimethylbenzene)	0.9982	1,	HUMAN	0.77	0.95
Styrene (Vinyl benzene)	0.9552	1,	HUMAN	0.83	0.96
Tetrachloroethene (Perchloroethylene)	0.8370	1,	HUMAN	0.23	0.23
Tetrachloromethane (Carbon tetrachloride)	0.7391	1,	HUMAN	0.05	0.05
Toluene	0.8573	1,	HUMAN	0.38	0.74
Tribromomethane (Bromoform)	0.7745	1, 3	HUMAN	1.39	1.23
Trichloroethene (Trichloroethylene)	0.7146	1,	HUMAN	0.35	0.50
Trichloromethane (Chloroform)	0.6167	1, 3	HUMAN	0.57	0.84
Isophorone (3,5,5-trimethylcyclohex-2-en-1-one)	1.2408	1,	HUMAN	2.85	3.28
OUTLIERS					
Sulphur hexafluride	0.4643	2,	HUMAN	-2.59	-1.52
Cyclopentadiene	0.6185	1,	HUMAN	1.22	0.23
CHF2OCF2CHFCl (Enflurane)	0.8009	2,	HUMAN	-0.24	0.63
2,3-Dimethylbutane	0.9540	1,	HUMAN	0.72	-0.11
Decane	1.5176	1,	HUMAN	1.53	0.71
Nonane	1.3767	1,	HUMAN	1.53	0.73

Descriptors E, S, A, B and L for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 18. This was used to correlate with the observed values for VOC air to urine distribution (log K).

[a]Calculated values for log K, obtained when correlating with the Abraham descriptors (E, S, A, B and L).

[b]REF= Literature reference

Table 18.0. Data set used to obtain a model to correlate air to olive oil distributions for VOCs

SOLUTE NAME	E	S	A	B	L	REF	C
(CF3)2CHOCH2F (Sevoflurane)	-0.465	0.232	0.080	0.147	1.688	1, 2	
1,1,1,2-Tetrachloroethane	0.542	0.630	0.100	0.080	3.641	1, 2, 3, 4, 9, 10	
1,1,1-Trichloroethane (Methyl chloroform)	0.369	0.410	0.000	0.090	2.733	1, 2, 3, 4, 5, 9, 10, 11	
1,1,2,2-Tetrachloroethane	0.595	0.760	0.160	0.120	3.803	1, 2, 4, 5, 9, 10, 11	
1,1,2-Trichloroethane	0.499	0.680	0.130	0.130	3.290	1, 2, 3, 4, 9, 10, 11	
1,1-Dichloro-1-Fluoroethane (HCFC-141b)	0.084	0.430	0.010	0.050	1.920	7,	
1,1-Dichloroethane	0.322	0.490	0.100	0.100	2.316	1, 2, 3, 4, 5, 9, 10, 11	
1,1-Dichloroethylene	0.362	0.340	0.000	0.050	2.110	2, 10	
1,2,3-Trimethylbenzene (Hemimellitene)	0.728	0.610	0.000	0.190	4.565	2, 9	
1,2,4-Trifluorobenzene	0.410	0.650	0.000	0.020	2.850	2,	
1,2,4-Trimethylbenzene (Pseudocumene)	0.677	0.560	0.000	0.190	4.441	2, 9	
1,2-Dibromoethane	0.747	0.760	0.100	0.170	3.382	2, 3, 4, 5, 10	
1,2-Dichlorobenzene (o-Dichlorobenzene)	0.872	0.780	0.000	0.040	4.518	1, 2, 3, 4	
1,2-Dichloroethane	0.416	0.640	0.100	0.110	2.573	1, 2, 3, 4, 5, 9, 10, 11	
1,2-Dichloropropane	0.371	0.680	0.000	0.150	2.866	1, 2, 3, 4, 10	
1,2-Diflurobenzene (o-Diflurobenzene)	0.390	0.630	0.000	0.060	2.843	2,	
1,2-Dimethoxyethane	0.116	0.670	0.000	0.680	2.654	3, 4	
1,3,5-Triflurobenzene	0.390	0.490	0.000	0.000	2.660	2,	
1,3,5-Trimethylbenzene (Mesitylene)	0.649	0.520	0.000	0.190	4.344	2, 9	
1,3-Dichlorobenzene (m-Dichlorobenzene)	0.847	0.730	0.000	0.020	4.410	1, 2, 3, 4	

1,4-Diflurobenzene (p-Diflurobenzene)	0.384	0.600	0.000	0.060	2.766	2,	2.82	2.69
1,4-Dioxane	0.329	0.750	0.000	0.640	2.892	3, 4, 5	2.83	2.90
1-Bromo-2-chloroethane (1-Chloro-2-Bromoethane)	0.572	0.700	0.100	0.090	2.982	2, 10	2.76	3.09
1-Butanol	0.224	0.420	0.370	0.480	2.601	2, 3, 4	2.99	3.02
1-Chlorobutane	0.210	0.400	0.000	0.100	2.722	1, 2, 3, 4, 5	2.48	2.52
1-Chloropentane	0.208	0.380	0.000	0.090	3.223	1, 2, 3, 4, 5	2.99	2.94
1-Chloropropane (Propyl chloride)	0.216	0.400	0.000	0.100	2.202	1, 2, 3, 4, 5, 10	2.06	2.06
1-Fluropropane	0.034	0.350	0.000	0.130	1.103	1, 3, 4	0.92	1.10
1-Heptanol	0.211	0.420	0.370	0.480	4.115	3, 4	4.26	4.35
1-Hexanol	0.210	0.420	0.370	0.480	3.610	2, 3, 4	3.90	3.91
1-Methoxy-2-propanol	0.218	0.610	0.350	0.620	2.655	2,	2.84	3.19
1-Nitropropane	0.240	0.950	0.000	0.310	2.894	2, 10	3.03	3.14
1-Pentanol	0.219	0.420	0.370	0.480	3.106	2, 3, 4	3.30	3.46
1-Propanol	0.236	0.420	0.370	0.480	2.031	1, 2, 3, 4	2.49	2.52
2,2,4-Trimethylpentane (Isooctane)	0.000	0.000	0.000	0.000	3.106	2, 10, 11	2.58	2.56
2,2-Dichloro-1,1,1-Trifluroethane (HCFC-123)	-0.160	0.400	0.220	0.000	1.746	6,	1.81	2.15
2,2-Dimethylbutane	0.000	0.000	0.000	0.000	2.352	1, 2	1.85	1.90
2,3,4-Trimethylpentane	0.000	0.000	0.000	0.000	3.481	2, 10	2.82	2.89
2-Butoxyethanol	0.201	0.500	0.300	0.830	3.806	2,	3.74	4.00
2-Chloropropane	0.177	0.350	0.000	0.120	1.970	2, 10	1.84	1.82
2-Ethoxyethanol	0.237	0.520	0.310	0.810	2.792	2,	2.98	3.14

2-Fluropropane	0.004	0.320	0.000	0.100	1.070	1, 3, 4	1.09	1.06
2-Heptanone (Methyl pentyl ketone)	0.123	0.680	0.000	0.510	3.760	2, 3, 4	3.85	3.67
2-Hexanone (Methyl butyl ketone)	0.136	0.680	0.000	0.510	3.286	1, 2, 3, 4, 5, 9	3.08	3.25
2-Isopropoxyethanol	0.196	0.470	0.300	0.910	3.170	2,	3.21	3.41
2-Methoxyethanol	0.269	0.500	0.300	0.840	2.490	2,	2.72	2.83
2-Methyl-1-propanol (Isobutanol)	0.217	0.390	0.370	0.480	2.413	1, 2	2.71	2.83
2-Methyl-2-propanol (t-butanol)	0.180	0.300	0.310	0.600	1.963	2, 3, 4	2.25	2.25
2-Methylpentane (Isohexane)	0.000	0.000	0.000	0.000	2.503	1, 2	2.01	2.03
2-Methylpyridine	0.598	0.750	0.000	0.580	3.422	3, 4	3.54	3.30
2-Nitropropane	0.216	0.920	0.000	0.330	2.550	2, 10	2.81	2.81
2-Pentanone (Methyl propyl ketone)	0.143	0.680	0.000	0.510	2.755	1, 2, 4, 5, 9	2.65	2.78
2-Propanol (Isopropanol)	0.212	0.360	0.330	0.560	1.764	1, 2, 3, 4	2.17	2.16
3-Carene	0.511	0.220	0.000	0.100	4.649	8,	3.70	3.96
3-Methyl-1-butanol (Isopentanol)	0.192	0.390	0.370	0.480	3.011	2,	3.00	3.36
3-Methylhexane	0.000	0.000	0.000	0.000	3.044	1, 2	2.49	2.51
3-Methylpentane	0.000	0.000	0.000	0.000	2.581	1, 2	2.07	2.10
3-Methylpyridine	0.631	0.810	0.000	0.540	3.631	3, 4	3.74	3.53
3-Pentanone (Diethyl ketone)	0.154	0.660	0.000	0.510	2.811	1, 2, 3, 4, 5, 9	2.68	2.81
4-Methyl-2-pentanone (Methyl isobutyl ketone)	0.111	0.650	0.000	0.510	3.089	1, 2	3.03	3.06
4-Methylpyridine	0.630	0.820	0.000	0.540	3.640	3, 4	3.75	3.55
Acetic acid	0.265	0.640	0.620	0.440	1.816	3,	3.64	2.95
Acetone (Propan-2-one)	0.179	0.700	0.040	0.490	1.696	1, 2, 3, 4, 5, 9	1.85	1.93

Acetylene (Ethyne)	0.190	0.600	0.060	0.040	0.140	1, 3, 4	0.24	0.55
Allyl chloride (3-Chloro-1-propylene)	0.327	0.560	0.000	0.050	2.109	2, 10	2.04	2.10
Allylbenzene	0.717	0.600	0.000	0.220	4.136	1, 2, 3, 4, 9	3.85	3.79
Argon	0.000	0.000	0.000	0.000	-0.688	1, 3	-0.82	-0.76
Benzene	0.610	0.520	0.000	0.140	2.786	1, 2, 3, 4, 5, 9, 10, 11	2.65	2.57
Benzyl alcohol	0.803	0.870	0.390	0.560	4.221	3, 4	4.73	4.71
Bromobenzene	0.882	0.730	0.000	0.090	4.041	3, 4, 5	4.14	3.79
Bromochloromethane (Chlorobromomethane)	0.541	0.800	0.010	0.060	2.445	2, 10	2.56	2.57
Butan-2-one (Methyl ethyl ketone)	0.166	0.700	0.000	0.510	2.287	1, 2, 3, 4, 5, 9	2.30	2.38
Butane	0.000	0.000	0.000	0.000	1.615	2, 3, 4, 5	1.27	1.26
Butyl acetate	0.071	0.600	0.000	0.450	3.353	2, 3, 4, 5	3.20	3.26
Butyl formate	0.121	0.630	0.000	0.380	2.958	3, 4	2.87	2.93
Butyl propanoate	0.058	0.560	0.000	0.470	3.833	3, 4	3.67	3.64
Butylbenzene	0.600	0.510	0.000	0.150	4.730	3, 4	4.46	4.26
Carbon Dioxide	0.000	0.280	0.050	0.100	0.058	1, 3	0.13	0.22
Carbon Disulfide	0.876	0.260	0.000	0.030	2.370	1, 3	2.18	1.91
Carbon Monoxide	0.000	0.000	0.000	0.040	-0.836	1, 3	-1.01	-0.89
Carbon tetrachloride (Tetrachloromethane)	0.458	0.380	0.000	0.000	2.823	1, 2, 3, 4, 5, 10, 11	2.56	2.53
CF2=CHCl (1-Chloro-2,2-difluroethylene) (1,1-Difluro-2-chloroethene)	-0.340	0.290	0.150	0.000	0.723	1, 3, 4	1.15	1.08
CF2ClCF2Cl	-0.190	0.050	0.000	0.000	1.427	1,	1.25	1.19

CF2HCF2CF2CF2H	-0.710	0.040	0.090	0.000	0.590	2,	1.48	0.74
CF2HCF2CFH2	-0.450	0.170	0.000	0.030	0.680	2,	1.01	0.71
CF2HCF2H (1,1,2,2-Tetrafluoroethane)	-0.280	-0.300	0.300	0.000	0.289	2,	0.67	0.41
CF2HCH3 (1,1-Difluroethane)	-0.250	0.490	0.040	0.050	0.517	2,	0.93	0.87
CF3CBrClH (Halothane)	0.102	0.380	0.150	0.050	2.177	1, 2, 3, 4, 6, 10	2.30	2.31
CF3CBrFH (Teflurane)	-0.070	0.210	0.200	0.020	1.370	1, 2, 3, 4	1.46	1.59
CF3CF2CF2CF2H	-0.780	-0.300	0.100	0.100	0.420	2,	0.61	0.31
CF3CFH2 (Norflurane) (1,1,1,2-Tetrafluoroethane)	-0.640	0.200	0.240	0.000	0.226	2,	0.48	0.80
CF3CFHCFHCF3	-0.710	-0.090	0.090	0.040	0.590	2,	1.34	0.62
CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	0.010	0.400	0.150	0.000	1.168	1, 2, 3, 4, 10	1.38	1.48
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	0.183	0.300	0.000	0.270	1.600	1, 2, 3, 4	1.68	1.44
CF3CH2OCH2CF3	-0.510	0.030	0.080	0.360	1.419	1,	1.67	1.35
CF4 (Perfluoromethane) (Carbon Tetrafluoride)	-0.550	-0.200	0.000	0.000	-0.819	2,	-1.28	-0.90
CFH2CH2CFH2 (1,3-Difluoropropane)	-0.190	0.270	0.050	0.050	1.280	2,	1.67	1.34
CHF2CF2CH2Br (Halopropane)	-0.070	0.280	0.200	0.000	2.030	1, 3, 4	2.51	2.23
CHF2OCF2CHFCl (Enflurane)	-0.230	0.400	0.120	0.130	1.750	1, 2, 4	2.01	1.99
CHF2OCHClCF3 (Isoflurane)	-0.240	0.500	0.100	0.100	1.576	1, 2, 3, 4, 10	1.96	1.90
CHF2OCHFCF3 (Desflurane)	-0.300	0.112	0.067	0.000	1.404	1,	1.27	1.37
Chlorobenzene	0.718	0.650	0.000	0.070	3.657	1, 2, 3, 4, 5, 10	3.44	3.43
Chlorodibromomethane	0.775	0.710	0.070	0.080	3.304	2, 10	3.43	3.27

Chloroethane	0.227	0.400	0.000	0.100	1.678	1, 2, 3, 5, 10	1.57	1.60
Chloroform (Trichloromethane)	0.425	0.490	0.150	0.020	2.480	1, 2, 3, 5, 10, 11	2.60	2.59
cis-1,2-Dichloroethylene	0.436	0.610	0.110	0.050	2.439	1, 2, 3, 9, 10	2.36	2.59
Cycloheptane	0.350	0.100	0.000	0.000	3.704	2,	3.44	3.08
Cyclohexane	0.305	0.100	0.000	0.000	2.964	1, 2, 3, 4, 5, 10, 11	2.47	2.44
Cyclohexene	0.395	0.280	0.000	0.090	2.952	11,	2.66	2.56
Cyclopentane	0.263	0.100	0.000	0.000	2.477	2, 4, 5	2.05	2.03
Cyclopentanone	0.373	0.860	0.000	0.520	3.221	3, 4	3.21	3.29
Cyclopropane	0.408	0.230	0.000	0.000	1.314	1, 2, 4	1.07	1.09
Decane	0.000	0.000	0.000	0.000	4.686	2, 3, 4, 5	3.98	3.94
Dibromomethane	0.714	0.690	0.110	0.070	2.886	2,	2.98	2.97
Dichloromethane (Methylene chloride)	0.387	0.570	0.100	0.050	2.019	1, 2, 3, 4, 5, 10, 11	2.15	2.18
Diethyl ether	0.041	0.250	0.000	0.450	2.015	1, 2, 3, 4, 5, 10	1.80	1.78
Difluorochloromethane	0.000	0.250	0.200	0.000	0.690	1, 3, 4	0.64	1.01
Difluromethane (CF2H2)	-0.320	0.490	0.060	0.050	0.040	2, 10	0.68	0.50
Di-isopropyl ether	-0.060	0.160	0.000	0.580	2.530	3, 4	2.15	2.16
Dimethoxymethane (Methylal)	0.099	0.460	0.000	0.520	1.894	3, 4	1.96	1.84
Dimethyl ether	0.000	0.270	0.000	0.410	1.285	1,	0.97	1.17
Dimethylacetamide (DMA)	0.363	1.350	0.000	0.770	3.638	3, 5	3.90	4.08
Dimethylformamide (DMF)	0.367	1.310	0.000	0.740	3.173	3, 5	3.46	3.63
Dimethylsulfoxide (DMSO)	0.522	1.720	0.000	0.970	3.459	3,	4.38	4.19
Di-n-butyl ether	0.000	0.250	0.000	0.450	3.924	3, 4, 5	3.42	3.46
Divinyl ether	0.259	0.390	0.000	0.130	1.760	1, 2, 3	1.78	1.65

