AN INVESTIGATION OF
C-MYC

Negative autoregulation
of the
c-myc oncogene

Edward Marc Laufer

A thesis submitted in partial fulfilment of the degree of
Doctor of Philosophy
University of London

Department of Biology
University College London
Gower Street
London

Imperial Cancer Research Fund
Lincoln's Inn Fields
London

October 1990
Abstract

C-myc protein, when expressed at sufficient levels, can repress the initiation of transcription from its own promoter. The DNA sequences within the human c-myc promoter through which the c-myc protein acts, and the proteins which bind these sequences in vitro, were investigated.

Chimaeric c-myc promoter:indicator constructs were stably transfected into Rat-1 fibroblasts, and the resultant colonies pooled. The cells were then infected with a retroviral vector which directs expression of myc protein, and polyclonal populations again pooled. The responsiveness of each construct to myc protein was assessed by quantitation of indicator mRNA. Deletion analysis revealed that sequences from -71 to +47 base pairs relative to the P2 promoter mRNA cap site are sufficient to impart myc sensitivity to P2 promoter-dependent transcription. Systematic mutation of this region suggests that it contains multiple, redundant response elements which are also involved in the control of basal transcription.

Several protein complexes were found to specifically interact with these sequences in gel retardation assays. Two of these, MAC II and MAC I were investigated in detail.

MAC II is recognised by two c-myc peptide-specific monoclonal antibodies, and may contain c-myc protein. MAC II recognises both sequence dependent and independent properties of a double stranded, plasmid-derived linear DNA fragment of the P2 promoter, maximally extending from positions -63 to +47. The sequence independent component appears to be stabilised by hydrogen bonding, and may involve the formation of a secondary structure.

MAC I is recognised by one c-myc peptide-specific monoclonal antibody, and contains myc related protein which is probably distinct from c-myc. It recognises a single stranded DNA sequence on the antisense strand, which spans the P2 promoter mRNA cap site and extends from approximately positions -30 to +31.

Possible models of the relationship between these complexes, c-myc autoregulation and myc protein function are discussed.
Acknowledgements

As the only tenured graduate student in the history of the Imperial Cancer Research Fund, it is with mixed enthusiasms that I say goodbye. I am thankful to everyone who has tolerated my idiosyncrasies for the past five years.

I would like to express my gratitude to Hucky for his guidance, patience and optimism that it would all work out in the end; and to Martin Raff for his encouraging words of advice.

I would also like to acknowledge those who have given me the excellent technical support I have received over the years. Particular thanks go to Graham Soundy, who taught me to drive a VAX; and to Iain Goldsmith and his crew of oligo synthesisers, who made a 132-mer for me that cloned first time.

My time at the ICRF has been greatly enriched by all of the people who have been generous of their own time, and willing to share their knowledge and ideas with me. Particular kudos here to Nic Jones, Steve Goodbourn, David Bentley, Mike Fried, Gerrard Evan and all the members of their and the numerous other labs through which I have wandered.

Special thanks as well to Jay, Anne, Gay, Tim, Linda, Mary, Xin, Shaun, Carolyn, Caroline, Jonathan, Sasha and Bruno: the generations in Hucky land with whom I have had the pleasure and privilege to work and learn and grow.

All the members of the extended Highbury family, especially Neil, Lisa, Mike and Lesley, and even Don Pietro, have given me the perspective to maintain my sanity.

Finally, I must acknowledge my family, for being there.
# Table of Contents

Title page .................................................................................................................. 1
Abstract ..................................................................................................................... 2
Acknowledgements ................................................................................................... 3
Table of Contents ...................................................................................................... 4
Figures ...................................................................................................................... 11
Tables ....................................................................................................................... 15
Abbreviations .......................................................................................................... 16

Chapter 1 Introduction .......................................................................................... 18
  1.1 Cancer and oncogenes ................................................................................ 19
  1.2 Oncogene cooperation .......................................................................... 21
  1.3 Roles of nuclear oncogenes ...................................................................... 23
  1.4 c-myc: Overview ....................................................................................... 24
    1.4.1 Gene structure ................................................................................. 24
    1.4.2 Protein ............................................................................................. 26
    1.4.3 Homology with transcription factors .............................................. 27
  1.5 Activation of the myc oncogene ............................................................... 29
    1.5.1 Retroviral transduction .................................................................. 29
    1.5.2 Retroviral insertion ........................................................................ 31
    1.5.3 Chromosomal translocation ............................................................ 31
    1.5.4 Gene amplification .......................................................................... 32
    1.5.5 Regulatory mutations are sufficient for oncogenic activation ....... 33
  1.6 Biological functions of c-myc ...................................................................... 34
    1.6.1 Role in development and differentiation in vivo .......................... 35
    1.6.2 Role in cellular differentiation in vitro ........................................... 35
    1.6.3 Mediator of mitogenic stimuli ....................................................... 37
    1.6.4 Role in cell cycle ............................................................................ 39
  1.7 Multiple levels of regulation of c-myc expression ..................................... 41
    1.7.1 During differentiation and development .................................... 41
    1.7.2 During mitogenic stimulation ....................................................... 43
  1.8 Mechanisms regulating c-myc expression ............................................... 44
    1.8.1 Cis-acting transcriptional control elements and trans-acting factors 45
      1.8.1.1 Minimal promoter elements .................................................. 45
1.8.1.2 Upstream and downstream elements ..........48
1.8.2 Control of transcript elongation ............................50
1.8.3 mRNA stability .........................................................51
1.8.4 Protein translation and degradation .............................52
1.9 Molecular functions of c-myc: Regulation of gene expression ...............................................................52
1.10 Autoregulation ............................................................53
1.11 Aims of the project .......................................................56

Chapter 2 Materials and Methods ........................................62
2.1 Materials .................................................................63
2.1.1 Reagents ...............................................................63
2.1.2 Enzymes ...............................................................66
2.1.3 Bacterial strains and genotypes .....................................67
2.1.4 Cloning vectors ........................................................67
2.1.5 Cell Lines ..............................................................67
2.1.6 Stock solutions ........................................................67
2.1.7 DNA constructs .........................................................70
2.1.7.1 Human c-myc gene promoter constructs:.........................70
2.1.7.2 RNAase protection probes ............................................78
2.1.8 Gel retardation probes ....................................................81
2.1.8.1 Plasmid fragments .......................................................81
2.1.8.2 Synthetic oligonucleotides .............................................81
2.1.9 Antibodies ............................................................83
2.2 Methods .................................................................84
2.2.1 Tissue culture ..........................................................84
2.2.1.1 General ...............................................................84
2.2.1.2 Transfection ..........................................................85
2.2.1.3 Retroviral manipulations .............................................87
2.2.2 General analytical techniques ......................................88
2.2.2.1 Quantification ......................................................... 88
2.2.2.2 Precipitation ........................................................... 89
2.2.2.3 Phenol extraction .................................................. 89
2.2.2.4 Agarose gel electrophoresis ................................. 90
2.2.2.5 Native polyacrylamide gel electrophoresis (PAGE) .......... 91
2.2.2.6 Denaturing PAGE ................................................. 92
2.2.2.7 Densitometry ....................................................... 93
2.2.3 DNA ........................................................................................ 93
2.2.3.1 Competent bacteria ............................................... 93
2.2.3.2 Bacterial transformation ..................................... 94
2.2.3.3 Plasmid minipreps ............................................... 94
2.2.3.4 Plasmid maxipreps ............................................... 95
2.2.3.5 Restriction digests ................................................. 97
2.2.3.6 Ligations.................................................................. 97
2.2.3.7 Oligonucleotide manipulations ........................ 99
2.2.3.8 Colony hybridisation............................................. 101
2.2.3.9 Deletions................................................................ 102
2.2.3.10 DNA Sequencing ................................................ 103
2.2.3.11 Southern blotting............................................... 104
    genomic DNA........................................................... 104
    cutting.................................................................. 104
    Transfer................................................................ 104
    Filter Hybridisations.............................................. 105
2.2.3.12 Random priming................................................ 105
2.2.3.13 Spun columns ..................................................... 105
2.2.4 RNA ........................................................................................ 106
2.2.4.1 RNA isolation ....................................................... 106
2.2.4.2 RNA purification: Guanidinium prep.......................... 107
2.2.4.3 RNA purification:
    Guanidinium/LiCl.................................................... 108
2.2.4.4 RNAase protection ............................................ 108
    RNA probe preparation........................................ 108
    hybridisations...................................................... 109
2.2.4.5 Northern blotting.............................................. 110
2.2.5 Protein:DNA interactions .............................................. 111
2.2.5.1 Probe isolation and labelling................................. 111
2.2.5.2 Whole cell extract preparation ..........................112
2.2.5.3 Gel retardation assays........................................113
   protocol ..................................................................113
   conditions ................................................................114
2.2.5.4 DMS footprinting ..............................................114
2.2.5.5 Formic acid modification ....................................116
2.2.5.6 Hydrazine interference ......................................116
2.2.5.7 Maxam and Gilbert markers ...............................116
2.2.6 Protein ....................................................................................116
   2.2.6.1 Protein quantification ..........................................116
   2.2.6.2 Immunoblot extract preparation .......................117
2.2.7 Computer assisted analysis ........................................117
   2.2.7.1 DNA.........................................................................117
   2.2.7.2 Protein database searches .................................117

Chapter 3 Promoter deletion analysis .............................................................123
3.1 Results..................................................................................................124
   3.1.1 Analysis of multiple indicator constructs ..................124
      3.1.1.1 Transfected genomic human c-myc
            constructs are appropriately regulated by
            v-myc in Rat-1 cells ..............................................127
      3.1.1.2 Human c-myc promoter fusion
            constructs also respond to v-myc ........................128
      3.1.1.3 5' end mapping of transfected fusion
            genes .....................................................................128
      3.1.1.4 v-myc response requires c-myc
            promoter sequences .............................................130
      3.1.1.5 Transfected indicator constructs
            respond to serum as well as to v-myc ...............131
      3.1.1.6 Conclusions .....................................................133
   3.1.2 Analytical strategy to identify a 'myc response
      element'...............................................................................134
   3.1.3 Results of large scale deletion series .......................136
      3.1.3.1 General observations .......................................136
      3.1.3.2 Deletion constructs -353 to -91 .......................138
            Effect of v-myc ....................................................138
            Promoter ratio .....................................................138
      3.1.3.3 Deletion constructs -91 to +430 .......................139
Chapter 3 Characterisation of P2 promoters

3.1 Characterisation of c-myc and a-globin signals
3.1.3.4 Characterisation of c-myc and a-globin signals.................................................140
3.1.3.5 Conclusions.............................................................141
3.1.4 Deletion constructs -88 through +97 .......................................................142
3.1.4.1 Introduction ...........................................................142
3.1.4.2 Effect of v-myc.......................................................142
3.1.4.3 Implications of passage dependent alterations of signal ..................................143
3.1.5 Reduction of non-specific upstream signal ..............................................144
3.1.5.1 Introduction ...........................................................144
3.1.5.2 Characterisation of upstream signals .......................................................145
3.1.5.3 Use of alternative vectors .......................................................146
3.1.5.4 Suppression upstream transcripts via SV40 terminator sequences ...............147
3.1.6 Clonal transfectant analysis.............................................................................148
3.1.7 Signal degradation timecourse.........................................................................151
3.1.8 -88SVter through +119SVter deletions ..............................................152
3.1.9 Linker scan series..........................................................................................155
3.1.10 Effect of SV40 enhancer ..............................................................................157
3.2 Discussion and Conclusions..........................................................................158
3.2.1 Response elements............................................................................................158
3.2.2 Upstream initiated transcripts ........................................................................162
3.2.3 Magnitude of repression ...............................................................................163
3.2.4 Elements required for minimal levels of transcription........................................164
3.2.5 Sequence homologies.....................................................................................166
3.2.6 Promoter usage..............................................................................................167
3.2.7 Position effects................................................................................................170
3.2.8 Future prospects..............................................................................................171

Chapter 4 Protein-DNA complexes proximal to the P2 promoter

4.1 Introduction ......................................................................................................239
4.2 Results ...............................................................................................................241
4.2.1 Protein extracts ..............................................................................................241
4.2.2 Complexes detected using salmon sperm DNA as carrier..............................242
4.2.3 Complexes 1 and 2 are probably E2F ................................243
4.2.4 Response of E2F to overexpression of v-myc ..................244
4.2.5 Complexes detected using poly d(IC)poly d(IC)  
as carrier .......................................................................................244
4.2.6 Comparison with Me1a1 ....................................................247
4.2.7 Comparison of control and v-myc Rat-1  
extracls ..........................................................................................249
4.3 Conclusions ........................................................................................249
   4.3.1 E2F...........................................................................................250
   4.3.2 B and β....................................................................................251

Chapter 5 The MAC II complex ..........................................................280
  5.1 Introduction ..................................................................................281
  5.2 Results ..................................................................................................282
     5.2.1 Effect of antibodies on specific complexes ......................282
     5.2.2 Relationship of MAC II complex, myc protein  
       levels and cell type ..................................................................283
     5.2.3 Peculiar probe characteristics .............................................285
     5.2.4 Intrinsic and extrinsic factors ............................................287
     5.2.5 Thermolability of MAC II binding capacity of  
       probe ..........................................................................................288
     5.2.6 Deletion mapping of MAC II binding site ......................291
     5.2.7 Competition mapping of MAC II binding site ................292
     5.2.8 Thermostability of MAC II in extracts .............................293
  5.3 Discussion and Conclusions ..............................................................294
     5.3.1 Presence of myc protein in MAC II ..................................294
     5.3.2 Nature of the MAC II binding structure .........................298
     5.3.3 MAC II binding site .............................................................302
     5.3.4 Relationship to negative regulation of myc  
       transcription ..............................................................................303

Chapter 6 The MAC I complex ..........................................................327
  6.1 Introduction ..................................................................................328
  6.2 Results ..........................................................................................328
     6.2.1 Effect of antibodies ...........................................................328
     6.2.2 Strand specificity ...............................................................329
     6.2.3 Antibody titration .................................................................330
     6.2.4 Effects of different carriers .................................................331
6.2.5 Relationship to myc protein levels and cell type specificity .................................................................331
6.2.6 Competition mapping of binding site ......................................................................................................333
6.2.7 Deletion mapping of binding site ...........................................................................................................334
6.2.8 Supershift competitions ..........................................................................................................................336
6.2.9 Thermostability of MAC I in extracts ......................................................................................................337

6.3 Discussion and Conclusions ......................................................................................................................337
  6.3.1 Does MAC I contain myc protein? .........................................................................................................338
  6.3.2 Stoichiometry of MAC I .........................................................................................................................339
  6.3.3 MAC I binding site .................................................................................................................................341
  6.3.4 Other sequence specific single strand binding proteins .........................................................................343
  6.3.5 Relationship of MAC I and myc autosuppression .................................................................................344

Chapter 7 Perspectives ........................................................................................................................................371
  7.1 Summary of conclusions ..........................................................................................................................372
  7.2 Mechanisms of myc suppression ............................................................................................................374
  7.3 Possible functions for MAC I and MAC II ...............................................................................................376
  7.4 Future prospects ........................................................................................................................................377

References ..........................................................................................................................................................381
Figures

I.1: Domains of human c-myc protein..................................................................57

I.2: DNA binding domains of proteins with homology to c-myc......................59

I.3: Sequence of the human c-myc gene..............................................................118

II.1: Schematic diagrams of human c-myc (genomic), β-interferon
and CAT indicator constructs...........................................................................174

II.2: Schematic diagrams of retroviral vector VM and LJ
proviruses.........................................................................................................176

III.1: Response of transfected human c-myc genomic constructs to
exogenous v-myc............................................................................................178

III.2: Response of transfected human c-myc promoter fusion
constructs and endogenous rat c-myc to exogenous v-myc.........................180

III.3: Mapping 5' ends of MycCAT and MycIFN fusion transcripts....................182

III.4: Schematic diagrams of plasmids pSV2CAT and
pUCSVterLTRIFN..........................................................................................184

III.5: Response of transfected pSV2CAT and pUCSVterLTRIFN and
endogenous rat c-myc to exogenous v-myc....................................................186

III.6: Response of endogenous Rat c-myc and a transfected human c-
mvc/CAT fusion construct to the presence of v-myc under
various serum and cell growth conditions......................................................188

III.7: Schematic diagrams of human c-myc 5' deletion series,
BSKS+(-353:+119/+208)IFN............................................................................190

III.8: Response of pBSKS+(-353:-91/+208)IFN constructs to v-myc................192

III.9: Response of pBSKS+(91:+119/+208)IFN constructs to v-myc..................194

III.10: Response of pBSKS+(91:+119/+208)IFN constructs to v-myc
(reduced exposure of Fig. III.11)....................................................................196

III.11: Response of endogenous rat c-myc and α-globin indicator
cotransfected with pBSKS+(91:+119/+208)IFN deletion
series to v-myc..................................................................................................198

III.12: Schematic diagrams of human c-myc 5' deletion series,
pUC(-88:+97/+208)IFN.................................................................................200

III.13: Response of pUC(-88:+97/+208)IFN constructs to v-myc........................202

III.14: Response of upstream initiated transcripts to v-myc and
serum..................................................................................................................204

III.15: Schematic diagrams of pUC and pGC1 myc:IFN fusion
constructs..........................................................................................................206

III.16: Response of pUC and pGC1 myc:IFN fusion constructs to
v-myc................................................................................................................208
III.19: Response of linear -88/+208IFN fragments to v-myc.............................210
III.20: Suppression of upstream non-specific transcript levels by
addition of SV40 transcription termination signals...............................212
III.21: Response of clonal pUCSVter-88/+208IFN transfectants to
v-myc.......................................................................................................214
III.22: Analysis of endogenous rat myc and p110v-gag-myc mRNA
levels in clonal pUCSVter-88/+208IFN transfectants.............................216
III.23: Reduction of transfected myc mRNA levels over time....................218
III.24: Schematic diagrams of human c-myc 5' deletion series,
pUCSVter(-88:+150/+208)IFN...................................................................220
III.25: Response of 5' deletion series, pUCSVter(-88:+150)IFN to
v-myc.......................................................................................................222
III.26: Schematic diagrams of human c-myc linker scan series,
pUCSVter(ls1-ls5:+150)IFN and pUCSVter(ls1-ls5:Pst)IFN.....................224
III.27: Response of linker scan series pUCSVter(ls1-ls5:+150)IFN to
v-myc.......................................................................................................226
III.28: Response of linker scan series pUCSVter(ls1-ls5:Pst)IFN to
v-myc.......................................................................................................228
III.29: Schematic diagrams of pUCSVter, pUCSVE,
pUCSVESVter(-88/+208)IFN and pJ123hcmyc constructs.........................230
III.30: Response of pUCSVE, pUCSVESVter(-88/+208)IFN and
pJ123hcmyc to v-myc..............................................................................232
IV.1: Detection of myc protein in Rat-1 whole cell extracts..........................254
IV.2: Complexes detected using salmon sperm DNA as carrier..................256
IV.3: Linker scan competitions vs. -66/+49 salmon sperm DNA
complexes...............................................................................................258
IV.4: E2F oligonucleotide competitions vs. -66/+49 salmon sperm
DNA complexes.....................................................................................260
IV.5: Comparison of Rat-1 LJ and VM complexes using -66/+49, ls1
and ls2 probes with salmon sperm DNA carrier....................................262
IV.6: Comparison of Rat-1 LJ and VM complexes using ls3, ls4 and
ls5 probes with salmon sperm DNA carrier..........................................264
IV.7: Complexes detected using poly d(IC)•poly d(IC) as carrier...............266
IV.8: Linker scan competitions vs. -63/+49 poly d(IC)•poly d(IC)
complexes............................................................................................268
IV.9: E2F oligonucleotide competitions vs. -66/+49 poly d(IC)•poly
d(IC) complexes....................................................................................270
IV. 10: 'Me1a1' oligonucleotides as probes and competitors vs. poly 
d(IC)*poly d(IC) complexes ................................................................. 272

IV. 11: Comparison of Rat-1 LJ and VM complexes using -66/+49, ls1 
and ls2 probes with poly d(IC)*poly d(IC) carrier ......................... 274

IV. 12: Comparison of Rat-1 LJ and VM complexes using ls3, ls4 and 
ls5 probes with poly d(IC)*poly d(IC) carrier ..................................... 276

V. 1: Supershifting Rat-1 complexes with anti-myc antibodies ...... 305

V. 2: Cell type distribution and response to myc levels of MAC II ... 307

V. 3: Supershifting HeLa complexes with anti-myc antibodies ...... 309

V. 4: Variations in MAC II binding capacity of different probe 
preparations ......................................................................................... 311

V. 5: Loss of binding upon melting MAC II probes ......................... 313

V. 6: Thermal stability of MAC II binding capacity of probes .......... 315

V. 7: Deletion analysis of MAC II binding (1) ...................................... 317

V. 8: Deletion analysis of MAC II binding (2) ................................. 319

V. 9: Competition mapping of MAC II binding ................................. 321

V. 10: Thermal stability of MAC II in extracts with or without probe .... 323

VI. 1: Supershifting single strand binding complex with anti-myc 
antibody ............................................................................................... 345

VI. 2: Self competition of MAC I binding: single and double strand 
competitors .......................................................................................... 347

VI. 3: Titration of antibody against MAC I ........................................ 349

VI. 4: Carrier specificity of MAC I binding .......................................... 351

VI. 5: Presence of MAC I in different cell types and in response to 
myc levels ............................................................................................ 353

VI. 6: Competition mapping of MAC I binding: antisense strand 
competitors .......................................................................................... 355

VI. 7: Competition mapping of MAC I binding: sense strand and 
unrelated competitors ........................................................................ 357

VI. 8: Deletion mapping of MAC I binding via supershifting (short 
exposure) ......................................................................................... 359

VI. 9: Deletion mapping of MAC I binding via supershifting (long 
exposure) ......................................................................................... 361

VI. 10: Mapping of MAC I binding: Competing supershifted 
complexes .............................................................................................. 363

VI. 11: Thermal stability of MAC I binding activity with or without 
probe .................................................................................................... 367

VI. 12: Secondary structure predictions of MAC I binding site ......... 369
VII.1: Regulatory elements and protein factors at the c-myc P2 promoter: Summary
Tables

III.1: Summary of promoter proximal construct suppression results .......... 234
III.2: Sequence homologies amongst mammalian c-myc P2 promoters ........................................................................................................ 236
IV.1: Summary of complex B and β binding and competition data .......... 278
V.1: MAC II binding and competition: Summary .................................... 325
VI.1: Deletion and competition mapping of MAC I binding:
       Summary ............................................................................................ 365
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>βIFN</td>
<td>β-interferon</td>
</tr>
<tr>
<td>DMEM</td>
<td>Dulbecco modified Eagle's medium</td>
</tr>
<tr>
<td>DMS</td>
<td>Dimethyl sulphate</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid (disodium salt)</td>
</tr>
<tr>
<td>EGTA</td>
<td>Ethylene glycol bis(β-aminoethyl ether) N,N,N',N'-tetraacetic acid</td>
</tr>
<tr>
<td>EtBr</td>
<td>Ethidium bromide</td>
</tr>
<tr>
<td>FCS</td>
<td>Fetal calf serum</td>
</tr>
<tr>
<td>FeLV</td>
<td>Feline leukemia virus</td>
</tr>
<tr>
<td>FSV</td>
<td>Feline sarcoma virus</td>
</tr>
<tr>
<td>GAPDH</td>
<td>Glyceraldehyde-3-phosphate dehydrogenase</td>
</tr>
<tr>
<td>GTC</td>
<td>Guanidinium isothiocyanate</td>
</tr>
<tr>
<td>LTR</td>
<td>Long terminal repeat</td>
</tr>
<tr>
<td>Mo-MuLV</td>
<td>Moloney murine leukemia virus</td>
</tr>
<tr>
<td>MMTV</td>
<td>Mouse mammary tumor virus</td>
</tr>
<tr>
<td>NBCS</td>
<td>New born calf serum</td>
</tr>
<tr>
<td>NP40</td>
<td>Nonionic detergent P-40</td>
</tr>
<tr>
<td>(d)NTP</td>
<td>(2'-deoxy)nucleotide 5'-triphosphate</td>
</tr>
<tr>
<td>(d)ATP</td>
<td>(2'-deoxy)adenosine 5'-triphosphate</td>
</tr>
<tr>
<td>(d)CTP</td>
<td>(2'-deoxy)cytosine 5'-triphosphate</td>
</tr>
<tr>
<td>(d)GTP</td>
<td>(2'-deoxy)guanosine 5'-triphosphate</td>
</tr>
<tr>
<td>dTTP</td>
<td>2'-deoxy thymidine 5'-triphosphate</td>
</tr>
<tr>
<td>UTP</td>
<td>Uridine 5'-triphosphate</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenyl-methyl-sulfonyl-fluoride</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>Sarkosyl</td>
<td>Sodium lauryl sarcosinate</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium dodecyl sulfate</td>
</tr>
<tr>
<td>SSC</td>
<td>Standard sodium citrate</td>
</tr>
<tr>
<td>STI</td>
<td>Soybean trypsin inhibitor</td>
</tr>
<tr>
<td>TPCK</td>
<td>n-Tosyl-L-phenylalanine chloromethyl ketone</td>
</tr>
<tr>
<td>g</td>
<td>gram</td>
</tr>
<tr>
<td>m(g)</td>
<td>milli(gram)</td>
</tr>
<tr>
<td>Symbol</td>
<td>Unit</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>( \mu(g) )</td>
<td>micro(gram)</td>
</tr>
<tr>
<td>( n(g) )</td>
<td>nano(gram)</td>
</tr>
<tr>
<td>( p(g) )</td>
<td>pico(gram)</td>
</tr>
<tr>
<td>( M )</td>
<td>molar</td>
</tr>
<tr>
<td>( cm )</td>
<td>centimeter</td>
</tr>
<tr>
<td>( bp )</td>
<td>base pair</td>
</tr>
<tr>
<td>( kb )</td>
<td>kilobase</td>
</tr>
<tr>
<td>( ^\circ C )</td>
<td>degrees Centigrade</td>
</tr>
<tr>
<td>( RT )</td>
<td>room temperature</td>
</tr>
<tr>
<td>( d )</td>
<td>day(s)</td>
</tr>
<tr>
<td>( hr )</td>
<td>hour(s)</td>
</tr>
<tr>
<td>( min )</td>
<td>minute(s)</td>
</tr>
<tr>
<td>( sec )</td>
<td>second(s)</td>
</tr>
</tbody>
</table>
Chapter 1 Introduction
1.1 Cancer and oncogenes

Within multicellular organisms proliferation of most cells is restricted both spatially and temporally, and complex interactive networks of inter- and intracellular communications pathways have evolved to maintain the control of cellular proliferation within the physiological needs of each organism. The disease of cancer occurs following the progressive loss of these controls, and consists of cells proliferating continuously and autonomously. Cancer cells display a number of growth characteristics, including independence of the requirement for external growth signals, which collectively define the transformed or oncogenic cellular state.

As the behavior of a cell is ultimately controlled by the genes which it expresses, it has long seemed probable that genetic mutations might be involved in cancer development. A genetic basis for cancer was first suggested from work on tumours in chickens, in which a virus (Rous sarcoma virus retrovirus) was isolated on the basis of its ability to induce transmissible sarcomas in infected birds (Rous, 1911). A combination of genetic and molecular biological analyses led to the demonstration that the transforming activity of this virus is due to the presence of a particular 'oncogene', v-src (Martin, 1970, Stehelin, et al, 1976, Vogt, 1971). A further advance in the understanding of cellular tumorigenesis occurred with the discovery that a counterpart for the v-src gene, c-src, exists in normal cellular DNA (Stehelin, et al, 1976), which led to the notion that oncogenes of acutely transforming retroviruses are transduced versions of cellular proto-oncogenes. More than two dozen retroviral oncogenes have been identified, each of which has a normal cellular counterpart (Reddy, et al, 1988). DNA tumour viruses such as the papova, papilloma and adenoviruses, were also found to contain oncogenes. Unlike the retroviral oncogenes, direct cellular counterparts of the oncogenes in DNA tumour viruses have not been identified.
While the discovery of oncogenes associated with tumour viruses was a major step in the elucidation of tumour induction, it was clear that viral oncogenes could not be the major cause of tumours in humans, as few human cancers have been associated with viral infection. Indeed, it was found that genomic DNA from chemically transformed cells contained a dominant transforming activity in in vitro cell transfection assays (Shih, et al, 1979). The transforming activities from several human tumours were subsequently isolated using a similar procedure (Murray, et al, 1981, Peruch, et al, 1981, Shih, et al, 1981) and identified as the cellular counterparts of two previously discovered retroviral oncogenes, v-Ha-ras and v-Ki-ras (Der, et al, 1982, Parada, et al, 1982, Santos, et al, 1982). The transforming capacity of these genes was later found to be due to oncogenic activation by a somatic mutation (Tabin, et al, 1982). Numerous other cellular oncogenes have since been identified, some of which have transduced viral counterparts, eg myc (Hayward, et al, 1981), and others of which do not, eg neu (Bargmann, et al, 1986, Bishop, 1987 for review).

Activation of cellular oncogenes has been documented to occur through many genetic mechanisms. These include transduction by a retrovirus, retroviral insertional mutagenesis, chromosomal translocation, point mutation, gene amplification or chromosomal deletion. The end result of these changes is generally either an alteration in the expression of the gene or a structural change of either regulatory or catalytic domains of the protein. Due to the biases of the original assay systems, oncogenes which act in a dominant fashion were the first detected. Predominantly through the use of reverse genetic techniques (Orkin, 1986), it has since been possible to identify various genes at loci which are homozygously mutated in certain tumors (Friend, et al, 1986, Lee, et al, 1987). These genes are known as anti-oncogenes or tumour-suppressor genes due to their postulated antiproliferative functions. Anti-
oncogenes can be altered by similar kinds of mutations as those which affect the dominant oncogenes, as well as by deletion from the cellular genome, leading to inactivation of the normal gene (Levine, 1990, Weinberg, 1989).

Once it became apparent that most of the more than sixty oncogenes identified to date are mutated versions of normal cellular genes, it became necessary to identify their normal functions and determine how the control of these functions is subverted in the oncogenic process. Noteworthy progress has been made over the past decade in elucidating the functions and activities of both proto- and activated oncogenes. The overwhelming conclusion is that these proteins perform critical tasks in regulating cell growth, inter- and intracellular communication, and many developmental processes. In particular, most of the oncoproteins for which functions or activities are known are implicated in various signal transduction pathways.

The protein products of oncogenes have been identified at various locations throughout cells and in extracellular spaces, and include extracellular growth factors, extracellular matrix associated growth factors, transmembrane tyrosine kinase growth factor receptors, guanine nucleotide binding (G) protein-linked growth factor receptors, membrane-associated nonreceptor tyrosine kinase signal transducers, membrane-associated G protein signal transducers and cytoplasmic serine/threonine kinases. Nuclear oncoproteins appear to be primarily involved in the regulation of transcription and/or progression through the cell cycle (Bishop, 1987, Reddy, et al, 1988, Weinberg, 1989, for reviews).

1.2 Oncogene cooperation

On the basis of genetic and epidemiological data it has long been apparent that cancer is a multistep process, requiring multiple rate-limiting steps during progression to malignancy (Knudson, 1986). With
the discovery of oncogenes this requirement became relatively easy to rationalise as discrete genetic lesions, each involving the mutation of a single gene involved in regulating cell growth. Using in vitro model systems, however, it was found that the introduction of a single oncogene such as ras was sufficient to transform an immortal cell line to a tumorigenic phenotype. This apparent paradox of 'single hit' in vitro transformation kinetics, as opposed to multistep in vivo kinetics, was partially reconciled when it was found that one activated oncogene is insufficient to impart a tumorigenic phenotype to primary cell cultures (Land, et al, 1983, Newbold and Overell, 1983, Rassoulzadegan, et al, 1982, Ruley, 1983). At least two oncogenes are required to transform these cells in most situations, and several pairs of 'cooperating' oncogenes have been identified (Land, et al, 1986, Land, et al, 1983, Ruley, 1983). Oncogenes which will function in this assay appear to fall into two classes, one member of each class being required for tumorigenic transformation. The first class consists of oncogenes whose gene products are located predominantly in the nucleus, such as the myc family (Birrer, et al, 1988, Land, et al, 1983, Ruley, 1983, Schwab, et al, 1985, Yancopoulos, et al, 1985), c-jun (Schütte, et al, 1989), adenovirus E1a (Ruley, 1983), and the Large T oncogenes of polyoma and SV40 viruses (Land, et al, 1983, Rassoulzadegan, et al, 1982). The second class contains proteins located predominantly in the cytoplasm or at the cell membrane, and includes members of the ras family (Land, et al, 1983), raf (Cleveland, et al, 1986), bcr-abl (Lugo and Witte, 1989) and polyoma virus middle T (Rassoulzadegan, et al, 1982).


1.3 Roles of nuclear oncogenes

Studies on nuclear oncogenes using the primary rodent fibroblast fibroblast assay revealed an additional activity of these proteins. While these cells normally have a limited lifespan in culture, constitutive expression of a nuclear oncogene was found in some cases to lead to the establishment of permanent cell lines, which do not display a transformed phenotype (Mougneau, et al, 1984, Ruley, et al, 1984). This activity has implicated the nuclear oncogenes as key components in the regulatory pathways controlling cellular proliferative potential.

Introduction of some cytoplasmic oncogenes, such as ras or raf, in the absence of any other oncogenes, into primary embryo fibroblast cultures (H Land, personal communication), Schwann cells (Ridley, et al, 1988) and the rat embryo fibroblast-derived REF-52 cell line (Kohl and Ruley, 1987) was found to inhibit their proliferation. This observation contrasts with the tumourigenic transformation observed when these genes are introduced into many established cell lines. Schwann cells and REF-52 cells expressing activated ras oncogenes arrest at both the G1 and G2/M phases of the cell cycle, and this arrest is prevented by the co-introduction of a nuclear oncogene, such as myc, E1a or SV40 LT (Kohl and Ruley, 1987, Ridley, et al, 1988). In addition to the prevention of a growth arrest, the combination of the two oncogenes leads to cellular transformation.
Nuclear oncogenes are thus capable of altering the response of a cell to an extranuclear signal, such as that emanating from the ras protein, from growth arrest to uncontrolled proliferation, a function which may be critical for tumour progression.

1.4 c-myc: Overview

The nuclear oncogene, c-myc, has been implicated in the genesis of numerous vertebrate tumours. It is the nuclear oncogene of cellular origin which is most potent in cooperative transformation or primary rodent fibroblast immortalisation assays. C-myc was first identified as the cellular homologue of the v-myc oncogene (Roussel, et al, 1979, Sheiness and Bishop, 1979), the transforming gene of the MC29 strain of avian myelocytomatosis virus (myc) (Ivanov, et al, 1964). MC29 is a replication defective retrovirus which induces primarily non-lymphoid leukemias in infected birds (Mladenov, et al, 1967). c-myc is a member of a family of related genes, including N-, L-, B- and S-myc (Ingvarsson, et al, 1988, Kohl, et al, 1983, Nau, et al, 1985, Schwab, et al, 1983, Sugiyama, et al, 1989), some of which also have oncogenic potential (Birrer, et al, 1988, Schwab, 1988, Yancopoulos, et al, 1985). The c-myc gene and its product have been highly conserved throughout vertebrate evolution, homologues having been cloned from chordate species as diverse as humans and fish. These traits and the close relationship between c-myc expression and the proliferative potential of cells have provided a strong impetus to view c-myc as an important regulator of cell growth. The c-myc gene, its oncogenic activation, biological activity, regulation and protein product have all been extensively investigated over the past decade, and are summarised below.

1.4.1 Gene structure

The c-myc genes of six vertebrate species have been cloned and sequenced. These include humans (Bernard, et al, 1983, Colby, et al, 1983,
rats (Hayashi, et al, 1987), birds (Nottenburg and Varmus, 1986,
al, 1989) and fish (van Beneden, et al, 1986). All of these genes contain a
similar three exon organisation, with the major open reading frame
contained in the second and third exons. The coding exons are the most
highly conserved, with 70% to 90% homology within the open reading
frames across species. The non-coding sequences are also well conserved
between human and rodent c-myc, with 70% homology of exon I (Hayashi,
et al, 1987), but are not well conserved amongst the more divergent
Considering the high degree of conservation of the known c-myc genes, it
is quite striking that no clear homologues have been identified in other
phyla. The related N-myc and L-myc genes are also found solely in
vertebrates, and share the three exon structure of c-myc, their major open
reading frames also being contained within their second and third exons.
The first exons of these genes have little homology with those of the c-myc

The c-myc genes often contain multiple transcription start sites.
Transcripts are initiated primarily from two promoters, P1 and P2, which
are separated by distances of the order of 160 bp in the mammalian genes
1983), and less in the c-myc genes of other species (Nottenburg and
Varmus, 1986, Vriz, et al, 1989). In most normal cells the downstream P2
cap site is preferred, accounting for up to 90% of the c-myc mRNA
approximately 600 bp upstream of P1, has been detected in certain human
myeloid lineages, and accounts for a small proportion of the total c-myc
mRNA in those cells (Bentley and Groudine, 1986b). Transcripts are
polyadenylated at one of two polyadenylation sites, which are usually located 150 to 200 bp apart. The second of the two sites is predominantly utilised, accounting for up to 85% of transcripts (Hayashi, et al, 1987, Laird-Offringa, et al, 1989). The combination of promoter and polyadenylation site choice results in approximately 2.2 kb P2-initiated and 2.4 kb P1-initiated mature transcripts in most mammalian species (Battey, et al, 1983, Bernard, et al, 1983). The mRNA is usually highly unstable, with a half-life on the order of 10-30 minutes depending on the cell type (Dani, et al, 1984).

1.4.2 Protein


The c-myc gene product is a nuclear phosphoprotein (Hann and Eisenman, 1984, Persson and Leder, 1984, Ramsay, et al, 1984). It is phosphorylated on serine and threonine residues, with the major phosphorylation sites located around the region corresponding to the exon 2/3 border and near the carboxyl terminus (Luscher, et al, 1989). The protein has a very short half life, usually approximately 20-30 minutes (Hann and Eisenman, 1984, Ramsay, et al, 1984), and is degraded by an energy dependent process (Luscher and Eisenman, 1988). C-myc protein is
distributed throughout the interphase nucleus, although it is excluded from the nucleoli, and in mitotic cells is not associated with condensed chromatin (Wingvist, et al, 1984). The protein is not abundant, with only 500-2000 molecules expressed in a proliferating normal cell. At the peak of expression following mitogenic stimulation approximately 5000-15,000 molecules can be expressed. Levels are often much higher in transformed cells, typically elevated by ten to one hundred fold (Evan, et al, 1988, Moore, et al, 1987).

The c-myc proteins are very highly conserved, displaying 60-90% amino acid homology across species throughout the entire open reading frame (Figure I.1). Certain regions of the protein show an even greater extent of similarity, including the regions in human c-myc which encompass amino acids 105-143 and 371-435, both of which are crucial for its transforming activity (Stone, et al, 1987). These sequences are also conserved in the N- (Kohl, et al, 1986, Stanton, et al, 1986), L- (DePinho, et al, 1987b) and S-myc (Okuda, et al, 1989) proteins. The B-myc gene, which is only distantly related to c-myc, has homologies only to the first c-myc coding exon, including amino acids 105-143 (Asker, et al, 1989). In addition to the regions of the protein for which some role has been identified, there are very strongly conserved regions to which no functional attributes have yet been assigned (See Figure I.1).

1.4.3 Homology with transcription factors

The approximately 80 amino acids located at the carboxyl terminus of the c-myc protein contain homologies with three structural motifs found in a large and rapidly growing superfamily of proteins (Penn, et al, 1990c for review, Prendergast and Ziff, 1989). Some of these proteins have been demonstrated to be sequence-specific DNA-binding proteins involved in transcriptional regulation (Landschulz, et al, 1988, Murre, et al, 1989a). The motifs consist of basic, helix-loop-helix (HLH) and leucine zipper (LZ)

The basic-LZ proteins include members of the C/EBP, fos/jun and ATF/CREB transcription factor families; basic-HLH proteins include numerous differentiation factors of vertebrate and insect origin (eg myoD1, myf-5, achaete-scute complex members and daughterless (Alonso and Cabrera, 1988, Braun, et al, 1989, Caudy, et al, 1988, Davis, et al, 1987, Hopwood, et al, 1989, Lin, et al, 1989, Villares and Cabrera, 1987)), enhancer binding factors E12 and E47 (Murre, et al, 1989a), and a S. cerevisiae centromere binding protein, CBF1 (Cai and Davis, 1990); basic-HLH-LZ proteins include the S. cerevisiae phosphatase regulon protein pho4 (Ogawa and Oshima, 1990) and immunoglobulin heavy chain enhancer binding protein BTF3 (Beckmann, et al, 1990). An alignment of some of the members of these families with the myc proteins is shown in Figure 1.2.

The degree of homology between the myc proteins and this large group of genes, many of which have been demonstrated to bind specific DNA sequences as dimers, strongly suggests that myc may behave similarly. Tetrramerisation of bacterially expressed c-myc protein has indeed been reported to occur through the HLH and LZ domains (Dang, et al, 1989a). However, c-myc does not appear to multimerise when translated in reticulocyte lysates or when isolated from growing cells, even under conditions in which the other proteins will readily multimerize (Kouzarides and Ziff, 1988, Murre, et al, 1989b, T. Littlewood and G. Evan,

1.5 Activation of the myc oncogene

The c-myc proto-oncogene has been found to be activated via numerous mechanisms in a wide variety of human and animal tumours. These mechanisms include retroviral transduction, retroviral insertion, chromosomal translocation and gene amplification. As will be seen, the common theme for activation appears to be the deregulation of normal c-myc expression.

1.5.1 Retroviral transduction

Acutely transforming retroviruses of the MC29 group (MC29, OK10, MH2 and CM II) induce the rapid growth of fatal polyclonal tumours in infected birds within a period of days or weeks following infection. The tumours tend to consist of aggressively proliferating cells of the non-lymphoid, primarily monocytic, hematopoietic lineages, though non-hematopoietic carcinomas and sarcomas have also been described (Beard, et al, 1976, Beard, et al, 1975, Chabot, et al, 1970, Mladenov, et al, 1967). All four members of the MC29 group of viruses contain transduced myc genes (Roussel, et al, 1979) and their genomes have been cloned and sequenced, and the oncoproteins encoded therein extensively characterised (reviewed
MC29, CM II, MH2 and OK10 (Roussel, et al, 1979, Sheiness, et al, 1978), all express different but highly related v-myc proteins which are distinct from the avian c-myc protein. Through the transduction process portions of the viral group-specific antigen (gag) genes have become joined in frame to the chicken myc sequences, resulting in the production of v-gag-myc fusion proteins. The specific nature of the fusion is unique to each virus: MC29 thus expresses a p110 Δgag-myc protein; CM II a p90 Δgag-myc protein; MH2 a p57 δgag-myc protein and OK10 p200 gag-Δpol-myc and p57 δgag-myc proteins. The Δgag proteins are encoded by genomic retroviral mRNAs while the δgag proteins are encoded by spliced sub-genomic mRNAs. The v-myc proteins, like their cellular homologues, are all relatively short-lived nuclear phosphoproteins which possess non-specific DNA binding activity in vitro. The viral proteins differ from their cellular counterparts by a number of point mutations. While some of these mutations appear to potentiate transformation of certain avian haematopoietic lineages, they are not absolutely required for activation of the oncogene (Frykberg, et al, 1987, Symonds, et al, 1989).

