A scaledown system for the rapid development of commercial inclusion body protein refold steps

A thesis submitted to the University of London
for the degree of
ENGINEERING DOCTORATE

by

James Philip Myers

September 2004

Department of Biochemical Engineering
University College London
Torrington Place
London WC1E 7JE
Abstract

Many therapeutic proteins, currently in development or manufacture, are expressed as inclusion bodies. Proteins must be refolded from the inclusion body into its active form. This refolding step can be difficult to develop and can give a low yield. This project was designed, in partnership with Avecia Biologics, with the purpose of improving the way in which commercial protein refold steps are developed in order to reduce development costs, manufacturing costs and time-to-market of new protein pharmaceuticals.

Development of protein refold steps for commercial manufacture was investigated. Methods for quantifying refolding yield were evaluated. It was determined that refold development experiments could best be improved by decreasing the amount of material required for experiments, increasing throughput of the experiments and designing the experiment efficiently. Techniques for performing refold reactions in microwells were developed and their usefulness was demonstrated with experiments to improve a refold step for IGF-1. Experiments using only 60mg of protein gave a refold step yield of 50% (up from 26%) and a 4-fold decrease in refold volume, using 20%v/v propylene glycol, 0.25M arginine, 0.25M GdnHCl as refold buffer additives.

Techniques were developed for performing rapid refold development experiments using laboratory automation, in particular a pipetting robot. The efficiency of automated microwell refold reactions was demonstrated with refold optimisation experiments, using lysozyme as an example protein. These experiments, which used only 130mg of protein and could be performed by one scientist in just two days, identified refold conditions giving a 58% increase in step yield. Another automated microwell-scale refold experiment using trypsinogen as the example protein established (in 1 day, using 3.1mg protein) that a 13% increase in yield could be achieved by including 20mM PEG300 and 10mM pentanol in the refold buffer. The application of these techniques in a commercial bioprocess development laboratory was further considered.
Acknowledgements

I would like to thank my supervisor, Nigel Titchener-Hooker for his invaluable guidance, help and patience throughout this project. Thanks also go to my industrial supervisors, John Liddell and John Hinton at Avecia, for their insights and efforts. I would also like to thank my advisor, Paul Dalby and all of the other staff at UCL Department of Biochemical Engineering for their help.

I’m very grateful to the Engineering and Physical Sciences Research Council and to Avecia Biologies for their funding. I am also grateful to all those people at Avecia who have helped during this EngD. Particular thanks go to Ian Hodgson and Matt Goodwin for their cooperation with the IGF-1 work, Allan Watkinson and Simon Walker for their antigen denaturation fluorescence data, Helen Briggs and Claire Johnston for taking me on for the commercial project.

I would also like to thank my MSc students at UCL: Andrew Roland-Price for his work on dye-binding, Disha Sharma and Tim Chow for their efforts with different scales of refolding reactors and Ihsan Al-Marzouki for his work on bench scale mixing of refolding reactions.

Thanks to all my friends and colleagues at the department who taught me how to use everything from rubber gloves to robots and who kept me sane, but not too sober, throughout my time there. Thanks also to Paul, Paul, Dave, Ben, Swin and, of course, John for the additional support, beer, fun and fine dining when it was needed most (usually lunchtime). Finally, I would like to thank my family for their practical and moral support.
Table of contents

LIST OF TABLES ............................................................................................................. 13
LIST OF FIGURES ....................................................................................................... 14
NOMENCLATURE ...................................................................................................... 18

1 INTRODUCTION ...................................................................................................... 22
  1.1 Introduction to inclusion body processes ......................................................... 22
    1.1.1 Significance of inclusion body protein refolding in therapeutic
    protein manufacture ............................................................................................. 22
    1.1.2 Advantages and disadvantages of inclusion body processes ................. 23
  1.2 Methods for production of proteins from inclusion bodies ......................... 24
    1.2.1 Release of inclusion bodies from cells ..................................................... 24
    1.2.2 Solubilisation and denaturation using high chaotrope
    concentration ........................................................................................................ 25
    1.2.3 Dilution refolding ...................................................................................... 25
    1.2.4 Dialysis or diafiltration refolding ............................................................. 27
    1.2.5 Column bound refolding ........................................................................... 28
      1.2.5.1 Ion exchange media bound refolding ............................................... 29
      1.2.5.2 Immobilised metal affinity media bound refolding ......................... 30
      1.2.5.3 Cellulose bound refolding ................................................................. 31
      1.2.5.4 Polyanionic matrix bound refolding ................................................. 31
    1.2.6 Size exclusion chromatography refolding .............................................. 31
    1.2.7 Other methods of solubilisation and refolding ....................................... 32
      1.2.7.1 High Hydrostatic Pressure ................................................................. 32
      1.2.7.2 pH change refolding ......................................................................... 33
    1.2.8 Techniques to enhance refolding yield ................................................. 33
  1.3 Protein refolding in commercial manufacture ............................................. 34
    1.3.1 Solubilisation and denaturation ............................................................... 35
    1.3.2 Refolding .................................................................................................... 38
  1.4 The need for parallel scaledown experiments for rapid
  development of refold steps ................................................................................. 43
1.5 The aims of the project ..................................................................45

2 DEVELOPMENT OF PROTEIN REFOLDING STEPS .......................48
2.1 Aims of refolding development experiments ....................................48
2.2 The objective function for optimisation ............................................51
2.3 Variables for refold optimisation (refold step parameters which can affect performance) ..........................................................53
  2.3.1 pH ...........................................................................................53
  2.3.2 Physical factors .......................................................................54
  2.3.3 Dilution factor and protein concentration .....................................54
  2.3.4 Refolding enhancing additives ...................................................55
  2.3.5 Redox agents for refolding ......................................................60
2.4 Regulatory compliance and cost driven changes in the scope of experimental work during a drug's development ..................................................63
2.5 The application of Design of Experiments (DOE) techniques to protein refolding experiments ..........................................................67
2.6 Conclusions ...................................................................................71

3 MEASURING THE YIELD OF THE REFOLDING REACTION IN REFOLD DEVELOPMENT EXPERIMENTS ..........................73
3.1 Potential problems with refolding yield calculation .........................73
3.2 Alternative approaches to refold yield calculation for development experiments .............................................................................75
3.3 Analytical methods for the analysis of refolded protein in refold development experiments ..........................................................77
3.4 Activity Assays ................................................................................78
  3.4.1 Lysozyme activity assay .............................................................80
  3.4.1.1 Materials and methods .......................................................80
  3.4.1.2 Results ..................................................................................81
  3.4.1.3 Discussion and Conclusions ...............................................82
3.5 Absorbance methods ........................................................................82
  3.5.1 Absorbance at 280nm (measuring total soluble protein) ..........82
  3.5.2 Turbidity (measuring precipitation) ..............................................83
3.6 Correlation between turbidity and recovered activity in refolded lysozyme .................................................................84

3.6.1 Method .............................................................................85
  3.6.1.1 Materials........................................................................85
  3.6.1.2 Preparation of unfolded lysozyme .................................85
  3.6.1.3 Microwell refolding.........................................................85
  3.6.1.4 10ml CSTR refolding.....................................................86
  3.6.1.5 Refolding conditions.......................................................87
  3.6.1.6 Assays........................................................................87

3.6.2 Results ................................................................................88
3.6.3 Discussion and conclusions of the experiment ....................89

3.6.4 Conclusions on absorbance methods.................................90

3.7 Chromatographic methods ....................................................90
  3.7.1 Size-exclusion chromatography ......................................90

3.8 Relationship between size exclusion chromatography and recovered activity in refolded lysozyme........................................91
  3.8.1 Materials and methods ....................................................92
  3.8.2 Results and discussion ....................................................94
  3.8.3 Adsorptive chromatographic methods..............................95
  3.8.4 Improvements in the throughput of an RP-HPLC assay for IGF 96

3.9 Fluorescence spectrometry methods ........................................98
  3.9.1 Fluorescent dye binding....................................................99
  3.9.2 Intrinsic fluorescence.......................................................100
  3.9.3 Experiment to determine the variation in intrinsic fluorescence of BP3 in different concentrations of chaotrope ..........................102
    3.9.3.1 Materials and methods .............................................102
    3.9.3.2 Results and discussion .............................................103
    3.9.3.3 Conclusions of the experiment .................................104
  3.9.4 Change in the intrinsic fluorescence of lysozyme upon refolding 105
    3.9.4.1 Materials and methods .............................................105
      3.9.4.1.1 Materials ..........................................................105
3.9.4.1.2 Fluorimeter control .....................................................105
3.9.4.1.3 Lysozyme refolding ....................................................107
3.9.4.2 Results ...........................................................................107
3.9.4.3 Discussion and Conclusions ...........................................108
3.10 Conclusions of the chapter ..........................................................109

4 THE USE OF MICROWELL PLATES FOR PROTEIN REFOLDING EXPERIMENTS ..............................................................................................................112
4.1 Rationale for using microwells as reaction vessels.............112
4.2 Practicalities of performing refolding experiments in microwells.114
  4.2.1 Performing the operations of a dilution refold reaction in a microwell plate ....................................................................................114
    4.2.1.1 Preparation of unfolded protein solution .........................114
    4.2.1.2 Preparation of refold buffer .............................................115
    4.2.1.3 Addition of unfolded protein to refold buffer .....................116
    4.2.1.4 Mixing of unfolded protein and refold buffer .....................117
  4.2.2 Control of variables affecting refolding ....................118
    4.2.2.1 Composition of refold buffer, unfolded protein solution and refolding protein concentration .......................................................118
    4.2.2.2 Mixing .............................................................................119
    4.2.2.3 Temperature ...................................................................120
    4.2.2.4 pH ...................................................................................121
    4.2.2.5 Refold time .....................................................................121
4.3 Characterisation of the effect of mixing methods on refold yield in microwells ...............................................................................................123
  4.3.1 Method ..................................................................................123
    4.3.1.1 Materials .........................................................................123
    4.3.1.2 Microwell protein refolding .............................................124
    4.3.1.3 Diffusion (no mixing) .......................................................124
    4.3.1.4 Magnetic stirrer mixing ...................................................124
    4.3.1.5 Pipette mixing .................................................................124
    4.3.1.6 Orbital shaker .................................................................125
    4.3.1.7 Activity assay ..................................................................125
  4.3.2 Results ..................................................................................126
4.3.3 Discussion and conclusions .................................................. 127

4.4 Loss of protein due to non-specific binding during refolding in microwell plates .................................................. 130
  4.4.1 Method .................................................................................. 131
    4.4.1.1 Solutions ........................................................................ 131
    4.4.1.2 Multiwell plates ............................................................... 131
    4.4.1.3 Refolding ........................................................................ 131
    4.4.1.4 Measurement of unbound protein concentration .......... 132
  4.4.2 Results .................................................................................. 132
  4.4.3 Discussion and Conclusions .................................................. 134

4.5 Scalability of refolding reactions from microwells ......................... 136
  4.5.1 Method .................................................................................. 137
    4.5.1.1 Materials ......................................................................... 137
    4.5.1.2 Preparation of unfolded lysozyme ................................. 138
    4.5.1.3 Refolding buffers and stock solutions ............................. 138
    4.5.1.4 Microwell plate refolding ................................................. 139
    4.5.1.5 10ml CSTR refolding ...................................................... 140
    4.5.1.6 200ml CSTR refolding ................................................... 140
    4.5.1.7 Activity assay .................................................................. 141
  4.5.2 Results .................................................................................. 141
  4.5.3 Discussion and conclusions of the experiment ..................... 147

4.6 Conclusions of the chapter .......................................................... 149

5 APPLICATION OF MICROWELL PLATE REFOLDING EXPERIMENTS: DEVELOPMENT OF THE PROTEIN REFOLDING STEP FOR THE PRODUCTION OF IGF-1 FROM E.COLI INCLUSION BODIES 150

5.1 Background .................................................................................. 150

5.2 Materials and Methods ................................................................. 151
  5.2.1 Chemicals .............................................................................. 151
  5.2.2 Microwell refold method ......................................................... 151
  5.2.3 Refold buffers ........................................................................ 152
  5.2.4 Additive screening experiment .............................................. 152
  5.2.5 Ethanol substitute experiment .............................................. 153
7.1.1.1 Materials .........................................................................187
7.1.1.2 Preparation of unfolded lysozyme ..................................187
7.1.1.3 Refolding buffers and stock solutions .........................188
7.1.1.4 Experimental design .......................................................189
7.1.1.5 Automated refolding reactions ........................................191
7.1.1.6 Bench scale refolding reactions ......................................192
7.1.1.7 Automated lysozyme activity assays .............................192

7.1.2 Results ..................................................................................195
7.1.2.1 Main effects of reported refolding enhancers on the refolding yield of lysozyme ..............................................................195
7.1.2.2 Two-factor interactions of reported refolding enhancers on the refolding yield of lysozyme ........................................................197
7.1.2.3 Response surface optimisation ......................................201

7.1.3 Discussion and conclusions of the experiment ....................203
7.1.3.1 Effects of refolding enhancers ........................................203
7.1.3.2 Experimental efficiency ..................................................204

7.2 Screening refolding enhancers to improve the refolding efficiency of trypsinogen using automated experiments in microwells ...................206

7.2.1 Materials and Methods ..........................................................207
7.2.1.1 Preparation of denatured trypsinogen ...........................207
7.2.1.2 Refolding buffers and stock solutions ............................208
7.2.1.3 Experiment design ..........................................................209
7.2.1.4 Automated refolding reactions .......................................209
7.2.1.5 Trypsinogen activation and assay .................................210

7.2.2 Results ..................................................................................210
7.2.2.1 Main effects of reported refolding enhancers on the refolding yield of trypsinogen ..............................................................211
7.2.2.2 Two-factor interactions of reported refolding enhancers on the refolding yield of trypsinogen ........................................................213
7.2.2.3 Significance of effects and Pareto chart .........................215
7.2.2.4 Experiment efficiency .....................................................217

7.2.3 Discussion and conclusions of the experiment ....................217

7.3 Discussion and conclusions of the chapter .................................218
8 THE APPLICATION OF AUTOMATED MICROWELL EXPERIMENTS FOR COMMERCIAL REFOLD DEVELOPMENT..........................219

8.1 The uses of automated microwell experiments in commercial dilution refold development........................................................................................................220
  8.1.1 Refold step creation...........................................................................221
  8.1.2 Screening for significant factors.....................................................223
  8.1.3 Optimisation..................................................................................224

8.2 Illustrations of the benefits of performing process development experiments using automated microwell refold reactions.................................225
  8.2.1 Lysozyme and trypsinogen refold development experiments using automated microwell refold reactions: time-efficiency and material-efficiency benefits.............................................................225
  8.2.2 An algorithm for rapid refold development using automated microwell scale experiments..............................................................227
  8.2.3 Comparison with conventional commercial refold step development........................................................................................................230

8.3 Scheduling of experiments..................................................................237
8.4 Capital equipment requirements..........................................................240
8.5 Procedures (software) for a unified system for rapid scaledown refold development..........................................................................................244
  8.5.1 Experiment design, scheduling and analysis....................................245
  8.5.2 Refold buffer planning.................................................................245
  8.5.3 System coordination......................................................................246

8.6 Conclusions of the chapter..................................................................246

9 COMMERCIAL AND FINANCIAL ISSUES IN A BIOPROCESS DEVELOPMENT AND MANUFACTURING BUSINESS..........................248

9.1 Introduction to the company ................................................................248
9.2 Background..........................................................................................249
9.3 Project costs, cost predictions and pricing............................................249
9.4 Project cashflow..................................................................................251
9.5 Outcomes............................................................................................253

10 CONCLUSIONS....................................................................................254
List of Tables

Table 2-1  Table of chemicals which have been reported to enhance the refolding of proteins. .................................................................57
Table 2-2  Table of redox reagents used in protein refolding.............62
Table 2-3  Typical stages in the development of a dilution refold step.....64
Table 3-1  Table of nine refold buffers used .....................................87
Table 4-1  Table of averages and standard deviations (from eight replicates) of yields obtained from step dilution refolds mixed by different methods in microwells. .........................................................126
Table 4-2  Composition of nine refold buffers. .................................139
Table 4-3  Table of correlation coefficients between data sets ..........146
Table 5-1  Additives tested in screening experiment ..........................153
Table 5-2  Additives tested in screening experiment ..........................153
Table 6-1  Table showing the aspirate-dispense flowrate and cycle times used to mix refolds in microwells. ......................................179
Table 7-1  Table of reported refolding enhancers used .....................189
Table 7-2  Table of reported refolding enhancers used .....................208
Table 8-1  Benefits of using automated microwell-scale refold experiments ..................................................................................236
List of figures

Figure 3-1 Calibration curve for lysozyme activity assay. Four replicates of each point were run. Mean activity values are plotted, error bars show ±1 standard deviation. A linear line of best fit was fitted, the equation and R² value are shown .................................................................81

Figure 3-2 Graph showing the relationship between turbidity of lysozyme solutions refolded in microwells and the activity of lysozyme solutions refolded in 10ml CSTRs under similar conditions. Each point represents the mean activity of three 10mL scale refolds and the mean turbidity of three microwell scale refolds, with error bars showing ±1 standard deviation. The points are labelled with the number of the refold condition to which they refer (see Table 3-1). Linear line of best fit to mean values is shown, with equation and R² value ......................... 88

Figure 3-3 Example of a SEC chromatogram of a sample of refolded lysozyme ...........................................................................................................93

Figure 3-4 Graph showing the relationship between SEC peak area and recovered activity of lysozyme solutions. Each point represents one SEC assay and the mean of three activity assays, with error bars showing ± one standard deviation. Linear line of best fit to mean values is shown, with equation and R² value .......................................94

Figure 3-5 Graph showing the maximum emission wavelength of a recombinant antigen as a function of urea concentration. .................... 101

Figure 3-6 Graphs showing the variation in magnitude and wavelength of the emission maximum of solutions of rHGF1BP3 in solutions containing different concentrations of urea and guanidine hydrochloride. .................................................................................................103

Figure 3-7 Graph showing the change in the fluorescence emission spectrum of a refolding lysozyme solution with time after the initiation of refolding, with excitation at 280nm .................................................. 107
Figure 4-1 Graph of yields obtained from step dilution refolds mixed by different methods in microwells .......................................................... 126

Figure 4-2 Graphs showing the reduction in concentration of unbound protein (both soluble and precipitated) with time. Plotted value is the concentration of unbound protein, relative to the initial concentration. ............................................................................................................ 133

Figure 4-3 Comparison plot showing the yield of recovered activity from dilution refold steps carried out under similar conditions at microwell and 10ml scales. Each point represents one condition. The yield of refolds at those conditions at microwell scale is plotted on the x-axis, the yield of refolds at those conditions at 10ml scale is plotted on the y-axis. ...................................................................................................... 143

Figure 4-4 Comparison plot showing the yield of recovered activity from dilution refold steps carried out under similar conditions at microwell and 200ml scales. Each point represents one condition. The yield of refolds at those conditions at microwell scale is plotted on the x-axis, the yield of refolds at those conditions at 200ml scale is plotted on the y-axis. ................................................................................................... 144

Figure 4-5 Comparison plot showing the yield of recovered activity from dilution refold steps carried out under similar conditions at 10ml and 200ml scales. Each point represents one condition. The yield of refolds at those conditions at 10ml scale is plotted on the x-axis, the yield of refolds at those conditions at 200ml scale is plotted on the y-axis. ...................................................................................................... 145

Figure 4-6 Graph of the yield of recovered activity from dilution refold steps carried out under similar conditions at microwell, 10ml and 200ml scales. .................................................................................................. 146

Figure 5-1 Standard curve: graph of peak magnitude against standard concentration. Linear line of best fit, with equation and $R^2$ value, is shown. ........................................................................................................ 156
Figure 5-2 Screening experiment results. Graph showing the concentration of correctly folded IGF-1 in solutions refolded using different buffers.

Figure 5-3 Ethanol substitute experiment results. Graph showing the concentration of correctly folded IGF-1 in solutions refolded using refold buffers containing different alcohols. Each bar represents the yield of correctly folded IGF-1 from one refold reaction.

Figure 5-4 Ethanol substitute experiment results. Graph showing the concentration of correctly folded IGF-1 in solutions refolded using refold buffers containing different alcohols and other additives.

Figure 5-5 Refold concentration results. Graph of correctly folded IGF-1 concentration in solutions refolded at different IGF-1 concentrations. Lines shown are second order polynomials fitted to data.

Figure 5-6 Refold concentration results. Graph showing yield of correctly refolded IGF-1 concentration in solutions refolded at different IGF-1 concentrations.

Figure 6-1 Yields from step dilution refolds mixed in a microwell plate by a pipetting robot under different conditions (see Table 6-1 for details of conditions). Each point represents the yield of one refold reaction.

Figure 6-2 Yields from step dilution refolds mixed in a microwell plate by a pipetting robot under different conditions (see Table 6-1 for details of conditions). Each point represents the yield of one refold reaction.

Figure 7-1 Standard curve for automated lysozyme assay (6 replicates at each concentration). Mean and median activity values for each concentration are plotted. Linear lines of best fit to both the mean and median values are shown, with equations and R² values.

Figure 7-2 Main effects of reported refolding enhancers on the enzyme activity of the refolded lysozyme solution.
Figure 7-3  Two-factor interactions of reported refolding enhancers on the refolding yield of lysozyme.................................................................199

Figure 7-4  Two-factor interactions of arginine, glucose and sucrose on the refolding yield of lysozyme.................................................................200

Figure 7-5  Contour plot showing the response of refolded lysozyme activity to different refolding buffer compositions. Black circles indicate refolding buffer compositions which were tested in the experiment. The scales on the three sides indicate the proportions of the three stock solutions which comprised the refold buffers used for each condition. ........................................................................................................202

Figure 7-6  Wireframe surface plot showing the response of refolded lysozyme activity to different refolding buffer compositions. The scales on the three sides of the base indicate the proportions of the three stock solutions which comprised the refold buffers used for each condition. ........................................................................................................202

Figure 7-7  Times required to perform a similar set of refolding optimisation experiments. ........................................................................................................205

Figure 7-8  Main effects of reported refolding enhancers on the activity of the refolded trypsinogen solution...............................................................212

Figure 7-9  Two-factor interactions of reported refolding enhancers on the refolding yield of trypsinogen.................................................................214

Figure 7-10  Pareto chart of main and 2-factor interaction effects of eleven reported refolding enhancers on the refolding yield of trypsinogen....216
Nomenclature

%v/v  percentage by volume
°C  degrees Celsius
2D  two dimensional
3D  three dimensional
A280  absorbance at 280nm
ANS  8-anilino-1-naphthalenesulphonate
AU  absorbance unit
BP3  human insulin-like growth factor 1 binding protein 3
CBD  cellulose binding domain
CD  circular dichroism
cGMP  current good manufacturing practice
CHAPS  3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulphonate
cm  centimetre
CMO  contract manufacturing organisation
COSHH  containment of substances hazardous to health
CSTR  continuously stirred tank reactors
CV  column volume
Cys^n  the cystine residue, being the n\textsuperscript{th} amino acid residue in a protein
DOE  design of experiments
DSP  downstream purification
DTE  dithioerythreitol
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>E.coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>EBITDA</td>
<td>earnings before interest, tax, depreciation and amortisation</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylene diamine tetraacetic acid</td>
</tr>
<tr>
<td>ELISA</td>
<td>enzyme-linked immunosorbant assay</td>
</tr>
<tr>
<td>EngD</td>
<td>engineering doctorate</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>GdnHCl</td>
<td>guanidine hydrochloride</td>
</tr>
<tr>
<td>GLP</td>
<td>good laboratory practice</td>
</tr>
<tr>
<td>GPC</td>
<td>gel permeation chromatography</td>
</tr>
<tr>
<td>GSSG</td>
<td>oxidised glutathione</td>
</tr>
<tr>
<td>HETP</td>
<td>height equivalent to theoretical plate</td>
</tr>
<tr>
<td>HEWL</td>
<td>hen egg white lysozyme</td>
</tr>
<tr>
<td>HIC</td>
<td>hydrophobic interaction chromatography</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>IB</td>
<td>inclusion body</td>
</tr>
<tr>
<td>IEC</td>
<td>ion-exchange chromatography</td>
</tr>
<tr>
<td>IGF-1</td>
<td>insulin-like growth factor 1</td>
</tr>
<tr>
<td>IMAC</td>
<td>immobilised metal affinity chromatography</td>
</tr>
<tr>
<td>kBar</td>
<td>kilobar</td>
</tr>
<tr>
<td>kg</td>
<td>kilogram</td>
</tr>
<tr>
<td>L</td>
<td>litre</td>
</tr>
<tr>
<td>L-BAPNA</td>
<td>N\textsubscript{\alpha}-benzoyl-L-arginine-p-nitroanilide</td>
</tr>
<tr>
<td>LPQ</td>
<td>laboratory process qualification</td>
</tr>
<tr>
<td>M</td>
<td>molar</td>
</tr>
</tbody>
</table>
m metre
mg milligram
min minute
ml millilitre
mM millimolar
mm millimetre
nm nanometres
OVAAT one variable at a time
PEG polyethylene glycol
PEG300 polyethylene glycol of average molecular weight 300 atomic mass units
PEG3350 polyethylene glycol of average molecular weight 3350 atomic mass units
PTFE polytetrafluoroethylene
QA quality assurance
QC quality control
R&D research and development
RM raw material
RPC reverse phase chromatography
rpm revolutions per minute
s second
SDS- sodium dodecyl sulphate polyacrylamide gel electrophoresis
PAGE
SEC size exclusion chromatography
SHE safety, health and environment
TCEP tris(2-carboxyethyl)phosphine
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tris</td>
<td>Tris (hydroxymethyl)-aminomethane</td>
</tr>
<tr>
<td>$\lambda_{\text{max}}$</td>
<td>wavelength of maximum emission magnitude</td>
</tr>
<tr>
<td>$\mu g$</td>
<td>microgram</td>
</tr>
<tr>
<td>$\mu l$</td>
<td>microlitre</td>
</tr>
<tr>
<td>$\mu m$</td>
<td>micrometre</td>
</tr>
</tbody>
</table>
1 Introduction

In this chapter, inclusion body protein refolding will be introduced. The position of inclusion body processes in the biopharmaceutical industry will be described. Methods for the production of active (folded) protein from inclusion bodies will be discussed and the operation of such refolding steps in commercial manufacture will be considered. The need for parallel, scaledown experiments in the development of commercial refolding steps will be discussed and the aims of this research project will be set out.

1.1 Introduction to inclusion body processes

1.1.1 Significance of inclusion body protein refolding in therapeutic protein manufacture

The value of the market for therapeutic proteins (pharmaceutical products which comprise a protein as an active ingredient) was calculated at $32 billion in 2003 and predicted to be $53 billion in 2010 (Datamonitor, 2004).

Recombinant DNA technology enables the production of heterologous proteins in microbial hosts, such as *E. coli*. This provides an economically attractive method of manufacturing therapeutic proteins. Unfortunately, high expression of recombinant proteins in microbial hosts often results in the product being produced in insoluble inclusion bodies within the cell (Rudolph and Lilie 1996). Inclusion bodies (IBs) form in the host cell as the protein is expressed. The expressed species bind together, forming solid lumps which can grow to the diameter of the cell. The product protein in these inclusion bodies is in an inactive, insoluble form and so must be solubilised and refolded into an active form during processing.
1.1.2 Advantages and disadvantages of inclusion body processes

During the development of a process to produce a protein product, it may be found that expression in inclusion bodies is the only possible system. Alternatively, a system expressing protein in inclusion bodies may be selected because it offers certain advantages over soluble (intracellular and excreted) expression. The advantages and disadvantages of inclusion body systems are described below.

Advantages:

- Proteins can be expressed in IBs to relatively high titres (Rudolph and Lilie 1996).
- High levels of a protein which is toxic to the host cell can be expressed because the protein is in an inactive form in the IB (Clark 2001).
- While in the IB, the protein has a high degree of protection against proteolysis (Clark 2001).
- The protein in the inclusion body is typically 40% - 80% pure and the inclusion body is typically large and dense. Consequently, centrifugation and washing of the inclusion bodies from cell homogenate offers a simple and effective first purification step (Cario & Villaverde 2002, Lilie et al 1998, Middelberg 2002).
- Separated inclusion bodies, when centrifuged down to a thick slurry or paste, contain very high concentrations of product protein in a stable form. It is therefore relatively easy to store in deep freeze and so can be stored for long periods. This allows a break (or "decoupling") in processing which can greatly facilitate scheduling in manufacturing facilities.

Disadvantages:

- Inclusion body processes require process steps which are not required by other systems. Systems expressing excreted protein do not require cell breakage (lysis or homogenisation) steps. Systems expressing
soluble protein do not require solubilisation and refolding steps. These additional steps can significantly increase processing costs.

- Inclusion body processes may have lower yields than processes with soluble expression. The yield of refolding reactions may be particularly low (see Chapter 3 for a discussion of refolding yields).

1.2 Methods for production of proteins from inclusion bodies

Methods for the production of folded proteins from inclusion bodies will be discussed in this section. Particular attention is paid to refold methods which have been used in the commercial manufacture of therapeutic protein products. (The use of these methods in commercial large scale manufacture are further discussed in section 1.3)

The most commonly used method (homogenisation of cells, centrifugation to separate inclusion bodies, denaturation by high concentration of chaotrope then refolding by reduction of chaotrope concentration through dilution) is discussed first. Other ways of effecting the reduction in chaotrope concentration for refolding are then presented. Alternative methods of solubilisation and refolding are outlined. Techniques for increasing the yield of refold steps are then analysed.

1.2.1 Release of inclusion bodies from cells

After fermentation inclusion bodies are released from the cells using either chemical or mechanical disruption means (Middelberg 2002). The most commonly used method is to disrupt the cells using a high-pressure homogeniser. The inclusion bodies are then separated and washed in a centrifuge to remove impurities. The washed inclusion bodies are then solubilised and refolded using the methods described below.
1.2.2 Solubilisation and denaturation using high chaotrope concentration

The separated inclusion bodies need to be solubilised to separate the protein molecules. The protein to be refolded (typically in the form of inclusion bodies) is added to a strong chaotrope such as 8M urea or 6M GdnHCl, to a protein concentration of typically 1-10mg/ml. The chaotrope disrupts the hydrophobic interactions which bind the protein molecules together, and which hold the protein in its shape (Rudolph and Lilie 1996). Consequently, inclusion bodies dissolve and the protein unfolds (denatures). If the inclusion body protein has disulphide bonds, a reducing agent such as 25mM DTT may be added to reduce (break) the disulphide bonds.

Strong detergents such as sodium dodecyl sulphate may be used, with or instead of urea or GdnHCl, for solubilisation and denaturation (Clark 2001). Detergents work in a different way to urea or GdnHCl, and solubilised protein may retain activity and some native structure. However, it is reported that the use of detergents may cause problems in manufacture. They can interfere with downstream chromatography and ultrafiltration steps, even after dilution. Assays for detergents can be problematic, which can make it difficult to validate their removal during purification.

1.2.3 Dilution refolding

The solution containing the denatured protein and chaotrope is then diluted into a refolding buffer to reduce the concentration of chaotrope. As the chaotrope concentration falls, the hydrophobic interactions reassert themselves, pulling the protein chain back into a structure (refolding). The protein first regains lost secondary structure, contracting from an extended chain to a compact globule. This is a very fast reaction – it has been observed to occur in less than 0.5seconds (Cleland and Wang, 1990). The protein then begins to regain its tertiary structure, domains of the protein moving relative to each other until they reach a stable structure, which is a
much slower reaction. Upon the reduction of chaotrope concentration, the protein rapidly assumes an intermediate state and then slowly reacts to reach its native state.

An oxidising agent such as 25mM cystine or 10mM oxidised glutathione may be included the refolding buffer to re-form disulphide bonds. Proteins are usually refolded at around pH8. The slightly alkaline conditions promote thiolate anion formation and so facilitates disulphide bond shuffling. As the protein folds up from its unfolded state, incorrect arrangements of disulphide bonds can form. These arrangements of disulphide bonds need to be shuffled into the correct arrangement for the protein to refold correctly.

Not all protein molecules will fold back into their active form. The native structure of a protein is typically the most thermodynamically stable structure for that protein. However, there is a degree of randomness in the refolding process and if the protein reaches a thermodynamically stable non-native state before it reaches its native state then it will stay in that non-native state. This results in a misfold – the protein fails to refold into the native state.

The hydrophobic interactions which refold the protein molecules may also cause those protein molecules to aggregate. A precipitate of aggregated protein often forms in the refolding solution. Proteins are most prone to aggregation after dilution into the refold buffer, when the chaotrope concentration is reduced, but before they become fully refolded. When fully folded, most of the hydrophobic sites on a soluble protein are buried, so the protein is much less prone to aggregation. The kinetics of aggregation have been shown to be approximately second order, in contrast to the kinetics of refolding which have been observed to be first order (Kiefhaber et al. 1991, Zettlemeissl et al. 1979). Low protein concentration during refold therefore favours refolding over aggregation, giving a higher refold yield.
The conditions in which the protein refolds can affect how likely it is to misfold, how likely it is to aggregate, and how likely it is to reach its native state. Refolding conditions can be optimised to minimise the amount of misfolding and aggregation, and so to maximise yield of correctly folded protein (see Chapter 2).

1.2.4 Dialysis or diafiltration refolding

In this method, the protein to be refolded is denatured in a strong chaotrope, as described in 1.2.2. The concentration of chaotrope in the denatured protein solution is then reduced using dialysis or diafiltration, causing the protein to refold.

Dialysis or diafiltration refolding is an attractive method for proteins which do not tend to aggregate during refold. This method does not reduce the concentration of the protein when it reduces the concentration of chaotrope. This has the advantage of producing a more concentrated protein solution with smaller process volumes than dilution refolding, giving lower processing costs. However this method reduces chaotrope concentration much more slowly than dilution refolding, and so the protein (at high concentrations) is exposed to intermediate chaotrope concentrations for significant lengths of time, which can lead to high levels of aggregation (London et al 1974).

Other problems with diafiltration refolding include binding of the refolding protein to membrane surface, precipitate fouling of the membrane, heating of the protein solution during pumping and higher equipment-related costs compared to dilution refolding.
1.2.5 Column bound refolding

Unfolded proteins can be immobilised on a solid support such as chromatography media, refolded while still bound, then eluted in the refolded form. Li et al (2003) have produced a thorough review article on chromatographic refold methods. The column-bound refolding methods which have the greatest potential for large scale manufacture are discussed here.

In column bound refolding methods, the protein to be refolded is denatured in a denaturing buffer, as described in section 1.2.2. The denaturing buffer must also be a suitable buffer for loading protein onto the solid support.

The solid support (typically a chromatography column) is equilibrated with denaturing buffer, then the denatured protein solution is loaded. The denatured protein binds to the solid support which is then washed with more of the denaturing buffer to remove impurities.

The chaotrope concentration in the column is then reduced in either of two ways:

Stepwise A low (or zero) chaotrope concentration refolding buffer is run down the column

Gradient A low (or zero) chaotrope concentration refolding buffer is prepared. A gradient is run down the column (in a similar way to a gradient elution), starting at 100% denaturing buffer, 0% refolding buffer, ending at 0% denaturing buffer, 100% refolding buffer.

The protein remains bound to the column during this buffer change. As the chaotrope concentration on the column falls, the protein refolds while still bound to the column in a similar manner to that described in section 1.2.3. The protein is left bound to the column while it refolds. The refolded protein is then eluted from the column in a similar way to normal chromatography.
Using either variant of this method, the protein remains immobilised on the column during refolding, consequently aggregation is inhibited and this method can give significantly higher yields compared to dilution refolding. This method of refolding can also provide a similar level of purification to a normal chromatography purification/capture step. The protein can be eluted from the column in a relatively concentrated form compared to dilution refolding.

A disadvantage of this method is that the denatured protein solution must be clarified by centrifugation or filtration prior to loading. Denatured protein solutions are typically quite viscous, so back pressure and/or low flow rates can be problems during filtration and loading. The denaturing buffer used during washing may also be quite viscous. Denaturing buffers may be expensive, so washing a column using volumes of them can be costly. The binding of the protein to the column may prevent correct folding. Any precipitates which form during loading, washing or refolding can foul the column, which may make it necessary to unpack the column and will reduce matrix lifetimes.

A number of different types of solid support may be used for this method. The use of these different matrices are discussed in the following subsections.

1.2.5.1 Ion exchange media bound refolding

In this variant of the method the unfolded protein is immobilised on ion exchange chromatography (IEC) media, refolded while still bound, then eluted in the refolded form (Creighton 1986).

The protein to be refolded is solubilised and loaded in a low ionic strength solution of urea (GdnHCl solutions are not used for this method, due to their high ionic strength.) After the urea concentration on the column has
been reduced (by applying a low ionic strength refolding buffer), the refolded protein is eluted from the column using a high ionic strength elution buffer or salt gradient as in normal ion exchange chromatography.

An alternative to this refolding and elution strategy is to do both in a single gradient. That is, to simultaneously reduce the chaotrope concentration and increase the salt concentration e.g. by running a gradient between denaturing buffer and a high ionic strength refolding buffer (Li, Zhang, Su 2002).

\subsection*{1.2.5.2 Immobilised metal affinity media bound refolding}

In this variant of the method, the unfolded protein is immobilised on immobilised metal affinity chromatography (IMAC) media, refolded while still bound, then eluted in the refolded form (Zouhar et al 1999). This method is typically used when the protein is expressed with a polyhistidine tag on one terminus, which binds particularly well to IMAC media.

The protein is denatured, loaded and refolded as described in section 1.2.5. After the denaturant concentration is reduced, the protein is then eluted from the column either using normal IMAC elution buffer, or by running an elution buffer containing an enzyme to cleave the polyhistidine tag from the protein so that the cleft protein elutes, leaving the tag behind.

This method of refolding can also provide a similar level of purification to a normal IMAC purification/capture step. Reducing agents used in the denatured protein solution to cleave disulphide bonds may cause reduction of the column metal ions, and so will need to be removed or oxidised prior to loading. Some metal ions, e.g. Cu, may catalyse the formation (oxidation) of incorrect disulphide bonds. Metals from the column can escape into the eluate, and validating their removal from the product may be expensive. If a protein has been produce with a polyhistidine tag to allow IMAC to be used and this tag must be removed then the enzyme to do this
must be produced or purchased, either of which are likely to be very expensive.

1.2.5.3 Cellulose bound refolding

Cellulose binding domain (CBD) from the *Clostridium thermocellum* cellulosome is active at urea concentrations of up to 6M, it may therefore be useful as a tag for column bound refolding (Berdichevsky et al 1999). In this method a protein may be expressed as a fusion with CBD. The protein-CBD fusion is solubilised in urea as described in section 1.2.2. The denatured protein-CBD in 6M urea is then loaded onto a column of beaded cellulose and is refolded and eluted in similar way to the IMAC method (section 1.2.5.2). This method has most of the same advantages and disadvantages as the IMAC method, however cellulose matrices are not vulnerable to reduction by the denaturing buffer, nor will they leach metal into the refolding protein solution.

1.2.5.4 Polyanionic matrix bound refolding

It has been reported that a protein expressed with a hexa-arginine tag, denatured in 8M urea, can be bound to polyanionic matrix (Stempfer et al 1996). In a similar way to IMAC and cellulose-bound refolding, these proteins can then be refolded, while still bound to the matrix, by reducing the urea concentration and then eluted in the folded form. Yields between 20-35% have been reported for two proteins refolded using this method, using loadings between 1-5mg protein per ml resin.

1.2.6 Size exclusion chromatography refolding

The buffer change (chaotrope concentration reduction) required for refolding can be done using a size exclusion chromatography (SEC) column, in much the same way as a desalting step (Werner et al, 1994, Batas and Chaudhuri, 1995).
The protein to be refolded is first denatured in a denaturing buffer, as described in section 1.2.2.

An SEC column is equilibrated in a suitable refolding buffer. Denatured protein solution is then loaded onto the column and eluted using refolding buffer. As the chaotrope concentration in the environment of the denatured protein decreases, the protein refolds as described in 1.2.3.

Variations on this method have been reported in the literature, involving the use of a chaotrope concentration gradient during elution, which increased refolding yield of lysozyme (Gu et al 2001).

SEC refolding can reduce aggregation (and so increase yields) compared to dilution refolding. It achieves this by restricting the movement of proteins while they refold. Another advantage of SEC refolding is that it can give a similar degree of purification to a normal SEC step. However, only a small volume of denatured protein solution can be loaded onto the column in each loading, so large scale processes may require the denatured protein to be added in aliquots, with a number of load-elute cycles.