Dodecane	0.000	0.000	0.000	0.000	5.696	4, 5	4.80	4.83
Ethane	0.000	0.000	0.000	0.000	0.492	1, 2, 4	0.26	0.27
Ethanol	0.246	0.420	0.370	0.480	1.485	1, 2, 3, 4	1.98	2.04
Ethene	0.107	0.100	0.000	0.070	0.289	1, 2, 3, 4	0.10	0.15
Ethyl acetate	0.106	0.620	0.000	0.450	2.314	2, 3, 4, 5	2.33	2.36
Ethyl formate	0.146	0.660	0.000	0.380	1.845	1, 4, 3	1.96	1.98
Ethyl propanoate	0.087	0.580	0.000	0.450	2.807	3, 4	2.71	2.76
Ethyl t-butyl ether (tert-Butylether)	-0.020	0.160	0.000	0.600	2.720	2,	2.28	2.32
Ethyl t-pentyl ether	0.030	0.230	0.000	0.370	3.200	2,	2.53	2.81
Ethylbenzene	0.613	0.510	0.000	0.150	3.778	1, 2, 3, 4, 5, 9	3.54	3.43
Ethyl-iodide (C2H5I)	0.640	0.400	0.000	0.140	2.573	3,	2.16	2.27
Fluorobenzene	0.477	0.570	0.000	0.100	2.788	2,	2.82	2.66
Fluoroethane	0.052	0.350	0.000	0.100	0.576	1, 3, 4	0.58	0.64
Fluoromethane	0.066	0.350	0.000	0.090	0.057	1, 3, 4	0.06	0.18
Fluorotrichloromethane (Trichlorofluromethane)	0.207	0.240	0.000	0.070	1.950	1,	1.60	1.70
Fluorochloromethane	-0.080	0.270	0.090	0.030	1.030	2, 10	1.35	1.16
Formic acid	0.343	0.750	0.760	0.330	1.545	3,	3.23	3.04
Helium	0.000	0.000	0.000	0.000	-1.741	1, 3	-1.76	-1.68
Heptane	0.000	0.000	0.000	0.000	3.173	1, 2, 3, 4, 5, 10, 11	2.61	2.62
Hexachloroethane (Perchloroethane)	0.680	0.680	0.000	0.000	4.718	2, 10	3.70	4.40
Hexadecane	0.000	0.000	0.000	0.000	7.714	4, 5	6.57	6.60
Hexaflurobenzene	0.088	0.560	0.000	0.010	2.345	2,	2.40	2.37
Hexane	0.000	0.000	0.000	0.000	2.668	1, 2, 3, 4, 5, 10, 11	2.16	2.18

Hydrogen	0.000	0.000	0.000	0.000	-1.200	1, 3	-1.31	-1.21
Isobutyl acetate	0.052	0.570	0.000	0.470	3.161	2,	3.11	3.07
Isopentylacetate (Aceticacidipentylester)	0.051	0.570	0.000	0.470	3.740	2,	3.47	3.57
Isopropyl acetate	0.055	0.570	0.000	0.470	2.546	2, 3, 4	2.69	2.53
Isopropyl bromide (2-Brompropane)	0.332	0.350	0.000	0.140	2.390	2, 10	2.21	2.15
Isopropylbenzene (Cumene)	0.602	0.490	0.000	0.160	4.084	1, 2, 3, 4, 9	3.75	3.68
Tricyclo[5.2.1.0]-decane (JP10)	0.590	0.450	0.000	0.060	4.840	10,	4.11	4.32
Krypton	0.000	0.000	0.000	0.000	-0.211	1, 3	-0.35	-0.34
Limonene	0.488	0.280	0.000	0.210	4.725	8,	3.76	4.08
MeOCF2CHCl2 (Methoxyflurane)	0.109	0.670	0.070	0.140	2.864	1, 2, 3, 4	2.94	3.03
Methanal (Formaldehyde)	0.220	0.620	0.000	0.330	0.730	3, 4	1.41	0.95
Methane	0.000	0.000	0.000	0.000	-0.323	1, 3	-0.51	-0.44
Methanol	0.278	0.440	0.430	0.470	0.970	1, 2, 3, 4	1.56	1.70
Methyl acetate	0.142	0.640	0.000	0.450	1.911	2, 3, 4, 5	2.00	2.01
Methyl formate	0.192	0.680	0.000	0.380	1.285	3, 4, 5	1.56	1.49
Methyl tert-butyl ether (MTBE)	0.024	0.210	0.000	0.590	2.380	2,	2.08	2.05
Methylchloride (Chloromethane)	0.249	0.430	0.000	0.080	1.163	2, 10	0.93	1.17
Methylcyclopentane	0.225	0.100	0.000	0.000	2.907	1, 2	2.31	2.41
Methylcyclohexane	0.244	0.060	0.000	0.000	3.319	11,	2.82	2.73
Methylpentaflurobenzene (2,3,4,5,6-Pentylfluorotoluene)	0.240	0.450	0.040	0.000	2.900	2,	3.17	2.79
m-Methylstyrene (1-vinyl-3-methylbenzene)	0.866	0.650	0.000	0.180	4.375	2, 10	4.17	4.00

m-Xylene (1,3-Dimethylbenzene)	0.623	0.520	0.000	0.160	3.839	1, 2, 3, 4, 9, 10, 11	3.56	3.49
Neon	0.000	0.000	0.000	0.000	-1.575	1, 3	-1.66	-1.54
Nitroethane	0.270	0.950	0.020	0.330	2.414	3, 4, 5	2.75	2.74
Nitrogen	0.000	0.000	0.000	0.000	-0.978	1,	-1.13	-1.02
Nitromethane	0.313	0.950	0.060	0.310	1.892	1, 3, 4, 5	2.44	2.34
Nitrous Oxide	0.068	0.350	0.000	0.100	0.164	1, 5	0.15	0.27
N,N-Dimethylaniline	0.957	0.810	0.000	0.410	4.701	3, 4	4.66	4.39
Nonane	0.000	0.000	0.000	0.000	4.182	3, 4, 5	3.48	3.50
Octane	0.000	0.000	0.000	0.000	3.677	1, 2, 3, 4, 5, 11	3.07	3.06
Oxygen	0.000	0.000	0.000	0.000	-0.723	1, 3	-0.94	-0.79
o-Xylene	0.663	0.560	0.000	0.160	3.939	1, 2, 3, 4, 9, 10, 11	3.64	3.60
p-Xylene	0.613	0.520	0.000	0.160	3.839	1, 2, 3, 4, 5, 9, 10, 11	3.55	3.49
Pentachloroethane	0.648	0.660	0.170	0.060	4.267	2, 10	3.83	4.28
Pentadecane	0.000	0.000	0.000	0.000	7.209	4,	6.13	6.15
Pentaflurobenzene	0.154	0.680	0.000	0.020	2.578	2,	2.59	2.67
Pentane	0.000	0.000	0.000	0.000	2.162	1, 2, 3, 4, 5, 11	1.71	1.73
Pentyl acetate	0.067	0.600	0.000	0.450	3.844	2, 3, 4	3.52	3.69
Phenol	0.805	0.890	0.600	0.300	3.766	3, 4	4.29	4.71
Piperidine	0.422	0.460	0.130	0.680	3.304	3,	3.91	3.19
p-Methylstyrene (1-vinyl-4-methylbenzene)	0.871	0.650	0.000	0.180	4.399	2, 10	4.14	4.02
Propane	0.000	0.000	0.000	0.000	1.050	5,	0.74	0.76
Propanoic acid	0.233	0.650	0.610	0.440	2.276	3,	3.94	3.35
Propyl acetate	0.092	0.600	0.000	0.450	2.819	2, 3, 4	2.75	2.78

Propyl bromide (1-Bromopropane)	0.366	0.400	0.000	0.120	2.620	2, 10	2.43	2.38
Propyl formate	0.132	0.630	0.000	0.380	2.433	4,	2.42	2.47
Propylbenzene	0.604	0.500	0.000	0.150	4.230	1, 2, 3, 4, 9	3.95	3.82
Pyridine	0.631	0.840	0.000	0.520	3.022	3, 4, 5	3.20	3.02
Styrene (Vinyl benzene)	0.849	0.650	0.000	0.160	3.856	1, 2, 3, 4, 9, 10, 11	3.65	3.56
Sulphur Hexafluoride	-0.600	-0.200	0.000	0.000	-0.120	1, 3	-0.58	-0.28
Tetrachloroethene (Perchloroethylene)	0.639	0.440	0.000	0.000	3.584	1, 2, 3, 5, 9, 10, 11	3.25	3.20
Tetradecane	0.000	0.000	0.000	0.000	6.705	4,	5.69	5.71
Tetrahydrofuran (THF)	0.289	0.520	0.000	0.480	2.636	3, 4, 5	2.39	2.50
Toluene	0.601	0.520	0.000	0.140	3.325	1, 2, 3, 4, 5, 9, 10, 11	3.10	3.04
trans-1,2-Dichloroethylene	0.425	0.410	0.090	0.050	2.278	1, 2, 3, 4, 9, 10	2.22	2.24
Trichloroethene (Trichloroethylene)	0.524	0.370	0.080	0.030	2.997	1, 2, 3, 9, 10, 11	2.79	2.79
Tridecane	0.000	0.000	0.000	0.000	6.200	4,	5.24	5.27
Triethylamine	0.101	0.150	0.000	0.790	3.040	3, 4, 5	2.83	2.54
Undecane	0.000	0.000	0.000	0.000	5.191	4, 5	4.36	4.39
Vinyl bromide (Bromoethene)	0.564	0.500	0.000	0.070	1.846	2, 10	1.75	1.75
Vinyl chloride (Chloroethene)	0.258	0.380	0.000	0.050	1.404	2, 10	1.39	1.34
Xenon	0.000	0.000	0.000	0.000	0.378	1, 4	0.24	0.17
α-Pinene	0.446	0.140	0.000	0.120	4.308	8,	3.46	3.61
β-Pinene	0.530	0.240	0.000	0.190	4.394	8,	3.63	3.74
OUTLIERS								
CF2HCFHCH2CF2H	-0.500	1.250	0.120	0.130	2.324	2,	0.12	2.74
N-Methylimidazole	0.589	0.950	0.000	0.800	3.805	3,	4.84	3.84

2-Methyl-1,3-butadiene (Isoprene)	0.313	0.230	0.000	0.100	2.101	2, 10	0.95	1.78

[a]Calculated values for log K, obtained when correlating with the Abraham descriptors (**E**, **S**, **A**, **B** and **L**)

[b]REF= Literature reference

Table 19.0. Data set used to obtain a model to correlate air to saline distributions for VOCs.

SOLUTE NAME	CONC %	REF[b]	SALINE LOG K	CALC LOG K[a]
(CF3)2CHOCH2F (Sevoflurane)		1,	-0.43	-0.39
1,1,1,2-Tetrachloroethane	0.90%	1, 2, 3	0.62	0.63
1,1,1-Trichloroethane (Methyl chloroform)	0.90%	1, 2, 3	-0.09	-0.20
1,1,2,2-Tetrachloroethane	0.90%	1, 2, 3	1.46	1.36
1,1,2-Trichloroethane	0.90%	1, 2, 3	1.14	1.13
1,1-Dichloro-1-Fluoroethane (HCFC-141b)	0.90%	5,	-0.29	-0.30
1,1-Dichloroethane	0.90%	1, 2, 3	0.34	0.51
1,1-Dichloroethylene	0.90%	1, 2	-0.46	-0.42
1,2,3-Trimethylbenzene (Hemimellitene)		1, 3	0.31	0.56
1,2,4-Trifluorobenzene		1,	0.26	0.09
1,2,4-Trimethylbenzene (Pseudocumene)		1, 3	0.26	0.43
1,2-Dibromoethane	0.90%	1, 2	1.24	1.52
1,2-Dichlorobenzene (o-Dichlorobenzene)		1,	0.95	0.44
1,2-Dichloroethane	0.90%	1, 2, 3	1.03	0.92
1,2-Dichloropropane	0.90%	1, 2	0.53	0.68
1,2-Diflurobenzene (o-Diflurobenzene)		1,	0.49	0.20
1,3,5-Triflurobenzene		1,	-0.14	-0.36
1,3,5-Trimethylbenzene (Mesitylene)		1, 3	0.15	0.33
1,3-Dichlorobenzene (m-Dichlorobenzene)		1,	0.74	0.24
1,4-Diflurobenzene (p-Diflurobenzene)		1,	0.38	0.14
1-Bromo-2-chloroethane (1-Chloro-2-Bromoethane)	0.90%	1, 2	0.95	1.00
1-Butanol		1,	3.13	2.94
1-Chlorobutane	0.90%	1,	-0.07	-0.29
1-Chloropentane	0.90%	1,	-0.15	-0.49
1-Chloropropane (Propyl chloride)	0.90%	1, 2	0.02	-0.17
1-Hexanol		1,	2.96	2.72
1-Methoxy-2-propanol		1,	4.09	3.90
1-Nitropropane	0.90%	1, 2	2.10	1.93
1-Pentanol		1,	3.03	2.83
1-Propanol		1,	3.32	3.07
2,2-Dichloro-1,1,1-Trifluroethane (HCFC-123)	0.90%	4,	-0.68	0.16

2,2-Dimethylbutane		1,	-2.02	-1.75
2,3,4-Trimethylpentane	0.90%	1,	-2.62	-1.99
2-Butoxyethanol		1,	3.85	4.07
2-Chloropropane	0.90%	1, 2	-0.09	-0.19
2-Ethoxyethanol		1,	4.36	4.31
2-Heptanone (Methyl pentyl ketone)		1,	2.16	1.86
2-Hexanone (Methyl butyl ketone)		1, 3	2.04	1.97
2-Isopropoxyethanol		1,	4.09	4.46
2-Methoxyethanol		1,	4.55	4.43
2-Methyl-1,3-butadiene (Isoprene)	0.90%	1, 2	-0.68	-0.51
2-Methyl-1-propanol (Isobutanol)		1,	3.01	2.90
2-Methyl-2-propanol (t-butanol)		1,	2.78	3.02
2-Methylpentane (Isohexane)		1,	-2.03	-1.78
2-Nitropropane	0.90%	1, 2	1.99	2.00
2-Pentanone (Methyl propyl ketone)		1, 3	1.88	2.09
2-Propanol (Isopropanol)		1,	3.08	3.14
3-Methyl-1-butanol (Isopentanol)		1,	2.93	2.76
3-Methylhexane		1,	-2.21	-1.90
3-Pentanone (Diethyl ketone)		1, 3	2.06	2.03
4-Methyl-2-pentanone (Methyl isobutyl ketone)		1,	1.94	1.92
Acetone (Propan-2-one)		1, 3	2.41	2.46
Allyl chloride (3-Chloro-1-propylene)	0.90%	1, 2	0.31	0.10
Allylbenzene		1, 3	0.42	0.74
Benzene	0.90%	1, 2, 3	0.45	0.42
Bromochloromethane (Chlorobromomethane)	0.90%	1, 2	0.94	0.84
Butan-2-one (Methyl ethyl ketone)		1, 3	2.24	2.25
Butane		1,	-1.74	-1.59
Butyl acetate		1,	1.51	1.46
Carbon tetrachloride (Tetrachloromethane)	0.90%	1, 2	-0.45	-0.62
CF2HCF2CF2CF2H		1,	-0.80	-1.37
CF2HCF2CFH2		1,	-0.49	-1.15
CF2HCF2H (1,1,2,2-Tetrafluoroethane)		1,	-0.18	-0.99
CF2HCH3 (1,1-Difluroethane)		1,	0.40	0.05
CF3CBrClH (Halothane)	0.90%	1, 2, 4	-0.30	0.12
CF3CBrFH (Teflurane)		1,	-0.49	-0.16
CF3CF2CF2CF2H		1,	-2.28	-1.75
CF3CFH2 (Norflurane) (1,1,1,2-Tetrafluoroethane)		1,	-0.64	-0.24

CF3CFHCFHCF3		1,	-0.89	-1.52
CF3CH2Cl (2chloro111trifluoroethane) (1-Chloro-2,2,2-trifluroethane)	0.90%	1, 2	-0.38	0.11
CF3CH2OCH=CH2 (Fluomar/ Fluroxene)		1,	-0.08	0.41
CF4 (Perfluoromethane) (Carbon Tetrafluoride)		1,	-2.39	-1.93
CHF2OCF2CHFCl (Enflurane)		1,	-0.13	0.25
CHF2OCHClCF3 (Isoflurane)	0.90%	1, 2	-0.25	0.32
Chlorobenzene	0.90%	1, 2	0.49	0.33
Chlorodibromomethane	0.90%	1, 2	0.87	0.92
Chloroethane	0.90%	1, 2	0.05	-0.06
Chloroform (Trichloromethane)	0.90%	1, 2	0.55	0.41
cis-1,2-Dichloroethylene	0.90%	1, 2, 3	0.38	0.68
Cycloheptane		1,	-1.12	-1.56
Cyclohexane		1,	-0.82	-1.44
Cyclopentane		1,	-1.06	-1.36
Cyclopropane		1,	-0.68	-0.70
Decane		1,	-2.39	-2.24
Dibromomethane		1,	1.16	1.04
Dichloromethane (Methylene chloride)	0.90%	1, 2	0.80	0.59
Diethyl ether	0.90%	1, 2	1.08	0.87
Difluromethane (CF2H2)	0.90%	1, 2	0.12	0.19
Divinyl ether		1,	0.15	0.05
Ethane		1,	-1.60	-1.36
Ethanol		1,	3.31	3.19
Ethene		1,	-1.05	-0.70
Ethyl acetate		1,	1.85	1.75
Ethyl t-butyl ether (tert-Butylether)		1,	0.92	1.10
Ethyl t-pentyl ether		1,	1.08	0.22
Ethylbenzene		1, 3	0.24	0.23
Fluorobenzene		1,	0.36	0.29
Fluorochloromethane	0.90%	1, 2	0.49	-0.36
Hexachloroethane (Perchloroethane)	0.90%	1, 2	-0.18	-0.14
Hexaflurobenzene		1,	-0.40	-0.28
Hexane	0.90%	1, 2	-1.71	-1.82
Isobutyl acetate		1,	1.41	1.50
Isopentylacetate (Aceticacidipentylester)		1,	1.35	1.38
Isopropyl acetate		1,	1.53	1.63
Isopropyl bromide (2-Brompropane)	0.90%	1, 2	0.03	-0.09

Isopropylbenzene (Cumene)		1, 3	0.03	0.16
Tricyclo[5.2.1.0]-decane (JP10)	0.90%	2,	-0.68	-0.53
MeOCF2CHCl2 (Methoxyflurane)		1,	0.62	0.74
Methanol		1,	3.49	3.58
Methyl acetate		1,	2.03	1.91
Methyl tert-butyl ether (MTBE)		1,	1.18	1.28
Methylchloride (Chloromethane)	0.90%	1, 2	-0.06	0.05
Methylcyclopentane		1,	-1.37	-1.48
Methylpentaflurobenzene (2,3,4,5,6-Pentylfluorotoluene)		1,	-0.37	-0.44
m-Methylstyrene (1-vinyl-3-methylbenzene)	0.90%	1, 2	0.29	0.74
m-Xylene (1,3-Dimethylbenzene)	0.90%	1, 2, 3	0.26	0.29
Octane		1,	-2.29	-2.03
o-Xylene	0.90%	1, 2, 3	0.52	0.40
p-Xylene	0.90%	1, 2, 3	0.26	0.29
Pentachloroethane	0.90%	1, 2	0.37	0.84
Pentaflurobenzene		1,	-0.13	0.05
Pentane		1,	-1.85	-1.71
Pentyl acetate		1,	1.38	1.36
p-Methylstyrene (1-vinyl-4-methylbenzene)	0.90%	1, 2	0.32	0.74
Propyl acetate		1,	1.72	1.59
Propyl bromide (1-Bromopropane)	0.90%	1, 2	0.16	-0.08
Propylbenzene		1, 3	0.20	0.11
Styrene (Vinyl benzene)	0.90%	1, 2, 3	0.44	0.76
Tetrachloroethene (Perchloroethylene)	0.90%	1, 2, 3	-0.12	-0.52
Toluene	0.90%	1, 2, 3	0.31	0.30
trans-1,2-Dichloroethylene	0.90%	1, 2, 3	0.16	0.13
Trichloroethene (Trichloroethylene)	0.90%	1, 2, 3	0.01	-0.18
Vinyl bromide (Bromoethene)	0.90%	1, 2	-0.36	0.25
Vinyl chloride (Chloroethene)	0.90%	1, 2	-0.37	-0.24
OUTLIERS				
CFH2CH2CFH2 (1,3-Difluoropropane)		1,	0.96	-0.59
Heptane	0.90%	1, 2	-0.96	-1.94
2,2,4-Trimethylpentane (Isooctane)	0.90%	1,	-2.85	-1.92
3-Methylpentane		1,	-2.72	-1.80
CF2HCFHCH2CF2H		1,	0.77	1.68

[a]Calculated values for log K, obtained when correlating with the Abraham descriptors (**E**, **S**, **A**, **B** and **L**)

[b]REF= Literature reference for the observed values for log K

Descriptors E, S, A, B and L for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 18.