The MH2 virus is unique in that it has transduced another oncogene, mht, in addition to myc (Jansen, et al, 1984, Kan, et al, 1984). The mht oncogene is the avian equivalent of the c-raf gene, a cytoplasmic serine/threonine kinase which is implicated in signal transduction (Moelling, et al, 1984). The presence of these two genes in the same virus suggests that they can cooperate in transforming their target cells, a postulate which has been confirmed experimentally (Graf, et al, 1986).

genes of three feline viruses, FeLV-CT4, FeLV-LC and FTT have been sequenced. Analysis of the putative v-myc coding sequences indicates that neither mutation of the myc sequences nor fusion with retroviral proteins is obligatory for activation (Braun, et al, 1985, Doggett, et al, 1989, Stewart, et al, 1986).

1.5.2 Retroviral insertion

Integration of retroviral proviruses at the c-myc locus has been well documented (Hayward, et al, 1981, Payne, et al, 1982, Varmus, 1984 for review). This mechanism appears to be a common cause of the induction of tumors by avian, rodent and feline non-acutely transforming viruses. Proviral insertion can occur either upstream or downstream of the c-myc locus, and in either orientation relative to the c-myc gene. In addition, insertion in either orientation within the first intron is often observed. When the provirus is inserted 5' of the major open reading frame, and in the same transcriptional orientation as the c-myc gene, transcription starts at the strong viral promoter. In instances of insertion in the opposite orientation or downstream of the c-myc locus, it is believed that the viral sequences act either as transcriptional enhancer elements or possibly to disrupt or displace normal c-myc cis controlling elements.

1.5.3 Chromosomal translocation

been documented. It is believed that the translocations occur through aberrant chain switching during B-cell ontogeny. The result is usually the positioning of immunoglobulin sequences upstream of the major c-myc open reading frame, although translocations to points 3' of the myc locus have also been observed. Breakpoints upstream of the c-myc locus have been mapped to both the 5' flanking sequences and within the non-coding first exon or intron. While most of the breakpoints occur within a few kb of the c-myc transcription unit, some have been identified several tens of kilobases from this locus. Transcription in those instances which delete the normal c-myc promoters is usually initiated from cryptic start sites within the c-myc first intron. The immunoglobulin sequences which are juxtaposed to the c-myc gene sometimes contain identified immunoglobulin enhancers, although they often do not. In some cases point mutations within the first or second exon have been described (Rabbitts, et al, 1983), which, while they may contribute to the activation process, do not appear to be obligatory (Lee, et al, 1985). The exact molecular basis of the regulatory changes which lead to the activation of the c-myc oncogene remain obscure, although they may either involve activation through the juxtaposition of specific immunoglobulin control elements or through the elimination of c-myc regulatory sequences (Spencer and Groudine, 1990 for review).

1.5.4 Gene amplification

Another mechanism whereby the myc gene family is implicated in carcinogenesis is through amplification of myc-specific sequences. The c-myc locus is amplified in a number of human tumors (Collins and Groudine, 1982, Dalla Favera, et al, 1982c), as are the N- (Kohl, et al, 1983, Schwab, et al, 1983) and L-myc (Nau, et al, 1985) genes. N-myc is amplified predominantly in neuroblastomas, and L-myc in small cell lung carcinomas. In both of these cases, the degree of amplification serves as a
Amplification of the myc loci may lead to higher levels of expression of
the myc proteins without altering the regulability of the genes, as some
regulatory functions remain active in cell lines which contain amplified c-
myc. However, as the normal modes of myc regulation remain
incompletely understood, this hypothesis is very difficult to confirm.

1.5.5 Regulatory mutations are sufficient for oncogenic activation

The common theme running through all of the mechanisms whereby
the c-myc gene is activated to its oncogenic potential in vivo is one of
deregulated expression. It is a striking fact that in each case of in vivo
activation of c-myc, while gross abnormalities of the DNA surrounding
the c-myc coding sequences may occur, there are very few documented
instances of gross alterations to the c-myc portion of the expressed protein
(Varley, et al, 1987). While in some instances mutations of the encoded c-
myc protein have been identified, eg (Rabbitts, et al, 1983), none is seen to
be consistently generated by all of the mechanisms of myc activation. This
has suggested that a loss of myc regulation is a necessary event in its
activation, while mutagenesis of the protein is probably neither necessary
nor sufficient.

This hypothesis has been confirmed through a number of studies both
in vitro and in vivo. The major in vitro assays which are used for
studying myc activation are its ability to extend the proliferative potential
of primary cell cultures, and its ability to participate in the transformation
of these cells in the presence of an additional 'cooperating' oncogene.
Using such systems it has been demonstrated that the addition of a cis
acting transcriptional enhancer sequence upstream of a genomic c-myc
fragment is sufficient to induce oncogenic activation. However, the same
c-myc fragment in the absence of any additional control elements or a
genomic fragment containing point mutations found in vivo are both

Several studies have made use of transgenic mice expressing a wild type c-myc open reading frame under the control of a heterologous promoter (Adams, et al, 1985, Alexander, et al, 1987, Stewart, et al, 1984b). These studies have clearly demonstrated that altered control is sufficient to predispose these animals to cancers, and furthermore, that the cell lineages of the tumors which develop are predominantly determined by the tissue specificity of the cis-acting control sequences used. In addition, these tumors are clonal and arise several weeks to months after birth, implying that additional rate-limiting steps are required for complete transformation.

In many tumors which contain activated c-myc, levels of c-myc protein are much higher than those seen in normal tissue (Moore, et al, 1987). This suggests that elevation of c-myc expression is sufficient to activate its transforming potential. However, myc levels in tumors are not always elevated compared to normal proliferating cells (e.g., in certain tumors with c-myc translocations) (Keath, et al, 1984b, Taub, et al, 1984). Thus temporally inappropriate, as opposed to merely excessive, expression of c-myc may be sufficient for activation.

1.6 Biological functions of c-myc

Clues to the normal biological function of c-myc protein have been gleaned from studies in the systems described above. The phenotypes of tumors which contain activated myc as well as the introduction of constitutively expressed c- and v-myc both in vivo and in vitro have all provided valuable information. Constitutive expression of c-myc appears to increase the proliferative potential of cells. In vivo this corresponds to hyperplasia within a distinct developmental compartment, and in vitro to phenotypes such as immortalisation, prevention of differentiation, and
reduction of growth factor dependence. While the specific timing of c-myc action in cellular growth and differentiation is not well understood, there is some evidence that myc functions may include regulation of various phases of the cell cycle, as well as entry into differentiation pathways.

1.6.1 Role in development and differentiation in vivo

C-myc expression has been studied in developing embryos and adult tissues of a number of species, including birds (Gonda, et al, 1982, Jaffredo, et al, 1989), toads (King, et al, 1986), rodents (Downs, et al, 1989, Schmid, et al, 1989, Stewart, et al, 1984a) and humans (Pfeifer-Ohlsson, et al, 1985). The distribution of c-myc transcripts is complex, involving a combination of spatial and temporal regulation. C-myc expression is detectable in many, but not all, proliferating and migratory tissues, is not generally found in differentiated organs and is sometimes seen to reduce during differentiation, although some expression in fully differentiated cells, avian post-mitotic neurons for example, has been observed (Jaffredo, et al, 1989).

Evidence for c-myc function in vivo comes from the study of c-myc associated tumors, transgenic mice and organ reconstitutions. C-myc associated tumors usually display a partially differentiated phenotype; transgenic mice expressing c-myc display expanded populations of cells in intermediate developmental compartments (Harris, et al, 1988, Schmidt, et al, 1988); and organ reconstitutions show hyperplastic growth of myc-infected tissues (Thompson, et al, 1989). All of these observations are consistent with constitutive myc expression either preventing terminal differentiation or stimulating cell growth by preventing complete withdrawal from the cell cycle.

1.6.2 Role in cellular differentiation in vitro

C-myc levels have also been studied in a number of cell lines which can be induced to differentiate in vitro. Myc levels are generally observed
to fall upon induction of differentiation, and remain low following terminal differentiation. In some instances there is a bi- or tri-phasic pattern to this reduction, with falls followed by the recovery of myc expression, with a subsequent drop to minimal levels (Endo and Nadal-Ginard, 1986, Lachman and Skoultchi, 1984, St-Arnaud, et al, 1988). Loss of c-myc expression upon differentiation has been observed in HL60 cells, a human promyelocytic leukemia line, induced along monocytic or granulocytic pathways (Reitsma, et al, 1983, Westin, et al, 1982); murine erythroleukemia (MEL) cells, induced along erythrocytic pathways (Lachman and Skoultchi, 1984); murine myeloid leukemia WEHI-3B cells induced into monocytes or granulocytes (Gonda and Metcalfe, 1984); F9 murine teratocarcinoma cells induced into parietal endoderm (Campisi, et al, 1984); 3T3-L1 preadipocytes induced to adipocytes (Freytag, 1988); and rat L6E9-B cells which are differentiable into myotubes (Endo and Nadal-Ginard, 1986). However, loss of c-myc upon differentiation is not always observed, as even five days after mouse keratinocytes have been induced to differentiate in vitro, they still express high levels of c-myc RNA, in spite of an observed 95% reduction in DNA synthesis (Dotto, et al, 1986).

Experimental evidence supports a role for c-myc in the regulation of the withdrawal from proliferation of a number of the differentiating cellular systems. Constitutive enforced c-myc expression prevents differentiation and withdrawal from the cell cycle in 3T3-L1 preadipocytes (Freytag, 1988), MEL cells (Coppola and Cole, 1986, Dmitrovsky, et al, 1986, Prochownik and Kukowsa, 1986), and myoblasts (Falcone, et al, 1985), although it does not prevent the commitment of MEL cells to the differentiation pathway (Coppola and Cole, 1986). Enforced removal of myc expression is sometimes sufficient to induce differentiation, as antisense c-myc will increase the kinetics of HL60 or MEL cell differentiation (Prochownik, et al, 1988, Yokoyama and Imamoto, 1987).
The effects c-myc expression on F9 teratocarcinoma cell differentiation are unclear, as differentiation and proliferation have been found to be both dependent on (Griep and DeLuca, 1986, Griep and Westphal, 1988) and independent of c-myc expression (Nishikura, et al, 1990). Permanent removal of c-myc is not always obligatory for maintenance of a terminally differentiated state, as c-myc expression in differentiated myotubes can be induced with no observable effect (Endo and Nadal-Ginard, 1986). A role for c-myc protein in preventing terminal differentiation by maintenance of a potentially proliferating state in many cell lines is thus strongly suggested.

While there is often an inverse correlation between c-myc expression and differentiation, and constitutive expression of c-myc can prevent terminal differentiation, it is difficult to completely dissociate the role of c-myc in differentiation from a role in maintaining proliferative potential. The observation that there is often a biphasic response of myc expression to inducers of differentiation suggests that myc expression may be required for fulfilment of the differentiation program. However, there is also frequently some cell growth following the recovery of myc levels after the initial differentiation stimulus (Lachman, et al, 1985). It may be that only transient removal of c-myc expression is sufficient for commitment to a differentiation pathway, while the reexpression of myc is necessary for the fulfilment of the additional cell divisions. Transient reductions in myc expression may also occur in cells such as keratinocytes, which differentiate even though long-term expression of c-myc appears to be maintained (Dotto, et al, 1986).

1.6.3 Mediator of mitogenic stimuli

C-myc appears to be involved in mediating extracellular mitogenic stimuli. Such a function is implied by the close correlation of c-myc expression and the presence of extracellular mitogens.
C-myc mRNA and protein levels are tightly regulated throughout the cell cycle, and appear to be closely coordinated, with c-myc protein levels following closely behind changes in c-myc mRNA (Hann, et al, 1985, Moore, et al, 1987, Rabbitts, et al, 1985b). In growing cell lines and cultures c-myc expression is usually detectable, while it is not in quiescent cell cultures (Campisi, et al, 1984, Kelly, et al, 1983). Following mitogenic stimulation of quiescent cells, there is a transient increase in the myc mRNA level which peaks after one to two hours (Campisi, et al, 1984, Kelly, et al, 1983). This stimulation is independent of novel protein synthesis, and its magnitude can be up to forty fold, depending on the cell line (Kelly, et al, 1983, Muller, et al, 1984). Levels decline thereafter, and if the cells are confluent, return to the initial values within 16 to 20 hours. If the cells are subconfluent, and therefore continue to proliferate, myc expression is maintained at a constant intermediate level throughout the cell cycle (Dean, et al, 1986, Hann, et al, 1985, Rabbitts, et al, 1985b, Thompson, et al, 1985). The maintenance of c-myc expression is dependent on the presence of serum growth factors, as myc levels rapidly fall upon their removal, approaching the limits of assay sensitivity within two to four hours (Dean, et al, 1986, and G Evan, D Hancock and T Littlewood, personal communication). C-myc levels rapidly rise following serum readdition, with a kinetic profile similar to that seen during stimulation from quiescence (Okuda, et al, 1989). A similar transient stimulation of c-myc expression is observed following mitogen addition to cycling cells, implying that the c-myc stimulatory pathways are constantly extant (C Waters and G Evan, personal communication). These observations are all consistent with a role for myc protein in the monitoring of extracellular signals.
1.6.4 Role in cell cycle

Regulatory roles for c-myc both during the progression into, and through, the cell cycle are suggested by its induction during the early mitogenic response and the correlation of its expression with the continued presence of mitogens throughout the cell cycle. It may exert its actions at multiple positions within the cell cycle by influencing immediate events required for progression through the cell cycle, as well as by influencing activities removed from the time of c-myc expression. Results of experiments which manipulate c-myc levels in a more direct fashion support these notions, although the mechanisms of c-myc action remain incompletely understood.


C-myc may also be directly involved in the control of DNA synthesis, although this remains a controversial issue. Some data does suggest such a role for c-myc protein. For instance, increased myc expression can lead to enhanced replication of SV40 replication origin based plasmids (Clas...
et al, 1987). In addition, a DNA fragment containing a putative cellular origin of DNA replication can be immunoprecipitated with anti-myc antibodies (Ariga, et al, 1989, Iguchi-Ariga, et al, 1987, Iguchi-Ariga, et al, 1988). However, not all of these findings have been reproduced (Gutierrez, et al, 1988). The strongest evidence for a direct function of c-myc protein in controlling DNA replication comes from experiments on Xenopus embryos. During the earliest phases of Xenopus embryonic development, when there is no requirement for novel RNA transcription (Bacharova, et al, 1966, Newport and Kirschner, 1982), microinjection of anti-myc antibodies inhibits cell proliferation, and these antibodies have been found to block the initiation of DNA synthesis in embryonic extracts (M. Mechali, personal communication). While these results strongly suggest an involvement of c-myc in the regulation of DNA synthesis, and thus entry into S phase, they do not conclusively demonstrate a direct role for c-myc in the synthesis of DNA.

There is additional evidence that c-myc has a function later in the cell cycle, during G2/M. This arises from the observation that Schwann cells and REF-52 cells which express an activated ras oncogene arrest during both G1 and G2/M. Coexpression of c-myc and ras prevents this growth arrest, suggesting c-myc functions at both these times to alleviate the arrest (Kohl and Ruley, 1987, Ridley, et al, 1988).

Activities of c-myc during the period from entry into S-phase through the end of G2/M may additionally affect the behaviour of cells during the early phases of the subsequent cell cycle. When fibroblasts are proliferating exponentially, cells transit from mitosis into S phase more rapidly than when they are stimulated to enter S phase from a quiescent state. The speed of this transition is thought to result from the post-mitotic cells traversing G1 without entering a G0 state. The control of this decision appears to be exerted during the previous cell cycle, not at the
transition from M to G\textsubscript{1}, and requires the apparently constant presence of specific growth factors (Scher, et al, 1979). A role for c-myc in this regulation is suggested by the correlation between its maintained expression from S through G\textsubscript{2}M, and the rapid entry of cells into S phase in the subsequent cell cycle (Okuda, et al, 1989, G. Evan, personal communication). Thus c-myc may function both in the progression from G\textsubscript{0} into G\textsubscript{1} and in the decision not to enter into G\textsubscript{0} following the completion of mitosis.

1.7 Multiple levels of regulation of c-myc expression

Since it has become clear that the inappropriate regulation of c-myc expression is sufficient to lead to its oncogenic activation, an immense effort has been expended in elucidating the pathways regulating c-myc expression (Spencer and Groudine, 1990 for review). A complex set of overlapping regulatory schemes have been found to be involved in controlling the level of c-myc protein expressed in cells. The profound effects of deregulated myc expression on normal cell growth and differentiation emphasise the importance of all of these methods of keeping c-myc levels under tight control. These include controls on transcription initiation and elongation, mRNA and protein stability and protein synthesis as well. The contribution of each of these controls varies with cell type, and in response to specific stimuli. Some examples of the combinatorial nature of these controls are presented below, and reported data on the molecular mechanisms through which these controls are instigated are described in the following section.

1.7.1 During differentiation and development

The mechanisms of c-myc control during development and differentiation have been studied in several systems. In HL60 cells there is a rapid decrease in c-myc mRNA upon induction of differentiation, which has been ascribed to a block in the elongation of nascent transcripts.
This elongation block persists for one to two days, after which levels of transcription initiation also fall, leading to a total cessation of transcription several days after induction of differentiation (Siebenlist, et al, 1988). Control of myc expression during WEHI 231 B-cell differentiation is achieved through a combination of alterations in rates of transcription as well as modulation of c-myc mRNA stability (Levine, et al, 1986).

In differentiating MEL cells, c-myc expression follows a biphasic pattern, with a 10-fold fall within five to six hours post-induction followed by recovery to initial levels by 24 hours. Several days after the initial induction, expression again falls to undetectable levels (Lachman and Skoultchi, 1984). This kinetic profile is produced by a number of changes in multiples levels of control. The initial fall is due to a block in transcription elongation similar to that seen with HL60 cells (Mechti, et al, 1986, Nepveu, et al, 1987). mRNA levels then recover to near initial levels, before falling following complete differentiation. After the commitment phase, at approximately 24 hours post-induction, the rate of protein synthesis is increased, and remains at this elevated level continuously. Around 24-48 hours post-induction, there is a dramatic dissociation of mRNA and protein levels, as the protein disappears completely while mRNA levels decline only three fold. These reductions occur through destabilisation of both the mRNA (Mechti, et al, 1986, Nepveu, et al, 1987, Watson, 1988) and protein (Spotts and Hann, 1990). Thus even though the mRNA is still synthesised, and the protein translated at slightly enhanced levels, the net result is a loss of any significant accumulation of steady state protein (Spotts and Hann, 1990). Post-transcriptional controls also appear to be utilised in avian hepatocytes. In situ detection of c-myc mRNA and protein has revealed that while the mRNA is highly expressed, no protein is detectable
Jaffredo, et al, 1989). The finding that myc protein and mRNA levels can be widely dissociated also makes it difficult to be confident of the quantity of c-myc protein in systems in which only mRNA levels have been assessed. While the regulation of c-myc protein at the translational and posttranslational levels have not been subjected to intense scrutiny in many developmental systems where c-myc mRNA has been studied, it is probable that such controls will be frequently found to overlap those which regulate the amounts of c-myc transcript.

Another regulatory scheme is evident during early Xenopus laevis development. In the oocyte there is a large pool of maternal c-myc mRNA which is translated and stored in a cytoplasmic pool during oocyte maturation (Godeau, et al, 1986, Taylor, et al, 1986). The protein is translocated into the nucleus immediately post-fertilisation (Gusse, et al, 1989). C-myc levels then drop precipitously as the embryo undergoes its first 12 rounds of cell division, a period during which there is no novel RNA polymerase II mediated transcription (Bacharova, et al, 1966, Newport and Kirschner, 1982). This reduction of the preexisting c-myc mRNA results in levels on a per cell basis which are similar to those seen shortly thereafter, when the embryonic c-myc genes are transcribed (Godeau, et al, 1986, Taylor, et al, 1986).

1.7.2 During mitogenic stimulation

The mitogenic stimulation of c-myc is also regulated by a combination of transcriptional and post-transcriptional controls, which vary depending on the specific mitogen, growth condition, and cell type under study. Treatment of quiescent, confluent Balb/c 3T3 A31 (Nepveu, et al, 1987) or NIH-3T3 (Greenberg and Ziff, 1984) fibroblasts with serum results in an increase of both c-myc transcription initiation and mRNA stability. Stimulation of subconfluent, mitogen deprived Balb/c 3T3 A31 cells with PDGF increases c-myc mRNA levels through purely post-transcriptional
mechanisms, while the same treatment of confluent cells leads to an increase in c-myc transcription (Rollins, et al, 1987). Stimulation of c-myc transcription by epidermal growth factor (EGF) and IMX (an inducer of intracellular cAMP levels) in the Balb/c 3T3 A31 cell line at confluence is due primarily to the release of a block to transcription elongation within the first exon, and not an increase in transcription initiation (Nepveu, et al, 1987). In another instance, the increase in c-myc mRNA levels in quiescent Chinese hamster ovary cells mitogenically stimulated with α-thrombin and insulin was found to be purely post transcriptional (Blanchard, et al, 1985). Reductions in c-myc mRNA levels following the mitogen-induced peak have also been found to utilise both transcriptional and posttranscriptional mechanisms (Dean, et al, 1986). In these experiments, as well as experiments on the control of c-myc expression during cellular proliferation, c-myc protein levels have always been observed to parallel changes in c-myc mRNA level (Evan, et al, 1988, Hann, et al, 1985, Moore, et al, 1987, Rabbitts, et al, 1985b, Thompson, et al, 1985). It thus appears that, in contrast to the regulatory controls observed in differentiation and developmental systems, the regulation of c-myc expression during entry into and throughout the cell cycle is achieved primarily through the control of c-myc mRNA levels.

1.8 Mechanisms regulating c-myc expression

Numerous cis-acting elements within the c-myc gene and mRNA have been implicated in the regulation of c-myc expression. Some progress has also been made in identifying the cellular factors which operate through these elements. The complexity of these systems is just beginning to be understood, and it is therefore not yet possible to draw a comprehensive scheme through which the contributions of all of these components can be appreciated.
1.8.1 Cis-acting transcriptional control elements and trans-acting factors

Transcription initiation at the c-myc promoters is regulated by a combination of positive and negative regulatory elements, which appear to be at least partially utilised in a cell type specific fashion. These modulatory sequences have been identified through studies of both the human and murine c-myc promoters. As the sequences of the c-myc genes of these species are highly homologous for several hundred bp upstream of the mRNA cap sites, and the genes appear to be regulated similarly in response to several stimuli, it is probable that many of the regulatory mechanisms are conserved between the two species. Each study has utilised different constructs, cell lines and methodologies for its analyses, which has resulted some consistent and many contradictory conclusions.

1.8.1.1 Minimal promoter elements

Both the P1 and P2 promoters have architectures centered around TATA boxes, sequences which probably act both to position the mRNA cap sites and contribute to basal levels of transcription. The sequences within the immediate vicinity of the cap sites have a high GC content, as has been seen in numerous other genes. The P2 promoter is predominantly utilised in most cell types in vivo, accounting for 75%-90% of the mRNA.

In vitro assays utilising both stable and transient transfections, as well as microinjection of c-myc DNA into Xenopus oocytes, have generally shown a ratio of promoter usage which reflects that seen in vivo, with approximately five to ten fold more transcript initiating at the P2 promoter. This ratio has been found to be relatively constant as long as sequences extending several hundred bp upstream of P1 have been included.

Essential sequences for P1 initiated transcription have been located between positions -101 and -293 relative to the P1 promoter cap site of the
human c-myc gene in both transfection (Hay, et al, 1987) and in vitro transcription (Postel, et al, 1989) analyses. These findings are in contrast with murine c-myc promoter transfections (Asselin, et al, 1989) and microinjection of the human c-myc gene into Xenopus oocytes (Nishikura, 1986) which defined a requirement only for sequences within 100 bp of P1. The differences between these results have not yet been reconciled.

DNA structural features as well as trans-acting factors which interact with some of these sequences have been identified. The sequences between -142 and -115 relative to the human c-myc P1 promoter contain a DNAase I hypersensitive site in vivo, the presence of which corresponds to positive activity of the promoter (Siebenlist, et al, 1988, Siebenlist, et al, 1984). This sequence is also S1 sensitive when present in supercoiled DNA in vitro, and the S1 sensitive structure is stabilised by the addition of RNA (Boles and Hogan, 1987). The S1 sensitive sequence contains a highly purine rich segment which may form a triple helical H-DNA structure, and an antisense oligomer containing this sequence will bind duplex DNA and repress transcription from the P2 promoter in vitro (Cooney, et al, 1988). Protein complexes which bind this sequence have also been detected, at least one of which is part of a ribonucleoprotein complex, and it has been suggested that the interaction of antisense RNA with this sequence may play a role in the negative regulation of c-myc transcription (Davis, et al, 1989, Kinniburgh, 1989, Postel, et al, 1989). This sequence is contained within a fragment which increases P2 mediated transcription, and may thus act to regulate transcription from both P1 and P2 in vivo (Lipp, et al, 1987).

The only transcription factor for which a binding site closer to the P1 promoter has been identified is Sp1. This protein may interact with a GC rich sequence found at position -44 from P1 (Asselin, et al, 1989). The
Spl protein is known to act as a basal transcription element in many promoters (Jones, et al, 1988).

Most groups have found that sequences within approximately 70 bp of the P2 cap site are necessary for efficient utilisation of this promoter, although there is some disagreement as to which sequences are most important (Asselin, et al, 1989, Hall, 1990, Lipp, et al, 1987, Nishikura, 1986, Thalmeier, et al, 1989). Two major elements have been identified. One is located from positions -66 to -56 upstream of the P2 promoter cap site, and has been implicated in control of the human c-myc P2 promoter (Lipp, et al, 1987, Nishikura, 1986, Thalmeier, et al, 1989), while the other extends from position -58 through -40 and has been identified through studies on the murine c-myc gene (Asselin, et al, 1989, Hall, 1990). A DNAase I hypersensitive site which correlates with transcription from the human c-myc P2 promoter has also been mapped to sequences approximately 60 bp upstream of P2 (Bentley and Groudine, 1986b, Siebenlist, et al, 1988, Siebenlist, et al, 1984). As these studies have utilised numerous human and murine cell types for analysis of both of these genes, and the sequences of the two genes are almost identical from position -70 through the P2 promoter TATA box (Bernard, et al, 1983), it is highly probable that the same elements throughout this region are involved in the regulation of both genes.

The sequence -66 GGCGGGAAAA -58 has been identified as a binding site for the transcription factor E2F, and this protein is presumably responsible for at least some of the positive regulatory attributes which have been attributed to this sequence (Hiebert, et al, 1989, Lipp, et al, 1989, Mudryj, et al, 1990, Thalmeier, et al, 1989). The E2F factor was originally identified as a mediator of the induction of transcription from the adenovirus E2 promoter by the adenovirus E1a protein (Kovesdi, et al, 1986). The E2F binding site of the E2 promoter consists of inverted repeats

The -57/-40 region has been identified as a binding site for a factor termed Me1a1 (Asselin, et al, 1989). While the binding sequences for this factor have not been well characterised (Asselin, et al, 1989), it has been suggested that sequences which are the murine promoter equivalent of human -51/-40 (5' GGAGGGAGGGA 3') are sufficient to mediate its effects (Hall, 1990).

Intriguingly, a protein, p55, has been reported to bind in a sequence specific fashion to RNA complementary to the sense strand of the human c-myc promoter between positions -66 and -40 (Parkin and Sonenberg, 1989). A portion of this sequence, GGGAGGG, is similar to that of the -115 P1 promoter DNA regulatory element described above, and this protein may be involved in regulation through that position as well, or in other transcriptional or posttranscriptional controls.

1.8.1.2 Upstream and downstream elements

Several other control regions of the c-myc promoter have been identified in addition to these promoter proximal elements. The functionality of these sequences appears to vary to a large degree with the cellular context in which they are assessed. These elements consist of both positive and negative regualtory sequences, and in some cases the proteins with which they interact have been characterised or identified.
Positive regulatory elements have been identified in several studies. Sequences upstream of the human c-myc P1 promoter between positions -2327 and -1257, -1257 and -353 acted positively on both promoters in one study (Hay, et al, 1987), while they were dispensible for maximal activity in another (Thalmeier, et al, 1989). An orientation dependent positive control element has been identified in the murine c-myc locus at the 3' end of the first exon (Yang, et al, 1986). These elements remain poorly understood, and the proteins which may interact with the regulatory sequences have yet to be identified.

Three elements which act negatively on transcription have also been identified within c-myc upstream regulatory regions. The first of these maps to the 26 bp sequence of the human c-myc gene from -343 to -318 5' of P1. Protein complexes which interact with these sequences have been shown to contain both members of the fos/jun transcription factor family, and an octamer sequence binding protein (Hay, et al, 1987, Hay, et al, 1989, Takimoto, et al, 1989). Another negative element has been identified within the murine c-myc promoter and acts specifically to repress transcription of c-myc in plasmacytoma cells. This is a purine rich sequence centered around position -290 upstream of P1. This sequence is bound by a plasmacytoma specific factor, myc-PRF. A general transcription factor, myc-CF1, which binds nearby may also be involved in this repression (Kakkis and Calame, 1987, Kakkis, et al, 1988, Kakkis, et al, 1989). The third negative element is present in sequences between positions -1188 and -428 upstream of the murine c-myc P1 promoter, but has not been characterised further (Remmers, et al, 1986).

The transforming growth factor β1 (TGF-β1) induces growth arrest of keratinocytes. C-myc transcription is repressed following this treatment, and an element which can mediate this repression in keratinocytes has
been mapped to sequences between -100 and +71 relative to the human c-myc P1 promoter (Pietenpol, et al, 1990).

Several sequences which are capable of modulating transcription from the c-myc promoters have thus been identified. In some instances proteins which may mediate these controls have been partially characterised. In most instances whereby c-myc transcription is regulated, however, very little is known either of the regulatory DNA sequences or the transcription factors with which they interact.

1.8.2 Control of transcript elongation

It has been well established that an important mechanism in the control of c-myc expression is in the ability to regulate the proportion of transcripts which are fully elongated. There is often a fraction of the nascent c-myc mRNA which is terminated before reaching the normal transcription stop sites, does not get polyadenylated, is rapidly degraded and is thus not translated. Modulation of the amount of attenuated transcript plays a role in the response of c-myc expression to growth as well as differentiation signals (Bentley and Groudine, 1986a, Mechti, et al, 1986, Nepveu, et al, 1987, Nepveu and Marcu, 1986).

The sequence requirements for attenuation have been investigated using DNA templates microinjected into oocytes, where the truncated transcripts are stable, and with transfected cell lines, which must be evaluated by nuclear run-on assays. The major sites of attenuation have been mapped to runs of T residues near the 3' end of the first c-myc exon. In hybrid promoter experiments it has been demonstrated that addition of a 95 bp sequence from the 3' end of the human c-myc first exon (Bentley and Groudine, 1988), or 180 bp from a similar region of the murine c-myc first exon (Miller, et al, 1989, Wright and Bishop, 1989) is sufficient to lead to attenuated transcription from some, but not all heterologous promoters. Furthermore, these sequences function in an orientation
dependent manner and while transcripts normally terminate at poly T tracts, these are not strictly necessary for attenuation to occur (Bentley and Groudine, 1988). Sequences within the 3' element have the potential to form stem-loop structures which might be involved in termination, although it is not clear if they exist in vivo. Studies on the murine c-myc gene transfected into cell lines have suggested that a 16 bp sequence just upstream of the P2 cap site is required for attenuation to occur. Loss of this sequence correlates with the inability to bind the Me1a1 factor (Wright and Bishop, 1989). It has been suggested that transcripts which initiate from P1 are incapable of being prematurely terminated (Spencer, et al, 1990). This does not appear to be strictly true, as a novel attenuation site upstream of the P2 mRNA cap site, and at which only P1 initiated transcripts can be terminated, has recently been identified. This site maps to sequences within approximately 40 bp of the P2 cap site, and can act autonomously on heterologous promoters (S Roberts and D Bentley, personal communication).

A DNAase hypersensitive site (IV) which maps within the 5' half of the first intron correlates with increased levels of transcript attenuation in differentiating HL60 cells (Bentley and Groudine, 1986a, Eick and Bornkamm, 1986, Siebenlist, et al, 1988). The significance of this change has not yet been determined.

1.8.3 mRNA stability

Levels of c-myc mRNA are also controlled by factors which modulate the stability of the RNA. The observation that transcripts from murine plasmacytoma cells, which have a truncated first exon, have increased stability, led to the hypothesis that sequences in this exon act as a tag for rapid RNA degradation (Eick, et al, 1985, Piechaczyk, et al, 1985, Rabbitts, et al, 1985a). Transfection and in vitro degradation studies have indicated that these sequences appear to contribute to, but are insufficient for, the
instability of the c-myc mRNA (Brewer and Ross, 1989, Jones and Cole, 1987, Pei and Calame, 1988). Sequences from the 3' end of the mRNA also appear to be required, including an AU rich sequence which is similar to sequences previously demonstrated to destabilise other mRNAs (Brewer and Ross, 1989, Jones and Cole, 1987, Pei and Calame, 1988, Shaw and Kamne, 1986).

1.8.4 Protein translation and degradation

There is some evidence that both translatability of the c-myc mRNA and the stability of the protein can be regulated. Very little is known of the mechanisms by which these controls are executed, except that a 240 bp region of the murine c-myc first exon is capable of inhibiting translation in in vitro extracts and upon microinjection into Xenopus oocytes, but not in several cell lines (Darveau, et al, 1985, Parkin, et al, 1988).

1.9 Molecular functions of c-myc: Regulation of gene expression

The nuclear localisation of the c-myc protein, its non-specific DNA binding ability and its structural homologies with many transcription factors have suggested that c-myc induces its pleiotropic effects by regulating gene expression. Several studies have correlated c-myc expression with modulations in gene expression, and have shown that some genes are regulated at either the transcriptional or post-transcriptional levels by increases in c-myc expression. Transcriptional regulation has been observed for the LFA-1 cell adhesion molecule α chain in EBV immortalised B-lymphoblasts, which is repressed by introduced c-myc (F Grignani and R Dalla-Favera, personal communication). The human heat shock 70 promoter is activated and the metallothionein I promoter repressed by transient cotransfection with c-myc (Kaddurah-Daouk, et al, 1987, Kingston, et al, 1984). Two additional genes, mr2 and plasminogen activator inhibitor I are post-transcriptionally increased when c-myc expression increases (Prendergast and Cole, 1989,
Prendergast, et al, 1990). An additional group of genes whose expression appears to be modulated by c-myc overexpression have been identified, and include class I HLA (Versteeg, et al, 1988), two histone genes (Cheng and Skoultchi, 1989), and two genes which are regulated at the G₀/G₁ transition (Schweinfest, et al, 1988). Although the expression levels of all of these appear to be altered in the presence of excess c-myc expression, the regulatory elements through which c-myc may act have not yet been identified, and it is therefore difficult to assess the directness of c-myc action.

1.10 Autoregulation

et al., 1984b, Stone, et al., 1987, Zerlin, et al., 1987). A possible cross regulatory network of myc gene expression is suggested by the repression of endogenous c-myc expression upon introduction of N-myc protein into cell lines (MW Brooks, LJZ Penn and H Land, personal communication), and by the lack of c-myc expression observed in tumors derived from N-myc transgenic mice (Dildrop, et al., 1989).

In one case it has been demonstrated that fusion of a cell line which autoregulates c-myc expression with one which does not results in the c-myc genes derived from both parental lines becoming down-regulable (Penn, et al., 1990a). This result implies that at least one trans-acting factor in addition to myc is involved in the autosuppression response. It also suggests a possible mechanism whereby non-responsive cell lines can be generated.

Although the initial impetus behind an autoregulatory theory of c-myc expression was the result found in Burkitt's lymphoma cell lines, it now appears that suppression of the unrearranged allele is probably not due to a negative feedback loop in this system. The strongest piece of evidence against a direct role for c-myc protein is the stability of repression observed in some Burkitt's derived cell lines even in the presence of cycloheximide, a treatment which can lead to the rapid induction of c-myc expression in fibroblast cell lines, presumably due to the lability of a repressor molecule, of which c-myc, as a highly unstable protein, would be a prime candidate. Suppression of the unrearranged c-myc allele in Burkitt's lymphoma cell lines is now thought to occur via a longer term regulatory mechanism related to the developmental stage of the untransformed target cells, in which c-myc may not normally be expressed. This mechanism may involve processes such as methylation or alterations in chromatin structure (Dunnick, et al., 1985, Nishikura and Murray, 1988, Siebenlist, et al., 1984), reviewed in (Spencer and Groudine, 1990).
The c-myc protein levels required for autosuppression in the Rat-1 fibroblast line have been quantified. It was found that the degree of repression of endogenous c-myc expression is proportional to the concentration of c-myc protein. Thus in proliferating cells which normally express 600-700 molecules of c-myc protein per cell, a ≥90% reduction in expression is observed upon the forced expression of ≥5000 molecules per cell of human c-myc protein, and it appears that doubling the c-myc protein concentration results in the subsequent reduction of c-myc expression by 50% (Penn, et al, 1990a). The levels of human c-myc protein required to induce a significant change in the expression of the endogenous gene fall well within those detected during normal cellular proliferation, and are consistent with autoregulation performing a homeostatic function.

The regions of the c-myc protein which are required for autosuppression have been identified (See Figure 1.1). They include two portions of the protein, amino acids 105-143 and 353-433 (Penn, et al, 1990b). Interestingly, the regions of the protein required for autoregulation overlap with those required for cooperation with the ras oncogene in transformation (Stone, et al, 1987). This suggests that the molecular mechanisms of these two activities may be similar. Thus myc protein may suppress the transcription of other genes involved in regulating normal cell growth, thereby exerting its oncogenic function when constitutively expressed. While amino acids 106-143 do not display homology with any known proteins outside of the myc family, amino acids 353-433 contain the basic, helix-loop-helix and leucine zipper regions discussed previously. In other proteins, these structures are involved in multimerisation and DNA binding, suggesting that c-myc may negatively regulate its transcription by directly or indirectly binding to its own promoter (Penn, et al, 1990c, for review).
1.11 Aims of the project

Autoregulation of c-myc expression is thus an activity which fuses two important issues in the study of the c-myc oncogene. The first of these concerns the control of c-myc expression, the loss of which have been shown to be sufficient to lead to its activation as an oncogene. Second, as it is the myc protein which is acting to regulate its own expression, if a direct interaction can be demonstrated between c-myc gene sequences and myc protein, it may provide clues about the biochemical basis of myc function.