1.2.7 Other methods of solubilisation and refolding

1.2.7.1 High Hydrostatic Pressure.

It has been reported that very high hydrostatic pressure can be used to both solubilise and refold aggregated proteins, including inclusion bodies. For example, aggregated protein was pressurised up to 2kBar for 24 hours in the presence of 1M GdnHCl. 100% recovery of active protein was reported using protein concentrations of up to 8.7mg/ml (St John et al 1999).

The high yields and high protein concentration reported for this method are appealing, but this method has not been attempted at manufacturing scale –
the costs involved in developing the required GMP pressurisation equipment would be high.

It should be noted that high pressure homogenisation exposes inclusion bodies to pressures up to ~1kBar for short periods. The solubilisation and refolding effect seen at 2kBar in the experiment described above may have some connection to the loss in inclusion body mass which is sometimes seen during repeated high-pressure homogenisation.

1.2.7.2 pH change refolding

Both high pH (Khan et al 1998) and, less commonly, low pH (Gavit and Better 2000) have been used to solubilise inclusion bodies. Protein refolding is then initiated by adjusting the pH back to moderate levels. This method has obvious advantages due to its simplicity and lack of dilution. The solubilisation can be accelerated by the addition of low concentrations of chaotropes, e.g. solubilise at pH12.5 with 2M urea, then refold by lowering pH. However, exposure to extremes of pH for long periods may cause chemical modifications to the protein.

1.2.8 Techniques to enhance refolding yield

The yield of refold steps can be improved by optimising the operating parameters of that step, e.g. protein concentration, buffer composition, temperature etc. (see section 1.3). There are also a number of novel techniques which have been used in laboratory refolding work to improve refolding yield.

Catalysts which aid refolding, such as the E. coli chaperones GroEL and GroES and the disulphide oxidation and shuffling catalyst DbsA can be added to the refolding buffer, giving significant improvements in yield for hard-to-refold proteins. However the high cost of these enzymes requires that they be re-used, and recovery of the enzymes from the process stream is inefficient (Clark 1998). Immobilising these catalysts by covalently
binding them onto a solid support offers a more cost-effective way of using them (Dong et al 2000, Altamirano 1999), though the cost is still significant and would probably be prohibitive for manufacture. Finding a regulation-compliant source for the enzyme may also be problematic.

Some additives, which are also used as refolding enhancers in dilution refolding, have also been shown to enhance refolding when bound to a solid support. Liposomes (Yoshimoto 1999) and PEG (Geng and Chang 1992) have been immobilised onto a matrix and used to enhance refolding yield. It has been speculated that dextrins, cyclodextrins, arginine and short synthetic peptides may also be used as refolding enhancers when immobilised on a solid support such as chromatography media (Li et al 2004).

### 1.3 Protein refolding in commercial manufacture

In this section, the operation of refold steps in commercial process development laboratories and manufacturing facilities will be discussed (the development of such refold steps will be discussed in Chapter 2).

Inclusion body expression, followed by solubilisation and refolding, is widely used in the commercial manufacture of protein products. For example, ribonuclease (Norgen), human growth hormones (Protheon and Lilly), human glucagon (Protheon), plasminogen activator (Roche) and human insulin (Lilly). In addition to those processes that are operated to produce marketed products, a much larger number of processes are developed and operated to produce material for preclinical and clinical trials. For example, Avecia Biologics has manufactured seven therapeutic proteins from inclusion bodies for preclinical and clinical trials in the last five years.

For obvious reasons of commercial confidentiality, there is scarce information in the public domain about the details of the development and
operation of commercial manufacturing processes. The discussions in this section are based upon the limited published information and upon information received from the contract biologics process development and manufacturing companies (CMOs) Avecia Biologics of Billingham, UK (sponsor of this project) and Dow Biopharmaceutical Contract Manufacturing Services (Dowpharma) of Stony Brook, USA. Conversations were held with Avecia staff who had significant experience of inclusion body process development and manufacture (John Liddell, John Hinton, Andrew Topping, Subhash Chaudhary, Eleanor Clements, Michael Burns, Allan Watkinson, Sylvia Grieve, Ray O'Donnel and Ann Harland). Correspondence was exchanged with David Watkins of Dowpharma.

Most commercial refold processes use the chaotrope denaturation - dilution refold method. This is due to a number of factors. The method is simple. It has been used many times, and is usually familiar to the process development and manufacturing staff. It has been found to scale up quite well. The method uses simple equipment (tanks, pumps and stirrers) which are relatively cheap and are usually already present in the manufacturing facility. For manufacture at small to intermediate scale, dilution refolding can be done using disposable process equipment (tank liners, peristaltic pumps and disposable hoses).

The dilution refold method also has characteristics which facilitate process development experiments. For example, unless multiple dialysis systems are available, it is only possible to test one dialysis refolding condition at a time. It is possible, however, to test many dilution refolding conditions in parallel, as each condition can be tested in a simple vessel (see Chapter 2).

1.3.1 Solubilisation and denaturation

Urea is reported to be the preferred chaotrope for solubilisation by two companies with extensive experience of developing and operating refold steps in commercial manufacture. It is relatively cheap and is compatible
with most process equipment. It does not greatly increase conductivity, and so does not need to be removed or diluted from the refolded protein solution before downstream ion exchange chromatography (IEC is often used as the first purification step after refold. Avecia has used urea denaturation successfully in a large number of cGMP manufacturing campaigns.

If inclusion bodies will not solubilise and denature satisfactorily in urea, it may be necessary to use GdnHCl as a chaotrope. However, GdnHCl has been found to be incompatible with many items of large-scale process equipment, particularly steel items. Denaturing GdnHCl solutions have higher ionic strength than urea solutions, so solutions of refolded proteins containing GdnHCl are more likely to need dilution or diafiltration to reduce conductivity prior to loading onto ion-exchange chromatography media.

The practicalities of preparing the solubilised protein solution can be problematic. Urea can be dissolved up to a concentration of approximately 10M at room temperature. GdnHCl can be dissolved up to a concentration of approximately 8M at room temperature. Urea solutions above approximately 8M and GdnHCl solutions above approximately 6.5M are avoided because there is a danger of crystals forming and damaging processing equipment.

There is a large change in volume on the addition of the chaotrope – e.g. adding 8 moles of urea to 1 litre of water gives approximately 1.5 litres of approximately 5M urea solution. The dissolution of urea or GdnHCl into water is strongly endothermic so heating may be needed to dissolve the required amount of urea in a reasonable time. The higher the concentration of chaotrope required, the longer the time needed to get the chaotrope dissolved. The time required is even longer at large scale, due to the need to get heat into the larger volume, and may be up to 24 hours even in a heated vessel.
If manufacture is being conducted at small scale then inclusion body paste can be prepared with a batch centrifuge. This can then be dissolved into a strong chaotrope solution. For example, 1L of inclusion body paste can be dissolved into 3L of 8M urea to give 4L of solubilised protein solution in 6M urea. Note that in this case the volume of solubilised protein solution is four times the volume of the inclusion body paste. This large increase in volume is necessary because of the lower limit on the urea concentration in the refolded protein solution (necessary to cause solubilisation) and the upper limit on the urea concentration in the urea buffer (due to the solubility limit of urea). The IB paste will contain a very high concentration of the product protein, so the solubilised protein solution may be diluted with more chaotrope solution (e.g. 6M urea for the example above) to achieve the desired solubilised protein concentration.

If the manufacture is now being carried out at larger scale then batch centrifugation becomes impractical and a continuous (e.g. disc-stack) centrifuge must be used to prepare the inclusion bodies. This will produce a much larger volume of less concentrated IB slurry. Due to the large volume of slurry it would be impractical to dissolve the slurry using an even larger volume of concentrated chaotrope solution as with the IB paste example above, therefore the chaotrope must be added as a solid directly to the slurry. This causes further problems as the solubilising IB slurry/solution must be held while the chaotrope dissolves. The solubilisation must be done in a jacketed vessel to supply heat to dissolve the chaotrope and the presence of the product protein limits the jacket temperature (due to the need to avoid exposing protein to high temperatures), so the chaotrope dissolution takes even longer.

If a continuous centrifuge is used to prepare inclusion body slurry, then the concentration of protein in the solubilised protein solution will largely be determined by the centrifuge performance. The maximum solubilised protein concentration possible will be the concentration of protein in the IB slurry divided by the volume-change which is caused by the addition of the
chaotrope. This solubilised protein concentration will typically be low, and so no further dilution with chaotrope solution will be done.

The reducing agent dithiothreitol (DTT) is reported to have a very short shelf life in solution and so is often added to the solubilised protein solution after the inclusion bodies have been solubilised. Alternative reducing agents to DTT include dithioerythreitol (DTE), tris(2-carboxyethyl)phosphine (TCEP) and 2-mercaptoethanol. 2-mercaptoethanol is extremely hazardous and so is unsuitable for use in large scale manufacture. TCEP is reported to be much more stable than DTT in the presence of metal ions, and slightly less stable in the presence of EDTA. The large-scale refolding of proteins from inclusion bodies solubilised in buffers containing TCEP has not been reported. If the protein has no disulphide bonds then reducing agents in denaturation (and reducing/oxidising agents in refolding) may not be necessary.

1.3.2 Refolding

A number of methods exist for refolding protein (see section 1.2) but, overwhelmingly, commercial processes use the dilution refolding method.

GMP manufacturing facilities are very expensive to build and validate. A launch product (a marketed therapeutic protein) may have a bespoke manufacturing facility built for its process. However, a process must be developed to manufacture product for clinical trials. Only a small percentage of products entering clinical trials are ever launched (most products fail in clinical trials) so most processes which are developed are only operated to manufacture trial material. Consequently it would be prohibitively expensive, in both time and money, to build a bespoke manufacturing facility for every product entering clinical trials because only a small percentage would reach the market and achieve a return on the investment.
It therefore follows that most of the processes which are developed are only ever operated in multiproduct facilities into which the process must be made to fit. New process equipment may be purchased for processes which are operated for late stage clinical trials (or which have rich backers) but this equipment must still fit into the existing facility. The majority of process development work is carried out to achieve processes which will fit into existing multiproduct facilities (and preferably into existing process equipment).

Large process volumes in refold steps are expensive (in terms of time and money). Larger volumes of purified water, larger amounts of chemicals, and longer times are required for buffer preparation and cleaning. Pumping operations take longer and may require larger pumps. Larger tanks may be required, with more powerful mixers. More dilute (larger volume) refolded protein solutions will lead to greater problems with downstream purification (e.g. longer chromatography loading times). There is therefore a strong driver to develop a refold step with the smallest possible refold volume, hence the lowest possible refold dilution factor and the highest refolded protein concentration.

Physical constraints in a manufacturing facility (e.g. tank size) may put an upper limit on the final volume of refolded protein solution (the refold volume) which can be produced. The downstream process may also constrain the refold volume e.g. a chromatography step will have a maximum flow rate for loading and there will be a maximum time for which product can be held prior to loading.

A lower limit on the dilution factor may sometimes be imposed by the need to reduce chaotrope concentration. If inclusion bodies are prepared using a continuous centrifuge, giving a dilute slurry, then the protein concentration in the solubilised protein solution will be low but the chaotrope concentration will still be high.
For example, it may be found that a washed IB slurry can be prepared to a maximum protein concentration of 3mg/ml. The IBs require 6M urea to solubilise, so after addition of solid urea up to this concentration the protein concentration in the solubilised protein solution is approximately 2mg/ml (due to the volume increase upon the addition of urea). It may be determined in process development experiments that the protein can be refolded at concentrations up to 1mg/ml (without significant losses to aggregation – see section 1.2.3). This would suggest that the refold step should comprise a 1+1 dilution of solubilised protein solution with refold buffer in order to give the smallest process volume and highest protein concentration possible. However, this dilution would give a final concentration of 3M urea. It may be determined that the highest concentration of urea in which the protein can refold is 1M (or, possibly, that the highest concentration of urea in which the refolded protein can be loaded onto a downstream IEC column is 1M). In this case the lowest possible dilution factor is 1+5, giving a refolding protein solution of 0.33mg/ml in 1M urea. This situation is relatively rare, however, and it is the need to reduce aggregation (and so increase yield) which typically leads to high dilution factors and low refolding protein concentrations.

Dilution factors used in dilution refolding steps are typically between 1+5 and 1+19 in commercial manufacture. Much larger dilution factors may be used at laboratory scale where process volume is not a problem. The danger there is that inappropriate laboratory processes will be transferred to pilot-scale manufacture.

Several different strategies exist to effect the refold dilution and are described below.
Step Dilution  The denatured protein solution is added to a larger volume of a refolding buffer in one step addition, then mixed to homogeneity. Due to its simplicity, this is the most commonly used strategy at small scale.

Fed Batch Dilution  The denatured protein solution is added to the refolding buffer over a significant period of time, with constant mixing. The denatured protein solution may be added continuously, or in steps ("pulsed refolding"). The protein begins to refold as soon as it is mixed into the refolding buffer. The duration of the addition is longer than the time required for the protein to refold, so the first protein to be added becomes folded long before the last protein is added. This has the effect that, at any point in time, only a small portion of the total protein is present in the unfolded (or partially folded) form. It is the unfolded (or partially folded) form of protein which has the greatest tendency to aggregate, so a fed batch dilution can give less aggregation (and so higher refolding yield) than a step dilution.
Staged Dilution

The denatured protein solution is added to a stirred vessel of refolding buffer dilution e.g. diluting 8M urea concentration down to 4M urea. This allows the protein molecules to partially refold, while still being flexible. This makes it easier for disulphide bonds to form in the native configuration, though it may also lead to increased aggregation. This refolding solution is then added to a larger stirred vessel containing more refolding buffer e.g. diluting 4M urea concentration down to 1M urea, allowing the protein molecules fully to form their native structure.

Continuous Dilution

The refolding buffer is slowly pumped into the denatured protein solution, with continuous mixing. The volume of refolding buffer added is usually much larger than the volume of denatured protein. At large scale it is difficult to mix the required wide range of volumes in a single vessel, so this strategy is not used.

At small scale, a vessel full of denatured protein solution may be tipped by hand into a larger vessel full of refolding buffer. This is impossible at large scale, where the denatured protein solution must be pumped into the vessel containing refolding buffer. Limited pumping speed can result in this addition taking significant time. Consequently some processes, which use a step dilution strategy when operated at small scale, are forced to use fed batch dilution at large scale.
1.4 The need for parallel scaledown experiments for rapid development of refold steps

There are strong financial incentives to get a novel therapeutic protein product to market quickly. The earlier a product reaches the market, the longer the product can be sold before other competing products are launched (or before the patent expires and generic versions appear) which will reduce profits. There is also an ethical incentive – the earlier to market, the more patients can be treated.

Processes must be developed before products can be manufactured, either for trials or for market. There is therefore a strong incentive to do process development early and rapidly.

Refold step development is typically on the critical path. The development of downstream purification steps may have to wait until the composition of the refolded protein solution (the stream leaving the refold step) is defined. The refold step may have to be performed early, at a preparative scale, in order to prepare material for downstream purification development experiments. There is therefore a particularly strong incentive to do refold step development experiments rapidly.

The operation of a protein refold step may be relatively simple (i.e. the dilution of a solubilised protein solution into a refold buffer) but refolding is complex at a molecular scale. Techniques exist to determine the native structure of a protein, but the pathways through which a protein may refold to reach this state are much more complex. There is no way of predicting, from structural knowledge of the protein, optimal conditions for refolding that protein. Refolding conditions must be determined experimentally.

There are many factors which can affect refolding (see Chapter 2). Knowledge of the protein allows some predictions to be made about refold conditions e.g. if a protein has disulphide bonds then redox reagents will be
needed. Some proteins will refold very easily, and may even refold fully at the first attempt. However for the majority of proteins a large number of refold conditions will need to be tested (a large number of refold reactions must be performed) in order to develop a satisfactory refold step. In general, the more refold reactions which can be performed, the better the refold step which can be developed.

A refold reaction will typically take between 1 and 24 hours. In order to develop rapidly a refold step it is therefore desirable to perform reactions in parallel in order to perform a large number in limited time. That is, it is desirable to initiate a series of refolding reactions at an interval significantly shorter than the duration of the reactions so that a number of reactions are proceeding at once. In this way the rate at which data can be accumulated is not limited by the relatively long duration of the reactions, but is only limited by the relatively short time needed to set up or to analyse the reactions.

During process development work, there will be a limited quantity of protein available. It is likely that refold development will be carried out in parallel with development of the fermentation and DSP steps. The process will not have been operated at large scale, so only a small quantity of inclusion body material will be available for development experiments. In order to perform a large number of refold reactions in refold development experiments, it is therefore necessary to perform the reactions at very small scale. The smaller the scale at which refold reactions can be performed, the more reactions are possible, and so the better the refold step which can be developed (experiments at smaller scale also consume less reagents, pose lower accidental chemical exposure risk and generate smaller volumes of hazardous waste for disposal).

To conclude, it is desired to perform large numbers of refold reactions at very small scale, operated in parallel, in order to rapidly develop refold steps. This has similarities to the high-throughput screening campaigns
which are used to identify target and drug candidates for pharmaceutical discovery. In these campaigns, very large numbers of reactions are performed in very small vessels (microwells, see Chapter 4) using automation (see Chapter 6) in order to generate and then screen large libraries of compounds. Refold development experiments could be performed in a similar manner, testing refold conditions instead of candidate compounds. This concept is at the core of this thesis.

1.5 The aims of the project

As discussed in this chapter, therapeutic proteins represent a large and growing part of the pharmaceutical industry (Datamonitor, 2004). Recombinant expression in microbial hosts is the method of choice for producing many of these therapeutic proteins, which often leads to the protein being produced in an inactive, insoluble form in inclusion bodies (Rudolph and Lilie, 1996) from which they must be released and refolded to an active form. The development of processes involving the refolding of protein from inclusion bodies is therefore an important area.

As discussed in sections 1.2 and 1.3, a number of different approaches exist to produce active protein from inclusion bodies however the most widely used method is chaotrope-solubilisation followed by dilution-refolding (Clark, 2001). The development of the chaotrope-solubilisation step is relatively straightforward however the dilution-refolding step is more complex (this is discussed further in Chapter 2) with widely varying yields and its development typically involves extensive experimentation (Tsumoto et al, 2002). The focus of this project will therefore be upon the process development of dilution refold steps for the production of active protein from chaotrope solubilised inclusion bodies.

Time is precious during process development and will limit the number of refold steps which can be tested (Middelberg, 2002) and inclusion body material for experimentation may also be limited. It is therefore important, in
order to gain maximum data for process development, to be able to perform
dilution refolding experiments which are efficient in their use of time and
inclusion body protein. Communications with Avecia Biologics and
Dowpharma, two contract biopharmaceutical manufacturing organisations
with significant experience of developing and operating inclusion body
processes, support these conclusions.

The purpose of this project is therefore to develop techniques to improve
the way in which commercial protein refold steps are developed. In
particular, to develop techniques which enable dilution refolding
experiments to deliver data for process design in a more time-efficient and
protein-efficient way. These improvements will bring one or more of the
following benefits to the manufacturer: reductions in development time,
reductions in development costs, increases in process yield, reductions in
process costs.

A significant requirement in every refold experiment will always be the need
to assess the degree of success or failure of each refold reaction tested.
Methods for doing this will therefore be studied with particular regard to their
impact on experiment efficiency.
Microwell scale reactions and laboratory automation have been used for many years to greatly improve the efficiency of biotech screening experiments. A key aim of this project will be to explore the application of these approaches to commercial dilution-refold step development experiments. Strategies to maximise the efficiency of such refold development experiments will also be considered.

The techniques which are developed will be applied to the development of a refold step for a protein or proteins, to demonstrate their value. The coordination of such techniques will also be considered, and proposals for a coherent system for rapid protein refold development, applicable to a wide range of proteins, will be made.
2 Development of protein refolding steps

In Chapter 1 the refolding of recombinant proteins from inclusion bodies was described, and the operation of such refold steps in commercial processes to produce therapeutic proteins was discussed. In this chapter, the development of protein refold steps, in the context of commercial bioprocess development, will be discussed.

The aims of commercial protein refolding development will be described. Ways in which these aims could be incorporated into objective functions for optimisation experiments will then be presented. The parameters which can be varied to optimise refold performance will be considered and the scope of refold development work which is carried out at various stages in the course of bioprocess development will then be discussed. Conclusions will then be drawn about the types of refolding development for which rapid microscale experimentation is most appropriate. Finally, the use of Design-of-Experiment (DOE) tools to enhance protein refold optimisation will be discussed, and conclusions drawn about their use.

2.1 Aims of refolding development experiments

The development of a refold step has several aims, which broadly fall into two categories – improving product process performance and minimising manufacturing cost (see section 1.3).

With respect to process performance, the developer will aim to:

Maximise refolding yield  Maximise the amount of correctly folded protein, or
of activity, produced per unit of unfolded protein.
Maximise refolded protein purity
Minimise the formation of aggregates and misfolded protein in order to minimise the duty on downstream purification.

Maximise refold step robustness
Develop a refold step which gives a consistent yield and purity.

In order to minimise manufacturing costs, the developer will aim to:

Use existing equipment
Design a refolding step which can be operated on existing equipment, preventing the need for expensive new facilities and equipment (e.g. determine the size of the largest refold tank available).

Reduce DSP requirements
Avoid doing things in the refold step which would require extra processing downstream, e.g. producing high process volumes which would require concentrating, or introducing contaminants which would require removal steps.

Reduce RM costs
Minimise the use of expensive materials, e.g. using cystine (~£50 per kg) instead of oxidised glutathione (~£5,000 per kg) to oxidise disulphide bonds if this does not have a large impact on yield.

Reduce processing time
Avoid sub batching, allowing the minimum time necessary for solubilisation and refolding.

Reduce validation requirement
Every extra raw material, piece of process equipment and processing step adds to the time, effort and cost of validation and regulatory compliance.
Reduce process volume  Minimise the volumes used in the process, e.g. by minimising dilution if this does not have a large impact on yield. This will reduce processing time (pumping time, filtering time, column loading time), minimise sub-batching, minimise the need for concentration steps, reduce RM costs.

Reduce SHE problems  Minimise buffer (and hence effluent) volumes. Use less hazardous reagents where possible, e.g. using DTT instead of mercaptoethanol for disulphide cleavage.

Avoid chemical incompatibility  Avoid using buffers which are incompatible with existing equipment. For example, try to avoid anything containing chloride ions at low pH in contact with steel.

There will be some conflict between these aims. Increasing the dilution factor in a dilution refold step usually increases the yield but adds to the process volume and processing cost. Adding an additive to the refolding buffer may increase yield but would need to be removed downstream. A judgement must made as to what compromise would be best (see section 2.2).
2.2 The objective function for optimisation

The quantity which is to be maximised (or minimised) in a programme of optimisation is the objective function.

In laboratory refolding optimisation, the objective function is typically the refolding yield (see Chapter 3). When optimising the performance of a refold step for commercial manufacture, the refolding yield is not the only consideration (see section 2.1). In this section, the objective function for commercial refold optimisation will be discussed.

By definition, the purpose of any commercial organisation is to make a profit and its aim is to maximise this. A bioprocess which is being developed for a commercial organisation is being developed to make a profit for that organisation. The purpose of the bioprocess development is to maximise the organisation's profit. Therefore the objective function for any refolding optimisation is profit (which, for the avoidance of doubt, is to be maximised).

The profit of an organisation is a very complex function. For example, a contract biopharmaceutical manufacturing organisation (CMO) will run a number of projects at a time. Each of which will run under contracts with a number of clients. The income to the CMO will be governed by the number of customers which the CMO can attract, and the terms of the contracts which are negotiated with them. This income will, directly or indirectly, be determined by the extent to which the CMO satisfies its customers. The costs to the CMO will be, in greatest part, the costs of employing development and manufacturing staff and the costs of maintaining GLP compliant development laboratories and GMP compliant manufacturing facilities. There will also be costs relating to materials, consumables and external services used in development and manufacture. Some or all of these costs may be passed through to the client according to the contract, nonetheless the CMO will seek to minimise these costs in order to please its clients and so win future business.
The effects that a protein refold step in a bioprocess can have on an organisation's profits are also complex. There are costs associated with the step itself. There are also costs associated with the impact that the refold step has on the rest of the process, particularly downstream purification.

It may be possible to develop a detailed, predictive, financial model for an organisation. It may also be possible to link a computational model of process development and manufacture to that financial model. Such a whole-process model would provide an ideal method for calculating an objective function for refold optimisation, but is beyond the scope of this research project. Therefore some approximation is needed in order to achieve a workable objective function.

Typically, the cost of development is (approximately) fixed by factors outside the development programme. It is largely determined by the amount of time spent on development (which determines the cost of the development staff and facilities used). Additional costs, such as consumables, are small in comparison. The amount of time spent on development is usually determined by the need to prepare a process for a given time slot in the manufacturing facility. There are two reasons for this. First, financial pressures give a strong incentive to manufacture a product quickly so that it can get to market as soon as possible. Clinical trials and drug launches are arranged in advance and are not flexible in their timing. Second, time in manufacturing facility is so expensive that time slots must be booked long in advance. Moving a manufacturing time slot is difficult and leaving it vacant is very expensive. Consequently, the greatest effect of refold step development on an organisation's costs is through manufacturing cost.
Unless a process requires the purchase of new bioprocess equipment (see section 1.3.2) the dominant manufacturing cost is the cost of time in the manufacturing facility. The two outputs of the refolding step which have the greatest effect on manufacturing time are the refold volume (larger volumes take longer to process) and the refolding yield (lower yields lead to more batches being needed). The actual time for which the protein refolds is small in comparison.

A reasonable approximation to profit, as an objective function for optimisation, is a function of refold yield and refold volume. Higher yields increase profit, and smaller process volumes increase profit. However the nature of the relationship between refold yield, refold volume and profit is itself complex and will vary greatly between organisations and between projects. The refolding development experiments in this work are performed outside the context of a commercial process development project. The relationship between refold yield, refold volume and profit is therefore undefined and so further simplification is necessary. In this work, the objective function for refold optimisation will normally be yield, with refold volume kept constant.

2.3 Variables for refold optimisation (refold step parameters which can affect performance)

In this section, parameters which can significantly affect the performance of the refold step will be discussed. These parameters are those variables which can be varied to optimise refold performance.

2.3.1 pH

The optimum pH for refolding varies from protein to protein, and must be determined experimentally. Higher pH in the refolding buffer can increase disulphide shuffling by promoting thiolate anion formation, which may give a higher refolding yield (Middelberg 2002). Too high a pH may prevent
refolding, or irreversibly damage the protein. Most refolds are carried out at a pH value between pH8 and pH 9.

2.3.2 Physical factors

Temperature can affect both the rate and the yield of refolding (Guise, West, Chaudhuri 1996). Most refolding is done at room temperature. Refolding at temperatures other than ambient temperature requires that the refold buffer is heated/chilled, and that the refold is done in a temperature-controlled vessel and so incurs significant extra costs. Refold temperature is therefore not normally investigated during refold development.

Under some refolding conditions, it is possible for poor mixing to have a negative effect on dilution refolding yield (Buswell et al 2001). Mixing in manufacture will typically be limited by the available equipment. The strongest mixing possible will typically be used, which leaves little scope for optimisation. Extremely vigorous mixing may aerate the refolding protein solution, which may have some effect on oxidation/reduction, and will probably cause foaming of the refolding protein solution. Aeration is unlikely to happen at large scale, due to the difficulty of vigorously mixing large volumes. Mixing is therefore a factor to be aware of, but not generally a variable for optimisation.

2.3.3 Dilution factor and protein concentration

The greatest loss of protein during refolding is to aggregation. Aggregation, being a second order (or higher) reaction, can be greatly reduced by using a lower protein concentration in the refolding buffer (see section 1.2.3). One way to achieve this is to reduce the protein concentration in the denatured protein solution. Another way is to increase the refold dilution factor, though this will also decrease the chaotrope concentration in the refolding protein solution. Reducing the protein concentration in the refolding protein solution will increase processing volumes, and so increase processing time and manufacturing cost. The lowest dilution factor (highest protein
concentration) that does not have an unacceptable impact on refolding yield is typically used.

2.3.4 Refolding enhancing additives

The composition of the refolding buffer is usually the most extensively investigated parameter in the optimisation of dilution refold steps. The refold buffer composition has many degrees of freedom – the concentration of each chemical in the refold buffer is a variable and there are many which can be used (redox reagents which are added to the refold buffer in order to reform disulphide bonds in the refolding protein are discussed in section 2.3.5). Chemicals which are added to the refold buffer to enhance refolding yield are discussed in this section.

A large number of chemicals have been found to enhance refold yield when included in the refold buffer. Many of these chemicals are very expensive, which makes them inappropriate for large-scale commercial manufacture. Table 2-1 shows some of the refolding enhancing additives which may be more appropriate for large-scale commercial manufacture. It should be noted that this list is not exhaustive. Further information on refolding enhancing additives may be found in publications by Guise, West Chaudhuri (1996), Lilie et al (1998), Clark (2001), Middelberg (2002), Tsumoto et al (2003) and references contained therein.

Different additives will work with different proteins and it will be necessary to do screening experiments to determine which additives, if any, work with a protein. Some additives which may have little effect on their own may have a large effect when used together. Examples of such interaction effects include: ethanol and sodium chloride in the refolding of IGF-1 (see Chapter 5), arginine and sucrose in the refolding of lysozyme (see Section 7.1). Where time and material constraints permit, additives should be tried in combinations. With the very wide range of possible additives, such screening experiments can get very large. It is unlikely that time and
material constraints will allow every possible additive to be tested in process development. A number of considerations can help to reduce the number of additives to be tested to a manageable size.

Raw material concerns should be considered. Refolding yields can be very sensitive to subtle changes in raw materials. Raw materials used in development should therefore be the same as those which will be used in manufacture. Some raw materials which are commonly used in the lab are unsuitable for use at large scale, or are not available in large quantities, or are not available with the required validation. Bringing a new raw material into a GMP manufacturing facility is expensive. Significant time and cost savings can be made by using a raw material for which the organisation already has a validated supplier, so existing raw materials should be tested in preference to other chemicals. Safety, health and environmental (SHE) issues should also be considered. Only additives which will not pose unacceptable hazards when used at large scale should be tested in development work. During late phase process development, regulatory considerations may make it impossible to add new chemicals to the process.

Some additives may have an adverse effect on the protein e.g. although arginine may improve the refolding of trypsinogen inclusion bodies, it also blocks the active site of the trypsin enzyme (see section 7.2.3).

Off-the-shelf “refolding kits”, which are designed to help determine effective refolding conditions, are available. These typically contain a number of buffers, with instructions, to help the user to screen a number of denaturing and refolding conditions. These kits are of some interest, but it should be noted that they are primarily designed for researchers trying to produce very small quantities of protein, e.g. for x-ray crystallography. They tend to produced poorly-optimised, low concentration refold steps, and often use chemicals which would be unsuitable for manufacture.
Each additive which is included in the refold buffer of the developed refold step will represent one extra variable (i.e. that additive's concentration) which must be considered during process validation. The additive will have been used because it has a significant effect on refold performance, therefore the concentration of that additive will be one of the critical parameters identified for testing in laboratory process qualification (LPQ – see Section 2.4) experiments. Each additional additive which is included in the refold buffer will significantly increase the amount of LPQ work which must be done. This is another incentive to minimise the number of additives used in the refold buffer.

Table 2-1 Table of chemicals which have been reported to enhance the refolding of proteins.

<table>
<thead>
<tr>
<th>Additive</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urea</td>
<td>Stabilises intermediate folding states. Reduces aggregation. Very commonly used. In any dilution-refold from protein solubilised in urea, the refolding protein solution will contain urea which is added with the solubilised protein. Concentrations used in the refold buffer are typically 0.3M - 2M. Too high a concentration will prevent refolding. Examples include Orsini and Goldberg, 1978, in which chymotrypsinogen A, refolded (at concentrations in the range 0.015-0.080mg/mL) into solutions containing different concentrations of urea, was found to have a yield up to 40% in 2M urea.</td>
</tr>
<tr>
<td>Guanidine hydrochloride (GdnHCl)</td>
<td>Stabilises intermediate folding states. Reduces aggregation. Very commonly used. In any dilution-refold from protein solubilised in GdnHCl, the refolding protein solution will contain urea which is added with the solubilised protein. Concentrations used in the refold buffer are typically in the range 0.1M - 1.5M. Too high a concentration will prevent refolding. GdnHCl can give a significant increase in conductivity of the refolded protein solution which can affect downstream IEC steps. Incompatible with a lot of steel processing equipment. Examples include Orsini and Goldberg, 1977, in which the yield of chymotrypsinogen A, refolded (at concentrations in the range 0.015-0.080mg/mL) into solutions containing different concentrations of GdnHCl, was found to have a maximum between 65% - 70% at concentrations of 1M - 1.3M GdnHCl.</td>
</tr>
<tr>
<td>L-arginine</td>
<td>Reduces aggregation. Very commonly used. Effective on many proteins. Concentrations used in the refold buffer are typically up to 0.8M. May give a viscous refold buffer. If added as arginine HCl in high concentrations, it may be incompatible with steel process equipment and may be conductive enough to cause problems with IEC loading. Examples include Ho et al, 2003, refolding lysozyme by dilution at 3mg/ml. The inclusion of 0.5M arginine in the refold buffer was found to increase the refold yield from ≈45% to ≈70%.</td>
</tr>
<tr>
<td>Osmolytes (e.g. sucrose, glucose, glycerol)</td>
<td>Stabilises native protein structure. May drive protein into native structure, decreasing misfolding or inhibit aggregation. Higher concentrations may give a viscous refold buffer. Concentrations used in the refold buffer are typically up to 0.1M - 4M, depending on solubility limit of the osmolyte. Examples include Hart et al, 1994, where it was reported that the inclusion of 1M sucrose in the refold solution had the effect of raising the average yield of correctly folded IGF-1 (at 0.67mg/mL, across a range of different refold solution compositions) from 51% to 60%. The inclusion of 4M glycerol had the effect of raising the yield to 64%.</td>
</tr>
<tr>
<td><strong>Polyethylene glycol (PEG)</strong></td>
<td>Reduces aggregation. May cause problems with downstream HIC steps. Effective with many proteins. Cleland et al (1992) found it enhanced dilution refold yields for four recombinant human proteins when used in molar excess at ratios of 2:1 to 20:1. For example, the yield of active interferon-gamma (at 1mg/mL) increased from 5% to 30% with the inclusion of 0.38mg/mL PEG3350 in the refold solution.</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>Salts (e.g. sodium chloride, ammonium sulphate)</strong></td>
<td>Can stabilise native protein structure. May increase rate of refolding, thereby favouring folding over aggregation. High conductivity may cause problems if refolded protein is to be IEC purified. Examples include Hart et al, 1994, where it was reported that the inclusion of 1M NaCl in the refold solution had the effect of raising the average yield of correctly folded IGF-1 (at 0.67mg/mL, across a range of different refold solution compositions) from 40% to 52%.</td>
</tr>
<tr>
<td><strong>Alcohols</strong></td>
<td>May cause problems with downstream HIC. Volatility / flammability may rule out alcohols on COSHH grounds. A very wide range of concentrations have been used. Hart et al (1994), who found that the inclusion of 20% ethanol in the refold solution increased the refold yield of IGF-1 from 23% to 41%. In contrast, Wetlaufer and Xie (1995) found the refold yield of carbonic anhydrase II increased from ~60% to ~76% with the inclusion of just 1mM n-hexanol.</td>
</tr>
<tr>
<td><strong>Mild detergent (e.g polysorbate, CHAPS, Triton)</strong></td>
<td>Reduces aggregation. May cause problems downstream. Binding to or permeation through UF or dialysis membranes may be unpredictable, making it difficult to remove. Examples include Wetlaufer and Xie, 1995. The inclusion of 10mM Triton X-100 in the refold buffer was found to increase the dilution refold yield of carbonic anhydrase II (at 4mg/mL) from ~60% to ~75%. CHAPS was found to have a similar effect.</td>
</tr>
</tbody>
</table>
If a protein, in its native form, binds to a molecule (e.g. an enzyme and a substrate, or an antibody and an antigen) then adding that substrate molecule to the refold buffer (in solution or immobilised) can give a big increase in refolding yield. Examples include Garboczi et al (1992). Recombinant human leukocyte antigen HLA-A2 is a protein comprising two subunits which presents peptide antigens on the surface of cells to elicit an immune response from T lymphocytes. It was found that HLA-A2, comprising two subunits separately expressed in E.coli inclusion bodies and solubilised in 8M urea, would refold (by dilution or dialysis) into the active soluble form only when an appropriate peptide antigen was present in the refold buffer to form the HLA-A2-peptide complex. Concentrations of the two subunits and peptide were 1μM, 2μM and 10μM respectively. Yields of 10-15% were achieved.

<table>
<thead>
<tr>
<th>Substrate or ligand</th>
<th>Samuel et al (2000) found that high concentrations of proline inhibited aggregation during refolding and suggested that high-order aggregates of proline (at proline concentrations &gt;1.5M) may behave as a protein folding chaperone. The inclusion of 2-5M proline in the refold solution increased the activity yield of a lysozyme dilution-refold (1mg/mL lysozyme) reaction from ≈10% to ≈25%.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proline</td>
<td>Cyclodextrin reduces aggregation. Cyclodextrins are large toroidal molecules with hydrophilic outer surface and a hydrophobic inner surface and as such are believed to have a similar effect to folding chaperones in enhancing refold yield. For example, in their 1998 patent, Sharma and Karrupiah report that the yield of a dilution-refold of GdnHCl-denatured carbonic anhydrase (at 17μM) could be increased from ≈40% to ≈95% with the inclusion of up to 100mM α-cyclodextrin in the refold solution.</td>
</tr>
</tbody>
</table>

### 2.3.5 Redox agents for refolding.

If the protein to be refolded has disulphide bonds in its desired structure, then it will be necessary to include an oxidising agent in the refold buffer to
re-form them (see section 1.2.3). Disulphide bonds re-form during the refolding process, and different arrangements of disulphide bonds may form, break and reform while the protein is passing through intermediate states from unfolded to folded. To aid this forming and breaking of disulphide bonds ("disulphide shuffling"), a redox couple (a reducing agent and an oxidising agent) may be used in the refolding solution instead of a simple oxidising agent.

If the conditions in the refolding solution are not sufficiently oxidising, correct disulphide bonds may not form. If the conditions are too oxidising, then disulphide shuffling may be inhibited, and disulphide bonds may become fixed in incorrect arrangements. Both the selection and the concentrations of redox reagents in the refolding protein solution may affect refolding yield. Note that some of the reducing agent used in solubilisation may be carried with the solubilised protein into the refolding protein solution. A number of redox systems which have been successfully used for the refolding of denatured and reduced proteins and a selection are listed in Table 2-2.
<table>
<thead>
<tr>
<th>Redox system</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxidised and reduced glutathione (GSSG and GSH)</td>
<td>Widely used at bench scale. Reduced glutathione is expensive. Oxidised glutathione is very expensive. Raman et al (1996) used various concentrations of this and other redox couples in experiments on the refolding of lysozyme. They achieved ≈80% refold yields (at 0.125mg/mL lysozyme) with 10mM GSH and 1mM GSSG. Interestingly, they also achieved ≈80% refold yields using just 1mM GSSG (no reducing agent) in the refold and observed that the yield was insensitive to the GSH:GSSG ratio.</td>
</tr>
<tr>
<td>Oxidised glutathione and dithiothreitol or dithioerythreitol (GSSG and DTT or DTE)</td>
<td>Oxidised glutathione is very expensive. Maachupalli-Reddy et al (1997) tested a range of concentrations of GSSG and DTT in the dilution-refolding of lysozyme and achieved a yield of 84% using 5mM GSSG and 2mM DTT (at 0.1mg/mL lysozyme).</td>
</tr>
<tr>
<td>Cystine and cysteine</td>
<td>Relatively low cost system. May give lower yield than glutathione for some proteins, but the lower cost can justify the slightly reduced yield. Buswell et al (2002), in bench scale experiments on the pulsed-dilution refolding (see section 1.3.2) of trypsinogen, used a refold buffer containing 3mM cysteine and 1mM cystine. The cysteine concentration increased to 5mM over time as the solubilised protein solution, containing 100mM cysteine, was added. Yields of up to 21% were achieved with trypsinogen refolded at a total concentration of 52μg/mL.</td>
</tr>
<tr>
<td>Cystine and dithiothreitol (DTT)</td>
<td>Low cost system, making it commercially appealing. May give lower yield than glutathione for some proteins, but the lower cost can justify the slightly reduced yield. Raman et al (1997) report experiments on the refolding of lysozyme, denatured in 6.5M GdnHCl with 80mM DTT and refolded by 100-fold dilution using refold buffers containing 1mM cystine. They experimented with the chaperone-like protein alpha-crystallin but found that it inhibited the oxidative refolding</td>
</tr>
</tbody>
</table>
of lysozyme. The yield of lysozyme, refolded at 0.115mg/mL, in the absence of alpha-crystallin was ≈80%.