Table 19.3. Data set used to obtain a model to correlate saline to olive oil distributions for VOCs.

SOLUTE NAME	V	SAL:OIL LOG P	CALC LOG P
(CF3)2CHOCH2F (Sevoflurane)	0.8548	2.13	2.16
1,1,1,2-Tetrachloroethane	0.8800	2.94	2.84
1,1,1-Trichloroethane (Methyl chloroform)	0.7576	2.58	2.61
1,1,2,2-Tetrachloroethane	0.8800	2.57	2.45
1,1,2-Trichloroethane	0.7576	2.17	1.98
1,1-Dichloro-1-Fluoroethane (HCFC-141b)	0.6529	2.03	2.12
1,1-Dichloroethane	0.6352	1.93	1.74
1,1-Dichloroethylene	0.5922	2.27	2.18
1,2,3-Trimethylbenzene (Hemimellitene)	1.1391	3.70	3.77
1,2,4-Trifluorobenzene	0.7695	2.53	2.77
1,2,4-Trimethylbenzene (Pseudocumene)	1.1391	3.70	3.78
1,2-Dibromoethane	0.7404	2.14	1.89
1,2-Dichlorobenzene (o-Dichlorobenzene)	0.9612	3.65	3.66
1,2-Dichloroethane	0.6352	1.59	1.62
1,2-Dichloropropane	0.7761	2.27	2.15
1,2-Diflurobenzene (o-Diflurobenzene)	0.7518	2.44	2.52
1,3,5-Triflurobenzene	0.7690	2.67	3.00
1,3,5-Trimethylbenzene (Mesitylene)	1.1391	3.76	3.80
1,3-Dichlorobenzene (m-Dichlorobenzene)	0.9612	3.73	3.78
1,4-Diflurobenzene (p-Diflurobenzene)	0.7518	2.44	2.55
1-Bromo-2-chloroethane (1-Chloro-2-Bromoethane)	0.6878	1.81	1.97
1-Butanol	0.7309	-0.14	-0.17
1-Chlorobutane	0.7946	2.55	2.61
1-Chloropentane	0.9355	3.14	3.25
1-Chloropropane (Propyl chloride)	0.6537	2.04	2.04
1-Hexanol	1.0127	0.94	0.97
1-Methoxy-2-propanol	0.7896	-1.25	-0.70
1-Nitropropane	0.7055	0.93	0.80
1-Pentanol	0.8718	0.27	0.40
1-Propanol	0.5900	-0.83	-0.73
2,2-Dichloro-1,1,1-Trifluroethane (HCFC-123)	0.6883	2.49	1.90
2,2-Dimethylbutane	0.9540	3.87	3.95

2,3,4-Trimethylpentane	1.2358	5.44	5.09
2-Butoxyethanol	1.0714	-0.11	-0.30
2-Chloropropane	0.6537	1.93	1.97
2-Ethoxyethanol	0.7896	-1.38	-1.37
2-Heptanone (Methyl pentyl ketone)	1.1106	1.69	1.72
2-Hexanone (Methyl butyl ketone)	0.9697	1.04	1.16
2-Isopropoxyethanol	0.9305	-0.88	-1.21
2-Methoxyethanol	0.6487	-1.83	-2.02
2-Methyl-1-propanol (Isobutanol)	0.7309	-0.30	-0.14
2-Methyl-2-propanol (t-butanol)	0.7309	-0.53	-0.49
2-Methylpentane (Isohexane)	0.9540	4.04	3.95
2-Nitropropane	0.7055	0.82	0.72
2-Pentanone (Methyl propyl ketone)	0.8288	0.77	0.59
2-Propanol (Isopropanol)	0.5900	-0.91	-0.96
3-Methyl-1-butanol (Isopentanol)	0.8718	0.07	0.41
3-Methylhexane	1.0949	4.70	4.52
3-Pentanone (Diethyl ketone)	0.8288	0.62	0.61
4-Methyl-2-pentanone (Methyl isobutyl ketone)	0.9697	1.09	1.17
Acetone (Propan-2-one)	0.5470	-0.56	-0.55
Allyl chloride (3-Chloro-1-propylene)	0.6106	1.73	2.02
Allylbenzene	1.0960	3.43	3.46
Benzene	0.7164	2.20	2.28
Bromochloromethane (Chlorobromomethane)	0.5469	1.62	1.61
Butan-2-one (Methyl ethyl ketone)	0.6879	0.06	0.01
Butane	0.6722	3.01	2.80
Butyl acetate	1.0284	1.69	1.70
Carbon tetrachloride (Tetrachloromethane)	0.7391	3.01	3.03
CF2HCF2CF2CF2H	0.8130	2.28	2.64
CF2HCF2CFH2	0.6190	1.50	1.97
CF2HCF2H (1,1,2,2-Tetrafluoroethane)	0.4610	0.85	1.39
CF2HCH3 (1,1-Difluroethane)	0.4258	0.53	0.84
CF3CBrClH (Halothane)	0.7410	2.60	2.24
CF3CBrFH (Teflurane)	0.6360	1.95	1.88
CF3CF2CF2CF2H	0.8310	2.89	2.52
CF3CFH2 (Norflurane) (1,1,1,2-Tetrafluoroethane)	0.4612	1.12	0.79
CF3CFHCFHCF3	0.8310	2.23	2.66
CF3CH2Cl (2chloro111trifluoroethane)	0.5659	1.76	1.67

CF3CH2OCH=CH2 (Fluomar/ Fluroxene)	0.7410	1.76	1.70
CF4 (Perfluoromethane) (Carbon Tetrafluoride)	0.3203	1.11	1.18
CHF2OCF2CHFCl (Enflurane)	0.8009	2.14	1.94
CHF2OCHClCF3 (Isoflurane)	0.8010	2.21	2.01
Chlorobenzene	0.8388	2.95	3.04
Chlorodibromomethane	0.7219	2.56	2.35
Chloroethane	0.5128	1.52	1.48
Chloroform (Trichloromethane)	0.6167	2.05	1.99
cis-1,2-Dichloroethylene	0.5922	1.98	1.73
Cycloheptane	0.9863	4.56	4.23
Cyclohexane	0.8454	3.29	3.62
Cyclopentane	0.7045	3.11	3.02
Cyclopropane	0.4227	1.75	1.85
Decane	1.5176	6.37	6.24
Dibromomethane	0.5995	1.82	1.79
Dichloromethane (Methylene chloride)	0.4943	1.35	1.36
Diethyl ether	0.7309	0.72	0.80
Difluromethane (CF2H2)	0.2849	0.56	0.18
Divinyl ether	0.6449	1.63	1.91
Ethane	0.3904	1.86	1.65
Ethanol	0.4491	-1.33	-1.30
Ethene	0.3474	1.15	1.14
Ethyl acetate	0.7466	0.48	0.56
Ethyl t-butyl ether (tert-Butylether)	1.0127	1.36	1.32
Ethylbenzene	0.9982	3.30	3.39
Fluorobenzene	0.7341	2.46	2.39
Fluorochloromethane	0.5300	0.86	1.58
Hexachloroethane (Perchloroethane)	1.1248	3.88	4.46
Hexane	0.9540	3.87	3.95
Isobutyl acetate	1.0284	1.70	1.62
Isopentylacetate (Aceticacidipentylester)	1.1693	2.12	2.19
Isopropyl acetate	0.8875	1.16	1.05
Isopropyl bromide (2-Brompropane)	0.7063	2.18	2.21
Isopropylbenzene (Cumene)	1.1391	3.72	3.93
Tricyclo[5.2.1.0]-decane (JP10)	1.1918	4.79	4.63
MeOCF2CHCl2 (Methoxyflurane)	0.8700	2.32	2.26
Methanol	0.3082	-1.93	-1.95
Methyl acetate	0.6057	-0.03	-0.01

Methyl tert-butyl ether (MTBE)	0.8718	0.90	0.77
Methylchloride (Chloromethane)	0.3719	0.99	0.98
Methylcyclopentane	0.8454	3.68	3.57
Methylpentaflurobenzene (2,3,4,5,6-Pentylfluorotoluene)	0.9458	3.54	3.56
m-Methylstyrene (1-vinyl-3-methylbenzene)	1.0961	3.88	3.70
m-Xylene (1,3-Dimethylbenzene)	0.9982	3.30	3.34
Octane	1.2358	5.36	5.09
o-Xylene	0.9982	3.12	3.33
p-Xylene	0.9982	3.29	3.34
Pentachloroethane	1.0024	3.46	3.33
Pentaflurobenzene	0.8049	2.72	2.71
Pentane	0.8131	3.56	3.37
Pentyl acetate	1.1693	2.14	2.27
p-Methylstyrene (1-vinyl-4-methylbenzene)	1.0961	3.82	3.70
Propyl acetate	0.8875	1.03	1.14
Propyl bromide (1-Bromopropane)	0.7063	2.27	2.27
Propylbenzene	1.1391	3.75	3.97
Styrene (Vinyl benzene)	0.9552	3.21	3.20
Tetrachloroethene (Perchloroethylene)	0.8370	3.37	3.50
Toluene	0.8573	2.79	2.84
trans-1,2-Dichloroethylene	0.5922	2.06	1.96
Trichloroethene (Trichloroethylene)	0.7146	2.78	2.68
Vinyl bromide (Bromoethene)	0.5224	2.11	1.79
Vinyl chloride (Chloroethene)	0.4698	1.76	1.57
OUTLIERS			
Ethyl t-pentyl ether	1.1530	1.45	2.81
Hexaflurobenzene	1.1248	2.80	4.02

Descriptors E, S, A and B for all solutes are found in table 9.0 from chapter 9 and table 18.0 from chapter 18.

Table 21.0. Blood/plasma/serum to brain distribution for drugs only.

SOLUTE NAME	SAMPLE	E	S	A	B	V	I	BRAIN REF	LOG P (B:B)	CALC
1-(3-Fluoropropyl)-2-nitroimidazole (FPN)	Blood	0.760	1.64	0.00	0.76	1.1519	0	24,	-0.24	-0.10
1-(8-fluorooctyl)-2-nitroimidazole (FON)	Blood	0.750	1.62	0.00	0.76	1.8555	0	24,	-0.17	0.39
4-Fluoroantipyrine	Blood	1.380	1.60	0.00	1.25	1.5023	0	25,	-0.05	0.06
4-Iodoantipyrine	Blood	2.100	2.05	0.00	1.31	1.7428	0	25,	-0.10	-0.03
5-Butyl-5-ethyl barbituric acid (Butethal)	Plasma	1.030	1.14	0.47	1.18	1.6557	0	7,	0.19	0.16
5-Heptyl-5-ethyl barbituric acid	Plasma	1.030	1.14	0.47	1.18	2.0784	0	7,	0.02	0.45
5-Hexyl-5-ethyl barbituric acid	Plasma	1.030	1.14	0.47	1.18	1.9375	0	7,	0.36	0.35
5-Methyl-5-ethyl barbituric acid	Plasma	1.030	1.17	0.46	1.18	1.2330	0	7,	-0.22	-0.14
5-Octyl-5-ethyl barbituric acid	Plasma	1.030	1.14	0.47	1.18	2.2193	0	7,	0.24	0.54
5-Pentyl-5-ethyl barbituric acid	Plasma	1.030	1.14	0.47	1.18	1.7966	0	7,	0.09	0.25
5-Propyl-5-ethyl barbituric acid	Plasma	1.030	1.14	0.47	1.18	1.5148	0	7,	0.09	0.06
9-OH Risperidone	Blood	3.050	2.75	0.32	2.02	3.0982	0	22,	-0.67	-0.03
AI-9 (YH1885)	Plasma	2.498	2.20	0.28	1.21	2.8039	0	5,	0.67	0.50
Alprazolam	Blood	2.896	2.50	0.00	1.55	2.2041	0	13,	0.04	-0.03
Alprazolam	Serum	2.896	2.50	0.00	1.55	2.2041	0	14,	-0.04	-0.03
Aminopyrine	Serum	1.680	1.74	0.00	1.60	1.8662	0	20,	0.00	0.09
Amobarbital (5-Ethyl-5-(3-methylbutyl)barbital)	Serum	1.030	1.11	0.47	1.23	1.7966	0	16,	0.04	0.25
Antipyrene (Phenazone)	Blood	1.300	1.83	0.00	1.37	1.4846	0	13, 18	-0.10	-0.21
Barbital (5-Ethyl-5-ethyl barbituric acid)	Plasma	1.030	1.14	0.47	1.18	1.3739	0	7,	-0.14	-0.03
BESP	Blood	2.550	3.41	0.00	2.00	3.4622	0	29,	-0.43	-0.16

Biperiden	Plasma	1.850	1.25	0.31	1.57	2.6196	0	6, 33	0.85	0.81
BMSP	Blood	2.520	3.41	0.00	2.00	3.3213	0	29,	-0.55	-0.26
BPSP	Blood	2.550	3.41	0.00	2.00	3.6031	0	29,	-0.01	-0.06
Bromperidol	Plasma	2.070	1.50	0.40	1.80	2.8506	0	33,	1.38	0.64
Bromperidol	Serum	2.070	1.50	0.40	1.80	2.8506	0	23,	1.38	0.64
BSP	Blood	2.660	2.95	0.24	1.89	3.3213	0	29,	0.07	0.01
C5H8N4S (SKB2)	Blood	1.305	1.00	0.75	0.80	1.1382	0	11,	-0.04	-0.05
Caffeine (1,3,7-Trimethylxanthine)	Blood	1.500	1.72	0.05	1.28	1.3632	0	13,	-0.06	-0.16
Caffeine (1,3,7-Trimethylxanthine)	Plasma	1.500	1.72	0.05	1.28	1.3632	0	17,	0.12	-0.16
Chlorpromazine	Blood	2.200	1.57	0.00	1.01	2.4056	0	13,	1.06	0.95
Chlorpromazine	Serum	2.200	1.57	0.00	1.01	2.4056	0	23,	1.06	0.95
Chlorpromazine	Plasma	2.200	1.57	0.00	1.01	2.4056	0	33,	1.06	0.95
Cimetidine (SKB1)	Blood	1.700	1.73	0.67	1.93	1.9563	0	11, 13	-1.42	-0.46
Clobazam	Serum	2.200	2.71	0.00	1.29	2.1326	0	14,	0.35	-0.25
Clonidine (SKB6)	Blood	1.600	1.34	0.55	1.16	1.5317	0	11, 13	0.11	-0.02
Clozapine	Plasma	2.560	1.52	0.25	1.70	2.4310	0	33,	1.30	0.58
Cocaine	Plasma	1.355	1.92	0.00	1.50	2.2977	0	32, 34	0.73	0.23
Codeine	Plasma	1.960	1.95	0.22	1.79	2.2057	0	8,	-0.22	-0.04
Codeine	Blood	1.960	1.95	0.22	1.79	2.2057	0	3, 13	0.57	-0.04
Cotinine	Plasma	1.049	1.49	0.00	1.38	1.3867	0	33,	-0.38	-0.06
Daidzein	Blood	2.170	2.05	0.92	1.13	1.7871	0	28,	-0.15	-0.54
Desipramine	Blood	1.620	1.82	0.09	0.91	2.2606	0	13,	1.20	0.54
Desmethylclobazam	Serum	2.240	2.25	0.25	1.24	1.9900	0	14,	0.36	-0.13
Desmethyldiazepam	Serum	2.210	1.68	0.26	1.16	1.9330	0	14,	0.50	0.30