The project undertaken in this thesis has been to determine the sequences within the human c-myc gene through which myc autoregulation is mediated, to characterise the protein complexes which can interact with these sequences, and to investigate their relationship to myc expression and the possibility that myc protein might itself be present in these complexes.
Figure I.1: Domains of human c-myc protein

Structural and functional domains of human c-myc. Secondary structure predictions based on amino acid sequence information suggest c-myc may be composed out of an α-helix/β-sheet domain (aa 1-203) and a predominantly α-helical domain at the carboxyl end (aa 238-439) which are separated by a less structured hinge region (aa 204-237). Additional structural motifs include: a highly acidic domain (XXX); a basic region (---); a helix-loop-helix domain (---); and a leucine zipper (---). Casein kinase II phosphorylation sites (*) and regions most highly conserved among the members of the myc gene family (---) are also indicated. The functional domains of c-myc identified to date are represented by the blackened regions in the specified schematic diagrams and map as follows: cotransformation activity together with ras oncogenes in rat embryo fibroblasts, residues 106-143 and 353-439; autosupression, residues 106-143 and 353-439; nuclear localisation, residues 320-328 and 364-374; and non-specific DNA binding, residues 290-318. II/III, indicates the border of exon II and exon III-encoded sequences.
HUMAN C-MYC PROTEIN

- Autosuppression
- Cooperation with ras
- Nuclear localization
- Non-specific DNA binding
Figure 1.2: DNA binding domains of proteins with homology to c-myc

Alignment of amino acid sequences of members of the basic leucine zipper (LZ) and basic helix-loop-helix (HLH) protein families with members of the myc family of proteins. The region of each proposed domain structure is indicated by the appropriate symbol above or below the sequences: basic (+), leucine zipper (Z), amphipathic helix (H), loop (Ω). Vertical lines delineate predicted domains. Spaces between residues are gaps added to improve the alignments, and all sequences shown are contiguous within one protein. The • symbols indicate highly conserved residues as they relate to the myc proteins. The alignments shown are based on those presented in (Kouzarides and Ziff, 1988, Murre, et al, 1989a, Prendergast and Ziff, 1989).

translocated human oncogene Lyl-1 (Mellentin, et al, 1989); the
myogenic control proteins MyoD (M: murine, X: Xenopus, CMD1:
(Braun, et al, 1989) and myogenin (Wright, et al, 1989); and the
Drosophila developmental regulatory proteins twist (Thisse, et al, 1988);
hairy (Rushlow, et al, 1989) and enhancer of split (E(spl)) m5, m7 and
m8 (Klämbt, et al, 1989).
Basic

Helix I

Loop

Helix II

----------------------------------------

Cys 208 REANVHAREERSEDEKAKRHU

Glu 88 LDLKRAAREDSEKAKRHU

Tyr 191 LUKASGLRGEQKEYK

Cre 290 LAKAVREERSEDEKAKRHU

Cys 359 LEANHAREERSEDEKAKRHU

Tag 200 REANVHAREERSEDEKAKRHU

Cre-Rh 144 REANVHAREERSEDEKAKRHU

\[ \text{Helix I} \]

\[ \text{Loop} \]

\[ \text{Helix II} \]
Chapter 2 Materials and Methods
2.1 Materials

2.1.1 Reagents

Acrylamide and N,N'-methylene bisacrylamide
Bio-Rad separate components, or BDH Acrylogel
5 premix

Activated charcoal
Sigma

Agarose and low melting point agarose
SeaKem, SeaPlaque: FMC BioProducts

Acids
Acetic (glacial): FSA
Boric: FSA
Hydrochloric: FSA

Ammonium persulphate
Bio-Rad

Antibiotics
Ampicillin (Na salt): Sigma
Chloramphenicol: Sigma
G418 (Genetecin): Gibco
Hygromycin B: Sigma
Kanamycin: Sigma
Puromycin: Sigma
Tetracycline: Sigma

Antifoam
Sigma

Antisera
Rat α v-Ha-ras: Oncogene Science

Aquasol
Dupont

β-mercaptoethanol
Sigma

Bacto agar
Difco

Brain heart infusion
Difco

BSA
Fraction V: Boehringer Manheim
For restriction analysis: BRL

Buffers
HEPES: Sigma
MOPS: Sigma
PIPS: Sigma
Tris acid: Sigma
Tris base: Sigma

CaCl₂
BDH

Chloroform
FSA

CsCl
BRL ultrapure

Dexamethasone
Sigma

Dextran Sulphate
Pharmacia

Digitonin
Sigma

63
<table>
<thead>
<tr>
<th>Chemical</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimethyl formamide</td>
<td>Sigma</td>
</tr>
<tr>
<td>DMS</td>
<td>Sigma</td>
</tr>
<tr>
<td>DMSO</td>
<td>BDH</td>
</tr>
<tr>
<td>DTT</td>
<td>Sigma</td>
</tr>
<tr>
<td>EDTA</td>
<td>FSA</td>
</tr>
<tr>
<td>EGTA</td>
<td>Sigma</td>
</tr>
<tr>
<td>EtOH</td>
<td>James Burrough</td>
</tr>
<tr>
<td>Ficoll</td>
<td>Pharmacia</td>
</tr>
<tr>
<td>Film</td>
<td>Fuji RX</td>
</tr>
<tr>
<td></td>
<td>Kodak XAR-5</td>
</tr>
<tr>
<td></td>
<td>Polaroid Type 55 and 57</td>
</tr>
<tr>
<td>Filters (0.2 &amp; 0.45 µm)</td>
<td>Millipore (syringe), Nalgene (100-500 ml)</td>
</tr>
<tr>
<td>Formaldehyde (38%)</td>
<td>BDH</td>
</tr>
<tr>
<td>Formamide</td>
<td>BDH (deionised with Bio-Rad AG 501-X8 ion-exchange resin)</td>
</tr>
<tr>
<td>Formic Acid</td>
<td>Sigma</td>
</tr>
<tr>
<td>Genescreen</td>
<td>Dupont - NEN Research Products</td>
</tr>
<tr>
<td>GeneScreen Plus</td>
<td>Dupont - NEN Research Products</td>
</tr>
<tr>
<td>Glass wool (siliconised)</td>
<td>Mandel Scientific</td>
</tr>
<tr>
<td>Glucose</td>
<td>BDH</td>
</tr>
<tr>
<td>Glutaraldehyde</td>
<td>BDH</td>
</tr>
<tr>
<td>Glycerol</td>
<td>FSA</td>
</tr>
<tr>
<td>Glycogen (molecular biology grade)</td>
<td>Boehringer</td>
</tr>
<tr>
<td>Guanadine thiocyanate</td>
<td>Fluka</td>
</tr>
<tr>
<td>8-hydroxyquinoline</td>
<td>Sigma</td>
</tr>
<tr>
<td>Intensifying screens</td>
<td>Du Pont Lightning Plus</td>
</tr>
<tr>
<td>IPTG</td>
<td>BRL</td>
</tr>
<tr>
<td>Isopropanol</td>
<td>FSA</td>
</tr>
<tr>
<td>Isotop II</td>
<td>Coulter Electronics</td>
</tr>
<tr>
<td>KCl</td>
<td>FSA</td>
</tr>
<tr>
<td>K₃Fe(CN)₆</td>
<td>Sigma</td>
</tr>
<tr>
<td>K₄Fe(CN)₆·H₂O</td>
<td>Sigma</td>
</tr>
<tr>
<td>KOAc</td>
<td>Sigma</td>
</tr>
<tr>
<td>Linkers, phosphorylated</td>
<td>Pharmacia</td>
</tr>
<tr>
<td>MeOH</td>
<td>BDH</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>FSA</td>
</tr>
<tr>
<td>MgSO₄·7H₂O</td>
<td>FSA</td>
</tr>
</tbody>
</table>
MnCl₂ FSA
Molecular weight markers λ Hind III and φX174 RF Hae III: BRL
Molecular weight V: Boehringer
123 bp and 1 kb ladder: BRL
NaCl BDH
NACS prepac cartridges BRL
Na₂HPO₄ FSA
NaH₂PO₄ FSA
NaOH BDH
Needles and Syringes Sabre
Nitrocellulose Millipore
Nonidet P40 Sigma
Nucleotide, deoxynucleotide, and dideoxynucleotide triphosphates Pharmacia
Oligonucleotides ICRF Clare Hall Services Unit using
Parafilm American Nat'l Can
an Applied Biosystems 380B synthesiser
Petri dishes Sterilin
PEG Koch Light
Phenol (redistilled) BRL ultrapure
Piperidine Sigma
Polybrene Sigma
Polyvinylpyrrolidone Sigma
Protease Inhibitors Aprotinin: Sigma
Benzamidine HCl: Sigma
Leupeptin: Sigma
PMSF: Sigma
STI: Sigma
TPCK: Sigma
Protein Assay Mix BioRad
Radiochemicals Amersham International
Random primers pd(N)₆ Pharmacia
RNAsin Promega
Salmon sperm DNA Sigma
SDS Sigma
Sephadex G50 Pharmacia
Serum

Spermidine

Spermine

Stains

Sucrose

TEMED

Tissue culture dishes

Triton X-100

Tryptone

Urea

X-gal

Yeast extract

Yeast tRNA

ZnCl₂

Foetal calf: Globepharm

Newborn calf: Gibco

Sigma

Sigma

Bromophenol blue: Sigma

Crystal violet: Sigma

EtBr: Sigma

Xylene cyanol: Sigma

Sigma

Bio-Rad

Falcon, Corning, Nunc

Sigma

Difco

BRL

BRL

Difco

Sigma

FSA

2.1.2 Enzymes

CIP (molecular biology grade) Boehringer

DNA polymerase I Klenow Fragment

Pharmacia

Exonuclease Bal 31 BRL

Exonuclease III BRL

Nuclease S1 Sigma

Proteinase K Boehringer

Restriction endonucleases Amersham, BRL, Boehringer, NBL, New England Biolabs, Pharmacia

RNase A Sigma

RNase T1 Sigma

Sequenase (kit) US Biochemicals

SP6 RNA polymerase New England Biolabs, Boehringer

T3 RNA polymerase Stratagene

T4 DNA ligase Pharmacia

T4 polynucleotide kinase Pharmacia

T7 RNA polymerase BRL, Stratagene
2.1.3 Bacterial strains and genotypes

DH1
- F-, hsdS20(rB-,mB-), recA13, ara-14, proA2, lacy1,
galk2, rpsL20(Smr), xyl-5, mtl-1, supE44, λ-

DH5

SCS-1 (Stratagene)
- F-, end A1, hsdR17(rK-,mK+), supE44, thi-1, λ-,
recA1, gyrA96, relA1

XL-1 blue (Stratagene)
- endA1, hsdR17(rK-,mK+), supE44, thi-1, λ-, recA1,
gyrA96, relA1, Δ(lac), [F, proAB, lacIqZΔM15,
Tn10 (tetR)]

2.1.4 Cloning vectors

pUC 13 (lab stocks)

Bluescript KS +/- and SK +/- Stratagene

pGCl (S. Goodbourn) (Myers, et al, 1985)
pGEM3 Promega
pJ6Ω (lab stocks) (Morgenstern and Land, 1990b)

2.1.5 Cell Lines

NIH 3T3 (lab stocks) (Jainchill, et al, 1969)

NIH 3T3-3 (lab stocks)

HeLa S3 (N. Jones) (Puck, et al, 1956)

ψ-2 (lab stocks) (Mann, et al, 1983)

PA317 (A.D. Miller) (Miller and Buttimore, 1986)

GP+E (S. Goff) (Markowitz, et al, 1988)

Rat-1 (lab stocks)

2.1.6 Stock solutions

PBS.A
- 1% w/v NaCl
- 0.025% w/v KCl
- 0.14% w/v Na₂HPO₄
- 0.025% w/v KH₂PO₄

PBS.ABC PBS.A with 6mM MgCl₂ and 10mM CaCl₂

Versene
- 0.54 mM EDTA in PBSA

1X SSC
- 150 mM NaCl
- 15 mM Na Citrate
pH 7.5 adjusted with NaOH

1X TAE
40 mM Tris OAc
2 mM EDTA

1X TBE
89 mM Tris base
89 mM boric acid
2 mM EDTA

1X MOPS
20 mM MOPS
1 mM EDTA
5 mM NaOAc
pH adjusted to 7.0 with glacial acetic acid

1X Denhardt's Solution
0.02% w/v Ficoll
0.02% w/v polyvinylpyrrolidone
0.02% w/v BSA (Fraction V)

TE
Tris Cl pH 7.5 or 8.0
1 mM EDTA

TEN
TE with 100mM NaCl

TES
TE pH 7.5 with 0.1% w/v SDS

fc
49.5% v/v redistilled phenol
49.5% v/v chloroform
1% w/v isoamyl alcohol
0.1% w/v 8-hydroxyquinone
saturated with TE

Plasmid Prep Solutions
Solution I
50 mM glucose
25 mM Tris Cl pH 8.0
10 mM EDTA

Solution II
200 mM NaOH
1% w/v SDS
Solution III  

5M KOAc pH 4.8

CsCl solution  

120g CsCl in 110 ml H₂O

CsCl/EtBr solution  

120 g CsCl in 100 ml dH₂O + 10 ml 5 mg/ml EtBr

BHI medium  

3.7% w/v brain heart infusion (Oxoid) adjusted to pH 7.5 with NaOH

BHI agar  

BHI media supplemented with 1.5% w/v bacto agar

M + G stop  

80% v/v formamide
0.02% w/v bromophenol blue
0.02% w/v xylene cyanol

SOC medium  

2% w/v tryptone
0.5% w/v yeast extract
10 mM NaCl
2.5 mM KCl
10 mM MgCl₂
10 mM MgSO₄
20 mM glucose

Tris Saline  

24mM Tris Base
0.8% w/v NaCl
0.38% w/v KCl
0.01% w/v Na₂HPO₄
0.1% w/v dextrose
0.001% w/v phenol red
adjusted to pH 7.7 with HCl

Trypsin  

Tris Saline plus
0.25% trypsin
100 u/ml Penicillin (Gibco)
0.01% w/v Streptomycin (Gibco)
TYM medium

- 2% w/v tryptone
- 0.5% w/v yeast extract
- 100 mM NaCl
- 10 mM MgSO₄

2X TY

- 1% w/v tryptone
- 1% w/v yeast extract
- 0.5% w/v NaCl

2X TY top agar

- 2X TY supplemented with 0.6% w/v bacto agar

2X TY agar

- 2X TY supplemented with 1.5% w/v bacto agar

2.1.7 DNA constructs

The following plasmids were constructed during the course of this project.

2.1.7.1 Human c-myc gene promoter constructs:

All references to sequence positions within the human c-myc gene are relative to the position of the P1 transcription initiation site (Battey, et al, 1983, Bernard, et al, 1983, Gazin, et al, 1984, Watt, et al, 1983) unless otherwise specified. The human c-myc genomic sequence from the HindIII site at -2329 through the EcoRI site at +5753 is shown in Figure II.1, and displays the exon/intron structure of the gene, the c-myc open reading frame, the endpoints of some of the deletion constructs described below, as well as selected protein binding sites proximal to the P1 and P2 promoter mRNA cap sites. Schematic diagrams of the various deletion series described here are presented as figures in Chapter 3 (Figures III.1, 9, 14, 17, 24, 26 and 29).

1 pBluescript vector constructs

pBSKS+HBCAT contains ~2.5 kb (-2329 to +208) of human c-myc promoter 5' sequences fused to bacterial CAT gene. The plasmid was constructed by first cloning a human c-myc genomic 2.5 kb HindIII-NaeI
fragment into the vector pCATb' which had been cleaved with BamHI,
made blunt and then cleaved with HindIII to yield pCATb'H(B). The 4.3
kb HindIII-Sall fragment of pCATb'H(B) containing the c-myc
promoter/CAT fusion was subcloned into HindIII-Sall digested Bluescript
KS+ vector.

pBSKS+HN277 contains the same human c-myc gene sequences as
pBSKS+HBCAT but fused to a fragment of the human β-IFN gene. The
HindIII-Nael human c-myc fragment and the 1.3 kb PvuII-EcoRI fragment
of pBPVIF* were cloned in a three-part ligation into HindIII-EcoRI cleaved
Bluescript KS+. This IFN fragment also contains some β-globin specific
sequences 3' of the β-IFN gene.

pBSKS+(2329X/+208B) contains the HindIII-NaeI -2329/+208 human c-
myc cloned into Bluescript KS+ vector. The HindIII site was made blunt
and linker to an XbaI (5'CTCTAGAG3') site. Following XbaI and Nael
digestion, the fragment was subcloned into XbaI/SmaI digested vector.
The final construct contains the human c-myc sequences in the same
orientation relative to the vector as are found in BSKS+HBCAT.

pBSKS+(2329X/+208)IFN is the result of removing the β-globin
sequences from the 3' end of pBSKS+HN277. This was done by subcloning
the NarI-BalI fragment of pBSKS+HN277 into NarI-EcoRV digested
pBSKS+(2329X/+208B) vector. The NarI sites in the human c-myc
promoter are found at positions -981 and -420. A NarI partial digest of the
pBSKS+HN277 plasmid before cleaving with BalI produced inserts which
contained both of these 5' endpoints. The vector DNA was digested to
completion, leaving only the -981 site as a 3' endpoint. Therefore it was
possible to construct pBSKS+ (2329X/-981:420/+208)IFN as well by using
the more truncated insert.
pBSKS+(2329B/+208)IFN was made by cleaving pBSKS+(2329X/+208)IFN with XbaI, blunting the ends, and inserting a BamHI (5'CCGGATCCGG3') linker.

pBSKS+(2329X/A981:420/+208)IFN when cleaved with XbaI and NarI, made blunt, BamHI (5'CCGGATCCGG3') linkered and recircularised, resulted in pBSKS+(420B/+208)IFN.

As there are multiple PvuII sites in pBSKS+(420B)IFN, the -353 deletion endpoint required a two-part ligation. pBSKS+(420B)IFN was cleaved with BamHI and Sall to produce a linear backbone, which was ligated to PvuII digested, BamHI (5'CCGGATCCGGG3') linkered, Sall digested insert to produce pBSKS+(353B/+208)IFN.

pBSKS+(420B)IFN cleaved with Smal, BamHI (5'CCGGATCCGGG3') linkered and digested, then recircularised yielded pBSKS+(101B+208)IFN.

Two deletion series were constructed using pBSKS+ (2329X/+208)IFN and pBSKS+ (2329X/Δ981:420/+208)IFN as starting material. The plasmids were cleaved with SstI and XbaI then digested with exoIII. The timepoint aliquots were made blunt and BamHI (5'CCGGATCCGG3') linkered, then recircularised. A number of colonies were chosen, the endpoints estimated by restriction mapping and confirmed by dideoxynucleotide sequencing using the M13 universal sequencing primer which will prime from just 5' of the Bluescript polylinker. The following constructs were isolated:

- pBSKS+(489/+208)IFN
- pBSKS+(420/+208)IFN
- pBSKS+(399/+208)IFN
- pBSKS+(348/+208)IFN
- pBSKS+(276/+208)IFN
- pBSKS+(181/+208)IFN
- pBSKS+(169/+208)IFN
pBSKS+(91/+208)IFN
pBSKS+(88/+208)IFN
pBSKS+(86/+208)IFN
pBSKS+(58/+208)IFN
pBSKS+(+119/+208)IFN

*pBSKS+(1034/Δ981:420/+208)IFN* this construct and the following one contains an internal deletion.

pBSKS+(1002/Δ981:420/+208)IFN

pBSKS+712IFN *this endpoint lies within the IFN gene and contains no human c-myc sequences.*

2 pGC1 vector constructs

pGC1H/R contains the 8.1 kb EcoRI/HindIII human c-myc genomic DNA fragment of phcmyc, a plasmid derived from the phage λ-LMC-41 (Dalla Favera, et al, 1982a), and cloned into the HindIII and EcoRI sites of the vector pGC1 (Myers, et al, 1985). It contains all of the human c-myc sequence from nucleotides -2328 through +5754.

pGC1H/RBam8 is pGC1H/R with an 8 bp BamHI linker (5'CGGATCCG3') inserted at the EcoRV site at position +2332. This results in the introduction of a frameshift mutation at amino acid 48 of the open reading frame, and presumably the incapacitation of any c-myc protein which is translated from the mRNA.

3 pUC13 vector constructs

The -88/+208 and +119/+208 human c-myc IFN fusion fragments were transferred into both the pUC13 and pGC1 (Myers, et al, 1985) vectors. The promoter fusion fragments were excised by cleaving the BSKS+ plasmids with BamHI and HindIII. The fragments were inserted into BamHI and HindIII cleaved vector DNA to produce pGC1(88/+208)IFN, pGC1(+119/+208)IFN, pUC(88/+208)IFN and pUC(+119/+208)IFN.
pUC(88/+208)IFN was cleaved with NotI, BamHI (5'CCGGATCCGG3') linker, BamHI digested and recircularised to create pUC(+23/+208)IFN.

The plasmid pUC(88/+208)IFN was cleaved with PstI, made blunt, BamHI (5'CCGGATCCGG3') linker, BamHI digested and recircularised to create pUCBamΔPstIFN. Note that this construct has also had 5 bp of IFN sequence removed.

Two exonuclease Bal31 deletion series were constructed from the pUC13 based constructs. Conditions were chosen such that ~5-10 bp/min would be removed. The pUC(88/+208)IFN plasmid was linearised with either BamHI or ApaI and digested with Bal31. Aliquots were made blunt and linker with BamHI (5'CCGGATCCGG3') ends. The BamHI series aliquots were then digested with BamHI and HindIII and subcloned into BamHI/HindIII digested pUC13. The ApaI aliquots were made blunt and BamHI (5'CCGGATCCGG3') linker and digested, then recircularised. Plasmid minipreps were screened by radiolabelling ends following BamHI/PstI digestion and sizing by denaturing PAGE. Plasmids with endpoints which appeared to be distributed at ~20 bp intervals were selected and the endpoints determined either by manual dideoxynucleotide sequencing or via an automated sequencing apparatus (Applied Biosystems ABI 370A DNA sequencer). The plasmids sequenced with the automated sequencer are marked with the symbol • below.

The following constructs were isolated from these preparations:

pUC(86/+208)IFN
pUC(64/+208)IFN
pUC(22/+208)IFN
pUC(13/+208)IFN
pUC(+43/+208)IFN
pUC(+53/+208)IFN •
pUC(+69/+208)IFN •
pUC(+97/+208)IFN •
pUC(+100/+208)IFN •
pUC(+112/+208)IFN •
pUC(+150/+208)IFN

4 pUC SVter vector constructs

pUCSVter(88/+208)IFN contains a head to tail duplicated SV40 polyadenylation site (de Wet, et al, 1987) cloned immediately 5' of the human c-myc sequences. The duplicated SV40 sequences were isolated from the plasmid p-210/+72IFNJ3Lucter, a gift of S. Goodbourn, by cleaving the plasmid with ApaLI, blunting the DNA, adding a BglII (5'CCAGATCTGG3') linker, then digesting with BamHI and BglIII. This fragment was cloned into the pUC(88/+208)IFN plasmid at the BamHI site. The construct used for further manipulations contains the insert oriented such that the BamHI site 5' of the SV40 sequences is fused to the BglIII site, and both are thus lost. The BamHI site at the 5' end of the human c-myc sequences is retained.

Several of the pUC based human c-myc /IFN fusions with endpoints 3' of -88 were transferred into the pUCSVter vector. The fusion fragments were excised with BamHI and HindIII and subcloned into BamHI and HindIII digested pUCSVter(88/+208)IFN. Plasmids with the following 5' human c-myc endpoints were constructed: +69; +97; +100; +112; +119; +140. The plasmid (pUCSVterBamAPstIFN) which contains no human c-myc sequences was made by inserting the homologous fragment of pUCBamAPstIFN the pUCSVter vector.

Two linker scan series of the human c-myc sequences covering the range of nucleotides from +94 through +132 were constructed. Five pairs of 54 nt complementary oligonucleotides containing the human c-myc sequences from +94 through +141 (-71 through -22, P2 relative) and BamHI compatible (5'GATC3') ends were synthesised. One set (ls1) contains only
wild type sequence. Each of the others contains a different nine contiguous nucleotides changed to the sequence 5'GAATGCCTC3', such that linkers spanning 36 nucleotides 5' of the TATAA box are introduced in total. Since many transcription factors recognise palindromic sequences, it was felt that the chances of introducing a new transcription factor binding site would be reduced by using an assymetrical sequence. This sequence is, however, a recognition site for the restriction endonuclease BsmI, which cleaves following the second C on the top strand and following the G which is complementary to the first C on the top strand, to produce a two nucleotide overhang. Thus any desired combination of half fragments can be constructed from the members of the linker scan series.

The region of sequence which contains each linker is as follows:

ls1: wild type: +94/+141:P1 or -71/-22:P2
ls2: +124/+132:P1 or -39/-31:P2
ls3: +115/+123:P1 or -48/-40:P2
ls4: +106/+114:P1 or -57/-49:P2
ls5: +97/+105:P1 or -66/-58:P2

and the specific sequences mutated are illustrated below:

ls2
ls3
ls4
ls5
ls1

(Capital letters indicate mutated bases and lower case letters those which are conserved with the wild type human c-myc sequences. Only the sense strand of each linker sequence is shown.)

The 150 ls series was constructed by cloning each of the oligonucleotide pairs into the pUCSVter(+151/+208)IFN plasmid at the BamHI site 5' of the human c-myc sequences. The oligonucleotides were chosen such that
when the introduced BamHI site is taken into account, the sequences to either side of it are maintained with their wild type spacings. The human c-myc sequences represented thus span the region from +91 through +208.

The \(\Delta\)Pst I series contains the oligonucleotides cloned into the pUCSVterBam\(\Delta\)PstIFN plasmid. Thus the constructs contain only the human c-myc sequences within the oligonucleotides, extending from +91 through +141, which is 2 bp 3' of the TATAA box.

An SV40 early region enhancer Eco RI fragment from positions 100 to 298 (Treisman, 1985, gift of R. Treisman) was cloned into both pUC(88/+208)IFN and pUCSVter(88/+208)IFN at the EcoRI site in the pUC polylinker 5' of either the human c-myc sequences or the SV40 polyadenylation sites. The orientation of the fragment was determined to be the same as the human c-myc sequences. The resulting plasmids are called pSVE(88/+208)IFN and pSVESVter(88/+208)IFN.

The plasmid pUCSVterLTRIFN, which contains the Moloney murine leukemia virus (MMLV) long terminal repeat (LTR) acting as a promoter for the IFN sequences within the pUCSVterBam\(\Delta\)PstIFN plasmid was constructed as follows: The plasmid pJ4\(\Omega\) (Morgenstern and Land, 1990b) was linearised with NheI, made blunt and BamHI (5'CCGGATCCGG3') linkered. The MMLV LTR was excised with BamHI and inserted into the BamHI site of pUCSVterBam\(\Delta\)PstIFN. This plasmid also contains some of the polylinker of pJ4\(\Omega\) which is 5' of the BamHI site.

The drug selectable plasmid J6\(\Omega\)hygro was constructed by inserting the 1.1 kb BamHI linkered hygromycin gene fragment of pBam hygro (gift of V. von Hoyningen-Huene) into the BamHI cloning site of the vector pJ6\(\Omega\). This construct expresses inserted genes from the Rat \(\beta\)-actin promoter, and appears to be expressed in almost all cell types (Morgenstern, 1989).
2.1.7.2 RNAase protection probes

The following plasmids were used as templates for in vitro transcription reactions. The expected size of the RNA transcript and protected bands are also presented.

1. Human c-myc 5': The plasmid pSP65 ex1 PvuII-anti (gift of D. Bentley) contains an 862 bp PvuII fragment of the human c-myc gene, cloned into the pSP65 vector. This fragment extends from -349 to +513 bp relative to the P1 promoter. When linearised with SmaI, which cleaves at -101 bp relative to P1, it directs the synthesis of a ~650 bp SP6 polymerase transcribed cRNA probe. When used to protect transcripts which terminate downstream of nucleotide +513, bands of 351 nt (P2 initiated) and 513 nt (P1 initiated) are detected. This probe should protect P1 initiated transcripts of 208 nt and P2 initiated transcripts of 47 nt when used to analyse fusion construct transcripts derived from constructs which end at the human c-myc gene exon I NaeI site. (ie most of the c-myc/β-IFN fusion constructs assayed.) Upstream initiated transcripts should protect 609 nt of probe.

2. Rat c-myc: The plasmid pBSKS+rmycexI (Penn, et al.,) contains a 930 bp PvuII-SacI fragment, which spans exon I and upstream flanking sequences of the rat c-myc gene, cloned into the bluescript vector BSKS+. When cleaved with DdeI this plasmid will direct T3 RNA polymerase to transcribe a probe of 520 nt. Rat c-myc P1 initiated transcripts will protect a fragment of 510 nt, while Rat c-myc P2 initiated transcripts protect one of 350 nt.

3. Myc/β-IFN fusion A: The plasmid pBSKS+ Δ3/1 contains a fragment of the human c-myc promoter from -169 bp to +208 bp relative to P1 fused to a fragment of human β-IFN coding sequence starting at +277 and ending at +639 relative to the normal transcript initiation point of the β-IFN gene. This plasmid was linearised with SmaI, which cleaves at -101
relative to human c-myc P1. It directs T3 polymerase to synthesise a
transcript of 403 nt. This transcript is complementary to 80 nt of β-IFN
sequence and 208 nt of P1 initiated or 47 nt of P2 initiated human c-myc
transcript. Thus this probe should protect 127 nt of P2 initiated or 288 nt of
P1 initiated transcript. Transcripts initiating upstream of the SmaI site
should protect 127 nt of P2 initiated or 288 nt of PI initiated transcript. When the transfected construct
contains less than 101 bp 5' of P1, and transcripts initiate upstream of the
promoter, correspondingly less of the probe will be protected.

Myc/β-IFN fusion B: The plasmid pBSKS+ A2/2 contains a fragment of
the human c-myc promoter from -169 bp to +208 bp relative to P1 fused to
a fragment human β-IFN coding sequence starting at +277 (of the β-IFN
gene) and ending at +688. This plasmid was linearised with SmaI, which
cleaves at -101 bp relative to P1 of the human c-myc gene. It directs T3
polymerase to synthesise a transcript of 442 nt. This transcript is
complementary to 129 nt of β-IFN sequence and 208 nt of P1 initiated or 47
nt of P2 initiated human c-myc transcript. Thus this probe should protect
176 nt of P2 initiated or 337 nt of P1 initiated transcript. Transcripts
initiating upstream of the SmaI site should protect ~426 nt of the probe.
When the transfected construct contains less than 101 bp 5' of P1, and
transcripts initiate upstream of the promoter, correspondingly less of the
probe will be protected.

4. v-myc: The plasmid pBSKS+gagmyc (Penn, et al.) contains the 430
bp BstEII-BamHI fragment of the MC29 retroviral v-gag-myc gene cloned
into the bluescript vector pBSKS+. This segment of the retroviral genome
contains coding sequences specific to the gag portion of the p110v-gag-myc
fusion protein (Reddy, et al, 1983). When cleaved with FnuDII this
plasmid will direct T3 RNA polymerase to transcribe a >135 nt probe. The
probe will protect 135 nt of the v-myc message.
5. β-IFN 3' end: The plasmid pSP3'IF contains a 629 bp NcoI-BglII fragment of the human β-IFN gene cloned into pSP64 (gift of S. Goodbourn). When linearised with PvuII this plasmid will direct the transcription of a >365 nt probe. This probe is complementary to a β-IFN specific, 364 nt region of the transfected myc promoter β-IFN fusion constructs used in this study.

6. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH): The plasmid pBSRatGAP (Penn, et al, 1988) contains a 450 bp ApaI-TaqI internal fragment of the rat GAPDH gene subcloned into the plasmid Bluescript KS+. This plasmid was linearised with the enzyme DdeI. T3 RNA polymerase generates a probe of ~300 nt, which is complementary to ~70 nt of the rat gene. It will also protect an ~35 nt murine fragment.

7. Chloramphenicol acetyl transferase (CAT): The plasmid pSPTK-CAT (Miksicek, et al, 1986) contains a 298 bp EcoRI-PvuII fragment of a herpes simplex virus thymidine kinase/bacterial CAT fusion construct cloned into an SP6 vector. When this plasmid is cleaved with EcoRI, SP6 RNA polymerase will transcribe a 347 nt probe, of which 145 nt are complementary to the CAT transcripts.

8. γ-actin: The plasmid pSP6 γ-actin (Enoch, et al, 1986) contains a BamHI-HindIII fragment from the 3' non-coding region of the human γ-actin gene cloned into the SP64 vector. When this plasmid is cleaved with HindIII, SP6 RNA polymerase will generate a transcript of 145 nt, which is complementary to ~70 nt of rodent or ~120 nt of human γ-actin transcripts.

9. α-globin: The plasmid pSP6α133 (Charnay, et al, 1984) contains human α-globin cDNA sequences -60 to +162 in a BamHI-AvaI fragment cloned into the SP6-PL1 vector. It is cleaved with BamHI. The 270 nt SP6 polymerase transcript overlaps the intron 1 boundary at position +133 and thus protects a 133 nt fragment if it initiates at the correct α-globin start site.
10. MLV LTR/β-IFN: The plasmid pGEM3'LTRIFN contains the SstI/BglII fragment of pUCSVterLTRIFN cloned into the pGEM3 vector. This plasmid can be linearised by digestion with EcoRI. It will direct the transcription by T3 RNA polymerase of an ~500 nt probe which will protect a 435 nt fragment initiating from the MLV LTR promoter.

11. BSKS+SspI: When digested with SspI, Bluescript KS+ acts as a template for T3 RNA polymerase to generate a transcript of ~750 nt. This probe is complementary to ~680 nt between the BamHI site in the polylinker and the SspI site at position 17 within the Bluescript vector.

2.1.8 Gel retardation probes

2.1.8.1 Plasmid fragments

Gel retardation probes were excised from plasmids by Bam HI/Pst I digestion. The coordinates of the fragments with respect to the human c-myc P2 mRNA cap site are indicated in brackets:

- pUCSVter(+97/+208)IFN [-66/+49]
- pUC(+97/+208)IFN [-66/+49]
- pUCSVter(+100/+208)IFN [-63/+49]
- pUCSVter(+112/+208)IFN [-51/+49]
- pUCSVter(+119/+208)IFN [-44/+49]
- pUCSVter(+150/+208)IFN [-13/+49]
- pUCSVter(+97/+181)IFN [-66/+22]

Double strand fragments were made single stranded by heating to 95°C for 2 minutes followed by rapid cooling on ice.

2.1.8.2 Synthetic oligonucleotides

Synthetic oligonucleotides were used as both single and double strand gel retardation probes and competitors. The sequence of each of these oligonucleotides and the ICRF synthesis reference numbers are indicated below:

<table>
<thead>
<tr>
<th>Name</th>
<th>ICRF ref #</th>
<th>Sequence, 5' to 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

81
Human c-myc P2 promoter oligonucleotides:

-66/+47 SENSE 8914 GATCCGG GCCGAAA
AGAACGGAGG GAGGGATCGC GCTGAGTATA AAAGCGGT TTTCAGGGCTT
TATCTAATCT GCTGTAGTA AAAGCGGT TTTCAGGGCTT

-66/+47 ANTISENSE 8915 GAGGGCTGCT GTGGGCCTGGCTT
CTGCTCTCTG CTGGAGTTAC TACACGGAGT TAGATAAAGC CCCGAAAACC
GGCTTTTATA CGCTCGGCTG GCTGAGTATA AAAGCGGT TTTCAGGGCTT

-45/+41 ANTISENSE 10836 GCTGCTCCCT CTGCTCTCTG
CTGCTCTCTG CTGGAGTTAC TACACGGAGT TAGATAAAGC CCCGAAAACC
GGCTTTTATA CGCTCGGCTG GCTGAGTATA AAAGCGGT TTTCAGGGCTT

-45/+31 ANTISENSE 10837 CAGGGT TAGATAAAGC CCCGAAAACC
TACAGGGGCTG TAGATAAAGC CCCGAAAACC

-21 to +47 SENSE 10158 GATCCGGGG CTTGGGGCCTT
TATCTAATCT GCTGTAGTA AAAGCGGT TTTCAGGGCTT

-21 to +47 ANTISENSE 10157 GAGGGCTGCT GTGGGCCTGGCTT
CTGCTCTCTG CTGGAGTTAC TACACGGAGT TAGATAAAGC CCCGAAAACC

-21 TO -1 SENSE 9884 GATCCGGGG CTTGGGGCCTT
TATCTAATCT GCTGTAGTA AAAGCGGT TTTCAGGGCTT

-21 TO -1 ANTISENSE 9887 TAGATAAAGC CCCGAAAACC

+1 TO +22 SENSE 9886 GATCCACTCG CTGTAATGTAAT
TACACGGGCTG CA

+1 TO +22 ANTISENSE 9885 CAGGGT TAGATAAAGC CCCGAAAACC
TACACGGGCTG CA

-20 to +23 SENSE 9961 GATCCGGGTT TTCGGGGCCTT
TATCTAATCT GCTGTAGTA AAAGCGGT TTTCAGGGCTT

-20 to +23 ANTISENSE 9960 CAGGGT TAGATAAAGC CCCGAAAACC

linker scan series:
LS1 SENSE 5085 GATCTGGCTT GCGGGAA
AGAACGGAGG GAGGGATCGC GCTGAGTATA AAAGCG

LS1 ANTISENSE 5086 GATCCCGTTC TATATCAGC
GCGATCCCTC CCTCGGTT TTTCAGGGC

LS2 SENSE 5087 GATCTGGCTT GCGGGAA
AGAACGGAGG GAGGGATCGC GCTGAGTATA AAAGCG

LS2 ANTISENSE 5088 GATCCCGTTC TATATCAGC
TTCATCCCT CCTCGGTT TTTCAGGGC
LS3 SENSE 5089 GATCTGGCTT GGCGGGAAAA
AGAACGGAGA ATGCCTCCGC GCTGAGTATA AAAGCG

LS3 ANTISENSE 5090 GATCCGCTTT TATACTCAGC
GCGAGGCATT CCTCCGTTCT TTTTCCCGCC AAGCCA

LS4 SENSE 5091 GATCTGGCTT GGCGGGAAAG
AATGCCTCGG GAGGGATCGC GCTGAGTATA AAAGCG

LS4 ANTISENSE 5092 GATCCGCTTT TATACTCAGC
GCGATCCCTC CGAGGCATTCC TTTTCCCGCC AAGCCA

LS5 SENSE 5093 GATCTGGCTT GAATGCCTCA
AGAACGGAGG GAGGGATCGC GCTGAGTATA AAAGCG

LS5 ANTISENSE 5094 GATCCGCTTT TATACTCAGC
GCGATCCCTC CCTCCGTTCT GAGGGATCGC AAGCCA

Melal site (Hall, 1990):
WT ME1A1 TOP 10160 GATCGGGAGGGGAGGGA
WT ME1A1 BOTTOM 10159 GATCTCCCTCCCCTCCC
MT1 ME1A1 TOP 10161 GATCGGGAGGGGATGTA
MT1 ME1A1 BOTTOM 10162 GATCTACATCCCCTCCC
MT2 ME1A1 TOP 10163 GATCGGGAGTGTAGGGA
MT2 ME1A1 BOTTOM 10164 GATCTCCCTA CACTCCC

Adenovirus E2 promoter oligonucleotides:
E2F TOP 3199 TCGACGTAGT TTTGCCGCTT
AAAGCATG

E2F BOTTOM 3198 CTTTAAGCGCGAAAACTACG

E2F-E2F TOP 3203 TCGACGTAGT TTTCCGCGCTT
AAATTTGAGA AAGGGCAGGA AACTAGCATG

E2F-E2F BOTTOM 3202 CTAGTTTCGC GCCCTTTCTCA
AAATTTAAGCG CGAAAACCTAC G

mE2F-mE2F TOP 3197 TCGACGTAGT TTTGGGGGCTT
AAATTTTGAAGA AAGGGCCCA AACTAGCATG

mE2F-mE2F BOTTOM 3196 CTAGTTTGGG GCCCTTTCTC
AAATTTAAGCG CCCAAAACCTAC G

2.1.9 Antibodies

The following antibodies were used:
mAb1: Monoclonal antibody PM3E7.C4 raised to the xmyc1 peptide APSEDIWKHELL (pep1) which corresponds to human c-myc protein amino acid positions 44-55 (Evan, et al, 1988). Both caprylic acid IgG and FPLC purified caprylic acid IgG fractions were used.

mAb2: Monoclonal antibody Myc13C7.B1 raised to the C59B peptide CSTSSLYLDLSAAASEC (pep2) which corresponds to human c-myc protein amino acid positions 173-188 (Evan, et al, 1985). The epitope which this antibody recognises has been mapped to the sequence SSLSYLDLSAAASEC (G Evan, personal communication). Caprylic acid IgG fractions were used.

Control mAb: An IgG preparation of mAb38 (Dransfield and Hogg, 1989), which recognises the LFA-1 protein α chain, which is biologically unrelated to myc protein, was used as a control mAb.

2.2 Methods
2.2.1 Tissue culture
2.2.1.1 General

All cell lines were cultured on plastic tissue culture dishes of various sizes from 6.4 mm diameter, 0.32 cm² growth area through 150 mm diameter, 707 cm² growth area, depending on the number and density of the cells being grown. All cell lines were cultured in Dulbecco's modified Eagle's media (DMEM) supplemented with 100 μg/ml kanamycin, 2 μg/ml gentamycin and 10% v/v foetal calf serum (or exclusively in the case of Ψ2 cells, with 10% new born calf serum). Incubators were maintained fully humidified at 37°C, 10% v/v CO₂ in air. All manipulations of living cells were performed in laminar flow hoods.

Cells were routinely passaged every three to four days. They were split by first washing 1x with PBS.A then briefly (1 - 2 mins) treated with 1:3 v/v trypsin:versene and subsequently diluted 1/10 or 1/20 into fresh media.
When counting was necessary, the cells were passed 2x through a 19 or 21 gauge needle following trypsinization and 0.2 ml diluted into 10 ml of isoton for counting on an automatic cell counter (Coulter Electronics). The appropriate number of cells were then manipulated as needed.

Cells were stored in 1 ml aliquots under liquid nitrogen. The freezing procedure consisted of a brief trypsinization followed by resuspension in an ice cold freezing mixture containing 50% v/v DMEM media, 40% v/v FCS and 10% v/v DMSO for all lines except NIH-3T3's. NIH-3T3's were resuspended in the DMEM/FCS mixture and then the DMSO was added dropwise. Cells were wrapped in paper towelling and slowly frozen at -70°C. After at least 2 days at -70°C they were transferred to liquid nitrogen.

Cells were recovered from liquid nitrogen freezing by rapidly thawing them to 37°C and slowly adding 25 ml of fresh media. After the cells had attached to the tissue culture dish (usually 2 - 3 hrs) the media was changed. Cells were then passaged as before.

Cell colonies were cloned by rinsing the dish 2x with PBS-A and placing a sterile 5-10 mm diameter glass cylinder coated on its bottom edge with sterile vacuum grease around the colony. Cells were removed from the dish by brief application of trypsin:versene and then placed into 1.5-3 cm diameter tissue culture dishes.

2.2.1.2 Transfection

Transfections were performed using a modification of the procedure of Wigler (1979b). In stable transfections, typically 10 μg of an indicator plasmid and 0.5 μg of a selectable marker plasmid were cotransfected into a subconfluent dish of cells. Transient transfections also contained 10 μg of DNA, but lacked the selectable marker, however they often contained a second plasmid which expressed a control gene.

The cells were seeded at a density of 7×10^5 cells per 10 cm diameter dish. Eighteen to twenty-four hours later, the media was replaced with 5 ml of
fresh DMEM/10% FCS. A calcium phosphate precipitate was formed by combining the DNA with 0.5 ml transfection buffer (20 mM Hepes (pH 6.95), 137 mM NaCl, 5 mM KCl, 0.7 mM Na2PO4, 6 mM glucose). Twenty-three to twenty-seven μl of 2.5M CaCl2 was added dropwise to the TFB/DNA mix while flicking the tube rapidly. Within 5-15 minutes a precipitate formed, as evidenced by a slight opalescent tinge to the mixture. This precipitate was then added to the cells, and gently mixed with a rocking motion to ensure even distribution across the plate.

The precipitate was allowed to remain on the cells for several hours, the specific time depending on the cell line used. (Rat-1's for 5 - 7 hrs, NIH-3T3's for 4 - 6 hrs, Ψ2's for 4 - 8 hrs and HeLa cells for 18 - 24 hrs.) The precipitate was removed by washing 2x with prewarmed PBS.A. The cells were then osmotically shocked by the addition of a prewarmed 15% v/v glycerol solution in PBS.A. The duration of the shock varied with the cell line being transfected. (Rat-1's, NIH-3T3's and Ψ2's for 3 mins, HeLa cells for 45 secs.) The cells were then washed an additional two times with PBS.A and fed with fresh media.

For stable colonies the cells were split into fresh media by a factor of 1/10 or 1/20 one day after transfection. On the next day (or two days later for puromycin selected colonies) fresh media was added containing the drug of choice. The following combinations of cell lines and drugs were found to give many surviving transfected colonies with no surviving background cells:

<table>
<thead>
<tr>
<th>Cell line</th>
<th>G418 (mg/ml)</th>
<th>Hygromycin (μg/ml)</th>
<th>Puromycin (μg/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rat-1</td>
<td>1</td>
<td>150</td>
<td>2.5</td>
</tr>
<tr>
<td>NIH-3T3</td>
<td>1</td>
<td>50 - 100</td>
<td>1.0</td>
</tr>
<tr>
<td>Ψ2</td>
<td>1</td>
<td>300</td>
<td>not done</td>
</tr>
<tr>
<td>HeLa</td>
<td>1</td>
<td>150</td>
<td>not done</td>
</tr>
</tbody>
</table>
Transfections were refed every three to four days for up to two weeks until colonies appeared. These were then picked or pooled as appropriate.

Transient transfections were usually harvested two or three days after application of the precipitate, and were split only if a portion of the cells were required as a control.

2.2.1.3 Retroviral manipulations

Retroviral preparations were prepared and stored using commonly established techniques (Morgenstern and Land, 1990a). Producer cell lines were made either by transfecting \( \Psi^2 \) cells directly, and then picking colonies, or by transiently transfecting an amphotropic producer line (GP+E or \( \Psi^{-}\)am) and infecting \( \Psi^2 \) cells with the virus containing supernatant. Infected \( \Psi^2 \) producers commonly give 10-fold higher titres on average than do transfected ones (J. Morgenstern, personal communication).

Retrovirus production

Producer cells were grown to confluence on 9- or 15-cm diameter dishes. Half the normal working volume of DMEM/10%NBCS media was put on the dishes. 18-24 hrs later the media was removed and passed through a 0.22- or 0.45 \( \mu \)m syringe filter. The viral supernatant was aliquoted and immediately frozen for storage at -70\(^\circ\)C. A small aliquot was taken for titration.

Retroviral infection

Cells which were to be infected for the isolation of stable colonies were seeded at low density, usually 1 to 2 \( \times \) 10\(^5\) cells/9 cm diameter dish, 18 to 24 hours prior to infection. Viral supernatant was diluted into an infection mixture containing DMEM/10% FCS and 8 \( \mu \)g/ml polybrene. The normal volume of media was replaced with 1/5 volume of the infection mixture. The cells were incubated with the virus for 1.5 to 2 hrs, after which the
supernatant was removed and replaced with fresh media. The next day
the media was changed to selective media. The cells were not usually split
again, being refed at three to four day intervals until colonies appeared.
Colonies were then pooled and passaged normally. After pooling the cells
were not normally maintained in selective media, but on an occasional
series of passages selection might be applied, and little additional cell death
was ever observed.