<table>
<thead>
<tr>
<th>Method</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>10mM cystamine and 1mM cysteamine</td>
<td>Relatively low cost system. May give lower yield than glutathione for some proteins, but the lower cost can justify the slightly reduced yield. Simmons et al (1997) refolded human low density lipoprotein receptor fragment by dilution at ≈60µg/mL in a solution containing 10mM cysteamine and 1mM cystamine. They estimated that a refold yield of 10% was achieved.</td>
</tr>
<tr>
<td>Metal catalysed air oxidation</td>
<td>Requires long refold times. Yield may be low. May require refold tank to be sparged, risking foaming problems. Hart et al (1994) used residual DTT from denaturation and 0.5µM copper chloride as an aerobic oxidation catalyst to refold IGF-1, refolding for 3-6hours. With IGF-1 refolding at a total concentration of 1.7mg/mL, the maximum yield obtained was 50%.</td>
</tr>
</tbody>
</table>

2.4 Regulatory compliance and cost driven changes in the scope of experimental work during a drug's development

The scope of bioprocess development projects will vary greatly from one product to another. The scope of development work will also vary over the product's development lifetime, according to regulatory and other requirements. The extent and focus of the development work done on a dilution refold step will vary in a similar way. These variations are described in this section.

As with process development generally, refolding development typically comprises a number of overlapping stages (see Table 2-3). The extent of work on each stage will be specific to the particular development project.
Table 2-3  Typical stages in the development of a dilution refold step.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Selecting a refold method</td>
<td>Selecting which type of refold step to use – finding a starting point for refolding development (a refolding reaction which works but may have low yield). In commercial manufacture, this is usually a dilution refolding method</td>
</tr>
<tr>
<td>2</td>
<td>Experimenting with protein concentration</td>
<td>Finding a reasonable trade-off between yield and process volume, from which further development can be done.</td>
</tr>
<tr>
<td>3</td>
<td>Screening for factors to improve yield</td>
<td>Determining what factors (i.e. additives, redox couple, pH etc.) can affect yield.</td>
</tr>
<tr>
<td>4</td>
<td>Optimisation</td>
<td>Identifying a combination of (protein concentration, pH, redox couple, additives concentration) which gives the best combination of yield and processing cost.</td>
</tr>
</tbody>
</table>

One factor affecting the scope of refold development work is the state of the existing refold step. If no refold step has yet been developed (e.g. if an entirely new process is being developed) or if the existing refold step is entirely impracticable then one must be developed de novo. Different proteins can behave very differently in refolding. Any knowledge of the protein will be applied to designing the refold step (e.g. if a protein is known to need a particular ion for its active form, then that will be included in the refold buffer), but developing a refold step from nothing can still require a great deal of experimental work. If the existing refold step is well developed, and is considered satisfactory for manufacture, then less development work will be done on the refold step and development resource will be concentrated elsewhere. The refold step will be considered to be optimised if the expected benefit of further optimisation work is less than the expected cost of that work (or if time or material limitations prevent further work).
The stage which the product has reached in its development life cycle will dictate the degree to which a process must be developed and so will strongly influence the scope of development work.

If a product is to be manufactured for preclinical trials then there will be very limited resources available for process development. The product will be manufactured in a non-GMP facility (e.g. a GLP laboratory). Only a relatively small amount of protein will be required. There will therefore be a strong emphasis on purity and less emphasis on yield. Development resources will be therefore concentrated on purification. The aims of refold development work will be to find a refold step that works for the protein and provides an acceptable yield.

If the product passes preclinical trials and is to be manufactured for Phase I clinical trials, then more resources will be available for development. The manufacture must be done in cGMP manufacturing facilities and so each batch will be much more expensive than in preclinical manufacture. There will therefore be a stronger incentive to improve batch yield. Refold development will focus on improving the performance of the refold step and optimisation experiments will be performed to achieve this. There may also be some characterisation experiments to evaluate the robustness and performance of the refold step.

If the product passes Phase I clinical trials and is to be manufactured for Phase II clinical trials, then yet more resource will be available for development. Phase II trials will require much more material than Phase I trials. The process may be operated at a larger scale and more batches may be required. Only minimal changes to the process may be made between Phase II and Phase III clinical trials. (If major changes were made, the product of the Phase III manufacture may not be considered equivalent to the product of the Phase II manufacture, and so Phase II trials may have to be repeated). The process which is operated for Phase II manufacture is
therefore essentially the process which will be operated for Phase III manufacture and then for launch product manufacture (if the product is successful). If the process performs poorly, then manufacturing costs will be high. The incentive to improve process performance before Phase II manufacture is therefore very strong. Refold development work will typically include substantial experiments to improve the performance of the refold step. Some laboratory process qualification (LPQ or characterisation) work will also be performed to characterise performance. For the refold step, operating windows – ranges for critical parameters within which the refold step performs satisfactorily – may be established.

If the product passes Phase II clinical trials and is to be manufactured for Phase III clinical trials, then process development work will focus on process characterisation. Only minimal changes to the process are desired. Any optimisation work will be very limited e.g. varying pH or buffer chemical concentrations within a small range.

Regulatory requirements mean that a large amount of LPQ (or process characterisation) work will be done prior to Phase III manufacture. All critical parameters will be identified by risk analysis and, for each critical parameter, an acceptable range will be determined within which the process is demonstrated to perform as expected. A narrower operating range will then be defined.

Critical parameters for a dilution refold step typically comprise IB slurry concentration, chaotrope concentration, dilution factor, refolding time, refold buffer pH and concentrations of refold buffer components. These parameters will be tested at a minimum of 3 levels (lower limit, target, higher limit) in combination in order to demonstrate that the refold step performs as expected within this range. The number of conditions for testing therefore increases rapidly as the number of critical parameters increases and statistical tools are often used to keep the experiments manageable (see Section 2.5).
After Phase III, little or no additional process development work should be
done. Some more LPQ work may be performed prior to a license
application (if Phase III trials are successful). If the license application is
approved and the product is successfully launched, then any future changes
to the process will be strictly controlled. If a change to the process
becomes necessary (e.g. a technological advance makes significant cost
reduction possible, a regulatory change requires alterations to the process,
or a specified raw material or consumable becomes unavailable) then the
implications of that change must be investigated (experimental work may be
necessary to support the investigation), regulators must be informed and
regulatory approval may be needed before future batches are released.

2.5 The application of Design of Experiments (DOE)
techniques to protein refolding experiments

In refold step development it is desired to optimise the performance of the
refold step. That is, the developer seeks to maximise the performance of
the refold step by changing that refold step.

The factors which may be changed to effect this optimisation (the
parameters of the refold step - see Section 2.3) are the independent
variables. The results of the refold step (e.g. yield, refold volume etc.) are
the dependant variables. The objective function is a function of the
dependant variables which represents an single combined measure of how
good the result of the refold step is (see Section 2.2). Thus, optimisation
attempts to set the independent variables to maximise the objective
function.

Historically, optimisation has been done using a one-variable-at-a-time
(OVAAT) approach. The objective function is maximised with respect to
one independent variable (i.e. that variable is varied to find a maximum), then that first variable is held constant and the objective function is maximised with respect to a second independent variable. This is then repeated for all independent variables. The advantage of this approach is that it is simple and easy to understand. However, there are usually interactions between the effects of these variables (particularly in protein refolding) and so this approach can be pseudo-convergent. The experimenter may perceive that an optimum has been found because varying one variable at a time does not lead to any increase in the objective function, when indeed this is some way from the optimum. The OVAAT approach is also inefficient – a large number of runs must be performed in each experiment.

Design of experiments (DOE) techniques were first developed by Fisher in the 1920s to improve the efficiency of experiments. DOE techniques were further developed more recently by Taguchi, Cox, Yule and others. There now exist a number of practical, relatively simple to apply, DOE techniques which are often described as a “tool box” for the efficient design of experiments. A large number of software packages are now readily available to make DOE techniques even easier to use.

DOE techniques facilitate the experimenter to design experiments in which all of the significant factors (variables) are varied systematically. In an experiment a number of conditions (e.g. a number of refolding conditions) are tested. One run (e.g. one refold reaction) may be performed at each condition, or a number of runs may be performed at each condition (e.g. triplicate experiments in which three replicate runs are performed at each condition) and an average of the results taken.

DOE techniques provide a mathematical framework for varying all factors simultaneously in an experiment. This helps to make the experiments efficient but can make them difficult for the experimenter to grasp intuitively. DOE techniques (and usually DOE software) are therefore used to plan
experiments and to analyse results. The analysis of the results can tell the experimenter which factors have significant effects, how these effects interact and can be used to identify optimal conditions.

There are a number of classes of experiment designs which DOE can produce. Different classes of design will be appropriate for different experiments. The decision of which class of design to use will be based on knowledge of the system being investigated and the type of information which is being sought. Two classes which are of particular interest for rapid refold step development are screening designs and response surface designs.

Screening designs are used where a large number of factors are to be tested in order to determine which factors have significant effects. This is particularly useful early in an investigation to reduce the number of factors which are then investigated in more detail in further experiments. Such screening experiments could be used early in refold step development, e.g. to determine which refold buffer additives enhance refolding or to determine which other factors have significant effects on yield. They could also be used in laboratory process characterisation work to help to determine which of the refold parameters are "critical parameters" (for which acceptable ranges must then be experimentally determined).

Response surface designs are used to investigate in more detail the dependence of dependent variables (results) on independent variables (factors). A mathematical model of this dependence can then be produced. Response surface designs are often used after screening experiments to find optimal conditions. This type of experiment design could be used in refold development to optimise refold conditions with respect to the significant factors determined in the screening experiment. Target operating conditions could then be set for the refold step to be operated at this optimum. The mathematical model could then be used to predict acceptable ranges for critical parameters, which would then be
experimentally tested. It is worth noting that despite the name, response
surface designs are not limited to two independent variables, although
producing a graph of the results is much simpler if only two are used.

Many of the experiments described in this thesis are of the simple single-
variable type. However, experimental design has been used in planning
some of the refold step development experiments, which were performed to
demonstrate the usefulness of microwell scale experimentation and
laboratory automation for refold step development.

Some simple two-factor matrix designs are used in the microwell-scale
refold development experiments described in Chapter 5. For example, one
of the ethanol substitute experiments (see Figure 5-4) in which one factor is
the alcohol in the refold buffer (20% ethanol or 20% propylene glycol or no
alcohol) and the second factor is the other additive included in the refold
buffer (no additive, arginine, PEG, GdnHCl, NaCl).

More complex experiment designs were used in the automated microwell-
scale refold development experiments described in Chapter 7. The use of
automation and microwell-scale experimentation allowed large numbers of
runs to be performed, making the use of statistical software tools particularly
valuable and the Minitab statistical software (Minitab Inc. of State College,
Pennsylvania, USA) was used in the planning and analysis of these
experiments. Screening experiments, seeking single and two-factor
interaction effects of possible refold enhancing additives, were performed
using the refolding of lysozyme and of trypsinogen as examples. A surface
response experiment was also performed, optimising the yield of lysozyme
refolding with respect to the concentrations of refold enhancing additives.
2.6 Conclusions

There are many aims of refold step development (see section 2.1) and the relative weight of each of these aims will be very particular to the individual development project – they will vary from organisation to organisation and from project to project. The composition of the notional objective function for a refold optimisation (see section 2.2) will therefore also vary from project to project and is complex. Many of the elements, such as process-related costs and time costs, will themselves be complex and often confidential. Whole process and financial modelling have the potential to provide an accurate objective function for refold optimisation, but are beyond the scope of this research project. Refolding yield will therefore be the primary aim of refolding optimisation experiments in this project. Refold volume (or refold concentration) will also be considered.

The requirements for refold development experiments are highly variable. Even for a single therapeutic protein product, developed by a single organisation, the scope of development experiments will change according to the stage of development (see section 2.4). It would be a mammoth task, and probably folly, to design a prescriptive system which could direct the development of protein refold steps in all circumstances. The exact requirements and aims of refold development experiments can be best judged by persons familiar with the particular development project and with the particular organisation doing it. This project will therefore focus on developing a system of flexible techniques to improve the efficiency and speed of refold development, which can then be applied by process development scientists.

Techniques for rapid, very small scale refold experiments (see section 1.5) will be most useful for early to mid stage process development (process development for preclinical, phase I and phase II manufacture). In these phases, a large number of refolding conditions must be tested in order to first create and then to optimise a refold step. To test these refolding conditions, a large number of refold reactions must be performed and in
early to mid phase manufacture there will be limited resources and material available for refold development.

A number of experiments must also be done in late-phase development during LPQ work for process validation but very small scale experimentation is less appropriate for these. There will be pressure to perform these experiments at larger scale in order to be confident of convincing regulators that the results are representative of manufacturing conditions. In late stage development there will also be more resource available (because the development will be more valuable) and there will be more inclusion body material available so there will be less need to perform experiments at very small scale.

Design of experiment (DOE) techniques are useful tools for further improving the efficiency of rapid, very small scale refold development experiments. In particular, screening designs are useful for early experiments to determine useful additives and significant factors and response surface designs are useful for optimising the refold step with respect to a number of factors. These DOE tools will therefore be applied to refold experiments in this research project. DOE tools may also be useful for refold LPQ experiments, even if they are carried out at larger scale.

In refold development experiments a large number of refold reactions must be performed and it will be necessary to make some measurement of the yield of each of these reactions. Measurement of the yield of refolding reactions will be explored in Chapter 3.
3 Measuring the yield of the refolding reaction in refold development experiments

In order to develop a protein refold step it is necessary to perform refold experiments. A number of different refolding reactions must be tested, and the relative success of those refold reactions assessed. The success (or failure) of a refold reaction is often assessed by calculating a value for the yield of that reaction. In order rapidly to develop a refold step it therefore is necessary rapidly to quantify the yield of refold reactions.

In this chapter, the particular difficulties which the refold step presents for yield calculation are discussed. Different approaches to estimating or calculating the yield of refold reactions are also discussed. The different analytical techniques which can be used in the determination of refold yield are then assessed, with particular reference to the suitability of these methods for use with rapid (high throughput) experimentation. Experiments are described to illustrate the usefulness of selected techniques. Conclusions are then drawn as to which analytical techniques are most useful for rapid refold development experiments.

3.1 Potential problems with refolding yield calculation

Refolding yield is commonly defined as the percentage (by mass) of protein entering the refold step which ends up as correctly folded soluble protein. This may be referred to as the “mass yield” or “yield by mass”.
In a typical refold reaction a slurry of inclusion bodies (IBs) is solubilised by the addition of a high concentration of chaotrope (e.g. urea or guanidine hydrochloride) to give a solution of unfolded protein which is then refolded by dilution into a refold buffer. To calculate refolding yield, two quantities must be measured. First, the concentration of product protein in the unfolded protein solution (or IB slurry) must be measured. Second, the concentration of correctly folded product protein in the refolded protein solution must be determined. Both of these concentration measurements are problematic. The relevant volumes or dilution factors must also be known, but this is relatively simple.

The concentration of protein in solution after the refold step may be measured, but there are often a number of isoforms of the protein in solution and each form may have different levels of activity. It has been found that protein standards themselves can contain different isoforms of protein (Clements, E., 2003, pers. comm.). One isoform may be identified as the native structure but that may not be the form with the highest specific activity and a decision must be made as to whether the native or the most active form is desired.

It is also difficult to calculate the mass of product protein entering the denaturing step. Assaying the protein content of an inclusion body slurry is difficult as the IB must first be dissolved and reducing SDS-PAGE can give unreliable results from inclusion bodies. Measuring dried IB weight is of limited use as there will be solids present other than the product protein.

The product protein in the IB slurry may be entirely dissolved after solubilisation, but it may still be difficult to measure protein concentration accurately. The denatured protein solution will be turbid, with many impurities, so A280 measurements are of limited use. The high denaturant concentration will prevent the use of most assays. Any kind of buffer change to prepare samples for assay is likely to cause some of the product protein to precipitate out, leading to erroneous readings.
Often, a number of analytical techniques are used to measure the mass yield of refold steps in refold development experiments. For example, in a process development project performed by Avecia Biologics, the protein concentrations before and after refolding were measured using SDS-PAGE, absorbance at 280nm and GPC-HPLC. Yields for a refold step, calculated from these measurements, were in the range of 80% to 110%, depending on the combination of measurements used. This illustrates the problematic nature of calculating refold yield.

Where protein activity is used as a measure of the amount of protein which has successfully refolded, an “activity yield” is often calculated instead of a “mass yield”. The activity yield is based upon the mass of solubilised protein entering the refold step and a standard activity of that protein (in activity units per unit mass). This raises the problem of how to determine the standard activity. The activity of a sample of protein standard may be used, but that protein standard may not be entirely active. In such cases the activity of refolded protein may be greater than the activity of the protein standard, giving an activity yield of greater than 100%.

3.2 Alternative approaches to refold yield calculation for development experiments

In experiments to develop a refold step, the aim is to identify the best refold step possible, rather than to determine the exact yield of that reaction. Due to the inherent difficulty in accurately calculating the yield-by-mass of a refold step (see section 3.1), alternative measures of the successfulness of a refolding reaction may be used during development.

In process development experiments, the aim is often to determine which of a set of refold conditions gives the highest yield. For example, an experiment may have the aim of determining the optimal refold buffer composition for a dilution refold reaction. Practical constraints may
determine the unfolded protein solution composition and dilution factor (see section 1.3). The experiment would therefore comprise a number of refold reactions, each of which had the same volume and into each of which had been added the same quantity of unfolded protein (aliquots of a fixed volume from a single solubilisation reaction). Communications with Avecia Biologics indicate that experiments of this form are very common in the development of commercial refold steps. In such an experiment it would not be necessary to calculate the yield of each refold reaction; it would only be necessary to determine the concentration of correctly folded protein from each reaction. The reaction conditions giving the highest concentration of correctly folded protein must be the reaction conditions giving the highest yield. The main benefit of this is to save the experimenter the time and resource necessary to determine accurately product protein concentration before refold.

Where development experiments examine both refold buffer composition and dilution factor (but, again, all refolds are done from the same solubilised protein solution), the yield may be calculated in terms of mass of correctly folded protein (or recovered protein activity) per unit volume of solubilised protein.

There may be experiments in which no attempt is made to calculate refold yields. Screening experiments, for example (e.g. to screen refold enhancing additives) may be performed very early in process development, and may be done in such large numbers or with such small volumes that full analysis is impractical. In such screening experiments, simple absorbance measurements may be used to identify conditions which cause significant levels of aggregation and precipitation and so which can be ruled out of further process development. The smaller number of conditions which do not cause great precipitation can then be investigated further using larger refolding reactions and deploying more analytical techniques.
3.3 Analytical methods for the analysis of refolded protein in refold development experiments

This project is particularly concerned with techniques for experiments to develop protein refold process steps. A large number of analytical methods have been applied to the analysis of protein refolding. In sections 3.3 to 3.9 the use of these analytical methods for refold development experiments (i.e. to determine the degree of success or failure of large numbers of refold reactions) will be considered.

In this section analytical methods which have been applied to protein refolding but which are of less interest for refold development experiments will be briefly discussed. In sections 3.4 to 3.9 those techniques which are more relevant to refold development experiments will be considered in more detail and their use in assessing the refolding of a number of proteins will be described.

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) is very commonly used in bioprocess development experiments to measure protein concentration and purity. In refolding experiments, it can be used to measure protein concentration and purity in the refolded protein solution. Sample preparation involves solubilising and denaturing proteins, so this method is of little use for studying the degree of folding of protein. The purity of refolded protein solutions is generally determined by the purity of the inclusion bodies from which protein is solubilised (see section 1.3), which is determined by the combination of fermentation, homogenisation and centrifugation steps and so is of limited interest for refold development experiments. Refolded protein solutions are typically very dilute (see section 1.3) and are often below the calibrated range (or standard curve) of SDS-PAGE assays which prevents accurate protein concentration measurement. SDS-PAGE assays are relatively slow, taking several hours to produce results, which is another disadvantage although a number of samples can be analysed in parallel.
Native PAGE gels have been used to determine the degree to which samples of refolded protein solution have aggregated. This method is even slower than SDS-PAGE and the low protein concentration of such solutions can prevent useful results from being obtained.

Particle sizing has been used to determine the degree of aggregation in refolded protein solutions (Watkinson A., 2003, Aveia pers.comm.). Refolded protein solutions are however, complex – that is, they can contain a broad population of particle sizes, which makes analysis difficult. Particle sizing methods are sensitive to the fluid properties of the protein solution. In development experiments, where a large number of solutions will be used, it would be very time-consuming to determine the fluid properties of each solution to allow accurate particle sizing. This method appears to offer no benefits over size exclusion chromatography.

Circular dichroism (CD) has been used to study the folding of proteins. CD machines are expensive and must be carefully maintained to retain sensitivity. CD typically requires that samples are very pure and relatively concentrated, which makes it difficult to apply to solutions of proteins refolded from inclusion bodies. CD is very slow, taking several hours to process one sample. This method may be useful for studying the nature of a protein but is of little use for refold development experiments where large numbers of samples will be generated.

3.4 Activity Assays

Refold steps in commercial manufacture are usually done to produce active soluble protein from inclusion bodies. The purpose of the refold step is to make the protein active and the purpose of refold step development is to optimise the amount of active protein produced. Therefore the definitive measure of the success of an experimental refold reaction is the amount of
active protein (or activity) which is produced per unit mass of inclusion body protein. It can, however, be difficult to measure that activity.

Some proteins have simple and quick activity assays, e.g. enzymes such as lysozyme (see section 3.4.1), alcohol dehydrogenase and alkaline phosphatase. These assays can be performed quickly and in large numbers, making them ideal for development experiments using large numbers of refold reactions. Where the product protein is an antigen, such as a vaccine, immunoassays such as ELISAs may be used as activity assay. These can also be performed quickly and in large numbers. Many activity assays can be performed in microwell plate format (see chapter 4) or automated to increase throughput.

Activity assays are not all quick and simple. Those involving living cells or organisms can be particular slow and difficult. Hsih et al (1997) describe experiments to develop the refolding conditions of a recombinant fish growth hormone, whose activity assay involved injecting young fish with the refolded protein and monitoring their growth over a number of weeks. The use of assays which involve animals also pose ethical problems.

If development experiments are being performed which involve refolding the denatured protein into different refold buffers, the components of the refold buffer may interfere with the activity assay (for example arginine in the refolding of trypsinogen – see section 7.2.3). If refold enhancing additives are being screened, it may be necessary also to screen the effects of the additives on the assay (i.e. to measure the activity of native protein in the presence of each of the additives). It may be possible to effect a buffer exchange into a standard buffer before performing the activity assay but this would complicate and slow the assaying and the buffer exchange itself may affect the refolded protein.
If the refold step of a protein is being developed and no activity assay exists for that protein or the activity assay is problematic then other indirect analytical methods must be used in development experiments.

### 3.4.1 Lysozyme activity assay

Many of the experiments described in this thesis involve the refolding of hen egg white lysozyme. One of the reasons for selecting lysozyme as the example protein was the existence of a simple activity assay for that protein. A calibration curve was generated for this assay.

The activity assay protocol used was that of Shugar (1952). The basis of the assay is the rate of lysis of the lysozyme sensitive bacteria *Micrococcus lysodeikticus*, as measured by the decrease in absorbance of a suspension of that bacteria. Shugar defined an activity unit as equal to a decrease in turbidity at 450nm of 0.001 per minute.

#### 3.4.1.1 Materials and methods

All chemicals, including protein and cells, were purchased from Sigma Aldrich of Gillingham, Dorset, UK. Absorbance readings were made using a Uvikon 922 spectrophotometer.

A solution of 1mg/ml lysozyme in 100mM Tris-HCl pH 8.2 was prepared. Dilutions of this lysozyme solution were performed using the same buffer to give solutions with lysozyme concentrations in the range 0.005mg/ml to 0.100mg/ml. An assay solution of 360 mg/L *Micrococcus lysodeikticus* in 0.1M potassium phosphate, pH 7.0 was prepared and kept at 25°C for a maximum of 1 hour.

The spectrophotometer was switched on and set to record absorbance at 450nm over time. The temperature control system was set to maintain cuvette and assay solution temperature at 25°C. The machine was left on for at least 30 minutes to warm up before use.
In practice, 100μl of diluted lysozyme solution was pipetting into a 3ml cuvette. 2900μl of assay solution was then added to the cuvette, and the contents were mixed by three aspirate-dispense cycles of 2900μl. The cuvette was then placed in the spectrophotometer and the absorbance at 450nm monitored. The activity of the lysozyme solution was taken to be the gradient of the linear part of the absorbance curve. Four replicate assays were performed for each lysozyme solution.

3.4.1.2 Results

![Calibration curve for lysozyme activity assay. Four replicates of each point were run. Mean activity values are plotted, error bars show ± 1 standard deviation. A linear line of best fit was fitted, the equation and R² value are shown.](image.png)
3.4.1.3 Discussion and Conclusions

The assay gives a linear response in the range 0.005mg/ml to 0.100mg/ml lysozyme with a small spread of activity readings. The spread of readings is approximately constant with respect to the magnitude of the readings, so the spread is more significant at lower readings. The lysozyme used had a specific activity of 20,000 activity units per mg, which was within the reported range of 8,000 to 50,000 for that product.

3.5 Absorbance methods

In this section the use of absorbance methods for determining the concentration of correctly folded protein in a refolded protein solution will be discussed.

3.5.1 Absorbance at 280nm (measuring total soluble protein)

Commonly, the greatest loss of product protein over a refold step is in the form of insoluble aggregates which precipitate out of the refolding protein solution while the correctly folded protein remains soluble. The concentration of correctly folded protein can therefore be estimated by measuring the absorbance at 280nm of the refolded protein solution (the extinction coefficient of the product protein having previously been determined). The advantages of this method are that absorbance measurements can be made quickly, and give a figure for the product protein concentration.

This method has a number of disadvantages. Unless the protein has been purified while in the solubilised form, the refolded protein solution will have similar impurities as were present in the inclusion bodies, which will also absorb at 280nm and so will be a source of error. Samples must be clarified before absorbance measurements are made, either by filtration or centrifugation. Absorbance readings at 280nm cannot be made using
normal polystyrene cuvettes or microwell plates – either quartz or specialised plastic must be used.

This method can only distinguish between insoluble aggregated protein and soluble protein. It only measures the total amount of soluble protein (correctly folded protein, misfolded protein and soluble aggregates) so the absorption may not reflect the concentration of correctly folded protein.

3.5.2 Turbidity (measuring precipitation)

The precipitated protein aggregates which may form upon refolding increase the turbidity of the refolding protein solution, which will often become noticeably cloudy. These aggregates commonly represent the greatest loss of product protein in the refold step, so measurements of the turbidity of the refolded protein solution can give an indication of how much protein is left in solution and so give an estimate of how much protein has correctly refolded. A higher turbidity indicates greater precipitation and hence less soluble protein which, in turn, indicates a lower concentration of correctly folded product protein.

This method has a number of disadvantages. Like the absorbance at 280nm method, the turbidity method only indicates the amount of soluble protein and does not distinguish between correctly folded protein, misfolded protein and soluble aggregates. The relationship of turbidity to soluble protein concentration is less direct than that of A280. It can be affected by factors such as the size of precipitate formed.

The advantages of the turbidity method are its speed and simplicity. No sample preparation may be required (unlike the A280 method, no clarification is done). No expensive quartz equipment is needed – measurements can be made through any transparent glass or plastic. Equipment to measure absorbance in the visible region of the spectrum is simpler and cheaper than equipment to measure absorbance at 280nm –
light sources can be ordinary incandescent lamps. Measurements can even be made without the use of filters or monochromators as solid precipitates block light over a great range of wavelengths.

### 3.6 Correlation between turbidity and recovered activity in refolded lysozyme

Turbidity measurements offer an appealingly simple method of comparing different refolding conditions (see section 3.5.1) in order to develop rapidly a dilution refolding step. An experiment was therefore performed to assess how useful are turbidity measurements as a predictor of recovered activities.

Two methods of measuring turbidity were available: measurement in a microwell plate using a plate reader and measurement in a cuvette using a conventional lab spectrophotometer. It was desired to measure the turbidity caused by all of the precipitate formed during refolding. To measure turbidity in a cuvette would require that refolded protein solution be pipetted from the reaction vessel into the cuvette. It would be difficult to ensure that a representative sample of solution and precipitate were transferred to the cuvette, as precipitate could settle out of suspension or adhere to surface of the reaction vessel. Also, a conventional spectrophotometer measures turbidity using a beam of light passing horizontally through a cuvette, so turbidity measurements could be affected by precipitate floating or sinking in the cuvette. A microwell plate reader measures turbidity using a beam of light passing vertically through a microwell, so sinking or floating would not move precipitate out of the path of the beam. For these reasons it was decided to make turbidity measurements in microwell plates, on solutions of protein which had been refolded in those plates. Refolds were also done at a larger scale (10ml) to measure recovered activity.

Hen egg white lysozyme (HEWL) was denatured and refolded by 1+19 dilution into nine different refold buffers. Similar refold reactions were
performed in microwells (which were then measured for turbidity in a plate reader) and in 10ml CSTRs (which were then measured for HEWL activity). The turbidity and recovered HEWL activity of refold reactions performed under similar conditions were then compared.

3.6.1 Method

Hen egg white lysozyme (HEWL) was denatured then refolded by 1+19 dilution refolding in microwells and 10ml stirred tanks (CSTRs). At both scales, nine different refold conditions were tested (using nine different refold buffers). All refolds were performed at room temperature (20±2°C), and refolding protein solutions were allowed to stand for 24 hours after dilution before assaying. All refolds were vigorously mixed upon dilution to minimise any effect of poor mixing upon refolding. All pipetting operations were done by hand using Gilson pipettes.

3.6.1.1 Materials

All chemicals, including protein, were purchased from Sigma-Aldrich Company Limited of Gillingham, Dorset, UK.

3.6.1.2 Preparation of unfolded lysozyme

Unfolded, reduced protein was prepared after the method of Goldberg et al (1991). Hen egg-white lysozyme (HEWL) was dissolved at 10 mg/ml in a denaturing solution containing 6M guanidine hydrochloride (GdnHCl), 25mM dithiothreitol (DTT). The solution was vortex mixed for 1 minute and left to stand for 2 hours at room temperature (20±2°C). It was then dialysed for 24 hours against an excess of 6M guanidine hydrochloride to remove dithiothreitol, and stored for up to 1 week in a refrigerator at 4°C. Unfolded protein solution was visually clear and colourless before refolding.

3.6.1.3 Microwell refolding

Microwell refolds were carried out in Corning 96-well polystyrene 350µl plates from Corning Inc. of Corning, New York, USA. These have a recommended working volume of 200µl.
Nine refolding conditions were tested at microwell scale, each condition being a dilution refold of unfolded HEWL solution into one of the nine refold buffers (see section 3.6.1.5). Each condition was tested in triplicate, making a total of 27 refolding reactions.

For each refold reaction, 10μl of unfolded HEWL solution was pipetted into one well of a 96-well plate, forming a droplet on the base of the well. 190μl of refold buffer was pipetted into the well, with the tip in contact with the droplet of unfolded HEWL solution. The contents of the well was immediately pipette-mixed by five cycles of aspirating and dispensing 190μl. Refolding protein solutions were allowed to stand in the plate for 24 hours before turbidity measurement.

3.6.1.4 10ml CSTR refolding

10ml CSTR refolds were carried out in 50ml polycarbonate centrifuge tubes (Beckman Coulter Ltd., California, USA), which were mixed using 7mm magnetic stirrer bars and a Variomag 9-position magnetic stirrer (H+P Labortechnik AG, Oberschleissheim, Germany).

Nine refolding conditions were tested at 10ml scale, each condition being a dilution refold of unfolded HEWL solution into one of the nine refold buffers (see section 3.6.1.5). Each condition was tested in triplicate, making a total of 27 refold reactions.

For each refold reaction, 9.5ml of refold buffer was stirred in a centrifuge tube at 200rpm. 0.5ml of unfolded HEWL solution was pipetted into the stirring refold buffer and allowed to mix for 10 minutes before being removed from the stirrer. Each refolding protein solution was left for 24 hours before assaying.
3.6.1.5 Refolding conditions

The nine refolding conditions tested were all 1+19 dilution into refold buffer at room temperature (20±3°C). Nine different refold buffers were used. Each refold buffer contained 0.12mM GSSG and 0.04mM DTT (as a redox couple to oxidise disulphide bonds), 1mM EDTA, 0.1M Tris pH8.2. The compositions of the nine refold buffers are given in the table below. All refold buffers were visually clear and colourless.

Table 3-1 Table of nine refold buffers used

<table>
<thead>
<tr>
<th>Condition number</th>
<th>Refold buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.12mM GSSG, 0.04mM DTT, 1 mM EDTA, 0.1M Tris, pH 8.2</td>
</tr>
<tr>
<td>2</td>
<td>7.5%v/v tert-butyl alcohol, 0.12mM GSSG, 0.04mM DTT, 1 mM EDTA, 0.1M Tris, pH 8.2</td>
</tr>
<tr>
<td>3</td>
<td>0.14M arginine, 1.5mM PEG3350, 0.12mM GSSG, 0.04mM DTT, 1 mM EDTA, 0.1M Tris, pH 8.2</td>
</tr>
<tr>
<td>4</td>
<td>0.2M sucrose, 0.12mM GSSG, 0.04mM DTT, 1 mM EDTA, 0.1M Tris, pH 8.2</td>
</tr>
<tr>
<td>5</td>
<td>0.14M arginine, 0.12mM GSSG, 0.04mM DTT, 1 mM EDTA, 0.1M Tris, pH 8.2</td>
</tr>
<tr>
<td>6</td>
<td>1.5mM PEG3350, 0.12mM GSSG, 0.04mM DTT, 1 mM EDTA, 0.1M Tris, pH 8.2</td>
</tr>
<tr>
<td>7</td>
<td>7.5%v/v tert-butyl alcohol, 1.5mM PEG3350, 0.12mM GSSG, 0.04mM DTT, 1 mM EDTA, 0.1M Tris, pH 8.2</td>
</tr>
<tr>
<td>8</td>
<td>0.14M arginine, 0.2M sucrose, 0.12mM GSSG, 0.04mM DTT, 1 mM EDTA, 0.1M Tris, pH 8.2</td>
</tr>
<tr>
<td>9</td>
<td>0.14M arginine, 1.5mM PEG3350, 0.2M sucrose, 0.12mM GSSG, 0.04mM DTT, 1 mM EDTA, 0.1M Tris, pH 8.2</td>
</tr>
</tbody>
</table>

3.6.1.6 Assays

Turbidity of the microwell refolds was measured at a wavelength of 405nm using a Spectracount absorbance plate reader from Perkin Elmer of Boston, Massachusetts, USA.
HEWL activity measurements were made using the technique described in section 3.4.

### 3.6.2 Results

Figure 3-2 shows the turbidity and activity readings of lysozyme samples refolded at each of the nine refolding conditions used.

![Graph showing the relationship between turbidity of lysozyme solutions refolded in microwells and the activity of lysozyme solutions refolded in 10ml CSTRs under similar conditions. Each point represents the mean activity of three 10mL scale refolds and the mean turbidity of three microwell scale refolds, with error bars showing ±1 standard deviation. The points are labelled with the number of the refold condition to which they refer (see Table 3-1). Linear line of best fit to mean values is shown, with equation and $R^2$ value.](image-url)

Figure 3-2 Graph showing the relationship between turbidity of lysozyme solutions refolded in microwells and the activity of lysozyme solutions refolded in 10ml CSTRs under similar conditions. Each point represents the mean activity of three 10mL scale refolds and the mean turbidity of three microwell scale refolds, with error bars showing ±1 standard deviation. The points are labelled with the number of the refold condition to which they refer (see Table 3-1). Linear line of best fit to mean values is shown, with equation and $R^2$ value.
3.6.3 Discussion and conclusions of the experiment

Precipitation of refolding protein causes turbidity in the refolded protein solution. In this experiment there was a fixed amount of protein in each refolding reaction (0.5mg HEWL per ml of refold) therefore it was expected that refolding reactions which had a higher turbidity would have a lower recovered activity. This relationship is indeed seen in the results.

Reactions using refold buffers which contained arginine gave enhanced refold yields and lower turbidity (conditions 3, 5, 8, 9) than the refold reaction using a refold buffer with no additives (condition 1). This result was expected as arginine is known to increase yield and decrease aggregation during lysozyme refolding. Ho et al (2003) found that the inclusion of arginine in the refold buffer both increased the yield of a lysozyme dilution-refold reaction and decreased the turbidity of the resulting refolded lysozyme solution. Reactions using refold buffers which contained sucrose also gave higher refold yields and lower turbidity (conditions 4, 8, 9) than the no-additive condition 1. Again, this result was expected. Sucrose enhances the refold yield for many proteins (see section 2.3.4) and has been reported to inhibit aggregation of lysozyme (Ueda et al, 2001) and stabilise native lysozyme (Lai et al, 2000). Low concentrations of PEG3350 in the refold buffer have been shown to enhance refold yield for a number of proteins (Cleland et al, 1992). In this experiment on the refolding of lysozyme, the presence of PEG3350 in the refold buffer appears to have different effects depending on which other additives are present, with very little effect when it is the only additive (condition 6, compared to condition 1). Reactions using refold buffers which contained tert-butyl alcohol gave lower refold yields and higher turbidity (conditions 2 and 7) than the refold reaction using a refold buffer with no additives (condition 1). Although the presence of alcohols enhances the refold yield of some proteins (see section 2.3.4), the presence of ethanol has been seen to decrease the stability of native lysozyme (Lai et al, 2000). If tert-butyl alcohol has a similar effect to ethanol then it could be expected to reduce the yield and increase aggregation in a refolding reaction.
The correlation between turbidity and recovered activity is not particularly high ($R^2=0.86$) for these refolding reactions – a value of $R^2=0.95$ would normally be required for the correlation to be considered good. It can therefore be concluded that turbidity would not be an accurate measure for predicting recovered activity. It could be useful for "screening out" poor refolding conditions i.e. if a large number of refolding conditions were tested (using a large number of refolding reactions) then turbidity measurements could be used to identify conditions which give low recovered activity (those with high turbidity). The conditions which give higher recovered activities could then be tested using activity assays.

3.6.4 Conclusions on absorbance methods

The speed and simplicity of absorption methods make them appealing methods for use in small scale, high throughput screening experiments. The disadvantages of absorption methods (see section 3.6) require that other techniques be used for more advanced experiments.

3.7 Chromatographic methods

In this section the use of chromatographic methods for determining the concentration of correctly folded protein in a refolded protein solution will be discussed.

3.7.1 Size-exclusion chromatography

Size exclusion chromatography (SEC), also known as gel filtration or gel permeation chromatography (GPC), is used to separate or distinguish between molecules of different sizes. It has been shown to be an interesting method of performing refolding (see section 1.3) but in this section is discussed only as an analytical method.
Upon refolding, many proteins can wholly or partially aggregate into an inactive form. These aggregates are often insoluble and so precipitate out of the refolding protein solution, giving a cloudy appearance. Aggregates may, however, be soluble. Absorbance methods can distinguish between soluble and insoluble protein (see section 3.5) but cannot distinguish between a soluble protein and soluble aggregates. Analytical size exclusion chromatography can distinguish between soluble protein and aggregates and so may be a better measure of the concentration of correctly folded protein.

SEC has a number of disadvantages. It is relatively slow – one chromatography system, with a single column, can analyse approximately one sample per hour (this may be faster if the assay is done in high performance liquid chromatography (HPLC) format). In order to distinguish between a protein and its dimer (the smallest aggregate), high resolution is required. This in turn requires a tall, well packed column and low flow rates. Although SEC may be able to distinguish between unaggregated protein and aggregated protein, it cannot distinguish between correctly folded unaggregated protein and incorrectly folded unaggregated protein. A protein may be entirely in the unaggregated state, but may be incorrectly folded and so have no activity. Conversely, it is possible for a protein to retain activity even when it has aggregated.

3.8 Relationship between size exclusion chromatography and recovered activity in refolded lysozyme

Samples from the CSTR refold reactions described in section 3.6 were analysed by size-exclusion chromatography in order to investigate the usefulness of this method as a means of determining the successfulness of refold reactions explored in the refold step development experiments.
3.8.1 Materials and methods

Materials and methods for the refold reactions are described in section 3.6. All chemicals used in chromatography were purchased from Sigma-Aldrich Company Limited of Gillingham, Dorset, UK.