Diazepam	Plasma	2.078	1.55	0.00	1.28	2.0739	0	1,	0.03	0.60
Diazepam	Serum	2.078	1.55	0.00	1.28	2.0739	0	14,	0.52	0.60
Didanosine	Plasma	1.810	1.78	0.58	1.90	1.5951	0	15,	-1.30	-0.65
Fentanyl	Plasma	1.830	1.75	0.00	1.81	2.8399	0	6,	0.56	0.67
Flunitrazepam	Serum	2.097	2.15	0.00	1.48	2.1433	0	14,	0.06	0.09
Fluoromisonidazole (1-(3-Fluoro-2-hydroxyproyl-2-nitroimidazole)(FMISO)	Blood	1.020	1.88	0.32	1.29	1.4092	0	24,	-0.01	-0.53
Fluoxetine	Plasma	1.000	1.30	0.10	0.93	2.2403	0	31,	0.36	0.80
Glycyrrhetinic acid	Plasma	1.560	2.17	0.93	1.60	3.8984	1	33,	-1.40	-0.74
Haloperidol	Plasma	1.900	1.39	0.40	1.76	2.7980	0	33,	1.34	0.68
Haloperidol	Serum	1.900	1.39	0.40	1.76	2.7980	0	23,	1.34	0.68
Hexobarbital	Serum	1.500	1.37	0.17	1.37	1.7859	0	16,	0.05	0.28
Hydroxyzine	Blood	2.000	2.21	0.10	1.89	2.9231	0	13,	0.39	0.29
Ibuprofen	Blood	0.730	0.59	0.59	0.81	1.7771	1	13,	-0.18	-0.50
Imipramine (SKB8)	Blood	1.150	1.60	0.00	1.15	2.4015	0	11, 13, 18	1.21	0.67
Indinavir	Plasma	3.500	3.83	1.12	4.00	4.8988	0	15, 27	-0.74	-1.04
Indomethacin	Blood	2.236	1.35	0.57	1.57	2.5299	1	13,	-1.26	-0.64
Indomethacin	Serum	2.236	1.35	0.57	1.57	2.5299	1	20,	-1.26	-0.64
Lorazepam	Plasma	2.510	1.28	0.45	1.63	2.1141	0	9,	0.47	0.44
Lorazepam	Serum	2.510	1.28	0.45	1.63	2.1141	0	14,	0.41	0.44
Lupitidine (SKB5)	Blood	2.960	3.39	0.40	2.20	3.1776	0	11,	-1.06	-0.63
Mannitol	Plasma	0.836	2.26	0.86	1.79	1.3062	0	8,	-1.60	-1.54
Mefloquine (+ve) (AI-8)	Plasma	0.999	1.34	0.43	1.30	2.3276	0	2,	0.63	0.43
Mepyramine (SKB7) (Pyrilamine)	Blood	1.819	1.92	0.00	1.59	2.3870	0	11, 13	0.50	0.33

Mesoridazine	Serum	2.790	2.90	0.00	1.51	2.9604	0	23,	-0.74	0.17
Methohexital	Serum	1.367	1.55	0.21	1.48	2.0903	0	16,	-0.06	0.24
Mianserin	Blood	2.070	1.67	0.00	1.09	2.1520	0	22,	0.99	0.64
Midazolam	Plasma	2.570	2.01	0.00	1.38	2.2629	0	6,	0.52	0.41
Midazolam	Blood	2.570	2.01	0.00	1.38	2.2629	0	13,	0.36	0.41
Midazolam	Serum	2.570	2.01	0.00	1.38	2.2629	0	14,	0.40	0.41
Miloxacin	Plasma	2.000	1.81	0.36	1.38	1.6675	1	33,	-0.92	-1.39
Mirtazapine	Blood	2.090	1.73	0.00	1.37	2.1109	0	22,	0.53	0.44
Morphine	Plasma	2.120	1.88	0.49	1.83	2.0648	0	3, 8	-0.65	-0.25
MSP	Blood	2.300	3.14	0.00	1.98	3.1462	0	29,	0.46	-0.20
Nalidixic acid	Plasma	1.558	1.80	0.59	1.25	1.6999	1	6,	-0.66	-1.54
Nevirapine	Plasma	2.780	2.22	0.30	1.44	1.9946	0	15, 27	0.00	-0.13
Nicotine	Plasma	0.865	0.88	0.00	1.09	1.3710	0	33,	0.32	0.51
Northioridazine	Serum	2.820	1.94	0.21	1.19	2.7608	0	23,	0.75	0.79
o-Ethoxybenzamide (AI-5)	Plasma	0.914	1.51	0.49	0.80	1.3133	0	1,	-0.05	-0.22
Olanzapine	Plasma	2.450	1.40	0.23	1.75	2.3742	0	4,	0.78	0.61
Org13011	Blood	1.070	1.98	0.00	1.96	2.6577	0	22,	0.16	0.16
Org30526	Blood	2.000	1.52	0.13	0.83	1.9515	0	22,	0.39	0.64
Org32104	Blood	2.130	2.00	0.30	1.20	2.1696	0	22,	0.52	0.15
Org34167	Blood	1.801	1.87	0.25	1.12	2.0924	0	22,	0.00	0.21
Org4428 (Beloxepin)	Blood	2.080	1.50	0.25	1.12	2.3105	0	22,	0.82	0.70
Org5222	Blood	1.960	1.80	0.00	0.86	2.0924	0	22,	1.03	0.59
Oxazepam	Blood	2.350	1.10	0.45	1.60	1.9917	0	13,	0.61	0.48
Oxazepam	Serum	2.350	1.10	0.45	1.60	1.9917	0	14,	0.55	0.48

Paracetamol (Acetaminophen)	Blood	1.060	1.63	1.04	0.86	1.1724	0	13,	-0.31	-0.79
Pentazocine	Plasma	1.400	1.15	0.60	1.25	2.4464	0	6, 33	0.64	0.63
Pentobarbital (5-Ethyl-5-(1-methylbutyl)barbital)	Blood	1.030	1.11	0.47	1.23	1.7966	0	13,	0.12	0.25
Pentobarbital (5-Ethyl-5-(1-methylbutyl)barbital)	Serum	1.030	1.11	0.47	1.23	1.7966	0	16,	0.16	0.25
Phenylbutazone	Blood	1.846	2.62	0.00	1.28	2.4329	0	18,	-0.52	-0.04
Phenytoin	Plasma	1.713	2.19	0.85	1.00	1.8693	0	6, 33	-0.16	-0.57
Pindolol	Blood	1.700	1.65	0.30	1.48	2.0090	0	30,	-0.15	0.10
Pindolol	Plasma	1.700	1.65	0.30	1.48	2.0090	0	30,	-0.12	0.10
p-Phenylbenzoic acid	Plasma	1.480	1.30	0.59	0.50	1.5395	1	6,	-1.26	-0.93
Promazine	Blood	2.050	1.70	0.00	1.01	2.2832	0	13,	1.23	0.74
Promazine	Plasma	2.050	1.70	0.00	1.01	2.2832	0	19,	0.67	0.74
Propofol	Plasma	0.815	0.88	0.32	0.51	1.6205	0	33,	0.91	0.71
Propranolol (Propanolol)	Plasma	1.840	1.43	0.44	1.31	2.1480	0	1,	0.96	0.38
Pyrene	Plasma	2.808	1.71	0.00	0.28	1.5846	0	26,	0.23	0.74
R-Etodolac	Plasma	1.803	2.17	1.05	1.15	2.2390	1	6,	-1.51	-1.70
Risperidone	Blood	2.830	2.45	0.00	2.34	3.0395	0	22,	-0.02	0.20
Salicylic acid	Plasma	0.890	0.84	0.71	0.38	0.9904	1	6,	-1.22	-1.08
Salicylic acid (2-Hydroxybenzoic acid)	Blood	0.890	0.84	0.71	0.38	0.9904	1	13,	-1.10	-1.08
S-Etodolac	Plasma	1.803	2.17	1.05	1.15	2.2390	1	6,	-1.34	-1.70
SKB15 (C12H13BrN4O2S)	Blood	2.360	2.36	0.40	1.60	2.1793	0	11,	-0.67	-0.34
SKB16 (C12H14N4O2S)	Blood	2.070	2.30	0.70	1.70	2.0043	0	11,	-0.66	-0.72
SKB17 (C19H20N4O2S)	Blood	2.671	2.68	0.40	1.83	2.7530	0	11,	-0.12	-0.25
SKB19 (C10H10N4S)	Blood	1.906	1.52	0.75	0.94	1.6051	0	11,	-0.18	-0.10
SKB20 (C10H11N5S)	Blood	2.245	1.95	0.98	1.24	1.7050	0	11,	-1.15	-0.60

SKB22 (C12H13N5OS)	Blood	2.776	2.92	1.25	1.61	2.0024	0	11,	-1.57	-1.41
SKB23 (C13H14N8S)	Blood	2.845	2.55	1.46	1.80	2.2980	0	11,	-1.54	-1.15
SKB24 (C14H20N4O3S)	Blood	1.960	3.28	0.26	1.55	2.3870	0	11,	-1.12	-0.87
SKB25 (C21H26N4O3S)	Blood	2.561	2.63	0.40	2.08	3.1365	0	11,	-0.73	-0.08
SKB26 (C17H18N4O3)	Blood	2.321	2.25	0.46	1.67	2.4094	0	11,	-0.27	-0.18
SKB29 (C18H19N5O2)	Blood	2.591	2.51	0.46	2.05	2.5484	0	11,	-0.28	-0.41
SKB30 (C17H26N2O2)	Blood	1.409	2.40	0.40	1.65	2.4317	0	11,	-0.46	-0.39
SKB31 (C22H28N2O2)	Blood	2.009	2.54	0.35	1.66	2.8986	0	11,	-0.24	-0.05
SKB34 (C15H23NO2)	Blood	1.255	1.52	0.37	1.41	2.0931	0	11,	-0.02	0.16
SKB36 (C20H27N3O)	Blood	1.989	2.30	0.20	1.35	2.7009	0	11,	0.69	0.25
SKB37 (C18H25N3OS)	Blood	2.159	2.34	0.28	1.53	2.6256	0	11,	0.44	0.06
SKB42 (C22H27N3O2)	Blood	2.694	2.69	0.40	1.40	2.8898	0	11,	0.22	0.04
SKF 101468	Blood	1.340	1.35	0.55	1.32	2.2321	0	12,	0.25	0.32
SKF 89124	Blood	1.530	1.86	0.87	1.60	2.2908	0	12,	-0.43	-0.35
Spiperone	Blood	2.240	2.80	0.24	1.96	3.0053	0	29,	0.26	-0.20
Stavudine	Plasma	1.257	1.76	0.75	1.61	1.5628	0	27,	-0.48	-0.74
Sulforidazine	Serum	2.720	2.86	0.00	1.63	3.0191	0	23,	0.18	0.17
Theobromine (3,7-Dimethylxanthine)	Plasma	1.500	1.60	0.50	1.38	1.2223	0	17,	-0.30	-0.52
Theophylline (1,3-Dimethylxanthine)	Blood	1.500	1.60	0.54	1.34	1.2223	0	13,	-0.29	-0.53
Theophylline (1,3-Dimethylxanthine)	Plasma	1.500	1.60	0.54	1.34	1.2223	0	17,	0.21	-0.53
Thiopental (Thiopentone)	Plasma	1.480	1.36	0.55	1.04	1.9014	0	6,	-0.15	0.25
Thiopental (Thiopentone)	Serum	1.480	1.36	0.55	1.04	1.9014	0	16,	-0.14	0.25
Thioridazine	Blood	2.700	2.10	0.00	1.30	2.9017	0	13,	0.24	0.83
Thioridazine	Serum	2.700	2.10	0.00	1.30	2.9017	0	23,	0.15	0.83

Thioridazine	Plasma	2.700	2.10	0.00	1.30	2.9017	0	31, 33	0.46	0.83
Tiotidine (SKB10)	Blood	2.305	1.98	1.18	2.23	2.2760	0	11,	-0.82	-0.82
Toliprolol	Blood	1.050	1.16	0.15	1.40	1.9199	0	30,	0.47	0.45
Toliprolol	Plasma	1.050	1.16	0.15	1.40	1.9199	0	30,	0.09	0.45
Triazolam	Serum	2.640	2.20	0.00	1.65	2.3265	0	14,	0.74	0.19
Trifluroperazine	Blood	2.000	1.80	0.00	1.50	2.8911	0	13,	1.44	0.84
Valproic acid	Plasma	0.140	0.57	0.60	0.50	1.3102	1	6,	-1.15	-0.77
Valproic acid	Blood	0.140	0.57	0.60	0.50	1.3102	1	13,	-0.22	-0.77
Y-G 14	Blood	0.750	1.00	0.08	1.08	1.1978	0	12,	-0.30	0.22
Y-G 15	Blood	0.750	0.80	0.00	1.25	1.3387	0	12,	-0.06	0.45
Y-G 16	Blood	0.970	1.20	0.16	0.89	0.9816	0	12,	-0.42	-0.01
Zidovudine (AZT)	Plasma	1.830	1.70	0.47	1.83	1.8192	0	15, 27	-0.70	-0.32
Zolantidine (SKB41)	Blood	2.689	2.64	0.40	1.38	2.9946	0	11,	0.14	0.16
OUTLIERS										
SKB13 ($C_{12}H_{16}BrN_5S$)	Blood	1.900	1.88	0.43	1.81	2.2703	0	11,	-2.15	-0.20
Y-G 19	Blood	1.571	1.70	0.16	0.95	1.5894	0	12,	-1.30	0.15
SKB4	Blood	3.170	2.52	0.35	2.50	3.4468	0	11,	-1.30	-0.04
Fluphenazine	Serum	2.160	2.30	0.26	1.80	3.0907	0	23,	1.49	0.24
Fluphenazine	Plasma	2.160	2.30	0.26	1.80	3.0907	0	33,	1.49	0.24
Y-G 20	Blood	1.390	1.65	0.16	1.24	1.2869	0	12,	-1.40	-0.21
Org12962	Blood	1.100	1.15	0.26	0.93	1.6463	0	22,	1.64	0.47
Temelastine	Blood	2.960	3.24	0.60	2.00	3.0660	0	21,	-1.88	-0.72
Icotidine (SKB3)	Blood	2.740	3.30	0.50	2.27	2.9711	0	11,	-2.00	-0.89
Ranitidine (SKB9)	Blood	1.600	1.63	0.25	2.33	2.3985	0	11, 13	-1.23	-0.13

2,4-Diclorophenoxyacetic acid (AI-4)	Plasma	1.206	1.20	0.72	0.56	1.3761	1	33,	0.15	-0.94
Acetylsalicylic acid (Asprin)	Blood	0.781	1.69	0.71	0.67	1.2879	1	13,	-0.50	-1.57
Terbinafine	Plasma	1.891	1.38	0.00	1.03	2.6061	0	10,	0.08	1.13
Promazine	Serum	2.050	1.70	0.00	1.01	2.2832	0	23,	1.80	0.75
Paraxanthine	Plasma	1.930	1.84	0.37	1.38	1.2223	0	17,	0.57	-0.48

Ref = Literature reference

Note the log BB data is for rats only

Table 22.0. Drug data set used in this model to obtain model to predict blood (or plasma) to fat distributions in rats.

SOLUTE NAME	SAMPLE	FAT REF	LOG P (B:F)
Biperiden	Plasma	2,	1.76
Cocaine	Plasma	5,	0.83
Daidzein	Blood	4,	-0.54
Fentanyl	Plasma	2,	1.43
Hexobarbital	Plasma	2,	0.20
Midazolam	Plasma	1, 2	0.95
Midazolam	Blood	1,	0.93
Nalidixic acid	Plasma	2,	-1.00
Pentazocine	Plasma	2,	0.40
Phenobarbital	Plasma	2,	-0.52
Phenytoin	Plasma	2,	0.26
p-Phenylbenzoic acid	Plasma	2,	-1.23
Procainamide	Plasma	2,	-0.89
Pyrene	Plasma	3,	1.07
R-Etodolac	Plasma	2,	-1.17
S-Etodolac	Plasma	2,	-0.77
Thiopental (Thiopentone)	Plasma	2,	0.89
Valproic acid	Plasma	2,	-0.82

Descriptors E, S, A, B, V and Ia for the rest of the solutes are found in table 21.0 (from chapter 21) and table 25.0 (from chapter 25).

Table 23.0. Drug data set used in this model to obtain model to predict blood (or plasma) to heart distributions in rats.