Retrovirus titration
Viruses were usually titred on NIH-3T3 cells. $7 \times 10^5$ cells/9 cm
diameter dish were infected with a series of dilutions of the virus in 2 ml
DMEM/10% FCS containing 8 $\mu$g/ml of the polyanion, polybrene.
Infection was allowed to proceed for 1.5 to 2 hrs, at which time the virus
solution was replaced with fresh media. 18 - 24 hrs later the cells were
split into selective media giving a final range of dilution of the virus from
10^{-3} to 10^{-6}. The plates were refed at 3 to 4 day intervals, until resistant
colonies grew out. Colonies were stained with 0.1% w/v crystal violet,
20% v/v methanol in H_2O for at least 5 mins then rinsed with tap water
and air dried. Virus titre/ml was calculated as (viral dilution)(cell
splitting factor)(# of colonies)/2 [for 2 ml volume of infection].

2.2.2 General analytical techniques
2.2.2.1 Quantification
DNA was quantitated using one of two methods, optical density (OD or
absorbance, A) or fluorimetry.

RNA was quantitated using OD.

Optical density: An appropriate dilution of nucleic acid solution
(usually 1/200) was made into H_2O. The OD at 260 and 280 nanometers
was then measured using a spectrophotometer. The concentration was
calculated using the formula [nucleic acid] = (A_{260})(dilution)(\Sigma), where
A_{260} is the absorbance at 260 nanometers wavelength, and \Sigma = 50 $\mu$g/ml for
dsDNA, 33μg/ml for ssDNA and 40 μg/ml for RNA (the concentration at an OD of A = 1.0). The relative purity of the nucleic acid was checked by calculating the $A_{260}/A_{280}$. If this value was less than 2.0 for DNA or 1.8 for RNA, the sample was considered pure enough to use.

Fluorimetry: DNA concentration was calculated by measuring the fluorescence of the sample in the presence of Hoechst dye at x mg/ml in 0.1x SSC. A standard curve using 1 to 100 ng DNA was used to calibrate the sample readings.

### 2.2.2.2 Precipitation

Nucleic acids were recovered from aqueous solution by one of the following protocols:

Ethanol precipitation (EtOH ppt): The nucleic acid was brought to 0.3M NaOAc, pH 5.2, 70% v/v ethanol, 20μg glycogen (a carrier), incubated on dry ice for 10 min and spun at 10,000g, 4°C for 5 to 10 min. The supernatant was discarded and the pellet washed with 70% v/v ethanol.

Ammonium acetate/ethanol precipitation (NH$_4$OAc/EtOH ppt): When it was necessary to precipitate nucleic acids but not free nucleotides, the NaOAc was substituted with 2M NH$_4$OAc, and the sample treated as above.

Spermine precipitation: This procedure was used when the DNA was not going to be resuspended in H$_2$O. The DNA solution was brought to a final concentration of 3mM spermine in the presence of 20 μg tRNA carrier. The sample was incubated for 10 min on ice and spun at 10,000g, 4°C, 10 min. The supernatant was discarded and the pellet resuspended in the appropriate solution.

### 2.2.2.3 Phenol extraction

Protein was removed from aqueous solution by phenol (f) and phenol/chloroform (fc) extraction. An equal volume of buffer saturated f or fc was added to the sample, it was mixed vigorously on a vortex mixer...
for 10 sec and centrifuged at 10,000g for 1 min. The upper, aqueous phase was saved, and the lower, organic one discarded. Excess phenol was removed by chloroform extracting the sample. This procedure was usually followed by ethanol precipitation.

2.2.2.4 Agarose gel electrophoresis

DNA molecules that ranged in size from 100 bp to 20 kbp were resolved using agarose gel electrophoresis. High (HMP) or low melting point (LMP) agarose was dissolved at the appropriate concentration (0.5% to 1.25% w/v depending on the sizes of the fragments to be resolved) in 1x TAE buffer, melted by heating in a microwave oven, and cast to a thickness of about 7 mm using a plexiglass horizontal gel tray and a teflon sample comb. After solidifying, the gel was submerged in 1x buffer and the comb then removed. Samples were prepared by the addition of 1/10th volume of stop and load dye (50 mM EDTA, 100 mM Tris 8.0, 50% v/v glycerol, 0.4% w/v bromophenol blue). Samples and markers were loaded into the wells and then subjected to a constant voltage electric field of 5V/cm or less, until the marker dyes had migrated the required distance. The gel was then stained with ethidium bromide (EtBr, 0.1 μg/ml), and the bands visualized on a UV light box. High background was reduced by destaining in H2O for at least 15 min. It was observed that bands could be tightened by running the gels for an additional 5 min after staining. This was especially true for LMP agarose gels, and others gels which had initially been run very fast. Gels were photographed using 254 nm illumination using Polaroid type 57 or type 55 black and white instant print film and a Red gelatin filter.

The DNA was recovered from preparative gels by one of two methods. If the presence of low concentrations of agarose was unimportant for subsequent use, as with most enzymatic reactions such as ligations or random priming, then the gel was made of LMP agarose and the gel
simply blotted dry and the band(s) excised with a razor blade while the gel was being illuminated from below on a long wave (365 nm) UV light box. If it was necessary to recover very pure fragments, (e.g. for bandshift reactions) then a high melting point agarose gel was used. The gel was blotted dry and a piece of DE81 paper inserted in a slit made just downstream of the band of interest. The gel was run for an additional few minutes, until the DNA had migrated onto the paper. The paper was then removed from the gel and incubated in 400 µl TE 7.5, 2M NaCl at 65°C, 5min, with periodic vortex mixing. The entire slurry was then transferred to the top of a 2 ml eppendorf tube which contained a 0.45 µm nitrocellulose filter in a plastic insert (Costar Spin-X). The sample was spun at 10,000g, RT for 1 min, by which time the DNA solution was in the bottom of the tube. The filter insert was discarded and the DNA precipitated by the addition of 1 ml ethanol and 20 µg glycogen. The sample was reprecipitated, ethanol washed, and resuspended in TE.

DNA standards used as size markers were as follows:

λ/HindIII digest (kb): 23.1, 9.4, 6.6, 4.4, 2.3, 2.0, 0.564, 0.125.

ϕX174/HaeIII digest (bp): 1353, 1078, 872, 603, 310, 281, 271, 234, 194, 118, 72.

1 kb ladder: (kb) 12.2, 11.2, 10.2, 9.2, 8.1, 7.1, 6.1, 5.1, 4.1, 3.1, 2.0, 1.6, 1.0, (bp) 517, 506, 396, 344, 298, 220, 201, 154, 134, 75.

123 bp ladder (bp): 4182 to 123 in 123 bp steps.


pBR322 Hae III (bp): 587, 540, 504, 458, 434, 267, 234, 213, 192, 184, 124, 123, 104, 89, 80, 64, 57, 51, 21, 18, 11.

2.2.2.5 Native polyacrylamide gel electrophoresis (PAGE)

Nondenaturing polyacrylamide gels were used for analytical and preparative separations of dsDNA fragments less than 500 bp in length,
and of protein:DNA complexes. 4% to 12% w/v polyacrylamide gels ((29:1 w:w acrylamide:N,N'-methylene bisacrylamide), 1/2x TBE, 0.1% v/v TEMED, 0.1% w/v ammonium persulphate) of 0.4 mm to 2.5 mm thickness were run at constant voltage of less than 10V/cm on a vertical gel apparatus.

Nonradioactive DNA bands were visualised as discussed for agarose gels. Required bands were cut out of the gel with a razor blade, and eluted from the gel fragment by overnight incubation at 37°C in 400 µl 700 mM NaCl, 10 mM Tris 7.5, 1 mM EDTA with occasional mixing. The gel fragments were spun out (10,000g, 1 min, RT), and the DNA precipitated from the supernatant by the addition of 3 volumes ethanol and 20 µg glycogen. Radioactive complexes were identified and eluted using the appropriate conditions

2.2.2.6 Denaturing PAGE

Denaturing polyacrylamide gels were used in a variety of procedures. These include DNA sequencing, RNA probe purification, RNAase protection analysis and oligonucleotide purification. 6% to 20% w/v polyacrylamide gels ((19:1 w:w acrylamide:N,N'-methylene bisacrylamide), 50% w/v urea, 1x TBE, 0.1% v/v TEMED, 0.1% w/v ammonium persulphate) were used on a vertical gel apparatus. Gels were run at constant power of approximately 50 milliwatts/cm² so as to maintain a surface temperature of 50 to 60°C. An aluminium plate was clamped to the front of the gel to ensure even heat distribution, and thus prevent "smiling" of the bands. Analytical gels were fixed (if required) for 10 min in 40% v/v methanol, 10% v/v glacial acetic acid and dried onto Whatman 3mm paper using a vacuum assisted, heated (80°C) gel drier. In preparative situations fragments were identified either by autoradiography or UV shadowing and eluted in the appropriate buffer as required.
2.2.2.7 Densitometry

Autoradiograms were scanned using an LKB laser densitometer... Parameters were usually set as follows: scan width: 1600 µm; Y-step: 100 µm; Baseline: average of 16 lowest points; Peak search: width 10, number 12, minimum 0.5%; Smoothing: 5 point moving boxcar average. Absolute peak heights were always used, as comparisons tended to be made between bands identified with the same probe in parallel reactions, and gave similar profiles. This also obviated the need to define arbitrary shapes to individual peaks. Any readings greater than 2.0 or less than 0.10 were treated with caution, as they tended to be outside of the linear range of the film or possibly subsumed within non-specific background. Whenever possible multiple exposures which had been preflashed before autoradiography were scanned to arrive at values within the middle range of intensities. Variations of the same samples between gels could be several fold in extreme circumstances.

2.2.3 DNA
2.2.3.1 Competent bacteria

This procedure for making competent bacteria was communicated by M. Scott (unpublished).

A single colony was picked from a fresh overnight BHI agar plate into 20 ml prewarmed TYM media in a 250 ml dH2O rinsed flask. It was grown (37°C, shaking vigorously) to an OD600 of 0.2-0.8. The culture was added to 100 ml prewarmed TYM in a 2 l rinsed flasked and grown (37°C, vigorous shaking) to an OD600 of 0.5-0.9. 500 ml prearmed TYM was then added to the flask, and growth continued to an OD600 of 0.6. The flask was then swirled in an ice water bath to rapidly cool the culture. The bacteria were pelleted (15 min, 4.2 k rpm, 0°C, Beckman J6 centrifuge with a swing-out rotor), and the supernatant discarded. The pellet was resuspended in 100 ml sterile Tfb I (30 mM KOAc, 50 mM MnCl2, 100 mM KCl, 10 mM CaCl2,
15% v/v glycerol, in dH2O) by gentle swirling on ice. The bacteria were pelleted again (as above, but an 8 min spin). The supernatant was again discarded and the pellet resuspended in 20 ml sterile, precooled Tfb II (10 mM Na-MOPS pH 7.0, 75 mM CaCl2, 10 mM KCl, 15% v/v glycerol in dH2O) by gentle swirling on ice. Aliquots of 0.1 to 0.5 ml were frozen on dry ice and stored at -70°C.

2.2.3.2 Bacterial transformation

Plasmids were transformed into bacteria by combining 50 µl to 100µl of transformation competent bacteria, which had been thawed and incubated on ice for not less than 15 min, with less than 1/5 volume of the DNA preparation, containing up to 50 ng of DNA. The mixture was incubated on ice for at least an additional 15 min, heat shocked for 5 min at 37°C or 45 sec at 42°C, and returned to ice for 1 min. 10 volumes of SOC media were added, and the suspension incubated at 37°C for at least 45 min. The bacteria were spun out of suspension (3000g, 1 min) and the pellet resuspended in 100 µl SOC media. This was then plated to BHI agar plates containing the appropriate drugs. 100 µg/ml ampicillin or 50 µg/ml kanamycin, final concentrations, were routinely employed. Plates were incubated at 37°C until colonies grew out (usually overnight).

2.2.3.3 Plasmid minipreps

Individual colonies were picked off of selective plates and grown (37°C, 200-300 rpm, overnight) in 2 ml selective BHI media. 1.5 ml were taken into an eppendorf tube, and the bacteria pelleted (10,000g, 1 - 2 min, RT). The pellet was resuspended in 100 µl Solution I (50 mM glucose, 25 mM Tris pH 8.0, 10 mM EDTA) by vigorous vortexing and incubated 5 min at RT. 200 µl Solution II (200 mM NaOH, 1% v/v SDS) was added, the contents mixed by rapid inversion/flicking, and the tube transferred to ice for 3 min. 150 µl ice cold Solution III (5M KOAc pH 4.8) was added, the tube again mixed well, and incubation continued for 3 min on ice. The
cellular debris was pelleted (10,000g, RT, 1 min) and 400 μl of supernatant precipitated with 1 ml ethanol. After mixing, the sample was immediately pelleted (10,000g, RT, 1 min), the supernatant discarded and the DNA pellet washed with 70% v/v ethanol. The pellet was resuspended in 30 μl TE 7.5, and stored at 4°C or used immediately. It was empirically found that so long as the Solution II treatment was kept short, and the sample spun immediately following precipitation, then further treatments such as phenol extraction were not required for most restriction enzymes to be able to cut the DNA. Certain enzymes always required additional purifications. These were generally those which cut in low salt buffers, or had a very GC rich recognition sequence. If an enzyme did not cut the first time, then the DNA was phenol/chloroform extracted and ethanol precipitated before redigesting.

Analytical restriction digests were performed using 1 μl of miniprep DNA if the plasmid had a high copy pUC replication origin (e.g. pUC, Bluescript, pGC1), or 3-5 μl of DNA with a low copy origin (e.g. pBR322, J vectors). Digests were performed in 10 μl total volume, using 3-5 fold excess restriction enzyme, for 1 hr at the appropriate temperature. Reactions were halted by the addition of 1/10 volume stop and load dye (50 mM EDTA, 100 mM Tris 8.0, 50% v/v glycerol, 0.4% w/v bromophenol blue) supplemented with 2 μg/ml RNAase A. They were then analysed via agarose or native PAGE.

2.2.3.4 Plasmid maxipreps

A fresh overnight colony of bacteria was picked into 2 ml BHI media with antibiotics, and grown overnight (37°C, shaking). 0.5 ml was inoculated into 100 ml BHI media without antibiotics in a 500 ml flask. The culture was grown overnight (37°C, shaking), the bacteria pelleted (2x 10 min, 4000g, 4°C, 50 ml plastic tube), and the supernatant discarded. The pellet was resuspended in 5 ml Solution I (50 mM glucose, 25 mM Tris pH
8.0, 10 mM EDTA) by vortexing, and incubated 5 min at RT. 10 ml Solution II (200 mM NaOH, 1% v/v SDS) was added, mixed by rapid inversion until the opacity cleared, and the lysate kept on ice no more than 3 min. 7.5 ml ice cold Solution III (5M KOAc pH 4.8) was added, and mixed by rapid inversion. The bacterial flocculant was pelleted (10 min, 4000g, 4°C), and the supernatant (about 20-25 ml), precipitated by adding to 20 ml isopropanol. The nucleic acid was recovered by centrifugation (10 min, 4000g, 4°C), the pellet rinsed with 70% v/v ethanol, drip dried, and resuspended in 5 ml EtBr/CsCl solution (120 g CsCl in 100 ml dH₂O + 10 ml 5 mg/ml EtBr). A precipitate of RNA and protein crud usually formed, and was cleared from suspension (10 min, 4000g, 4°C). The supernatant was transferred into a heat-sealable tube (Beckman), and spun in an ultracentrifuge to band the DNA (65,000 rpm 4 hr or 55,000 rpm overnight in a Vti65.2 Beckman vertical rotor, 20°C). The supercoiled DNA was identified as the lower of two bands which appeared in the tube under long wave (365 nm) UV illumination. The band was pulled by first punching a hole in the top of the tube, and then extracting the DNA with a 19 g needle. The DNA was reband under the same centrifugation conditions. The tube being topped up this time by the addition of CsCl solution without additional EtBr (120g CsCl in 110 ml H₂O). The plasmid was recovered by dripping the band into a 15 ml polypropylene tube. (Hole punched in bottom first, then in top). The volume (usually 0.5-1 ml) was brought to 2 ml by adding TE 7.5. 5 ml ethanol was added to precipitate the DNA. It was pelleted (10 min, 6000g, RT), 70% v/v ethanol washed, and resuspended in 400 µl TE 7.5. The solution was phenol/chloroform extracted, ethanol precipitated/washed, and the pellet resuspended in TE 7.5. The plasmid yield was quantitated by OD, and diagnostic restriction digests performed. Plasmid DNA was stored at 4°C, short term, or -20°C, long term.
2.2.3.5 Restriction digests

Analytical digests were performed using 0.5 to 1 μg of DNA and a 3 to 5 fold unit excess of enzyme in a total volume of 10 μl. Buffer conditions were as specified for each enzyme in the New England Biolabs catalog. If New England Biolabs did not stock a particular enzyme, the buffer conditions used were those provided by the manufacturer. If multiple enzymes were to be used on the same sample, the enzyme with the conditions requiring the lowest salt concentration was used first, and the digestion allowed to go to completion. NaCl was added to adjust conditions to those of the next enzyme, enzyme added, and the digestion was completed. If multiple incompatible reaction conditions were required, the volume was increased to 100 μl with TE 7.5, the sample phenol extracted and ethanol precipitated. The sample was then resuspended in the appropriate buffer for the next enzyme, and digested accordingly. Reactions were allowed to proceed for at least an hour, and then stopped by the addition of 1/10 volume of stop and load dye. Samples were then analyzed by gel electrophoresis.

Preparative reactions were performed similarly to analytical ones, except that the volumes were scaled up such that the final glycerol content was never more than 5% v/v (enzyme stocks contain 50% v/v glycerol). Final DNA concentrations were between 50 and 200 μg/ml. For very large preparative digests of >100 μg of DNA, if the enzyme was stable enough (New England Biolabs catalog), less enzyme was used and the reaction allowed to proceed overnight. Digested DNA was phenol extracted, ethanol precipitated, and the desired fragments isolated by the appropriate procedure.

2.2.3.6 Ligations

DNA fragments for ligation were usually prepared by digesting plasmid DNA with the appropriate restriction endonuclease modifying the DNA
as required (see below) and then purifying the fragments using a low melting point agarose gel. Enough starting DNA was used to provide approximately 1 µg of isolated fragment in 50 µl of LMP agarose. The specific strategy used for ligating DNA molecules together depended on both the vector and insert ends.

For compatible, overhanging ended ligations, 1 µl (~20 ng) vector was combined with 9 µl (~200 ng) insert in 1x ligation mix (50 mM Tris pH 7.8, 10 mM MgCl₂, 1 mM ATP, 20 mM DTT) in a total volume of 20 µl, to which 0.1 µl ligase (6 Weiss units) was added. The LMP agarose was melted at 65°C and cooled to 37°C before addition, and the final LMP agarose concentration was kept to <0.4% w/v. Incubation was at RT, from 2 hr to o/n. The mix was then heated to 65°C to melt residual agarose and 10 µl used to transform 50 to 75 µl competent bacteria.

Blunt ended molecules were either ligated directly [using 1 µl (60 Weiss units) ligase and incubating o/n, RT], or they were modified by the addition of oligonucleotide linkers, which would provide compatible overhanging ends. Incompatible ends were made flush and then linkerered to provide compatible ends.

Linkered ends were prepared as follows: 5 µg DNA was cut with the appropriate restriction enzymes, phenol extracted, and ethanol precipitated/washed. If the above restriction digest generated blunt ends the following step was omitted, otherwise 5' and 3' overhang ends were made flush by resuspending in 20 µl 1x blunting mix (20 mM Tris pH 7.5, 5 mM MgCl₂, 10 mM β-ME, 200 µM each dNTP) and adding 1 µl (10-15 units) *E. coli* DNA polymerase I (Klenow fragment). For 5' overhanging ends, incubation was for 30 min, RT. 3' overhangs were incubated for 30 min at 37°C. The blunting reaction was stopped by heating to 65°C, 5 min. Linkers (0.33 µg), ligase (1 µl, 60 Weiss units) and 10x post blunting mix (4 µl, 500 mM Tris pH 7.8, 50 mM MgCl₂, 10 mM ATP, 200 mM DTT) were
added as the reaction was brought to 40 µl total volume. Ligation of the linkers was allowed to proceed o/n, RT. The ligase was denatured by heating, 68°C, 10 min, the volume brought to 100 µl with TE 7.5, and the sample phenol extracted. The solution was then brought to 1 ml in the appropriate NACS loading buffer, and applied to a preequilibrated NACS column (Nucleic Acid Chromotography System, an ion exchange resin, BRL, protocol as supplied by the manufacturer) Following washing and elution from the NACS column the DNA was ethanol precipitated/washed and digested with the appropriate restriction enzyme for the sites contained within the linkers (~100-150 units enzyme, >2 hr). The sample was then phenol extracted, ethanol precipitated/washed and purified on a LMP agarose gel. The linked fragment was then used in an overhanging, compatible end ligation.

If the ends of the vector could ligate together and form a closed circle in the absence of inserts, they were dephosphorylated by Calf Intestinal Alkaline Phosphatase (CIP) treatment. 3 unit CIP/µg DNA were added to the restriction digest. The reaction was incubated at 37°C, 30 min and stopped by the addition of EGTA to a final concentration of 50 mM (500mM stock). Incubation was continued for 15 min, 68°C, following which the reaction volume was brought to 100 µl with TE 7.5, and phenol extracted 2x, ethanol precipitated/washed, and gel purified.

2.2.3.7 Oligonucleotide manipulations

Oligonucleotides were synthesized on an Applied Biosystems Model 370 oligonucleotide synthesizer. Oligonucleotides were resuspended at 4 mg/ml in TE prior to purification. Typically, 20 µg was purified at one time. 5 µl of DNA was combined with an equal volume of deionised formamide. This was then denatured at 95°C, 5 min, and placed on ice. Full length oligo was resolved from incomplete strands using denaturing PAGE. Acrylamide concentrations from 6% to 10% w/v were used.
depending on the length of the oligo. Marker dyes were usually loaded into lanes next to the sample, as they often ran in the same position as the DNA and would otherwise obscure its visibility. Following electrophoresis the gel was transferred between two pieces of Saran Wrap, shadowed with a 254 nm UV light source on white paper towelling, the band of interest marked and excised from the gel. The oligo was eluted from the gel slice by overnight incubation in 500 mM NH₄OAc, 10 mM MgCl₂, 1 mM EDTA, at 37°C and then ethanol precipitated. DNA was resuspended in TE and the concentration measured by UV spectrophotometry (1 OD₉₀₀ unit=33 mg/ml).

Oligos which were used for cloning were annealed by combining equimolar amounts (usually about 1 μg) of each strand in TE, 10 mM MgCl₂ in the smallest feasible volume. The sample was heated to 90°C for 3 min, transferred to a beaker containing about 50 ml of the hot water, and allowed to slowly cool to room temperature (usually 1 to 2 hr). If the oligo to be cloned had the same sequence at either end, it was left unphosphorylated and added to the phosphorylated vector in vast molar excess, thus ensuring that only one insert would be cloned into each vector.

If necessary, oligos were phosphorylated before cloning. 100 ng of annealed DNA was combined in a final reaction mix containing 66 mM Tris 7.5, 10 mM MgCl₂, 1 mM ATP, 15 mM DTT, 1 mM spermidine, 10-20 units T4 PNK in a total volume of 20 μl and incubated at 37°C for one hour. The reaction was stopped by phenol extracting and ethanol precipitating the DNA. Alternatively, equimolar amounts of single stranded oligonucleotide were combined in the same reaction mixture described above, the reaction allowed to proceed to completion, and then heated to 95°C followed by slow cooling to allow the strands to anneal.
Using this procedure of phosphorylating then annealing, it was also possible to kinase only one of the strands before annealing.

2.2.3.8 Colony hybridisation

Occasionally positive clones were present at too low a frequency to be easily found by random colony picking followed by miniprep analysis. In these instances colony hybridisation to a radiolabelled probe was used to identify the desired clones.

Bacteria were either plated at low density onto drug selected plates or drug resistant colonies picked to a fresh plate in a gridded pattern. After fresh colonies had grown out, the plate was cooled at 4°C for a few hours. Colonies were lifted off the plate by overlaying a dry nylon membrane filter (Amersham Hybond N) onto the cold plate using the bottom of a glass beaker to help achieve firm and even contact in all regions of the filter. Pinholes were made through the filter and agar plate to allow reorientation later. The filter was then placed colony side up onto 3MM paper saturated with, but not swimming in, 0.5 M NaOH for 5 min. The filter was transferred to a second sheet of 3MM soaked with 10X SSC, 0.5 M Tris 7.4 for 7 min, and this step repeated once. Finally, it was blotted on 3MM soaked with 2X SSC for 5 min. The filter was air dried and rinsed in 2X SSC, 0.1% w/v SDS. The residual bacterial slime was then wiped off, and the filter stored until needed for hybridisation.

Filters were prehybridised and hybridised as described in section xx.

Following hybridisation, filters were washed in 2X SSC, RT 1x; 0.2X SSC, 0.1% w/v SDS, 60°C, 30 min 2x, then exposed to autoradiographic film. Positive colonies were picked by reorienting the reference marks on the filter with those on the plate and choosing those colonies which matched spots on the film that had a tell-tale comet-like tail. Cultures were then grown up and miniprep DNA analysed.
2.2.3.9 Deletions

Promoter deletion series were made using either Exo III or Bal 31 exonuclease. Exo III was used for large scale deletions as it digests DNA at rates of several hundred bases/minute. Finer deletions were made with Bal 31 which digests DNA at a much slower rate around 10 bp/min.

The Exo III protocol is a variation of that described by Henikoff (1985). Exo III will only digest 5' overhanging or blunt ends. Twenty-five μg of the plasmid to be deleted was first digested with a four bp 3' overhang generating restriction endonuclease and then with a 5' or blunt cutter. The DNA was resuspended in 60 μl Exo III buffer and 500 units (a vast excess) enzyme added. Digestion was allowed to proceed at a temperature from RT to 37°C depending on the reaction rate required. 2.5 μl aliquots were removed into 7.5 μl S1 mix on ice at the appropriate timepoints.

After all samples had been taken, the S1 reactions were incubated at RT, 30 min to make the ends flush. 1 μl S1 stop was then added and each sample brought to 100 μl with TE and phenol extracted, ethanol precipitated and resuspended in a small volume of TE. The extent of the deletions was assayed by agarose or native polyacrylamide gel electrophoresis. At this stage the linear DNA molecules of the required aliquots either recircularised directly, recircularised following the addition of a restriction site linker, or the other end of a required fragment cut, the fragment gel purified and subsequently cloned into another vector.

All of these procedures were done using the standard blunting, linkering and cloning protocols detailed above.

Bal 31 deletions were performed as described in Maniatis et al (1982). As Bal 31 digests bidirectionally, the deletions were subcloned into a new vector. 10 μg DNA was cut with the endpoint restriction enzyme. The sample was resuspended in 100 μl total volume containing 200 mM NaCl, 12 mM CaCl₂, 12 mM MgCl₂, 20 mM Tris 8.0, 1 mM EDTA, 100 μg/ml BSA

* (66 mM Tris HCl pH 8.0, 0.66 mM MgCl₂)
and incubated at 30°C. After 3 min, 1 u Bal 31 was added and 10 μl aliquots
removed at the appropriate intervals to 40 μl TE + 50 μl
phenol/chloroform and quickly vortexed. Following ethanol
precipitation and resuspension the samples were made blunt with klenow
and linkered with phosphorylated oligonucleotides (section xx). Samples
were digested with a second enzyme, gel purified, and cloned into a
suitably digested vector.

The endpoints of fine deletions were mapped by radiolabelling a 5' end
and then resolving the fragments on a native acrylamide gel. Miniprep
DNA was digested with the appropriate restriction enzyme(s). Several
hundred nanograms of DNA were combined in a reaction mixture
containing 200 mM each of dGTP, dATP, dTTP; 2 mCi a-32P 3000
Ci/mmole dCTP, 100 mM NaCl, 10 mM MgCl2, 10 mM Tris 7.5, 1.25 u
Klenow. The reaction was incubated at RT for 5 min and stopped by the
addition of 1/10 volume stop and load dye, and the samples then resolved
on a native polyacrylamide gel.

2.2.3.10 DNA Sequencing

DNA samples were sequenced manually using the dideoxy chain
termination method. In all cases supercoiled DNA was used as the
template. To 1 μg DNA in 20 μl TE, 5 μl 1M NaOH, 1 mM EDTA was
added and incubated 5 min, RT. The sample was neutralised by addition
of 2.5 μl 2 M NH4OAc pH 4.6 and precipitated with 60 μl ethanol. The
washed denatured pellet was then treated as per the protocols in the
Sequenase manual (US Biochemicals).

Some plasmids were sequenced using an automated DNA sequencing
apparatus by R. Brown, ICRF (Applied Biosystems ABI 370A DNA
sequencer). Samples for analysis were prepared by ethanol precipitating 8
μg of CsCl gradient pure DNA, washing thoroughly with 70% v/v ethanol,
and resuspending in TE 7.5 at 1 mg/ml.
2.2.3.11 Southern blotting

Genomic DNA

Genomic DNA was isolated from confluent 15 cm dishes of cells. 2x PBSA washed cells were lysed in DNA lysis buffer (100 mM NaCl, 10 mM EDTA, 10 mM Tris pH 8.0, 0.5% SDS, 2 µg/ml proteinase K, freshly added) and incubated overnight at 37°C. The viscous lysate was phenol extracted 2x, and chloroform extracted 2x, then ethanol precipitated. The DNA was removed from the supernatant by fishing it out with the sealed, hooked end of a pasteur pipette, rinsed in 70% v/v ethanol and resuspended in TE 7.5. Yields were quantitated by UV spectrophotometry.

cutting

An appropriate amount of DNA was restriction digested checked for completion of the reaction by agarose gel electrophoresis. The digested DNA was ethanol precipitated and resuspended in a small volume of TE (usually less than 30 µl). The concentration of DNA in this solution was directly quantitated by fluorimetry to ensure equivalent loadings on the transfer gel.

Transfer

The DNA was resolved on an agarose gel of the appropriate concentration and stained with ethidium bromide. A photographic record of the gel next to a ruler was made to provide reference points for sizing bands, and to verify equal loadings of the samples. The DNA was nicked on a 254 nm UV transilluminator for ~15 sec to ensure complete transfer of the larger fragments. Following two denaturing washes in 0.5 M NaOH, 1.5 M NaCl, the DNA was transferred to a Genescreen nylon membrane in the same buffer. Complete transfer was determined by photographing a restained gel following transfer. The filter was neutralized by 2x 2 min washes in 1.5 M Tris, pH 7.5. Following air drying and a 45 min, 80°C bake, the DNA was crosslinked to the filter by UV treatment (5 min, 254 nm).
The crosslinking was probably not strictly necessary, as alkaline transfer should result in a permanent attachment of the DNA to the membrane.

**Filter Hybridisations**

Filters were prehybridised in 5x SSC, 25 mM NaHPO$_4$ pH 6.5, 10% w/v dextran sulphate, 50% v/v formamide, 0.02% w/v BSA, 0.02% w/v Ficoll 400, 0.02% w/v polyvinylpyrrolidone, 100 µg/ml tRNA, 0.5% w/v SDS, 0.01% v/v Antifoam A for 1 hour to overnight at 42°C, in the minimum volume required to completely cover the filter. Probe (10$^6$ cpm/ml, >5x10$^8$ cpm/µg Cerenkov) was boiled with sheared salmon sperm DNA (final conc 25 µg/ml) and added directly to the prehybe. Hybridisation was allowed to proceed at least overnight at 42°C. Filters were rinsed in 2xSSC, RT, 5 min twice, then transferred to washes of greater stringency at 60-65°C, culminating in a final wash of 0.1-0.2xSSC, 0.1% w/v SDS. Filters were kept wet and sealed into plastic bags for exposure to autoradiographic film (-70°C if with intensifying screen, RT if without)

**2.2.3.12 Random priming**

DNA probes for filter hybridisations were uniformly labelled using the random priming DNA synthesis procedure (Feinberg and Vogelstein, 1983). Ten to fifty µg of purified DNA fragment were combined with 1 µg of random 6-mer phosphorylated oligonucleotides (Pharmacia, pdN$_6$) in 20 µl total volume. The sample was boiled for 2 min and cooled to room temperature. 3 µl 10x buffer (500 mM Tris pH 7.4, 100 mM MgCl$_2$, 50 mM DTT), 5 µl (50 µCi) α-32P-dCTP (3000 Ci/mmol), 1 µl dA,T,GTP (33 mM) and 10 units of Klenow DNA polymerase fragment were added. The reaction was incubated for 30 min to overnight, 37°C and the probe recovered from the unincorporated nucleotides using a spun column.

**2.2.3.13 Spun columns**

Spun columns were used to separate unincorporated nucleotides from DNA. A 1 ml syringe was plugged with a small quantity of siliconised
glass wool and tamped down so that the plug was 2 to 3 mm thick. Sephadex beads (G-50, medium, equilibrated with TE, 100 mM NaCl) were loaded into the syringe and packed by centrifugation (1000g, 1 min, placing the column in a 15 ml conical tube). The process was repeated until 0.8 to 1.0 ml of beads were in the column. The reaction mix was brought to 200 µl with TE, 100 mM NaCl and loaded onto the column. If there was low melt agarose present in the mixture, it was heated to 65°C to melt the agarose before loading. The column was spun at 1000g, 1 min, and the flow through collected in an eppendorf tube. If less than 200 µl was recovered, an additional wash was performed using the same buffer. Unincorporated nucleotides remained trapped in the beads. Probe for blots was usually used directly after quantitation by scintillation counting.

2.2.4 RNA

2.2.4.1 RNA isolation

Total cellular RNA was isolated using one of two methods, both adaptations of the method of Chirgwin, et al (1979). The first uses a guanindinium isothiocyanate/SDS cell lysis protocol, followed by ultracentrifugation through a cesium chloride cushion. The second is a further modification by Seed (S. Goodbourn, personal communication) of the same procedure, in which LiCl is added to accelerate the precipitation of the RNA and thus allow a shorter ultracentrifugation spin.

RNA was isolated from cells growing under a variety of conditions. Sub confluent conditions meant that the cells were not contact inhibited, and capable of growing through several additional divisions. Typically Rat-1 cells were seeded at 3x10^6 cells/15cm diameter dish and harvested 18-24 hrs later. If the Rat-1 cells had been previously infected with a v-myc expressing retrovirus, only 2x10^6 cells were seeded, and the cells harvested 18-24 hrs later. When confluent, serum starved conditions were used, cells were grown to confluence and then maintained there for at least two
days, before being media changed to fresh media with only 0.5% FCS present. In the case of Rat-1 cells, RNA was harvested a further 2 or 3 days later. NIH-3T3's could only withstand overnight exposure to 0.5% serum, and were thus treated accordingly. Confluent, serum stimulated conditions were the same as serum starved except fresh FCS was added to a final concentration of 10% v/v and left on the cells for the required amount of time. Yields of total RNA were usually of the order of 20μg/10^6 Rat-1 cells or 10μg/10^6 NIH-3T3 cells.

2.2.4.2 RNA purification: Guanidinium prep

Cells were grown as described above. The media was removed and the cells lysed with 1.5 - 2.0 mls of RNA lysis buffer (4.23 M guanidinium isothiocyanate, 0.5% w/v sarkosyl, 25 mM Na citrate, 0.1% w/v antifoam, 50 mM β-mercaptoethanol in aqueous solution, pH 7.0). The genomic DNA was sheared by passaging the mixture 6x through a 21 gauge needle. The mix was subsequently layered onto a 1.5 ml CsCl cushion (5.7 M CsCl, 25 mM EDTA, in H₂O; pH 7.0) in a 5 ml (Beckman SW50 or SW55 rotor) ultracentrifuge tube. This preparation was then spun at 35 k rpm for 16 hrs using a swing out ultracentrifuge rotor. The supernatant was carefully removed and the tube inverted to drain completely. The tube top was then cut off with a sterile razor blade approximately 1 cm from the bottom. The top was discarded and the pellet in the tube bottom washed 1x very gently with 200 μl 75% v/v ethanol. Following removal of the ethanol wash, the pellet was resuspended in 300 μl TES and immediately ethanol precipitated. After a 10,000g, 4°C microfuge spin, the pellet was washed with 75% v/v ethanol and resuspended in TES. The RNA was then quantitated.
2.2.4.3 RNA purification: Guanidinium/LiCl

This procedure is similar to the previous total RNA preparation procedure, except for modifications to the cell lysis and ultracentrifugation steps.

Media was aspirated, and the cells lysed with a solution of guanidinium and lithium chloride (50% w/v GTC, 18.5% w/v LiCl, 2% v/v β-mercaptoethanol). The maximum capacity of this solution is $5 \times 10^7$ cells/ml. The lysate was then spun though the same CsCl cushion as described previously, except at 50,000 rpm for 2 hrs at 15°C using a Beckman SW55 swing out ultracentrifuge rotor. The RNA pellet was then processed as described above.

2.2.4.4 RNAse protection

RNA probe preparation

RNA probes were transcribed in vitro using standard transcription protocols (Zinn, et al, 1983). They were then gel purified for use in RNAase protection assays.

One microgram of template DNA, cleaved with a 5' overhang or blunt end generating restriction enzyme, was combined in 10 µl of reaction mix (50 µCi of 600Ci/mmol γ-32P GTP; 1mM each of CTP, ATP, UTP; 10mM DTT; 40 mM Tris pH 7.5; 6mM MgCl$_2$; 2mM spermidine; 20 units Rnasin; 10 units RNA polymerase) and incubated at 37°C for 60 to 90 min. (Some probes were required at reduced specific activity and this was achieved by adding non radioactive GTP to the reaction e.g. γ-actin reactions were performed with 5 µCi of radioactive GTP and 0.225mM non radioactive GTP.) 25 units RNAase free DNAase were added and the reaction continued for 15 min at 37°C. The reaction was then brought to 100 µl with TES, FC extracted 1x and NH$_4$OAc/EtOH precipitated.

Purified probes were isolated from denaturing PAGE gels. The pellet was resuspended in 10 µl M + G loading dye (80% v/v deionised
formamide, 1mM EDTA, 0.1% w/v bromophenol blue, 0.1% w/v xylene cyanol) and run on a 6% w/v 19:1 acrylamide:bis-acrylamide, 50% w/v urea gel in 1x TBE buffer. Following electrophoresis the gel was covered with saran wrap and orientation marks of radioactive ink applied to pieces of tape on its edges. A 15 sec x-ray exposure was taken of the gel. The bands on the gel corresponding to full length transcripts were excised and then both the gel and the film rinsed with ethanol. After they had dried, the film was used as a template to remove the appropriate gel slices, which were eluted in 400 µl elution buffer (0.5 M NH₄OAc, 0.1% w/v SDS, 0.1 mM EDTA) at 37°C overnight. The supernatant was precipitated by the addition of 20 µg glycogen and 1.1 ml ethanol, and the recovered pellet 75% v/v ethanol washed and resuspended in TES. The yield was calculated by combining 1 µl with 1 ml liquid scintillant (Aquasol) and counting in a liquid scintillation counter.

**hybridisations**

Ten micrograms of total RNA was ethanol precipitated, 75% v/v ethanol washed and resuspended in 30 µl hybridisation buffer (80% v/v formamide, 300 mM NaCl, 40 mM Na Pipes pH 6.2, 0.1% w/v SDS) to which probe had already been added. At least 100,000 cpm of each full specific activity probe, and a similar amount, corrected for its reduced specific activity, of each other probe, was used in each hybridisation. The sample was heated to 85°C for 10 min to completely denature it, and then incubated for at least 8 hrs at 45°C to 52°C, depending on the specific probes being used. Following this incubation, 350 µl RNAase solution (300 mM NaCl, 10 mM Tris pH 7.6, 5 mM EDTA, 40 µg/ml RNAase A, 2 µg/ml RNAase T1) was added, and the sample incubated at 30°C for 30 to 60 min. 10 µl 20% w/v SDS was added, followed by 10 µl PK solution (10 mg/ml Proteinase K in TE 7.5). Following a 37°C, 15 min incubation, the sample was fc extracted 1x, and 350 µl ethanol precipitated by the addition of 1 ml
ethanol and 20 μg glycogen carrier. The pellet was washed in 75% v/v ethanol and resuspended in 4 μl M + G loading dye. Samples were then subjected to denaturing PAGE, the gels dried down and exposed to preflashed autoradiographic film. Exposure times ranged from overnight to several weeks.

2.2.4.5 Northern blotting

Northern blots were performed by separating RNA species by denaturing (formaldehyde) agarose gel electrophoresis, transferring the RNA to a nylon membrane, and hybridising it with a radioactive probe.

10 μg of each RNA sample were ethanol precipitated/washed and resuspended in 20 μl RNA load dye (1x MOPS (20 mM MOPS pH 7.0, 5 mM NaOAc, 1 mM EDTA), formaldehyde, 0.1% w/v xylene cyanol, 0.1% w/v bromophenol blue). Samples were heat denatured, 68°C, 10 min, then rapidly cooled on ice. They were electrophoresed through an agarose/formaldehyde gel (1% w/v agarose, 1x MOPS, 2.2 M formaldehyde, 0.1 μg/ml EtBr) in running buffer of the same composition (except the agarose) until the bromophenol blue had reached the end of the gel. The gel was then rinsed in H2O and photographed with a ruler next to the gel for size calibration purposes on a 254 nm lightbox. The UV treatment (approximately 20 sec) serving to help break the RNA into smaller pieces. The RNA was then transferred to Genescreen Plus (Dupont). This was accomplished by making a sandwich of the following components on top of a glass plate: (closest to the glass first) two sheets of 3MM paper prewetted in 10x SSC with several inches overlapping two opposite edges; the gel, top side face down; a sheet of Genescreen Plus, cut to fit the gel, concave side contacting the gel, prewetted first in H2O, 10 sec, then in 10x SSC, 5 min; two additional sheets of prewetted (10x SSC) 3MM paper; a 5 cm stack of dry paper towels; another glass plate; a 500 g to 1 kg weight. This entire stack was placed on top of a shallow tray.
containing 1 l 10x SSC, with the wicks in the liquid. Pieces of Parafilm
were snugly arranged around the gel so that there was no contact between
the 3MM wicks and the paper towels. Transfer was allowed to proceed at
least overnight.

The filter was then removed from the stack and rinsed once in 2x SSC.
It was baked at 80°C for 1 hr and the RNA cross linked to the filter by
illumination with 254 nm UV light (5 min from 25 cm using a UVG54
mineralight). The filter was prehybridised in the smallest reasonable
volume of hybe mix (50% v/v deionised formamide, 4x SSC, 10x
Denhardt’s solution, 10% w/v dextran sulphate, 15 mg/ml tRNA, 50 mM
NaPi pH 7.2, 1 mM EDTA) in a sealed plastic bag at 42°C for at least 30 min.
Radioactive probe and low molecular weight (<500 bp) DNA were
denatured by boiling and added to the bag at final concentrations of 10^6
Cerenkov cpm/ml and 25 μg/ml respectively. The contents were
thoroughly mixed and incubated at 60°C overnight. The filter was
removed from the bag, washed at the appropriate stringency, usually 2x at
2x SSC, RT, followed by 2x at 0.2x SSC, 0.1% w/v SDS, 65°C, and
autoradiographed.