Size exclusion chromatography was performed using a Biocad system (Applied Biosystems, California, USA). A 10mm diameter Pharmacia HR column was packed to a bed height of 285mm with Pharmacia Superdex 75 Prep Grade (Amersham Biosciences, Freiburg, Germany). HETP measurements were made using the manufacturer’s protocol. The column was then repeatedly unpacked and then re-packed until the HETP value met the manufacturer’s specifications. A value of N=9500 plates/m was achieved which was considered acceptable.

The column was then equilibrated at 1ml/min with 3 column volumes of running buffer (100mM Tris-HCl, pH8.2). 100µl of native lysozyme solution in running buffer was loaded onto the column and run at 0.5ml/min, with the eluate monitored at 280nm, to determine the point at which native lysozyme eluted. The column was then kept refrigerated and brought back to room temperature (20±3°C) immediately prior to use.

Samples were taken from the refolded protein solutions at 24hours ± 5hours after the initiation of refolding. 1ml of refolded protein solution was diluted with 4ml of running buffer (to minimise the formation of precipitates on the column) and filtered through 0.2µm syringe filters before being loaded. 100µl of each sample was loaded and run at 0.5ml/min, with the eluate monitored at 280nm. An example chromatogram is shown in figure Figure 3-3.
Figure 3-3 Example of a SEC chromatogram of a sample of refolded lysozyme.

For each refolded sample, the area of the peak occurring at the same position as that for the native lysozyme was taken to indicate the concentration of unaggregated lysozyme.

In order to allow nine samples to be analysed in one day, the column was not cleaned in between samples. Samples had been prepared from purified lysozyme and filtered so inter-sample cleaning was considered unnecessary. After use, the column was cleaned with 1M NaOH and stored in 100mM NaOH.
3.8.2 Results and discussion

Figure 3-4 is a graph of the SEC peak area and activity values measured for each of the nine refold conditions used.

![Graph showing the relationship between SEC peak area and recovered activity of lysozyme solutions. Each point represents one SEC assay and the mean of three activity assays, with error bars showing ± one standard deviation. Linear line of best fit to mean values is shown, with equation and $R^2$ value.](image)

Figure 3-4
Lysozyme is a monomeric protein in its native state. It was therefore expected that refolded protein solutions with the greatest concentration of unaggregated protein (i.e. the greatest SEC peak eluting at the same position as native lysozyme) would have the greatest activity. This relationship is seen in the results, however the correlation between SEC peak area and recovered activity is not good ($R^2=0.84$)

Size exclusion chromatography is a relatively slow analytical technique. In this experiment it took approximately 50 minutes to analyse each sample. This would cause great problems for any experiment which generated a large number of samples, such as a refold screening or optimisation experiment.

Refolded protein in the samples experienced two buffer changes (first on dilution with running buffer, second on application to the column) which may have altered the degree of aggregation of the protein. It may have been possible to tailor the running buffer for each sample, so that the refolded protein did not experience any buffer change. This would require buffer to be prepared and the column re-equilibrated for each sample, which would be prohibitively time-consuming in any experiment generating a large number of samples. Such an analysis was beyond the scope of this study.

3.8.3 Adsorptive chromatographic methods.

Ion exchange chromatography and reverse phase chromatography (RPC) have both been used as analytical methods in protein refold experiments. The conformation of a protein can influence its binding characteristics, so these methods can give information about the concentration of different isoforms of a protein. They have been successfully used to distinguish between and quantify the concentration of different isoforms of refolded proteins in the development of commercial inclusion body protein refold steps (Liddell J., Clements E., 2003 Avécia pers. comms.).
IEC and RPC methods have the advantage of being able to quantify the concentration of the desired isoform of a refolded protein, as compared to size exclusion chromatography which can only quantify the concentration of unaggregated protein.

IEC and RPC have a number of disadvantages as analytical methods for refold development experiments. They are slow compared to absorbance and some activity assays – rates of analysis are typically between 1 sample per hour and 4 samples per hour if done in HPLC format. Both methods require sample preparation which can change the state of the refolded protein. Refold buffers often contain significant concentrations of salts, so samples for IEC may need to be heavily diluted in order to bind to the column. Refolded protein solutions typically have low concentrations of protein, so this dilution can reduce the concentration to levels where it is almost undetectable.

3.8.4 Improvements in the throughput of an RP-HPLC assay for IGF

In commercial bioprocess development businesses, assay protocols are often developed by QC departments. Assays are therefore typically developed to give maximum accuracy and robustness, with little or no consideration given to throughput of samples (speed of assaying). For process development experiments, where a large number of process steps may be performed and analysed, speed of assaying is a major issue. The experiment described in this section illustrates the increase in throughput of an HPLC assay (reduction in cycle time) which can be achieved when that is considered a priority.

An assay protocol to determine the concentration of correctly folded IGF-1 was supplied by Avecia Biologics of Billingham, Cleveland, UK, for use in IGF-1 refold development experiments (see Chapter 5). (Specific details of the assay are confidential.)
The protocol followed a standard pattern. The RP-HPLC column was first equilibrated with aqueous buffer. The sample was then loaded onto the column and unbound material washed off. A linear gradient between an aqueous buffer and an acetonitrile buffer was then run, during which the IGF-1 eluted. The column was then cleaned using a high concentration of acetonitrile.

The assay had been optimised to give the best separation between isoforms of IGF-1 (the native isoform and a second, inactive isoform) which eluted close together. The assay had a cycle time of 23 minutes per sample. It was anticipated that the assay would be the rate-limiting-step in a set of refold experiments to be carried out. It was therefore desired to reduce this cycle time if it were possible to do so without compromising the separation of isoforms. A study was performed in order to achieve this.

A number of RP-HPLC runs were performed. The parameters for these runs were variations on the parameters given in the Avecia assay protocol. The samples run were all identical, a standard protein solution supplied by Avecia which contained both isoforms of IGF-1. Blank runs, which followed the Avecia assay protocol but with no sample loading, were also performed to determine whether the column cleaning was effective.

The following parameters were investigated:
- Equilibration time
- Elution gradient
- Elution start % and end %
- Cleaning time

Flowrate was not investigated. Only one column was available and increasing the flowrate was considered to present too high a risk of irreversible damage.
As expected, it was found that the elution gradient could not be increased without impairing separation. The elution start % was raised and the end % was lowered without impairing separation, giving a shorter elution gradient. It was found that significant reductions in the equilibration time impaired the binding of the sample protein. It was found that the cleaning time could be significantly reduced while still giving effective column cleaning – no carry over of protein and no degradation in column performance were observed.

It was decided to use a revised assay protocol for the IGF-1 refolding experiments described in Chapter 5, which used the shorter elution gradient and cleaning time. These improved parameters reduced the cycle time from 23 minutes per assay to 12 minutes per assay, which allowed higher throughput in the refolding experiments. This illustrates the importance, when developing an assay during bioprocess development, of considering not only the assay’s performance but also the time and materials required to use the assay.

3.9 Fluorescence spectrometry methods

The use of two classes of fluorescence spectrometry; using a protein’s intrinsic fluorescence and using fluorescent dyes, to monitor the refolding of proteins in refold development experiments is discussed in this section with experimental results to investigate the utility of these methods in the context of this thesis.
3.9.1 Fluorescent dye binding

Fluorescent dyes, particularly those which bind to protein's hydrophobic patches, have been used to study the refolding of protein (Matulis et al, 1999). John et al (2001) describe a method for quantitative analysis of protein folding using ANS and bis-ANS, the most commonly used dyes. The fluorescent dye is introduced to the protein solution and binds to available hydrophobic patches on the protein. As the protein folds or unfolds, the environment experienced by the fluorescent dye changes and so the fluorescence of the dye changes. This can then be monitored by measuring fluorescence spectra of the protein solution.

This method has some disadvantages for refold development experiments. Before the refold development experiments for a protein can be done, experiments must be performed to determine the effects that the degree of folding of that protein has on the dye's fluorescence. It is likely that a large number of different buffer compositions will be tested in development experiments. These buffers may have effects on the dye's fluorescence (independent of protein folding effects) which must be investigated, so that the protein folding effects can be isolated and determined.

The fluorescent dyes bind to the protein. This binding could affect the conformation of the protein, or affect the refolding characteristics of the protein. For example, Kundu and Guptasarma (1999) describe ANS inhibiting aggregation during the unfolding and refolding of carbonic anhydrase. These effects could confuse the results of the refold development experiments by the formation of artefacts.
3.9.2 Intrinsic fluorescence

If a protein contains suitable fluorophores (e.g. tryptophan) then changes in its structure may result in changes in its intrinsic fluorescence. For some proteins, these changes may be sufficiently strong that unfolding and refolding can be monitored by observing changes in the fluorescence spectrum of the protein solution.

In such cases fluorimetry may not indicate the concentration of correctly folded protein in solution, but it can give a measure of the degree of folding of the population of protein molecules in the solution. Fluorescence measurements are fast compared to other analytical methods. A fluorescence measurement at a fixed wavelength can be made in less than one second. A fluorescence spectrum can be measured in less than one minute. Fluorescence readings can be taken from samples in a microwell plate, which gives a convenient way of handling large numbers of samples.

Figure 3-5, which was obtained from Avecia Biologics, shows the wavelength of the emission peak ($\lambda_{\text{max}}$) of samples of a protein in buffers containing different concentrations of a chaotrope (urea), when excited by light at 280nm. The protein contained a significant number of tryptophan residues and so it was expected that the protein would fluoresce strongly and that this fluorescence would change with the conformation of the protein. Higher chaotrope concentrations increasingly disrupt the structure of the protein, changing the environments of its tryptophan fluorophores and so changing its emission spectrum. This property of the protein allows a measure of the degree of folding of the protein to be made using fluorimetry.
In this particular case the protein is a recombinant antigen and is the active ingredient in a vaccine being manufactured for clinical trials. The protein was expressed in inclusion bodies, solubilised in a strong urea solution then refolded by dilution to produce correctly folded protein. Fluorimetric measurements formed part of the acceptability criteria for the process. Under these tests the solubilised inclusion body solution must have a $\lambda_{\text{max}}$ value above 350nm (indicating that the protein had substantially unfolded) and the refolded protein solution had to achieve a $\lambda_{\text{max}}$ value below 338nm (close to value for the protein standard, indicating that the protein had substantially refolded).

Figure 3-5 Graph showing the maximum emission wavelength of a recombinant antigen as a function of urea concentration.
3.9.3 Experiment to determine the variation in intrinsic fluorescence of BP3 in different concentrations of chaotrope

Recombinant human insulin-like growth factor 1 binding protein 3 (rhlGF1BP3 or BP3) is a binding protein for IGF1, a hormone (see Chapter 5). A complex of IGF1 and BP3 was being manufactured by Avecia Biologics for clinical trials. In the manufacturing process, BP3 was expressed in inclusion bodies, solubilised using a strong chaotrope and refolded to its native form.

An experiment was performed to measure differences intrinsic fluorescence (see section 3.9.2) of BP3 in various degrees of folding. This experiment had two purposes. First, to further investigate the usefulness of intrinsic fluorescence as an analytical method for protein refolding. Second, to determine whether it is a useful method for measuring the degree of folding of BP3.

3.9.3.1 Materials and methods

BP3 was produced by Aveceia Biologics of Billingham, Cleveland, UK. Other chemicals were purchased from Sigma Aldrich of Gillingham, Dorset, UK.

Solutions were prepared containing 11.8μg/ml BP3, 1mM EDTA, 25mM DTT, 50mM sodium acetate pH8.0 and various concentrations of either urea or guanidine hydrochloride (GdnHCl). Fluorescence measurements were made using a Perkin Elmer LS45 spectrophotometer with excitation wavelength set to 280nm and emission wavelengths scanned from 200nm to 425nm. Excitation and emission slit widths were set to 10nm. Scan speed was set 25nm/minute. 10 scans of each solution were generated, and the average was calculated. Scans of buffer blanks (containing all components but the protein) were also taken and subtracted from the scans of protein solutions. The magnitude and wavelength of the maximum of the resulting blanked fluorescence spectrograms were then calculated.
3.9.3.2 Results and discussion

Figure 3-6 Graphs showing the variation in magnitude and wavelength of the emission maximum of solutions of rIGF1BP3 in solutions containing different concentrations of urea and guanidine hydrochloride.

The wavelength of the fluorescence emission maximum ($\lambda_{\text{max}}$) does not appear to change as BP3 unfolds. Any changes in wavelength are small compared to experimental inaccuracies. It was concluded therefore that monitoring $\lambda_{\text{max}}$ was not a useful method for monitoring the degree of folding of BP3. The sequence of BP3 contains many fewer fluorophores than the antigen described in section 3.9.2, which may explain the lack of a strong
signal. It appears that intrinsic fluorescence may be a useful technique where a protein has a suitable large number of fluorophores, but will not be useful for every protein.

There appears to be a significant change in the magnitude of the fluorescence emission maximum with increasing concentrations of GdnHCl, but not with increasing concentrations of urea. This may be due to the much higher polarity of GdnHCl compared with urea. Both chemicals are strong denaturants, which have been used to denature BP3 for refolding. The fact that the decrease in emission magnitude is observed with GdnHCl but not with urea suggests that this decrease is not due to unfolding, but to other effects.

If resources had been available to perform further experiments then similar studies would have been performed examining the effects of chaotrope concentration on the wavelength of the excitation wavelength giving maximum emission magnitude. Fluorescence emissions reflect the state of the molecule in its excited state. Fluorescence excitation spectra may give more information about the state of the molecule in its unexcited state (i.e. the state from which it becomes excited), which is its normal state in solution and so may give more information about the foldedness of the protein.

3.9.3.3 Conclusions of the experiment

It was concluded that intrinsic fluorescence was not a useful method for monitoring the degree of folding of BP3. It may be a useful method for measuring the degree of folding of some proteins (e.g. the antigen described in section 3.9.2), but it is not useful for the development of refold steps for all proteins. Buffer components can interfere with the fluorescence (as seen in the differences between BP3 in urea and GdnHCl) and this interference could cause problems in refold development experiments where a large number of different buffers would be used.
3.9.4 Change in the intrinsic fluorescence of lysozyme upon refolding

An experiment was performed to determine the change in the intrinsic fluorescence emission spectrum of hen egg white lysozyme during refolding. This experiment had two purposes. First, to investigate the use of fluorescence as an analytical method for future refold development experiments with lysozyme. Second, to demonstrate that legacy laboratory equipment, for which no computer-control software was available, could be controlled and monitored using software written by the user.

3.9.4.1 Materials and methods

3.9.4.1.1 Materials

Chemicals, including protein, were purchased from Sigma Aldrich of Gillingham, Dorset, UK. Fluorescence measurements were made using a Perkin Elmer LS-40 fluorimeter (Perkin Elmer, Massachusetts, USA) with a quartz flow-through cell and integral sample pump. Control software was developed and operated using LabView G programming language (National Instruments, Texas, USA) on a Compaq Armada V300 laptop computer (Compaq, California, USA).

3.9.4.1.2 Fluorimeter control

The Perkin Elmer LS-40 is a microprocessor controlled fluorimeter. Parameters are normally set using a set of keys on the front of the machine, and results are produced via a small thermal printer built into the top of the machine. In this way, single frequency fluorescence measurements can be made, and emission or excitation scans can be produced. The instrument also has an RS232 serial port through which commands and information can be passed to and from it, though no software for external computer control was available.
It was desired to monitor the change in fluorescence spectrum with time of refolding lysozyme solution. To do this using the built-in key/printer control system would require that a large number of (paper) fluorescence spectrograms be generated over the refolding time (6 hours), and this set of paper spectrograms then be analysed. It was considered to be more practicable to control the instrument using an external computer and to gather the fluorescence data electronically via the RS232 serial interface.

Software was developed using LabView visual basic to control the fluorimeter, and to collect fluorescence data from the fluorimeter. The software sent serial command codes to the instrument and collected data from the fluorimeter via the serial interface (serial codes for the instrument were provided by the manufacturer). The software was controlled by the user using a graphical interface through which the following control parameters could be set: type of scan (emission or excitation), excitation or emission fixed wavelength (nm), scan start wavelength (nm), scan end wavelength (nm), interval between scans (minutes) and duration of data collection (the number of minutes over which fluorescence data would be gathered). At the start of measurement, the fluorimeter drew refolding protein solution into the cell and measured the fluorescence spectrum. After the set interval, the cell was flushed through with more refolding protein solution and the fluorescence spectrum measured again. This was repeated for the specified duration. The generated fluorescence data were compiled by the software and displayed in graphical form or exported to a spreadsheet file.

For this experiment, the machine was set to make an emission fluorescence scan every 2 minutes over the range 310nm to 470nm, with an excitation wavelength of 280nm, for 6 hours.
3.9.4.1.3 Lysozyme refolding

Unfolded lysozyme solution was prepared using the method described in section 3.6.1.2. 1ml of this solution was diluted into 99ml of refold buffer (0.3M GdnHCl, 0.12mM GSSG, 0.04mM DTT, 1mM EDTA, 100mM Tris pH8.2) with vigorous mixing. This solution was immediately presented to the fluorimeter and measurements initiated.

3.9.4.2 Results

Figure 3-7 show the results of the experiment. The fluorescence emission spectra recorded during the experiment are plotted against time, with t=0 being the initiation of refolding (the dilution of the denatured protein solution into the refolding buffer). Results are plotted using the Labview programme as a pseudo-3D graph with contours of constant emission magnitude at a spacing of 400 fluorescence units.

Figure 3-7 Graph showing the change in the fluorescence emission spectrum of a refolding lysozyme solution with time after the initiation of refolding, with excitation at 280nm.
3.9.4.3 Discussion and Conclusions

The high fluorescence values seen at wavelength below 320nm (seen as a ridge at the “back” of the graph) is likely to be due to scattering of the excitation beam.

There is an intense fluorescence emission at around 350nm, which decreases rapidly after the initiation of refolding to less than 10% of its initial value within 120 minutes. This is consistent with the (hydrophobic) tryptophan fluorophores in the lysozyme protein becoming buried as the protein refolds.

This intense fluorescence signal would be useful for identifying unfolded protein. It could be used in experiments screening large numbers of refold conditions to identify those conditions in which the protein remained unfolded. For example, it would be particularly useful in experiments testing refold buffers which contained various concentrations of chaotropes, or in testing different refold dilution factors to screen out those conditions in which chaotrope concentrations were high enough to prevent refolding.

Developing the computer programme to control the fluorimeter took approximately one week. This included “debugging” and overcoming problems caused by inaccuracies in the fluorimeter manual. (Approximately one week of study using LabView’s tutorial programme was also required to become proficient with the package, but LabView has many applications beyond the control of this one instrument.) The programme which was developed could collect fluorescence data with respect to time, and allowed data to be collected over long periods of time (in this case 6 hours) while the fluorimeter ran unattended, which enhanced the functionality of the fluorimeter. It also made it easy to link the control of and data from the fluorimeter to other programmes, such as control systems for chromatography systems and liquid handling robots (see Chapters 5 and 8) in order to construct an integrated system for automated experiments. (This programme was also used by other researchers for other research.
projects). This illustrates that legacy laboratory equipment can be computer
to enhance its usefulness and could then be integrated with other
laboratory systems, which would reduce the cost of implementing laboratory
automation systems (see Chapter 8).

3.10 Conclusions of the chapter

To enable effective refold development experiments to be performed for a
wide range of proteins (see section 1.5) a number of analytical techniques
must be available.

For early screening experiments, where a large number of different refolding
reactions are tested in order to identify classes of conditions for further
investigation, fast analytical techniques are required. This is analogous to
high-throughput screening of candidates for drug discovery. Turbidity
measurements can be used to identify conditions which give high levels of
precipitation. For proteins with suitable fluorophores, fluorescence
measurements can be used to identify conditions in which the protein
remains unfolded. These conditions can then be “screened out” and further
work can be done based on the remaining conditions (either by further
experiments or more detailed analysis).

For later optimisation experiments, where more accurate information is
required but where fewer refold reactions are performed and so a lower
throughput is acceptable, more detailed but slower techniques (such as
chromatographic methods) will be used.

The techniques which are used in any given set of refold development
experiments will depend on the properties of the protein being refolded. If a
protein is found to aggregate readily upon refolding then those methods
which distinguish between aggregated and unaggregated protein (such as
turbidity and SEC) will be heavily used. If a protein is found to refold to a
number of stable isoforms, where only one of which is the desired product,
then methods which can distinguish the correct isoform (such as RP-HPLC or IEC-HPLC) will be needed.

The selection of assay to use for refold development experiments will be influenced by the assay development that has already been done. To use an assay which has already been developed will be an appealing alternative to developing a new assay for refold development.

A coherent system to develop refold steps for a wide range of protein product will require access to a variety of analytical equipment. This is discussed in Chapter 8.

Provided a practicable activity assay exists for the protein, then activity assays are the preferred analytical technique because they directly measure the desired property of the protein product whereas other analytical methods only measure a surrogate of this. If the assay is fast (or if assays can be performed in parallel, e.g. in microwell plate format) then it can be used throughout refold development.

Further refolding experiments carried out in this research project (Chapters 4 to 7 inclusive) will use activity assays, where this is practicable, to calculate the degree of success or failure of refolding reactions. Where activity assays are impracticable other analytical methods will be used, based upon the investigations reported in this chapter.

In refold development experiments the desired outcome is information about the performance of the refold reactions being tested. Each refold reaction is being performed in order to provide refolded protein material for assaying and only a very small quantity is required for this purpose. The limited amount of inclusion body material which is typically available for refold development experiments gives a strong incentive to perform refold experiments at the smallest scale possible (subject to the results being representative and other practical constraints). The potential of performing
refolding experiments at very small scale in microwell plates (at “microscale”) will be explored in the next chapter.
4 The use of microwell plates for protein refolding experiments

To develop a commercial protein dilution refold step, a great deal of dilution refolding experimentation must be done. This development would be facilitated if the dilution refolding experiments could be achieved at small scale and in parallel (see section 1.4). It is proposed that microwells, in the form of standard microwell plates, represent an attractive class of vessel in which to perform these refolding experiments.

This chapter discusses the rationale for using microwell plates, and how dilution refolding reactions can in practice be performed in them. It then describes work to evaluate the utility of microwells as vessels for refolding reactions.

4.1 Rationale for using microwells as reaction vessels

Microwell plates (or multiwell plates) are rectangular plates of standard dimensions (approximately 120mm by 80mm) with a number of small wells in their top surface. Each well in the plate can be used as a separate vessel. This makes these plates a convenient way of handling large numbers of reactions.

Many different types of plate are available. The most commonly used types have 24, 96, 384 or 1536 wells. These have wells with volumes between 10ml (for a deep, 24 well plate) and just a few microlitres (for a 1536 well plate). The most commonly used type of plate has 96 wells, each with a volume of 350μl. Refold development might be greatly facilitated if refolding experiments could be carried out at this small scale and still generate useful data for scale-up or selection purposes.
The most common type of plate is made of polystyrene, though plates are also available in polypropylene, nylon, PTFE, glass, quartz and other materials. Plastic plates are usually cheap enough to be disposable (though more expensive plates can be cleaned and reused). Plates are also available with different coatings, e.g. to promote cell binding or to inhibit binding. Plates are available with different well geometries e.g. round well, square well, baffled well, conical bottomed well, flat bottomed well. The low cost and variety of available plates makes them attractive experimental equipment.

A wide variety of laboratory equipment is available to work in conjunction with microwell plates e.g.:

- **Multichannel pipettes** – hand operated pipettes allowing liquids to be aspirated or dispensed to a number of microwells in one operation.
- **Robotic pipettors** – computer controlled machines to aspirate and dispense liquids into microwells.
- **Solids dispensers** – computer controlled devices to dispense a defined volume or mass of solids into wells in a microwell plate.
- **Autosamplers** – peripherals for analytical equipment which allow samples to be drawn automatically from microplates.
- **Plate handlers** – robot arms to move microplates around, stacking systems to store and dispense microplates.
- **Plate readers** – spectrophotometers, fluorimeters, nephelometers which can make readings from microplate wells instead of cuvettes.
- **Mixers** – devices to agitate the contents of microwells by shaking or using magnetic stirrers.
- **Filtration equipment** – systems which can perform filtration operations in microplate format, allowing many samples in a microplate to be filtered in parallel.
The plates are of standard geometries so many different types of plate can be used with any one piece of equipment. This range of equipment offers great scope for facilitating high throughput process development experimentation, including protein refolding experiments.

4.2 Practicalities of performing refolding experiments in microwells

In order for microwell plates to be a useful format in which to perform dilution refolding reactions, two things are necessary. First, it must be practical to perform in a microwell all the necessary operations for a refold reaction. Second, it must be possible to control in the microwell all of the variables which affect refolding. Without these characteristics, results obtained from such experiments are likely to contain artefacts. These issues are considered in this chapter.

4.2.1 Performing the operations of a dilution refold reaction in a microwell plate

This section describes how the operations necessary for a dilution refold reaction may be performed in a microwell plate.

4.2.1.1 Preparation of unfolded protein solution

Preparation of unfolded protein solution typically requires that inclusion body paste be dissolved in a strong chaotrope solution (see section 1.3). While this preparation could be performed in a microwell plate, it would be difficult due to the stickiness and viscosity of inclusion body pastes and the very small quantity of paste which would be required for one microwell-full of refold.

A more practicable approach would be to prepare a quantity (e.g. between 1ml and 10ml) of unfolded protein solution using conventional methods,
then to distribute this solution to the microwells by pipette. In this way, one aliquot of unfolded protein solution could be used for a large number of refolding reactions. This approach would rule out using microwell methods to develop inclusion body solubilisation steps. However, developing the inclusion body solubilisation step is much simpler than developing the refold step (see section 1.3), so there is not the same need for high-throughput experimentation.

4.2.1.2 Preparation of refold buffer

In the development of a commercial inclusion body protein refolding step it is likely that the variable which is most explored is the composition of the refolding buffer (see section 1.3). It is therefore important that refold buffer of different compositions can be prepared for refolding experiments.

Refold buffers are typically produced by mixing solutes (e.g. redox reagents, refolding enhancing additives, buffering agents) with water. Water and any reasonable volumes of liquid solutes can be easily added to microwells by pipette, however most of the solutes to be added will be solids. Adding solid solutes to microwells presents more difficulties. The mass of solid to be added to a microwell-full of refold buffer will be very low, typically on the order of one microgram. Adding this small mass of solid to a number of microwells in a plate would be very difficult, even with the aid of an automated solids-dispensing machine.

A more practicable approach would be to prepare a number of “stock solutions”, each being an aqueous solution containing a high concentration of a solute, with one stock solution containing no solute to be used as a diluent. Different combinations of these stock solutions could be mixed together by pipette to produce different refold buffers. In this way a large number of different refold buffers could be produced from a relatively small number of stock solutions.
For example, 250μl of refold buffer, containing 0.1M arginine, 25mM oxidised glutathione, 25mM Tris, pH7.5 could be prepared by mixing 50μl of Buffer A, 62.5μl of Buffer G and 137.5μl of Diluent, where:

- **Buffer A** is 0.5M arginine, 25mM Tris pH7.5
- **Buffer G** is 100mM oxidised glutathione, 25mM Tris, pH7.5
- **Diluent** is 25mM Tris, pH7.5

This approach would impose some artificial restrictions on the compositions of refold buffers which could be used in the experiment. The stock solutions could not contain concentrations of solute higher than the solubility limit of that solute, and so the compositions of the possible refold buffers would be limited to containing solutes at linear combinations of those solubility limits.

In practice, this restriction would not cause a great problem as it seems unlikely that a refold buffer would be used which contained a number of solutes at concentrations high enough to be close to their solubility limits. To use such a near-saturated refold buffer in manufacture would be risky – any evaporation or changes in temperature could result in solutes crystallising out of solution, altering the concentrations achieved as well as damaging pumps, filters and chromatography columns.

### 4.2.1.3 Addition of unfolded protein to refold buffer

It is simple to mimic the addition of unfolded protein solution to refold buffer for a step dilution refold in a microwell. Unfolded protein solution is pipetted into a microwell, then the refold buffer is pipetted in (or vice versa).

To mimic this addition for a continuous dilution, staged dilution or fed-batch dilution refold would be more complex. The rate and timing of the additions would need to be controlled carefully, possibly necessitating the use of a pipetting robot or a small-volume pump.
4.2.1.4 Mixing of unfolded protein and refold buffer

There are a number of ways in which unfolded protein solution and refold buffer can be mixed in microwells (see section 4.3). It is worth noting that unfolded protein solution will be considerably more viscous than refold buffer due to the high concentration of protein and chaotrope.

Shaker platforms can be used to agitate the contents of microwell plates. There are two concerns with using this method. First, vigorous shaking can cause liquid to spray from the microwells. Second, this method often produces a very poor and unpredictable degree of mixing (see section 4.3).

Magnetic stirrers, similar to conventional bench-scale magnetic stirrers can be used to agitate the contents of microwell plates. A 96-position magnetic stirrer block has been produced by H+P Labortechnik of Oberschleissheim, Germany (which can drive 96 magnetic stirrer bars in a standard 96-well plate in position on top of the stirrer block (a similar device has been produced for 24-well plates). The mixing times achievable with these microwell magnetic stirrers are controllable over a broad range, in a similar way to bench-scale magnetic stirrers. However, the handling of 96 tiny magnetic stirrer bars is a concern. The dispensing of these magnetic stirrer bars to the 96-well plate, the removal, cleaning and storage of the stirrer bars would all be difficult operations.

It is possible to conceive of a device to facilitate the handling of these stirrer bars. This device would resemble a 96 well plate with an electromagnet on it's base. The stirrer bars would be stored in this device when not in use. When needed to stir reactions in a 96-well reaction plate, the electromagnet would be activated, securing the magnetic stirrer bars to the bottom of the wells in the device. The device would then be inverted, positioned on top of the 96-well reaction plate, and the electromagnet switched off. The 96 bars would then fall (driven by gravity or magnetism) out of the 96 wells in the device and into the 96 wells in the reaction plate, where they could be driven by the stirrer block. The process would be reversed to remove the
stirrer bars from the 96-well reaction plate, and the bars could be cleaned in
the device while secured in the wells by the electromagnet.

The simplest method of mixing in microwells is pipette-mixing. In this
method, a pipette is used repeatedly to aspirate and then dispense a
volume of liquid in a microwell. Mixing can be controlled by varying the
volume, number and speed of aspirate-dispense cycles. The drawback of
this method is that the microwell can only be mixed while the pipette is at
that microwell (i.e. with one pipette, only one microwell can be mixed at one
time). This is only a minor drawback as the refold solution needs only to be
mixed to homogeneity at the beginning of the refolding reaction. The
characteristic times for these two steps are very different with the mixing
time being much shorter than that for refolding. Weiss et al (2002)
demonstrated that mixing times of less than 5 seconds can easily be
achieved by pipetted mixing in a microwell plate.

4.2.2 Control of variables affecting refolding

For a refold experiment to be useful, the conditions of the refold reactions in
that experiment must be controlled according to the experiment design that
is being followed. This section contains a discussion of how parameters
which can affect refolding (see Section 1.3) can be controlled for dilution
refold reactions in a microwell format.

4.2.2.1 Composition of refold buffer, unfolded protein solution and
refolding protein concentration

The composition of the refold buffer, the composition of the unfolded protein
solution and the concentration of the refolding protein in the refolding
protein solution can all have very strong effects on the refold reaction. All of
these factors are determined by liquid transfer operations.
Composition of the refold buffer is determined by the way in which the refold buffer is prepared (see Section 4.2.1.1). Likewise the composition of the unfolded protein solution is determined by the way in which the unfolded protein solution is prepared (see Section 4.2.1.2), while the concentration of refolding protein in the refolding protein solution is determined by the concentration of protein in the unfolded protein solution and the dilution factor used, i.e. the volumes of unfolded protein solution and refold buffer which are mixed together (see Sections 4.2.1.2 and 4.2.1.3)

**4.2.2.2 Mixing**

The intensity of mixing applied during the refolding of the protein solution, e.g. the speed at which the unfolded protein solution is mixed into the refold buffer, may have some effect on the refolding reaction (see Section 4.3). Depending on the type of mixing used in microwells (see Section 4.2.1.4), this mixing can be controlled in different ways.

Using magnetic stirring, the stirring speed and the geometry of the magnetic stirrer bar and the microwell, along with the properties of the liquid, will determine the mixing time. Using shaking as a means of mixing, the speed and throw of the shaker, the microwell geometry and the liquid properties will determine the mixing time. Using pipette mixing, the speed and volume of aspiration and dispensing, the pipette tip diameter and the well geometry will be the determining factors. Using any of the above methods of mixing in microwells, the volume of liquid in the microwell and the length of time for which the microwell is mixed will have an effect on the degree of mixing achieved.

All of the microwells in a microwell plate are fixed relative to the plate and to each other. Using a shaker it is therefore only possible to mix all of the microwells in a plate at the same time, with the same speed and the same throw.
4.2.2.3 Temperature

Microwells are small (typically 350μl), so liquid in a microwell will have little thermal inertia and so will only require a small amount of heat transfer to change temperature.

There are typically 96 microwells in a single microwell plate, so independent temperature control of each microwell would be impractical. It is unlikely that refold temperature will be extensively investigated during refold development, because refold temperature is also difficult to control at manufacturing scale (see Section 2.3.2) so this will not be considered as a variable. Instead, all microwells in a plate would be controlled to the same temperature i.e. the whole plate would be kept at one temperature. There are two conventional ways of doing this. First, the microwell plate can be kept in a temperature controlled environment such as an air conditioned laboratory, an incubator or a refrigerator. Second, the microwell plate can be held on a temperature controlled surface, e.g. a ThermoMixer or a heating/cooling block. This second method relies on good conduction of heat between the flat bottom of the microwell plate and the temperature controlled surface, and on poor conduction of heat between the top surface of the microwell plate and the environment (the top of the plate may need to be covered or insulated to achieve good temperature control with this method).

Microwells are shallow, typically containing ≈6mm depth of liquid. If a microwell plate is uncovered for long periods, a significant fraction of liquid can evaporate from the plate giving a significant change in concentration in the microwell. This evaporation is greatly increased if the microwell is heated, though evaporation can be decreased by covering/enclosing the plate or by keeping the air around the plate at high humidity.
4.2.2.4 pH
To measure the pH of the contents of a microwell by conventional methods is difficult, as the microwell is smaller than conventional pH probes. To adjust the pH of the microwell by the conventional method of adding small amounts of concentrated acid or alkali is also difficult due to the very small volumes of liquid involved.

The most practicable way of setting the pH of a refold reaction in a microwell is to set the pH of the solutions to be put into the microwell. E.g. if a refold buffer is prepared from three stock solutions, each of pH 7.5, then the refold buffer will have a pH of approximately 7.5. It may be possible to set the pH of a refold buffer by mixing together high pH buffer and low pH buffer in specific and pre-calibrated quantities to achieve intermediate pH values.

4.2.2.5 Refold time
Refold time in a microwell refold experiment is the same as refold time in a conventional bench scale refold experiment i.e. the time between the initiation of refolding (the mixing of unfolded protein solution with refold buffer) and the analysis of the refolded protein solution e.g. by activity assay or RP-HPLC. Accurate control of refold time becomes a greater concern with microwell refold experiments because of the scope for high-throughput experimentation. If it is desired to perform, for example, a microwell refold experiment involving 96 refold reactions, each of which is to have a refold time of exactly 4 hours, then each refolded protein solution must be assayed 4 hours after the initiation of refolding in each well. It may be possible to initiate 96 refold reactions in 10 minutes. However, 4 hours later, it may be impossible to perform the 96 subsequent assays that are needed within 10 minutes.

In order for each refold reaction to have the same refold time, the rate at which refold reactions are initiated must be the same as the rate at which
refolded protein solutions can be assayed. For experiments involving a variety of refold times, the initiation of refold reactions must be carefully scheduled so that each refolded protein solution is ready for assaying at a time when it can be assayed, e.g. at a time when the HPLC machine is available (see Section 6.3). This makes establishing an assay schedule crucial for microwell format experimentation.
4.3 Characterisation of the effect of mixing methods on refold yield in microwells

It has been seen that sufficiently poor mixing in a dilution refold step can lead to a reduction in refold yield. It is proposed that refold experiments for refold step development can be performed in microwells. There are a number of different methods of mixing the contents of microwells, as outlined earlier in this chapter. A decision must be made as to which method of mixing is to be used in the experiments. It is important that the mixing method which is used does not have a spurious impact on refold yield.

The purpose of the experiment described in this section was to determine which methods of microwell mixing produce adequate mixing conditions such that these do not affect refold yields.

4.3.1 Method

Hen egg white lysozyme (HEWL) was chosen as a suitable model protein for refolding. Denatured HEWL was refolded by step dilution into refolding buffer. 12 refolds were performed using each of 6 different mixing methods: diffusion (no mixing), magnetic stirring, pipette mixing and three variations of orbital shaking, making a total of 72 runs (72 refold reactions) in 72 microwells.

4.3.1.1 Materials

All chemicals, including protein, were purchased from Sigma-Aldrich Company Ltd (Gillingham, Dorset, UK). The microwell plates used were Corning Costar Ultra Low Attachment 96 well microwell plates (Corning Inc., Corning, New York, USA). These have a recommended working volume of 200µl.
4.3.1.2 **Microwell protein refolding**

Unfolded HEWL was prepared according to the method of Section 4.5.1.2. 96-well plates were prepared by soaking in refold buffer for one hour before use to hydrate the well's hydrogel coating.

HEWL was refolded in microwells by the dilution of 10μl of unfolded HEWL solution with 190μl of refold buffer (25mM cysteine, 1mM EDTA, 0.1M Tris-HCl, pH 8.2). All pipetting operations were done by hand using Gilson pipettes. Refolding protein solutions were allowed to stand for 2 hours at room temperature (20±3°C) after mixing. Refolded protein solutions were pipette mixed by 5 cycles of 190μl aspirate-dispense before being sampled. 6 methods of mixing were tested, with 12 runs being performed using each method, making a total of 72 runs.

4.3.1.3 **Diffusion (no mixing)**

For each run (each refold reaction), 190μl of refold buffer was pipetted into the microwell. 10μl of unfolded HEWL solution was then carefully pipetted onto the base of the microwell, and allowed to stand for 2 hours. The microwell plate was kept still throughout.

4.3.1.4 **Magnetic stirrer mixing**

For each run (each refold reaction), a 4mm PTFE coated magnetic stirrer bar was placed in the microwell using plastic tweezers, and the microwell plate was positioned on a Variomag digital magnetic stirrer set to 1400rpm. 190μl of refold buffer was pipetted into the microwell, and the magnetic stirrer was switched on. 10μl of unfolded protein solution was then pipetted into the well. The microwell was stirred for 5 minutes then the stirrer bar was removed using a magnet, and the refolding protein solution was allowed to stand for 2 hours.

4.3.1.5 **Pipette mixing**

For each run (each refold reaction), 10μl of unfolded HEWL solution was pipetted into one well of a 96-well plate, forming a droplet on the base of the well. 190μl of refold buffer was pipetted into the well, with the tip in contact
with the droplet of unfolded HEWL solution. The contents of the well was immediately pipette-mixed by three cycles of aspirating and dispensing 190µl. Refolding protein solutions were allowed to stand in the plate for 2 hours.

### 4.3.1.6 Orbital shaker

A 96-well plate was shaken using a Thermomixer orbital shaker platform. 10µl unfolded HEWL solution and 190µl refold buffer were added to wells, and the orbital shaker was switched on. The speed of the shaker was increased until refolding protein solution was seen to be close to spraying out of the wells. This represented the maximum agitation possible and was in the region of 1400rpm. The shaker speed was then stepped back to 700rpm, well below the point at which loss of liquid by spraying could have been an issue. The plate was discarded after each experiment and the shaker cleaned prior to the next.

Three methods of mixing were tested using the orbital shaker. In the first method, 10µl of unfolded HEWL solution was pipetted onto the bottom of the well, then 190µl of refold buffer was pipetted into the well and the plate shaken at 700rpm. In the second method, 190µl of refold buffer was pipetted into the well, then 10µl of unfolded HEWL. The plate was then shaken at 700rpm. In the third method the plate was shaken at 700rpm and, while the plate was being shaken, 10µl of unfolded HEWL was pipetted into the well followed by 190µl of refold buffer (this was possible because the throw of the orbital shaker was less than the radius of the microwell). In each method, refolding protein solutions were shaken for 5 minutes at 700rpm, then allowed to stand for 2 hours at room temperature (20±3°C).

### 4.3.1.7 Activity assay

The yield of recovered activity from each refold reaction was assayed using the technique described in Section 3.4. Refolded protein solutions were pipette mixed by 5 cycles of 190µl aspirate-dispense before being sampled.
4.3.2 Results

Table 4-1 Table of averages and standard deviations (from eight replicates) of yields obtained from step dilution refolds mixed by different methods in microwells.