SOLUTE NAME	SAMPLE	REF	HEART LOG P (B/P:H)	CALC
1-(3-Fluoropropyl)-2-nitroimidazole (FPN)	Blood	11,	-0.02	-0.14
1-(8-fluorooctyl)-2-nitroimidazole (FON)	Blood	11,	-0.05	-0.07
5-Methyl-5-ethyl barbituric acid	Plasma	6,	-0.26	0.24
5-Ethyl-5-ethyl barbituric acid (Barbital)	Plasma	6,	-0.16	0.27
5-Propyl-5-ethyl barbituric acid	Plasma	6,	0.03	0.28
5-Butyl-5-ethyl barbituric acid (Butethal)	Plasma	6,	0.27	0.29
5-Pentyl-5-ethyl barbituric acid	Plasma	6,	0.37	0.31
5-Hexyl-5-ethyl barbituric acid	Plasma	6,	0.36	0.32
5-Heptyl-5-ethyl barbituric acid	Plasma	6,	0.21	0.33
5-Octyl-5-ethyl barbituric acid	Plasma	6,	0.21	0.35
5-Nonyl-5-ethyl barbituric acid	Plasma	6,	0.59	0.36
AI-6	Plasma	9,	0.98	0.45
AI-9 (YH1885)	Plasma	10,	0.20	0.36
Biperiden	Plasma	1, 5	0.85	0.57
Cocaine	Plasma	8,	0.60	0.07
Cefazolin	Plasma	1, 5	-0.96	-0.86
Cotinine	Plasma	1,	-0.32	0.06
Diazepam	Plasma	1, 2	0.36	0.38
Digoxin	Plasma	1,	0.13	0.32
Dicloxacillin	Plasma	1, 5	-1.14	-0.72
Doxorubicin	Plasma	15,	0.76	0.59
Fentanyl	Plasma	1, 5	0.66	0.34
Fluoromisonidazole (1-(3-Fluoro-2-hydroxyproyl-2-nitroimidazole)(FMISO)	Blood	11,	0.04	-0.05
Fluoxetine	Plasma	14,	0.09	0.19
Glycyrrhetinic acid	Plasma	1,	-0.92	-0.65
Hexobarbital	Plasma	1, 5	0.11	0.30
Lorazepam	Plasma	7,	0.60	0.72
Midazolam	Plasma	3, 4, 5	0.62	0.36
Midazolam	Blood	4,	0.59	0.36
Nalidixic acid	Plasma	5,	-0.31	-0.79
Nicotine	Plasma	1,	0.10	0.23
N-Acetylprocainamide (Acecainamide)	Plasma	1,	0.34	0.25
o-Ethoxybenzamide (AI-5)	Plasma	1,	0.01	0.06

Pentazocine	Plasma	1, 5	0.73	0.50
PenicillinV	Plasma	1, 5	-1.01	-0.50
Phenobarbital	Plasma	1, 5	0.06	0.25
Phenytoin	Plasma	1, 5	-0.15	0.15
Pindolol	Blood	13,	0.22	0.30
Pindolol	Plasma	13,	0.26	0.30
p-Phenylbenzoic acid	Plasma	1, 5	-0.64	-0.68
Propranolol (Propanolol)	Plasma	2,	0.70	0.45
Procainamide	Plasma	1, 5	0.40	0.03
Pyrene	Plasma	12,	0.27	0.41
R-Etodolac	Plasma	5,	-0.74	-0.74
R-Carvedilol	Plasma	5,	0.54	0.47
Salicylic acid (2-Hydroxybenzoic acid)	Plasma	5,	-0.72	-0.70
S-Carvedilol	Plasma	5,	0.87	0.47
S-Etodolac	Plasma	5,	-0.35	-0.74
Terbinafine	Plasma	16,	0.26	0.43
Thiopental (Thiopentone)	Plasma	1, 5	0.07	0.36
Thioridazine	Plasma	14,	0.26	0.41
Toliprolol	Blood	13,	0.40	0.27
Toliprolol	Plasma	13,	0.03	0.27
Valproic acid	Plasma	5,	-0.37	-0.78

REF = reference

LOG P (rat data for blood or plasma to heart distribution)

Doxorubicin: E = 3.630, S = 3.50, A = 1.30, B = 3.64, V= 3.7284 and Ia = 0

N-Acetylprocainamide: E = 1.389, S = 1.75, A = 0.42, B = 1.95, V= 2.3153 and Ia = 0

Descriptors E, S, A, B, V and Ia for the rest of the solutes are found in table 21.0 (from chapter 21) and table 25.0 (from chapter 25).

Table 24.0. Drug data set used in this model to obtain a model to predict blood (plasma or serum) to lung distributions in rats.

SOLUTE NAME	SAMPLE	LUNG REF	LOG P (B:L)	CALC
5-Methyl-5-ethyl barbituric acid	Plasma	14,	-0.21	0.20
5-Ethyl-5-ethyl barbituric acid (Barbital)	Plasma	14,	0.00	0.24
5-Propyl-5-ethyl barbituric acid	Plasma	14,	0.16	0.29
5-Butyl-5-ethyl barbituric acid (Butethal)	Plasma	14,	0.18	0.33
5-Pentyl-5-ethyl barbituric acid	Plasma	14,	0.22	0.38
5-Hexyl-5-ethyl barbituric acid	Plasma	14,	0.08	0.42
5-Heptyl-5-ethyl barbituric acid	Plasma	14,	0.12	0.47
5-Octyl-5-ethyl barbituric acid	Plasma	14,	0.49	0.51
5-Nonyl-5-ethyl barbituric acid	Plasma	14,	0.35	0.56
Acebutalol (Acebutolol)	Blood	7,	0.73	0.42
AI-2 (SCH442416)	Blood	2,	0.62	0.69
AI-6	Plasma	5,	0.60	0.11
AI-9 (YH1885)	Plasma	18,	1.34	0.69
Antenolol	Blood	7,	0.28	0.44
Azithromycin	Serum	4,	2.08	1.85
Cocaine	Plasma	6, 20	1.06	0.80
Cefazolin	Plasma	3, 13	-0.11	-0.36
Cotinine	Plasma	3,	-0.18	0.55
Diazepam	Plasma	3, 10	0.53	0.76
Dicloxacillin	Plasma	3, 13	-0.92	-0.28
Doxorubicin	Plasma	9,	0.53	0.72
Erythromycin	Serum	4,	1.62	1.68
Fentanyl	Plasma	3, 13	1.14	1.10
Fluoxetine	Plasma	1,	1.24	0.67
Glycyrrhetinic acid	Plasma	3,	-0.66	-0.15
Hexobarbital	Plasma	3, 13	0.45	0.59
Hydroquinone (1,4-Dihydroxybenzene)	Blood	15,	-0.16	-0.50
Lorazepam	Plasma	16,	0.44	0.63
Metoprolol	Blood	7,	1.06	0.84
Midazolam	Plasma	11, 12, 13	0.65	0.79
Midazolam	Blood	12,	0.62	0.79
Miloxacin	Plasma	3,	-0.29	-0.46
Nalidixic acid	Plasma	13,	-0.48	-0.65
Nicotine	Plasma	3,	0.18	0.57

o-Ethoxybenzamide (AI-5)	Plasma	3,	-0.01	0.05
Olanzapine	Plasma	17,	1.43	0.86
Oxeprenolol	Blood	7,	1.27	0.77
PenicillinV	Plasma	3,13	-0.80	-0.39
Phenobarbital	Plasma	3, 13	-0.06	0.07
Phenytoin	Plasma	3, 13	-0.14	-0.06
Pindolol	Blood	21,	0.99	0.56
Pindolol	Plasma	21,	1.03	0.56
p-Phenylbenzoic acid	Plasma	3, 13	-0.55	-0.80
Practolol	Blood	7,	0.48	0.46
Pyrene	Plasma	19,	0.35	0.35
Salicylic acid (2-Hydroxybenzoic acid)	Plasma	13,	-0.72	-1.02
Terbinafine	Plasma	8,	0.41	0.90
Theophylline (1,3-Dimethylxanthine)	Plasma	3,	-0.15	0.12
Thiopental (Thiopentone)	Plasma	3, 13	0.12	0.30
Thioridazine	Plasma	1,	1.13	0.96
Toliprolol	Blood	21,	1.08	0.68
Toliprolol	Plasma	21,	0.70	0.68
Valproic acid	Plasma	13,	-0.38	-0.80
OUTLIERS				
Propranolol (Propanolol)	Plasma	10,	1.74	0.70
Propranolol (Propanolol)	Blood	7,	1.53	0.65
Pentazocine	Plasma	3, 13	1.43	0.61
Biperiden	Plasma	3, 13,	1.79	0.99

REF = reference

LOG P (rat data for blood, plasma or serum to lung distribution)

Acebutalol: E = 1.600, S = 2.42, A = 0.90, B = 2.10, V= 2.7556 and Ia = 0

Azithromycin: E = 2.242, S = 3.20, A = 1.19, B = 4.90, V= 5.9980 and Ia = 0

Doxorubicin: E = 3.630, S = 3.50, A = 1.30, B = 3.64, V= 3.7284 and Ia = 0

Erythromycin: E = 2.900, S = 3.73, A = 1.25, B = 4.96, V= 5.7730 and Ia = 0

Hydroquinone: E = 1.063, S = 1.27, A = 1.06, B = 0.57, V= 0.8338 and Ia = 0

Metoprolol: E = 1.170, S = 1.33, A = 0.17, B = 1.76, V= 2.2604 and Ia = 0

Oxeprenolol: E = 1.310, S = 1.49, A = 0.17, B = 1.62, V= 2.2174 and Ia = 0

Practolol: E = 1.450, S = 1.90, A = 0.60, B = 1.84, V= 2.1763 and Ia = 0

Descriptors E, S, A, B, V and Ia for the rest of the solutes are found in table 21.0 (from chapter 21) and table 25.0 (from chapter 25).

Table 25.0. Drug data set used in this model to obtain a model to predict blood (or plasma) to muscle distributions in rats.

SOLUTE NAME	E	S	A	B	V	I	SAMPLE	MUSCLE REF	LOG P (B:M)	CALC
3,3',5,5'-Tetrachlorobiphenyl	1.960	1.44	0.00	0.11	1.8138	0	Blood	12,	0.00	0.18
4-Chlorobiphenyl	1.500	1.05	0.00	0.18	1.4466	0	Blood	13,	0.00	0.17
4,4'-Dichlorobiphenyl	1.640	1.18	0.00	0.16	1.5690	0	Blood	13,	0.30	0.17
2,2',4,5,5'-Pentachlorobiphenyl	2.040	1.61	0.00	0.13	1.9362	0	Blood	13,	0.00	0.19
2,2',4,4',5,5'-Hexachlorobiphenyl	2.180	1.74	0.00	0.11	2.0586	0	Blood	13,	0.60	0.20
5-Methyl-5-ethyl barbituric acid	1.030	1.17	0.46	1.18	1.2330	0	Plasma	7,	-0.22	0.09
5-Ethyl-5-ethyl barbituric acid (Barbital)	1.030	1.14	0.47	1.18	1.3739	0	Plasma	7,	-0.09	0.10
5-Propyl-5-ethyl barbituric acid	1.030	1.14	0.47	1.18	1.5148	0	Plasma	7,	0.14	0.11
5-Butyl-5-ethyl barbituric acid (Butethal)	1.030	1.14	0.47	1.18	1.6557	0	Plasma	7,	0.12	0.13
5-Pentyl-5-ethyl barbituric acid	1.030	1.14	0.47	1.18	1.7966	0	Plasma	7,	0.29	0.15
5-Hexyl-5-ethyl barbituric acid	1.030	1.14	0.47	1.18	1.9375	0	Plasma	7,	0.31	0.16
5-Heptyl-5-ethyl barbituric acid	1.030	1.14	0.47	1.18	2.0784	0	Plasma	7,	0.16	0.18
5-Octyl-5-ethyl barbituric acid	1.030	1.14	0.47	1.18	2.2193	0	Plasma	7,	-0.09	0.19
5-Nonyl-5-ethyl barbituric acid	1.030	1.14	0.47	1.18	2.3602	0	Plasma	7,	0.10	0.21
AI-2 (SCH442416)	3.988	3.20	0.25	1.69	2.7696	0	Blood	1,	-0.05	0.17
AI-6	2.250	1.96	0.95	1.97	1.6648	0	Plasma	9,	0.43	-0.03
Biperiden	1.850	1.25	0.31	1.57	2.6196	0	Plasma	2, 6	0.49	0.24
Cocaine	1.355	1.92	0.00	1.50	2.2977	0	Plasma	8,	0.60	0.31

Cefazolin	3.620	4.00	0.75	2.68	2.8265	1	Plasma	2, 6	-0.94	-0.92
Clobazam	2.200	2.71	0.00	1.29	2.1326	0	Plasma	2,	0.41	0.25
Cotinine	1.049	1.49	0.00	1.38	1.3867	0	Plasma	2,	-0.05	0.23
Daidzein	2.170	2.05	0.92	1.13	1.7871	0	Blood	15,	0.08	-0.03
Diazepam	2.078	1.55	0.00	1.28	2.0739	0	Plasma	2, 3	0.15	0.24
2,3-Dideoxyinosine (Didanosine)	1.810	1.78	0.58	1.90	1.5951	0	Plasma	2,	-0.16	0.08
Digoxin	4.240	5.70	1.80	5.25	6.2117	0	Plasma	2,	0.15	0.27
Dicloxacillin	2.900	3.05	0.70	2.04	3.0085	1	Plasma	2, 6	-1.30	-0.87
Fentanyl	1.830	1.75	0.00	1.81	2.8399	0	Plasma	2, 6	0.49	0.35
Fluoxetine	1.000	1.30	0.10	0.93	2.2403	0	Plasma	17,	0.44	0.28
Flurazepam	2.040	1.68	0.00	1.90	2.8959	0	Plasma	2,	0.69	0.35
Glycyrrhetinic acid	1.560	2.17	0.93	1.60	3.8984	1	Plasma	2,	-1.00	-0.77
Hexobarbital	1.500	1.37	0.17	1.37	1.7859	0	Plasma	2, 6	-0.11	0.20
Medazepam	1.898	1.77	0.00	0.89	2.0582	0	Plasma	2,	0.34	0.24
Methotrexate	4.000	3.70	1.60	3.37	3.2197	1	Plasma	10,	-0.82	-1.09
Midazolam	2.570	2.01	0.00	1.38	2.2629	0	Plasma	4, 5, 11	0.11	0.24
Midazolam	2.570	2.01	0.00	1.38	2.2629	0	Blood	5,	0.08	0.24
Morphine	2.120	1.88	0.49	1.83	2.0648	0	Plasma	6,	0.40	0.13
Nalidixic acid	1.558	1.80	0.59	1.25	1.6999	1	Plasma	6,	-0.44	-0.94
Nicotine	0.865	0.88	0.00	1.09	1.3710	0	Plasma	2,	0.19	0.22
o-Ethoxybenzamide (AI-5)	0.914	1.51	0.49	0.80	1.3133	0	Plasma	2,	-0.09	0.09
Pentazocine	1.400	1.15	0.60	1.25	2.4464	0	Plasma	2, 6	0.77	0.16
PenicillinV	2.200	1.90	0.80	1.89	2.4358	1	Plasma	2, 6	-1.22	-0.93

Pentobarbital (5-Ethyl-5-(1-methylbutyl)barbital)	1.030	1.11	0.47	1.23	1.7966	0	Plasma	2,	-0.10	0.15
Phenobarbital	1.630	1.80	0.73	1.15	1.6999	0	Plasma	2, 6	0.06	0.04
Phenytoin	1.713	2.19	0.85	1.00	1.8693	0	Plasma	2, 6	-0.16	0.02
p-Phenylbenzoic acid	1.480	1.30	0.59	0.50	1.5395	1	Plasma	2, 6	-1.10	-0.98
p,p'-Dichlorodiphenylsulfone	1.880	2.20	0.00	0.54	1.8499	0	Blood	16,	0.58	0.21
Propranolol (Propanolol)	1.840	1.43	0.44	1.31	2.1480	0	Plasma	3,	0.34	0.15
Procainamide	1.268	2.15	0.57	1.50	2.0178	0	Plasma	2, 6	0.49	0.15
Prazepam	2.486	1.90	0.00	1.29	2.3880	0	Plasma	2,	0.26	0.25
Pyrene	2.808	1.71	0.00	0.28	1.5846	0	Plasma	14,	0.20	0.11
R-Carvedilol	3.152	2.83	0.84	1.96	3.1032	0	Plasma	6,	-0.10	0.11
Salicylic acid (2-Hydroxybenzoic acid)	0.890	0.84	0.71	0.38	0.9904	1	Plasma	2, 6	-0.85	-1.04
S-Carvedilol	3.152	2.83	0.84	1.96	3.1032	0	Plasma	6,	0.20	0.11
Terbinafine	1.891	1.38	0.00	1.03	2.6061	0	Plasma	11,	0.00	0.30
Tetracycline	3.500	3.50	0.69	3.29	3.0992	0	Plasma	2,	0.29	0.17
Theophylline (1,3-Dimethylxanthine)	1.500	1.60	0.54	1.34	1.2223	0	Plasma	2,	-0.22	0.05
Thiopental (Thiopentone)	1.480	1.36	0.55	1.04	1.9014	0	Plasma	2, 6	-0.12	0.11
Thioridazine	2.700	2.10	0.00	1.30	2.9017	0	Plasma	17,	0.26	0.30
Thiobarbital	1.407	1.18	0.50	1.10	1.4787	0	Plasma	2,	-0.22	0.08
Valproic acid	0.140	0.57	0.60	0.50	1.3102	1	Plasma	6,	-0.80	-0.93

REF = reference

LOG P (rat data for blood and plasma to muscle distribution)

Table 26.0. Drug data set used in this model to obtain a model to predict blood (plasma or serum) to liver distributions in rats.