2.2.5 Protein:DNA interactions

2.2.5.1 Probe isolation and labelling

Fragments isolated from plasmids and synthetic oligonucleotides were
used as gel retardation probes. Fragments were prepared by digesting
plasmid DNA with the appropriate restriction endonuclease and resolving
the insert by native PAGE. Following EtBr staining, the appropriate band
was excised from the gel, eluted in 0.5 M NH₄OAc, 1 mM EDTA, 37°C
overnight, ethanol precipitated and resuspended in TE. Recovery was
quantitated by either UV spectrophotometry or PAGE. Oligonucleotides
were prepared via denaturing PAGE purification followed by elution as for
the fragment preparations.
Single strand oligonucleotide probes were 5' phosphate labelled by polynucleotide kinase. DNA, usually 50 to 100 ng, was combined in a reaction mix containing 66 mM Tris 7.5, 10 mM MgCl$_2$, 50 µCi $\gamma^{32}$P ATP 3000-5000 Ci/m mole, 15 mM DTT, 1 mM spermidine, 10-20 units T4 PNK in a total volume of 10-20 µl and incubated at 37°C for one hour. If single strand probes were required the reaction mix was combined with an equal volume of deionised formamide, heated to 95°C and loaded directly on a denaturing PAGE gel. For double strand probes equimolar quantities of the complementary strand oligonucleotide were heated with the labelled fragment (95°C, 5 min) and slowly cooled to RT. Following annealing the double strand probes were purified by native PAGE as described above. Yields of probes following gel purification were typically 50% of theoretical maxima.

Double strand fragment probes were labelled at both ends with $\alpha^{32}$P-dNTPs (typically A and C) by the Klenow fragment of DNA polymerase (5-10 min RT). The labelled DNA was usually phenol extracted and ethanol precipitated from NH$_4$OAc and then NaOAc. Occasionally unincorporated nucleotides were removed by Sephadex G-50 spun column chromatography or NACS-52 ion exchange chromatography (see DNA methods section above).

Incorporation of radioactivity was monitored by liquid scintillation counting a small aliquot of probe.

Sequences of fragment and oligonucleotide probes used are described in the Methods section above.

2.2.5.2 Whole cell extract preparation

Extracts for gel shift assays were prepared using a whole cell lysate procedure from G. Evan (personal communication), optimised to allow good recovery of myc protein. Cells were grown to subconfluent density, and usually harvested 18 to 24 hours after seeding. Following thorough
rinsing with ice cold PBS.A, cells were scraped into 15 ml conical tubes using a rubber policeman. Samples were pelleted by centrifugation (5 min, 1000g, 4°C), the supernatant discarded, and the pellet gently resuspended at a density of 3 x 10^7 cells/ml by pipetting in lysis buffer (0.1% w/v digitonin, 200 mM NaCl, 25 mM Tris pH 8.0, 1 mM DTT, 1 mM EDTA, 50 μg/ml phenyl-methyl-sulfonyl- fluoride (PMSF), 10 μg/ml n-tosyl-L-phenylalanine cloromethyl ketone (TPCK), 10 μg/ml soybean trypsin inhibitor, 1 μg/ml aprotinin, 1 μg/ml leupeptin, 1 mM benzamidine HCl). The sample was then transferred to a precooled Dounce, and homogenised with 10 strokes of the 'B' pestle. The sample was transferred to an eppendorf tube, incubated for 10 min, 0°C, and the nuclei and cellular debris pelleted by centrifugation (10000g, 4°C, 5 min). The supernatant was aliquoted to precooled eppendorf tubes, samples rapidly frozen on dry ice and stored at -70°C. One small aliquot was used to quantify the protein yield using a BioRad protein assay kit. Protein yields for Rat-1 cells were typically 100-200 μg/10^6 cells at 5-10 mg/ml. It was empirically determined that the 'B/β' complex appeared to degrade after multiple freeze thawings, so aliquots were made small enough (usually 50-75 μl) to prevent this. MAC I and MAC II complexes were both relatively insensitive to multiple freeze-thawings.

2.2.5.3 Gel retardation assays

protocol

Protein extract was combined with gel retardation buffer (100mM NaCl, 10mM Tris 7.6, 1mM EDTA, 1mM Benzamidine HCl, 1mM DTT, 5% (v/v) glycerol final concentration) and preincubated with non-specific carrier 5' on ice before addition of antibody, specific competitor or probe. Antibody and/or specific competitor DNA were added 5' prior to the addition of radiolabelled probe. 4-10μg of extract and 25-50,000 cpm (0.5-5 fmol) probe were used in standard analytical reactions. Following probe addition,
samples were kept on ice for a further 20 minutes, and complexes then resolved by native PAGE. Gels were typically poured between 20cm x 20cm glass plates with 1.5mm thick teflon spacers and combs using 4% (w/v) 30:1 acrylamide:bisacrylamide with 1/2x TBE buffer. They were prerun for >30' at 200V until the current stabilised between 10mA and 20mA. Samples were resolved at 200V constant voltage and gels run for typically 1-2hr until the free probe migrated approximately to the end of the gel. Gels were dried onto Whatman 3MM paper at 80°C using a vacuum assisted gel drier and exposed to Kodak X-OMAT AR-5 or Fuji RX x-ray film.

conditions

Nonspecific carrier DNA or synthetic nucleotide analogs were used at the following concentrations:

- Salmon sperm DNA (sonicated): 50 μg/ml
- poly d(IC)•poly d(IC): 75 μg/ml
- single strand Bluescript DNA: ~50 μg/ml
- Bacterial DNA (sonicated): 50 μg/ml
- poly dI•poly dC: 75 μg/ml
- poly dA•polydT: 75 μg/ml

Chemical compounds were added at the following final concentrations:

- MgCl₂: 10 mM
- ZnSO₄: 10 μM-2.5 mM
- EtBr: 5x10⁻⁶-5.0 mg/ml
- Spermidine: 5 mM
- Formamide: 5-10%

4 NTP and dNTPs in combination: 1mM each

2.2.5.4 DMS footprinting

Nucleotides directly involved in the contact of protein:DNA complexes were identified by DMS interference and protection assays. Probes were
labelled on one strand either by 'filling in' with $^{32}$P α- dNTP's and E. coli DNA polymerase I large (Klenow) fragment or kinasing with $^{32}$P γ-ATP and T4 polynucleotide kinase.

DNA was methylated for interference assays by suspension of ~5x10$^5$ - 1x10$^6$ cpm probe in 200μl 50 mM Na Cacodylate pH 8.0, 20 μg glycogen. 1 μl DMS was added and the reaction incubated 3-4 min at RT. 50μl DMS stop and 750μl ethanol were added, and the DNA precipitated from ethanol three times. The modified probe was then used in a band shift reaction containing 50 μg extract in a total volume of 50-100μl. Complexes were resolved via native PAGE using a gel slot of increased width to accommodate the larger amounts of protein in the reaction. Following electrophoresis, the gel was equilibrated in 0.5xTBE, 0.1% SDS for 30min, RT. The DNA was then electroblotted onto Whatman DE81 (DEAE-cellulose) paper using a standard electroblotting transfer apparatus in the same buffer. Transfers were performed at 200mA constant current for 0.4hr at RT. These condition were empirically determined to allow almost quantitative transfer of probe to the paper. Following exposure to film, the bands of interest were excised, eluted from the DE81 paper and ethanol precipitated. Methylated residues were cleaved by 1M piperidine treatment (95°C, 30min) and the piperidine removed by repeated ethanol precipitation. Sequencing ladders were resolved via denaturing PAGE on 8% (w/v)19:1 acrylamide:bisacrylamide, 50%(w/v) urea, 1xTBE wedge gradient gels run at ~23W constant power to maintain a surface temperature of ~42°C. Gels were fixed in 10% (v/v) methanol:10% (v/v) acetic acid, dried at 80°C on a vacuum assisted gel dryer and exposed to x-ray film.

DMS protection reactions were performed as described for interference assays, except that the protein:DNA bandshift mixtures were treated for 1 min with 0.4μl DMS immediately before loading onto the preparative gel.
2.2.5.5 Formic acid modification

Probes were depurinated using formic acid and then subjected to preparative bandshifts as described for DMS interference assays. Probe (~5x10^5 - 1x10^6 cpm) in 20μl H_2O, 20μg glycogen was treated with 50μl formic acid, 10min, RT after which 200μl HZ stop was added. The DNA was repeatedly precipitated from ethanol and then used in a preparative bandshift reaction.

2.2.5.6 Hydrazine interference

Pyrimidines were cleaved with hydrazine and probe then used in interference reactions as described for DMS interference. Probe (~5x10^5 - 1x10^6 cpm) in 30μl H_2O, 20μg glycogen was treated with 50μl hydrazine, 10min, RT after which 200μl HZ stop was added. The DNA was repeatedly precipitated from ethanol and then used in a preparative bandshift reaction.

2.2.5.7 Maxam and Gilbert markers

The products of the formic acid and hydrazine reactions described above modify DNA at purines and pyrimidines respectively. Immediately following modification piperidine cleavages were performed (see DMS interference protocol) and the products used as markers on sequencing gels.

2.2.6 Protein

2.2.6.1 Protein quantification

Protein content was measured using a BioRad Protein Assay kit, the procedure is as recommended by the manufacturer. 5x stock assay mix was made 1x by the addition of 4 volumes of H_2O. 1-5 μl extract was added to 1 ml assay mix in a disposable plastic cuvette, mixed well, incubated (5 to 60 min, RT), and the A_{595} measured. A standard curve of 1 to 20 μg γ-globulin (BioRad) was used to calibrate the readings. Any OD >0.5 was
considered out of the linear range of the assay, and a greater dilution made.

2.2.6.2 Immunoblot extract preparation

Cell extracts were prepared for immunoblotting by boiling 100 µg of cell extract as used in the gel retardation assays in an equal volume of protein gel loading buffer (125 mM Tris pH 6.8, 2 mM EDTA, 4% w/v SDS, 20% v/v glycerol, 10% v/v β-mercaptoethanol, 0.002% Bromophenol blue).

2.2.7 Computer assisted analysis

2.2.7.1 DNA

Restriction maps, translations of open reading frames, database searches and other DNA sequence manipulations were performed using the Intelligenetics suite of programs on a VAX 8700 computer.


2.2.7.2 Protein database searches

Homology searches of protein sequences against a database were performed using the PROSEARCH program (Collins, et al, 1988) on an AMT-600 Distributed Array Processor.
Figure II.1: Sequence of the human c-myc gene

Annotated 8082 bp sequence of the HindIII:EcoRI genomic DNA fragment (Intelligenetics/GenBank sequence file HUMMYCC.2).

The 5' endpoints of some of the deletion constructs as described in the text are indicated by bold-faced type with the sequence position relative to the P1 promoter cap site shown. The +47 (relative to P2)/+208 position is at the 3' end of many of the constructs.


Exon/intron boundaries are indicated as EX#/IVS#. A solid vertical line is present to the right of the exon sequences.

The major polyadenylation signals are indicated by horizontal lines through the sequences.

The predicted translation product of the major open reading frame, initiated from the major ATG or minor CTG translation start sites is presented in three letter code underneath the DNA sequence.

Within the c-myc protein sequence several relevant sequences are indicated. The peptides to which the monoclonal antibodies, mAb1 and mAb2 were raised are indicated by single solid underlines. Domains of the protein which are required for c-myc autoregulatory and transforming activities are also shown in [] brackets (Penn, et al, 1990b, Stone, et al, 1987). These include amino acids 106-143 and 353-434. Residues within the putative dimerisation and DNA binding helix-loop-helix, leucine zipper and basic regions are indicated by single, double and dotted underlines (Penn, et al, 1990c).
CGCTGCCAGG ACCCGCTTCT CTGAAGGTG CTCTTTGCAG CTGCTTAGAC

Alternate start
GCTGGATTGG TTTCCGGTAG
(METAspPhe PheArgValV)

EX1/IVS1
TGAAAAACCC/AGGTAGACCCA GAAGTCCACT TGCCCTTTAAA TTTATTTTT TATCACTTAA ATGCCTGAGAT

a1GlutAASG

GAGTCTGAATG CCTAAAATGG GTGCTTTTTC TCCCATTTCT CGGCTATTGA CACTTTTTC AGAGTAGTAA
TGAGAACTTGG GCTGGGGGATGGG GGGGTAAATGC CAGAATGCT TGGGGTTAAG TGGACTGCTG AGAGGACAGG
AGGACGAAAG CCAGAAAGTGA CTTTTAAGAG CCCTCTTGGCA TTGGGATGGG CAGCAGCTGG
ATTCCGCTGG ACCTGGCAAGC GTTCGATGCA TCCCGCTTCC CCCGGAGCGGCAT CCTCCTGGCT GTTGGGGGGG
GCTGGGGGGT GCTTCTGGG CCCCAGGACT CACGGACTT GGGGTGTTTG GCTTGGGGGG
GGTCTTGGTG AGCCAGATGG CTCGCCAGCC GTCGATTGTT CCCCCTCTCC CCGAGAGCGG CATAACTTTG CAGGGCGATT
CCCAACCCCG CTGATCTTCT TAAGAAGTGG CAGATTTGCT TTTAAAAAG CAATTAAACA ATTTAAAAAC
TGCTGCTGTA GAAGGTGTTAG CAGCTGAGGA TGGGTAGGCG CAGCAGGGCGG AAAAAAGGGG CAGAGGTGAG
TCCGATGCTC TCTGGAAGCT GCTACTGGAA AAAACAGGGGC GAATCTCCGC ACCAGCCTTC GACTCCCCG
CCGGCGCCGC CTTGGGCGTG CTCGCCTGCG GAGATGCGGA GGAAGTGCAG GGGGCGGGGC TCTGGGCGGT
TCCAGAACAG CTGCTACCTCC TGGTGGGCTG GTCGCGGCGG AGTTAGCCCA GGGGGTCTCG TGGCGCGATT
GGATCTGACTG AAAGGGATGG CCCCATTATAT TATTGACAC CCCCCCTGTA TTTATGAGGG
GTGTTAAAGG CCGCGGCGTG AGTCGGCCAC TCGAGGCGGA GAGAAAGAAGA AAGAAGCGT GCAAAAAGGG
TGTTGACGGG GGGGTGTTAC ACGGGGCGGG CAGAAAGGGA TGGAGGAGAG TGGACGCTCGT
GGCGGGCGGG TGGAGACAGC GCTAGGGCAG CAGGTGAAAG CACCCGCAGC GAGGGGCGG GCGGGCGGG
GGGGGCGGGT GCCGGCCACA TGAAGAGCGG CACGTACAGG ATAGCGAGGG ACCTGCAAAG GGGGGTGAAG
GGGTGTCCTT TTATTTCCAC ACAAAGACC ACCAAGCGG TTTAGGGGAT AGCTCTGCAA GGGGGAGAGT
TCCGAGACTGG GGGCGCGACT GGGCGCGCTT CAGGTTTGG GCTCGGGGAT CCCCAGGGGC GGGGGCGTCG

IVS1/EX2
TCCCGCCTTG TGGCGCCCGC TCCAGC/AGCC TCCCGCGAGG ATG CCC CTC AAC GGT AGC TTC

inPr oProAlaThr) MET Pro Leu Asn Val Ser Phe

ACC AAC AGG AAC TAT GAC TCT GAC TAG TCG GTG CAG CCG TAT TTC TAC TGC Thr Asn Arg Asn Tyr Asp Leu Asp Tyr Asp Ser Val Gin Pro Tyr Phe Tyr Cys

GAC GAG GAG GAC AAC TTC TAC CAG CAG CAG CAC GAG CTG CAG CCC CCG Asp Glu Glu Glu Asp Phe Tyr Gin Gin Gin Glu Ser Glu Leu Gin Pro Pro aa 44 mAb1 peptide aa 55

GGC CCC ACC AGG GAT ATC TGG AAG AAA TTC GAG CTG CTG CCC ACC CCG CCC CTG Ala Pro Ser Arg Asp Ile Asp Ile Tyr Phe Gin Gin Gin Ser Gin Gin Gin Gin Ser Glu Leu Pro Thr Pro Pro Leu

TCC CCT ACC CGC CGC TCC GGG CTC TGC TCG CCC TCC TAC GTT GCG GTC ACA CCC Ser Pro Ser Arg Arg Ser Gly Leu Cys Ser Pro Ser Tyr Val Ala Val Thr Pro

TTC TCC CCT GGG GGA GAC AAC GAC GGC GGT GGC GGG AGC TTC ACG GCC GAC Phe Ser Leu Arg Gly Asp Asp Arg Ser Tyr Gin Gin Gin Gin Pro Pro Thr Ala Asp (aa 106

CAG CTG GAG ATG CTG ACC GAG CTG CTG GGA GGA GAC ATG GTG AAC CAG AGT TTC Gin Leu Glu Met Val Thr Glu Leu (Leu Gly Asp Met Val Asn Gin Ser Phe

ATC TGG GAC CCG GAC GAG ACC TTC ATC AAA AAC ATC ATC TTC CAG GAC TGC Ile Cys Asp Pro Asp Glu Thr Phe Ile Tyr Asn Ile Ile Ile Gin Asp Cys aa 143

ATG TGG AGC GGC TTC TCG GCC GCC GCC AAC CTC GTC TCA GAG AAG CTG GCC GCC GCC Met Thr Ser Gly Phe Ala Ala Ala Ala Ala) Leu Val Ser Glu Lys Leu Ala Ser

TAC CAG GCT GCG AAA GAC AGC GGC AGC CAG CCC GCC GGC GGC CAC AGC Tyr Gin Ala Ala Arg Lys Asp Ser Gly Ser Pro Asn Pro Ala Arg Gly His Ser aa 173 mAb2 peptide aa

GTC TGC TCC ACC TCC AGC CTG TAC CAG GAT CTG ACC GCG GCC GCC TCA GAG Val Cys Ser Ser Tyr Leu Tyr Leu Gin Asp Leu Ser Ala Ala Ala Ser Glu
427  439
CAG TTG AAA CAC AAA CTT GAA CAG CTA CGG AAC TCT TGT GCG TAA GGAAA AGTAA
Gln Leu Lys His Lys Leu Glu Gin Leu Arg Asn Ser Cys Ala

GGAAAAACGAT TCCTTCTACG AGAAAATGTCCT GAGCAATCA CCGATGAACCT TGGTTCAAT TGAATGATCAA
ATGCAAACCTC ACAACCTTGG CTGAGTCTTG AGACTGAAAG ATTTAGCCAT AATGTAACCT GCCTCAAATT
GGACTTTGGG CATATAAGAA CTTTTATATG CTTACACTCT TTATTTTTTC TTTAAACAGAT TTTATTAA
GAAATTGTTT TTAAAAATT TTAGATTTAC ACAATGTTTTC TCTGTAATAA TGCCATTTA AAAGTTAAT
EX3/
CTTTTAAGAA ACGTTTATAG CAGTTACACA GAATTTCAAT CCTAGTATAT AGTACCTAGT ATTATAGGTA
CTATAAACCC TAACTTTTT TATTTAGTA CTATTTGCTTT TTTAAAGTTG ATTTTTTTCT ATTTTTTTA
EX3/
GAAAAAAAATA AAAAAAGCC AAAATAATCA TTAGGGCAAAA TCCTAAAGTG TGAATTGTTT GTTTGTTTGC
TTCTCTCTTCC CAAACACACAC AATTCCTGTT TGGTTCAAAC ATTGGCCTT TGCAGGGCCG GCTTTAGAAA
AGGCAAGAGT TTTCTCTGTT TGAATGGGTT TGAGGGGGCT TAAGTTCTTTT AAGTTCTTGG AGGTCTTAAG
ATGGCTCTGT GAGACTAAGA TAAACCCAG AGTTGACAGT TAGAAGGAAAT GGCAGAAGGC AGGTGAGGAG
GCTGAGACGA TTGCAAAGGAG TACAAGGGCT TCAAGGCGTA CAGTTAAGTA CACAAAGGAG CATAAGGACT
GGGGAGGTTGG GAGGAAGGTTG AGGAAGAAAC TCTGTTACT TTAGTTAACG AGTCCAGTC CTCCTCCTAC
TCAAAAACCA GGAATT
Chapter 3  Promoter deletion analysis
3.1 Results

3.1.1 Analysis of multiple indicator constructs

Reporter plasmids were constructed which contain genomic fragments of the human c-myc gene and upstream sequences fused to two different indicator genes. Each plasmid was analysed for the response of its indicator gene to the overexpression of v-myc protein from an independent vector in Rat-1 cells. The plasmid constructs assayed include pGC1H/R, pGC1H/RBam8, pBSKS+HBCAT and pBSKS+HN277IFN (see Figure III.1 for maps). pGC1H/R contains an 8.1 kb HindIII/EcoRI genomic fragment of the human c-myc gene which includes 2.3 kb of sequences upstream of the P1 transcription initiation site, and about 5.8 kb of downstream sequences, including all three exons and approximately 550 bp downstream of the polyadenylation site. RNA transcribed from this gene is capable of directing the production of human c-myc protein as it contains the complete human c-myc open reading frame (Land, et al, 1986). pGC1H/RBam8 contains an eight bp linker inserted into the human c-myc open reading frame at the 48th codon, which should lead to the production of a frameshifted and presumably inactive protein, thus controlling for any effects of the human c-myc derived from the pGC1H/R vector. The other two plasmids contained human c-myc sequences from -2329 to +208 bp relative to P1 fused to the bacterial chloramphenicol acetyl transferase (CAT) gene (pBSKS+H(B)CAT) or a fragment of the human β-IFN gene (pBSKS+HN277IFN). The β-IFN gene sequences used contain only a portion of the open reading frame, deleting regions of the protein required for biological activity, and would thus not be expected to have any effect on cell growth (De Maeyer and De Maeyer-Guignard, 1988).

These plasmids were stably cotransfected into Rat-1 cells in combination with a selectable marker plasmid, pJ6Ωhygro, which
expresses a hygromycin resistance gene under the control of the rat β-actin promoter, and approximately 150 hygromycin resistant colonies from each transfection pooled. The pooled colonies were maintained in selective media until they were split into parallel cultures and infected either with VM (Dotto, et al, 1985), a retrovirus which expresses the v-myc gene from its LTR and a G418 resistance gene from an internal SV40 promoter, or LJ (Korman, et al, 1987), a control retrovirus which expresses only a G418 resistance gene from an internal SV40 promoter (see Figure III.2). The VM virus has been previously characterised for its expression properties in Rat-1 cells, and has been found to produce levels of p110v-gag-myc protein which are sufficient to cause >90% suppression of endogenous rat c-myc RNA levels (Penn, et al, 1990a). Following selection in G418-containing media, several hundred drug resistant colonies were pooled. Total RNA was harvested from subconfluent growing cultures of these cells, 24-48 hours after seeding.

Expression from the transfected plasmids was analysed by RNAase protection. In this procedure RNA is hybridised to a uniformly labelled antisense RNA probe, and then treated with RNAase. Only those regions of the probe which are present as RNA:RNA hybrids remain intact, and can be resolved by denaturing polyacrylamide gel electrophoresis and visualised via autoradiography. The bands thus detected are quantitatively indicative of the mRNA present in the sample. Use of multiple probes allows the comparison of expression levels of different sequences in one sample, or normalisation of the loading between samples with an mRNA known to be expressed invariantly. A schematic of the protocol is shown below:
Cotransfect myc promoter construct and hygromycin resistance gene.

Hygromycin selection \(~10-14\) days

Pool >100 stable hygromycin resistant colonies.

1 day

Infect with v-myc or control retrovirus.

G418 selection \(~10\) days

Pool >100 G418 resistant colonies.

\(~24\) hours

Harvest total cellular RNA from subconfluent, growing cells.

Quantitate myc promoter dependent RNA by RNAase protection.
3.1.1.1 Transfected genomic human c-myc constructs are appropriately regulated by v-myc in Rat-1 cells

Results for the pGC1H/R and pGC1H/RBam8 transfectants are shown in Figure III.3. RNA was hybridised to the human c-myc 5' and γ-actin probes (see Materials and Methods for descriptions of the probes). If the human c-myc promoter derived transcripts are initiated properly from the P1 or P2 initiation sites, bands of either 509 nt or 348 nt should be protected. The γ-actin probe should detect a band of ~70 bp. Bands corresponding in size to those predicted for human c-myc specific transcription are present, with both the P1 and P2 promoters being utilised in the control cells infected with LJ (see Figure III.3A). The P2-dependent signal is always stronger than the P1-dependent signal, and after correction for variation in signal intensity due to probe length, the mean ratio of the P2:P1 signals in 17 samples (including some data not shown) is 2.4:1 (sd 0.74). Human c-myc specific signals from the colony pools infected with the v-myc virus are reduced compared to those of the relevant control lanes, though the ratio of the P1 and P2 bands is unchanged. The frameshifted construct displays a similar degree of activity and responsiveness to v-myc as the intact gene (see Figure III.3B). When the results from eight pairs of samples (the four shown in the figure, and data not shown) were quantitated by densitometry the mean reduction of P2-dependent signal was 5.2 (sd 2.4) fold and the P1-dependent signal 6.3 (sd 3.3) fold. The signals seen with both of these constructs are particularly weak, requiring several weeks of exposure to appear with even moderate intensity. Since relatively high levels of myc protein are required to significantly repress the endogenous c-myc gene (Penn, et al, 1990a), the amount of human c-myc protein being expressed from the pGC1H/R construct is probably not sufficient to significantly alter its own expression.
Thus, an 8 kb fragment of the human c-myc gene when transfected into Rat-1 cells appears to initiate transcription appropriately from both the P1 and P2 promoters, and its expression is down regulated following infection with a retroviral vector capable of suppressing endogenous c-myc RNA expression in these cells (Penn, et al, 1990a). However RNA levels from these constructs are quite poor relative to endogenous Rat c-myc mRNA levels, requiring at least 50 fold longer exposure time to achieve the same intensity of signal (as seen in e.g. Figure III.4 below).

3.1.1.2 Human c-myc promoter fusion constructs also respond to v-myc

Results of the human c-myc promoter fusion constructs, pBSKS+HBCAT and pBSKS+HN277IFN, are shown in Figure III.4. These samples were analysed using the β-IFN 3' end or CAT RNAase protection probes, which map only the indicator gene, and not the human c-myc specific component of the mRNAs (see Materials and Methods). In addition to the response of the transfected constructs, this figure also shows the levels of RNA expression from the endogenous rat c-myc P2 promoter in the BSKS+H(B)CAT transfected, VM or LJ infected, samples. The CAT specific probe should protect a band of 145 nt, the β-IFN probe one of 363 nt and the rat c-myc probe a P2 specific band of 350 nt. In each cast the appropriate band is repressed in the VM infected samples relative to the controls. The IFN and c-myc signals are each repressed 10-15 fold, and the CAT signal is repressed ~3 fold. Thus, while all three sets of samples display a suppression of signal, the β-IFN and endogenous c-myc appear to be more responsive than the CAT indicator.

3.1.1.3 5' end mapping of transfected fusion genes

As the probes for the β-interferon and CAT messages did not map the 5' end of their respective transcripts, it was important to ascertain that the mRNA was both initiating and responding to the v-myc protein appropriately. The same RNA samples which were analysed in Figure
III.4 were probed with the human c-myc 5' end probe. In this instance correctly initiated transcripts should be protected by a fragment of 47 nt for P2 initiation or 208 nt for P1 initiation. Figure III.5 shows both BSKS+HN277IFN and BSKS+H(B)CAT samples either probed with the human c-myc 5' end and γ-actin probes, or only the human c-myc 5' end probe. A 47 nt P2 band is detectable in both of the LJ infected control pools and is reduced ~4 fold in the VM infected pools. Weak bands of ~310, 210 and 150 nt are also visible, primarily in the LJ infected pools. The 310 nt band probably corresponds to transcripts which initiate upstream of position -101 and thus protect through to the 3' end of the probe. The 210 nt band may be indicative of P1 specific transcripts, but due to the weakness of the signal, and the presence of other bands of approximately the same intensity but the inappropriate size, it is difficult to be certain.

The 5' ends of the HN277IFN transcripts were also mapped using the Myc/βIFN fusion A probe which should protect a P2-dependent band of 127 nt and a P1 dependent band of 288 nt. Bands of the appropriate sizes were detected, with the repression due to v-myc being 3-4 fold for transcripts initiating from either promoter. The ratio of P2:P1-specific signals after correction for variation in probe length was ~10:1. An additional band, which corresponds to the full length of the homology between probe and transfected gene was also detected. The intensity of this signal was at least equal to that of the P2-dependent signals, and was reduced in the v-myc infected samples by approximately the same degree as the appropriately initiated transcripts.

Thus the promoter fusion constructs transcribe from both the P1 and P2 promoters of the human c-myc gene, and these correctly initiated transcripts respond to the presence of high levels of myc protein in the same fashion as the endogenous rat c-myc gene. The properly initiated constructs respond equally well to the presence of v-myc protein; both the
IFN and CAT plasmids are repressed 3-4 fold. The difference in the degree of repression between the two constructs seen with this probe compared to those used in Figure III.4 is probably due to the differential in levels of transcript which initiate upstream of position -101 and which are also repressed in the presence of v-myc.

3.1.1.4 v-myc response requires c-myc promoter sequences

It was next necessary to establish that the response shown was due to the presence of the human c-myc sequences rather than to the indicator genes themselves. To test this possibility, two different constructs were tested in the assay. The first was pSV2CAT (Gorman, et al, 1982), a plasmid which expresses the CAT gene from the SV40 early promoter/enhancer. The second was a construct containing the βIFN sequences under the control of the Moloney murine leukemia virus LTR (pUCSVterLTRIFN). Maps of these constructs are shown in Figure III.6, and the results of this experiment in Figure III.7. The probes used for the RNAase protection assays should detect either indicator-specific sequences in the case of the CAT transfection, or the 5' ends of the IFN containing mRNAs. The endogenous c-myc signal was also monitored in the pSV2CAT samples. Neither the pSV2CAT or LTRIFN transfectants show a significant change of the transfectant signal in the presence of v-myc, while the endogenous rat c-myc RNA displays an ~8 fold reduced signal. Interestingly, as discussed below, the presence of SV40 promoter and enhancer sequences does not prevent the response of a myc promoter construct to respond to v-myc. Thus the response to high v-myc levels seen with the c-myc-IFN and -CAT hybrid constructs seems to be determined by the presence of human c-myc promoter-specific sequences.
3.1.1.5 Transfected indicator constructs respond to serum as well as to v-myc

We were interested to see if our transfected indicator constructs would mirror the response of the endogenous c-myc gene to v-myc under a number of cell growth conditions, and in particular following serum renewal, since it has been reported that transcription of the c-myc gene can be stimulated several fold by the application of fresh serum to quiescent cells in culture (Greenberg and Ziff, 1984). Pooled cells transfected with the CAT indicator construct and infected with either the v-myc or control retrovirus were grown under a variety of serum and cell density conditions. The growth parameters tested include subconfluent proliferating cells maintained in 10% serum and harvested 18 to 24 hours after seeding, and subconfluent or confluent cells which had been maintained in 0.5% serum for 2-3 days, then harvested with or without a 2 hr treatment of 10% serum.

The samples were probed for endogenous c-myc, transfected CAT, and control γ-actin messages (Figure III.8). Since γ-actin is itself responsive to both serum and cell density it is not strictly valid to use this indicator as a loading control, except among samples harvested under the same growth conditions (Greenberg and Ziff, 1984).

The serum response of endogenous c-myc gene was analysed along with that of the transfected H(B)CAT construct (Figure III.8). When the cells were maintained in 0.5% serum in a subconfluent state, endogenous c-myc expression did not appear to be greatly affected compared to cells in 10% serum. This signal was slightly stimulated by the addition of serum. This is somewhat surprising since c-myc levels in Balb/c 3T3 cells maintained under very similar conditions have previously been reported to be significantly reduced (Dean, et al, 1986). More recent studies on the particular Rat-1 fibroblast line used here have indicated that endogenous
c-myc protein levels do not alter significantly until serum concentrations are reduced to 0.05%, and also that if either c- or v-myc is constitutive expressed these cells will not growth arrest (G Evan, D Hancock and T Littlewood, personal communication). When the cells were made confluent and then serum starved, there was a 2-3 fold reduction in the level of c-myc expression. Addition of serum led to a 4-16 fold induction. The serum response observed in each case was similar in the presence of v-myc, though the absolute levels of mRNA were repressed.

CAT-specific signals were detectable under all of the growth conditions tested. The transfected H(B)CAT gene displayed a response to serum which paralleled that of the endogenous gene in some, but not all cases. The strongest response was seen when cells were maintained at confluence, with a maximal stimulation of 6 fold. However, this response was highly variable, with some samples showing little or no increased signal following serum application. The presence of v-myc did not have a significant effect on the serum responsive behaviour of the CAT message, as a similar magnitude of response was found in almost all cases regardless of whether control or v-myc infected pools were assayed.

Both the endogenous c-myc and CAT RNA levels were reduced in all cases in the presence of v-myc. While the extent of repression of the transfected gene was seen to vary somewhat, no correlation between the specific growth conditions and the degree of its repression was apparent. The reduction of endogenous c-myc expression by v-myc was strongest in the proliferating cells, and least when the cells were maintained at subconfluence in 0.5% serum.

The transfected hybrid indicator constructs thus appear to be stimulable by serum as well as repressible by v-myc under most growth conditions. Additional experiments described in more detail below have indicated that the serum response is not easily separable from an apparent serum
response of the non-myc sequences. A more extensive analysis of the human c-myc serum response was therefore not attempted. These results on the response of the endogenous c-myc gene are largely concordant with a more detailed study of the serum and myc responsiveness of Rat-1 c-myc expression performed by LJZ Penn and MW Brooks. They found an average maximal induction of ~4 fold, seen 1.5 hours following serum stimulation, and roughly the same absolute extent of stimulation in the presence or absence of exogenous myc (personal communication).

3.1.1.6 Conclusions

These results established that it was possible to transfect a human c-myc promoter construct into the Rat-1 cell line and observe regulation of the transcribed indicator in response to both exogenous myc expression and serum application. The transcripts protected the appropriate sized probe fragments to have initiated from the human c-myc P1 and P2 promoters. Interestingly, the ratio of promoter utilisation appears to differ in the constructs used. The P2 promoter was utilised 2-3 fold more often when constructs containing ~8 kb of genomic sequence were tested. When ~2.5 kb of upstream sequence through P2 were assayed, greater than 90% of transcripts appeared to initiate from the P2 promoter. The promoter utilisation found with the larger construct is similar to that seen when endogenous human c-myc transcripts from HeLa cells are analysed (not shown), while a much greater utilisation of the P2 promoter is reflective of endogenous c-myc expression in the Rat-1 host cell line.

The specificity of the response to the human c-myc gene fragments was demonstrated by the inability of either the pSV2CAT or the LTR-IFN constructs to respond to increased levels of myc expression, while both the human c-myc promoter-IFN and -CAT fusion constructs show a similar repression in response to exogenous myc. The combination of nonresponsiveness of the control constructs, and the similar
responsiveness of human c-myc promoter sequences with two different indicators, one of eukaryotic and the other prokaryotic origin, imply that the ability to be repressed is due to the presence of the human c-myc sequences, and not to sequences in the indicators.

3.1.2 Analytical strategy to identify a 'myc response element'

Since 2.5 kb of human c-myc promoter proximal sequences were capable of mediating a response to v-myc, we decided to concentrate on identifying a minimal response element for down regulation within this region. The most straightforward method for identifying such elements has typically been to produce a set of nested deletions of the promoter in question, transfet them into cells, and identify a construct which would no longer respond to the stimulus. This also seemed the most promising approach in this situation, so a set of 5' deletions was constructed for assaying.

The same constructs which can be used for determining the boundaries of a myc response element can also be used for monitoring other regulatory phenomena. For instance, the mechanism of determination of promoter usage or 'basal' levels of transcription may involve sequences in the region of interest. However, as our primary goal was the identification of a myc response element with the ultimate goal of understanding myc protein function as opposed to myc promoter function, it was felt prudent to monitor changes in these activities, but only insofar as they would lead to progress in elucidating the autoregulatory mechanism.

In order to monitor the veracity of the transfection/infection protocol, a construct containing the human α-globin transcription unit (Treisman, et al, 1983) was used in subsequent transfections. As there is no a priori reason to expect that this construct would respond to levels of myc protein in the cells, it was hoped that monitoring levels of α-globin mRNA would be useful to control for an equal division of transfected cells between the
retrovirally infected pools derived from a single transfection, and would control for any drift in signal once the pool had been divided. This plasmid has been successfully used as a transient transfection expression control in previous gene expression studies (Treisman, 1985). The modified protocol is shown below:
Cotransfect promoter deletion, alpha globin, and hygromycin resistance genes.

Hygromycin selection  ~10-14 days

Pool >100 stable hygromycin resistant colonies.

1 day

Infect with v-myc or control retrovirus.

G418 selection  ~10 days

Pool >100 G418 resistant colonies.

~24 hours

Harvest total cellular RNA from subconfluent, growing cells.

Quantitate myc promoter dependent RNA by RNAase protection.

3.1.3 Results of large scale deletion series

3.1.3.1 General observations

The constructs assayed in this next series of experiments are shown schematically in Figure III.9. The 5' deletion series was constructed using
residue -2329 relative to P1 as the start point. The 3' end of the constructs is uniformly at +208 bp relative to P1 (or +47 bp relative to P2). The deletions were constructed using a combination of available restriction endonuclease sites and exonuclease digested fragments. As it was impractical to test every deletion using two indicators, and the interferon constructs initially appeared to give a much stronger signal than the CAT constructs (see Figure III.4, compare the CAT +/- VM lanes with the IFN +/- VM lanes), all of the deletions were fused to the β interferon indicator sequences. The deletions tested had the following 5' endpoints relative to P1 as determined by sequence analysis: -353, -348, -276, -181, -169, -101, -91, -88, -86 and +119. As a control, a construct deleted to position +430 in the interferon sequences was also assayed. RNA from pools of colonies isolated and manipulated as described above was analysed for expression of both the transfected human c-myc/βIFN fusion construct and the α-globin gene, in addition to the endogenous rat c-myc and γ-actin genes (Figures III.10-13).

Transfected fusion construct signals were analysed using the human c-myc/βIFN probe A (Figures III.10, III.11, III.12). This probe should protect 288 nt P1 dependent transcripts and 127 nt P2 dependent transcripts (see Materials and Methods). Duplicate independent transfections are shown, with the P1 and P2 dependent signals indicated. In order to facilitate comparisons between gels, the -91 sample on the -353 to -91 gel was also analysed as the first of the pair of -91 deletions on the -91 to +430 gel. Within the constraints of a given deletion endpoint, myc promoter specific bands are detectable for each sample. For example, P1 specific bands are seen in all lanes up to the -58 deletion, after which, as there are no longer sequences 5' of P1 present, the signal disappears. Similarly, the +430 deletion, which contains no human c-myc sequences, has no detectable signal. In addition to the appropriately initiated transcripts, in
all cases a strong signal which corresponds to upstream initiated transcription is indicated. In those deletions in which the 5' end of the transfected human c-myc sequences is shorter than in the probe (ie deletions 3' of position -101) this 'readthrough' signal is seen to be shortened accordingly.

3.1.3.2 Deletion constructs -353 to -91

1 Effect of v-myc

The presence of v-myc leads to some reduction of both the P1 and P2 dependent signals in all of the deletions analysed (Figure III.10). There does not appear to be any consistent variation in the responsiveness of the P1 signal for constructs with 5' endpoints from -353 to -91, which demonstrate a mean degree of repression of ~6 fold and a range from 2 to 14 fold. The P2 signal in these same samples is also repressed, though to a lesser degree, with a mean of ~5.5 fold and a range from 2.5 to 20 fold. As was seen with the construct containing 2.5 kb of human c-myc promoter sequences, (Figure III.5) the upstream initiating signals also appear to be repressed in the presence of v-myc, though to a lesser degree than the correctly initiating transcripts (mean 4.8, range 1.3-22). Some samples (eg the second -276 deletion) display a more dramatic reduction than most of the other samples, and this excessive repression is paralleled by a general reduction in the non-specific background seen. However this increased response is not reproducible, as the duplicate samples in each case display a response much closer to the mean.

2 Promoter ratio

The ratio of promoter usage was measured for each of the samples. No consistent change was observed to result from the presence of v-myc. The ratio was ~10:1 (sd 4.5) P2:P1 for each of the deletions except -101. In the transfections with this construct, the P1 specific signal was noticeably reduced, leading to a ratio of promoter usage in favor of P2. It is difficult
to quantitate this value due to low signal strength and high background levels.

3.1.3.3 Deletion constructs -91 to +430

1 Effect of v-myc

RNAase protection analysis of the transfected human c-myc gene fusions with endpoints from position -91 through +430 (into the βIFN sequences) is shown in Figures III.11 and III.12. The two figures respectively correspond to long and short exposures of the same gel. Figure III.11 optimises the P2 promoter response, and III.12 that of the PI promoter and upstream initiated transcripts. The effect of v-myc expression on these samples is more ambiguous than on those of the previous figure. Both the PI and P2-specific signals of the -91 and -88 endpoint samples demonstrate clear repression. The first sample with a -86 endpoint shows a drastic degree of repression in the presence of v-myc, an observation which is not repeated in the second sample, which appears less repressed than many of the other samples, especially with regard to the PI specific signal. The signals in the -58 sample are less consistent, with only one of the two transfected pools displaying clear repression. A similar situation is seen with the +119 endpoint deletion, which appears to repress less well than the other constructs. This deletion contains the P2 but not the P1 promoter, and the P2 dependent signals appear to be only minimally repressed by v-myc.

Once again the expression level of the upstream initiated transcripts appears to mirror that of the promoter-specific signals (Figure III.12). The -91 and -88 samples demonstrate a reduction in signal upon the introduction of v-myc, while the remaining three deletions show reduced degrees of responsiveness.

2 Promoter usage
The absolute amount of P2 initiated transcript is approximately constant in each of the control virus infected pools; none of the deletions produces a consistent change in mRNA levels (Figure III.11). This is not the case for the P1 initiated transcripts, where the constant level of signal seen in the -91 and -88 samples is reduced to a different, but constant, level in the -86 and -56 lanes (Figure III.12). This change is not very large, perhaps 2-3 fold. Surprisingly the ratio of promoter usage is different in the first -91 deletion sample of Figure III.12 than in the -91 samples of Figure III.10. The same RNA sample was analysed in each gel using the same probe, so the results should not differ. Yet the ratio for the -91 construct in Figure III.10 is ~4:1 P2:P1 and in the Figure III.11/12 samples it is ~1.4:1 P2:P1 after correcting for the size of the bands.

3.1.3.4 Characterisation of c-myc and α-globin signals

The same RNA samples analysed in Figures III.10, III.11 and III.12 (5' endpoint deletions -348 to +430) were assayed in parallel for the expression of the endogenous c-myc gene and the transfected α-globin control (Figure III.13). The probe used to measure endogenous rat c-myc gene expression levels maps the 5' ends of both P1 and P2 initiated transcripts. In the 22 sets of pools analysed, a 5-20 fold reduction of both the P1 and P2 signals was detected in all samples following infection with v-myc.