<table>
<thead>
<tr>
<th>Method</th>
<th>Mean</th>
<th>Median</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Mixing</td>
<td>0.33</td>
<td>0.32</td>
<td>0.020</td>
</tr>
<tr>
<td>Magnetic stirrer (1400rpm)</td>
<td>0.51</td>
<td>0.52</td>
<td>0.060</td>
</tr>
<tr>
<td>Pipette Mixing</td>
<td>0.55</td>
<td>0.55</td>
<td>0.026</td>
</tr>
<tr>
<td>Shaker: protein then buffer then mix</td>
<td>0.30</td>
<td>0.29</td>
<td>0.054</td>
</tr>
<tr>
<td>Magnetic stirrer (1400rpm)</td>
<td>0.23</td>
<td>0.23</td>
<td>0.028</td>
</tr>
<tr>
<td>Shaker: protein then buffer during mixing</td>
<td>0.52</td>
<td>0.52</td>
<td>0.020</td>
</tr>
</tbody>
</table>

Figure 4-1 Graph of yields obtained from step dilution refolds mixed by different methods in microwells.
4.3.3 Discussion and conclusions

The yield of dilution refold reactions can be affected by the mixing of the denatured protein solution with the refolding buffer. Like many proteins, lysozyme irreversibly aggregates during refolding, which is the major cause of loss of refold yield. This aggregation has been shown to be a fast reaction (commitment to the aggregation pathway is essentially complete within 1 minute) of higher than second order, (Goldberg et al, 1991, Zettlmeisl et al, 1979) while correct refolding is a competing slower, first order reaction (with a half-time of minutes to hours, depending on folding conditions). Lower protein concentration therefore favours correct refolding.

In a dilution-refold reaction, the protein is initially at high concentration in the denatured protein solution. The dispersal of the protein into the refold buffer reduces the protein concentration during refolding which reduces aggregation and hence increases yield (relative to the yield which would be obtained if the protein were refolded at the concentration in the denatured protein solution). However, when the denatured protein solution is added to the refold buffer, a high protein concentration will transiently exist until the protein is dispersed. If the time required for this dispersal is significant (relative to the irreversible aggregation reaction) then this may increase aggregation thus reducing refold yield. Lee et al (2002) showed that increasing mixing intensity increased refold yield of lysozyme up to a maximum when dilution-refolding is performed in a stirred-tank reactor or in an oscillatory-flow reactor. However, they only found this effect when the protein had been denatured in urea, not GdnHCl. In contrast, Buswell et al (2002) found that the yield of trypsinogen from a oxidative dilution refold reaction performed in a stirred tank reduced when the stirrer speed was increased. However they only found this effect in the absence of baffles, when the higher stirrer speed induced vortexing. With baffles, vortexing was reduced and the increase in stirrer speed had no effect on yield. The refold buffer used in this experiment was 50mM CaCl₂, 5mM Tris, 3mM cysteine, 1mM cystine, pH9. The presence of metal ions in the refold
buffer, combined with the absence of chelator and the low concentration of redox reagents suggests that air oxidation may have had an effect on refolding. In which case, the increased aeration provided by the vortexing may have sufficiently changed the redox conditions to affect refolding and the reduction in yield seen at higher impeller speed may have been a redox effect rather than a mixing effect.

In the experiments reported in this section, refold yields are seen to be higher when the dilution-refold is well-mixed. The mixing conditions which gave reduced refold yields (the “no mixing” method and two of the shaker-mixing methods) appear to give mixing times which are significant relative to the rate of the irreversible aggregation reaction. The results reported in this section support the conclusion that poor mixing in dilution-refold reactions can give reduced refolding yields and demonstrate that this effect can occur when the protein has been denatured in GdnHCl.

The refolds without deliberate mixing gave low yields, as expected. The spread of yields from this group was also very small.

The refolds mixed using the magnetic stirrer gave higher yields, suggesting that the mixing was fast enough to prevent poor mixing affecting yield (i.e. the mixing time was fast relative to the rate of irreversible aggregation). Although no obvious material loss was seen from the magnetically stirred microwells, it may be that there was a small amount, and this could have contributed to the large spread of yields.

The refolds which were mixed using the first and second orbital shaker methods gave low refold yields, while those mixed using the third orbital shaker method gave higher refold yields. This may be because orbital shaking gives better mixing when there is a smaller volume of liquid in the microwell. In the first two methods 200μl of liquid is present in the microwell before shaking is started but in the third method, liquid is added to the well while the plate is being shaken. Consequently there will be a short period
when the microwell is being shaken with a small amount of liquid in it and when mixing will be more intense and so faster. The higher yield for the third orbital shaker method may also be due to the combination of shaking together with the jet of refold buffer entering the microwell (effectively a component of pipette mixing) giving significantly better mixing than shaking alone. Weiss et al (2002) report that mixing the contents of a microwell with pipette aspirate-dispense cycles gives much faster mixing than a microplate shaker. They found that mixing times of less than 5 seconds (effectively their lower limit of detection) were achieved using pipette mixing, compared to mixing times of several minutes using shaking at 350 rpm without pipette mixing. This indicates that mixing by microplate shaking would give a mixing time which was significant (or long) compared to the characteristic time of the irreversible lysozyme aggregation reaction and so it is unsurprising that mixing by shaking gives reduced refold yields. The higher yield for the third orbital shaker method may suggest that mixing is critical only in a short period of time after addition of the liquids. In the third method, the plate is already being shaken when the two liquids are pipetted in. In the first and second method there is short period of a few seconds between the two liquids coming into contact with each other and the shaking beginning. The refolds which were pipette mixed gave the highest yields and had a small spread, indicating that this is a good and reproducible method of mixing and will be the preferred method of mixing microwell dilution refold reactions for refold development experiments.

The fact that all of the groups of refolds had standard deviations of yield of at least 0.02 confirms that there is a degree of inherent variability in either the pipette volume accuracy, the refold reaction or the HEWL activity assay. This level of standard deviation is relatively small however and enables direct comparison between the outputs of the different contacting methods investigated.
4.4 Loss of protein due to non-specific binding during refolding in microwell plates

Refolding reactions are typically allowed to run for between 1 and 24 hours and are done in aqueous solutions at very low protein concentrations (typically between 10μg/ml and 100μg/ml protein). During refolding the structure of the protein changes and hydrophobic patches may become temporarily or permanently exposed. It may therefore be expected that some protein will become hydrophobically bound to surfaces in the refold vessel.

It is proposed that optimisation of refold steps be done using refolding experiments which are run in microwell plates. The working volumes of the wells in these plates are in the range of 0.1ml to 10ml. Such small volumes have very large surface-area:volume ratios compared to manufacturing scale process equipment. For example, a 1000L refold in a typical process vessel such as a jacketed stainless steel tank or a disposable bag may have a wetted surface area of approximately 4.5m², giving a surface area to volume ratio of 4.5m²/m³. A 1ml refold in one well of a 24-well plate would have a surface area of approximately 5cm², giving a surface area to volume ratio of 500m²/m³. Even if non-specific binding does not lead to significant loss of protein in large-scale processes, it may do at the microwell scale. This would adversely affect the results of the experiments.

An experiment was performed to investigate the binding of protein to microwell plates during refolding. Protein was denatured in a strong chaotrope then refolded by dilution in a number of different microwell plates with different surface chemistries.

The experiment had two purposes. First, it would allow an estimate to be made of the amount of refolding protein lost to the surface of the well in any future experiments using microwell plates. Second, it would allow the
selection of the best microwell plate to minimise loss by non-specific binding in future experiments.

4.4.1 Method

4.4.1.1 Solutions
Hen egg white lysozyme (HEWL) was chosen as a suitable model protein for refolding. A sample of HEWL was denatured, reduced and dissolved at 10 mg/ml in 6M guanidine hydrochloride (GdnHCl) with 0.1M acetic acid according to the method of Roux et al (1997) and Goldberg et al (1991). Two refolding buffers were prepared: Buffer A containing 0M GdnHCl, 0.1M Tris-HCl, pH 8.2, 1mM EDTA, 20μM DTT, 60μM GSSG. Buffer B containing 0.3M GdnHCl, 0.1M Tris-HCl, pH 8.2, 1mM EDTA, 20μM DTT, 60μM GSSG.

4.4.1.2 Multiwell plates
Samples of Greiner Cellstar Tissue Culture, Greiner Cellstar Suspension culture and Corning Costar Ultra Low Attachment 24 well multiwell plates were obtained and prepared according to manufacturers instructions. The wells each had a wetted area of 430mm² when filled with 1ml of solution.

4.4.1.3 Refolding
A 10μl aliquot of the denatured HEWL in 6M GdnHCl was placed in each well of the multiwell plate. 990μl of refolding buffer was vigorously added to each well and immediately mixed using 5 pipette aspirations of 990μl. All pipetting operations were done by hand using Gilson pipettes. Buffer A was used in 12 wells in each plate, and buffer B used for the other 12. The plate was maintained at 5°C during refolding to prevent microbial spoilage. Absorbance measurements were taken at 1 minute, 1 hour, 4 hours and 24 hours after the initiation of refolding. At each time point, samples were
taken from six wells, three containing HEWL refolded in buffer A and 3 containing HEWL refolded in buffer B.

4.4.1.4 Measurement of unbound protein concentration

The quantity of HEWL bound to the well was determined by observing the disappearance of HEWL in solution and in unbound aggregates. Samples of the well contents were taken and made up to 6M GdnHCl solutions to dissolve and unfold all of the unbound HEWL.

Each cuvette was blanked using the appropriate solution. The contents of the well were agitated by 5 pipette aspirations of 250μl immediately before absorbance measurements were taken, to resuspend any aggregates not bound to the well. A 250μl sample was drawn from the well and mixed with 750μl of 8M GdnHCl. The sample was allowed to equilibrate in the cuvette for 30 minutes. An absorbance reading was then taken at 280 nm using Kartell UV grade Semi Micro Cuvettes (Kartell S.p.A., Novilio, Italy) in a Uvikon 922 spectrophotometer (Uvikon, Bunnik, Netherlands).

4.4.2 Results

The six plots in Figure 4-2 show the decrease in the concentration of unbound HEWL (both soluble and precipitated) with time after the initiation of refolding, in the three microwell plate types and in both refold buffers.
Figure 4-2 Graphs showing the reduction in concentration of unbound protein (both soluble and precipitated) with time. Plotted value is the concentration of unbound protein, relative to the initial concentration.
4.4.3 Discussion and Conclusions

Hen egg white lysozyme is a soluble 14.4kD protein containing 4 disulphide bonds which make a significant contribution to the structural stability of the protein. In its native form it is a two lobe globular structure separated by a cleft in one side of the molecule which contains the active site. One lobe is dominated by α-helices and the second by a three stranded anti-parallel β-sheet.

The oxidative folding of reduced lysozyme has been widely studied and discussed (Goldberg et al (1991), Radford et al (1992), Wildegger and Kiefhaber (1997), Roux et al (1997), Buswell and Middelberg (2002)) and has been found to be complex. Upon initiation of refolding, the lysozyme undergoes very rapid hydrophobic collapse. An intermediate state is formed. Radford et al (1992) showed that, in this intermediate state, the α-domain is fully folded but the β-domain, comprising β-strands, is not fully folded. This intermediate folds to the correctly folded structure in a much slower reaction, taking minutes to hours depending on conditions. Lysozyme also has a direct folding pathway, without an intermediate and ≈20% of lysozyme may fold along this direct pathway. The intermediate state, being partially folded, is more stable than the unfolded state so the free energy of activation of the transition from intermediate to folded is greater than that of the direct transition from unfolded to folded states. An alternative model for the folding of lysozyme has the protein passing from the unfolded state to the folded state via a number of intermediate states, some of which are partially active.

Intermediate states, which have some hydrophobic surfaces exposed, are the most prone to aggregation. The surface of the native protein is hydrophilic. If binding of the refolding protein to the surface of the vessel (in this case, the microwell) were hydrophobic then this would lead to the expectation that the rate of loss to surface-binding would be highest when

- 134 -
the intermediate states were most populated and would tail off as the lysozyme refolding goes to completion. However, Buswell and Middelberg (2002) demonstrated that native (folded) lysozyme could also become incorporated into aggregates during refolding which indicates that native lysozyme can go into aggregates, possibly via an intermediate.

The rate of decrease in unbound protein concentration (reflecting the rate of surface-binding) is seen to slow with time (note that the time-scales in Figure 4-2 are logarithmic) and is substantially complete after 4 hours which supports the theory that it is the unfolded or intermediate states which bind to the vessel surface, whereas the native state is not prone to surface-binding. The fact that the surface-binding decreases with time may also be due to the potential binding sites on the well surface becoming saturated with protein (i.e. the well surface becomes blocked).

Tissue culture plates are made of polystyrene. Their surfaces have been treated by the addition of charged functional groups. This encourages tissue cells to bind to the surface. Suspension culture plates are untreated polystyrene plates whose surfaces are very hydrophobic. The Ultra Low Attachment plates are coated with a hydrogel layer which is hydrophilic and neutrally charged. The tissue culture plate (Greiner Cellstar) bound more HEWL than the suspension culture plate (Greiner Cellstar), which suggests that the binding of the protein to the plate may be at least partly ionic. The tissue culture plates also continued to bind protein after 4 hours, when the lysozyme refold reaction would be expected to have substantially gone to completion, suggesting that this surface can continue to bind protein even when it is in forms with no exposed hydrophobic patches. Unsurprisingly the Corning Costar Ultra Low Attachment plates, having uncharged hydrophilic hydrogel surfaces, bound the least protein.

One conclusion from this experiment is that, of the three types of surface examined, the hydrogel coated surface appears the most promising for the testing of refolding reactions since it binds the least HEWL. The difference
between the amounts of protein that it binds in the two buffers is also the smallest. During experimentation to determine the optimum conditions for refolding, a wide variety of different buffers will be used. A surface which binds more consistently in different buffers will reduce experimental uncertainties.

A survey of the available literature revealed no publications on the loss of protein due to binding to the vessel surface during dilution refolding.

Regardless of the surface chemistry of the microwell plate, the presence of 0.3M GdnHCl in the refold buffer appears to prevent the binding of significant quantities of protein to the well surface. The means by which low concentrations of chaotrope inhibits this loss to binding may be similar to the way in which low concentrations of chaotrope inhibits loss to aggregation (see section 2.3.4). Another conclusion drawn from this experiment is that in refolding reactions where there will be low concentrations of chaotrope (e.g. 0.3M GdnHCl) in the refolding protein solution, the surface chemistry of the plate may have little effect and in such cases, material selection is less critical. The refolding protein solution produced in any dilution refolding reaction which uses a moderate dilution factor (typical of those used in commercial manufacture) will contain a low concentration of chaotrope carried over from the chaotrope solubilised protein solution (e.g. a protein solubilised in 7M urea, refolded by 1+19 dilution into a chaotrope-free refold buffer, will be refolding in a solution containing 0.35M urea).

### 4.5 Scalability of refolding reactions from microwells

It is proposed that development of commercial inclusion body dilution refold steps be done using refolding experiments performed in microwell plates. This can only be done if the results obtained at microwell scale are representative of results obtained at manufacturing scale, i.e. if the refolding yield obtained in a refolding step performed under certain conditions at
microwell scale is similar to the refolding yield obtained under similar conditions at large scale.

Dilution refold steps for commercial manufacture are carried out at very large scale (typically 100L-50,000L). Unfortunately, experiments at this scale would be prohibitively expensive. Dilution refold experiments for commercial refold step development are conventionally carried out at bench scale (100ml-1L scale). The results from experiments at this scale have been found to agree well with results at manufacturing scale. (Liddel, J., 2001 Avecia pers.comm. and Watkins, D., 2000 Dowpharma pers.comm.). The purpose of the experiment described in this section is to determine whether refold yields obtained at microwell scale agree well with refold yields at scales up to 200ml, and hence to determine whether microwell experiments may be used in place of conventional bench scale experiments for refold step development.

4.5.1 Method
Hen egg white lysozyme (HEWL) was denatured then refolded by 1+19 (i.e. 1 to 20) dilution refolding in three scales of vessel: 200µl microwell, 10ml stirred tank (CSTR) and 200ml stirred tank (CSTR). At each scale, nine different refolding conditions were tested (using nine different refold buffers). All refolds were done at room temperature (20±3°C), and refolding protein solutions were allowed to stand for four hours after dilution before assaying. All refolds were vigorously mixed upon dilution to minimise any effect of poor mixing upon refold yield. All pipetting operations were done by hand using Gilson pipettes.

4.5.1.1 Materials
All chemicals, including protein, were purchased from Sigma-Aldrich Company Ltd (Gillingham, Dorset, UK).
4.5.1.2 Preparation of unfolded lysozyme

Hen egg-white lysozyme (HEWL) was dissolved at 10 mg/ml in a denaturing solution containing 6M guanidine hydrochloride (GdnHCl), 25mM dithiothreitol (DTT). The solution was vortex mixed for 1 minute and incubated for 2 hours at 25 °C.

4.5.1.3 Refolding buffers and stock solutions

Refolding buffers were prepared by mixing together volumes of refolding stock solutions. Four refolding stock solutions were prepared. Each stock solution contained 25mM cysteine, 1mM EDTA, 0.1M Tris-HCl, pH 8.2. Three of the refolding stock solutions each contained a refold enhancing additive (20g/L PEG3350, 100g/L arginine-HCl and 300g/L sucrose respectively). One refolding stock solution contained no additive. Thus, by mixing together these stock solutions, refolding buffers could be prepared which could contain various quantities of the reported refolding enhancing additives, but which would always contain 25mM cysteine, 1mM EDTA, 0.1M Tris-HCl, pH 8.2. Nine refold buffers were prepared in this way (see Table 4-2).
Table 4-2 Composition of nine refold buffers.

<table>
<thead>
<tr>
<th>Refold buffer number</th>
<th>PEG3350 stock solution</th>
<th>Sucrose stock solution</th>
<th>Arginine stock solution</th>
<th>Additive-free stock solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>100</td>
</tr>
<tr>
<td>2</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>75</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>75</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>25</td>
<td>75</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>25</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>25</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>7</td>
<td>25</td>
<td>0</td>
<td>25</td>
<td>50</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>50</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

4.5.1.4 Microwell plate refolding

Microwell refolds were carried out in Corning 96-well polystyrene 350μl plates from Corning Inc. (Corning, New York, USA). These have a recommended working volume of 200μl.

Nine refolding conditions were tested at microwell scale, each condition being a dilution refold of unfolded HEWL solution into one of the nine refold buffers (see Table 4-2). Each condition was tested in triplicate, making a total of 27 runs.

For each run (each refold reaction), 10μl of unfolded HEWL solution was pipetted into one well of a 96-well plate, forming a droplet on the base of the well. 190μl of refold buffer was then pipetted into the well, with the tip in contact with the droplet of unfolded HEWL solution. The contents of the well was immediately pipette-mixed by five cycles of aspirating and dispensing 190μl. Refolding protein solutions were allowed to stand in the plate for four hours before assaying.
4.5.1.5 10ml CSTR refolding

10ml CSTR refolds were carried out in 50ml polycarbonate centrifuge tubes (Beckman Coulter Ltd., California, USA), which were mixed using 7mm magnetic stirrer bars and a Variomag 9-position magnetic stirrer (H+P Labortechnik AG, Oberschleissheim, Germany).

Nine refolding conditions were tested at 10ml scale, each condition being a dilution refold of unfolded HEWL solution into one of the nine refold buffers (see Table 4-2). Each condition was tested in triplicate, making a total of 27 runs.

For each run (refold reaction), 9.5ml of refold buffer was stirred in a centrifuge tube at 200rpm. 0.5ml of unfolded HEWL solution was pipetted into the stirring refold buffer. Each refolding protein solution was mixed for 5 minutes in the centrifuge tube on the magnetic stirrer before the tube was removed from the stirrer and the solution allowed to stand for four hours before assaying.

4.5.1.6 200ml CSTR refolding

200ml CSTR refolds were carried out in a 250ml baffled glass beaker stirred using a 19mm diameter Rushton turbine driven by a Eurostar Labortechnik digital stirrer.

Nine refolding conditions were tested at the 200ml scale, each condition being a dilution refold of unfolded HEWL solution into one of the nine refold buffers (see Table 4-2). Due to time constraints, only condition 1 was tested in triplicate (making a total of 11 runs).

For each run (each refold reaction), 190ml of refold buffer was stirred in the vessel at 200rpm. 10ml of unfolded HEWL solution was poured into the
stirring refold buffer. Each refolding protein solution was mixed for 5 minutes in the stirred vessel. The stirrer was then stopped and the refolding protein solution was poured out into a normal 250ml glass beaker and allowed to stand for four hours before assaying. The baffled beaker and Rushton turbine were then cleaned prior to the next run.

4.5.1.7 Activity assay

The yield of recovered activity from each refold reaction was assayed using the technique described in section 3.4.

Pearson product moment coefficient calculations were carried out using Microsoft Excel software. This coefficient is an index that indicates the strength of a linear relationship between two sets of data sets. This coefficient has a range of -1 to +1, -1 indicating a very strong negative correlation, +1 indicating a very strong positive correlation, and 0 indicating no correlation at all.

4.5.2 Results

Figure 4-3, Figure 4-4 and Figure 4-5 are comparison plots which compare the yields of recovered HEWL activity from refold. Figure 4-3 compares the yields of microwell and 10ml CSTR refolds. Figure 4-4 compares the yields of microwell and 200ml CSTR refolds. Figure 4-5 compares the yields of 10ml and 200ml CSTR refolds. 9 points are plotted on each graph, each point representing the refolding yield under 1 condition at two different scales (the yield at one scale on the y-axis, the yield at a second scale on the x-axis).

Figure 4-6 shows the yields from each of the 3 scales at each of the 9 conditions.
For the yields from the microwell and 10ml CSTR, error bars on the graph show one standard deviation each side of the mean (each condition was tested in triplicate at each scale). For yields from the 200ml CSTR, error bars are only shown for condition 1, which was the only condition tested in triplicate at this scale.

Table 4-3 shows the Pearson product moment correlation coefficients between the sets of yield data at the three scale.
**Figure 4-3** Comparison plot showing the yield of recovered activity from dilution refold steps carried out under similar conditions at microwell and 10ml scales. Each point represents one condition. The yield of refolds at those conditions at microwell scale is plotted on the x-axis, the yield of refolds at those conditions at 10ml scale is plotted on the y-axis.
Figure 4-4 Comparison plot showing the yield of recovered activity from dilution refold steps carried out under similar conditions at microwell and 200ml scales. Each point represents one condition. The yield of refolds at those conditions at microwell scale is plotted on the x-axis, the yield of refolds at those conditions at 200ml scale is plotted on the y-axis.
Figure 4-5 Comparison plot showing the yield of recovered activity from dilution refold steps carried out under similar conditions at 10ml and 200ml scales. Each point represents one condition. The yield of refolds at those conditions at 10ml scale is plotted on the x-axis, the yield of refolds at those conditions at 200ml scale is plotted on the y-axis.

\[ y = 0.9277x + 0.0543 \]

\[ R^2 = 0.9575 \]
Figure 4-6 Graph of the yield of recovered activity from dilution refold steps carried out under similar conditions at microwell, 10ml and 200ml scales.

Table 4-3 Table of correlation coefficients between data sets

<table>
<thead>
<tr>
<th></th>
<th>Microwell and 10ml CSTR</th>
<th>10ml CSTR and 200ml CSTR</th>
<th>Microwell and 200ml CSTR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson product moment correlation coefficient, $r$</td>
<td>0.98</td>
<td>0.96</td>
<td>0.98</td>
</tr>
</tbody>
</table>
4.5.3 Discussion and conclusions of the experiment

A survey of the available published literature revealed no studies on the scalability of dilution refolding. However, Middelberg (2002) described dilution refolding from chaotrope-denatured protein as “a simple and approximately scale-invariant method of protein refolding” which agrees well with information received from Avecia Biologics during discussions about their experience of bench-scale and manufacturing-scale refolding. The operation of dilution refolding is simple: the dilution of one liquid into another. The only factors which would therefore be expected to complicate scaling are mixing (it being easier to mix small volumes than large volume) and surface effects (smaller volume reactors having much larger surface area : volume ratios than larger reactors). It has been seen that mixing can affect refold yield if the mixing is sufficiently slow, but that refold reactions in microwells can easily be mixed sufficiently rapidly to prevent this (see section 4.3). It has also been seen that loss of protein to surface binding is minimal if hydrogel-coated microwell plates are used or if a low concentration of chaotrope is present in the refolding protein solution (see section 4.4). The dilution refold experiments described in this section were rapidly mixed and (in common with most commercial dilution-refold development experiments) contained a low concentration of denaturant in the refolding protein solution because there was a high concentration of denaturant present in the denatured protein solution. It would therefore be expected that refolding would be seen to scale well i.e. that refold yields obtained at the microwell scale (<1mL) would agree well with yields at the 10mL scale and bench scale (200mL).

If refolding scaled perfectly between microwell, 10ml CSTR and 200ml CSTR scales, using these methods, and the yields achieved were perfectly repeatable, then the comparison plots would show 9 data points all on a line of $y=x$. The results obtained were close to these, which indicates that refold yields obtained at microwell scale agree well with refold yields obtained at bench (200ml) scale. This confirms that microwell refold
experiments could be used in place of conventional bench scale refold experiments.

The three types of vessel used for these refold experiments had slightly different geometries, in addition to having different volumes. The microwell and 200ml CSTR vessels were both flat bottomed cylinders of similar aspect ratio. The centrifuge tube which was used as the 10ml CSTR vessel was also a cylinder with a similar depth:width ratio, however it had a curved bottom. This would give the vessel a slightly different surface-area:volume ratio and could be expected to affect mixing patterns in the vessel. However, in spite of this difference in geometry, good agreement was still found between the yields obtained at different scales, which would support the conclusion that dilution refolds under these conditions are insensitive to differences in surface-area:volume ratio and small differences in mixing.

Pearson correlation coefficients between the data sets were close to 1, indicating good correlation between results. This is particularly significant for process optimisation experiments. In such cases the experimenter is seeking the greatest yield and it is less important that the experiment predict exactly the yield at large scale and more important that the condition giving the best yield at small scale should also give the best yield at large scale.

The refold yields at microwell scale showed a greater spread than the yields achieved at the larger scales (standard deviations for microwell scale were greater than standard deviations for 10ml and 200ml scale). This may be due to the difficulty in accurately dispensing 10μl of unfolded protein solution, which is a viscous liquid, leading to variations in the quantity of unfolded protein added to each refold. It may be due to mixing by human-operated pipette in the microwell refolds being more variable than mixing by digitally-controlled stirrer in the CSTR refolds.
4.6 Conclusions of the chapter

As discussed in Section 4.1 and Section 4.2, microwell plates are an attractive and feasible format in which to perform dilution refold experiments for the development of commercial inclusion body refold steps. The experiments described in Section 4.3 show that either pipette mixing or orbital shaker mixing are suitable methods for the mixing of dilution refold reactions in microwells, with pipette mixing being the preferred method. The experiments described in Section 4.4 show how to perform experiments without the non-specific binding of refolding protein interfering with results. The experiments described in Section 4.5 show that refold yields obtained from microwell refolding reactions (typical of those which would be performed in development of a commercial dilution-refold step) reflect the refold yields obtained from refolding reactions at larger scale, and so can be used for the optimisation of commercial refold steps.

The use of microwell plates allows such experiments to be performed using a very small quantity of protein, which satisfies one of the aims of the research (see Section 1.5). This approach also facilitates parallel experimentation, which facilitates the rapid acquisition of data and so partly satisfies another aim of the research (see Section 1.5). In Chapter 5, experiments will be described which demonstrate the use of this approach to generate useful data for the process development of a commercial protein refold step.
5 Application of microwell plate refolding experiments: development of the protein refolding step for the production of IGF-1 from E.coli inclusion bodies

At the time of the experiments described in this chapter, Avecia was developing a process to produce recombinant human insulin-like growth hormone 1 (IGF-1) for a client company for use in early phase clinical trials. IGF-1 is a monomeric, disulphide-bonded soluble protein which was expressed in E.coli in the form of inclusion bodies. A downstream process had previously been developed for IGF-1, but process development work was being done to increase the yield and decrease the cost of the process.

The experiments described in this chapter had two purposes. First, to provide information about the refolding of IGF-1 to aid Avecia in process development. Second, to demonstrate the usefulness of microwell refold experiments in providing refold data using very small quantities of protein.

5.1 Background

The existing refold step, which was used as the “starting point” for development, was a dilution refold step. Unfolded IGF-1 solution (solubilised IGF-1 inclusion bodies) at ≈8.5mg/ml IGF-1 in ≈7M urea was diluted into refold buffer containing 5mM cystamine to oxidise disulphide bonds, giving a final volume of 20litres of refolded IGF-1 solution per litre of unfolded IGF-1. Among the issues that were being addressed in process development were the low yield of native IGF-1 and the large process volume of the IGF-1 refold step.

Previous experiments performed by Avecia had shown that good refold yields could be obtained by using a refold buffer containing 20%v/v ethanol and a salt. SHE concerns led Avecia to believe that it would be impractical
to use an ethanol-containing refold buffer at manufacturing scale. Also, with the existing process, each batch of refolded protein would be large (due to the 20-fold dilution) leading to high processing costs. Avecia therefore sought three areas of information from these microwell scale refold experiments:

1. Which potential refolding enhancing additive would give the highest refold yield?
2. Did a less volatile alternative exist, which can be used in place of ethanol in the refold buffer, but which gave similar refold yields?
3. Can the refold be done at higher protein concentrations without causing a significant drop in refold yield?

IGF-1 had not yet been produced at large scale in Avecia, so only a small amount of unfolded protein was available for experiment.

Experiments were designed in consultation with Avecia.

5.2 Materials and Methods

5.2.1 Chemicals

Unfolded IGF-1 at 8.5mg/ml in 7M urea and IGF-1 1mg/ml standard solution was supplied by Avecia Biologies of Gillingham, Cleveland. All other chemicals were purchased from Sigma-Aldrich Company Limited of Gillingham, Dorset, UK.

5.2.2 Microwell refold method

In order to filter and load samples into the available HPLC machine, sample volumes of at least 1ml were required. Refolds were therefore performed at 2ml scale in microwells.
Microwell refolds were carried out in Corning Costar Ultra Low Attachment 24 well plates (see Section 4.4). These were soaked in refold buffer for 1 hour before use to hydrate the hydrogel coating of the well.

For each run (each refold), unfolded IGF-1 solution (between 100\(\mu\)l and 400\(\mu\)l, depending on the dilution factor used) was pipetted into a well. Refold buffer (between 1600\(\mu\)l and 1900\(\mu\)l) was then added to that well to make the volume in the well up to 2ml. The contents of the well was immediately pipette mixed by 5 cycles of aspirating and dispensing 1ml, and allowed to stand at room temperature (20\(\pm\)3°C) for four hours. 1ml of the refolded protein solution was then taken for assaying.

5.2.3 Refold buffers

Refold buffers were prepared from stock solutions using a similar method to that described in Section 4.4.1. All stock solutions contained 25mM Tris, pH8.0. 50\(\mu\)l of 0.2M cystamine solution (final concentration 5mM cystamine) was added to each refold to oxidise disulphide bonds.

5.2.4 Additive screening experiment

A screening experiment was designed in which 13 potential refold enhancing additives were tested in the presence or absence of ethanol. Some of these additives have previously been tested with IGF-1 refolding (see Section 5.4 for further discussion), others were chosen based on their use in the refolding of other proteins (see Section 2.3.4) and on suggestions from Avecia. 28 refold buffers in total were used in the screening experiment, 14 contained 20%v/v ethanol and 14 contained no ethanol (1 of each set of 14 contained no other additive). All refold buffers contained 5mM cystamine, 25mM Tris, pH8.0.
Table 5-1 Additives tested in screening experiment

<table>
<thead>
<tr>
<th>Additive</th>
<th>Concentration in refold buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>arginine</td>
<td>0.25M</td>
</tr>
<tr>
<td>aspartate</td>
<td>10mM</td>
</tr>
<tr>
<td>glycine</td>
<td>10mM</td>
</tr>
<tr>
<td>propanol</td>
<td>20%v/v</td>
</tr>
<tr>
<td>glycerol</td>
<td>20%v/v</td>
</tr>
<tr>
<td>PEG300</td>
<td>2mg/ml</td>
</tr>
<tr>
<td>PEG3350</td>
<td>2mg/ml</td>
</tr>
<tr>
<td>sucrose</td>
<td>300mg/ml</td>
</tr>
<tr>
<td>trehalose</td>
<td>0.01M</td>
</tr>
<tr>
<td>GdnHCl</td>
<td>0.25M</td>
</tr>
<tr>
<td>NaCl</td>
<td>0.25M</td>
</tr>
<tr>
<td>urea</td>
<td>0.25M</td>
</tr>
<tr>
<td>reduced glutathione</td>
<td>1mM</td>
</tr>
</tbody>
</table>

Table 5-2 Additives tested in screening experiment

Refold reactions were initiated by diluting 100µl of unfolded IGF-1 solution with 1900µl of refold buffer and were allowed to proceed for 4 hours before analysis.

5.2.5 Ethanol substitute experiment

Three potential substitutes for ethanol in the refold buffer were tested: n-propanol, glycerol and propylene glycol. Propanol and glycerol were selected because they were less hazardous than ethanol and had been shown in other studies (Hart et al, 1994) to have a beneficial effect on IGF-1 refold yield under some conditions. Propylene glycol was selected because it is chemically similar to the other two and has even lower volatility and flammability. All three potential substitutes are cheap, with compliant suppliers and are soluble in water. Refold buffers were prepared, each containing 5mM cystamine, 25mM Tris pH8.0 and 20%v/v of one of the
substitutes. Refold reactions were initiated by diluting 100μl of unfolded IGF-1 solution with 1900μl of refold buffer and were allowed to proceed for 4 hours before analysis.

The substitute which gave the highest yield (propylene glycol) was then tested in the presence of the four additives which gave the highest yield, in order to test whether it had similar effects to alcohol in the presence of the additives. Four refold buffers were prepared, each containing 20% v/v propylene glycol, 5mM cystamine, 25mM Tris pH8.0. Each of the four buffers also contained one of: 0.25M arginine, 2mg/ml PEG3350, 0.25M GdnHCl, 0.25M NaCl.

Refold reactions were initiated by diluting 100μl of unfolded IGF-1 solution with 1900μl of refold buffer and were allowed to proceed for 4 hours before analysis.

5.2.6 Refold IGF-1 concentration experiment

Refold reactions were performed using five different dilution factors, hence giving five different concentrations of IGF-1 in the refolding protein solution. Two runs were done at each dilution factor, using two different refold buffers, making a total of ten runs.

The five dilutions used were 1+4, 1+7, 1+9, 1+11, 1+14, 1+19. Refold buffer A contained 5mM cystamine, 25mM Tris pH8.0. Refold buffer B contained 20% v/v propylene glycol, 0.25M arginine, 0.25M GdnHCl, 5mM cystamine, 25mM Tris pH8.0. Refold reactions were allowed to proceed for 4 hours before analysis.

5.2.7 IGF-1 assay

The concentration of correctly folded IGF-1 in the refolded protein solution was determined using a reverse-phase high performance liquid
chromatography (RP-HPLC) assay, which could distinguish between correctly folded IGF-1 and a misfolded form of IGF-1 which was also produced by the refold step.

An assay protocol was supplied by Avecia Biologics of Billingham, Cleveland. Details of the assay protocol are confidential. This protocol was further developed to improve throughput (see Section 3.8.4). Assays were performed using an RP-HPLC column received from Avecia and a BioCad chromatography system. Samples of refolded protein solution were mixed with sample buffer, filtered through a 0.2μm syringe filter and loaded onto the column. The concentration of the correctly folded form of IGF-1 in the sample was indicated by the amplitude of the peak in the absorbance trace at the same position as the peak from the standard. A calibration curve of IGF-1 standard diluted to different concentrations was plotted prior to assaying protein.

5.3 Results

The additive screening experiment used 2.8ml of unfolded IGF-1 solution (23.8mg of unfolded IGF-1). The ethanol substitute experiments used another 0.5ml of unfolded IGF-1 solution (4.25mg of unfolded IGF-1). The refold concentration experiment used another 2.5ml of unfolded IGF-1 solution (21.25mg of unfolded IGF-1). Including material that was wasted, a total of approximately 7ml of unfolded IGF-1 solution (60mg IGF-1) was used in these experiments.
Figure 5-1 Standard curve: graph of peak magnitude against standard concentration. Linear line of best fit, with equation and $R^2$ value, is shown.
Figure 5-2 Screening experiment results. Graph showing the concentration of correctly folded IGF-1 in solutions refolded using different buffers.
Figure 5-3  Ethanol substitute experiments results. Graph showing the concentration of correctly folded IGF-1 in solutions refolded using refold buffers containing different alcohols. Each bar represents the yield of correctly folded IGF-1 from one refold reaction.
Figure 5-4  Ethanol substitute experiment results. Graph showing the concentration of correctly folded IGF-1 in solutions refolded using refold buffers containing different alcohols and other additives.
Figure 5-5 Refold concentration results. Graph of correctly folded IGF-1 concentration in solutions refolded at different IGF-1 concentrations. Lines shown are second order polynomials fitted to data.

Figure 5-6 Refold concentration results. Graph showing yield of correctly refolded IGF-1 concentration in solutions refolded at different IGF-1 concentrations.
5.4 Discussion and conclusions

5.4.1 Overview of IGF-1 refolding

Insulin-like growth factor-1 (IGF-1) is a protein of 70 amino acid residues. In its native form it is a globular protein comprising three α-helices, a β-turn and an extended structure containing three disulphide bonds with a number of exposed apolar residues which are involved in its functional binding. In the native form the disulphide bonds are Cys\textsuperscript{6}-Cys\textsuperscript{48}, Cys\textsuperscript{18}-Cys\textsuperscript{61} and Cys\textsuperscript{47}-Cys\textsuperscript{52}. When IGF-1 undergoes oxidative refolding it can also form another stable structure (a misfold) which differs from the native structure in having a different tertiary structure, less helical structure and having the disulphide bonds Cys\textsuperscript{6}-Cys\textsuperscript{47}, Cys\textsuperscript{18}-Cys\textsuperscript{61} and Cys\textsuperscript{48}-Cys\textsuperscript{52}. It has been found that, in the presence of thiol reagents, these two isomers can interconvert and the ratio of the two is thermodynamically controlled (Miller et al, 1993). IGF-1 can also form disulphide-bonded multimers and this aggregation is enhanced by higher refolding protein concentrations. This type of aggregation upon refolding is observed with many proteins and involves the hydrophobic association of folding intermediates (Cleland and Wang, 1990).

Upon the initiation of oxidative refolding, IGF-1 rapidly forms an intermediate state containing the Cys\textsuperscript{18}-Cys\textsuperscript{61} disulphide bond (Hober et al, 1992, Miller et al, 1993). The IGF-1 can pass between a number of intermediate states, having different numbers and arrangements of disulphide bonds ("disulphide shuffling"), but all having the Cys\textsuperscript{18}-Cys\textsuperscript{61} bond (Rosenfeld et al, 1997). It is known that one of the intermediates on the pathway to native IGF-1 contains the Cys\textsuperscript{18}-Cys\textsuperscript{61} and Cys\textsuperscript{6}-Cys\textsuperscript{48} disulphide bonds, with the Cys\textsuperscript{47} and Cys\textsuperscript{52} unbonded. The formation of the final disulphide bond from this intermediate (to form native IGF-1) is slow; the Cys\textsuperscript{47}-Cys\textsuperscript{52} is a less favourable high-energy bond, anchoring an α-helix against the hydrophobic core. Milner et al (1998) proposed another intermediate on the pathway to native IGF-1, which has Cys\textsuperscript{18}-Cys\textsuperscript{61} and Cys\textsuperscript{47}-Cys\textsuperscript{52} bonds formed. The formation of the Cys\textsuperscript{6}-Cys\textsuperscript{48} bond from this
intermediate (to form the native IGF-1) is fast. These two intermediates on the path to native IGF-1 can readily interconvert, via the initial intermediate with the single Cys\textsuperscript{18}-Cys\textsuperscript{51} disulphide bond.

Hober et al (1992) postulated that disulphide bonding determined the pathway for IGF-1 refolding and hence the refold yield and consequently studied the refolding of IGF-1 under different redox conditions. Hart et al (1994) proposed that forces governing non-covalent folding (as opposed to disulphide bonding) of the protein govern the selectivity of IGF-1 refolding. The results of their experiments testing refolding in solutions containing different salts, chaotropes, osmolytes and alcohols supported this and this agrees with the widely accepted view that disulphide bond formation, though significant, does not determine the pathway for protein refolding. The results shown in this chapter also tend to support this view: refold reactions which were performed with different refold buffers, all containing the same redox reagents, gave very different refold yields.