SOLUTE NAME	SAMPLE	LIVER REF	LOG P (B:L)	CALC
1-(3-Fluoropropyl)-2-nitroimidazole (FPN)	Blood	15,	0.92	0.29
1-(8-fluorooctyl)-2-nitroimidazole (FON)	Blood	15,	0.47	0.43
3,3',5,5'-Tetrachlorobiphenyl	Blood	16,	0.78	0.48
4-Chlorobiphenyl	Blood	17,	0.00	0.40
4,4'-Dichlorobiphenyl	Blood	17,	0.48	0.42
2,2',4,5,5'-Pentachlorobiphenyl	Blood	17,	0.78	0.51
2,2',4,4',5,5'-Hexachlorobiphenyl	Blood	17,	1.08	0.53
5-Methyl-5-ethyl barbituric acid	Plasma	9,	0.44	0.31
5-Ethyl-5-ethyl barbituric acid (Barbital)	Plasma	9,	0.57	0.34
5-Propyl-5-ethyl barbituric acid	Plasma	9,	0.47	0.37
5-Butyl-5-ethyl barbituric acid (Butethal)	Plasma	9,	0.47	0.39
5-Pentyl-5-ethyl barbituric acid	Plasma	9,	0.51	0.42
5-Hexyl-5-ethyl barbituric acid	Plasma	9,	0.44	0.45
5-Heptyl-5-ethyl barbituric acid	Plasma	9,	0.15	0.48
5-Octyl-5-ethyl barbituric acid	Plasma	9,	0.21	0.50
5-Nonyl-5-ethyl barbituric acid	Plasma	9,	0.33	0.53
AI-2 (SCH442416)	Blood	1,	0.94	0.94
AI-6	Plasma	3,	0.58	0.46
AI-9 (YH1885)	Plasma	7,	1.10	0.77
Azithromycin	Serum	2,	2.20	1.58
Cocaine	Plasma	14, 19	0.08	0.68
Cefazolin	Plasma	8,	-0.10	-0.02
Daidzein	Blood	20,	0.06	0.35
Dicloxacillin	Plasma	8,	-0.37	-0.07
Doxorubicin	Plasma	4,	0.64	1.06
Erythromycin	Serum	2,	1.43	1.57
Fentanyl	Plasma	8,	0.58	0.90
Fluoromisonidazole (1-(3-Fluoro-2-hydroxyproyl-2-nitroimidazole) (FMISO)	Blood	15,	0.28	0.35
Fluoxetine	Plasma	22,	0.77	0.56
Hexobarbital	Plasma	8,	0.78	0.57
Hydroquinone (1,4-Dihydroxybenzene)	Blood	11,	-0.12	-0.02

Lorazepam	Plasma	12,	1.08	0.72
Methotrexate	Plasma	10,	0.48	0.00
Midazolam	Plasma	5, 6, 8	0.77	0.79
Midazolam	Blood	6,	0.57	0.79
Morphine	Plasma	8,	0.08	0.64
Nalidixic acid	Plasma	8,	-0.23	-0.48
Olanzapine	Plasma	23,	1.33	0.83
Pentazocine	Plasma	8,	0.36	0.57
PenicillinV	Plasma	8,	-0.60	-0.23
Phenobarbital	Plasma	8,	0.26	0.35
Phenytoin	Plasma	8,	0.36	0.30
p-Phenylbenzoic acid	Plasma	8,	-0.46	-0.59
Procainamide	Plasma	8,	0.51	0.44
Pyrene	Plasma	18,	0.37	0.54
R-Etodolac	Plasma	8,	-0.92	-0.52
R-Carvedilol	Plasma	8,	0.64	0.81
Salicylic acid (2-Hydroxybenzoic acid)	Plasma	8,	-0.64	-0.78
S-Carvedilol	Plasma	8,	1.08	0.81
S-Etodolac	Plasma	8,	-0.37	-0.52
Terbinafine	Plasma	13,	0.18	0.77
Thiopental (Thiopentone)	Plasma	8,	0.36	0.44
Thioridazine	Plasma	22,	0.94	0.91
OUTLIERS				
p,p'-Dichlorodiphenylsulfone	Blood	21,	1.34	0.52
Valproic acid	Plasma	8,	0.26	-0.57

REF = reference

LOG P (rat data for blood, plasma or serum to liver distribution)

Azithromycin: E = 2.242, S = 3.20, A = 1.19, B = 4.90, V= 5.9980 and Ia = 0

Doxorubicin: E = 3.630, S = 3.50, A = 1.30, B = 3.64, V= 3.7284 and Ia = 0

Erythromycin: E = 2.900, S = 3.73, A = 1.25, B = 4.96, V= 5.7730 and Ia = 0

Hydroquinone: E = 1.063, S = 1.27, A = 1.06, B = 0.57, V= 0.8338 and Ia = 0

Descriptors E, S, A, B, V and Ia for the rest of the solutes are found in table 21.0 (from chapter 21) and table 25.0 (from chapter 25)

Table 27.0. Data set used in this model to obtain model to predict drug blood (& plasma) to skin distributions in rats

COMPOUND NAME	SAMPLE	REF	SKIN LOG BS[a]	CALC LOG BS[a]
2,2',4,4',5,5'-Hexachlorobiphenyl	BLOOD	9	1.48	1.05
2,2',4,5,5'-Pentachlorobiphenyl	BLOOD	9	0.85	1.00
3,3',5,5'-Tetrachlorobiphenyl	BLOOD	10	0.85	0.98
4,4'-Dichlorobiphenyl	BLOOD	9	1.00	0.87
4-Chlorobiphenyl	BLOOD	9	1.00	0.83
5-Butyl-5-ethyl barbituric acid (Butethal)	PLASMA	8	0.14	0.12
5-Heptyl-5-ethyl barbituric acid	PLASMA	8	0.04	0.18
5-Hexyl-5-ethyl barbituric acid	PLASMA	8	0.42	0.16
5-Methyl-5-ethyl barbituric acid	PLASMA	8	0.05	0.07
5-Nonyl-5-ethyl barbituric acid	PLASMA	8	0.31	0.21
5-Octyl-5-ethyl barbituric acid	PLASMA	8	0.09	0.20
5-Pentyl-5-ethyl barbituric acid	PLASMA	8	0.05	0.14
5-Propyl-5-ethyl barbituric acid	PLASMA	8	0.19	0.10
Acrylic acid	BLOOD	12	0.01	-0.49
Barbital (5-Ethyl-5-ethyl barbituric acid)	PLASMA	8	0.08	0.08
Biperiden	PLASMA	2, 7	0.60	0.20
Cefazolin	PLASMA	2, 7	-0.53	-0.84
Decane	BLOOD	11	0.68	0.73
Diazepam	PLASMA	2, 4	0.54	0.46
Fentanyl	PLASMA	7	0.32	0.30
Glycyrrhetinic acid	PLASMA	2	-0.80	-0.55
Hexobarbital	PLASMA	2, 7	-0.03	0.24
Midazolam	PLASMA	5, 6, 7	0.15	0.51
Midazolam	BLOOD	6,	0.11	0.51
Nalidixic acid	PLASMA	7	-0.46	-0.55
Nicotine	PLASMA	2	0.04	0.31
o-Ethoxybenzamide (AI-5)	PLASMA	2, 4	0.02	0.27
p,p'-Dichlorodiphenylsulfone	BLOOD	1,	1.21	0.83
Pentazocine	PLASMA	2, 7	0.67	0.16
Phenobarbital	PLASMA	2, 7	0.14	0.12
Phenytoin	PLASMA	2, 7	-0.03	0.20

p-Phenylbenzoic acid	PLASMA	7	-0.82	-0.25
Salicylic acid	PLASMA	2, 7	-0.58	-0.40
Thiopental (Thiopentone)	PLASMA	2, 7	0.07	0.23
Valproic acid	PLASMA	7	-0.33	-0.44
OULIERS				
Terbinafine	PLASMA	3	1.61	0.75

[a] blood (or plasma) to skin distribution of drugs

Acrylic acid: E = 0.357, S = 0.58, A = 0.60, B = 0.43, V= 0.5627 and Ia = 1

Decane: E = 0.00, S = 0.00, A = 0.00, B = 0.00, V= 1.5176 and Ia = 0

Descriptors E, S, A, B, V and Ia for the rest of the solutes are found in table 21.0 (from chapter 21) and table 25.0 (from chapter 25).

Table 28.0. Drug data set used in this model to obtain a model to predict blood (plasma or serum) to kidney distributions in rats.

SOLUTE NAME	SAMPLE	REF	KIDNEY LOG P (B:K)	CALC
1-(3-Fluoropropyl)-2-nitroimidazole (FPN)	Blood	16,	0.61	0.37
1-(8-fluorooctyl)-2-nitroimidazole (FON)	Blood	16,	0.47	0.59
5-Methyl-5-ethyl barbituric acid	Plasma	12,	0.10	0.31
5-Ethyl-5-ethyl barbituric acid (Barbital)	Plasma	12,	0.25	0.35
5-Propyl-5-ethyl barbituric acid	Plasma	12,	0.59	0.40
5-Butyl-5-ethyl barbituric acid (Butethal)	Plasma	12,	0.64	0.44
5-Pentyl-5-ethyl barbituric acid	Plasma	12,	0.46	0.49
5-Hexyl-5-ethyl barbituric acid	Plasma	12,	0.32	0.53
5-Heptyl-5-ethyl barbituric acid	Plasma	12,	0.31	0.58
5-Octyl-5-ethyl barbituric acid	Plasma	12,	0.40	0.62
5-Nonyl-5-ethyl barbituric acid	Plasma	12,	0.93	0.66
AI-2 (SCH442416)	Blood	2,	0.67	0.70
AI-9 (YH1885)	Plasma	10,	1.23	0.75
Azithromycin	Serum	3,	2.19	2.00
Biperiden	Plasma	11,	1.04	0.88
Cefazolin	Plasma	11,	0.45	0.19
Daidzein	Blood	19,	-0.06	0.12
Diazepam	Plasma	6,	0.38	0.75
Dicloxacillin	Plasma	11,	0.11	0.27
Doxorubicin	Plasma	5,	0.71	0.89
Erythromycin	Serum	3,	1.42	1.83
Fentanyl	Plasma	11,	1.08	1.09
Fluoromisonidazole (1-(3-Fluoro-2-hydroxyproyl-2-nitroimidazole)(FMISO)	Blood	16,	0.23	0.37
Fluoxetine	Plasma	1,	0.97	0.73
Hexobarbital	Plasma	11,	0.18	0.63
Hydroquinone (1,4-Dihydroxybenzene)	Blood	13,	-0.16	-0.25
Lorazepam	Plasma	14,	0.62	0.64
Methotrexate	Plasma	18,	0.48	0.08
Midazolam	Plasma	7, 8, 11	0.66	0.77
Midazolam	Blood	8,	0.64	0.77
Morphine	Plasma	11,	0.98	0.59
Nalidixic acid	Plasma	11,	-0.27	-0.09

Olanzapine	Plasma	9,	0.76	0.85
Pentazocine	Plasma	11,	1.30	0.63
PenicillinV	Plasma	11,	0.57	0.15
Phenobarbital	Plasma	11,	-0.14	0.23
Phenytoin	Plasma	11,	0.20	0.15
p-Phenylbenzoic acid	Plasma	11,	-0.52	-0.24
p,p'-Dichlorodiphenylsulfone	Blood	20,	0.67	0.45
Propranolol (Propanolol)	Plasma	6,	0.58	0.59
Procainamide	Plasma	11,	0.81	0.45
Pyrene	Plasma	17,	0.38	0.35
R-Etodolac	Plasma	11,	-0.92	-0.21
R-Carvedilol	Plasma	11,	0.43	0.64
Salicylic acid (2-Hydroxybenzoic acid)	Plasma	11,	-0.36	-0.43
S-Carvedilol	Plasma	11,	0.85	0.64
S-Etodolac	Plasma	11,	-0.41	-0.21
Terbinafine	Plasma	15,	0.45	0.89
Thiopental (Thiopental)	Plasma	11,	0.49	0.41
Thioridazine	Plasma	1,	1.19	0.94
Valproic acid	Plasma	11,	0.18	-0.20

OUTLIERS				
AI-6	Plasma	4,	1.71	0.50

REF = reference

LOG P (rat data for blood, plasma or serum to kidney distribution)

Azithromycin: E = 2.242, S = 3.20, A = 1.19, B = 4.90, V= 5.9980 and Ia = 0

Doxorubicin: E = 3.630, S = 3.50, A = 1.30, B = 3.64, V= 3.7284 and Ia = 0

Erythromycin: E = 2.900, S = 3.73, A = 1.25, B = 4.96, V= 5.7730 and Ia = 0

Hydroquinone: E = 1.063, S = 1.27, A = 1.06, B = 0.57, V= 0.8338 and Ia = 0

Descriptors E, S, A, B, V and Ia for the rest of the solutes are found in table 21.0 (from chapter 21) and table 25.0 (from chapter 25).

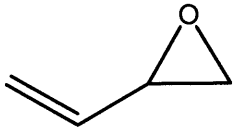
Appendix

Structures

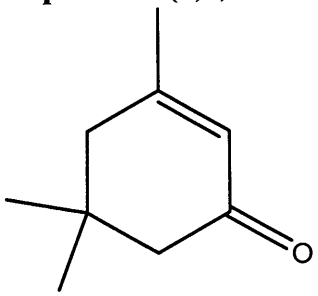
Structures for trivial names

Structures for VOCs

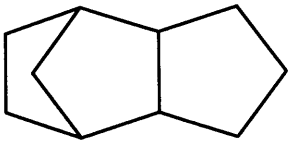
1,2-epoxy-3-butene (BMO)



Isophorone (3,5,5-trimethylcyclohex-2-en-1-one)

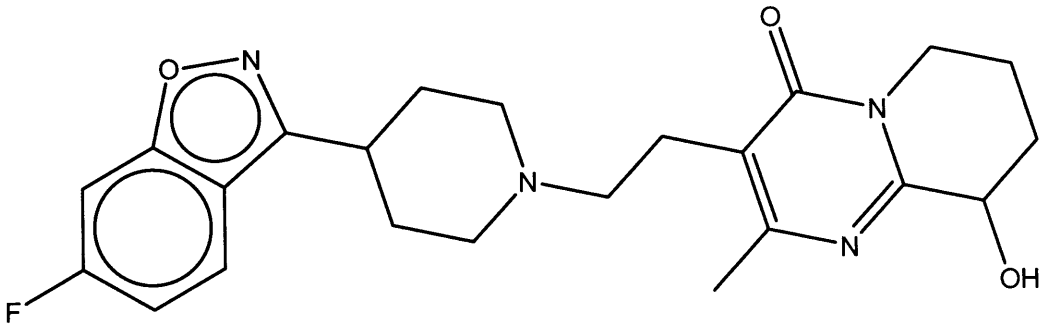


JP-10

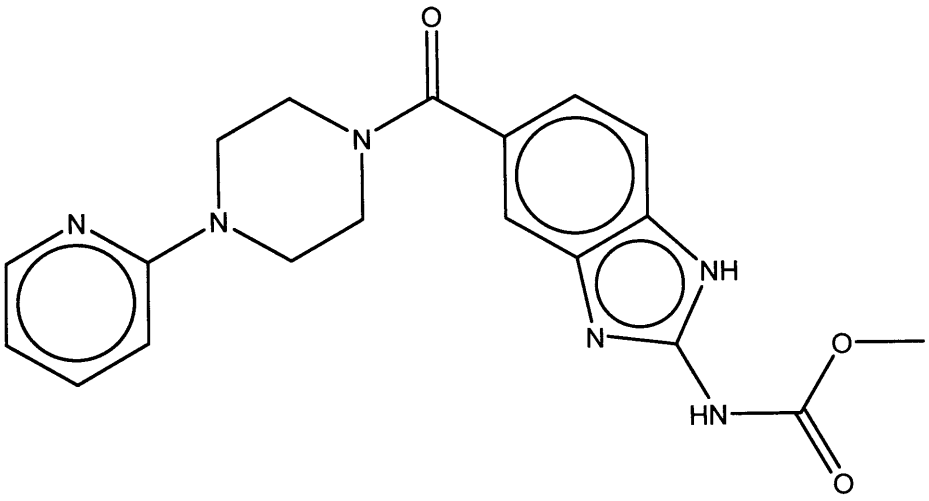


Structures for drugs

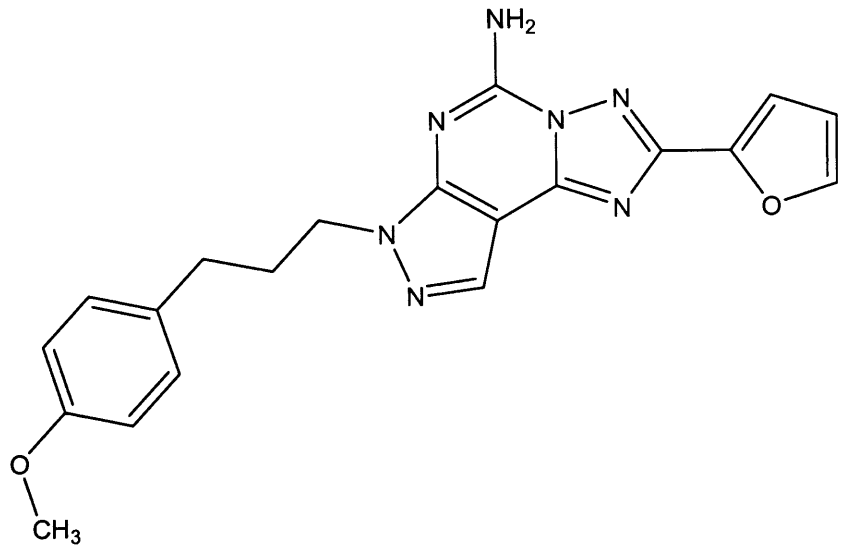
9-OH Risperidone



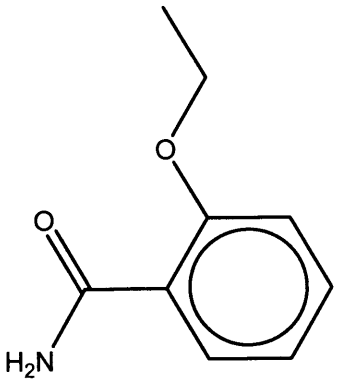
AI-1 (CDRI-81/470)



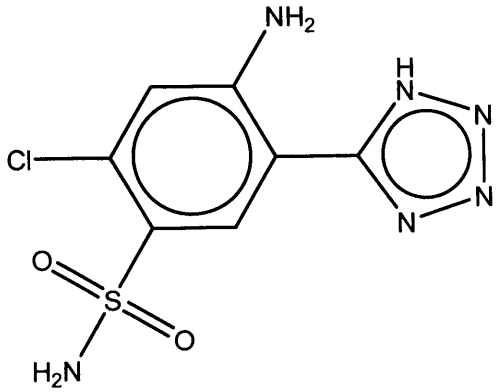
AI-2 (SCH442416)



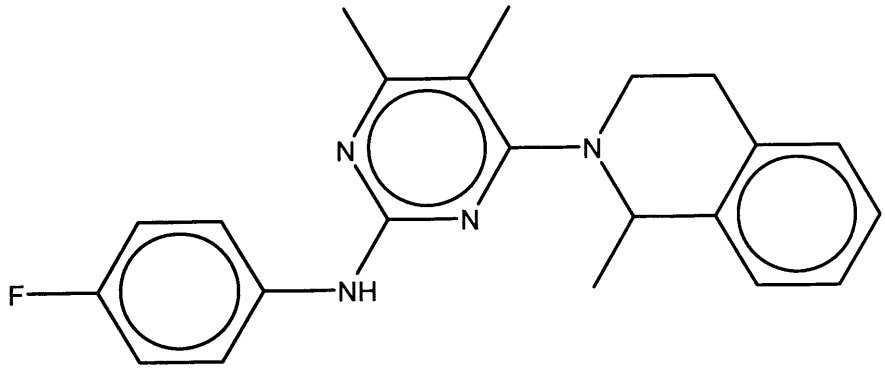
AI-5 (o-Ethoxybenzamide)