Contrary to expectations, the α-globin signals do not provide a simple, consistent reference signal, but rather vary dramatically amongst transfections. The variability seen can be at least 100 fold between individual transfections. For example the signals detected in the -88 deletion endpoint samples are barely detectable, while those of the +119 transfection saturate the film in the same exposure. There appears to be no correlation between α-globin signal levels and any parameter. Even in duplicate experiments with the same human c-myc promoter deletion the α-globin expression showed remarkable diversity (eg the -276 or -86...
endpoints). In addition there is no straightforward effect of v-myc expression on the levels of α-globin expression. In some cases the control and v-myc virus-infected pools show equivalent levels of α-globin expression (eg -348 pair 1, -91 pair 1). In others the v-myc infected signals are stronger (eg -348 pair 2). While in yet others the control virus-infected pools have a stronger signal (eg -169 pair 2). The average change in α-globin signal in the presence of v-myc as quantified by densitometry, bearing in mind that some of the extreme signals are beyond film linearity, is a 1.1 (sd 0.8) fold decrease when all of the pairs (n=22) are considered. This result does suggest that even though there is great diversity in expression levels, there is no effect on expression attributable to v-myc.

Variations in the α-globin signal between two duplicate transfections do not seem to correlate with any corresponding variations in the control virus infected myc/βIFN signals. For example the -91 deletion endpoint transfections have a significant difference in the amount of α-globin signal between the two transfections, and minimal variation between the control or v-myc virus infected lanes (Figure III.13). The same samples display very consistent basal levels of either P1 or P2-dependent myc/βIFN signal, and also have similar, though repressed, levels of expression in the presence of v-myc. When the two types of signal are compared across all of the samples the general trend is that the human c-myc promoter dependent bands are relatively invariant, regardless of the 5' endpoint of the construct, while the α-globin signals, which are all transcribed from the same construct, vary greatly.

3.1.3.5 Conclusions

The implication of the results from this deletion series imply that in order for a human c-myc promoter construct to be repressed by the presence of v-myc, it is sufficient for a fusion construct to contain the
approximately 300bp of human c-myc sequence from positions -88 through +208. These sequences appear sufficient for correct initiation and regulation of transcripts from both the P1 and P2 promoters. While other deletions with endpoints closer to the promoters have been assayed (-86, -58, +119), the results are too ambiguous to definitively assign any of them as responders.

3.1.4 Deletion constructs -88 through +97

3.1.4.1 Introduction

To determine the 5' boundary of the response element(s) to greater resolution, an additional 5' deletion series was constructed. The start point of this series was the pUC -88/+208 IFN construct. Serial deletions of approximately 20 to 30 bp were synthesised for analysis (Figure III.14). In total, deletions with 5' endpoints at -88, -64, -13, +23, +43, +69 and +97 were assayed.

3.1.4.2 Effect of v-myc

Data for deletions through +97 are shown in Figure III.15. In addition to the RNA from the infected transfections, samples were also prepared from the parental transfected pools at the time of infection. These samples were probed for myc/IFN, α-globin and γ-actin. In each case the IFN signal is reduced in the presence of exogenous myc. Surprisingly the signal is also reduced between the parental cells and those infected with the control virus. The degree of signal reduction is of a similar magnitude as the loss of signal due to the presence of v-myc. The α-globin signals in the pools behave variably. In some transfections (eg -88, +69), there is a progressive loss of signal that parallels that seen between the parental, control virus, and v-myc infected samples. This contrasts with others (eg -64, -13, +97) in which the α-globin band is of the same intensity in all three lanes.
3.1.4.3 Implications of passage dependent alterations of signal

The loss of signal from the parental to the control infected samples was seen in each of the seven pools assayed. This variation may be systemic, since the time between infecting with the viruses and harvesting RNA samples was about two weeks. It is conceivable that there is a time-dependent loss of signal from transfected DNA in these cells, a phenomenon which has been documented previously (Pellicer, et al, 1980, Scangos, et al, 1981, Wigler, et al, 1979a). As it is known that introduction of v-myc into Rat-1 cells leads to an acceleration in their growth rate (LJZ Penn, personal communication), the two infected populations, while cultured, selected, passaged, seeded and harvested in parallel, may undergo a different number of cell divisions which correlates with the presence of v-myc. If that were the case, we might expect to see the reduction of signal observed. Thus the effect of v-myc on indicator expression may be due to an indirect activity of the oncogene. If this were the case, signal loss might be expected to be observed for all transfected indicators. Thus the α-globin signal, derived from a cotransfected but independent gene, should vary in concert with the myc dependent signal. This does not, however, appear to happen, as the α-globin and myc dependent signals do not vary in concert in many of the pools analysed. This result argues against a completely general time dependent signal degradation phenomenon. However it does not exclude the possibility that it is specific to the myc promoter.

The α-globin signals were found to fluctuate temporally within individual pools of transfectants, in addition to the variation in absolute magnitude of α-globin signal noted previously. Because of the many inconsistencies of this signal, its value as a control for the myc-dependent expression was greatly compromised. To avoid further complications, it was omitted from subsequent experiments.
Bearing in mind the caveat that the reduction in myc/IFN signal may be due to some less than direct effect of v-myc expression on cell growth, the combined results of the constructs described above suggest that all of the constructs with 5' deletions through +97 are capable of imparting myc responsiveness to the human c-myc promoter. The one 5' deletion more proximal to the P2 cap site which has been studied, +119, remains unassigned in its responsiveness.

3.1.5 Reduction of non-specific upstream signal

3.1.5.1 Introduction

Upon the introduction of exogenous genes into cells via transfection, it is not unusual to observe transcripts which do not initiate from the normal start points, eg (Goodbourn, et al, 1985, Treisman, 1985). However, in most instances these erroneous transcripts have been found to be a small fraction of the appropriately initiated transcripts under study, and their source is usually not investigated. Because the transcripts which map 5' of the end of the human c-myc probe were as prominent as the properly initiated ones, and because they also appeared to be regulated by v-myc, we felt it prudent to investigate their origin, and, if possible, eliminate them. The maintenance of specific, myc responsive signals in the absence of these signals would increase confidence that the myc response already observed was not artefactual. While it was desirous to reduce these transcripts, one interesting possibility suggested by their regulability is that mechanistically, down-regulation may occur through a physical block to the passage or entry onto the DNA of the transcription complex. There are several alternative explanations for the origin of these transcripts. One possibility is that the transfected DNA contains cryptic promoter sequences. Alternatively, the lack of specific eukaryotic control elements in plasmids of bacterial origin may lead to random initiation throughout the transfected DNA. Another explanation may be that
transcripts initiating within the host genome are being transcribed through the transfected DNA. Several approaches were taken to differentiate between these alternatives.

3.1.5.2 Characterisation of upstream signals

To address possible initiation within the vector upstream of the human c-myc sequences, a mapping experiment was performed with an RNAase protection probe (BSKS+SspI) derived from the ~650 nt of the Bluescript vector immediately 5' of the site into which the human c-myc promoter fragment was subcloned. Samples probed with this fragment include RNA derived from cells transfected with a -101 5' endpoint deletion as well as the +430 deletion which does not contain any human c-myc sequences. Colony pools were either infected with the LJ or VM viruses or, alternatively, made confluent and serum starved and then parallel cultures treated with 10% serum or left unstimulated (Figure III.16). The probe protects a strong band which corresponds to the entire region of homology between the cloning site and the 3' end of the probe in all of the samples. This indicates that the majority of the transcripts are initiating outside of this region of the vector but does not distinguish between plasmid or cellular origins. The smear beneath the band is primarily due to degradation of the probe, but may also reflect some heterogeneity of transcription start sites within the vector sequences.

The LJ virus infected samples of both constructs display signals of similar intensity. Upon v-myc infection the human c-myc promoter-containing samples display some repression, though not as much as the appropriately initiated signals. This response is not seen in the +430 deletion samples, especially when the slight underloadings of the VM samples are taken into account. The serum stimulated signals show a strong induction in the -101 lanes that is echoed in the +430 lanes. (NB The γ-actin signals differ as γ-actin is itself serum responsive.) This is
somewhat surprising as the +430 samples contain no myc promoter sequences, and would thus not be expected to display a serum response. Additionally, the magnitude of the 'basal' signals is greater in the samples derived from confluent, serum starved cells than in those from LJ infected, proliferating cells (the two halves of the figure have different exposure durations).

Thus these transcripts primarily initiate at some point greater than 600 bp upstream of the point at which the inserts have been cloned into the Bluescript vector. It is not possible from these results to distinguish between an origin within the genomic sequences or even farther into the plasmid. As expected from previous data (eg Figure III.11 vs III.7) the response to exogenous myc is dependent on the presence of c-myc promoter sequences. Surprisingly, the serum response seems to be independent of these same sequences; it appears to be a completely autonomous event, requiring no human c-myc promoter elements. As a result of this finding, the serum response investigation was pursued no further.

3.1.5.3 Use of alternative vectors

To determine if the upstream transcripts were specific to the Bluescript vector itself, the myc promoter interferon fusion fragment was transferred into two additional vectors, pUC13 and pGC1 (Figure III.17). These vectors are similar to the Bluescript vector in that they contain the same bacterial origin of replication and ampicillin resistance marker, but do not contain all of the other components. The -88/+208 fragment was used as it contained the shortest unambiguous feedback response element thus far identified, and the +119/+208 fragment in order to clarify its response to v-myc. The orientation of the inserts was different in pGC1 and pUC13 to further increase their dissimilarity.
RNA from cells transfected with these constructs was assayed following the introduction of exogenous myc (Figure III.18). In the pGC1-88/+208 and pUC-88/+208 samples there is a band which corresponds to promoter specific RNA, as well as one which corresponds to non-specific, upstream initiated RNA. Both of these messages are down regulated in the v-myc infected cells. The pGCl+119/+208 and pUC+119/+208 constructs both display strong non-specific, upstream initiated signals and almost undetectable P2-specific signals. Due to the weak intensity of the P2-specific bands it is not possible to determine whether they are repressed in response to v-myc expression.

It thus appears that it is not possible to eliminate the non-specific band merely by swapping to a different vector backbone. This implies that the transcripts most likely initiate within adjacent cellular sequences. To strengthen this conclusion, linear DNA containing only the myc promoter and interferon sequences was transfected into cells which were then used in the assay. Figure III.19 shows that P2 specific initiation as well as upstream transcripts can still be detected, and are both repressed in the presence of v-myc. While the upstream transcripts may indeed initiate within the cellular DNA, it is surprising that they should be responsive to the presence of exogenous myc.

3.1.5.4 Suppression upstream transcripts via SV40 terminator sequences

As the vector swapping approach had proved unsuccessful, an alternative method of reducing the upstream initiated, non-specific transcripts was attempted. A sequence from the SV40 virus containing a tandemly duplicated polyadenylation signal has been found to reduce upstream transcripts in a different vector (de Wet, et al, 1987). This fragment (called SVter) was subcloned into the pUC -88/+208 IFN construct immediately upstream of the myc sequences. The pUC vector
was chosen as it appeared to have a weaker upstream signal than the pGC1 plasmid.

Analysis of the RNA from these constructs is shown in Figure III.20. The upstream transcripts detectable in the SVter lanes were found to be dramatically reduced compared to those in the control lanes and those which remain no longer down regulate. The magnitude of the P2-dependent bands, however, is not significantly affected by the presence of the SVter sequences. In addition, they continue to respond to the presence of v-myc. Because of these results it was decided to assay subsequent promoter constructs in this type of vector.

3.1.6 Clonal transfectant analysis

Although at first sight, it would appear that all of the c-myc promoter constructs containing sequences with a 5' end from 2.3 kb through 88 bp upstream of the P1 promoter and ending 208 bp downstream of this promoter are capable of being repressed by exogenous myc, some of the control experiments performed do not provide satisfactory evidence to allow this interpretation to be presented with total confidence. A transfected myc promoter does not appear to be regulated in the same manner as a transfected α-globin promoter with regard to both the efficiency of its use between transfections, and the constancy of its activity within one pool of transfected colonies. To resolve some of these issues, it was decided to analyse the performance of one construct in clones of transfected cells, to facilitate our understanding of the behavior of the system. The questions we wished to address include: Is this expression level related to transfectant copy number? What percentage of drug resistant cells contain the indicator construct? Is the endogenous c-myc gene of all clones repressed by v-myc? If so, do all clones which express the transfected construct respond to exogenous myc in the same manner?
Can we use these results to design a better assay or to help interpret the previously acquired data?

Twelve colonies which were cotransfected with the pUC SVter -88/+208 IFN construct and a selectable marker were ring cloned and expanded. These cell populations were then infected with the v-myc or control viruses. RNA was isolated from the cells before infection and from each of the infected pools. Analysis of RNA from four clones is shown in Figure III.21. The samples shown were chosen as representatives of the different types of behaviour seen. All four clones show no change in expression between the parental and control virus infected samples, in notable contrast to the variations seen with the colony pools. The myc promoter-dependent IFN signal of clones 2, 7 and 11 is reduced in the presence of v-myc, while it is unchanged in clone 6. In those samples where a clear P1-dependent signal is seen, it responds in concert with the P2-dependent signal. The 'basal' level of expression for each of the clones is highly disparate, as are the levels of readthrough transcript (though in each case this is a fraction of the P2-dependent specific signal). Clones 6 and 7 show strong IFN bands and clone 2 a moderate signal while that of clone 11 is quite weak. Of the twelve clones analysed, three had undetectable, four weak, three moderate and two relatively high basal levels of expression. Of these, six appeared to down-regulate transfectant mRNA and three not in the presence of exogenous myc, with no apparent correlation between the level of basal signal and the myc response. Furthermore, none of the clones showed an increased signal after v-myc infection. The responsiveness to v-myc of the endogenous c-myc gene, as well as the presence of gag-myc specific RNA, for each the clones is shown in Figure III.22. All four clones behave identically, with a reduction of endogenous c-myc signal upon introduction of v-myc.
Southern blot analysis was performed on DNA from each of the clones. The DNA was digested with BamHI and PstI, enzymes which precisely excise the -88/+208 human c-myc promoter fragment, and filters were probed with this same fragment (data not shown). Ten of the twelve clones had a detectable signal, which should be proportional to the number of copies of integrated DNA. The signal varied from approximately one copy equivalent for six of the samples, through five to ten copies for two samples, with two clones having fifty to greater than one hundred copies of integrated DNA. There was no correlation between copy number and response to exogenous myc. For instance, the non-responsive clone 6 and responsive clone 7, each have approximately one copy. The only correlation which may exist is an inverse one between copy number and basal expression level of the transfected gene. The two clones with the highest copy numbers each have barely detectable levels of expression. These results are summarised in the table below:
<table>
<thead>
<tr>
<th>Clone number</th>
<th>Copy number</th>
<th>Relative basal expression level</th>
<th>Myc Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>++</td>
<td>+</td>
</tr>
<tr>
<td>3</td>
<td>~50</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>5</td>
<td>~100</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>+++</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>++++</td>
<td>+</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>++</td>
<td>-</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>+++</td>
<td>+/-</td>
</tr>
<tr>
<td>11</td>
<td>3</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>N/A</td>
</tr>
</tbody>
</table>

3.1.7 Signal degradation timecourse

When infected immediately following pooling, pools of colonies transfected with human c-myc promoter fusion constructs demonstrated a reduction in transfected myc gene expression even in control virus infected cells, yet none of the cell clones analysed demonstrated this phenomenon. If the stability of expression could be engineered into the assay, it would be possible to be quite confident of the deletion mapping results. We hypothesised that the strength of the transfected signal stabilises within the time required to expand a clone to approximately 2x10⁶ cells, which is ~2 weeks following isolation. To determine if pools behaved in a similar manner, a transfected population containing the pUCSVter-88/+208IFN construct was passaged twice a week for several weeks, with RNA harvested following each passage. After 6 passages the pool was infected with control or v-myc virus and RNA isolated following
the standard selection protocol. Analysis of the myc promoter dependent signal is shown in Figure III.23. A gradual reduction in the P1 and P2 signals is apparent, with minimal change between passages 5 and 6. The signal is reduced upon addition of v-myc compared to the LJ infected control, and to a greater degree than is seen between passages. Thus the major mechanism appears to be attributable to promoter responsiveness, while there may be a small extent of signal loss due to the increased cell growth caused by v-myc.

3.1.8 -88SVter through +119SVter deletions

The information obtained in the previously described series of experiments was subsequently used to refine the assay for myc-dependent repression: Deletion constructs were exchanged into vectors containing upstream SVter sequences, and pooled colonies were passaged for extended periods to allow maximal stabilisation of signal. The modified protocol is shown below:
Cotransfect myc promoter construct and hygromycin resistance gene.

Hygromycin selection ~10-14 days

Pool >100 stable hygromycin resistant colonies.

1 day

Passage

>2 weeks

Infect with v-myc or control retrovirus.

G418 selection ~10 days

Pool >100 G418 resistant colonies.

~24 hours

Harvest total cellular RNA from subconfluent, growing cells.

Quantitate myc promoter dependent RNA by RNAase protection.
From the previous deletion analyses, including the results of the pools and clones it was apparent that human c-myc sequences contained between -88 and +208 are sufficient for a v-myc response, and that as small a region as +97/+208 may be sufficient. The higher resolution deletion results required further investigation, and thus 5' deletions with endpoints to +69, +97, +100, +112, +119 and +150 were constructed into the pUCSVterIFN vector (Figure III.24), and analysed via the procedure described above. The results using these constructs are shown in Figure III.25.

P2 specific transcripts are reduced in the presence of v-myc in each of the deletions through +112. The +119 deletion appears to show no change between the control and v-myc infected cells, though the samples have a low signal/noise ratio. The +150 construct shows no P2-specific signal, as expected. Basal transcript levels are constant through the +97 deletion, and drop sharply for the remainder. However this pattern of basal transcription was not reproducibly seen (see eg Figure III.27 below, and data not shown). In order to attempt to resolve whether the +119/+208 construct would respond to v-myc expression, four additional pools of transfectants were assayed. The P2-specific signal in each of these cases was again very weak and could not be clearly resolved from the background due to excessive levels of upstream initiated transcript (data not shown). Whether the +119/+208 deletion contains v-myc responsive sequences thus remains unclear, though it seems that this is not the case.

The behaviour of the upstream initiated transcripts from these constructs was totally unexpected. As seen previously, they were quite weak in the -88 endpoint deletion. However, as the SVter sequences were moved in closer proximity to the P2 cap site, the magnitude of the upstream transcripts increased markedly, until upstream signal became significantly stronger than the specific signal in the +100, +112 and +119
constructs, and the SVter sequences appeared to be completely ineffectual. It is perhaps significant that these readthrough transcripts appear myc-responsive for all of the deletions except the +119 deletion, in which the P2-dependent signal may not be myc responsive either.

3.1.9 Linker scan series

The combined results of the 5' deletions from -88 through +119 suggest that sequences between +112 and +208 are required for a v-myc response. Whether the +119/+208 fragment will suffice is ambiguous as in some experiments (Figure III.11) it appears to weakly respond to v-myc, while in others (Figure III.25) it does not. This is still a rather large region of the promoter within which a response element might reside. If the +112:+119 region actually delimits the 5' end of an element, further deletions would not contribute much greater resolution. It was therefore necessary to determine how much additional sequence is required to maintain a v-myc responsive phenotype. Two simultaneous approaches were undertaken to address this issue. A series of five pairs of complementary oligonucleotides spanning the human c-myc sequences from positions +94 through +141 (-71:-13 relative to P2) were synthesised. One pair contained the wild type sequences, while the remaining four contained a series of linker scanning mutations. These mutant oligonucleotides contain 9 bp linkers composed of the non-palindromic Bsm I recognition sequence (GAATGCCC↓TC). The linkers are positioned such that they consecutively mutate the 36 bp immediately 5' of the P2 TATA box, from positions +97 to +132 (-66 to -31). The 3' end of each pair contains a Bam HI linker which maintains the correct spacing of the human c-myc sequences when cloned 5' of deletion +150. The sets of oligonucleotides were cloned upstream of either the +150 deletion (+150 Is series) or a Bam HI linkered Pst I site (+283 in IFN) of the IFN gene (Pst Is series), in the pUCSVterIFN vector (See Figure III.26). These two linker scan series thus serve the dual
purpose of creating a 3' deletion while simultaneously providing for the mutation of most of the sequences upstream of the P2 tata box which are conserved in the minimal responsive fragment identified thus far (+112/+208).

Results of at least duplicate transfections of the +150 ls series and repeats of the +97 and +112 deletions are shown in Figure III.27. P2-dependent signals are observed for each of the linker scanning constructs, though their intensity is much weaker than the +97 or +112 signals. A v-myc induced reduction of signal is seen for at least one experiment with each of the constructs. However, one transfection of both the ls2 and ls3 constructs does not appear to be repressed, and in both of these cases the signal is noticeably stronger than most of the other transfections.

Results from duplicate transfections of the Pst ls series are shown in Figure III.28. P2-dependent signals are visible for the all of the constructs, though the ls1 and ls2 signals are stronger than those of ls3, ls4 and ls5, which in some samples have no observable signal. In each case where there is detectable expression, and most clearly with ls1 and ls2, it is reduced following v-myc addition, though the ls3, ls4 and ls5 results must be treated with caution as the starting signals are so poor.

Taken together, these two series of linker scanning constructs do not identify any sequence which is absolutely required for a v-myc response. The Pst series results indicate that no human c-myc sequences 3' of the P2 TATA box are essential for v-myc responsiveness. However the basal levels of expression from these constructs, especially when any of the sequences between +97 and +123 are mutated, are very difficult to detect, and imply that there is some activity attributable to the sequences between positions +141 and +208. The inability to detect a loss of v-myc responsiveness with any of these mutations suggests that either there are multiple, redundant responding factors which interact within this region,
or that the TATA box sequences, which are not mutated in any of these constructs, mediate the response directly.

3.1.10 Effect of SV40 enhancer

We decided to test the effect of adding an SV40 enhancer sequence 5' of the human c-myc promoter fragment. Two questions were addressed: First, does the addition of enhancer sequences increase the ease with which myc promoter dependent signals can be detected, and second, does the enhancer abrogate the response of the human c-myc promoter fragment to v-myc?

An SV40 early promoter fragment of ~200 bp containing the 72 bp repeats of the SV40 enhancer (Treisman, 1985) was cloned into the pUC-88/+208IFN construct directly 5' of the -88 position (pUCSVE-88/+208IFN, 'SVE') or 5' of the SVter sequences (pUCSVESVter-88/+208IFN, 'SVESVter'). The plasmid pJ123hcmyc (Land, et al, 1986) which contains the SV40 early promoter, including the enhancer and promoter proximal sequences, cloned 5' of the 8.1 kb human c-myc HindIII/EcoRI genomic DNA fragment used in the experiments described previously (eg plasmid pGC1H/R), was also assayed (Figure III.29).

Analysis of RNA from pools of colonies transfected with these constructs and infected with the appropriate viruses is shown in Figure III.30. The P1 and P2 dependent signals from the pJ123hcmyc transfectants are readily detectable, and are repressed ~2 fold in the presence of v-myc. The signal from this construct is significantly stronger than that seen from a similar construct in the absence of the SV40 sequences (ie only an overnight exposure was required to detect the signal). Even though this construct contains an intact SV40 early promoter, upstream initiated transcripts which might have initiated from the SV40 early cap site are much weaker than the human c-myc P1 initiated transcripts. While direct
quantitation of the human c-myc protein levels expressed from this construct have not been performed, since this construct will cooperate with ras proteins in an REF-cotransformation assay (Land, et al, 1986), it is probable that sufficient levels of myc protein are expressed to allow autorepression, which may reduce the extent of the response to v-myc.

The SVESVter and SVE transfections both display readily detectable P1 and P2 signals which are reduced several fold in the v-myc infected pools. There is a dramatic increase in the magnitude of the P1 signal in the pUCSVE pool, when compared with either the SVESVter pool or a pUCSVter-88/+208IFN pool shown for comparison. The magnitude of the P2 signals is similar between all three pools. Despite the change in P1-dependent transcript levels, they are still repressed in the presence of v-myc. Thus the ratio of P1:P2 transcripts is altered by the SVE sequences from <1:10 to ~1:1, and this change is prevented when the SVter sequences are placed between the SVE and human c-myc sequences. It remains unclear whether the SVter sequences are specifically required for this difference, or whether it is solely a distance effect. Furthermore, the presence of the SV40 enhancer sequence does not compromise the response of the human c-myc promoter sequences to v-myc expression.

3.2 Discussion and Conclusions

The data presented in this chapter have established that sequences from the human c-myc promoter are sufficient to drive the expression of a heterologous indicator in Rat-1 cells, and allow the suppression of this expression by v-myc protein when it is present at high levels in the same cells. Sequences which can impart this responsiveness have been mapped within 70 bp of the human c-myc P2 promoter mRNA cap site.

3.2.1 Response elements

The sequences identified as functional targets for myc autoregulation include two overlapping fragments proximal to the P2 promoter. These
sequences span positions -71 through -22, as well as positions -51 through +47 (see Table III.1 for a summary; positions relative to the P1 promoter cap site are in parentheses). Experiments deleting further sequences from the 5' end of the c-myc promoter constructs to position -44, while retaining sequences through position +47, reduced the level of P2-dependent transcript to below the limits of detection. Mutation of the -71 to -22 sequences in four consecutive nine bp regions (the +150 linker scan series), in the presence of downstream sequences through position +47, demonstrated that none of the four linker-scanning mutations was individually capable of eliminating responsiveness to myc levels. In the absence of the downstream sequences (the Pst linker scan series), it also appears that none of the linker scanning mutations affects repression of the P2-dependent transcripts, although the linker scan constructs ls3, ls4 and ls5 all reduced the basal levels of transcription to an extent impairing a clear interpretation.

Whether the sequences from -21 to +47 are involved in the response is not clear. The responsiveness of the -71/-22 linker scan constructs (Pst series), which do not contain the -21 to +47 sequences, implies that they are not absolutely required (Table III.1). However, it has not been demonstrated that these sequences are dispensable for the response when all of the sequences between positions -71 and -51 are mutated or deleted, as the linker scan constructs ls3, ls4 and ls5 individually contain mutations of only nine bp in a region spanning twenty bp. It is similarly possible that an independent suppressor element is located in the -21 to +47 region. Thus, while no single sequence element has yet been identified as essential for the response, it remains possible that multiple redundant, and possibly overlapping, elements, within the P2 proximal sequences act to mediate the downregulation.
The only sequence which is unchanged in all of the constructs assayed is located between positions -30 and -22 (5' TATAAAAGC 3'). This sequence consists primarily of the P2 TATA box, and was left unaltered as it is required for basal transcription from the P2 promoter (Asselin, et al, 1989). It is thus formally possible that this is the only sequence required for a myc response. However, if the non-responsiveness of the 5' deletion to position -44, proves to be reproducible, then sequences upstream of position -44 would also be required for suppression. Furthermore, since the linker scan data indicate that mutation of one single nine bp upstream sequence will not completely impair myc responsiveness, it is possible that the TATA box in combination with at least one of a number of possible additional elements is required. Redundancy in the mechanisms governing such a central control point as a negative feedback loop would make biological sense, as it would protect the cell from the potentially deleterious effects of single mutations.

TATA boxes were originally identified as AT rich regions which are very important in the correct initiation of transcription of a number of eukaryotic genes (Breathnach and Chambon, 1981). TATA boxes are involved in both the positioning of the mRNA cap site and in mediating basal levels of transcription. In higher eukaryotes transcription initiates approximately 30 bp downstream of the TATA box. TATA boxes have also been implicated in the regulation of some genes either as direct targets for regulatory signalling pathways, or as indirect targets which mediate the interactions of other proteins which interact with upstream or downstream sequence elements (Greene, et al, 1987; Wu, et al, 1987). The consensus TATA sequence is TATAA, although many variants exist (Bucher and Trifonov, 1986). There are genetic data which support a functional distinction of TATA sequences; such specificities have been documented for both direct and indirect regulation (Chen and Struhl,
For instance, the hsp70, c-fos and adenovirus E1b genes can all be activated by adenoviral E1a protein through their respective TATA boxes, and the sequence of the TATA boxes is important in determining responsiveness (Simon, et al, 1988, Simon, et al, 1990, Wu, et al, 1987). Similarly, activation of the S. cerevisiae his3 gene by the GCN4 protein, or the human myoglobin gene by an upstream muscle specific enhancer, are both dependent on the presence of specific TATA sequence, although this sequence is not the direct target for these activating mechanisms (Chen and Struhl, 1988, Wefald, et al, 1990). The sequence TATAAAA, which is identical to that of the c-myc P2 promoter, has been identified as one through which some of these stimuli can be mediated (Simon, et al, 1988, Simon, et al, 1990, Wefald, et al, 1990). It is thus possible that this sequence defines a general class of regulatable TATA elements. It seems unlikely, however, that it is a direct target for myc autosuppression as this sequence is present in numerous genes, including 'housekeeping' genes such as β-actin, which are unlikely candidates for regulation by myc protein. Furthermore, although the sequence TATAAAA, is the most conserved of the TATA sequences (Bucher and Trifonov, 1986), regulatory targets of myc protein have not been frequently identified. It will be interesting to discover the role of the myc TATA box sequences in the autoregulatory response.

The co-mapping of the putative myc responsive sequences with those which are required for detectable levels of c-myc P2 promoter-dependent transcription has made it difficult to discern an absolute specific requirement for any single sequence in autosuppression. Separation of these two activities, if possible, would facilitate the dissection of the mechanisms governing myc repression.
3.2.2 Upstream initiated transcripts

Transcripts which initiate upstream of the human c-myc sequences were frequently detected. The level of these transcripts was generally found to be down-regulated by myc protein coordinately with the human c-myc promoter-dependent transcripts, in a manner which appears to be dependent on the presence of human c-myc promoter sequences. The responsiveness of these transcripts to elevated myc levels may be a valid indicator of the responsiveness of the human c-myc promoter sequences. If this is so, then the construct which contains human c-myc sequences from position -44 to +47 (+119 to +208 relative to P1), the responsiveness of which is still in question, could be considered nonresponsive, as the upstream transcripts were expressed at constant levels regardless of the myc protein concentration (Figure III.25), and conversely, the linker scan constructs of the Pst series, ls3, ls4 and ls5 all demonstrate a repression of their upstream initiated transcripts, even though the P2-specific transcripts are almost undetectable (Figure III.28).

The efficient down-regulation of upstream-initiated transcripts in a human c-myc promoter-dependent fashion suggests a possible model for the mechanism of repression. This model would postulate that a physical block to the progress of all transcription complexes forms on the promoter, preventing not only the productive transcription of c-myc promoter dependent transcripts, but also the elongation of the upstream initiated transcripts. The prematurely terminated upstream transcripts would not be polyadenylated, and thus would be unstable and undetectable, as is observed.

Although it has been possible to extract some potentially significant mechanistic information from observing the regulation of the upstream initiated transcripts, the main focus of these studies must be the control of transcripts which initiate at the correct c-myc mRNA cap sites. The
upstream initiated transcripts were often found to be a problem, as their magnitude was often significantly stronger than those of the P1- or P2-dependent transcripts, and degradation products of the protected probes often obscured the specific bands of interest, particularly when the human c-myc promoter fragment was promoting only weak levels of specific transcript. This problem could be circumvented by the inclusion of transcription termination sites derived from the SV40 virus, which reduced the detectable levels of the upstream initiated transcripts. The extent of this blockage, however, was found to be dependent on the distance of the terminator sequences from the human c-myc P2 promoter cap site, with a greater distance between the P2 cap site and the terminator sequences resulting in a more efficient reduction in upstream initiated transcript (Figure III.25). Thus in the case of the weakest human c-myc P2 promoter dependent signals, even in the presence of the terminator sequences, the upstream initiated transcripts were extremely deleterious to the ability to measure a regulatory response. The reason for the distance dependence of the terminator effect is unclear. It will be interesting to learn if the placement of a heterologous 'stuffer' sequence upstream of short 5' flank sequences is sufficient to block these transcripts, thus eliminating the background problem.

Levels of upstream transcripts could also be reduced by the presence of SV40 early promoter and enhancer, or enhancer sequences alone, upstream of the human c-myc promoters. As these sequences did not prevent repression of the transfected promoter construct (Figure III.30), their inclusion may be a useful alternative approach for reducing the upstream transcript signals.

3.2.3 Magnitude of repression

The endogenous Rat-1 c-myc gene is repressed approximately ten-fold upon introduction of v-myc protein using the VM virus. In contrast, the
transfected constructs used in the assay presented here were repressed approximately three-fold in the vast majority of cases. The difference between the magnitude of repression of the transfected and endogenous genes can be attributed to several reasons. The first is the demonstrated variation in responsiveness of various transfected cell clones (Figure III.21). Only approximately 75% of the clonal transfectants analysed were found to respond to an increase in myc protein expression, and of those which did respond, not all were repressed to the same extent. This variability in behavior is probably due to the positions in the genome into which the reporter plasmid has integrated. The amount of repression in a pooled population is thus the sum of the responsiveness of many individual clones.

The clonal heterogeneity in responsiveness poses an additional burden on the assay since the signal being assessed is repressed and not stimulated, leading to a high background signal. For instance, if 75% of signals are repressed ten-fold, and the remainder not at all, we would expect to detect a residual signal of 1\-(0.1)(0.75)+0.25=0.325, which is approximately one-third of the initial signal or a three fold reduction. In an inducing situation, with the same relative inductions we would expect to detect 1+(10)(0.75)+(0.25)=8.75, an almost ten-fold induction. Thus it is not unreasonable for a three-fold reduction to be considered significant within the constraints of the assay system.

3.2.4 Elements required for minimal levels of transcription

The basal level of expression of the chimaeric human c-myc/indicator constructs in proliferating cells was relatively invariant whether several kb or only approximately one hundred bp of human c-myc P2 promoter are present (See Figure III.10-12, 25). However, within 70 bp 5' of the P2 cap site there appear to be multiple elements which are required for maintenance of basal transcription in addition to potential target
sequences for myc regulation. The basal signal was maintained in deletions through position -66 relative to the P2 promoter cap site so long as the constructs extended to position +47. Expression from deletions to -63 and -51 was detectable, although it was sometimes weaker than that observed with the -66 5' endpoint construct. Further deletions to position -44 produced barely detectable levels of transcription. The linker scan (Pst) constructs, which contain human c-myc sequences from -71/-22, revealed that replacement of any of the sequences between -66 and -40 in nine bp blocks led to a decrease in specific signal (Figure III.28). As this region spans 27 bp, and many transcription factor binding sites are on the order of ten bp in length, it is probable that multiple factors interact with these sequences. The basal signals from these constructs were more difficult to detect than those from the +150 linker scan series, which contain human c-myc sequences through position +47 with the same linker insertions. This may be the result either of positive regulatory elements located within the sequences downstream of the TATA box, or from the use of the normal c-myc mRNA cap site in the +150 construct series.

The identification of significant control elements located within the sequences between -66 and -40 is consistent with the findings of several other studies. The region from -66 to -58 has been identified as a positive control element in the human c-myc promoter in both cell line transfections (Lipp, et al, 1987, Thalmeier, et al, 1989) and Xenopus oocyte microinjections (Nishikura, 1986). These sequences have also been implicated stimulation of the c-myc promoter by serum growth factors (Mudryj, et al, 1990) and the adenovirus E1a oncogene (Hiebert, et al, 1989, Lipp, et al, 1989, Thalmeier, et al, 1989). The -57 to -40 region has been identified as a positive regulatory sequence of the murine c-myc promoter in transfection and in vitro transcription assays (Asselin, et al, 1989, Hall,
1990), and also contains sequences implicated in the serum response of the human c-myc gene (Mudryj, et al, 1990).

The constant levels of transcript detected whenever constructs contained more than 100 bp of upstream human c-myc sequence implies that there are no functional control elements outside of the P2 promoter proximal sequences. This finding is consistent with those reported in studies of human and murine c-myc promoter regulation both in cell lines and Xenopus oocytes (Lipp, et al, 1987, Nishikura, 1986, Thalmeier, et al, 1989), but contrasts with others which have identified positive and negative control regions both upstream and downstream of the c-myc promoters (Hay, et al, 1987, Remmers, et al, 1986, Yang, et al, 1986). The cause of the inconsistencies between these findings are not obvious, but it is possible that each of the different systems studied uses a different set of transcriptional and posttranscriptional controls to regulate c-myc expression.

3.2.5 Sequence homologies

The c-myc P2 promoter-proximal sequences of several mammalian species are highly conserved (Table III.2). If one assumes that the regulatory controls on c-myc expression which are mediated through these sequences are similar in each of these species, then the regions of highest conservation may give some indication of the functionally most important continuous sequences. Both up- and downstream of the mRNA cap sites are several distinct regions of almost absolute identity. Using the numbering of the human c-myc P2 promoter, the first of these extends from -70 to -54 which is identical amongst the human, cat, mouse and rat c-myc genes. The next highly conserved sequence occurs in the purine rich sequence which extends from approximately -50 to -39. There is then some variability until the P2 TATA boxes (-31 to -23), which are identical in all four species, and specify a third conserved region. A sector
of increased diversity then extends until the mRNA cap sites. Fourteen of
the next fifteen nucleotides are identical in all of the species, sequences
which would be present in both the DNA and the 5' end of the P2-
dependent mRNA. Finally, another stretch of thirteen of fourteen
identical nucleotides extends downstream from position +23.

As discussed above, the upstream conserved sequences through the
TATA box have been genetically implicated in the control of basal levels
of transcription. Some of these sequences also appear to be required for
premature termination of transcription at sites at the 3' end of exon I
(Miller, et al, 1989). The sequences downstream of the mRNA cap site,
which are also well conserved, have not yet been shown to have a
regulatory role. Some of the data presented in this chapter have
implicated these sequences as potential mediators of the myc
autoregulatory response, a function which would be consistent with their
high degree if conservation. It will be interesting to learn the significance
of these conserved sequences for this control pathway.

3.2.6 Promoter usage

The ratio of stable transcripts initiating from the P1 or P2 promoters
was found to be approximately 10:1 in favor of P2 initiation, which is
similar to the promoter ratio seen for endogenous Rat-1 c-myc expression,
for all constructs which contained human c-myc sequences ending at
position +47 relative to P2. When genomic constructs extending through
approximately 5000 bp of c-myc downstream sequence were assayed, the
ratio was altered to approximately 2-3:1 in favor of P2 (compare Figures
III.10 and III.3). These results are similar to findings reported by Hay et al
(1987) who observed that the ratio of promoter usage shifted towards P2
with the loss of downstream sequences. This alteration in promoter usage
upon loss of the downstream c-myc sequences suggests that they function
to maintain a specific ratio of transcripts. In both of these studies, addition

167
of the SV40 enhancer and promoter to the genomic fragment more than 2 kb upstream of P1 stimulated transcription from both promoters by an equivalent amount (Figure III.30). However, Hay et al (1987) also reported P1 transcription to be dependent on sequences positioned between -293 and -101, which contrasts with the insensitivity of the P1-dependent transcripts to the loss of sequences upstream of -101 observed in the present study (Figure III.10).

A shift in promoter usage was observed in the present study when an SV40 enhancer was placed upstream of position -88 (relative to P1), in a construct which extended to +47 (+208 relative to P1), resulting in almost equal proportions of cap site usage (Figure III.30). The ability to change this ratio by the addition of enhancer sequences close to the promoters could be due either to a differential specificity of the enhancer for the closest promoter, or may indicate an insensitivity of the P2 promoter to enhancer function. The former interpretation seems more likely, as the enhancer stimulated transcription from both promoters when placed at a much greater distance.

One intriguing alternative possibility is that the SV40 enhancer affects premature termination of the nascent transcripts and not their initiation when placed close to the promoters. Support for this notion comes from several observations. It has been reported that the sequence requirements for premature termination of the human c-myc gene include sequences which extend to approximately position +520 relative to P1 (Bentley and Groudine, 1988). These sequences are absent in all of the constructs which display a strong P2 promoter bias both in the present study and in that of Hay et al (1987). It has furthermore been reported that the vast majority of P1 initiated transcripts do not terminate prematurely at this downstream site (Spencer, et al, 1990). However, a block to elongation of P1 initiated transcripts has been found to occur between the P1 and P2 promoter cap
sites, and the promoters appear to be utilised for transcription initiation with approximately equal efficiency in oocyte microinjection assays (S Roberts and D Bentley, personal communication). As opposed to the sequences required for premature termination at the end of exon I, the sequences required for termination of the P1-initiated transcripts are present in all of the constructs containing the SV40 enhancer sequences (Figure III.29 and (Hay, et al, 1987). Thus it is possible that the SV40 enhancer sequences increase the magnitude of the P1 initiated signal to approximately equal that of the P2 initiated signal by alleviating a transcriptional block upstream of the P2 cap site. Furthermore, the total increase in signal is somewhat less than that observed when the SV40 enhancer sequences are placed much farther upstream of the promoters, suggesting that the enhancer may be functioning differently in the two situations. An additional observation suggests that the effect of the proximally situated enhancer is not typical of enhancers in general. When the SV40 transcription termination sequences are placed between the enhancer and the c-myc promoter sequences at position -88, all effects of the enhancer are eliminated. This result is unexpected as enhancers in general, and the SV40 enhancer in particular, were originally described as capable of exerting the effects in position and orientation dependent fashion (Banerji, et al, 1981, Serfling, et al, 1985). Placing the termination sequences downstream of the enhancer should be topologically equivalent to placing the enhancer downstream of the promoter, which would not be expected to affect its function. It will be interesting to learn at what stage of transcription control the enhancer is operating, and whether this varies with distance or the specific c-myc promoter under study.

If the inclusion of enhancer sequences upstream of the P2 promoter alone is found to increase the levels of P2-dependent transcript, then it may provide a way to assess the contributions of the non-TATA sequences
to the mechanism of myc autoregulation while still maintaining a significant level of signal.

3.2.7 Position effects

The absolute magnitude of transcripts detected from the promoter fragments assayed in this study is significantly less than that seen emanating from the endogenous c-myc promoters. For instance, even when eight kb of the human c-myc locus was transfected, the signals detected were approximately fifty-fold weaker than the endogenous rat transcripts. When transfected clones were studied, a large variation in the amount of signal was found with some clones expressing almost endogenous levels of transcript, and others almost none (Figure III.21). In addition the magnitude of the signals did not correlate with the number of introduced indicator sequences. These findings suggest that the position within the genome at which the transfected DNA integrated, as well as the introduced sequences, are critical in determining the usage and responsiveness of the promoter. Position effects frequently have been observed for other introduced DNA fragments (Mantei, et al, 1979, Palmiter and Brinster, 1985, Wold, et al, 1979).