In developing a refold step for IGF-1 it is desired to find conditions which favour the formation of the native state, reducing the formation of the misfolded state and aggregates, thereby maximising yield of native protein.

5.4.2 Standard curve
The data shows a very good fit to a straight line (Figure 5-1). This agrees with information received from Avecia about the assay method, indicating that the assay was working well. The concentration of correctly folded IGF-1 in the sample is directly proportional to the amplitude of the chromatogram peak in this range with a constant of proportionality of 5.96mg IGF-1 / mIAU.

5.4.3 Additive screening experiment
The same quantity of unfolded IGF-1 was used in each refold and each refold was of the same volume, so the concentration of correctly refolded
IGF-1 in each refolded protein solution was directly proportional to the yield of that refold.

Of those tested in the screening experiment, the refold buffers containing the additives arginine-HCl, GdnHCl or NaCl in combination with 20% ethanol gave the greatest refold yields (see Figure 5-2). These additives were therefore chosen for use in the subsequent ethanol-substitute experiment. Arginine and GdnHCl are known to increase refold yield of proteins by inhibiting aggregation of refolding protein (Tsumoto et al, 2002). For a protein folding in aqueous solution, the burial of a charged residue is unfavourable. If the solution contains salt (i.e. increased ionic strength), this effect is enhanced, which can increase folding rates thereby favouring correct folding over aggregation. The inclusion of an alcohol in the refold buffer enhances IGF-1 refold yield by inhibiting aggregation and promoting correct folding over misfolding (see section 5.4.4) but one effect of the alcohol is to reduce the polarity of the refold buffer which would be expected to make the burial of charged residues less unfavourable. The addition of arginine-HCl, GdnHCl or NaCl may mitigate this effect of the alcohol, which could explain the beneficial effect of using these additives together with an alcohol.

Hart et al (1994) published a study on the effect of many factors upon the refolding of IGF-1 by dilution, using a metal-catalysed air-oxidation system. They found that the addition of GdnHCl or urea to the refold buffer enhanced IGF-1 refold yield. They also found that the inclusion of 0.3M salt (either NaCl, MgCl₂, Na₂SO₄, or MgSO₄) in the refold solution enhanced the refold yield. With either chaotrope or salt, the inclusion of 20% ethanol enhanced the beneficial effect. The results reported in this chapter agree well with this, illustrating that refold experiments performed in microwells can give useful, reproducible results.

The extremely low yield of the refold into the ethanol/aspartate buffer is more difficult to explain and may be an experimental error, but given the
limited amount of IGF-1 available it was not possible to repeat this experiment.

5.4.4 Ethanol substitute experiments

The same quantity of unfolded IGF-1 was used in each refold and each refold was of the same volume, so the concentration of correctly refolded IGF-1 in each refolded protein solution was directly proportional to the yield of that refold.

Hart et al (1994) concluded that the results of their IGF-1 folding experiments supported the view that decreasing solution polarity (e.g. by the addition of alcohols) stabilises hydrophobic intermediates and thereby inhibit aggregation. Results shown in this chapter also agree with this view; IGF-1 refolds in the presence of 20% alcohol gave higher yields than those without alcohols. In native IGF-1 there are a number of apolar residues on the surface of the protein (Val\textsuperscript{11}, Phe\textsuperscript{23}, Tyr\textsuperscript{24}, Phe\textsuperscript{25}) which are involved in its functional binding (Cooke et al, 1991). These, along with other residues which may also be exposed only in partially folded intermediates, will contribute to the hydrophobic association which produces aggregates and so reduces refold yield. A reduction in solution polarity, by the presence of alcohol in the buffer, would be expected to make these residues more stable in solution, thereby inhibiting aggregation and so increasing refold yield.

In addition to this aggregation-inhibiting role, the inclusion of alcohol in the refold buffer may also enhance IGF-1 refold yield by improving the selectivity of the refold reaction. It has been found that the formation of $\alpha$-helix structure is favoured by lower solution polarity, for some proteins (Jackson and Mantsch, 1992). Correctly folded IGF contains more $\alpha$-helix structure than the misfolded form (Hober et al, 1992). Hart et al (1994) therefore speculated that reduced solution polarity enhances $\alpha$-helix formation during IGF-1 refolding, thereby increasing the yield of correctly-
folded IGF-1 relative to misfolded IGF-1. This is another possible mechanism for the refold-enhancing effect of 20% alcohol which is seen in the results reported in this chapter.

Hart et al (1994) found that 20% n-propanol had a similar beneficial effect to 20% ethanol on the refold yield of IGF-1. In contrast, in the ethanol substitute experiment, 20% ethanol was also found to have a beneficial effect but 20% n-propanol had little or no effect (see Figure 5-3). This difference in findings may be due to the fact that Hart and colleagues were using a different basal refold reaction (using refold buffers containing 2M urea, 20mM glycine, pH10.5 and air-oxidation catalysed by 0.5µM CuCl2). Glycerol (1,2,3-propanetriol) and propylene glycol (1,2-propanediol) both appeared to enhance IGF-1 refold yield, with propylene glycol giving a similar increase to ethanol. Glycerol has been used as a refold enhancing additive with other proteins. For example, Meng et al (2001) found that the inclusion of 10% glycerol in the refold buffer increased the activity yield of a creatine kinase dilution refold reaction from 55% to 69%. The addition of glycerol or propylene glycol to an aqueous refold buffer increases the viscosity of that buffer. Apart from the beneficial effects that glycerol and propylene glycol may have in common with ethanol upon refolding yields, this increase in viscosity may also inhibit aggregation by reducing the diffusivity of proteins in solution.

Of the three possible substitutes for ethanol in the refold buffer which were tested (propanol, glycerol, propylene glycol), propylene glycol gave the highest refold yield. Propylene glycol is very much less volatile and flammable than ethanol and so is an attractive substitute for ethanol in large-scale manufacture. Propylene glycol was therefore tested to determine what effects it would have in the presence of other additives (the additives tested here were selected based upon the results of the additive screening experiment – see Figure Figure 5-2 ).
With 0.25M arginine, 2mg/ml PEG3350 and 0.25M NaCl, the yield from the 20% propylene glycol containing refold buffers were close to the yields from the 20% ethanol containing refold buffers. With 0.25M GdnHCl, the 20% propylene glycol buffer gave a lower yield than the 20% ethanol buffer (Figure 5-4). It was concluded that the further investigation into the use of propylene glycol and these additives would be worthwhile.

5.4.5 Refold IGF-1 concentration experiment

The refolding yield appears to be insensitive to the concentration of IGF-1 in the refolding solution. Refold yield remained approximately constant over the range of 0.425mg/ml to 1.7mg/ml IGF-1. (Figure 5-5 and Figure 5-6). This result is surprising. It had been expected that higher concentrations of IGF-1 in the refolding solution would lead to more aggregation and so give much lower yields (see Section 1.3). In common with many proteins, the aggregation of IGF-1 during refolding follows approximately second-order kinetics, whereas refolding is first-order. Hart et al (1994) found that increasing IGF-1 concentration in the refold over the range 0.1mg/ml to 2mg/mL increases aggregate formation and so decreases yield of correctly folded IGF-1 from 50% to 25%. However, final urea concentration was held constant at 2M in their experiments. All refolds in the IGF-1 concentration experiment reported in this chapter were performed by dilution from an 8.5mg/ml unfolded IGF-1 solution in 7M urea into refold buffer. The different refold IGF-1 concentrations were achieved by using different dilution factors, so the refolds with higher IGF-1 concentrations also had higher urea concentrations (0.35M urea and 0.425mg/ml IGF in the 1+19 dilution refold, up to 1.4M urea and 1.7mg/ml IGF-1 in the 1+4 dilution refold).

The presence of urea in the refolding solution is known to inhibit aggregation during refolding and thereby enhance refold yield for many proteins, including IGF-1 (see section 2.3.4). In a separate experiment,
Hart and colleagues found that increasing the concentration of urea in the refold solution increased the refold yield of IGF-1 refolding at 1.7mg/mL protein from 16% yield at 0.25M urea to 24% yield at 4M urea. It therefore seems likely that the higher urea concentrations in the lower dilution refolds prevented aggregation, giving the approximately constant yield.

5.4.6 Experimental efficiency

The purpose of this research is to develop techniques to facilitate the rapid evaluation and optimisation of protein refold steps using small quantities of protein. The primary purpose of the experiments described in this chapter was to demonstrate the utility and efficiency of performing dilution refold development experiments at microplate scale.

The experiments reported in this section used only 60mg of unfolded IGF-1, but provided much useful information for the process design team. At a modest fermentation titre of 2g/ml, this amount of IGF-1 could have been produced from 30ml of broth (assuming efficient preparation of the solubilised protein). To perform similar experiments at normal bench scale, with refold reactions performed at the 100ml – 1L scale, would have required between 3g and 30g of IGF-1 which would require between 1.5L and 15L of fermentation broth.

The speed at which the experiments could be run was limited by the RP-HPLC assay. One column and chromatography machine were available, so only ≈4 samples could be analysed per hour. With the 4 hour refold time, and the time required for preparation, cleaning and shutdown, one laboratory worker working a normal eight hour day could perform approximately 20 refolds and the associated assays in one day (this number was actually achieved on one day during the experiment). The experiments described in this section involved 22 refolds and RP-HPLC assays, plus another 11 assays for the standard curve. A dedicated laboratory worker could have performed the experiments in 2 or 3 days. Access to automated
devices, which could perform experiments and assays day and night, would have allowed the experiments to be performed in less than 24 hours.

5.4.7 Further work

Had more IGF-1 and more time been available, further experimentation would have been possible. Replicates of each run could have been performed so as to increase confidence in the data. Further experiments could have been performed to determine what concentrations of propylene glycol and other additives gave optimum refold yields. Experiments could also have been carried out to investigate the effects of IGF-1 and urea concentration in the unfolded protein solution, the refold dilution factor and the redox reagents upon refold yield.

As discussed in Chapter 2, the final design of a refold step is chosen based on many factors, not just the refold yield and the batch volume (or protein concentration). Although adding additional chemicals to the refold buffer may increase the yield or allow the use of higher protein concentrations, the added chemicals can cause unacceptable problems. The cost and time required to obtain cGMP-compliant raw materials must be considered. The added chemicals can cause SHE problems, the volatility of ethanol being a good example. The additives may be incompatible with manufacturing facilities – although both GdnHCl and NaCl have been used in large scale manufacture, they can lead to corrosion of steel manufacturing equipment and so must be used with caution.

Further experiments performed by Avecia showed that the inclusion of either ethanol or propylene glycol in the refold buffer caused unacceptable loss of yield in a downstream IEC purification step. High concentrations of arginine and sugars also led to increases in filtration and chromatography pressure. A tight manufacturing schedule prevented further process development work being done on the refold step prior to the manufacturing campaign but this itself illustrates one of the potential advantages of a
microwell system (in terms of speed) to the generation of reliable process data and insight. It is proposed that automation of such experiments offers further improvements in time-efficiency. The automation of microwell refolding experiments, using a pipetting robot, is discussed in the next chapter.
6 Refolding reactions using a pipetting robot

To develop a commercial protein dilution refold step, a large amount of dilution refolding experimentation must be done. This development would be facilitated if the dilution refolding reactions could be performed at small scale and in parallel (see Section 1.4). It was shown in Chapters 4 and 5 that performing these experiments in microwell plates was both practicable and offered significant advantages over performing them at conventional bench-scale. It is further proposed that performing these experiments using laboratory automation, in the form of a pipetting robot, offers further advantages.

In this chapter, the rationale for using a pipetting robot is discussed, as are the practicalities of performing refold dilution reactions using a pipetting robot. The use of automation to assay refolded protein (refolding yield) and the coordination of refold reactions and assays are then discussed. An experiment is then described, to investigate the effects that the operating parameters of the pipetting robot have on a dilution refold reaction yield.

6.1 Rationale for using a pipetting robot

Performing a refolding experiment covering a large number of potential protein refold reactions requires carrying out a large number of liquid transfer operations. If the refold reactions are being conducted at very small scale, such as in microwells, then those liquid transfers will be done by pipetting. For each different condition being tested, slightly different pipetting operations will be required -- i.e. pipetting different volumes of liquid between different reservoirs and microwells. It would be difficult for a human operator to perform this large number of different pipetting operations. Due to their small size, it is difficult to label visibly microwells so it would be easy for the operator to make a mistake. Operator boredom, caused by the repetitive work, would increase the number of errors and may
cause other problems. A pipetting robot will not get bored and, if correctly programmed, will not make these mistakes.

The use of a pipetting robot has other advantages. The flowrate of aspiration and dispensing can be accurately controlled, as can the position of the pipette tip, which could allow better control of pipette mixing. The volume which a robotic pipettor aspirates and dispenses can be set and quickly altered by the controlling software, unlike a hand pipette where volume is set using an adjusting wheel and where each alteration may take several seconds. Robotic pipettors can have multiple pipettes, each with independently controlled aspirate/dispense volumes, which can make robotic pipetting much faster than hand pipetting when performing large numbers of different-volume liquid transfer operations. Laboratory automation can, in principle, run 24 hours per day. If an experiment is robustly-programmed into the controlling software of robotic pipettors, plate handlers, plate readers etc., then it can be left running unattended. An experimenter could prepare an experiment, leave the automation to perform the reactions and assays overnight, and return the next day to study the results.

Pipetting robots are expensive compared to hand pipettes, however they have many uses beyond protein refolding experiments. Robots can be gainfully used to perform experiments involving a number of bioprocess steps as well as their more conventional biochemistry uses (e.g. high throughput assaying).

### 6.2 Mimicking the refolding dilution step using a pipetting robot

In section 4.2, the practicalities of performing dilution refolding experiments in microwell plates was discussed. In this section, the practicalities of using a pipetting robot to perform those refolding experiments will be advanced.
A pipetting robot can perform a much more limited range of actions than a human laboratory worker. Some robots have additional abilities such as being able to detect contact between the pipette tip and a liquid, and robots can be synchronised with other pieces of computer controllable laboratory equipment (e.g. plate readers, shaker platforms, plate handlers). A basic pipetting robot can perform only three actions. It can move a pipette tip over a work area, it can aspirate fluid into that tip and it can dispense fluid from that tip. A pipetting robot is good for moving small volumes of liquid, and not much else. Therefore, if a dilution refolding experiment can be reduced to a set of liquid transfers, then it can be performed using a pipetting robot.

At the core of a step dilution refolding reaction is the mixing of a small volume of an unfolded protein solution with a larger volume of refold buffer. This can readily be performed in microwell plate format using a pipetting robot, as follows. A small volume of unfolded protein solution is aspirated from a container and dispensed into a microwell. A larger volume of refold buffer is then aspirated from a container and dispensed into the well containing the unfolded protein solution. The contents of the microwell are then immediately mixed by pipette mixing, thereby reducing the chaotrope concentration in the environment of the protein and so initiating refolding. If the pipetting robot has a number of pipettes which can work in parallel, then a number of refolding reactions can be initiated together, to refold in parallel.

A problem may be encountered with dispensing the unfolded protein solution into the microwell. The unfolded protein solution will be viscous, and only a small volume of it will be dispensed, so it may remain attached to the pipette as a drop on the tip, rather than falling into the microwell. This problem can be addressed by careful choice of pipette tip, dispensing speed, refold volume and by touching off the drop onto the base of the microwell. Use of a pipetting robot becomes more complicated when considering the operations which are necessary before the initiation of refolding.
The preparation of unfolded protein solution, as discussed in Section 4.2, would be difficult to perform in microwell plates using automation. Development of the refold step is likely to focus on the refolding conditions, rather than the unfolding conditions (see Section 1.3). Hence in process development there would not be a need to do a large number of different protein unfolding reactions and hence there would be less need to use automation to study the unfolding protein.

In the development of a commercial inclusion body protein refolding step it is likely that the variable which is most explored is the composition of the refolding buffer (see Section 1.3). It is therefore important that large numbers of refold buffers, of different compositions, be prepared for refolding experiments. As discussed in Section 4.2, a large number of refold buffers can be prepared from a smaller number of stock solutions by pipetting. The stock solutions themselves could not be prepared using the pipetting robot, as this would require the handling of solids, and so the smaller number of stock solutions would need to be prepared by hand.

A large dilution refolding experiment (comprising a large number of refold reactions) could be performed using a pipetting robot to perform the majority of operations. The human operator would need to prepare a small number of stock solutions and unfolded protein solutions, but the number of operations required to do this would be small compared to the number of operations required to prepare a large number of refold buffers and initiate a large number of refold reactions. Thus, a large array of different protein refolding reactions could be set up with only a small amount of human work.

6.3 Assaying using a pipetting robot

For a refold experiment to be useful, it is not only necessary to set up a large number of different refold reactions, it is also necessary to determine the yield of those refold reactions (to quantify the success or failure of those
Using a pipetting robot a large number of refold reactions may quickly be set up. However, data can only be produced as quickly as these refold reactions can be assayed. This section discusses how this assaying may be done with the use of a pipetting robot, and how the assay affects the speed with which refold experiments can be achieved.

When performing experiments to develop a protein refolding step, unless the stability of that refolded protein has been extensively studied, it will be necessary to determine the yield of refolding reactions (to assay for correctly folded protein) immediately after the protein has been allowed to refold for the prescribed refold time. It is likely that assaying for correctly folded protein concentration will take a significant length of time, and so it is likely that assaying will be the rate-limiting step in performing refold experiments. For example, it is pointless to initiate 96 refold reactions every hour if only four refolded protein solutions can be assayed every hour.

The different methods for determining the yield of protein refold reactions were discussed in Chapter 3. The most widely used of these methods fall into two categories: plate based assays and HPLC based assays.

6.3.1 Plate-based assays
A number of the assays discussed in Chapter 3 can be performed in microwell format using a pipetting robot. Turbidity assays (see Section 3.5.1) could be performed using any transparent microwell plate and an absorbance plate reader. A280 assays could be performed using a UV-transparent plastic or quartz microplate and an absorbance plate reader. Where filtration of samples is required, this could be done through a 96-well filter, although to do this would require either human intervention or a plate handling arm as the filter plates need to be lifted onto and off the vacuum filtration rig. Any liquid phase activity assays (see Section 3.4) can be performed at small scale in microwells, prepared by pipette and read using a plate reader.
Plate-based assays involve a lot of pipetting, but each sample from a refold experiment will typically be treated in the same way (unless there is reason to expect that some samples will give very high readings, in which case some samples may be pre-diluted). Therefore, to prepare a plate full of assays, each microwell in that plate will undergo the same pipetting operations, of the same volume, and in the same order. In this way, a plate full of assays could fairly easily be prepared by hand using a multi-channel pipette and so there is less incentive to use a pipetting robot. Use of a pipetting robot may still give some advantages. For example, assay preparation may be faster with the robot, or the robot could be programmed to perform the refolding reactions and the assays overnight.

If it is practicable to assay for correctly refolded protein (refold yield) using a microwell plate-based assay, then it is possible to assay 96 samples in parallel. In this case it will be practicable to perform at least 96 refold reactions in parallel, allowing refold experiments to be performed very quickly.

6.3.2 HPLC assays

HPLC assays for refolded protein (reverse-phase HPLC, ion-exchange HPLC etc. – see Section 3.7) are performed using a suitable HPLC column and HPLC system. The assays themselves cannot be performed by a pipetting robot, but a pipetting robot can prepare and dispense samples to the HPLC system.

One HPLC system, with a single column, can assay only one sample at a time. Each assay takes, typically, between 10 and 60 minutes to perform. Therefore if each refolded protein solution is be assayed immediately after it’s prescribed refold time, then the initiation of each refold reaction must be scheduled so that the refold is ready to be assayed when the HPLC system
is ready to accept another sample. This constrains the number of refold reactions which can be performed in parallel.

For example, if a refold experiment is performed in which each refold reaction is to have a refold time of 2 hours, and each reaction is be assayed using an HPLC assay which takes 15 minutes per sample, then each refold reaction must be initiated at least 15 minutes after the previous refold reaction (and so a maximum of 8 refold reactions can be proceeding at any one time).

The restriction that this type of assay can place on the rate at which refold experiments are performed negates one of the advantages of using a pipetting robot (the high rate at which refold reactions can be set up). However the use of a pipetting robot has another key advantage in this case – refold reactions can be initiated 24 hours a day, therefore the HPLC system can be supplied with samples 24 hours a day, maximising the number of refold reactions that can be performed. In the example above, a pipetting robot, operating 24 hours a day, could initiate, sample and assay up to 96 refold reactions per day (4 refolds per hour for 24 hours). A human laboratory worker, working an 8 hour day could initiate, sample and assay a maximum of 24 refold reactions per day.

It may be found, during refolding development experiments, that the variability of yield of refolds performed in microwells is significantly greater than the variability of results from an assay. In this case, if assaying a sample takes significantly longer than initiating a refold reaction, then a tactic of averaging-by-mixing may be used. That is, samples from a number of replicate refold reactions are mixed together, and this mixture is assayed for correctly refolded protein concentration. In this way, an average refolded protein concentration of the replicates is obtained using a single assay.
6.4 Experiment to investigate different ways of mixing dilution refolds using a pipetting robot

The way in which the unfolded protein solution is mixed into the refold buffer can have a significant effect on the yield of dilution refolding reactions (Section 4.3). The use of a pipetting robot to perform microwell protein refold experiments allows the experimenter more control over pipette mixing than can be achieved with hand pipetting. The flowrate at which liquid is aspirated and dispensed, the time between aspirate-dispense cycles and the position of the pipette tip in the microwell can all be accurately and repeatably controlled.

The purpose of the experiment described in this section is to investigate what effect the aspirate-dispense flowrate and number of cycles used for pipette-mixing the refold dilution have on refold yield.

6.4.1 Method

In the first part of the experiment, 11 different mixing conditions were defined, each having a different combination of aspirate-dispense flowrate and number of aspirate-dispense cycles. 8 refolds of hen egg-white lysozyme (HEWL) were carried out at each condition, giving a total of 88 runs (88 refolding reactions). In the second part of the experiment, 12 further different mixing conditions were defined, giving a total of 96 further runs.

The two parts of the experiment were carried out in separate 96-well plates. For each run, one refolding reaction was set up by the robot. After two hours of refolding at room temperature (this was 14°C during the first part and 18°C during the second part), samples were taken from each refolded solution by the robot and assayed to determine refold yield.
6.4.1.1 Chemicals

All chemicals, including protein, were purchased from Sigma-Aldrich Company Ltd. (Gillingham, Dorset, UK).

6.4.1.2 Preparation of unfolded lysozyme

Hen egg-white lysozyme (HEWL) was dissolved at 10mg/ml in a denaturing solution containing 6M guanidine hydrochloride (GdnHCl), 25mM dithiothreitol (DTT). The solution was vortex mixed for 1 minute and incubated for 2 hours at 20±3°C.

6.4.1.3 Automated refolding reactions

Refolding reactions were performed on a Perkin Elmer Multiprobe II four tip pipetting robot from Perkin Elmer of Boston, Massachusetts, USA. The robot was controlled using WinPrep software, also from Perkin Elmer. The 192 refold reactions were carried out in two Costar 96-well polypropylene 1ml assay blocks from Corning Inc. of Corning, New York, USA. These polypropylene plates were selected for their low cost and their availability. All liquid transfer operations to and between plates were effected by the pipetting robot, unless otherwise stated.

Each refolding reaction was initiated by the mixing of 40μl of denatured lysozyme solution with 760μl of refolding buffer (25mM cysteine, 1mM EDTA, 0.1M Tris-HCl, pH 8.2), in the following manner. The robot dispensed volumes of the denatured protein solution into the wells of the refolding reaction plate. The robot then aspirated volumes of the refolding buffers from a reservoir and dispensed them into wells in the refolding reaction plate. The robot then immediately mixed the contents of the wells in the refolding reaction plate, initiating the refolding reaction. The flowrate at which the robot aspirated and dispensed liquid, and the number of times that the robot aspirated and dispensed to mix are given in Table 6-1. For
comparison, eight refolds were diluted and mixed by hand using a Gilson pipette (condition number 12).

The robot flushed out the inside of the tips and rinsed the outside of the tips after each operation. Refolding reactions were allowed to proceed for 2 hours before being assayed.

**Table 6-1** *Table showing the aspirate-dispense flowrate and cycle times used to mix refolds in microwells.*

<table>
<thead>
<tr>
<th>Condition Number</th>
<th>Plate 1</th>
<th>Plate 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Flowrate of aspiration and dispensing for refold buffer addition and mixing (µl/s)</td>
<td>Flowrate of aspiration and dispensing for refold buffer addition and mixing (µl/s)</td>
</tr>
<tr>
<td></td>
<td>Number of aspirate-dispense cycles after addition of refold buffer</td>
<td>Number of aspirate-dispense cycles after addition of refold buffer</td>
</tr>
<tr>
<td>1</td>
<td>50</td>
<td>10</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>25</td>
</tr>
<tr>
<td>3</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>4</td>
<td>200</td>
<td>100</td>
</tr>
<tr>
<td>5</td>
<td>400</td>
<td>10</td>
</tr>
<tr>
<td>6</td>
<td>800</td>
<td>25</td>
</tr>
<tr>
<td>7</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>8</td>
<td>100</td>
<td>50</td>
</tr>
<tr>
<td>9</td>
<td>200</td>
<td>75</td>
</tr>
<tr>
<td>10</td>
<td>400</td>
<td>100</td>
</tr>
<tr>
<td>11</td>
<td>800</td>
<td>200</td>
</tr>
<tr>
<td>12</td>
<td>approx.400 (hand mixed)</td>
<td>800</td>
</tr>
<tr>
<td></td>
<td>5 (hand mixed)</td>
<td>2</td>
</tr>
</tbody>
</table>

- 179 -
6.4.1.4 Automated lysozyme activity assays

The robot pipette-mixed refolded HEWL solutions by 3 aspirate-dispense cycles of 600μl, immediately before taking samples for activity assays. Lysozyme activity assays were carried out according to the method of Section 7.1.1.7.
6.4.2 Results

Figure 6-1 Yields from step dilution refolds mixed in a microwell plate by a pipetting robot under different conditions (see Table 6-1 for details of conditions). Each point represents the yield of one refold reaction. (Plate 1).

Figure 6-2 Yields from step dilution refolds mixed in a microwell plate by a pipetting robot under different conditions (see Table 6-1 for details of conditions). Each point represents the yield of one refold reaction. (Plate 2).
6.4.3 Discussion and conclusions of the experiment

As was discussed in section 4.3.3, the yield of a dilution-refold reaction can be influenced by the mixing of the denatured protein solution with the refold. Results shown in section 4.3.2 agree with the findings of previous groups (Lee et al, 2002) that if this mixing is too slow, the yield of native, unaggregated protein can be adversely affected. The results shown in this section also agree with this – those reactions which were mixed using low aspirate/dispense flowrates showed reduced refold yields.

Mixing by pipette in a small vessel (e.g. the microwells used in this experiment) can give very short mixing times. Weiss et al (2002) found that mixing times of less than 5s were achieved. Using a pipetting robot, some control over this mixing can be exerted. Increasing the rate at which liquid is aspirated or dispense will increase the intensity (the rate of energy dissipation, the degree of turbulence) of the mixing. Increasing the number of aspirate/dispense cycles will extend the duration of this active mixing. Geometry of the well and the pipette tip and the physical properties of the liquids will also affect mixing.

The refolding yield achieved in reactions in this experiment appears to "plateau out" with increasing speed and duration of pipette mixing. Where refold reactions have been poorly mixed (i.e. low aspirate-dispense flowrate and number of cycles e.g. conditions 1, 2, 13, 14, 17, 18), the degree of mixing appears to have a strong effect on refold yield. Where refold reactions have been well mixed (i.e. at moderate or higher aspirate-dispense flowrate and number of cycles e.g. conditions 4 to 10 and 21 to 23), the degree of mixing appears to have little or no effect on refold yield. It may be that under the conditions examined in this experiment, there is a critical degree of mixing below which poor mixing has an adverse impact on yield, but above which mixing is sufficiently fast to avoid these effects and so extra mixing has little effect on yield.
It is worth noting that the lower aspirate/dispense flowrates used in this experiment (a volume of 760μL at a flowrate of 100μL/s or lower) would be very difficult to achieve by hand with a hand pipette. To do so would require that the operator aspirate in one smooth movement for more than 7 seconds, dispense in one smooth movement at the same rate and then repeat the aspirate/dispense cycle. The aspirate/dispense flowrates commonly used in hand-pipetting a volume of ≈800μL would normally be much higher than this.

Conditions with very high aspirate-dispense flowrates (conditions 11 and 24, flowrate=800μl/s) gave a very wide spread of refold yields. Liquid was seen to spray from these microwells during dispensing and mixing. This spraying would have ejected liquid from the microwell before it was fully mixed and so may account for the spread of refold yields.

Similar refold conditions appear to give different refold yields on plate 1 and plate 2. The highest refold yield obtained on plate 2 (≈65%) is higher than that on plate 1 (≈53%). This difference in calculated yield values between the two plates illustrates the significance of the temperature dependence of the HEWL assay. The two sets of experiments using the two plates were performed on two separate days, when the temperature in the lab may have been different. The lysozyme activity assay is known to be temperature sensitive. The assay was also performed at room (ambient) temperature, so the different values of measured activity are likely to be due to the temperature differences in the lab: 14°C during the first part and 18°C during the second part.

It was concluded that, with this system, pipette mixing using an aspirate/dispense flowrate of 200μL/s or faster and at least 2 aspirate/dispense cycles (after the initial dispensing of refold buffer into the well containing the denatured protein solution) is sufficient to prevent poor mixing affecting refold yield and that an aspirate/dispense flowrate of
800μL/s or greater can lead to spraying. Further dilution refold experiments performed in similar microwell plates using this pipetting robot will therefore be pipette-mixed using these parameters.

6.5 Discussion and conclusions of the chapter

A survey of the available literature revealed only one publication reporting refold experiments performed using a laboratory robot. Scheich et al (2004) describe an automated refolding screen which they evaluated using two model enzymes, carbonic anhydrase II and malate dehydrogenase, then used to find refolding conditions of a third protein, the p22 subunit of human dynactin. The refold screen which they describe tests a set of 30 refold conditions – 22 simple dilution refold conditions and a further 8 IMAC-resin bound conditions for His-tagged proteins (see section 1.2.5.2). Their refolds were performed in 1.5mL-2mL volume tubes using a pipetting robot and ancillary equipment. In their refold experiments, mixing was achieved by placing the tubes on a shaker platform. Their folding screen comprised a relatively small set of fixed refolding conditions (in a similar way to off-the-shelf refolding kits - see section 2.3.4) and was designed for the preparation of proteins for structural and functional analysis. It allowed the experimenter to very quickly find conditions under which a small quantity of folded protein could be produced. As such, it was suitable for applications where a very high throughput of proteins are required to be produced in native form in small quantities e.g. for structural and functional characterisation of newly identified proteins. This folding screen had similar drawbacks to off-the-shelf refolding kits. The small range of refold conditions which were tested for each protein, and the fixed set of conditions, made it unlikely that the refold reaction produced would be suitable for larger-scale production. However, their work did demonstrate the potential of laboratory automation to expedite refolding experiments. The effects of mixing in the automated refold reactions were not considered in their publication.
As discussed in sections 6.1 and 6.2, the use of a pipetting robot to perform dilution refold experiments is both attractive and feasible. As discussed in Section 6.3, the pipetting robot can also be used to facilitate and accelerate the assaying required to determine the yield of these experimental refold reactions and the reactions and assays can be synchronised to maximise throughput. The experiments described in Section 6.4 show that a pipetting robot can mix unfolded protein solution and refold buffer in a microwell to a sufficiently high degree that poor mixing does not adversely affect refold yields.

The use of laboratory automation, such as the pipetting robot, greatly facilitates the performance of refold reactions in microwell plates. The use of such automation, together with microwell scale refold reactions facilitates the rapid performance of refold step development experiments using a very small quantity of protein, which satisfies two of the aims of the project (see Section 1.5). In Chapter 7, experiments will be described which demonstrate the use of these techniques to generate data for refold step development.
7 Application of automated microwell experiments for the rapid development of protein refold steps

In Chapter 5, the use of microwell refolding reactions for the development of a protein refold step was demonstrated. This illustrated the benefits of microwell refolding techniques. In Chapter 6 the use of laboratory automation equipment to further improve the efficiency of microwell refolding experiments was explored.

In this chapter, experiments are described in which microwell refolding experiments are performed using laboratory automation to develop refold steps for two proteins: lysozyme and trypsinogen. These experiments were performed in order to demonstrate the usefulness of automated microwell experiments for the rapid development of protein refold steps.

7.1 Optimising lysozyme refolding using automated experiments in microwells

During the early process development stage of the production of a therapeutic protein there is very little protein material available for experimentation, and speed of development is critical. Design information about the refolding step must be gathered as quickly as possible and using as small a quantity of inclusion body protein as feasible.

The experiments described in this section were carried out to test the hypothesis that significant improvements in the yield of an inclusion body protein step-dilution refold step can be achieved rapidly, and using very little material, by the use of automated refolding experiments in microwells. Thus the primary purpose of this work is not to improve the refolding yield of
the protein, rather it is to study how efficiently this improvement can be made using the techniques described in Chapters 4 and 6 of this thesis.

In these experiments the refolding step being optimised is the step dilution refolding of solubilised, unfolded hen egg white lysozyme (HEWL). In commercial processes of interest the protein being refolded would be solubilised, unfolded protein from washed inclusion bodies. Solubilised, unfolded HEWL is used here because it is cheap and has a simple activity assay.

The starting point for these optimisation experiments is a simple step dilution refold. A volume of 10 mg/ml solution of denatured protein in 6M GdnHCl, 25mM DTT is mixed with 19 volumes of a refolding buffer of 25 mM cysteine, 1mM EDTA, 0.1 M Tris-HCl, pH 8.2. This refold step is quite typical of the type of refold step in a therapeutic protein production process as it goes into early process development. Improvements in the refolding yield achieved by varying the composition of the refolding buffer are sought (i.e. the objective function for optimisation is yield and the independent variables investigated are the concentration of additives in the refold buffer).

7.1.1 Materials and Methods

7.1.1.1 Materials
All chemicals, including protein, were purchased from Sigma-Aldrich Company Ltd (Gillingham, Dorset, UK).

7.1.1.2 Preparation of unfolded lysozyme
Hen egg-white lysozyme (HEWL) was dissolved at 10 mg/ml in a denaturing solution containing 6 M guanidine hydrochloride (GdnHCl), 25 mM dithiothreitol (DTT). The solution was vortex mixed for 1 minute and incubated for 2 hours at 25 °C.
7.1.1.3 Refolding buffers and stock solutions

Refolding buffers were prepared by mixing together volumes of refolding stock solutions. Twelve refolding stock solutions were prepared. Each stock solution contained 25 mM cystine, 1 mM EDTA, 0.1 M Tris-HCl, pH 8.2. Eleven of the refolding stock solutions contained high concentrations of chemicals which have been reported to enhance the refolding yield of various proteins (see Section 2.3.4). One refolding stock solution contained no additive. Thus, by mixing together these stock solutions, refolding buffers could be prepared which could contain various quantities of the reported refolding enhancing additives, but which would always contain 25 mM cysteine, 1 mM EDTA, 0.1 M Tris-HCl, pH 8.2. The refolding enhancer concentrations are given in Table 7-1.
### Table 7-1 Table of reported refolding enhancers used

<table>
<thead>
<tr>
<th>Refolding enhancing additive</th>
<th>Conc. in stock solution (M)</th>
<th>Conc. in stock solution (g/l)</th>
<th>&quot;High&quot; conc. in refolding buffer (M)</th>
<th>Volume fraction of stock in refolding buffer at &quot;High&quot; level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 L-arginine</td>
<td>0.574</td>
<td>100</td>
<td>0.25</td>
<td>0.436</td>
</tr>
<tr>
<td>2 ammonium sulphate</td>
<td>1.136</td>
<td>150</td>
<td>0.50</td>
<td>0.440</td>
</tr>
<tr>
<td>3 sucrose</td>
<td>0.876</td>
<td>300</td>
<td>0.25</td>
<td>0.285</td>
</tr>
<tr>
<td>4 glucose</td>
<td>0.832</td>
<td>150</td>
<td>0.25</td>
<td>0.300</td>
</tr>
<tr>
<td>5 glycerol</td>
<td>5.28</td>
<td>500</td>
<td>1.32</td>
<td>0.250</td>
</tr>
<tr>
<td>6 PEG 300</td>
<td>0.33</td>
<td>100</td>
<td>0.10</td>
<td>0.300</td>
</tr>
<tr>
<td>7 PEG 3350</td>
<td>0.006</td>
<td>20</td>
<td>0.0020</td>
<td>0.335</td>
</tr>
<tr>
<td>8 ethanol</td>
<td>21.7</td>
<td>100</td>
<td>1.0</td>
<td>0.046</td>
</tr>
<tr>
<td>9 2-propanol</td>
<td>0.167</td>
<td>10</td>
<td>0.010</td>
<td>0.060</td>
</tr>
<tr>
<td>10 NaCl</td>
<td>0.40</td>
<td>23.4</td>
<td>0.10</td>
<td>0.250</td>
</tr>
<tr>
<td>11 cyclohexanol</td>
<td>0.101</td>
<td>10</td>
<td>0.010</td>
<td>0.099</td>
</tr>
<tr>
<td>12 none</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
</tr>
</tbody>
</table>

#### 7.1.1.4 Experimental design

Experiment designs were produced and experimental data were processed using Minitab v.13.30 statistical software from Minitab Inc. of State College, Pennsylvania, USA and Excel spreadsheet software from Microsoft Corp. of Seattle, Washington, USA.

The purpose of these experiments is to find a buffer composition containing not more than two refolding enhancers, which gives an optimum refolding yield (not more than two refolding enhancers are sought because it is desirable to keep the refold buffer as simple as possible – see sections
2.3.4 and 2.4). The first, screening, experiment allows the selection of the two most promising refolding enhancers from the eleven tested. The second, response surface, experiment finds the buffer composition using those two refolding enhancers which gives an optimum refolding yield.

A screening experiment was designed to determine the effect of each of the eleven reported refolding enhancing additives and to determine what interactions there were between the effects of any two additives. Each condition to be tested in the experiment comprised each of the additives at either a “high” or “low” concentration. The “high” concentration for each additive was selected, based upon reports in the literature of that (or a similar) additive’s refold-enhancing effect (see Section 2.3.4), not to be the concentration which had given optimal yield for a protein, but to be at a lower concentration at which beneficial effect had been seen. Zero was selected as the “low” concentration.

The screening experiment design comprised 67 conditions to be tested. The conditions were tested in triplicate, giving 201 refolding reactions to perform and assay for the screening experiment. The order of the reactions was randomised to prevent extraneous factors affecting results.

The two most promising refolding enhancing additives were selected and a response surface experiment was designed to determine the optimum concentrations of each of those two additives in the refolding buffer.

A mixture experiment design was selected so that the refolding buffer compositions to be tested could be produced by mixing various volumes of the stock solutions. In a mixture experiment, the factors being studied are proportions of different components of a blend. In this case the refolded buffers to be tested were produced by blending three stock solutions in various proportions – two stock solutions containing near-saturated solutions of the two most promising refolding enhancing additives and one stock solution, acting as a diluent, containing no additive.
The response surface mixture experiment design was a simplex lattice design of degree 6, augmented with centre point and axial points, with 3 replicates (triplicate) for each design point (condition). This gave 31 conditions to be tested, 93 refolding reactions to be run for the response surface experiment. Again, the order of the reactions was randomised to prevent extraneous factors affecting results.

7.1.1.5 Automated refolding reactions

Refolding reactions were performed on a Perkin Elmer Multiprobe II four tip pipetting robot from Perkin Elmer of Boston, Massachusetts, USA. The robot was controlled using WinPrep software, also from Perkin Elmer. The reactions were carried out in Costar 96-well polypropylene 1ml assay blocks from Corning Inc. of Corning, New York, USA. Polypropylene plates were selected for their low cost and their availability.

The robot was programmed to perform the experiment one plate (i.e. up to 96 reactions) at a time. Refolding buffers were prepared in one 96-well plate, refolding reactions were carried out in a second 96-well plate. All the plates required for each experiment were loaded onto the work surface of the robot at the beginning of the experiment. All liquid transfer operations to and between plates were effected by the pipetting robot.