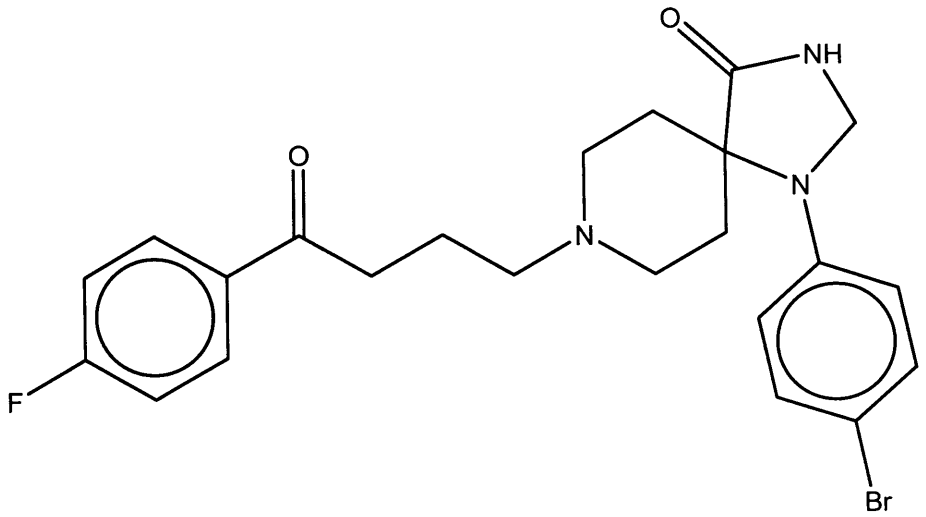
AI-6



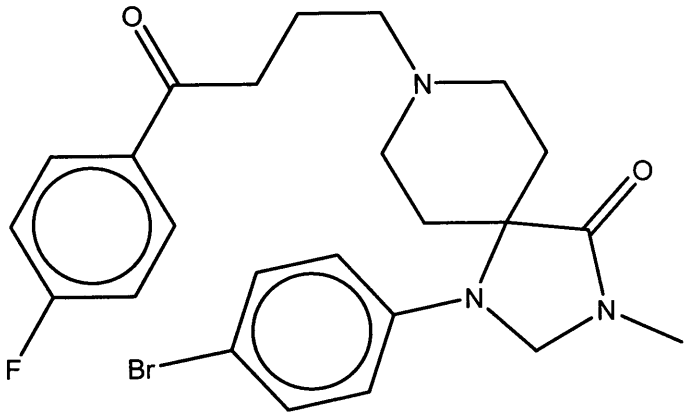
AI-9 (YH1885)



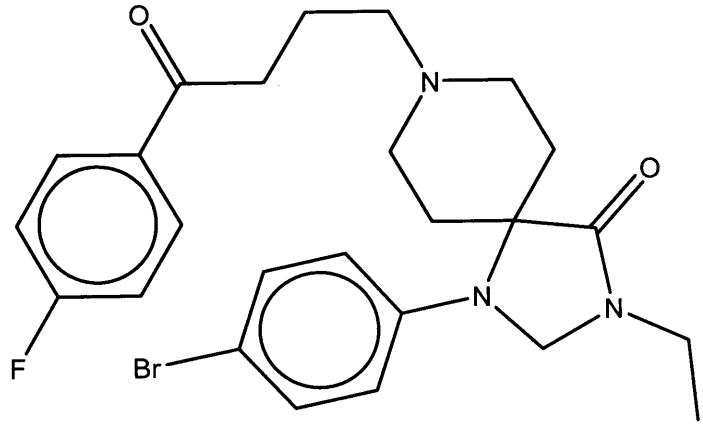
BSP



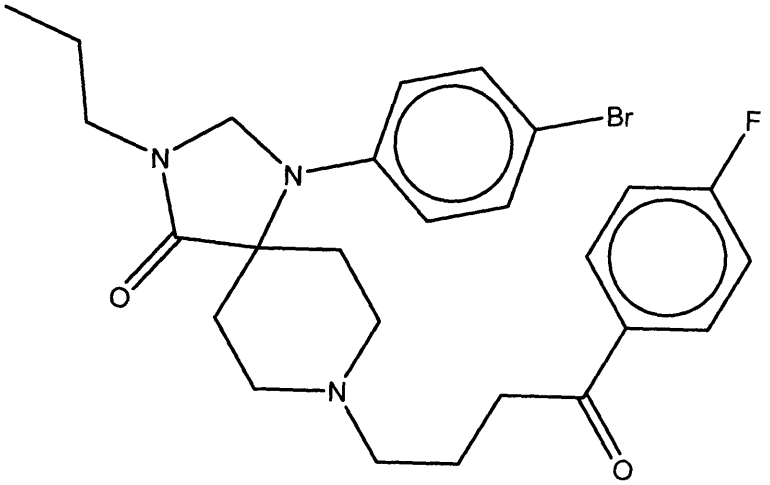
BMSP



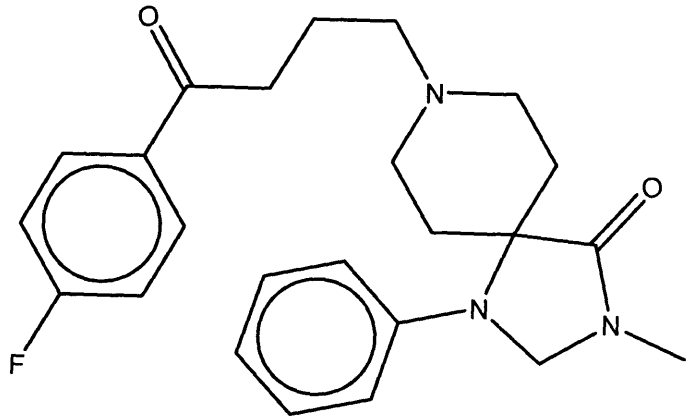
BESP



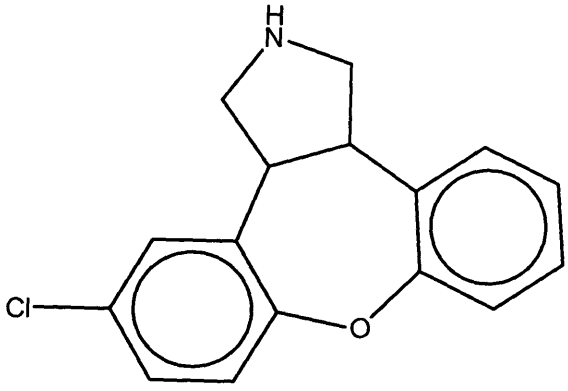
BPSP



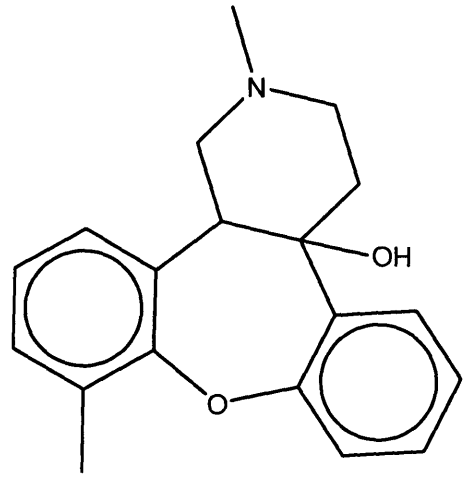
MSP



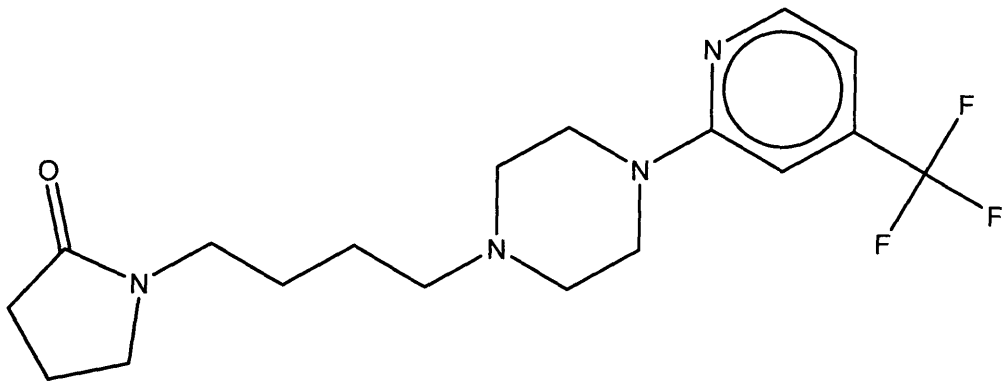
Org30526



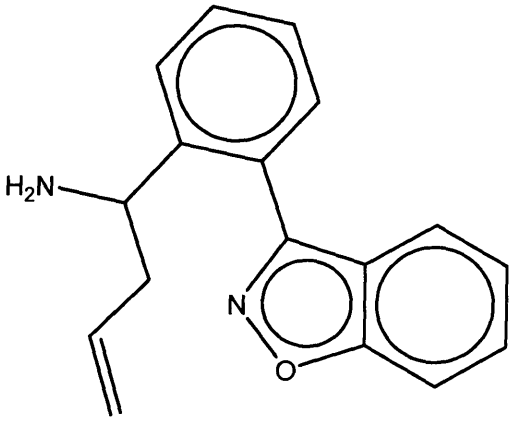
Org4428 (Beloxepin)



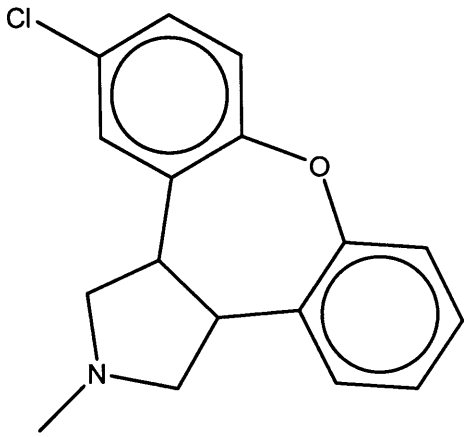
Org13011



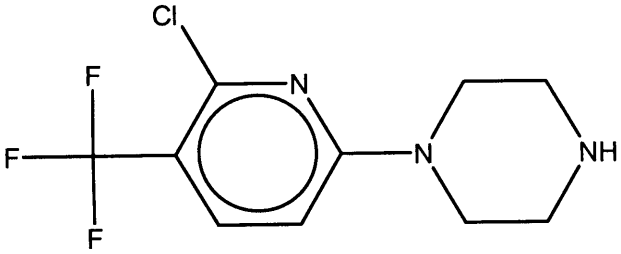
Org34167



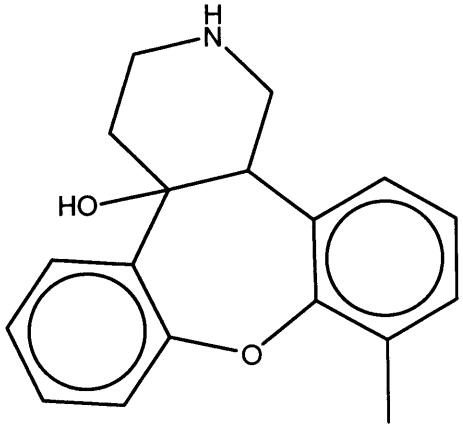
Org5222



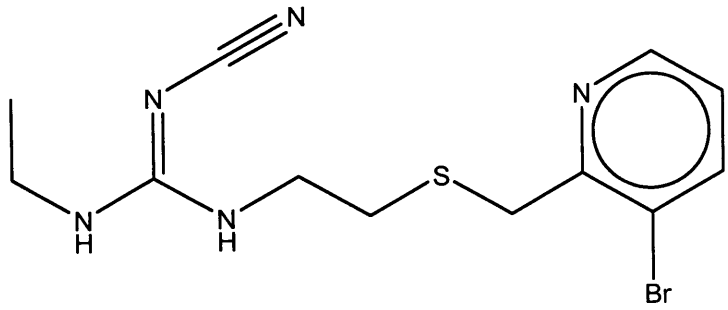
Org12962



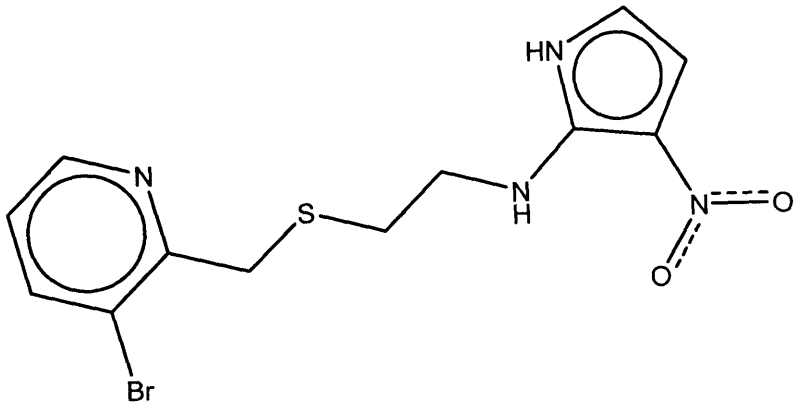
Org32104



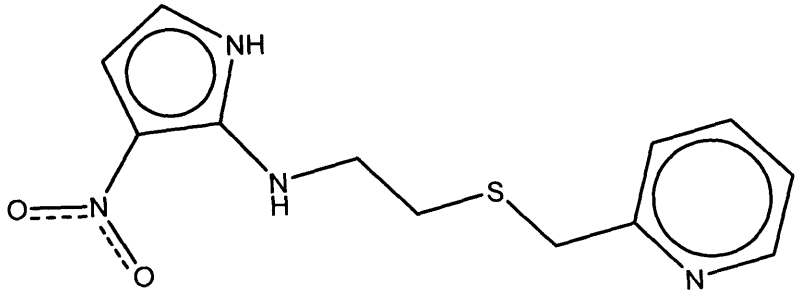
SKB13 (C12H16BrN5S)



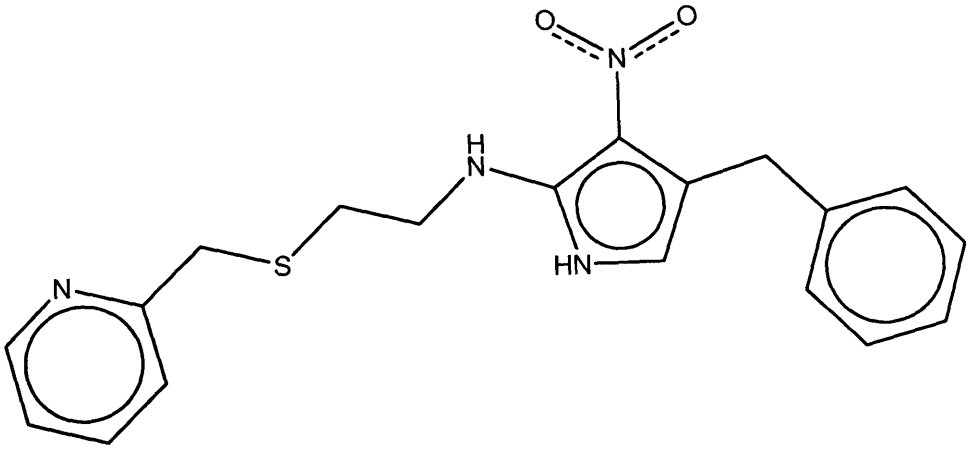
SKB15 (C12H13BrN4O2S)



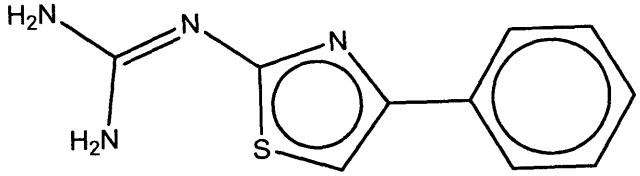
SKB16 (C12H14N4O2S)



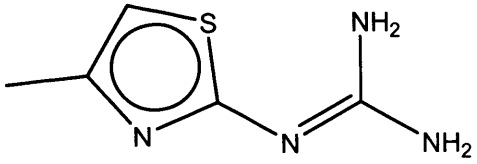
SKB17 (C19H20N4O2S)



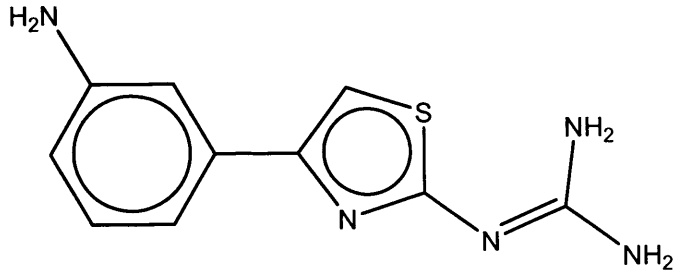
SKB19 (C10H10N4S)



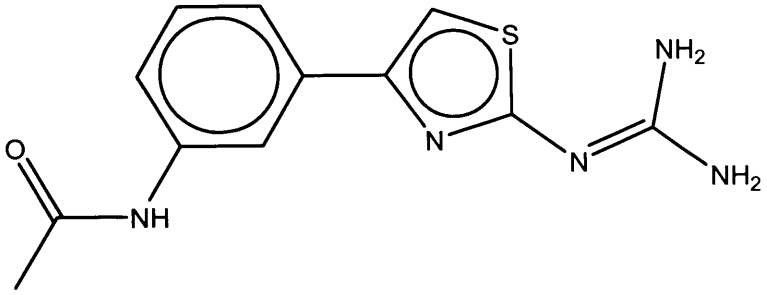
SKB2 (C5H8N4S)