Sequences which can enable a transfected gene to be expressed at levels approximating that of the endogenous locus, and in a dose dependent fashion have been identified for some genes (Forrester, et al, 1986, Greaves, et al, 1989, Grosveld, et al, 1987, Stief, et al, 1989, Tuan, et al, 1985). Perhaps the best studied of these is the locus activating region (LAR) of the multigene cluster which contains the human beta globin locus (Townes and Behringer, 1990 for review). Sequences located 50-65 kb upstream of the gene cluster appear to organise it into an 'open' chromatin structure, and act as potent transcriptional enhancer sequences. These sequences are constitutively active and do not appear to be involved in controlling the temporal or spatial regulation of expression of genes within the locus.
These controls appear to be mediated by enhancer and promoter elements closer to each individual gene. Some proteins which interact with hypersensitive sites within the LAR sequences have been identified, but it is incompletely understood how these sequences function to control transcription. It seems probable that sequences located many kb from the c-myc locus may also be involved in the regulation of its expression, and that their identification will be required before either high levels of expression or completely appropriate regulatory actions can be consistently reproduced in an experimentally accessible context.

In the absence of such c-myc specific sequences, it may be that the use of an appropriate LAR domain from a different gene may result in an increase in the efficiency of c-myc promoter usage. Experiments have been performed assessing the utility of such sequences to increase human c-myc promoter dependent transcription. The LAR sequences used in these experiments originated from the chicken lysozyme gene (Stief, et al, 1989). These sequences have been reported to increase transcription by at least ten-fold, and in the presence of appropriate enhancer sequences, support transcription in a copy number-dependent, additive fashion. Inclusion of the human c-myc sequences from -88/+208, relative to P1, in a construct containing the lysozyme gene elements was not found to have any effect on the low levels of human c-myc dependent signal (data not shown). In order to attain LAR function on the human c-myc promoter it may be necessary to include an enhancer along with the LAR elements, or to use LAR sequences from a different gene.

3.2.8 Future prospects

The data presented in this chapter have demonstrated that sequences proximal to the human c-myc P2 promoter are sufficient for imparting myc protein responsiveness to a heterologous indicator. While it has been possible to map these elements to a relatively small region of the
promoter, the identification of a single, necessary sequence has not been possible. The identification of such a sequence must be one of the primary aims of any future work. The greatest impediment to such a result has been the low level of P2 promoter dependent transcription observed when using constructs which contain minimal human c-myc promoter sequences, further obscured by the relatively high levels of upstream initiated transcription.

Considered a posteriori, there are a number of alternative approaches which might lead to an increase in the levels of transcription while allowing mutation of a response element. One of these would be to use constructs which contain enhancers or other regulatory elements upstream of the P2-specific sequences, but in the absence of the P1 promoter, or alternatively, to investigate the response of the P1 promoter in the presence of upstream enhancer sequences. Additionally, it may be possible to confer myc responsiveness to a different promoter using some of the human c-myc promoter sequences in the absence of a myc TATA box. These experiments may be difficult as it was found that the myc promoter directed most of the transcription when an SV40 promoter was placed two kb upstream of the human c-myc promoter cap sites, to the exclusion of SV40 promoter usage (see Figure III.30).

Another difficulty impairing the generation and interpretation of data in this study has been the effects of prolonged growth of the parallel pools of retrovirally infected cells before harvesting. The relatively lengthy selection time required for the isolation of infected cells has necessitated the use of control indicators such as the α-globin gene to monitor drifting expression in the transfected pools. It has also led to a requirement for extended passaging of colonies in order to stabilise expression of the transfected indicators before infection with the viruses. The use of cell lines which express an inducible myc protein may help alleviate some of
these problems. Unfortunately, inducible promoter constructs regulating myc protein expression have not allowed the isolation of cell lines which display stable, regulable down-regulation of endogenous c-myc expression (MW Brooks, LJZ Penn and H Land, personal communication, and data not shown). More encouragingly, a human c-myc/estrogen receptor fusion protein has recently become available which allows myc protein function to be regulated by the application of hormones, while the protein is synthesised constitutively (Eilers, et al, 1989). When expressed in rodent fibroblasts this protein has been demonstrated to impart a conditionally transformed phenotype to the cells. Preliminary experiments have indicated that myc autosuppression in Rat-1 cells can also be made conditional with this protein (MW Brooks and H Land, personal communication). It will be interesting to investigate whether this system will allow appropriate regulation of the chimaeric human c-myc promoter/indicator constructs.
Figure III.1: Schematic diagrams of human c-myc (genomic), β-interferon and CAT indicator constructs

pGC1H/R: Genomic human c-myc HindIII/EcoRI (-2329/+5753; relative to P1) fragment inserted into the plasmid vector pGC1 (Myers, et al, 1985). Boxes indicate exons, darkened boxes the c-myc coding region, arrows the P1 and P2 promoters.

pGC1H/RBam8: pGC1H/R with frameshifting 8 bp BamHI linker inserted at the EcoRV site in exon II.

pBSKS+HBCAT: HindIII/NaeI(+208) fragment of human c-myc cloned upstream of the BamHI/Sall bacterial chloramphenicol acetyl transferase (CAT) gene fragment, inserted into the plasmid vector Bluescript KS+ (Stratagene).

pBSKS+HN277IFN: HindIII/NaeI fragment of human c-myc cloned upstream of the PvuII/BalI (+277/+1150) β-interferon gene fragment, inserted into the plasmid vector Bluescript KS+.

Some restriction endonuclease sites are marked on the inserts, significant sites are highlighted in bold face and those in parentheses were destroyed during cloning.

Inserts are to scale.

For further details of the constructions, see Materials and Methods.
Figure III.2: Schematic diagrams of retroviral vector VM and LJ proviruses

VM (Dotto, et al, 1985) expresses the p110^{v-gag-myc} protein from the Moloney murine leukemia virus LTR (U3.R.U5), and the G418 resistance marker (neo) from an internal SV40 early region promoter/enhancer.

LJ (Korman, et al, 1987) expresses the G418 resistance marker from the internal SV40 promoter/enhancer.

The position of the mutant splice donor sequence (SD\textsuperscript{-}), and plasmid origin of replication (pBRori) in each virus is indicated. The horizontal lines below the proviral maps indicate the transcripts which originate from the two promoters.

The vector sequences but not the insert p110^{v-gag-myc} insert are drawn to scale.

Typical retroviral vector titres were:
- VM: $1 \times 10^6$ colony forming units (cfu)/ml
- LJ: $2 \times 10^5$ cfu/ml
VM: v-gag myc

LJ: LTR SV40 neo pBR Ori LTR

RNA: ____________________________
Figure III.3: Response of transfected human c-myc genomic constructs to exogenous v-myc

10 μg RNA from pools of Rat-1 cells transfected with genomic human c-myc plasmids and subsequently infected with control (LJ) or v-myc (VM) viruses were subjected to RNAase protection analysis with the human c-myc 5' and γ-actin probes.

Pools transfected with pGC1H/R (Human c-myc, panel A) or pGC1H/RBam8 (Human c-myc frameshift, panel B) are indicated. Sizes of single strand DNA size markers are indicated at the left, in nt.

A map of the expected lengths of the protected c-myc probes is displayed at the bottom of the figure. The PvuII sites are at positions -353 and +509 relative to promoter P1. P1, P2 indicate the positions of human c-myc mRNA cap sites P1 and P2
A. Human c-myc

- c-myc P1
  - 622
  - 527
- c-myc P2
  - 404

B. Frameshift human c-myc

- c-myc P1
  - 862 bp ss RNA probe
  - Protected fragment
- c-myc P2
  - 509 bp
  - 348 bp

\[ \text{Human c-myc Exon I} \]

\[ \text{γ-actin} \]
Figure III.4: Response of transfected human c-myc promoter fusion constructs and endogenous rat c-myc to exogenous v-myc

10 μg RNA from pools of Rat-1 cells transfected with human c-myc promoter fusion constructs and subsequently infected with v-myc (VM, v-myc) or control (LJ, control) retroviruses was analysed by RNAase protection. Each set of three lanes contains samples from duplicate VM infected pools and one control pool.

The transfected indicator constructs pBSKS+H(B)CAT (MycCAT) or pBSKS+HN277IFN (MycIFN) are indicated. Bands detected with γ-actin (γ-actin) CAT (CAT), human β-IFN3' (IFN) or rat c-myc exonI (Rat c-myc P2) probes are indicated by arrows. Numbers to the left of the figure indicate the position of single strand DNA size markers, in nt.

The BSKS+H(B)CAT transfected, control sample which was probed for CAT signal was also simultaneously probed for rat c-myc message, and the smear in this sample is due to degradation of this probe. However, the two corresponding VM infected samples were not probed for endogenous c-myc gene expression. In order to avoid confusion, the top of the figure was cropped in these lanes. The weak bands which are seen in all of the samples in this figure are due to undigested γ-actin probe as determined in a tRNA control lane (not shown).
Figure III.5: Mapping 5' ends of MycCAT and MycIFN fusion transcripts

10 μg RNA from pools of Rat-1 cells transfected with human c-myc promoter fusion constructs pBSKS+H(B)CAT (C) or pBSKS+HN277IFN (β) and subsequently infected with v-myc (VM, +) or control (LJ,-) retroviruses was analysed by RNAase protection using probes which map the 5' ends of the human c-myc derived sequences.

Samples in the left hand panel were probed either with the human c-myc 5' and γ-actin probes or the human c-myc 5' end probe only (two rightmost lanes). Samples in the right hand panel were probed with the MycIFN fusion A and γ-actin probes.

Arrows indicate bands which correlate with putative human c-myc P1 (P1) or P2 (P2) promoter initiated transcripts, as well as those of unknown (°) or full length probe overlap (RT) origin. Rat γ-actin transcripts are also indicated. Numbers to right of figure mark position of single-strand DNA markers, in nt.
Figure III.6: Schematic diagrams of plasmids pSV\textsubscript{2}CAT and pUCSVterLTRIFN

The plasmid pSV\textsubscript{2}CAT (Gorman, et al, 1982) contains the SV40 early promoter/enhancer controlling expression of the bacterial chloramphenicol acetyl transferase gene cloned into the plasmid pBR322.

pUCSVterLTRIFN contains duplicate SV40 transcription termination sequences (SVter) immediately upstream of the Moloney murine leukemia virus (MoMLV) LTR, which controls the transcription of the human $\beta$-interferon indicator, cloned into pUC13.

Relevant portions of the constructs are indicated, restriction sites in parentheses were destroyed in the cloning.

The figures are not to scale. See Materials and Methods chapter for further details of the pUCSVterLTRIFN construct.
SV40 early promoter/enhancer

CAT

pSV2CAT

MoMLV LTR

U3

R

U5

[SalI] (Pmol)

[BalI] (Pmol)

[pUC SVter LTR IFN]

pUC13
Figure III.7: Response of transfected pSV$_2$CAT and pUCSVterLTRIFN and endogenous rat c-myc to exogenous v-myc

RNAase protection analysis of 10 µg RNA from cell pools stably transfected with pSV$_2$CAT or pUCSVterLTRIFN and then infected with either v-myc-carrying or control retrovirus.

(left panel) pSV$_2$CAT samples were protected with three probes simultaneously: Rat c-myc exonI (P2), CAT (CAT) and γ-actin (γ). (right panel) pUCSVterLTRIFN samples were protected with two probes simultaneously: LTRIFN (IFN) and Rat GAPDH (GAPDH).

Bands corresponding to RNA protected by each of these probes are indicated by the labelled arrows. Numbers to the right of the panels indicate the size of end labelled single-strand DNA markers, in nt.

For complete descriptions of the protection probes, refer to the Materials and Methods chapter.
Figure III.8: Response of endogenous Rat c-myc and a transfected human c-myc/CAT fusion construct to the presence of v-myc under various serum and cell growth conditions

10 µg of RNA were analysed by RNAase protection analysis. The growth conditions of the cells from which the RNA was prepared is indicated at the top of the figure. Rat-1 cell pools which had been stably transfected with the plasmid pBSKS+H(B)CAT and subsequently infected with either a retrovirus carrying both v-myc and G418 resistance genes (+) or only the G418 resistance gene (-) were harvested following a variety of treatments.

The following growth conditions were used: Either cells were harvested at sub-confluence 24 hours following seeding (SS, 24 hours post serum); or media on subconfluent cultures was changed to 0.5% serum 24 hours after seeding, and the cells maintained in 0.5% serum for 3 days (SC), following which they were either harvested or treated with 10% fresh serum for 2 hrs (0 or 2 hours post serum); or cells were grown to confluence, then changed to 0.5% serum for 3 days (C), and harvested or treated with 10% fresh serum for 2 hours, then harvested (0 or 2 hours post serum).

Samples were protected with three probes simultaneously: Rat c-myc exonI (Rat c-myc P2), CAT (CAT) and γ-actin (γ-actin). Numbers to the right of the figure indicate the size of single-strand DNA markers, in nt.
Figure III.9: Schematic diagrams of human c-myc 5' deletion series, BSKS+(-353:+119/+208)IFN

Human c-myc promoter/β-interferon constructs of the 5' deletion series pBSKS+-353 through +119/+208 IFN.

Endpoints of the constructs are illustrated by the length of the thickened line. Constructs with 5' ends at -353, -348, -276, -181, -169, -101, -91, -88, -86, -58 and +119 are shown, human c-myc sequences are to scale.

See Materials and Methods for details of the constructions.
Human c-myc

-353:+119 deletions

P1

-353
-348
-276
-181
-169
-101
-91
-88
-86
-58
+119

P2

+208

BamHI
NaeI/PvuII
β-Interferon
HindIII

pBSKS+ Δ5' myc IFN

Bluescript KS+
Figure III.10: Response of pBSKS+(-353:-91/+208)IFN constructs to v-myc

10 μg RNA from stable pools of Rat-1 colonies cotransfected with a pBSKS+(Δmyc/+208)IFN deletion construct, an α-globin control and pJ6Ωhygro then superinfected with either VM (+) or LJ (-) retroviral vectors was analysed by RNAase protection analysis. Samples were hybridised to the Myc/β-IFN fusion A and human γ actin probes. Bands corresponding to transcripts initiating from human c-myc P1 and P2 promoters are indicated (P1 and P2). ◦ symbols indicate bands corresponding to transcripts protecting the full length of homology between the Myc/β-IFN fusion A probe and the 5' end of the human c-myc sequences. • symbol indicates a band of unknown provenance which is frequently detected.

Numbers at the top of the lanes indicate the 5' end of each deletion relative to P1. Independent duplicate transfection pools are shown. Numbers to the right of the figure indicate the sizes of single strand DNA marker fragments in nt.
Figure III.11: Response of pBSKS+(-91:+119/+208)IFN constructs to v-myc

10 µg RNA from stable pools of Rat-1 colonies cotransfected with a pBSKS+ΔmycIFN deletion construct, an α-globin control and pJ6Ωhygro then superinfected with either VM (+) or LJ (-) retroviral vectors was analysed by RNAase protection analysis. Samples were hybridised with to the Myc/β-IFN fusion A and human γ actin probes (see Materials and Methods). Bands corresponding to transcripts initiating from human c-myc P1 and P2 promoters are indicated (P1 and P2). ° symbols indicate bands corresponding to transcripts protecting the full length of homology between the Myc/β-IFN fusion A probe and the 5' end of the human c-myc sequences. • symbol indicates a band of unknown provenance which is frequently detected.

The +430 deletion maps to position +430 of the β-interferon sequences, 153 bp from the 5' end of the β-interferon sequences in the other indicator constructs.

Numbers at the top of the lanes indicate the 5' end of each deletion relative to P1. Independent duplicate transfection pools are shown. Numbers to the left of the figure indicators the sizes of single strand DNA marker fragments in nt.
<table>
<thead>
<tr>
<th>5' endpoint:</th>
<th>-91</th>
<th>-88</th>
<th>-86</th>
<th>-58</th>
<th>+119</th>
<th>+430 (IFN)</th>
</tr>
</thead>
<tbody>
<tr>
<td>v-\text{myc}:</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

![Image of gel electrophoresis result with bands at positions 309, 217, 201, 190, 180, 160, 147, 123, 110, 90, and 404.](image)
Figure III.12: Response of pBSKS+(-91:+119/+208)IFN constructs to v-myc (reduced exposure of Fig. III.11)

This figure contains a shorter exposure of the autoradiogram used in Figure III.11 to emphasise the P1-dependent signals. See legend of Figure III.11 for details.
5' endpoint:

<table>
<thead>
<tr>
<th>v-myc:</th>
<th>-91</th>
<th>-88</th>
<th>-86</th>
<th>-58</th>
<th>+119</th>
<th>+430 (IFN)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-+</td>
<td>+</td>
<td>+</td>
<td>-+</td>
<td>+</td>
<td>-+</td>
</tr>
</tbody>
</table>

**Legend**

- **P1**: Markers at positions 404 and 309.
- **P2**: Markers at positions 242, 238, 217, 201, 190, 180, 160, 147, 123, 110, and 90.

**Gene Expression**

- **γ-actin**
Figure III.13: Response of endogenous rat c-myc and α-globin indicator cotransfected with pBSKS+(-91:+119/+208)IFN deletion series to v-myc

10 μg RNA samples from the same colony pools analysed in Figures III.10-12 were probed for expression of their endogenous rat c-myc and cotransfected α-globin genes. 5' ends of the transfected pBSKS+(Δx/+208)IFN constructs are indicated.

Rat c-myc P1 and P2 initiated (rat c-myc exon I probe; P1 or P2), α-globin (α-globin probe; α) and rat γ-actin (human γ-actin probe; γ) mRNA levels were assessed.
Figure III.14: Schematic diagrams of human c-myc 5' deletion series, pUC(-88:+97/+208)IFN

Human c-myc promoter/β-interferon constructs of the 5' deletion series pUC-88 through +97/+208 IFN.

Endpoints of the constructs are illustrated by the length of the thickened line. Constructs with 5' ends at -88, -64, -13, +23, +43, +69 and +97 are shown; human c-myc sequences are to scale.

See Materials and Methods for details of the constructions.
Figure III.15: Response of pUC(-88:+97/+208)IFN constructs to v-myc
RNAase protection analysis of 10 μg RNA of colony pools (*)
cotransfected with pUC(Δx/+208)IFN deletion constructs (5' end
indicated), α-globin and pJ6Ωhygro and subsequently infected with v-
myc (VM, +) or control (LJ, -) retroviral vectors.

The MycIFN fusion A probe was used for analysis of transfected
human c-myc P1 (P1) and P2 (P2) dependent transcripts; α-globin probe
for α-globin (α) transcripts and γ-actin probe for γ-actin (γ) transcripts.
Bands which correspond to the 5' end of the transfected human c-myc
construct (read-through transcripts) are indicated (°).

Numbers to the left of the figure indicate the size of single strand
DNA markers in nt.
Figure III.16: Response of upstream initiated transcripts to v-myc and serum

10 µg RNA from Rat-1 colony pools transfected with the constructs pBSKS+(-101/+208)IFN or pBSKS+(+430)IFN, were probed for upstream initiated myc promoter-independent transcript levels (readthrough; BSKS+SspI probe) and γ-actin (γ; γ-actin probe).

Left panel: Colony pools superinfected with the VM (+) or LJ (-) retroviral vectors.

Right panel: Colony pools made confluent and serum starved (-) or stimulated for 2 hours (+).

Numbers to the periphery of the panels indicate the positions of single strand DNA markers, in nt.
Figure III.17: Schematic diagrams of pUC and pGC1 myc:IFN fusion constructs

- pUC-88/+208IFN: BamHI/HindIII human c-myc β-interferon fragment inserted into the pUC13 vector polylinker.
- pUC+119/+208IFN: BamHI/HindIII human c-myc β-interferon fragment inserted into the pUC13 vector polylinker.

All positions are relative to the human c-myc P1 promoter mRNA cap site. Positions of human c-myc P1 and P2 promoters are indicated.

Arrows indicate relative positions and orientations of ampicillin resistance gene (amp), plasmid origins of replication (Ori), F1 M13 phage origin of replication (M13 ori), and lac operon lac I (lacI) and lac Z (lacZ) genes.

For further details of the constructions, see Materials and Methods.
Figure III.18: Response of pUC and pGC1 myc:IFN fusion constructs to v-myc

10 μg RNA from colony pools transfected with pGC1 (pGC1) or pUC (pUC) (-88/+208) (-88) or (+119/+208) (+119) IFN and superinfected with control (LJ, -) or v-myc (VM, +) retroviral vectors.

Human c-myc P1 (P1) or P2 (P2) dependent transcripts were detected with the MycIFN fusion A probe. γ-actin transcripts (γ) were detected with the human γ-actin probe.

°: Bands corresponding to the 5' end of the human c-myc sequences.

Numbers to the left of the figure indicate the positions of single strand DNA size markers, in nt.
Figure III.19: Response of linear -88/+208IFN fragments to v-myc

RNAase protection analysis of transcripts from Rat-1 colony pools cotransfected with gel-purified linear -88/+208IFN fragment (excised by BamHI:HinDIII digestion of pUC-88/+208IFN), α-globin and pJ6Ωhygro, and superinfected with control (−; LJ) or v-myc (+; VM) retroviral vectors.

Human c-myc P1 (P1) or P2 (P2) dependent transcripts were detected with the MycIFN fusion A probe; α-globin transcripts (α) with the α-globin probe; and γ-actin transcripts (γ) with the γ-actin probe.

◦: Bands corresponding to the 5' end of the human c-myc sequences.

•: Bands of unknown provenance

Numbers to the left of the figure indicate positions of single strand DNA markers, in nt.
Figure III.20: Suppression of upstream non-specific transcript levels by addition of SV40 transcription termination signals

RNAase protection analysis of transcripts from Rat-1 colony pools (*) cotransfected with pUC-88/+208IFN (-88) or pUCSVter-88/+208IFN (SVpA-88) and pJ6Ωhygro, and superinfected with control (-; LJ) or v-myc (+; VM) retroviral vectors. The pUC-88/+208IFN transfection also included the α-globin plasmid.

Human c-myc P1 (P1) or P2 (P2) dependent transcripts were detected with the MycIFN fusion A probe; α-globin transcripts (α) with the α-globin probe; and γ-actin transcripts (γ) with the γ-actin probe.

Readthrough: Bands which correspond to the non-human c-myc promoter-dependent transcripts initiating upstream of the homology between the 5' ends of the human c-myc indicator constructs and the MycIFN fusion A probe.

*: Band of unknown provenance
Readthrough

-88

virus:

* - +

SVpA

-88

* - +

P1

P2

α

γ

404
309
242/238
217
201
190
180
160
147
123
110
90
Figure III.21: Response of clonal pUCSVter-88/+208IFN transfectants to v-myc

RNAase protection analysis of transcripts from Rat-1 colonies (*) cotransfected with pUCSVter-88/+208IFN and pJ6Ωhygro, and colony pools of these clones superinfected with control (-; LJ) or v-myc (+; VM) retroviral vectors.

Human c-myc P1 (P1) or P2 (P2) dependent transcripts were detected with the MycIFN fusion A probe; γ-actin transcripts (γ) with the γ-actin probe; and rat GAPDH transcripts (G) with the GAPDH probe.

°: Bands corresponding to the 5' end of the human c-myc sequences.
•: Band of unknown provenance in clone 2 which may correspond to human c-myc promoter specific cryptic initiation.

Numbers to the left of the figure indicate the positions of single strand DNA markers, in nt.
Figure III.22: Analysis of endogenous rat myc and p110\textsuperscript{v-gag-}myc mRNA levels in clonal pUCSVter-88/+208IFN transfectants

Transcripts of 10 μg RNA from the colonies and colony pools described in Figure III.21 probed by RNAase protection analysis for endogenous rat c-myc (P2; rat myc probe), p110\textsuperscript{v-gag-}myc (gag; v-myc probe) and γ-actin (γ; γ-actin probe) expression. The p110\textsuperscript{v-gag-}myc expression derives from the integrated VM provirus.

Numbers to the left of the figure indicate the positions of single strand DNA markers, in nt.
Figure III.23: Reduction of transfected myc mRNA levels over time

A pool of Rat-1 colonies transfected with pUCSVter(-88/+208)IFN was passaged at 3-4 day intervals, and RNA prepared 24 hours after seeding each passage. Pools were infected with control (LJ) or v-myc (VM) retroviral vectors at passage 6, and RNA prepared for analysis from the subsequent polyclonal populations.

Human c-myc P1 (P1) or P2 (P2) dependent transcripts were detected with the MycIFN fusion A probe; rat GAPDH (G) transcripts with the GAPDH probe.

°: Bands corresponding to the 5' end of the human c-myc sequences.

•: Probe specific background bands.
Figure III.24: Schematic diagrams of human c-myc 5' deletion series, pUCSVter(-88:+150/+208)IFN

Human c-myc promoter/β-interferon constructs of the 5' deletion series pUCSVter-88 through +150/+208 IFN.

Endpoints of the constructs are illustrated by the length of the thickened line. Constructs with 5' ends at -88, +69, +97, +100, +112, +119 and +150 are shown, human c-myc sequences are to scale.

Sense strand sequences of the human c-myc component of the deletions +69, +97, +100, +112, +119 and +150 are presented at the bottom of the figure. The P2 TATA box and mRNA cap site (+1) are underlined.

See Materials and Methods for details of the constructions.
Human c-myc

-88:+150 deletions

SVter

pUCSVter Δ5' myc IFN

pUC13
Figure III.25: Response of 5' deletion series, pUCSVter(-88:+150)IFN to v-myc

Pools of colonies transfected with the pUCSVter(Δx/+208)IFN construct indicated were passaged for >7 passages then infected with either control (LJ;-) or v-myc (VM;+) retroviral vectors. RNA was prepared from pools of these colonies and analysed by RNAase protection for human c-myc promoter P1 (P1) and P2 (P2) dependent transcription (MycIFN fusion A probe) as well as rat GAPDH (G; GAPDH probe) transcription.

*: Bands corresponding to the 5' end of the human c-myc sequences.

•: Probe specific background bands.
Figure III.26: Schematic diagrams of human c-myc linker scan series, pUCSVter(ls1-Is5:+150)IFN and pUCSVter(ls1-Is5:Pst)IFN

Sense strand sequences of the linker scan series oligonucleotides ls1 through ls5 as inserted at the BamHI site of pUCSVter+150/+208IFN and BamHI linker Pst I site of the β-IFN gene of pUCSVterIFN.

Underlined sequences within the wild type (WT) human c-myc indicate the P2 TATAAA box and mRNA cap site (+1). Sequences in capitals indicate positions which differ from the WT myc sequences. Underlined sequences within the oligonucleotides indicate the nonamer linkers.

Endpoints of the wild type (ls1) oligonucleotide sequences or positions of the mutant linkers (ls2-ls5) with reference to both the human c-myc P1 and P2 promoters are indicated.

See Materials and Methods for details of the constructions.
WT HUMAN C-MyC

P2

+30

+1

+49

tcagggctgctggggagggggagggatcgggtctgagyATATAAAAgcgcgtttccggcctgttaattccagcagagggagcagcagcccctaggctgggaa...

+150 LS SERIES

1s1 (+94/+121 P1 or -71/-22 P2)

GATCTggcttgccgagggaatagccgagggatcgggtctgagtataaaaaagGATCCGGggcggctttttatatttcggtctgtgtcgatattccagcagagggagcagcagcccctaggctgggaa...

1s2

GATCTggcttgccgagggaatagccgagggatcgggtctgagtataaaaaagGATCCGGggcggctttttatatttcggtctgtgtcgatattccagcagagggagcagcagcccctaggctgggaa...

1s3

GATCTggcttgccgagggaatagccgagggatcgggtctgagtataaaaaagGATCCGGggcggctttttatatttcggtctgtgtcgatattccagcagagggagcagcagcccctaggctgggaa...

1s4

GATCTggcttgccgagggaatagccgagggatcgggtctgagtataaaaaagGATCCGGggcggctttttatatttcggtctgtgtcgatattccagcagagggagcagcagcccctaggctgggaa...

1s5

GATCTggcttgccgagggaatagccgagggatcgggtctgagtataaaaaagGATCCGGggcggctttttatatttcggtctgtgtcgatattccagcagagggagcagcagcccctaggctgggaa...

PST I LS SERIES

1s1

GATCTggcttgccgagggaatagccgagggatcgggtctgagtataaaaaagGATCCGGggcggctttttatatttcggtctgtgtcgatattccagcagagggagcagcagcccctaggctgggaa...

1s2

GATCTggcttgccgagggaatagccgagggatcgggtctgagtataaaaaagGATCCGGggcggctttttatatttcggtctgtgtcgatattccagcagagggagcagcagcccctaggctgggaa...

1s3

GATCTggcttgccgagggaatagccgagggatcgggtctgagtataaaaaagGATCCGGggcggctttttatatttcggtctgtgtcgatattccagcagagggagcagcagcccctaggctgggaa...

1s4

GATCTggcttgccgagggaatagccgagggatcgggtctgagtataaaaaagGATCCGGggcggctttttatatttcggtctgtgtcgatattccagcagagggagcagcagcccctaggctgggaa...

1s5

GATCTggcttgccgagggaatagccgagggatcgggtctgagtataaaaaagGATCCGGggcggctttttatatttcggtctgtgtcgatattccagcagagggagcagcagcccctaggctgggaa...

SVTer

β-Interferon

EcoRI BamHI/BglII BamHI HindIII

pUCSVTer ls series IFN

pUC13
Figure III.27: Response of linker scan series pUCSVter(1s1-1s5:+150)IFN to v-myc

Pools of colonies transfected with the pUCSVter(+150)IFN linker scan series constructs (or pUCSVter+97/+208IFN or pUCSVter+112/+208IFN) indicated were passaged for >7 passages then infected with either control (LJ;-) or v-myc (VM;+) retroviral vectors. RNA was prepared from pools of these colonies and analysed by RNAase protection for human c-myc promoter P2 (P2) dependent transcription (MycIFN fusion B probe) as well as rat γ-actin (γ; γ-actin probe) transcription.

°: Bands corresponding to the 5’ end of the human c-myc sequences or the BamHI site at the 3’ end of the inserted oligonucleotides.
Figure III.28: Response of linker scan series pUCSVter(ls1-1s5:Pst)IFN to v-myc

Pools of colonies transfected with the pUCSVter(+150)IFN linker scan series constructs indicated were passaged for >7 passages then infected with either control (LJ;−) or v-myc (VM;+) retroviral vectors. RNA was prepared from pools of these colonies and analysed by RNAase protection for human c-myc promoter P2 (P2) dependent transcription (MycIFN fusion B probe) as well as rat γ-actin (γ; γ-actin probe) transcription.

°: Bands corresponding to the 5' end of the human c-myc sequences or the BamHI site at the 3' end of the inserted oligonucleotides.
Figure III.29: Schematic diagrams of pUCSVter, pUCSVE, pUCSVESVter(-88/+208)IFN and pJ123hcmyc constructs

The relative positions of the SV40 terminator (SVter), SV40 enhancer (SVE) and human c-myc -88/+208 (Human c-myc) fragments upstream of the β-interferon indicator sequences in the pUC13 vector for the pUCSVter, pUCSVE, pUCSVESVter(-88/+208)IFN constructs is shown.

The J123hcmyc (Land, et al, 1986) construct contains SV40 early promoter and enhancer (SV40 E/P) sequences cloned upstream of the same HindIII to EcoRI 8.1 kb genomic human c-myc fragment used in the pGC1H/R construct (see Figure III.1).
Figure III.30: Response of pUCSVE, pUCSVESVter(-88/+208)IFN and pJ123hcmyc to v-myc

Pools of colonies transfected with the pUCSVESVter(-88/+208)IFN (SVter-88IFN), pUCSVE(-88/+208)IFN (SVESVter-88IFN) or pJ123hcmyc (SV123myc) constructs as indicated were infected with either control (LJ; -) or v-myc (VM; +) retroviral vectors. RNA was prepared from pools of these colonies and 10 μg analysed by RNAase protection for human c-myc P1 (P1) or P2 (P2) promoter dependent as well as rat γ-actin (γ; γ-actin probe) transcription.

RNA from an uninfected pool of pUCSVter(-88/+208)IFN (SVter-88IFN) transfectants is also shown.

The pUC0IFN samples were probed with the MycIFN fusion A probe, and the pJ123hcmyc samples with the human c-myc 5' probe.

°: Bands corresponding to the 5' end of the human c-myc sequences.

•: Probe specific background bands.

Numbers to the right of the figure indicate single strand DNA markers, in nt.
Table III.1: Summary of promoter proximal construct suppression results

The basal expression and response to elevated myc protein levels of constructs containing human c-myc sequences within approximately 70 bp of the P2 promoter cap site are compiled.

Levels of transcript (TXN) and responsiveness (REP) are approximately quantified as +, readily detectable; +/-, weakly detectable; +/?, apparently positive, but ambiguous; ?/- apparently negative, but ambiguous; and -, undetectable.

Sequence coordinates are presented relative to the P2 promoter mRNA cap site (relative to P1 in parentheses). Lower case letters indicate human c-myc sequences and upper case letters non-myc sequences. Single underlined uppercase letters are β-interferon sequences, double underlined letters positions mutated by the linkers in the linker scan series, and non-underlined sequences linkers added at the ends of fragments.
Table III.2: Sequence homologies amongst mammalian c-myc P2 promoters

Alignment of the human, mouse, rat and cat c-myc genes in the region surrounding the P2 promoter mRNA cap site. The sequences are aligned relative to the human c-myc gene, and are numbered with respect to the human P2 mRNA cap site (underlined +1). CON: Nucleotide positions which are absolutely conserved in all four genes.
-70 -60 -50 -40 -30 -20 -10  +1  +10  +20  +30  +40

HUMAN:  GAGGCTTGGCGGGAAAAAGA...ACGGA.CGGAGGGAC...CGCTGAGTATAAAGACGGTGGTCGCTATGAGTATAGGAGGGAGACGGAGGAG

MOUSE:  GACGCTTGGCGGGAAAAAGA...AGGGAGGGAGGATCTGAGTCG...AGTATAAAAAGACGGTGGTCGCTATGAGTATAGGAGGGAGACGGAGGAG

RAT:   GAGGCTTGGCGGGAAAAAGA...AGCGAGGGAGGATCCGGAGTGC...AGTATAAAAAGACGGTGGTCGCTATGAGTATAGGAGGGAGACGGAGGAG

CAT:    GGGGCTTGGCGGGAAAAAGAGAAAGAGGGAGGATCGTGTGCTAGCTGATAAAAGACGGTGGTCGCTATGAGTATAGGAGGGAGACGGAGGAG

CON:    G...GCTTGGCGGGAAAAAGA...A.GAGGAGGGATC...G...GTATAGAAGGTTGAGTGTAGTAGTAAT...CTAAGCGAGGAGCAGGCAGGGAGCGGAG
Chapter 4  Protein:DNA complexes proximal to the P2 promoter
4.1 Introduction

In the previous chapter it was demonstrated that it was possible to maintain a response to exogenous myc using sequences immediately proximal to the P2 promoter. However it was not possible to determine precisely which elements are responsible for this response, as no single mutation was capable of eliminating responsiveness while maintaining a detectable level of transcription. This suggests that either redundancies in the responsive sequences exist, or the response is mediated directly through the tata box, which was not altered in the transfection studies. As an alternative to the molecular genetic approach we decided to analyse the interactions between the promoter proximal sequences and cellular proteins more directly. The proteins which interact with this region of the promoter are the obvious first candidates through which myc might act. Myc protein may itself be present in one or all of these complexes, and thus interact directly with the transcriptional machinery. Alternatively, the effects of myc overexpression may be mediated indirectly, through modulation of the interactions of positive and negative factors at the P2 promoter.

The technique chosen to investigate these interactions was the electrophoretic gel retardation assay, in which cell extracts are combined with naked, radioactively labelled DNA fragments in vitro, and specific DNA:protein complexes identified by their ability to retard the migration of the DNA probe through a native polyacrylamide gel (Fried and Crothers, 1981). The advantages of the assay are its sensitivity, precision, ease and speed of use. Its utility has been demonstrated repeatedly, as witnessed by the numerous DNA binding proteins which have been identified and characterised over the past decade. Certain limitations must, however, be borne in mind. The procedure measures long term occupancy of a DNA site, and thus may not always detect transient or
unstable interactions. Furthermore the conditions utilised for in vitro
binding reactions do not necessarily reflect those encountered within the
normal cellular environment, which may result in the inability to detect
all naturally occurring or physiologically relevant interactions.

Typical gel-retardation binding conditions usually consist of a
hypotonic, tris-buffered, reducing solution optionally supplemented with
various metal salts such as MgCl₂ or ZnSO₄, the cations of which are
required cofactors for some DNA binding proteins. Heterologous DNA or
synthetic homo- or heteropolymers such as poly d(IC)•poly d(IC) or poly
d(A)•poly d(T) are generally added as carriers to compete non-specific
DNA binding proteins. Additional factors such as the non-ionic detergent
NP-40 or spermidine are often also used to increase binding specificity or
reduce non-specific background. Glycerol or ficoll are usually added to
increase solution density and thus facilitate gel loading. The choice of
each of these components can often greatly affect which complexes are
detected, with particularly dramatic effects being associated with the choice
of non-specific carrier. The initial characterisations presented here utilise
binding conditions (50 mM NaCl, 10 mM Tris pH 7.5, 1 mM EDTA, 1 mM
DTT, 5% glycerol) chosen because they detect members of the helix-loop-
helix family of proteins, with which myc protein might be expected to
share DNA binding properties based on their amino acid sequence

A general characterisation of the detectable protein:DNA interactions
between extracts of Rat-1 cells and various DNA fragments extending from
71 bp 5' to 49* bp 3' of the human c-myc P2 cap site (+94:+208 relative to
P1) was undertaken. This range of probes, which is broader than the

*Due to a clerical error, the 3' end of this sequence was mislabelled as +49 even though
it is actually +47. For consistency with the figures, it will be referred to as +49 from this
point.
minimal sequences required for a myc response, was used because these studies were initiated before the linker scan series results were produced. Due to the proximity of all of the probe sequences to the P2 cap site, all sequence references are given relative to P2 rather than P1. The major DNA:protein complexes were characterised through a combination of probe mutations and competition analyses. Correlation of these characteristics with those of previously known DNA binding activities were used to assign probable identities where possible. The relative abundance and mobility of each complex was also compared between extracts prepared from control or v-myc overexpressing Rat-1 cells. As a further test for the presence of myc protein, the effect of monoclonal anti-myc antibodies on the complexes was assayed.

4.2 Results

4.2.1 Protein extracts

Extracts were prepared by lysis of proliferating whole cells in 0.1% digitonin under slightly hypertonic salt conditions followed by diffusion of the nuclear proteins into the extraction buffer. This procedure has previously been optimised for its ability to extract myc protein from cell nuclei and was used for preparing all cell extracts tested in this study (GE Evan, personal communication). To ascertain whether the extracts contained myc proteins, an immunoblot using control and v-myc extracts was probed with polyclonal anti-myc antiserum (Figure IV.1). The antiserum used was raised against the most conserved region of the myc family of proteins, and has been shown to detect both rat c- and chicken v-myc proteins (Evan, et al., 1988, Penn, et al., 1990a). Bands of 62 and 110 kd were detected in the control and v-myc extracts, respectively. As expected, a strong reduction of the endogenous c-myc protein signal is observed in the v-myc infected cell extracts.
4.2.2 Complexes detected using salmon sperm DNA as carrier

Probes were tested for their ability to form complexes under a variety of conditions. The variable which most strongly distinguished between different complexes was the choice of non-specific DNA carrier used in the assay. When a -66/+49 fragment was used as probe, in conjunction with salmon sperm DNA as non-specific carrier, two major complexes (labelled 1 and 2, Figure IV.2) of similar mobility and intensity were detected. A series of weaker, slower migrating bands (marked by asterisks) were also often, but not always, seen. The same DNA fragment blocks all of these bands when used as an unlabelled competitor. Competition is incomplete at 20 fold molar excess, and almost total at 200 fold excess. Use of the -63/+49 probe results in the detection of one very weak complex (#), which has a greater mobility than that of complex 2, and is only marginally compatible (Figure IV.2 and data not shown). These data imply that some of the nucleotides between positions -66 and -63 are required for formation of complexes 1 and 2.

The linker scan series oligonucleotides, ls1 through ls5 (spanning -71:-22, see legend, Figure IV.2 for sequences), were used both as competitors and DNA probes to further define the sequence requirements of these complexes. These fragments were also used in the transfection analyses described in Chapter 3, and contain mutations through the 36 base pairs immediately 3' of position -66. When used as probes, the ls fragments bind two major complexes and three minor ones, with slightly increased mobilities relative to the bands seen with the larger -66/+49 probe. -66/+49 and ls1 through ls3 all bind with a similar affinity, while ls4 detects all of the same complexes but with reduced intensity, and ls5 is not retarded under these conditions (Figure IV.2). In competitions against the -66/+49 probe, the ls1 through ls4 oligonucleotides all compete both complexes 1 and 2 very efficiently, blocking almost completely at 20 fold
molar excess (Figure IV.3). Ls5, however, is quite impaired in its ability to compete, displaying only a marginal effect on either complex 1 or 2 even at 200 fold excess. The pattern of band mobilities as well as their specificities is consistent with the linker scan probes binding the same complexes as the larger DNA fragment. These data also confirm that all of these complexes are sequence specific, as they can be bound and competed by only a subset of the fragments used.

The combination of 5' end deletions, internal linker scan probes and competitions allows the partial definition of necessary sequences for the formation of these complexes. The 5' sequence minima falls between positions -66 and -63. The ls5 fragment, which is mutated at seven of the eight nucleotides between positions -65 and -58, does not compete or bind any of the complexes. Ls4, however, which contains mutations scattered throughout the next nine nucleotides, is fully capable of competing these complexes and binds them too, albeit with slightly reduced affinity. Thus the binding site of all the detected complexes extends through at least position -57 and may extend to position -49, although the extent of the sequence requirements into the ls4 mutated region is unclear, but is probably not extensive. This analysis allows the current assignment of the binding site as follows: 5' (-66) **GCGGGAAAAAGAACGGA** (-49) 3', where doubly underlined sequences are those defined by the 5' deletions and ls5, and singly those identified by ls4. It is interesting to note that in the -63/+49 construct the juxtaposition of BamHI linker sequences to position -63 results in just two point mutations in the GC rich region from GCGGGG to GCGGG which completely abolish binding.

4.2.3 Complexes 1 and 2 are probably E2F

During the course of this work, it was reported that the human c-myc promoter contains a binding site for the transcription factor E2F, which maps to the sequence GGCGGGAAAA, from position -66 to -57 5' of P2.
The binding site of E2F protein was originally defined through gel retardation and footprinting analyses of the adenovirus 5 E2a early gene promoter, and consists of an inverted repeat of the sequence 5' GCGCGAAA 3' (Kovesdi, et al, 1986, SivaRaman and Thimmappaya, 1987). To determine whether the complexes seen in the Rat-1 extracts have a similar binding specificity to E2F, oligonucleotides containing a single, a double or a point mutated double E2F binding site from the Adenovirus 5 E2a promoter were used as competitors against the -66/+49 human c-myc promoter probe (Figure IV.4). Both the single and double sites compete efficiently for binding of complexes 1 and 2, while the point mutated oligonucleotide does not. Complexes 1 and 2, and the minor complexes of slower mobility, are thus indistinguishable from E2F based on their binding and competition properties (Thalmeier, et al, 1989). Therefore these complexes will in future be referred to as E2F.