At the commencement of a trial, or after each change of plate, the robot aspirated volumes of refolding stock solutions from their reservoirs, dispensed them into the wells of the buffer preparation plate and then mixed the resulting buffers by repeated aspirate-dispensing. The volumes of each of the refolding stock solutions dispensed into each of the wells in the buffer preparation plate was controlled by a data table which was produced from the experiment design and was linked to the controlling software. A total volume of 850μl of refolding buffer was prepared for each refolding reaction.
Each refolding reaction was initiated by the mixing of 40μl of denatured lysozyme solution with 760μl of refolding buffer, in the following manner. Once all of the buffers were prepared in the buffer preparation plate, the robot dispensed volumes of the denatured protein solution into the wells of the refolding reaction plate. The robot then aspirated volumes of the refolding buffers from wells in the buffer preparation plate and dispensed them into wells in the refolding reaction plate. The robot then immediately mixed the contents of the wells in the refolding reaction plate, using the method described in Section 6.4, initiating the refolding reaction.

The robot flushed out the inside of the tips and rinsed the outside of the tips after each operation. Refolding reactions were allowed to proceed for 2 hours before being assayed.

7.1.1.6 Bench scale refolding reactions
Bench scale refolding reactions (performed to confirm the yield of the optimised refolding reaction) were carried out in a cylindrical glass beaker of diameter 63mm, with an electrically-driven, 6 bladed, steel Rushton turbine of diameter 24mm, height 6mm, centrally located 21mm above the base of the beaker. 190ml of refolding buffer was poured into the beaker, and was stirred at 400rpm. 10ml of the denatured HEWL solution was then poured into the refolding buffer, initiating the refolding reaction. The refolding reaction was allowed to proceed for 2 hours before being assayed.

7.1.1.7 Automated lysozyme activity assays
Assays of refolded lysozyme solutions from the microwell and bench scale refolding reactions were carried out using the Perkin Elmer Multiprobe II four tip pipetting robot, and the Spectracount absorbance plate reader from Perkin Elmer of Boston, Massachusetts, USA. The robot was controlled using WinPrep software and the plate reader was controlled using the Spectracount software, both from Perkin Elmer. Lysozyme activity values
were calculated from absorbance data using Excel spreadsheet software from Microsoft Corp. of Seattle, Washington, USA.

Sample dilutions were carried out using Costar 96-well polypropylene 1ml assay blocks. Absorbance measurements were made in Corning 96-well polystyrene 350μl plates. Both were supplied from Corning Inc. of Corning, New York, USA.

The activity assay used was based on that of Shugar (1952). The basis of the assay is the rate of lysis of the lysozyme sensitive bacteria Micrococcus lysodeikticus, as measured by the decrease in absorbance of a suspension of that bacteria.

An assay solution of 300mg/l Micrococcus lysodeikticus in 0.1M potassium phosphate, pH 7.0 was used. Assays were carried out at ambient temperature (approximately 20°C). All pipetting operations were carried out using the pipetting robot.

In practice 20μl of each refolded protein solution was mixed with 780μl of 0.1M potassium phosphate, pH 7.0. 20μl of this diluted solution was then mixed with 780μl of assay solution. 200μl of this assay mix was dispensed into each well of the 96-well polystyrene 350μl plates. The assay plate was then loaded into the plate reader, and the absorbance of each well was measured every minute for ten minutes at a wavelength of 405nm. The rate of decrease of absorbance of the sample was taken as being indicative of the lysozyme activity.

It is in the nature of pipetting operations that, occasionally, a drop of liquid will stray. For example, a drop may remain on the pipette tip at the end of a dispensing step, then drop into a different well as the pipette travels over the plate. Or, a droplet may be sprayed out of a well during an aspirate-dispense mixing operation, and land in a different well. The size of these drops will be small, but can be significant when compared to the very small
size of the microwell. These dropping events can occur during robotic pipetting in the same way as during hand pipetting, and can lead to one reaction (one microwell) having an apparent yield which is very much higher or very much lower than the other replicates of that reaction. These outliers are small in number but they can have a strong effect on mean values because they lie out so far. This can be seen on the standard curve (Figure 7-1). For this reason, the median value of a set of replicates was taken to be representative of the average of that set, rather than the mean.

Figure 7-1 Standard curve for automated lysozyme assay (6 replicates at each concentration). Mean and median activity values for each concentration are plotted. Linear lines of best fit to both the mean and median values are shown, with equations and $R^2$ values.
7.1.2 Results

The first part of this section describes the main effects of the reported refolding enhancers, as determined from the screening experiment. The second part describes the two-factor interactions of the reported refolding enhancers. Based on the results of the screening experiment, two refolding enhancers were selected for the surface response experiment, whose results are described in the third part of this section. The optimum refolding condition was determined from the results of this surface response experiment.

7.1.2.1 Main effects of reported refolding enhancers on the refolding yield of lysozyme

Figure 7-2 shows eleven main effect plots, which together show the main effects of the presence or absence of each of the reported refolding enhancers on the enzyme activity of the refolded lysozyme solution. Main effect plots are used to display the influence of each of the factors on the response. In this case the factors are the presence or absence of each of the eleven reported refolding enhancing chemicals in the refolding buffer, and the response is the enzyme activity of the refolded protein solution.

The "0" point on each plot indicates the average enzyme activity with that reported refolding enhancer absent from the refolding buffer (i.e. at the "Low" level), and the "1" point indicates the average enzyme activity with it present (i.e. at the "High" level). The experiment design was symmetric, so that for each factor, the only difference between the group of reactions with that factor present and the group of reactions with that factor absent, is the presence or absence of that factor. The difference between these two average activities indicates the effect of that reported refolding enhancer on the refolding yield of lysozyme. Figure 5.2 shows that arginine, glucose and sucrose have the most positive effect on the refolding yield of lysozyme.
Figure 7-2 Main effects of reported refolding enhancers on the enzyme activity of the refolded lysozyme solution.
7.1.2.2 Two-factor interactions of reported refolding enhancers on the refolding yield of lysozyme

Interaction plots are used to determine if there is any interaction between the effects of two factors on a given response, i.e. they indicate whether one factor enhances or diminishes the effect of another factor. By comparing the interaction plots of different pairs of factors, it is possible to compare the strengths of the interactions between those pairs of factors, and so to determine which pair of factors have the greatest effect when operating in concert.

Each interaction plot displays the level of the response for each combination of levels of two factors. In this case the factors are the presence or absence of each of the eleven reported refolding enhancing chemicals in the refolding buffer, and the response is the enzyme activity of the refolded protein solution.

Figure 7-3 shows fifty-five interaction plots, which display the fifty-five interactions between the eleven factors. This plot is complex and can be difficult to read so, for clarity, Figure 7-4 shows the interaction plots of just the three factors - arginine, sucrose and glucose - which had been determined from the screening experiment (see section 7.1.2.1) to yield the most positive (and therefore interesting) effects on the refolding yield of lysozyme.

For ease of interpretation, Figure 7-4 displays the data such that the two points on the left side of the plot are at the lower level of the factor labelled below the plot. The two points on the right side of the plot are at the higher level of the factor labelled below the plot. The two points plotted as black circles and connected by an unbroken line are at the lower level of the factor labelled to the left of the plot. The two points plotted as red squares and connected by a broken line are at the higher level of the factor labelled to the left of the plot.
An interaction is present if the change in the response from the low to the high level of one factor depends on the level of a second factor. This is easily seen in an interaction plot as a convergence or divergence of the two lines. The greater the departure from parallel, the greater the interaction.

The interaction plot for glucose and sucrose shows a slight convergence of the lines, indicating a negative interaction at the levels studied i.e. the effect of the presence of one is diminished by the presence of the other. The interaction plot for arginine and sucrose shows a slight divergence of the lines, indicating a positive interaction, or “synergy”, at the levels studied i.e. the effect of the presence of one is increased by the presence of the other. This suggested that further experimentation using arginine and sucrose would be more fruitful than further experimentation using sucrose and glucose or using glucose and arginine. Therefore the effect of arginine and sucrose concentration on the refold yield was examined in the surface response experiments. (Experimental design described in section 7.1.1.4, experimental results described in section 7.1.2.3.)
Figure 7-3 Two-factor interactions of reported refolding enhancers on the refolding yield of lysozyme
Figure 7-4  Two-factor interactions of arginine, glucose and sucrose on the refolding yield of lysozyme
7.1.2.3 Response surface optimisation

A response surface optimisation experiment was carried out to determine the optimum concentrations of arginine and sucrose in the refolding buffer. The experiment design, materials and methods were as described in section 7.1.1. Figure 7-5 and Figure 7-6 show the results of the surface optimisation experiment. They show how the activity of the refolded lysozyme solution depends on the concentration of refolding enhancers in the refolding buffer. For ease of reading, this response surface is displayed both as a contour plot (Figure 7-5) and as a pseudo-3D wireframe plot (Figure 7-6).

The refolding buffers tested were prepared, as described in section 7.1.1.3, by mixing various proportions of the arginine, sucrose and additive-free stock solutions, accordingly the response surface is shown over a mixture plot. In the mixture plots in Figure 7-5 and Figure 7-6, the composition of the refolding buffer is described in terms of the proportions of the three stock solutions of which it was composed. For example, a point at the "arginine" vertex refers to a refolding buffer containing only the arginine stock solution, and a point in the centre of the plot refers to a refolding buffer containing equal proportions of the three stock solutions.

An optimum refolding buffer condition is observed where the refolding buffer consists entirely of the arginine stock solution. A bench scale refolding reaction was performed using the arginine stock solution as the refolding buffer. This gave a refolded lysozyme solution activity of 0.00605AU/min, which agrees well with the microwell-scale average result of 0.00616AU/min.
Figure 7-5 Contour plot showing the response of refolded lysozyme activity to different refolding buffer compositions. Black circles indicate refolding buffer compositions which were tested in the experiment. The scales on the three sides indicate the proportions of the three stock solutions which comprised the refold buffers used for each condition.

Figure 7-6 Wireframe surface plot showing the response of refolded lysozyme activity to different refolding buffer compositions. The scales on the three sides of the base indicate the proportions of the three stock solutions which comprised the refold buffers used for each condition.
7.1.3 Discussion and conclusions of the experiment

7.1.3.1 Effects of refolding enhancers

Of the eleven reported refolding enhancing chemicals tested, three chemicals (arginine, glucose and sucrose) were seen to have a substantial positive effect on the refolding efficiency of lysozyme. This supports reports that these chemicals can enhance the refold yield of some proteins. Arakawa and Tsumoto (2003) concluded that arginine improves the refolding efficiency of proteins, not by facilitating the refolding itself, but by inhibiting the aggregation of proteins during refolding and thereby allowing more of the protein molecules to remain free in solution and to refold.

Ammonium sulphate was seen to have a large negative effect on the refolding efficiency of lysozyme. The wells in the microwell plates which contained lysozyme refolded in the presence of ammonium sulphate were inspected and were seen to contain large quantities of precipitates, suggesting that most or all of the lysozyme had aggregated and precipitated out of solution. This may have been due to the relatively high ionic strength of the ammonium sulphate containing refolding buffer. The same driving forces which lead to a protein precipitating during salting-out may have driven this precipitation. This suggests that if such a refolding enhancer screening experiment were carried out in the future, a lower concentration of ammonium sulphate should be tested.

Glucose and sucrose showed a slight negative interaction at the levels studied. That is to say that the effect of the presence of one was diminished by the presence of the other in the refolding buffer. This result is probably a reflection of the fact that the compounds are chemically very similar and so may work in similar ways, each making the other somewhat redundant.
7.1.3.2 Experimental efficiency

A key purpose of this research project is to develop techniques to facilitate the rapid evaluation and optimisation of protein refold steps using small quantities of protein. It is desirable to have generic techniques which can be applied to different proteins. Hence, the purpose of the results reported in this chapter was not to improve lysozyme refolding yield per se. Lysozyme is merely used as a convenient model protein. With appropriate adaptation (e.g. for different assays), these techniques could be applied to the process development of any protein dilution refold step.

The potential value of the techniques used in this set of experiments lie in (1) the short timescale in which the experiments can be performed, (2) the small amount of operator-time required and (3) the small quantity of protein used. These features are considered in more detail in this section.

Figure 7-7 is a timeline showing how long it would take to perform refolding experiments similar to those reported in this section for a different protein. It assumes that a refold time is fixed at 2 hours and that an automated parallel assay is used (see section 8.3). Although 2 hours refold time would be sufficient for many proteins, some proteins would require longer.

The lysozyme refolding optimisation experiments described in this section took considerably longer to run, due to the time required to develop methods and robot programs. Hence, the timeline shown in Figure 7-7 does not show how long it took to prepare and perform the experiments in section 7.1, it only shows how long it would take to do similar experiments in the future.

The automated microwell experiments described in this section used only 130 mg of protein, and achieved an increase in refolding yield of 195%. This contrasts sharply with the single refolding reaction carried out in the conventional way at bench scale (to confirm the result of the microwell-scale...
reaction) which used 100 mg of protein but provided only one data point, and hence would have been of very limited use for process optimisation.

Figure 7-7  Times required to perform a similar set of refolding optimisation experiments.
7.2 Screening refolding enhancers to improve the refolding efficiency of trypsinogen using automated experiments in microwells

The lysozyme experiments described in section 7.1 demonstrated that automated refolding experiments in microwells can be used to rapidly generate data for refold step development using a very small amount of protein. This chapter describes experiments which apply these techniques to the refolding of trypsinogen, a commercially interesting protein which gives very low refold yields.

Experiments were carried out, in microwell plates, using a pipetting robot, to determine if the step dilution refolding yield of trypsinogen could be increased by the inclusion of additives in the refolding buffer, and to select which additives would have the greatest effect on the refolding yield.

These experiments had two aims. First, to provide information about the refolding of trypsinogen in order to aid process development. Second, to demonstrate that these automated microwell techniques can be applied to a commercially interesting protein (and not just lysozyme) so as to rapidly provide useful information using a very small amount of protein.

The experiments described in this section were performed with Edward Hibbert, a PhD student who was doing research at UCL Department of Biochemical Engineering in collaboration with Eli Lilly Company. Eli Lilly had a protocol for the production of trypsin. Trypsinogen inclusion bodies were produced in an *E. coli* fermentation. The inclusion bodies were separated and washed. The trypsinogen was solubilised in a denaturing buffer, then refolded by dilution into a refolding buffer. Trypsin was then produced by enzymatic cleavage of the pro-sequence from the trypsinogen.
For reasons of commercial confidentiality Eli Lilly did not release the details of their denaturation and refolding steps. However, a paper co-written by a chief engineer at Eli Lilly (Buswell, Ebtinger, Middelberg) describes experiments on the denaturation and refolding of trypsinogen. Washed trypsinogen inclusion bodies were solubilised in a denaturing buffer of 5.5 M urea, 100 mM cysteine, 10 mM EDTA, 5mM Tris, pH 9.5 at 8°C for 1 hour. Refolding was initiated by injection of the denatured trypsinogen solution into a stirred vessel containing a refolding buffer of 50mM CaCl₂, 5mM Tris, 3mM cysteine, 1mM cystine, pH 9.

The experiments described in the paper tested the effect on the refolding yield of different impeller speeds and of different concentrations of trypsinogen in the denaturing buffer. Two different impeller speeds were tested, at two denaturing buffer trypsinogen concentrations. The dilution ratio was varied so that, for each denaturing buffer trypsinogen concentration, the final concentration of total trypsinogen (native, misfolded and aggregated) in the refolding buffer was 60 μg/ml.

Impeller speed was found to have little effect, but a higher concentration of trypsinogen in the denaturing buffer was found to give a significant increase in refolding yield. A concentration of 1.4 mg/ml trypsinogen in the denaturing buffer gave refolding yields from 11.0% to 11.6%, while a concentration of 2.8 mg/ml trypsinogen in the denaturing buffer gave refolding yields from 13.6% to 14.5%.

7.2.1 Materials and Methods
Enterokinase was provided by Eli Lilly. All other chemicals were purchased from Sigma-Aldrich Company Limited of Gillingham, Dorset.

7.2.1.1 Preparation of denatured trypsinogen
The denatured trypsinogen was prepared according to a method used in previous work by Edward Hibbert. Purified native trypsinogen was
dissolved at 1.4 mg/ml in a denaturing buffer of 5.5M urea, 10mM EDTA, 10mM DTT, pH 9.5. The solution was vortex mixed for 1 minute and allowed to denature for 2 hours at 17 °C.

### 7.2.1.2 Refolding buffers and stock solutions

Refolding buffers were prepared in a similar way to section 5.5.1.3, by the mixing of combinations of 12 refolding stock solutions. All buffers contained 50mM CaCl₂, 2.5mM oxidised glutathione (GSSG), 0.1M Tris, pH 8.2 (this was based on a refold buffer previously used by Edward Hibbert). Eleven of the twelve stock solutions contained reported refolding enhancers, the twelfth was used as a diluent.

<table>
<thead>
<tr>
<th>Refolding enhancer</th>
<th>Conc. i stock solution (M)</th>
<th>Conc. in stock solution (g/l)</th>
<th>&quot;High&quot; molarity in refolding buffer (M)</th>
<th>Volume fraction of stock in refolding buffer at &quot;High&quot; level</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 L-arginine</td>
<td>0.574</td>
<td>100</td>
<td>0.250</td>
<td>0.436</td>
</tr>
<tr>
<td>2 ammonium sulphate</td>
<td>1.136</td>
<td>150</td>
<td>0.500</td>
<td>0.440</td>
</tr>
<tr>
<td>3 sucrose</td>
<td>0.876</td>
<td>300</td>
<td>0.250</td>
<td>0.285</td>
</tr>
<tr>
<td>4 glucose</td>
<td>0.832</td>
<td>150</td>
<td>0.250</td>
<td>0.300</td>
</tr>
<tr>
<td>5 glycerol</td>
<td>0.217</td>
<td>20</td>
<td>0.050</td>
<td>0.250</td>
</tr>
<tr>
<td>6 PEG 300</td>
<td>0.067</td>
<td>20</td>
<td>0.020</td>
<td>0.300</td>
</tr>
<tr>
<td>7 PEG 3350</td>
<td>0.006</td>
<td>20</td>
<td>0.002</td>
<td>0.335</td>
</tr>
<tr>
<td>8 ethanol</td>
<td>2.174</td>
<td>100</td>
<td>0.434</td>
<td>0.200</td>
</tr>
<tr>
<td>9 L-BAPNA (substrate)</td>
<td>0.001</td>
<td></td>
<td>0.0001</td>
<td>0.100</td>
</tr>
<tr>
<td>10 2-pentanol</td>
<td>0.040</td>
<td>23.4</td>
<td>0.010</td>
<td>0.250</td>
</tr>
<tr>
<td>11 cyclohexanol</td>
<td>0.101</td>
<td>10</td>
<td>0.010</td>
<td>0.099</td>
</tr>
<tr>
<td>12 none</td>
<td></td>
<td></td>
<td></td>
<td>n/a</td>
</tr>
</tbody>
</table>

Table 7-2 Table of reported refolding enhancers used.
7.2.1.3 Experiment design

Experiment designs were produced and experimental data were processed using Minitab v.13.3 statistical software from Minitab Inc. of State College, Pennsylvania, USA and Excel spreadsheet software from Microsoft Corp. of Seattle, Washington, USA.

The purpose of the experiment was to determine what effect the addition of each of the eleven reported refolding enhancing compounds to the refolding buffer might have on the refolding yield of trypsinogen. Also to determine what effects interactions between pairs of these compounds might have on the yield. The experiment was designed in the same way as the screening experiment in section 7.1.1.4. This gave 67 conditions to be tested. Due to material constraints, only the additive-free condition was replicated. The final design contained 72 refolding reactions to be run.

7.2.1.4 Automated refolding reactions

Refolding reactions were performed on a Perkin Elmer Multiprobe II four tip pipetting robot from Perkin Elmer of Boston, Massachusetts, USA. The robot was controlled using WinPrep software, also from Perkin Elmer. The reactions were carried out in Costar 96-well ultra-low binding 350µl well plates from Corning Inc. of Corning, New York, USA. The ultra-low binding plates were selected for their low protein binding characteristics (see section 4.4).

The automated microscale refolding reactions were carried out in the same way as the lysozyme refolding reactions described in section 7.1.1.5, with three differences, as follows. First, a total of 300 µl of refolding buffer was prepared for each refolding reaction. Second, the refolding reaction was initiated by the mixing of 15 µl of denatured trypsinogen solution with 285 µl of refolding buffer (thus giving a 1 in 20 dilution). Third, because trypsinogen refolding had been shown to be insensitive to mixing conditions (Buswell,
Ebtinger, Middelburg), the robot control software's standard settings were used for the pipette mixing.

7.2.1.5 Trypsinogen activation and assay

Samples of trypsinogen were assayed using a method previously developed by Hibbert. Trypsinogen samples to be assayed first needed to be activated by the removal of a prosequence from the protein using 5μg/ml enterokinase. The basis of the assay was the hydrolysis of 1mM Nα-benzoyl-L-arginine-p-nitroanilide (L-BAPNA) by the protein, releasing free 4-aniline whose accumulation was monitored by absorbance readings at 405nm.

Activation and assaying were carried out in a polystyrene 96-well 350μl plate. In practice, 20μl of 30μg/ml enterokinase was added to a well, followed by a 100 μl sample of refolded trypsinogen. The contents of the well were mixed by aspirating and dispensing the contents three times and then allowed to stand at 17 °C for 5 minutes. 100μl of 2mM L-BAPNA solution was then added to the well and the contents mixed again. The absorbance of the wells at 405 nm was then recorded for 10 minutes, blanked against wells containing no trypsinogen. The rate of increase of absorbance of the sample was taken as being indicative of the trypsin activity.

7.2.2 Results

The first part of this section describes the main effects of the reported refolding enhancers, as determined from the screening experiment. The second part describes the two-factor interactions of the reported refolding enhancers, as determined from the screening experiment.
7.2.2.1 Main effects of reported refolding enhancers on the refolding yield of trypsinogen

Figure 7-8 provides eleven main effect plots, which show the main effects of the presence or absence of each of the reported refolding enhancers in the refolding buffer on the activity of the refolded trypsinogen solution. Main effect plots are used to display the influence of each of the factors on the response. In this case the factors are the presence or absence of each of the eleven reported refolding enhancing chemicals in the refolding buffer, and the response is the activity of the refolded protein solution.

The "0" point on each plot indicates the average enzyme activity with that reported refolding enhancer absent from the refolding buffer, and the "1" point indicates the average enzyme activity with it present.
Figure 7-8 Main effects of reported refolding enhancers on the activity of the refolded trypsinogen solution.
7.2.2.2 Two-factor interactions of reported refolding enhancers on the refolding yield of trypsinogen

Figure 7-9 shows fifty-five interaction plots, which display the fifty-five interactions between the eleven factors. The two points on the left side of each plot are at the lower level of the factor labelled below the plot. The two points on the right side of the plot are at the higher level of the factor labelled below the plot. The two points plotted as black circles and connected by an unbroken line are at the lower level of the factor labelled to the left of the plot. The two points plotted as red squares and connected by a broken line are at the higher level of the factor labelled to the left of the plot.

An interaction is present if the change in the response from the low to the high level of one factor depends on the level of a second factor. This is easily seen in an interaction plot as a convergence or divergence of the two lines. The greater the departure from parallel, the greater the interaction.
Figure 7-9 Two-factor interactions of reported refolding enhancers on the refolding yield of trypsinogen.
7.2.2.3 Significance of effects and Pareto chart

Figure 7-10 is a Pareto chart of the main effects and interaction effects of the eleven factors tested. It can be used to compare the relative magnitude and statistical significance of those effects. Each bar on the chart represents an effect, and the dashed red line indicates an alpha value of 0.1, that is, all effects extending past the line are held to be significant at the 10% level. Only the 30 most significant effects are shown.

The most significant effects were those involving arginine, ammonium sulphate and ethanol (labelled A, B and H on the chart) which were strong negative effects, giving little or no trypsinogen activity in the refolded protein solution. The most significant, positive effect was an interaction of PEG 300 and 2-propanol. The most significant, positive, main effect was that of 2-propanol.
Pareto Chart of the Standardized Effects
(response is Activity, Alpha = .10, only 30 largest effects shown)

Figure 7-10 Pareto chart of main and 2-factor interaction effects of eleven reported refolding enhancers on the refolding yield of trypsinogen.
7.2.2.4 Experiment efficiency

The set of experiments on the refolding of trypsinogen described above took under 9 hours to perform. 1.51mg of trypsinogen was used in the reactions though, including wastage, 3.1mg of trypsinogen was consumed by the experiment.

The highest trypsinogen refolding yield achieved in the experiment was with a refolding buffer containing the refolding enhancers 20 mM PEG 300 and 10mM 2-pentanol. The refolded trypsinogen activity with these enhancers was 0.0272 absorbance units/minute which, compared to the additive-free refolded trypsinogen activity of 0.0241 absorbance units/minute, represents an increase of 13%.

It should be noted, though, that this screening experiment was designed to determine which compounds acted as refolding enhancers, not to optimise the refolding yield. It is likely that a further increase in refolding yield could be achieved by a further optimisation experiment using the identified refolding enhancers.

7.2.3 Discussion and conclusions of the experiment

Of the refold buffer additives tested in this experiment, PEG300 showed the greatest positive effect on refold yield. The greatest recovered trypsinogen activity was achieved with the refold buffer containing 20mM PEG300 and 10mM 2-pentanol.

Arginine, ammonium sulphate and ethanol all had a strong negative effect on the yield of active trypsinogen. Precipitates were observed in the microwells containing trypsinogen refolded with ammonium sulphate, suggesting that this concentration of the salt caused "salting-out" of the protein. A lower concentration of the salt may have given a higher yield. Arginine is known to bind to the active site of trypsin. It may be that
refolding in the presence of arginine gave a high yield of correctly folded trypsinogen, but that this trypsinogen had a molecule of arginine blocking it’s active site so that when it was activated to trypsin, little activity was detected.

Had more of the trypsinogen been available, further experiments could have been carried out to screen other potential refolding enhancing additives and to identify optimal refold buffer compositions.

7.3 Discussion and conclusions of the chapter

The experiments in this chapter have shown that automated microwell refold experiments can provide useful process development information in a very time- and material-efficient manner.

In a process development laboratory which deals with a relatively high turnover of new processes (e.g. in a contract biologics development and manufacturing company, where dozens of new processes may be investigated per year) it is desirable to have a systematic approach to refolding development. That is, a system which can be swiftly adapted to any inclusion body protein provided by a client, which can be used to efficiently develop a refold step.

Automated microwell refolding techniques would allow the experimental data required for such refold development to be rapidly gathered. The next chapter will discuss the practicalities of using these automated microwell refold techniques in such a system for refolding development, in the context of a commercial process development laboratory. The equipment requirements and procedures needed to best implement such a system will be discussed.
The application of automated microwell experiments for commercial refold development

The purpose of this research project was to allow the development of inclusion body protein dilution refold steps to be done more efficiently, in order to reduce the time-to-market and the manufacturing cost of recombinant protein drugs. To this end, techniques have been developed to allow dilution refold experiments to be performed with higher throughput and using a smaller quantity of inclusion body material than is required for currently used techniques. This higher throughput of experiments allows design data to be produced more quickly. The small scale of experiments allows more efficient use of inclusion body material.

In Chapter 2 the development of protein refold steps in a commercial bioprocess development laboratory was discussed. In this chapter, lessons learnt during the course of this research project are used for a discussion of the practicalities of applying automated microwell refold experiment techniques in such a laboratory. The ways in which these techniques can best be used are considered. The equipment and procedures which would be necessary to apply these techniques are then discussed.

Automated microscale refolding experiments will be most useful in process development laboratories which have a relatively high throughput of bioprocesses (i.e. a contract biologics manufacturer or a large biopharm company which may deal with 10 or more new processes per year, rather than a small biotech business which may work on only 1 or 2). It is the application of automated microscale techniques in these laboratories which is considered in this chapter.
With future work it will be possible for these techniques be employed as parts of a unified system (coordinated by software or SOPs) which can be applied to developing refold steps for the majority of inclusion body proteins. The planning, coordination and control of experiments performed with such a system is discussed.

8.1 The uses of automated microwell experiments in commercial dilution refold development

It is proposed that automated microwell experiments be used for the commercial development of dilution refold steps (from solubilised inclusion body protein). In this section an approach to development which makes best use of such techniques will be put forward.

Commercial protein refold development is typically done as part of a larger process development effort in order to develop a manufacturing process for cGMP manufacture (see Chapter 2). The scope of the refold development work will depend on the stage which the product has reached in its development life cycle and the existing refold step. Very small scale, high throughput experimentation is most useful for the creation and optimisation of the refold step (see Chapter 2), and it is these applications which are considered in this section.

The development scientist will use any available knowledge of the particular properties of the protein (and the development project) in applying these techniques. The influence of commercial and regulatory factors upon refold development experiments is discussed in section 2.4. Examples of ways in which knowledge of the protein can guide experiment planning include:

- If the protein is known to aggregate upon refolding, but only refold to one stable final structure (i.e. it forms no unaggregated misfolds) then there will be an early emphasis on additive screening experiments with aggregation-inhibiting additives (see section 2.3.4).
• If the protein is known to be more stable under certain solution conditions (pH, ionic strength, additives – perhaps determined through early stability experiments) then those conditions may be used as a starting point for refold buffer optimisation.

• If the protein is known to require a particular metal ion in its native structure then refold buffers will contain that metal ion.

• If the protein is known to have a convenient substrate or binding partner then the inclusion of this in the refold solution will be investigated.

• If the protein’s sequence contains fluorophores then intrinsic fluorescence analytical methods may be investigated.

• Experiments will avoid any additives which are known to irreversibly block the active site of the protein.

• If the protein is known to have no disulphide bonds then there will be little or no experimentation with redox conditions.

• The isoelectric point of the protein will be avoided during refolding, as with other process steps (proteins at their isoelectric point have no net charge – intermolecular repulsion is minimised, which favours unproductive aggregation).

It should be noted that, very early in process development, very little may be known about the protein structure and folding pathway.

Refold steps which are designed using very small scale experiments (such as those described in this chapter) will subsequently be demonstrated at a larger scale before being used in manufacture. This demonstration will require only a very small number of refold reactions (typically three) compared with development experiments.

8.1.1 Refold step creation

In the case of a protein for which no refold step has been developed, or one for which the developed refold step is completely unworkable for manufacture, a refold step must be created de novo. To this end, an
experiment will be performed to find refolding conditions in which some of the protein will refold.

A range of refold buffers will be prepared, comprising solutions at different pH, with different additives (and, for a disulphide bonded protein, different redox reagents). These may be prepared fresh for each experiment. Alternatively an array of standard stock solutions may be prepared (see Section 4.2.1.2), aliquotted into small volumes and then frozen. A set of these would then be thawed for use with each new protein. Any unstable chemical solutions (e.g. solutions of DTT) would need to be prepared fresh for each experiment. Solutions of chemicals which are chemically stable but prone to microbial spoilage (e.g. solutions of sucrose or glycerol) could be made in advance but would require cold storage. Some solutions (e.g. strong solutions of ethanol or NaCl) could be prepared in advance and stored in sealed containers for long periods without cold storage.

An experiment will then be performed in which solubilised protein solution is diluted into these refold buffers using a range of dilution factors (e.g. from 1+9 to 1+999). These refold reactions will be performed by pipetting robot using a similar method to that described in Chapter 7. The refolded protein solutions will then be analysed to determine the degree of success or failure of each refolding reaction (see Chapter 3) and this analysis may also be done using lab automation (see Section 6.3).

The first analyses performed will be turbidity measurements. Those refolds which have very high turbidity, indicating very high levels of precipitation and so low yield, will be screened out. Those refolding conditions which give lower turbidity will then be further analysed using slower, more detailed techniques to assess the refold performance.

The number of runs (the number of refold reactions) which can be performed will be limited by the available time, resources and materials. This initial experiment is similar to experiments carried out using “refolding
kits" (see Section 2.3.4) but the use of automation allows many more refolding conditions to be tested, giving a much higher probability of finding a useful refold reaction. The use of very small scale reactions allows more refolding conditions to be tested using a given amount of protein.

One or more of the most successful refolding conditions are then carried forward for further development. Alternatively, if time or material limitations mean that further development is impossible, or if the performance of one of those reactions is considered to be good enough, then those refolding conditions may be used for manufacture.

8.1.2 Screening for significant factors

Where a refold step exists and is to be developed further, for example a refold step which has just been created, an experiment will be carried out to determine which factors can give a significant improvement in refold performance (note that similar experiments may be carried out during LPQ work to identify critical parameters).

The experiment will investigate the effects that factors such as dilution, pH, refold buffer additives, redox reagent concentrations and refolding time have on the performance of the refold reaction. The effects of interactions between these factors may also be investigated.

DOE tools will be used to create an efficient screening experiment design (see Section 2.5). The screening experiment may use a similar design to those used in Chapter 7. A large number of refold reactions will be performed, testing a large number of refolding conditions, so the use of automated microscale refolding reactions will give considerable time-efficiency and material-efficiency benefits. A large number of refold buffers will be used so lab automation (a robotic pipettor) will be used to prepare these from a smaller number of stock solutions (see sections 4.2.1.2, 6.2 and 7.1.1.3). Dilution refold reactions will be performed using a robotic
pipettor, which will allow simple control of the dilution factor used for each reaction and of factors relating to the composition of the refold buffer.

Analysis of the large number of refolded protein solutions will be performed using automation and the results processed using DOE software. Results of such experiments will be used to identify which factors have the most significant impact on refold performance, which will then be further investigated in optimisation experiments.

### 8.1.3 Optimisation

The final stage in refold step development will be optimisation. An experiment will be carried out to find an optimum refold condition. For all of the parameters which have a significant effect on refold performance (the significant factors identified in a screening experiment) the combination of values which give the best refold performance will be identified.

DOE tools will be used to generate a response surface experiment design (see Section 2.5). The independent variables for this experiment will be the parameters which were determined to be significant in the screening experiment (note that there may be more than two of these factors). The dependent variables (the results of the refolding reactions being tested, e.g. yield, refolded protein concentration) will be calculated together and assessed to determine the performance of each refold condition tested (see Section 2.2).

It is likely that more than one of the parameters investigated in this response surface experiment will be part the refold buffer composition (e.g. additive concentrations, pH). Therefore a large number of similar but different refold buffers will be required. This large number of buffers will be prepared using a robotic pipettor from a small number of stock solutions by a similar method to that described in Section 7.1. Refold reactions will then be performed using the robotic pipettor and analysed using automated
assays. The results will be examined to identify the optimum refolding conditions and to predict operating windows.

8.2 Illustrations of the benefits of performing process development experiments using automated microwell refold reactions

The purpose of this research project was to allow the development of inclusion body protein dilution refold steps to be done more efficiently, in order to reduce the time-to-market, development and manufacturing cost of recombinant protein drugs. To this end, techniques have been developed to allow dilution refold experiments to be performed with higher throughput and using a smaller quantity of inclusion body material than is required for currently used techniques. This higher throughput of experiments allows design data to be produced more quickly and the small scale of experiments allows more efficient use of inclusion body material. The purpose of this section is to illustrate the benefits of using this approach for refold development experiments. The demonstrations of automated microwell scale refold development experiments which are described in Chapter 7 are further discussed, with respect to their material-efficiency and time-efficiency. An algorithm for refold development using this approach is proposed and the theoretical benefits of the application of this to the development of a dilution refold step is discussed, using the example of a protein which underwent process development for Phase I clinical manufacture.

8.2.1 Lysozyme and trypsinogen refold development experiments using automated microwell refold reactions: time-efficiency and material-efficiency benefits

The refold development experiments described in Chapter 7 were performed to demonstrate the usefulness of automated microwell refold experiments for the rapid development of refold steps. The first set of
experiments were on the refolding of lysozyme. The starting point for these refold development experiments was a simple refold reaction, refolding lysozyme at a concentration relevant to commercial manufacture. A screening experiment was performed to identify refold buffer additives which would enhance the yield of active lysozyme. A surface optimisation experiment was then performed to optimise the yield of active lysozyme with respect the concentrations of the two most promising refold buffer additives. In these experiments with lysozyme, 98 different refold conditions were tested in triplicate, giving a total of 294 refold reactions which were tested and analysed. Due to the small scale of these experiments, only 130mg of lysozyme was required. It was estimated that a similar set of experiments, using the same microwell scale refold reactions and the same laboratory automation, could be performed by one scientist in two days.

If similar experiments were performed and analysed by hand at conventional bench scale then they would require considerably more time and materials. In order to estimate these requirements, the following assumptions were made:

- Refold reactions are performed at 200mL scale (c.f. the microwell scale reactions performed at <1mL scale) by dilution of denatured protein solution into a stirred beaker of refold buffer.
- Only one refold reaction is performed under each condition tested (c.f. the automated microwell refold reactions which were performed in triplicate).
- One scientist can prepare, perform and assay 20 refold reactions per day.

Thus, 98 refold reactions are performed, each comprising the addition of 10mL of denatured protein solution (10mg/mL protein in 6M GdnHCl, 25mM DTT) to 190mL of stirred refold buffer. These bench scale experiments require 9.8g of protein (c.f. 130mg of protein required for the automated microwell scale experiments). One scientist could perform these experiments in five days (c.f. two days for the automated microwell refold experiments).
The automated microwell-scale trypsinogen refold buffer additive screening experiment described in section 7.2 comprised 72 refold reactions (testing 67 conditions). If similar experiments were performed and analysed by hand at conventional bench scale (using the same assumptions as above) then 67 refold reactions would be performed, each comprising the addition of 10mL of denatured protein solution (1.4mg/mL protein in 5.5M urea, 10mM EDTA, 10mM DTT) to 190mL of stirred refold buffer. These bench scale trypsinogen refold experiments require 938mg of protein (c.f. 3.1mg protein used in the automated microwell trypsinogen refold experiment). One scientist could perform these experiments in four days (c.f. 1 day was required for the automated microwell trypsinogen refold experiment).

As can be seen by these two illustrations, automated microwell refold experiments offer considerable time and material efficiency benefits, compared to more conventional bench scale refold experiments performed by hand.

8.2.2 An algorithm for rapid refold development using automated microwell scale experiments

As discussed in Section 8.2.1, the use of automated microwell-scale experiments can bring significant time and material-efficiency benefits to refold development experiments. In this section, an algorithm is proposed for refold development experiments, performed during the development of a process for commercial manufacture, using automated microwell-scale refold reactions. The objective of such refold development experiments will be determined by the development scientists, based upon a number of issues (see Section 2.2) but generally the main aim will be to maximise the yield of active protein while also keeping the refold at a practicable volume. If a potential therapeutic protein has progressed far enough through the development pipeline that a process is being developed for commercial manufacture, then it is likely that some protocol already exists to refold the
protein and, though this protocol will probably be very inefficient, it will provide a starting point for refold development (if none exists then one must be created *de novo* – see Section 8.1.1). Determining the yield of a protein refold reaction can be difficult (see Chapter 3) and the method that is used will have a strong influence on the planning of refold development experiments (see Section 8.3). Development of a refold step will, to some extent, be an iterative exercise and the process development team will need to use judgement to decide when to stop optimisation experiments and declare the refold step “developed”.

The proposed algorithm is set out below. Steps in the algorithm are numbered. Each development project is different, with different starting points and different constraints and objectives. The development scientist will therefore use judgement and experience to apply the algorithm flexibly to each project.

A set of simple preliminary experiments, for which automation offers no great advantage, should first be performed to set parameters for further experiments:

1. Compare a continuous-dilution refold reaction (diluting over 24 hours) with a fast-dilution refold reaction (which is then allowed to fold for 24 hours). Decide whether to proceed with continuous dilution (slower and more expensive) or fast-dilution (simpler, cheaper, but possibly giving poor yield).

2. Perform a time-course experiment to determine a sufficient refold time for further experiments.

These preliminary experiments will then be followed by screening and optimisation experiments, for which microwell-scale experimentation, automation and DoE tools offer significant advantages. If a fast dilution refold is selected then the refold will be performed by a pipetting robot using a similar method to those described in Chapter 8. If a continuous-dilution
refold is selected then the pipetting robot will be used to mimic this by diluting in stages, with mixing in-between each dilution. Refold reactions will be performed in microwell plates and experiments will be designed and analysed using DoE software tools, as appropriate.

(3) Select a set of factors which may be expected to influence refold yield (see Section 2.3). Any existing knowledge of the protein or its refolding will be used to select these factors. If no useful information about the protein is available then the following factors should be considered: Refold dilution, refold buffer pH, step or fed-batch dilution, redox reagents (see Section 2.3.5), refold buffer additives (see Section 2.3.4).

(4) Perform screening experiments to determine factors which can have a significant positive effect on the performance of the refold reaction. If time and resource constraints permit, these factors should be tested in combination.

(5) Using the results of the screening experiment select, from those factors tested, a smaller set of factors which have the most significant positive effect.