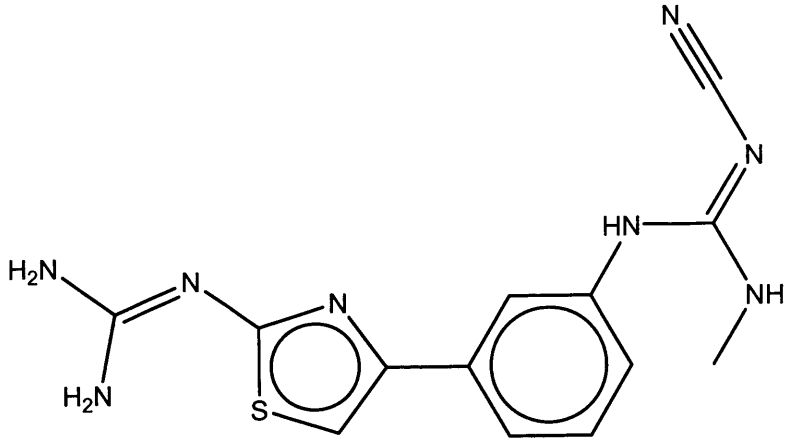
SKB20 (C10H11N5S)



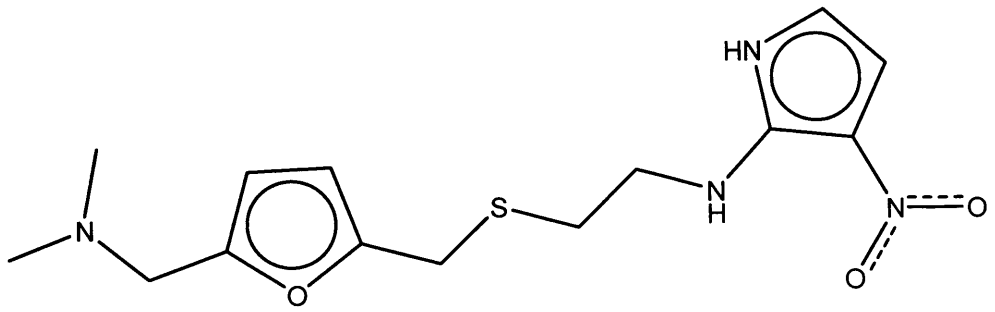
SKB22 (C12H13N5OS)



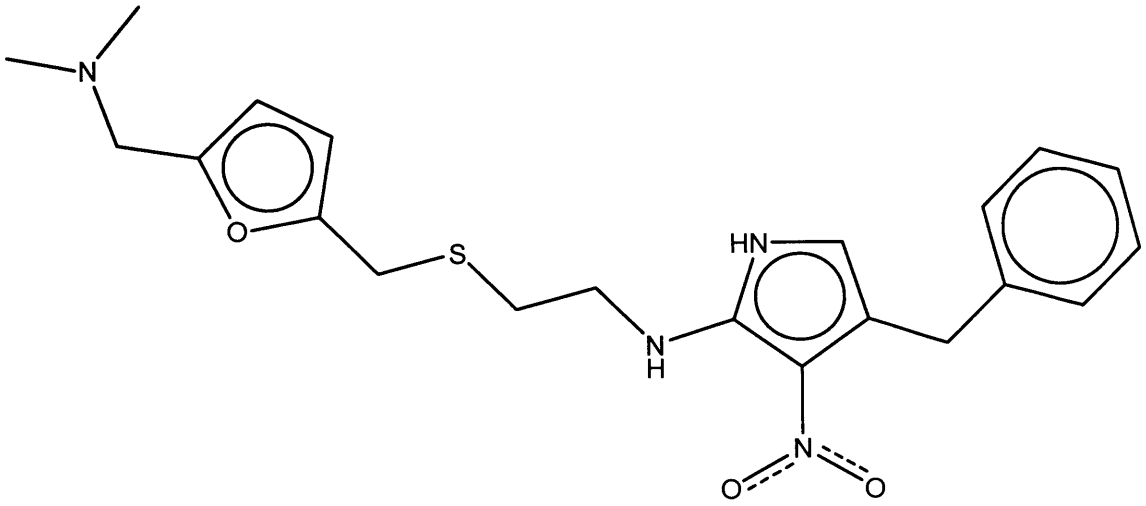
SKB23 (C13H14N8S)



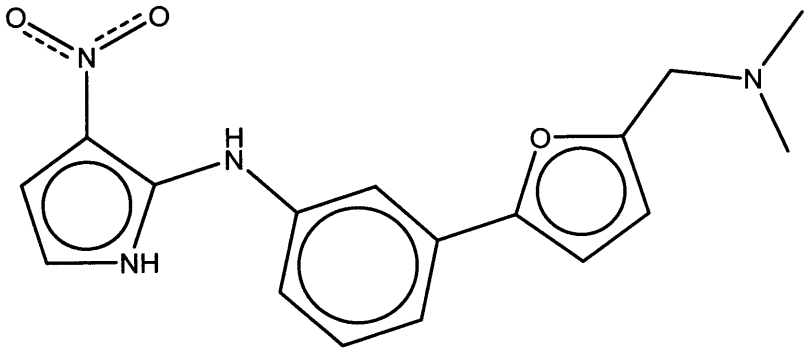
SKB24 (C14H20N4O3S)



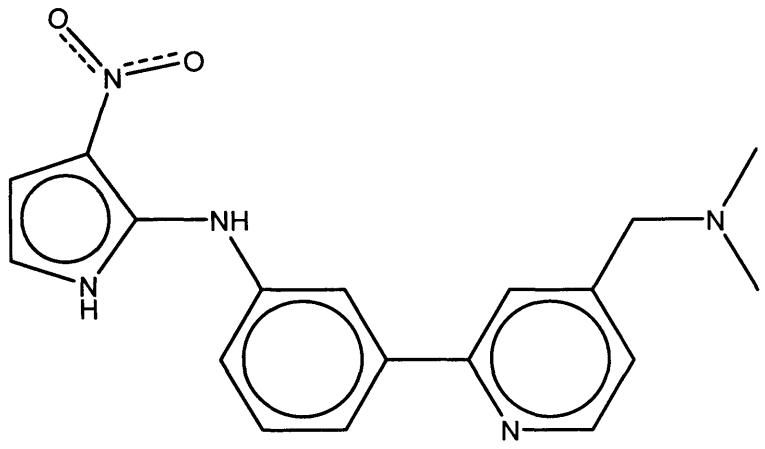
SKB25 (C21H26N4O3S)



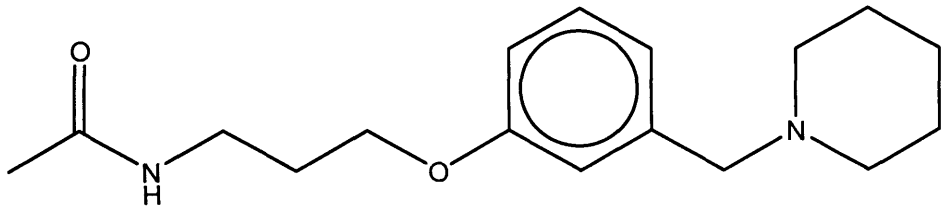
SKB26 (C17H18N4O3)



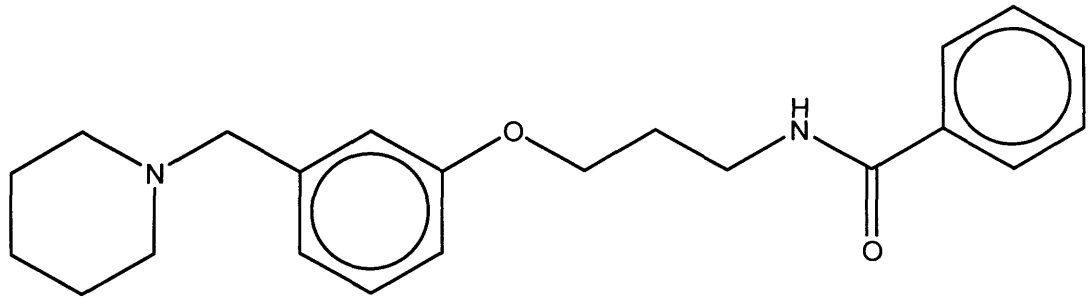
SKB29 (C18H19N5O2)



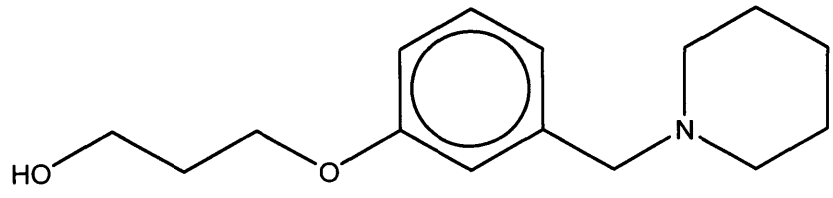
SKB30 (C17H26N2O2)



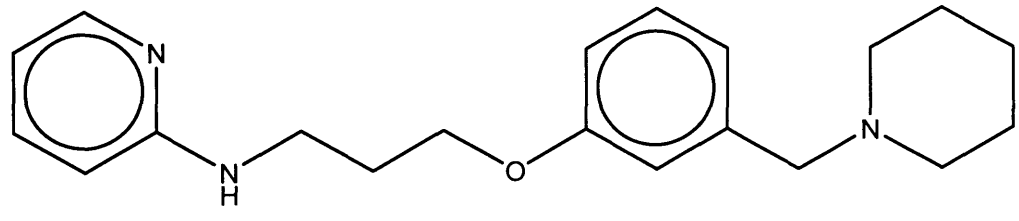
SKB31 (C22H28N2O2)



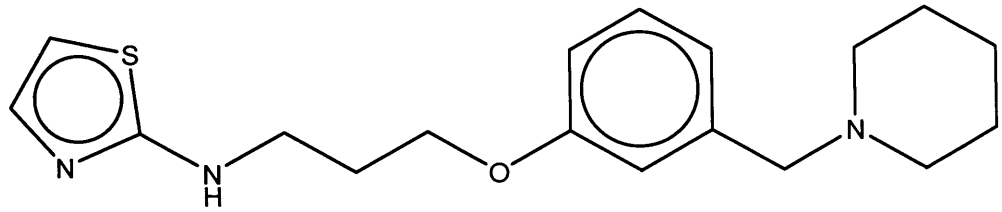
SKB34 (C15H23NO2)



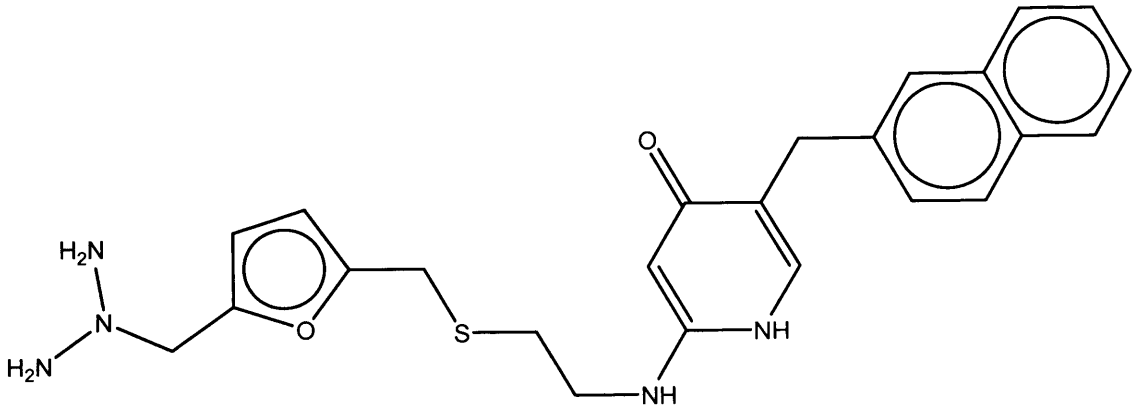
SKB36 (C20H27N3O)



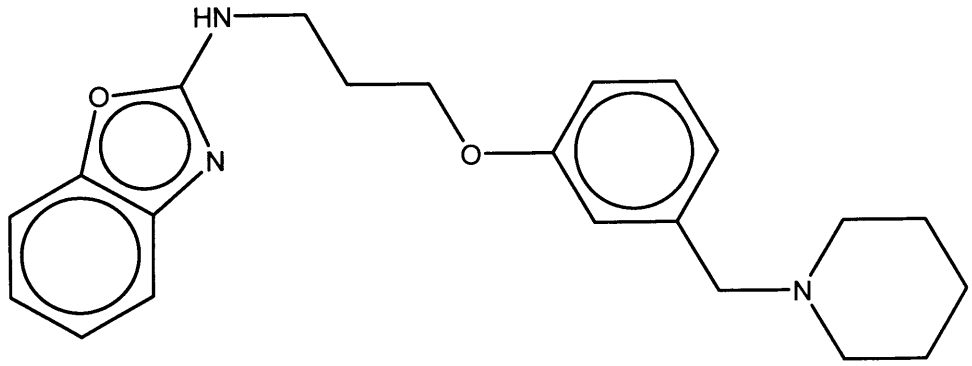
SKB37 (C18H25N3OS)



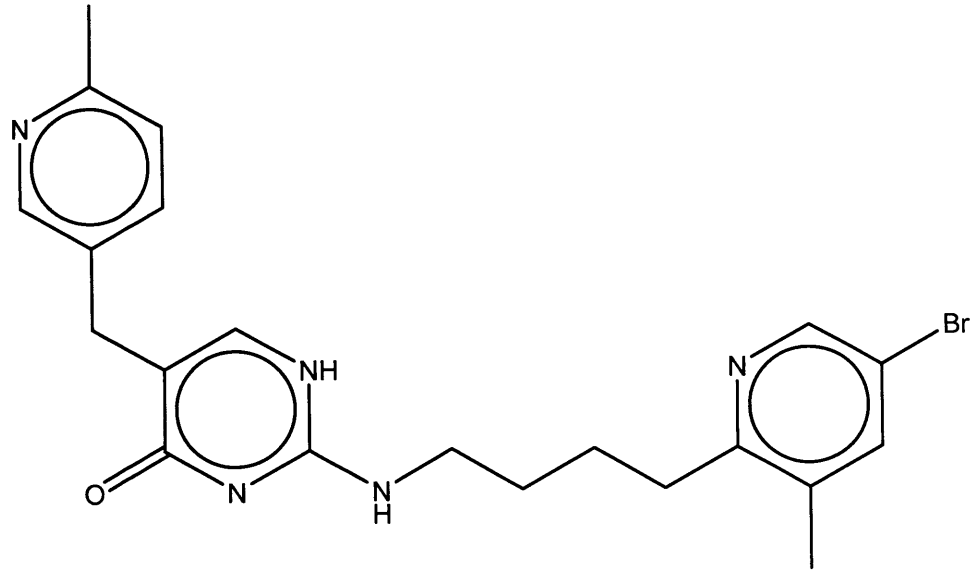
SKB4 (C24H27N5O2S)



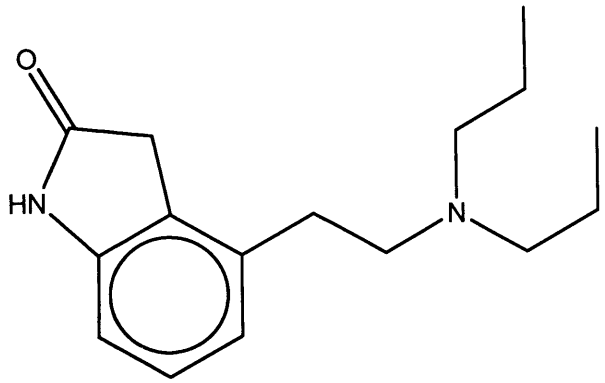
SKB42 (C22H27N3O2)



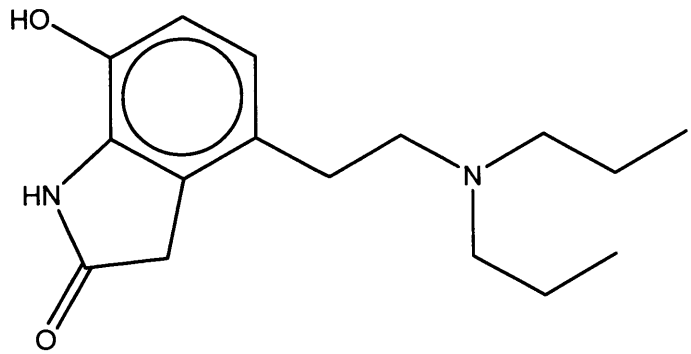
SKB45 (Temelastine)



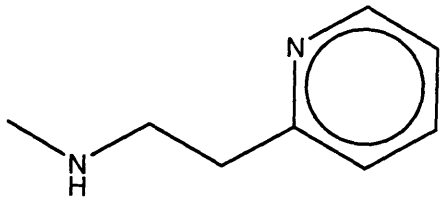
SKF 101468



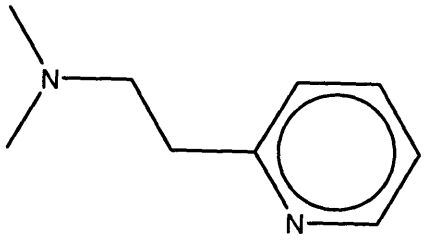
SKF 89124



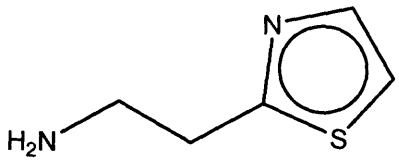
Y-G 14



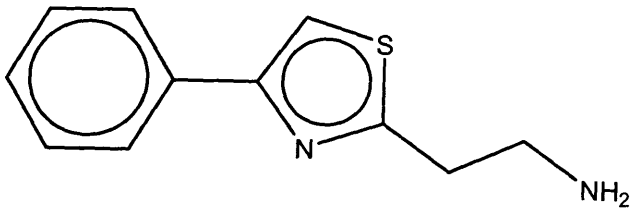
Y-G 15



Y-G 16



Y-G 19



Y-G 20