(6) For each of these factors, select a range of possible values which should be tested. This range should be based upon any prior knowledge of the protein or the specific factor (see Section 2.3) and upon what is feasible in manufacture.

(7) Perform an optimisation experiment, systematically varying the factors across their ranges. The number of refold conditions to be tested will be decided by development scientist based upon time and resource constraints.

(8) Decide whether the “optimum” refolding reaction (from the above optimisation experiment) is acceptable for use as a refold step in the proposed commercial manufacture, whether further refold development experiments are possible or desirable. If it is decided to develop the step further, then perform more development
experiments (starting at step 3), using this first "optimum" refolding reaction as a starting point.

(9) When an acceptable refold step has been defined from microwell-scale experiments, confirm the performance of this refold step at a larger scale.

In the following section the theoretical application of this algorithm to a real-world refold step development will be discussed.

8.2.3 Comparison with conventional commercial refold step development

To further illustrate the usefulness of an automated microwell scale approach to refold development experiments, an example of commercial refold development (as part of a bioprocess development project) will be considered and the benefits which could have been achieved by applying the algorithm in Section 8.2.2 using automated microwell-scale refold experiments will be discussed.

The example taken here is a recombinant disulphide-bonded human protein, expressed in *E.coli* in the form of cytoplasmic inclusion bodies, which forms the active ingredient of an experimental drug product. The contract manufacturer Avecia Biologies was developing the process, with a customer who owned the experimental drug, which it would then operate in its multiproduct GMP manufacturing facility to produce protein for Phase I clinical trials and stability studies. The customer had determined that 70g of GMP manufactured product was required for these. For commercial confidentiality, the customer and the product cannot be named.

At the time of these refold development experiments, the development of the fermentation protocol was well advanced and the downstream process had been outlined and partially developed. Fermenter titres of 5-10g protein per litre of broth were being achieved, and fermentations at development
scale (1L – 5L) had a turn-around time of one week. Fermentation broth was homogenised and the inclusion bodies were harvested and washed using batch centrifugation, producing a paste of washed inclusion bodies. These harvest and wash steps took approximately 5 hours and were very efficient: one litre of fermentation broth produced approximately 50g-80g of washed inclusion body paste, containing approximately 5-10g of the product protein. Inclusion body paste was dissolved in solubilisation buffer (6M GdnHCl, 40mM DTT, 1mM EDTA) at 150mL solubilisation buffer per 1L of fermentation broth, giving a solubilised protein concentration of approximately 30-60mg/mL which was then clarified by centrifugation to remove insoluble material. Attempts had been made to solubilise the inclusion bodies using chemicals other than GdnHCl, but these were unsuccessful. This solubilised inclusion body protein was refolded by dilution. An ion exchange chromatography (IEC) step followed the refold step in the process, so the refolded protein solution had to be conditioned by filtration (to remove precipitates) and dilution (to reduce conductivity) prior to loading. The IEC step and the rest of the downstream process was typical and easy to develop, compared to the refold step.

The most practicable assay available for process development experiments was a RP-HPLC method. The assay had a throughput of approximately two samples per hour. An SDS-PAGE method was used to quantify fermenter titre and protein concentration in solubilising buffer. A microwell-plate based ELISA was possible, but a suitable method had not been developed at the time of the refold development experiments.

The basal refold condition (i.e. the suboptimal refold reaction conditions from which a refold step was to be developed) was a dilution of solubilised protein into refold buffer, with the dilution ratio (calculated from the solubilised protein concentration) to give a final product protein concentration of 0.3mg/mL. It was observed that during refolding, the majority of the product protein was lost to aggregation. Lower refolding concentrations reduced this aggregation, but would require even larger
process volumes. It was also known that the protein required salts in the refold buffer to refold correctly. Due to the inherent difficulty in measuring the concentration of the desired protein in a solution of dissolved inclusion bodies, it was decided to move from a defined refolding protein concentration to a defined dilution ratio (so that refolding could proceed without the need to wait for assay results on the solubilised protein concentration). Defining the dilution factor would also allow the process development team to ensure that the refold step would fit into the process vessels available in the manufacturing facility. Analysis of the process and the manufacturing facility determined that this refold dilution should be 1+23. Smaller dilution factors would increase aggregation, hence reducing yield. Larger dilution factors would decrease aggregation but would give impractically large process volumes (see Section 2.3.3). An early time-course experiment showed that the refold reaction was substantially complete after 5 hours. Refolds were normally allowed to stand for longer than this for practical reasons (e.g. they were left to refold overnight).

The refold development experiments which were performed were:

(a) to select the salts to be added to the refold buffer (and their concentrations).

(b) to determine what other additives should be added to the refold buffer to inhibit aggregation.

(c) to select the redox reagents to be added to the refold buffer (and their concentrations).

Most of the refold reactions performed in these development experiments were performed scaled to 1L fermentation volume i.e. inclusion bodies were prepared from approximately one litre of broth, were solubilised in 150mL solubilising buffer, clarified, then refolded with 23 volumes of refold buffer, giving a total refold volume of approximately 5.5L. Each of these refold reactions took approximately 1 day to prepare. Following refold and analysis time, the refold yield value was available the following day. In some of the later experiments, the solubilised protein inclusion body
solution was aliquotted and a number of smaller refolds were performed using the solubilised inclusion body solution from 1L batch of fermentation broth. The smallest refold volume used in these experiments was 100mL.

The developed (optimised) refold step comprised a dilution of one volume of solubilised protein into a stirred vessel containing 23 volumes of refold buffer whose composition was 0.75M arginine-HCl, 4mM cystine, 0.58g/L NaCl, 0.41g/L MgCl₂, 0.29g/L CaCl₂, 0.2g/L PEG3350, 50mM Tris pH8.

The process development programme prior to manufacture lasted approximately five months. It was estimated that approximately 3.5 man-months were spent on refold development experiments, over the course of four months. Due to the production of broth from high-titre 1-5L scale fermentation experiments, the availability of inclusion body protein for experiments was not a constraint in this case. However, it was estimated that between 300g and 500g of inclusion body protein (i.e. the product of approximately 60L of fermentation broth) was consumed by these refold development experiments.

In order to calculate the benefits which could have been achieved by applying the algorithm in Section 8.2.2 using automated microwell-scale refold experiments to this refold development work, the following assumptions were made:

(i) Suitable automated laboratory equipment is available, as described in Section 8.4, to allow a robotic pipettor to be coordinated with an HPLC system to perform refold experiments.
(ii) As the refold is known to require salts, a salt screening experiment is performed, testing salts for inclusion in the refold buffer, comprising 100 runs. The results of this experiment show that NaCl, MgCl₂ and CaCl₂ have the most positive effect on yield.
(iii) A second screening experiment is performed, testing 15 possible refold enhancing additives singly and in combination, comprising 196 runs (similar to the two screening experiments used in Chapter 7).
The results of this experiment indicate that arginine-HCl and PEG3350 are the most promising yield-enhancing additives.

(iv) A redox-conditions screening experiment is performed, comprising 25 runs. This experiment is not automated, due to the poor shelf-life of some redox reagents. The best yields from this experiment are achieved with cystine in the refold buffer.

(v) A "response surface" optimisation experiment is then performed, comprising 366 runs. The refold yield is optimised with respect to the concentrations of arginine-HCl, PEG3350, cystine, NaCl, MgCl₂ and CaCl₂.

(vi) The automated refold reactions are performed at 10mL scale (i.e. using a deep 24-well plate) because the availability of inclusion body protein is not a significant constraint for this set of refold development experiments. The pipetting operations required to do refold experiments at this scale are expected to be more consistent than those for the smaller 1mL or 0.2mL refold reactions described in Chapter 7.

(vii) The performance of the optimised refold reaction is confirmed by performing the step at larger scale. Washed inclusion bodies from a 5L fermentation are solubilised and refolded (a 28L refold volume).

(viii) It is further assumed that the hypothetical automated microwell-scale refold experiments lead to the same conclusions as the conventional refold experiments which were actually performed, thus leading to the same optimised refold step. This assumption is made because the results of the automated microwell-scale refold reactions cannot be predicted.

In total, these hypothetical development experiments comprise 662 microwell refold reactions, 25 bench scale (100mL) refold reactions and one larger scale (28L) refold reaction.

The assay for the product protein involves an HPLC method with a throughput of approximately two samples per hour. Only one sample can
be assayed at a time, so the experiments are scheduled as described in the “sequential analysis” case described in Section 8.3. The robot operates around the clock, performing refold reactions so that each reaction is timed to finish shortly before the HPLC system is ready for the next sample. Therefore, allowing for reasonable down-time, 40 runs (40 refold reactions) can be performed and assayed per day. With a normal 5-day working week, 200 runs can be performed per week. The 662 microwell refolds can be performed by one development scientist in approximately four weeks, with another week for the bench scale and larger scale refolds; a total of 1.2 man-months.

The microwell experiments (total refold volume 6.62L) consume the equivalent of the inclusion bodies produced in 1.2L of fermenter broth, i.e. between 6g and 12g. The bench scale experiments (total refold volume 2.5L) consume between 2.3g and 4.5g. The larger scale experiment to confirm the performance of the optimised refold reaction (28L refold volume) consumes between 25g and 50g. In total, between 33.3g and 66.6g of inclusion body protein are used, produced from 6.7L of fermenter broth.

The benefits which could hypothetically be achieved by applying automated microwell scale refold experiments to this refold development are summarised in Table 8-1.
In this example, with a ready supply of fermenter broth, the main benefits of using automated microwell-scale refold experiments are in human resource and in time. If a direct cost of £4000 per man-month is assumed, then a saving of £9,200 in human resource costs is achieved for this project. Estimating the value of the time saving is more difficult. In this example, refold development was on the critical path so shortening the refold development work shortens the whole development programme. It is tempting to assume that shortening the refold development by 2.8 months enables the product begin clinical trials 2.8 months earlier, thus getting the product to market 2.8 months early. However the real savings in the length of the overall programme would probably somewhat smaller than this.

In this example, it has been assumed that the hypothetical automated microwell refold experiments achieved the same results as the conventional experiments which were performed. With the time-efficiency and material-efficiency of this type of experiment it becomes possible to perform more refold reactions, allowing experiments which more extensive (i.e. which comprise more runs). Therefore, it seems likely that, in reality, these automated microwell refold experiments would achieve a more efficient refold step, bringing further benefits.
8.3 Scheduling of experiments

In commercial refold step development it is desirable to maximise the throughput of refolding experiments and hence maximise the rate at which data is produced. There are a number of reasons for this. The time available for refold development is often limited (e.g. by manufacturing schedules) and so faster data production means more data can be gathered and so a better refold step can be developed. Even in cases where there is no such limit, it is desirable to minimise the time required to develop the refold step because process development resource and laboratory time is expensive (see Chapter 2). Scheduling of experiments is crucial to maximising throughput. In this section the scheduling of automated microwell refold experiments is considered.

Issues to be considered when planning a schedule for refolding experiments include the time required to initiate refold reaction, the refolding time, the shelf life of refolded protein samples (the time after refolding within which samples must be analysed), the time required to analyse refold reaction and laboratory working hours.

One of the advantages of automated refold experiments is that refold reactions can be performed and analysed unattended, so laboratory working hours is not limiting. The shelf life of the solutions of refolded protein which are produced in the experiment is more problematic. The refolded protein solutions will be in different buffers, at different concentrations and so the way in which those protein solutions will change/degrade during storage will be unpredictable. Freezing or chilling samples to preserve them can lead to changes in the refolded protein, particularly aggregation and precipitation, so these are to be avoided. It is therefore desirable to analyse each refolded protein solution after its refolding time i.e. to analyse the refolded protein solution at the time at which it would be fed to the next step in the process. This couples the rate
at which refolding reactions are initiated to the rate at which refolded protein solutions are analysed. Of these two rates, it is often the analysis rate which is the slower, so the rate at which refolding experiments are performed can be limited by the rate at which samples can be analysed (see Chapter 3). Therefore, the scheduling of refolding experiments will fall into one of two modes depending on the nature of the analytical method being used.

Where only one sample can be analysed at a time, e.g. using an HPLC assay, the initiation of refolding reactions will be scheduled such that refolded protein solutions are ready to be analysed when the instrument is available. This situation will be referred to as “sequential analysis”. Therefore only a small number of reactions will be running at any one time. For example, in the experiments described in Chapter 5, refold time was kept constant at 2 hours (each refolding reaction was allowed to proceed for 2 hours before being analysed) and the analytical equipment could analyse one sample every 15 minutes. Therefore a maximum of eight refold reactions were proceeding at any time.

Where large numbers of samples can be analysed at a time, e.g. an ELISA or activity assay performed in microwell plates, then large batches of refolding reactions will be initiated together then analysed together (each batch will typically comprise one or more microwell plates). This situation will be referred to as “parallel analysis”. In this way, large numbers of reactions are performed in parallel. For example, in the experiment described in Section 7.1, refold time was 2 hours and the assays could be performed in 96-well plates and read using a plate reader. One plate of assays could be prepared and read in 30 minutes. It was therefore possible to have all 201 reactions of the screening experiment proceeding in parallel and so all of the experiment could be performed in one afternoon.

As can be seen from these two examples, experiments using parallel analysis can produce data at a much faster rate than experiments using
sequential analysis. However, sequential analysis methods (e.g. HPLC) typically provide measurements with a smaller error than parallel methods (e.g. ELISA) so fewer replicates may be required.

Experiment scheduling and the types of analysis will therefore be considered when preparing experiment designs. Consider the example of a surface response experiment to optimise the performance of the refold step for a particular protein with respect to a number of parameters. The experimental space in which an optimum refold condition is to be found will be limited by a number of constraints (e.g. pH values between 7 and 9, dilution factors between 5 and 30).

If an activity assay is to be used (parallel analysis) then large batches of refold reactions can be performed at once. An experiment design which covers the experimental space with a fine grid (i.e. a large number) of conditions to be tested, and tests each condition with a number of replicates, may therefore be chosen. An optimum condition can thus be found with a high degree of precision and confidence, using only one experiment.

If an HPLC assay is to be used (sequential analysis) then only smaller numbers of refold reactions can be performed at once. An experiment design which tests a smaller number (i.e. a sparser grid) of conditions to be tested, with fewer or no replicates, may be chosen. This experiment will only find the approximate position of an optimum. A second stage of the experiment may then be performed which focuses on the region of interest (and which uses more replicates) to find the optimum with more precision and confidence.

In the first case (parallel analysis, single experiment) an optimum is found using only one experiment but using a large number of reactions. This approach is chosen because in the first case many reactions can be run in parallel. An experiment with a large number of reactions takes only slightly
more time than an experiment with a small number of reactions and takes much less time than two experiments in sequence. This single stage experiment therefore identifies the optimum faster.

In the second case (sequential analysis, staged experiments) two experiments are required to find the optimum. However a smaller number of reactions are required (than in the first case) because only the region around the optimum is tested in detail. This approach is chosen because in the second case only a low throughput of reactions is possible. Performing two experiments takes only slightly more time than performing one experiment with the same number of reactions and takes much less time than performing one experiment with many more reactions. The staged experiments therefore identifies the optimum faster.

A possible aim for future work would be the development of software which integrates Design-of-Experiments tools with tools for scheduling experiments.

8.4 Capital equipment requirements

It is proposed that dilution refolding steps for commercial bioprocesses be developed using automated microwell techniques. In sections 8.1 and 8.3 the way in which these techniques would be used was described. In this section the equipment requirements for applying these techniques are discussed.

The core piece of capital equipment needed to perform automated microwell refolding experiments is a pipetting robot, also known as a automatic pipettor or a liquid handling robot (see chapters 6 and 7). This performs similar operations to a hand pipette, but the pipette tip is at the end of a “robot arm” which allows it to move to any position within a fixed workspace. The movements of the pipette tip (or tips) and the aspiration/dispensing actions are electronically controlled using computer
software. The movement of the arm around the workspace allows the pipettor to aspirate from, dispense to and mix liquids in almost any open vessel in the workspace. The most commonly used vessels are microwell plates (see chapters 4 and 5) and the workspace will typically hold a number of these, along with larger vessels to act as reagent reservoirs and other equipment. Electronic control of the aspiration and dispensing allow greater control of flowrates than is possible with hand pipetting (see Section 6.4) and allows aspirate/dispense volumes to be changed much faster, which is a great advantage when performing large numbers of slightly different refolding reactions.

The robotic pipettor is used to prepare refold buffers, to initiate refolding reactions and to take samples for analysis. The analysis itself will require different equipment depending on the assays being used (see Chapter 3 and Section 6.3). As was concluded in Chapter 3, the analytical methods which will be most useful for refold step development are turbidity and fluorescence measurements, chromatographic (particularly HPLC) methods and activity assays.

For the purposes of automated microwell refold experiments, turbidity and fluorescence measurements can best be performed using microplate readers. Separate plate readers may be purchased, although units are available which can read microplates using both absorbance (including turbidity) and fluorescence. For the purposes of automation, samples can be dispensed directly into microplates which are sitting on the plate carriers of these machines (for this, the machines must be within the workspace of the robot). Alternatively, microwell plates full of samples may be transferred from the robot's workspace into the plate reader using a plate handler (see below).

Chromatographic assays can be performed using a normal HPLC system, however the transfer of refolded protein samples from the robot to the HPLC is worth considering. There are, broadly speaking, two methods. In one
method, a tube connects the HPLC system’s sample inlet port to a sample receiver in the robot’s workspace. Samples are prepared by the robot and dispensed into the sample receiver from which they are sucked up by the HPLC system. Issues with this method include flushing and dead volume in the transfer line. In the other method, samples are prepared by the robot and dispensed into a carrier (typically either a carousel of vials or a microwell plate) and then samples are aspirated from the carrier using an autosampler integral to the HPLC system. Note that this approach will require either that the HPLC system’s autosampler is in the robot’s workspace or that a plate handler (see below) moves the carrier between the robot and the HPLC system.

In addition to the issues of transferring samples to the HPLC system, clarification of samples prior to loading will also require consideration. If a microwell plate vacuum filtration system and a plate handling system are available, these may be used to clarify samples. Alternatively, an inline filter may be fitted in the HPLC system, either in the sample line, or in line with the column. Settling can also be used to ease the burden on an inline filter – protein precipitates tend to sink and over the course of a refold reaction most precipitated protein will settle to the base of the microwell. Aspirating samples from just below the liquid surface will therefore help. To further enhance this settling effect, plates can be centrifuged using a centrifuge fitted with plate-carrying rotor.

Activity assays vary from protein to protein. Most of these assays involve mixing of assay solutions and samples and measuring a resultant change in a property of the solution (for example the assays for lysozyme and trypsinogen described in sections 7.1.1.7 and 7.2.1.5) which can easily be done using a pipetting robot and the analytical equipment described above. ELISA and other surface-binding assays are usually performed in microwell plates and so are relatively easy to automate.
Temperature control equipment may be needed, either to perform refold experiments at defined temperatures, to keep reagents refrigerated prior to use or to perform activity assays at a controlled temperature. Equipment is available which maintains a solid surface at a set temperature. The surface is shaped so that microwell plates fit snugly onto it and so can be cooled or warmed by it. Alternatively, equipment is available in which temperature-controlled water is circulated around sets of vessels (in a similar manner to a water bath). The entire robot and workspace may be kept in a temperature and humidity controlled environment, in order to minimise spoilage or evaporation of samples (see Section 4.2.2.3). Sample heating equipment is much smaller and cheaper than sample cooling equipment. If temperature control is needed it may be decided to keep everything in a cold environment (e.g. a refrigerated cabinet or cold-room) and achieve temperature control of individual items using only heating.

In order to extend the capabilities of the automated system, a plate handler may be used. This is a robot arm which can pick up and move microwell plates (and other small items of labware) around. It can be used to transfer microwell plates full of samples (or sample carousels) between the robotic pipettor’s workspace and analytical equipment. It can move microwell plates between the workspace and storage equipment, dramatically increasing the number of plates which can be used (this would otherwise be limited to the number which could fit on the pipetting robot’s workspace). Plate handlers are particularly useful if automated systems are to be operated around the clock, as they perform many of the tasks which would otherwise be done by hand.

If filtering of samples is required, then a vacuum filtration station may be used. This comprises an electronically controlled vacuum pump connected to a vacuum chamber. A filter plate (similar to a microwell plate, but the bottom of the wells are made of filter material) is filled with samples, placed on top of the vacuum chamber and the vacuum applied. Liquids in the wells are sucked through the filter materials and drip into corresponding wells in a
A process development laboratory wishing to implement automated microscale refolding techniques will probably decide to initially purchase only a basic pipetting robot and any analytical equipment needed. Additional items such as plate handlers could then be purchased at a later date, after the system had proven its worth, in order to increase capabilities.

The cost of purchasing a basic but flexible single tip pipetting robot is in the region of £10,000 (a four-tip pipetting robot, such as the one described in chapters 6 and 7 is in the region of £20,000). For comparison, this is similar to the cost of employing one process development scientist for 30 days, or operating a cGMP manufacturing plant (suitable for producing material for Phase I clinical trials) for 12 hours. Servicing and operating costs for such a robot are estimated to be in the region of £2,000 per year.

The cost of purchasing a microplate reader is around £20,000 and an HPLC system is around £50,000. It is likely, however, that any process development laboratory which might consider using automated microwell refold techniques would already use these items of equipment.

In addition to refolding experiments, a pipetting robot has a number of other uses in the process development lab, the most obvious of which is preparing samples and performing repetitive assays.

8.5 Procedures (software) for a unified system for rapid scaledown refold development

The pieces of automation and analytical equipment described in section 8.4 will each be controlled by computer using software supplied by the manufacturer. This section contains a brief discussion of the additional software and procedures needed to:
• Design and schedule experiments and analyse results
• Plan refold buffer make-up
• Coordinate the control of all of the equipment

8.5.1 Experiment design, scheduling and analysis
A number of statistical software packages exist to aid in the efficient design of experiments (DOE software - see Section 2.5). These programmes can produce experiment designs in the form of array (or spreadsheet) files. These designs can then be followed by equipment-controlling programmes. Analytical data can then be fed back into the statistical software for processing into results. As concluded in Section 8.3, it would be useful to develop software to consider and plan scheduling of experiments alongside experiment design.

8.5.2 Refold buffer planning
Refold development experiments will require the preparation of large numbers of different refold buffers. These can be prepared by robotic pipettor from a smaller number of stock solutions (see sections 4.2.1.2, 6.2 and 7.1.1.3). The problem of which stock solutions to use, and the combination of different volumes of different solutions to make up the required range of refold buffers, can be complex. Computer software to facilitate these calculations will save time and effort. The inputs to the software will be an array detailing all of the refold buffers required. There will be two outputs from the software. The first is a list of required stock solutions which the development scientist must either make up or retrieve from storage. The second is an array (or spreadsheet) of data to tell the robotic pipettor how much of each stock solution to mix together to form each refold buffer.
8.5.3 System coordination

In order for the robotic pipettor, analytical and other equipment to work together it will be necessary for them to be coordinated by a single computer programme. This master programme will control the separate programmes which control the individual pieces of equipment.

Each piece of equipment (robot, plate reader, HPLC system, plate handler) will typically have its own controlling programme. For new equipment these programmes will be supplied by the manufacturer. For legacy equipment, these programmes may be written by the user (see Section 3.9.4). A separate computer system may be used to run the equipment programme for each piece of equipment, or it may be possible to run a number of programmes on one computer system. One master programme (coordinating the whole experiment) will in turn control each of these equipment programmes, sending instructions to them and receiving information from them. This master programme may be specifically written for the task. Alternatively, the software controlling one piece of equipment may be made to control the other equipment programmes. For example, the WinPrep software which is used to control the Multiprobe II could be used to direct the software which controlled the Spectracount plate reader (see chapters 6 and 7). This master programme will direct the automated equipment to perform the experiment, following the experiment design.

8.6 Conclusions of the chapter

It will be practicable to use computer software to coordinate automated equipment to perform microscale refold development experiments. The use of this system will allow automated microscale refold experiments to be performed with efficient use of time and material, and with only a small amount of operator time.

This system will first be developed by a company which develops a large number of protein refold steps, e.g. a biopharmaceutical contract
manufacturing organisation (CMO) or a large pharma company. The system will initially be developed using a limited amount of equipment, (a robotic pipettor, a plate reader and an HPLC system) and so may still require intermittent operator intervention e.g. to transfer microwell plates full of samples. Once the system has proven its worth, capabilities may be expanded e.g. by incorporating a plate handler to extend the scope for unattended (hence round-the-clock) operations. This will further enhance the time and material efficiency of automated microscale refold experiments.

The process development laboratory, in which the system discussed in this chapter would be used, will be one part of a whole company. The manufacturing facility in which the developed refold step will be operated will be another part of the company. The usefulness of an automated microscale experimentation system will be judged in terms of its impact upon the profitability of the whole company. An understanding of the financial and operating structure of a biologics process development and manufacturing company was developed by studying the profitability of recent projects performed by such a company - this work is described in the next chapter.
9 Commercial and financial issues in a bioprocess development and manufacturing business.

In order to develop further an understanding of the commercial environment in which protein refold development work is performed, a project was undertaken for the business team at the sponsoring company (Avecia Biologics). In this project, knowledge and skills which had been developed during the business and transferable skills course elements of the EngD programme were applied to address real business problems. That commercial project is described in this chapter.

9.1 Introduction to the company

Avecia Biologies is a contract therapeutic protein process development and manufacturing organisation (CMO). Previously, the business had only done development work (no manufacture). At the time of this study (Autumn 2001), it had recently opened a cGMP manufacturing facility with two streams, each equipped with 100L – 1000L scale fermenters and associated downstream processing equipment. This facility, named the Advanced Biologies Centre (ABC), is used to produce products for preclinical and early phase clinical trials. Only a handful of manufacturing projects had been performed since the facility was opened.

Avecia Biologics was a relatively small company, employing approximately 120 people. It was planning to expand its process development and manufacturing capabilities, increasing its staff to around 300 people and building a new cGMP manufacturing facility (3000L fermentation scale with two independent streams) to manufacture late stage clinical trial and launch product.
9.2 Background

A handful of process development and manufacturing projects had been performed for customers. Avecia’s management were concerned that some projects had been less profitable than predicted, which had adversely affected the business’ profits. It was suspected that underestimates had been made when development and manufacturing costs were predicted prior to contract negotiation. In addition, a need was identified for a mechanism to track the cashflow of projects on a monthly basis. Management were keen to address these problems before the planned business expansion the following year.

As part of the commercial element of the EngD course, it was decided to investigate recent development and manufacturing projects, to present the resulting observations and also to develop an administrative system for tracking project costs. What follows is the outcome of that investigation.

9.3 Project costs, cost predictions and pricing

In order to investigate the causes of the low profitability of projects, information was gathered about current and recently completed process development and manufacturing projects. Interviews and meetings were held with project managers and other key personnel. Reports (written for customers) on completed development and manufacture work were studied. Cost predictions, made before contracts were negotiated, were retrieved. Data from the business’ monthly accounts were compiled and examined and retrospective accounts for each project to-date assembled. These accounts were then compared with cost predictions for those projects. From this collected information a number of observations were made.

Project cost predictions were made using a simple cost model – estimates were made of the number of man-months of work required from each function (R&D, QA, QC) for each part of the project and of the number of weeks that would be required in the manufacturing facility. These figures
were then entered into a spreadsheet and multiplied by cost figures for each man-month and manufacturing-week. Overhead costs were then added to calculate a total predicted cost for the project. This predicted cost was communicated to the sales team and, with a profit margin added, became the price basis for contract negotiations. Any underestimates of costs in these predictions would therefore lead to low prices being offered in contract negotiations, which would lead to low or even negative project profitability.

Project proposals, including project cost predictions, were prepared by project managers, who were all from the R&D department. Perhaps because of this, cost predictions for R&D work were very accurate. Project managers had only limited contact with other departments (i.e. quality assurance (QA), quality control (QC), process engineering and manufacturing). Consequently they sometimes lacked an appreciation of the amount of work done by these departments, which led to underestimates of the cost of their contribution to projects (project managers were aware of what tasks were being done by the other departments but not of the amount of work required to perform them).

Predictions for the amount of time required in the manufacturing facility were typically 20% too low and consequently manufacturing costs were underestimated. This led to manufacturing projects giving much lower % profit than R&D-only projects which was a particular problem because manufacturing projects were of much higher value. The greatest cause of this underestimate of manufacturing time was the estimator not considering the time required for pre-campaign and post-campaign cleaning. The amount of supporting work done by R&D and QC personnel during manufacturing campaigns was also significantly underestimated. In addition, the people writing proposals (and also the customers) did not always fully appreciate the amount of validation work that was required for manufacture of product for different phases of clinical trials.
The business had only recently begun doing projects which involved manufacturing. It was expected that manufacturing cost predictions would swiftly become more accurate as staff gained experience of the requirements of GMP manufacture.

For the purpose of cost prediction, project length was estimated as a number of weeks and the manufacturing facility usually worked on a weekly cycle. However the cost of operating the facility had been calculated on a monthly basis (i.e. annual cost divided by twelve). A conversion rate of 4 weeks = 1 month was normally used in cost prediction, which exacerbated inaccuracies.

9.4 Project cashflow

Avecia Biologics' existing accounting systems collated cost information on a monthly basis against a number of cost centres. Elements of the business each had a cost centre to which charges were allocated whenever the business made payments. For example, there were separate cost centres for:

- each section's salaries
- each area's material, consumables and maintenance expenditure
- manufacturing facility's running costs

There were also cost centres for each Project, though these could only be used for charging consumables and salaries which were directly charged to that project (i.e. staff who were entirely dedicated to that project - typically only R&D staff)

One of the key tasks of the Business Accountant was to collate these cost figures at the end of each month, along with figures for income (typically comprising milestone payments from customers), to compare those against the business' current account and, from that, to calculate earnings before interest, tax, depreciation and amortisation (EBITDA), the preferred measure of profit, as well as gross earnings for the month, the year to-date
and projections for the rest of the year. These figures were then presented to senior management.

The Business Accountant monitored the company's cashflow according to normal accounting practice. It was, however, difficult to monitor the cashflow of each project individually. The monthly profit of the company was transparent but it was desired to also monitor the monthly profitability of each project and to predict each project's cashflow for the remainder of the year in order to control better the whole company's cashflow. To this end, a spreadsheet system was developed with the cooperation of the Business Accountant.

The spreadsheet took the following inputs each month:

- Income from each project (received milestone payments) from current accounts.
- Direct costs for each project's cost centre, from the existing accounting system.
- Manufacturing occupancy (amount of time which that project occupied the manufacturing facility) in the current month.
- Total business costs for that month, from the existing accounting system
- Predictions (from Project Managers) of future resource and manufacturing time requirements, and of future milestone payments.

From these inputs the spreadsheet calculated the total costs of each project (including manufacturing costs and overhead costs allocated to that project) and profits for each project for the current month and the year-to-date. The total of the costs for all projects was then compared against the total costs of the business, in order to allow the allocation of overhead costs to be checked. The spreadsheet then calculated predictions for project costs and profits for the remaining months of the financial year, and for future years. These could then be compared with the business' budget in order to better monitor and predict business performance.
9.5 Outcomes

The observations made about project cost prediction were presented to a meeting of managers including the R&D Manager, the Business Technology Manager, the Business Development (i.e. sales) Manager, the Strategy Manager and the Business Accountant. A committee was set up to develop an improved system of cost prediction. One result of this process was that a member of the commercial team was assigned to perform all future project proposal writing (including cost prediction) tasks, consulting with each department as appropriate. This specialisation allowed the proposal writer to develop an understanding of projects and costs across functional boundaries. The proposal writer could quickly gain experience of writing a large number of proposals and, by monitoring the costs of those projects, continuously refine the cost prediction system. These helped to increase the accuracy of cost predictions and consequently improve profitability.

The projects cost tracking spreadsheet concept was taken and developed further, as part of a larger project to improve business systems. A unified “project tracker” system was thereby developed. In addition to recording each project’s progress (in each of the functional areas) against the contracted plan and recording present and forecast cost, the full system also uses this information to facilitate financial projection and resource planning and scheduling for the whole business.
10 Conclusions

The aims of this Eng.D. research project were to develop techniques which can improve the way in which commercial protein refold steps are developed, in order to deliver data for process design in a more time-efficient and cost-efficient manner. In particular, the project focussed on the process development of dilution refold steps for the production of soluble proteins from chaotrope solubilised inclusion bodies, as this is the most common commercial refold step.

10.1 Refold step development

It was determined that the most efficient refold experiments would be performed using very small scale reactions running in parallel. Microwell scale reactions and automation technologies (similar to those which are commonly used in drug discovery) were identified as offering a promising approach to improving refold development experiments.

The practice of protein refolding in commercial bioprocesses was studied. The way in which such refold steps are developed was investigated. The aims and scope of commercial refold development projects were explored. The applications of design-of-experiments (DOE) tools to design efficiently refold development experiments was then set out.

10.2 Analysis of refold reactions

The difficulties which are inherent in measuring the yield of protein refolding reactions were considered. Analytical methods which could be used to quantify the degree of success or failure of refolding reactions were examined. It was concluded that the analytical methods which would be used in refold development experiments would be different for each protein and that different methods may be appropriate for different stages of
development. For early screening experiments, where a large number of different refolding reactions are tested in order to identify classes of conditions for further investigation, fast analytical techniques such as turbidity or intrinsic fluorescence measurements are suitable. For later optimisation experiments, where more accurate information is required but where fewer refold reactions may be performed, more detailed but slower techniques such as HPLC methods will be used. Activity assays are the preferred analytical method for refold development experiments, but it was recognised that practicable activity assays are not available for many proteins.

10.3 The microwell format

The use of microwell plates as a format in which to perform protein refolding experiments was investigated. Pipette-mixing was selected as the best method of mixing dilution refold reactions. It was found that non-specific binding of the refolding protein to the microwell surface could lead to significant losses but that the use of hydrogel-coated microwells or chaotrope-containing solutions could prevent this. It was demonstrated that dilution refold reactions scaled well between the microwell scale (200μl) and the bench scale (200ml) at which conventional refold development experiments are performed. It was therefore concluded that microwell scale protein refolding reactions could be used to efficiently provide useful data for process development. The use of microwell refolding experiments was demonstrated for the development of a refold step for the commercial manufacture of insulin-like growth factor (IGF-1). A significant amount of valuable information was gained using a very small amount of the inclusion body protein.
10.4 Automation of microwell scale refold experiments

The use of laboratory automation, in particular a pipetting robot and automated analytical equipment, to facilitate rapid microwell scale refolding experiments was investigated. It was concluded that such automated equipment could perform all of the operations necessary to perform refold development experiments. The mixing of dilution refold reactions using the pipetting robot was tested and it was concluded that the robot easily gave sufficient mixing control. It was concluded that the use of laboratory automation, with microwell scale reactions, could significantly increase the time-efficiency of protein refold development experiments.

The usefulness of automated microwell scale refolding techniques was demonstrated in experiments to develop the refold steps of two proteins. Refold optimisation experiments were carried out using lysozyme as an example protein because it is well characterised and has a convenient activity assay. Experiments were also performed to develop the refold step of trypsinogen, a commercially interesting protein. It was shown that these techniques allowed refolding experiments to be performed much more quickly, using far less protein, than in conventional bench scale experiments.
The way in which these automated microwell refolding techniques could be brought together in a coherent system for commercial refold step development (suitable for use in a commercial process development laboratory) was then discussed. The application of this system to different stages in refold step development was considered. The best mode of scheduling experiments, according to the type of experiment and the analytical method used, was explored. The purchasing of capital equipment needed to implement this automated microwell refold development system, and the software required to make best use of it was then discussed. It was concluded that such a system for the rapid automated microscale development of protein refold steps would first be implemented by a company which deals with a high rate of inclusion body processes e.g. a contract process development and manufacturing organisation (CMO) or large pharma company.

10.5 Commercial awareness

A project was described, in which problems relating to project cost predictions and cash-flow in a (CMO) were investigated. This project casts light on the commercial environment in which biopharmaceutical manufacture and process development is performed and in which an automated microscale system for the rapid development of protein refold steps will be used.
10.6 Summary

The work presented in this thesis has developed techniques to improve the efficiency of commercial refold step development experiments. Through the use of microwell refold reactions, laboratory automation and DOE tools one or more of the following advantages may be gained. Refold development work can be done in less time, using fewer man-hours and requiring less inclusion body material, which will reduce development costs and time-to-market of new therapeutic proteins. Within fixed limits of time, resource or available material, more information can be obtained from refold development experiments, leading to more efficient refold steps which will in turn reduce protein manufacturing costs.
11 Future Work

The work presented in this thesis indicates a number of opportunities for future work. Summaries of possible future work are given in this chapter.

11.1 Connection of development work with process and financial modelling

In Section 2.2 an objective function for refold step optimisation was considered i.e. a function which combines all of the outcomes of a tested refold step (yield, refold volume etc) into a single measure of how good a refold step is. It was concluded that the ideal objective function would be profit. Software which combines bioprocess and financial models would allow the developer to estimate the effect which any tested refold step would have on profit (indeed this could be applied to the development of other process steps as well). Results from experiments could then be analysed with reference to these combined process and financial models in order to maximise profit. The development of such software would allow more profitable bioprocess development.

11.2 Intrinsic Fluorescence

In Sections 3.9.2, 3.9.3 and 3.9.4 the use of variation in intrinsic fluorescence of proteins undergoing chaotrope-induced unfolding and refolding was explored. In these experiments, excitation wavelength was held constant and the emission wavelength was scanned (i.e. to find $\lambda_{\text{max}}$ the wavelength of the emission maximum for the fixed excitation). This approach was found to be useful for measuring unfolding only for some proteins. An alternative approach would be to find the wavelength of the excitation wavelength giving maximum emission magnitude. That is, to monitor emission at a fixed wavelength and to scan through excitation wavelengths to find the excitation wavelength which gave the greatest
emission. The way in which this wavelength changed during unfolding or refolding of proteins could then be investigated. Excitation is from the normal state (whereas emission is from the excited state) so it may be that excitation data give more information about the state of the protein in the solution. Development of this method could offer a different way of measuring the degree of folding of a protein solution, which may be more widely applicable.

11.3 Equipment to facilitate high throughput experiments

The automated equipment used in this work can significantly improve the throughput of refold step development and other experiments. More equipment could be developed in future work to work with this automation equipment in order to facilitate or extend the scope of such experiments.

As described in Section 4.2.1.4, a device to place and remove numbers of magnetic stirrer bars from microwell plates would facilitate any experiment requiring mixing (magnetic stirring could be used in place of pipette mixing). In order to prevent problems associated with grinding between the stirrer bar and the base of the well (e.g. cell damage or breaking up of precipitates) a stirrer bar could be developed in which the magnet is connected to a float such that the stirrer floats below the surface of the liquid instead of sitting on the base of the well.

During long automated experiments using a pipetting robot, evaporation from reservoirs and wells could become a problem. Enclosing the robot in a temperature controlled cabinet can reduce this problem. Development of a device to control the humidity in such a cabinet (analogous to a thermostat controlling temperature) could further reduce evaporation problems. Vessels could developed which are fitted with lids (or flaps) which are held closed (up) with light springs and open downwards when the pipette tip pushes them aside (i.e. when the pipette tip moves to dispense/aspirate from the liquid under the lid).
Devices exist to measure and control pH in a liquid using a pH probe and acid/alkali addition pumps (such pH controllers are commonly fitted to fermenters). The development of such a pH controller which is small enough to adjust pH in a microwell and which could be mounted alongside a pipette tip on a pipetting robot would facilitate automated experiments (e.g. refold buffer preparation for refold development experiments).

11.4 Refold buffer planning software

Refold development experiments will require the preparation of large numbers of different refold buffers. These can be prepared by robotic pipettor from a smaller number of stock solutions (see sections 4.2.1.2, 6.2 and 7.1.1.3). The problem of which stock solutions to use, and the combination of different volumes of different solutions to make up the required range of refold buffers, can be complex. Computer software to facilitate these calculations, as described in Section 8.5.2, will save time and effort.

11.5 Design and Scheduling of Experiments

Scheduling is crucial to the efficiency of development experiments. As concluded in section 8.3, it would be useful to develop software to consider and plan scheduling of experiments alongside experiment design.
12 References


Ho J.G.S., Middelberg A.P.J., Ramage P., Kocher H.P., 2003. The likelihood of aggregation during protein renaturation can be assessed using the second virial coefficient. Protein Science 12 pp.708-716


- 267 -


Ueda T., Nagata M., Imoto T., 2001. *Aggregation and chemical reaction in hen lysozyme caused by heating at pH6 are depressed by osmolytes, sucrose and trehalose.* Journal of Biochemistry Vol130, pp491-496