4.2.4 Response of E2F to overexpression of v-myc

The identification of the E2F DNA:protein complexes was performed using extracts made from Rat-1 cells infected with the v-myc virus, VM. It was of interest to determine if the levels or mobility of these complexes were altered in these extracts relative to those from Rat-1 cells infected with the control virus, LJ. The -66/+49 fragment and each of the Is oligonucleotides were used as probes in self competition experiments with both LJ and VM infected cell extracts (Figure IV.5 and IV.6). The mobility of the complexes was identical between the two extracts, as was self-competed. The absolute level of binding activity was, however, moderately increased in the VM infected cell extracts at equal protein concentrations. This increase was seen with all of the probes used. When normalised against a complex which shows little apparent variation in a parallel bandshift (complex B, see below), the increase was quantitated to

244
~1.5-2-fold, though the significance of such a variation is not immediately obvious.

4.2.5 Complexes detected using poly d(IC)·poly d(IC) as carrier

Other complexes in addition to E2F have previously been reported to bind sequences proximal to the P2 promoter of the murine c-myc gene. Some of these map to sites closer to the mRNA cap site than E2F. For example, a factor called Me1a1 has been mapped to a site centered around position -50 of the murine gene. It should be comparable by ls5, but not ls3, and thus be distinguishable from E2F (Asselin, et al, 1989). None of the complexes described above has these characteristics. Scrutiny of the methods used in reports of the E2F complex versus the Me1a1 complex revealed that E2F is detected when salmon sperm DNA or the heteropolymer poly dA·poly dT is used as non-specific competitor, while reports of the Me1a1 factor use the synthetic heteropolymers poly dI·poly dC or poly d(IC)·poly d(IC) (Asselin, et al, 1989, Jones, et al, 1988, Thalmeier, et al, 1989). This prompted us to reexamine the complexes bound by the same probes when using poly d(IC)·poly d(IC) as non-specific competitor.

Several DNA fragments as well as the ls1 through ls5 probes were utilised in gel retardation assays that contained poly d(IC)·poly d(IC) as non-specific carrier (Figure IV.6). The fragments consist of members of the 5' deletion series with endpoints at positions -66, -63, -51 and -44, which all terminate at position +49, and a 3' deletion fragment, -66/+22.

Multiple complexes are detected with probes of the deletion series. Two major complexes, labelled A and B, of which A sometimes appears as a doublet, are detectable with the -66/+49 probe. A series of higher mobility complexes are also seen with this probe, though they are less intense than A or B. The 5' deletions through position -51, as well as the -66/+22 probe detect a complex with a mobility very similar to B. Only
-66/+49 and -63/+49 detect complex A. These complexes are both specific, as they can be significantly self competed with unlabelled DNA (Figure IV.6 and data not shown).

The linker scan series of probes also detect several complexes, with the major one, labelled β, migrating just slower than the B complex. β binds selectively to the ls1, ls2 and ls5 probes, but not to ls3 or ls4, implying that it requires residues which are mutated in either of these probes, which span the region from positions -57 through -40, or 5' AAGAACCGAGGGAGGGAT 3'. The mobility of β does not correspond to that which might be expected of either the A or B complex using a shortened probe: it migrates too quickly for A and too slowly for B (Compare with the positions of the E2F complexes when using the -66/+49 or ls1 probes (Figure IV.3), which change in proportion to the difference in probe lengths).

The linker scan series was also used in competition experiments against the -66/+49 and -63/+49 probes. The results are essentially identical, so only those for the -66/+49 probe are shown (Figure IV.8). None of the linker scan oligonucleotides will compete complex A, while ls1, ls2 and ls5 all compete complex B, and do so slightly more efficiently than the probe fragment in self competitions. Ls3 and ls4 are both inefficient competitors, though some reduction in signal at 200 fold molar excess is visible. It may be noteworthy that while complex B can be competed by some sequences, none is as efficient as in competitions against the E2F complexes (compare with Figure IV.5). From the data of Figure IV.7, the 5' end of the B complex maps between positions -51 and -44, and 3' end maps to within +22, though no 3' boundary has been unambiguously determined. The competition data with the linker scan mutants suggest that the binding site lies between positions -57 and -40, inclusively, the same region which binds complex β. The intersection of
the deletion and competition data suggest a binding site for the B complex from -51 to -40, or 5' GGAGGGAGGGAT 3'. While the B and β binding sites appear to overlap, it is not yet clear whether they have the same binding specificity (see below).

Since the -66/+49 and -63/+49 probes both detect complexes A and B, these complexes must be distinct from the E2F complex, which is only seen using the -66/+49 probe. This deduction is supported by Figure IV.9, in which the E2F oligonucleotides are used as competitors for the bands detected with the -66/+49 probe. The oligonucleotides do not compete efficiently for A or B binding, nor for any of the minor complexes.

The properties of complex A are quite unique and relatively complicated. They are explored more fully in Chapter 5. The characterisation of the B/β complexes is described below.

4.2.6 Comparison with Me1a1

The Me1a1 complex as defined by Marcu's group maps to the sequence 5' AAAAAGAAGGAGGGGAGGGAT 3' as assayed by DNAse I footprinting on the murine c-myc promoter, which corresponds to positions -59/-44 of the human c-myc promoter, which consists of the sequence 5' AAAAAGAAGGAGGGGAGGGAT 3' (the differences in these sequences are underlined). Methylation interference experiments identify all of the G nucleotides within the murine G3AG4AG3A sequence as essential for binding, although a minimal binding element has not been demonstrated (Asselin, et al, 1989). More recently, Hall has demonstrated that the sequence 5' GGGAGGGGAGGGA 3', which he also terms Me1a1, will bind a specific protein complex in gel retardation assays using whole cell extracts of HeLa cells (Hall, 1990). He additionally identifies two sets of point mutations which specifically block the formation of this complex. While both of these papers describe a complex called Me1a1, it has not been determined if they are identical.
The sequences to which the B and β complexes map are consistent with one or both of these complexes being the same as either of these Me1a1 complexes. To determine if B is the same as the Hall Me1a1 complex, synthetic oligonucleotides containing the wild type (wt) or a mutant Me1a1 binding site as defined by Hall (Hall, 1990) were used as competitors against both the -66/+49 and wt Me1a1 probes (Figure IV.10). Surprisingly, none of the oligonucleotides was found to compete any of the complexes detected with the -66/+49 probe. The Me1a1 probe did detect a complex, b, which was competable by the wt sequence in a self competition assay, as well as by both of the mutant oligonucleotides (Figure IV.10). The mutant oligonucleotides will also detect a complex of the same mobility, which is similarly competable. Varying the non-specific carrier conditions also failed to detect a complex with the characteristics described by Hall, although he finds no difference with either poly(dIdC) (sic) or salmon sperm DNA carriers (data not shown). Thus the B complex clearly does not have the same DNA binding specificity as the Hall Me1a1 complex, nor do these extracts appear to contain any activity which is similar to that which he describes. No additional experiments were performed to characterise the b complex, which is also distinct from the Hall Me1a1 complex.

While the B and β complexes (see Figure IV.7) share some DNA binding and competition attributes, it is unclear whether they contain the same protein components, or if either is the same as the Marcu Me1a1 complex. The aberrant migration of B and β relative to their probe lengths of 58 bp (β) and 130 bp (B) suggests that these may be different complexes. The B complex is probably distinct from Me1a1 (Marcu) as it binds a probe (-51/+208) which does not contain the 5' half of the Marcu binding site (AAAAAGAA), and is not competed by the other half of this sequence (the Hall Me1a1 oligonucleotide). It is possible that β corresponds to the
Marcu Me1a1 complex, and is actually a distinct complex from B. Additional experimentation is required to resolve these paradoxes.

4.2.7 Comparison of control and v-myc Rat-1 extracts

The ability of Rat-1 LJ and VM extracts to support the formation of the A, B or $\beta$ complexes was investigated. The -66/+49 and Is series of probes were used in self competition assays with extracts made from both Rat-1 cell pools (Figures IV.11 and IV.12). One major difference between the two sets of extracts is apparent. The intensity of the A complex signal is dramatically reduced in the v-myc extracts, though its mobility is unaltered with respect to the control. The complex is well competed by unlabelled probe in both extracts. The B and $\beta$ complexes, as well as all of the minor complexes detected with the various probes, are detectable at similar levels and compete identically in both extracts. In fact, the invariance of the B complex has proved to be a very useful internal control when comparing extracts or other bands (eg the E2F complexes described above). As mentioned previously, these changes are investigated more completely in Chapter 5.

4.3 Conclusions

In Chapter 3 sequences proximal to the human c-myc P2 promoter were identified as potential regulatory elements in the response to levels of myc protein. In this chapter protein complexes which are capable of forming on these DNA sequences in vitro were assessed. It was found that several distinct sequence-specific DNA binding complexes could be detected. The levels of two of these appear to be altered in extracts from cells expressing elevated amounts of myc protein. One of these, complex A, appears to be novel, while the other is similar to a previously described complex, E2F (Kovesdi, et al, 1986, Thalmeier, et al, 1989).

Complex A has been found to bind sequences spanning the human c-myc P2 promoter from positions -63 to +47. Binding is reduced
significantly in extracts prepared from v-myc infected Rat-1 cells. The properties of this complex are explored more fully in Chapter 5.

4.3.1 E2F

The E2F complex binds the sequence 5' GGCGGGAAAA 3' located from positions -66 to -57 upstream of the P2 promoter. This complex bound DNA only when salmon sperm DNA was used as carrier. Levels of binding activity were increased up to two-fold in the v-myc containing Rat-1 cell extracts. The E2F binding site may be involved in the response of the human c-myc P2 promoter to elevated levels of myc protein, as well as in the maintenance of basal levels of transcription (See Chapter 3; Table III.1).

The E2F factor has been implicated in the regulation of several genes, including c-myc, in the response to both mitogenic stimulation of growth and during adenovirus infection (Hiebert, et al, 1989, Kovesdi, et al, 1986, Mudryj, et al, 1990, Thalmeier, et al, 1989). Levels of E2F binding activity have been observed to increase in response to these stimuli (Kovesdi, et al, 1986, Mudryj, et al, 1990). In the case of the adenoviral E2 gene promoter, the namesake of this binding activity, the factor binds to an inverted repeat sequence in a cooperative manner (Hardy, et al, 1989, Kovesdi, et al, 1986), dependent upon the presence of both the adenoviral E1a and E4 19 kd proteins (Hardy, et al, 1989, Huang and Hearing, 1989). The increased levels of transcription resulting from this interaction are dependent on the cooperative binding to at least dimerised binding sites (Hardy, et al, 1989). It has been found that the E1a protein facilitates the dissociation of the E2F factor from preexisting cellular multiprotein complexes. The E4 protein then interacts with the E2F factor, inducing increased stability of the complex on the E2 promoter DNA (Bagchi, et al, 1990). In situations where binding to a single element is observed, the 54 kd E2F factor appears
to bind as a monomer, albeit in complexes with other cellular proteins, yielding an overall less stable protein:DNA complex (Bagchi, et al, 1990).

In the human c-myc promoter there exists only one high affinity binding site for E2F (Thalmeier, et al, 1989). It has been suggested that sequences centered around position -40, 3' CTAGCGCG 5', constitute an inverted repeat of the principal binding site, (-66 to -58; 5' GCGGAAA 3') as they will weakly compete binding at high degrees of excess DNA fragment (Hiebert, et al, 1989, Mudryj, et al, 1990). The 3' CTAGCGCG 5' sequences have not, however, been demonstrated to bind the E2F factor in either gel retardation or DNA footprinting assays (Hiebert, et al, 1989). The data presented here do not show any apparent binding or competition for E2F binding by the sequences from positions -48 to -31, which are mutated by either linker scan fragments ls2 or ls3, and which span the potential second binding site. Furthermore, competition with a single E2 promoter E2F binding site is as efficient as that with a dimeric binding site. Thus the E2F factor binding to the human c-myc promoter as detected in Rat-1 extracts appears to be to only a single binding site. These results suggest that if the -40 sequences mediate responses in which E2F has also been implicated, then they do so through interactions with factors other than E2F, such as complexes β or B.

The E2F complexes in the Rat-1 extracts consist of two major species and several minor, lower mobility ones, although the binding occurs to only one sequence. This suggests that either there are multiple forms of E2F protein, or it is interacting with different partners, resulting in heterodimers of different mobility.

4.3.2 B and β

Complexes B and β were found to bind the c-myc promoter when poly d(IC)•poly d(IC) was used as non-specific carrier, but not in the presence of salmon sperm DNA. The relationship between these two complexes is
not fully understood at present, particularly with regard to whether they are the same or different complexes (see Table IV.1 for a summary of the binding and competition data). The major piece of evidence suggesting that they are distinct complexes is that their mobilities do not vary in direct proportion to the length of the DNA fragment on which each has formed. Thus when the DNA fragment which binds complex B is shortened, and binds β, the mobility of this complex is reduced relative to that of complex B formed on a longer DNA probe (Figure IV.7). The minimal sequences to which these complexes bind are not absolutely clear, but apparently involve the polypurine rich sequences centered around position -40. If the complexes are distinct, it would appear that β should bind to a sequence from -57 through -40, while B does not require sequences upstream of position -51. The genetic data presented in Chapter 3 (see Table III.1) implicate these sequences as potential elements for myc protein-mediated repression, as well as basal levels of transcription. However, neither the binding activities nor mobilities of these complexes have been found to vary with respect to myc protein expression, implying that if they are involved in mediating the myc response, they may do so through indirect or unstable interactions with myc protein.

The identity of the B and β complexes with respect to Me1a1, which maps to sequences which correlate with human positions -57 to -40, is unresolved. It is possible that the β complex, if it is distinct from the B complex, is the same as the previously described Me1a1 complex, as it appears to share a similar minimal binding sequence (Asselin, et al, 1989). The B complex may be a novel myc promoter binding complex, as its formation does not appear to require the most upstream sequences of the Me1a1 binding site, and is also not competed by a sequence which binds another complex called Me1a1, which may itself be distinct from the Me1a1 complex described previously (Asselin, et al, 1989, Hall, 1990).
Previous reports have indicated differences between the proteins which interact with the human and murine c-myc P2 promoter proximal sequences. The human c-myc P2 promoter sequences have been reported to interact only with E2F (Thalmeier, et al, 1989), while the murine c-myc P2 promoter has been found to bind only the Me1a1 complex (Asselin, et al, 1989). In this work it has been possible to demonstrate the binding of complexes to regions of the human c-myc promoter from positions -66 through -40 by varying the non-specific carrier DNA used. It thus appears that the specificity of the previously reported complexes resides in the conditions under which the analysis is performed rather than on the source of the c-myc P2 promoter being analysed.

An additional complex, identified in this chapter as complex A, has not previously been described. A characterisation of some of its properties is presented in the next chapter.
Figure IV.1: Detection of myc protein in Rat-1 whole cell extracts

Immunoblot of 100 µg extract prepared from Rat-1 cells infected with control (LJ, -) or v-myc (VM, +) carrying retroviruses. The blot was probed with polyclonal rabbit antiserum (pan-myc) raised against the x-myc1 peptide (Evan, et al, 1988).

xreact = cross reacting non-myc bands

numbers to right of figure indicate relative molecular weight, kD x 10^-3.
Figure IV.2: Complexes detected using salmon sperm DNA as carrier

Gel retardation analysis of complexes which interact with human c-myc promoter probes. 10 µg Rat-1 v-myc cell extract and 1 µg salmon sperm DNA were used in each reaction. The indicated fold molar excess of unlabelled probe was used in self-competitions. Two major (1 and 2) and three minor (*) complexes are indicated. A weak complex which may be non-specific (#) is also indicated.

-66/+49 and -63/+49 probes are Bam HI/Pst I plasmid-derived fragments containing human c-myc promoter sequences with the indicated endpoints relative to the P2 cap site. The ls1 - ls5 probes are annealed oligonucleotides containing sequences between -71 and -22 relative to the human c-myc mRNA P2 cap site. The ls2-ls5 oligonucleotides contain a nine bp linker replacing the wild type sequences as indicated below:

<table>
<thead>
<tr>
<th>ls1</th>
<th>ls2 (-30)</th>
<th>ls3 (-39)</th>
<th>ls4 (-48)</th>
<th>ls5 (-57)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5'</td>
<td>5'</td>
<td>5'</td>
<td>5'</td>
<td>5'</td>
</tr>
<tr>
<td>GATCTggcttgccgggaaaaagacccggagggatcgcgtgagtat attenuatedG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3'</td>
<td>AcgacaaggcccccttttcttgccccctcctagcgcgcactcatattttcgcCAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GAATGGCCCTC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CTTACGGAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>gAAATgCCTC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>cTTACGGAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GaATGocCTC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CTTACGGGAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>GAAATgCCTC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CTTACGGAG</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Capital letters indicate sequences which differ from the wild type human c-myc gene, and the number in parentheses indicates the nt position 3' of the linker.

F: Free probe
carrier: Salmon Sperm DNA

<table>
<thead>
<tr>
<th>probe</th>
<th>-66/+49</th>
<th>-63/+49</th>
</tr>
</thead>
<tbody>
<tr>
<td>competitor</td>
<td>self</td>
<td>self</td>
</tr>
<tr>
<td>fold excess</td>
<td>0</td>
<td>20 200</td>
</tr>
</tbody>
</table>

[Image of gel electrophoresis with bands labeled 1, 2, * for each probe condition.]
The indicated molar excess of unlabelled double strand oligonucleotides was used as competitor in gel retardation assays against the complexes formed on the plasmid-derived -66/+49 human c-myc P2 promoter fragment.

10 µg Rat-1 v-myc cell extract and 1 µg salmon sperm DNA were used in each reaction. The two major (1 and 2) and three minor (*) complexes are indicated.

The $ls_1$ - $ls_5$ fragments are annealed oligonucleotides containing sequences between -71 and -22 relative to the human c-myc mRNA P2 cap site. The $ls_2$-$ls_5$ oligonucleotides contain a nine bp linker replacing the wild type sequences as indicated below:

<table>
<thead>
<tr>
<th>Fragment</th>
<th>5' Sequence</th>
<th>3' Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ls_1$ (WT)</td>
<td>5' GATCTggcttggcggaaagacggaggagtacgccgtgactaatataaagcG</td>
<td>3' AccgaacgcccttttttctgcatccccctccccctgagctgacccatattttccGctcG</td>
</tr>
<tr>
<td>$ls_2$ (-30)</td>
<td>5' GAATGCCTC</td>
<td>3' CTTACGGAG</td>
</tr>
<tr>
<td>$ls_3$ (-39)</td>
<td>5' GaATgCCTC</td>
<td>3' cTTAcGGAG</td>
</tr>
<tr>
<td>$ls_4$ (-48)</td>
<td>5' GaATGcCTC</td>
<td>3' CTTAcGGAG</td>
</tr>
<tr>
<td>$ls_5$ (-57)</td>
<td>5' gaATgCCTC</td>
<td>3' cTTAcGGAG</td>
</tr>
</tbody>
</table>

Capital letters indicate sequences which differ from the wild type human c-myc gene, and the number in parentheses indicates the nt position 3' of the linker.

**F**: Free probe
<table>
<thead>
<tr>
<th>probe: -66/+49</th>
<th>competitor:</th>
<th>ls1</th>
<th>ls2</th>
<th>ls3</th>
<th>ls4</th>
<th>ls5</th>
</tr>
</thead>
<tbody>
<tr>
<td>fold excess:</td>
<td>0</td>
<td>20</td>
<td>200</td>
<td>0</td>
<td>20</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>20</td>
<td>200</td>
<td>0</td>
<td>20</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>20</td>
<td>200</td>
<td>0</td>
<td>20</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>20</td>
<td>200</td>
<td>0</td>
<td>20</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>20</td>
<td>200</td>
<td>0</td>
<td>20</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>20</td>
<td>200</td>
<td>0</td>
<td>20</td>
<td>200</td>
</tr>
</tbody>
</table>

* * *

1

2

F
Figure IV.4: E2F oligonucleotide competitions vs. -66/+49 salmon sperm DNA complexes

Synthetic oligonucleotides of single, double or point mutated double adenovirus E2A promoter E2F sites were used in competitions against labelled -66/+49 human c-myc P2 promoter fragment probe. 10 μg Rat-1 v-myc extract, 1.0 μg salmon sperm DNA and the indicated molar excess of cold oligonucleotide was used in each gel retardation assay. Major (1 and 2) complexes are indicated.

Oligonucleotide sequences are shown; underlined sequences indicate E2F sites and 'x's point mutations

F: Free probe
carrier: Salmon Sperm DNA

competitor: 1x E2F  2x E2F  pm E2F

fold excess:  0  20  200  0  20  200  0  20  200

1x E2F

5' TCGACGTAGTTTTTCGCGCTTAAAGCATG 3'
3' GCATCAAAGCCGGAATTTT

2x E2F

5' TCGACGTAGTTTTTCGCGCTTAAATTTGAGAAAGGGCGCGAACTAGCATG 3'
3' GCATCAAAGCCGGAATTTAAACTCTTTCCCGCCGGCTTTGATC

pm E2F

5' TCGACGTAGTTTTGGGGCTTAAATTTGAGAAAGGGCCCCAAACTAGCATG 3'
3' GCATCAAACCCCGGAATTTAAACTCTTTCCCGGGGTTTGATC
Figure IV.5: Comparison of Rat-1 LJ and VM complexes using -66/+49, ls1 and ls2 probes with salmon sperm DNA carrier

A variety of probes were used in gel retardation assays with both Rat-1 LJ and VM extracts. All binding reactions were performed using 1μg salmon sperm DNA as carrier and 10 μg cell extract. The indicated molar excess of unlabelled probe was used in self-competitions. The two major (1 and 2) and three minor (*) complexes are indicated.

The plasmid-derived -66/+49 human c-myc P2 promoter fragment and ls1 and ls2 double strand oligonucleotides were used as probes and self-competitors.

The ls1 (wild type) oligonucleotide contains sequences from -71:-22 relative to the human c-myc P2 promoter. Ls2 has a nonamer linker replacing the wt sequences as indicated:

```plaintext
ls1 (WT) 5' GATCTggcttcggtggagggagggatcgctgatgtataaagcG
3' AacgaaccgcctttttttttttgtccctcctctagcgcagtctatatatccagCcTAg
ls2 (-30) 5' GAATGCCCTC
3' CTTCCGGAG
```

Capital letters indicate sequences which differ from the wild type human c-myc gene, and the number in parentheses indicates the nt position 3' of the linker.

F: Free probe
<table>
<thead>
<tr>
<th>carrier:</th>
<th>SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>probe:</td>
<td>-66/+49 ls1 ls2</td>
</tr>
<tr>
<td>extract:</td>
<td>LJ VM LJ VM LJ VM</td>
</tr>
<tr>
<td>fold excess:</td>
<td>0 20 200 0 20 200 0 20 200 0 20 200 0 20 200</td>
</tr>
</tbody>
</table>

* * *

1 2

F
Figure IV.6: Comparison of Rat-1 LJ and VM complexes using ls3, ls4 and ls5 probes with salmon sperm DNA carrier

A variety of probes were used in gel retardation assays with both Rat-1 LJ and VM extracts. All binding reactions were performed using 1μg salmon sperm DNA as carrier and 10 μg cell extract. The indicated molar excess of unlabelled probe was used in self-competitions. The two major (1 and 2) and three minor (*) complexes are indicated.

The plasmid-derived -66/+49 human c-myc P2 promoter fragment and ls3, ls4 and ls5 double strand oligonucleotides were used as probes and self-competitors.

The ls3, ls4 and ls5 oligonucleotides contain sequences from -71:-22 relative to the human c-myc P2 promoter with nonamer linker mutations:

<table>
<thead>
<tr>
<th></th>
<th>-71</th>
<th>-66</th>
<th>TATA</th>
<th>-22</th>
</tr>
</thead>
<tbody>
<tr>
<td>ls1 (WT)</td>
<td>GATCTggctgtgcagggaaagacaacggagggatcgcgtgatataaaagcG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3'</td>
<td>TTTGCTTCGCT CTCGCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ls3 (-39)</td>
<td>gAAgCCTC</td>
<td>gAAgCCTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3'</td>
<td>TTTGCTTCGCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ls4 (-48)</td>
<td>gAAgCCTC</td>
<td>gAAgCCTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3'</td>
<td>TTTGCTTCGCT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ls5 (-57)</td>
<td>gAAgCCTC</td>
<td>gAAgCCTC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3'</td>
<td>TTTGCTTCGCT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Capital letters indicate sequences which differ from the wild type human c-myc gene, and the number in parentheses indicates the nt position 3' of the linker.

F: Free probe
Figure IV.7: Complexes detected using poly d(IC)•poly d(IC) as carrier

Gel retardation analysis of complexes which interact with human c-myc promoter probes was performed. 10 μg Rat-1 cell extract and 1.5 μg poly d(IC)•poly d(IC) were used in each reaction. 10 μg Rat-1 v-myc extracts were used in the first 6 lanes, and 10 μg Rat-1 control extracts in the others. The indicated fold molar excess of unlabelled probe was used in self-competitions. Three major complexes (A, B and β) are indicated.

The -66/+49, -63/+49, -51/+49, -44/+49 and -66/+22 probes are Bam HI/Pst I plasmid-derived fragments containing human c-myc promoter sequences with the indicated endpoints relative to the P2 cap site. The ls1 - ls5 probes are annealed oligonucleotides containing sequences between -71 and -22 relative to the human c-myc mRNA P2 cap site. The ls2 - ls5 oligonucleotides contain a nine bp linker replacing the wild type sequences as indicated below:

| ls1 (WT) | -71 | 6′GATCTggcttgcgggaaaaagactgcgtgaagtaaaacG |
| ls2 (-30) | 5′ | AccgaaccgcctttttccttcctccctcctcctagccgactcatatatttcgCcTaG |
| ls3 (-39) | 5′ | gAAATgCCTC |
| ls4 (-48) | 5′ | cTTAcGGAG |
| ls5 (-57) | 5′ | gAAATgCCTC |

3′ Ac cc ga a c cc c t t t t c t t g c c t c c c t a g c g c g a c t c a t a t t t t c g C c T a G

Capital letters indicate sequences which differ from the wild type human c-myc gene, and the number in parentheses indicates the nt position 3′ of the linker.

F: Free probe
Figure IV.8: Linker scan competitions vs. -63/+49 poly d(IC)•poly d(IC) complexes

The indicated molar excess of unlabelled double strand oligonucleotides was used as competitor in gel retardation assays against the complexes formed on the plasmid-derived -63/+49 human c-myc P2 promoter fragment.

10 µg Rat-1 v-myc cell extract and 1.5 µg poly d(IC)•poly d(IC) polymer were used in each reaction. The two major (A and B) complexes are indicated.

The ls1 - ls5 fragments are annealed oligonucleotides containing sequences between -71 and -22 relative to the human c-myc mRNA P2 cap site. The ls2-ls5 oligonucleotides contain a nine bp linker replacing the wild type sequences as indicated below:

| lsl (WT)     | 5'GATCTggcttgccggacagggcatcgctgatgtataaaagcG 3' |
| lsl (-30)    | 5'GAAATGCTTC 3' |
| lsl (-39)    | 5'gATGcCTTC 3' |
| lsl (-48)    | 5'gATGcCTTC 3' |
| lsl (-57)    | 5'gATGcCTTC 3' |

Capital letters indicate sequences which differ from the wild type human c-myc gene, and the number in parentheses indicates the nt position 3'.

F: Free probe
<table>
<thead>
<tr>
<th>probe:</th>
<th>-63/+49</th>
</tr>
</thead>
<tbody>
<tr>
<td>carrier:</td>
<td>pdIC•pdIC</td>
</tr>
<tr>
<td>competitor:</td>
<td>ls1</td>
</tr>
<tr>
<td>fold excess:</td>
<td>0</td>
</tr>
</tbody>
</table>

[Image of gel electrophoresis with bands labeled A, B, and F]
Figure IV.9: E2F oligonucleotide competitions vs. 
-66/+49 poly d(IC)•poly d(IC) complexes

Synthetic oligonucleotides of single, double or point mutated double adenovirus E2A promoter E2F sites were used in competitions against labelled -66/+49 human c-myc P2 promoter fragment probe. 10 μg Rat-1 v-myc extract, 1.5 μg poly d(IC)•poly d(IC) and the indicated molar excess of cold oligonucleotide was used in each gel retardation assay. Major (A and B) complexes are indicated.

Oligonucleotide sequences are shown; underlined sequences indicate E2F sites and 'x's base changes in the mutant oligonucleotide

F: Free probe
carrier: pu(I)C-pu(I)C

competitor: 1xE2F 2xE2F pmE2F

fold excess: 0 20 200 0 20 200 0 20 200

1xE2F

5' TCGACGTAGTTTTCGCGCTTAAAGCATG 3'
3' GCATCAAAAGCGCGAATTTC 5'

2xE2F

5' TCGACGTAGTTTTCGCGCTTAAATTTGAGAAAGGGCGCGAAACTAGCATG 3'
3' GCATCAAAAGCGCGAATTTAAACTCTTTCCCGCGCTTTGATC 5'

pmE2F

5' TCGACGTAGTTTTCGCGCTTAAATTTGAGAAAGGGCCCAACTAGCATG 3'
3' GCATCAAAACCAGCTAAACTCTTTCCCGGGTTTGATC 5'
Figure IV.10: 'Me1a1' oligonucleotides as probes and competitors vs. poly d(IC)•poly d(IC) complexes

Synthetic oligonucleotides of the murine c-myc 'Me1a1' sites (Hall, 1990) were used in competitions against labelled -66/+49 human c-myc P2 promoter fragment probe. The wild type (WT) oligonucleotide was also used as probe and self-competed or competed with the two mutant 'Me1a1' oligonucleotides (MT1 and MT2). 10 μg Rat-1 control extract, 1.5 μg poly d(IC)•poly d(IC) and the indicated molar excess of cold oligonucleotide was used in each gel retardation assay. Major (B and b) complexes are indicated.

The sense strand of the wild type 'Me1a1' oligonucleotide probe is shown; underlined sequences indicate the myc sequences and T residues base changes in the mutant sequences.

F: Free probe
probe: -66/+49

<table>
<thead>
<tr>
<th>competitor</th>
<th>WT</th>
<th>MT1</th>
<th>MT2</th>
</tr>
</thead>
<tbody>
<tr>
<td>fold excess</td>
<td>0</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>

'melal'

<table>
<thead>
<tr>
<th>competitor</th>
<th>WT</th>
<th>MT1</th>
<th>MT2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>10</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>10</td>
<td>100</td>
</tr>
</tbody>
</table>

GATCGGGACGGGAGGGAGATC

WT GATCGGGACGGGAGGGAGATC
MT1 T T
MT2 T T
Figure IV.11: Comparison of Rat-1 LJ and VM complexes using -66/+49, lsl and ls2 probes with poly d(IC)•poly d(IC) carrier

Complexes detectable with various human c-myc P2 promoter probes were self competed and compared between Rat-1 v-myc (VM) and control (LJ) extracts. 10 μg extract, 1.5 μg poly d(IC)•poly d(IC) and unlabelled competitor in the molar excess indicated were used with each probe. Major complexes (A, B and β) are indicated. The probes used in these reactions were plasmid-derived -66/+49 fragment, lsl and ls2 oligonucleotides.

The lsl (wild type) oligonucleotide contains sequences from -71:-22 relative to the human c-myc P2 promoter. Ls2 has a nonamer linker replacing the wt sequences as indicated:

```
-71  -66
lsl (WT) 5'GATCTggcttgcccggaaagaacggaggggtacgcgctgatgtataaaagcG
   3' AacgaaccggccttttctttgcccctccctaggcgacatcctctctctctctcG
ls2 (-30) 5' GAATGCCCTC
          3' CTTACGGAG
```

Capital letters indicate sequences which differ from the wild type human c-myc gene, and the number in parentheses indicates the nt position 3' of the linker.

F: Free probe

The right hand panel contains an extended exposure of the -66/+49 probe complexes.
<table>
<thead>
<tr>
<th>carrier:</th>
<th>pd(IC) \cdot pd(IC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>probe:</td>
<td>-66/+49</td>
</tr>
<tr>
<td></td>
<td>ls1</td>
</tr>
<tr>
<td></td>
<td>ls2</td>
</tr>
<tr>
<td>extract:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LJ</td>
</tr>
<tr>
<td>fold excess:</td>
<td>0</td>
</tr>
</tbody>
</table>

[Image of gel electrophoresis results with bands labeled A, B, F, and β.]
Figure IV.12: Comparison of Rat-1 LJ and VM complexes using ls3, ls4 and ls5 probes with poly d(IC)•poly d(IC) carrier

Complexes detectable with various human c-myc P2 promoter probes were self competed and compared between Rat-1 v-myc (VM) and control (LJ) extracts. 10 μg extract, 1.5 μg poly d(IC)•poly d(IC) and unlabelled competitor in the molar excess indicated were used with each probe. The major complex (β) is indicated. The probes used in these reactions were ls3, ls4 and ls5 annealed double strand oligonucleotides.

The ls3, ls4 and ls5 oligonucleotides contain sequences from -71:-22 relative to the human c-myc P2 promoter with nonamer linker mutations:

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>Sequence (5'→3')</th>
<th>Position 3' of the linker</th>
</tr>
</thead>
<tbody>
<tr>
<td>ls1 (WT)</td>
<td>5′GATCTggcttgccggaaaagcgggagggaggggctcgctgggtataaaaaagG</td>
<td>3′</td>
</tr>
<tr>
<td>ls3 (-39)</td>
<td>5′Accgaaccggccctttttttttctgtcctccctctcttagccgactcatttttcgg</td>
<td>3′</td>
</tr>
<tr>
<td>ls4 (-48)</td>
<td>5′GaATGcCTC</td>
<td>3′</td>
</tr>
<tr>
<td>ls5 (-57)</td>
<td>5′gAATGcCTC</td>
<td>3′</td>
</tr>
<tr>
<td></td>
<td>gAATGcCTC</td>
<td>cTTAcGGAG</td>
</tr>
</tbody>
</table>

Capital letters indicate sequences which differ from the wild type human c-myc gene, and the number in parentheses indicates the nt position 3' of the linker.

F: Free probe
<table>
<thead>
<tr>
<th>carrier</th>
<th>probe</th>
<th>extract</th>
<th>fold excess</th>
</tr>
</thead>
<tbody>
<tr>
<td>pd(IC)•pd(IC)</td>
<td>ls3</td>
<td>LJ</td>
<td>0 20 200</td>
</tr>
<tr>
<td></td>
<td>ls4</td>
<td>VM</td>
<td>20 200 0</td>
</tr>
<tr>
<td></td>
<td>ls5</td>
<td>LJ</td>
<td>20 200 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>VM</td>
<td>20 200 0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>20 200</td>
</tr>
</tbody>
</table>

β

F
Table IV.1: Summary of complex B and β binding and competition data

Data defining human c-myc P2 promoter sequence probes which will bind (Bind) complexes B or β, as well as sequences which will compete (Comp) the formation of complex B.

Consensus sequences are presented at the bottom of the figure. Capital letters indicate those sequences which appear critical for binding, while underlines indicate the relative importance of a sequence for complex competition. The extended sequence of the complex B consensus indicates that the 3' boundary of this site is poorly defined. (See text for further details.)

Within the various probe sequences, lower case letters indicate wild type human c-myc sequences, capital letters non-myc sequences and double underlined sequences those positions which are mutated in the linker scan series oligonucleotides.

Sequence coordinates are given relative to the P2 promoter cap site.

Hall: Me1a1 complex binding site as defined by Hall (1990)
ls1 to ls5: Linker scan series double strand oligonucleotides
| -66 | GATCCGGgagggaggggaaggagtctagctgatataaaagcagtttcggggtcttataactcgcctgatatactcgcctgtaaattccagcgaagggcgaagggcgcgggctgctgca | + | + |
| -65 | GATCCGGgagggaggggaaggagtctagctgatataaaagcagtttcggggtcttataactcgcctgtaaattccagcgaagggcgaagggcgcgggctgctgca | + | + |
| -64 | GATCCGGgagggaggggaaggagtctagctgatataaaagcagtttcggggtcttataactcgcctgtaaattccagcgaagggcgaagggcgcgggctgctgca | + | + |
| -63 | GATCCGGgagggaggggaaggagtctagctgatataaaagcagtttcggggtcttataactcgcctgtaaattccagcgaagggcgaagggcgcgggctgctgca | + | + |
| -62 | GATCCGGgagggaggggaaggagtctagctgatataaaagcagtttcggggtcttataactcgcctgtaaattccagcgaagggcgaagggcgcgggctgctgca | + | + |
| -61 | GATCCGGgagggaggggaaggagtctagctgatataaaagcagtttcggggtcttataactcgcctgtaaattccagcgaagggcgaagggcgcgggctgctgca | + | + |
| -60 | GATCCGGgagggaggggaaggagtctagctgatataaaagcagtttcggggtcttataactcgcctgtaaattccagcgaagggcgaagggcgcgggctgctgca | + | + |

**β consensus**

| GGAAGGAGCCGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAGAGGAAGAg
Chapter 5 The MAC II complex
5.1 Introduction

Two major differences have been observed in the formation of protein:DNA complexes on the human c-myc P2 promoter proximal sequences between extracts from Rat-1 cells expressing v-myc, and thus down-regulating endogenous c-myc expression, and control extracts. First, E2F binding activity, which has been observed only in the presence of salmon sperm carrier DNA, is increased by up to two fold in the v-myc extracts. Second, complex A, seen in the presence of the synthetic poly d(IC)•poly d(IC) heteropolymer, is greatly reduced in binding activity in the v-myc extracts.

In light of the intricate relationships between intracellular signalling pathways, myc levels, and cellular growth state it was important to address if either of these changes in DNA binding activity could be attributable to myc protein levels in the extracts, which might be expected if these complexes are involved in c-myc autoregulation. If myc is present as a limiting factor in one of the DNA binding complexes, then a direct correlation between myc protein levels and signal intensity would be expected. Conversely, its presence might not be limiting for complex formation, and thus no change in signal intensity would be expected. The presence of myc protein in a complex, even as a non-limiting component, might, however, be expected to alter complex mobilities between the LJ and VM extracts, as they express myc proteins of different sizes (62 vs. 110 kd). Protein size-dependent complex mobilities have been demonstrated previously using in vitro analyses of other DNA binding proteins (eg GCN4, (Hope and Struhl, 1986)). If myc levels indirectly regulate complex formation they would be expected to lead to either an increase or decrease in signal intensity, but without a concomitant change in band mobility. The most direct way of distinguishing amongst these alternatives was to add antibodies raised against the myc protein to the gel retardation assay.
If myc is involved in complex formation, it would be expected that the antibody would either reduce the mobility of the DNA:protein complex (a 'supershift'), or alter the signal intensity (via blockage of myc activity).

Two monoclonal anti-myc antibodies, raised against peptide antigens, were used for these studies. PM-3E7 (mAb1) was raised to a peptide of sequence APSEDIWKFELL (pep1) which is most highly conserved amongst all members of the myc family, and corresponds to positions 44-55 of the human c-myc protein (Evan, et al, 1988). It specifically detects c-, v-, N- and L-myc proteins from a wide range of vertebrate species including humans, mice and quails in immunoprecipitations, western blots, and ELISA assays. The monoclonal anti-myc antibody Mycl-3C7 (mAb2) was raised to an epitope in the peptide CSTSSLYLQDLSAAASEC (pep2), which has been mapped to the carboxy-terminal 15 amino acids of the peptide, and corresponds to human c-myc amino acids 173-188. This sequence is moderately conserved amongst myc protein family members, and the antibody has been shown to immunoprecipitate and immunoblot p62c-myc from murine and human sources, and will detect avian p110y-gag-myc in ELISA assays (Evan, et al, 1985, G. Evan, personal communication).

5.2 Results

5.2.1 Effect of antibodies on specific complexes

The effect of each antibody was determined by incubating it with the control extract/DNA mix and then resolving the resultant complexes by PAGE. The antibodies were added either before or after addition of the DNA probe. The specificity of any antibody-induced effects was assessed by preincubating each antibody with its cognate antigenic peptide or with the non-cognate peptide. In addition, an antibody raised to the biologically unrelated LFA-1 α chain (mAb38) was also tested (Dransfield and Hogg, 1989).
The results obtained when complex A was analysed using the control Rat-1 extracts are quite striking (figure V.1). A supershift of all of complex A is seen when either mAb1 or mAb2 is used, the alteration in mobility being greater with mAb1 than with mAb2. The supershift is completely blocked in each case when the cognate peptide is used, and unaffected by the incorrect peptide, as expected for the specific recognition of epitopes in the complex by each antibody. A supershift is detectable whether the antibodies are added preceding or following probe addition, with a slight increase in complex intensity seen in the post-probe addition samples. Complex A is the only one which is specifically affected by the anti-myc antibodies; none of the other complexes detected with either poly d(IC)•poly d(IC) or salmon sperm DNA carriers is specifically affected by antibody addition (see below, and data not shown).

This result is quite unexpected, as complex A is reduced in intensity but not altered in mobility in the v-myc extracts, and the v-myc extracts contain more myc protein than the control extracts, in the form of the larger v-myc protein (see Figure IV.1). The most straightforward interpretation of these results is, however, that myc protein is present in this complex, and the antibodies are binding to exposed epitopes on the protein, leading to a decrease in its electrophoretic mobility. Furthermore, the results with v-myc suggest that this form of the protein is unable to enter complex A, which is thus proportional to c-myc levels. The only firm conclusion, however, is that the complex shares at least two epitopes with myc protein, in one or more of its components. It has thus been renamed MAC II, for Myc-related Antigen Complex, double strand DNA binding. (See Chapter 6 for a description of the MAC I complex).

5.2.2 Relationship of MAC II complex, myc protein levels and cell type

The somewhat paradoxical relationship between v-myc protein levels and MAC II levels becomes more unusual when extracts from Rat-1 cells
expressing high levels of other myc proteins are compared. Extracts from multicolony pools expressing wild type human c-myc, human c-myc deletion mutant Δ41-53 (which spans epitope 1), and a clone which expresses high levels of human c-myc deletion mutant Δ145-262 (which spans epitope 2) (Penn, et al, 1990b) were assayed (Figure V.2). A complex of the appropriate mobility was detected in each of the human c-myc protein expressing lines. The two colony pool extracts have slightly increased signals relative to control, while the Δ145-262 clonal cell extract has the highest levels of activity. The v-myc extract in this figure clearly contains less MAC II activity than the other extracts. The inability to detect much MAC II activity in v-myc expressing Rat-1 extracts has to date been observed with four independent extract preparations (data not shown).

This figure also shows that extracts made from murine NIH-3T3 and human HeLa cells both contain an activity which approximately comigrates with MAC II. The NIH-3T3 v-myc extracts are made from cells which express high levels of v-myc protein but do not down regulate endogenous c-myc expression, and HeLa cells contain very high levels of c-myc protein (Moore, et al, 1987, Penn, et al, 1990a). A similar amount of binding activity is detectable in the control and v-myc NIH-3T3 extracts. The Rat-1 complex usually consists of two bands, the upper of which is more intense. They cannot, however, always be resolved and sometimes appear as a single broader band. Both the NIH-3T3 and HeLa complexes seem to migrate as single bands, which may correspond to the upper of the two Rat-1 bands. A band corresponding to the B complex is also seen in each of these extracts.

Both the NIH-3T3 and HeLa complexes are also specifically supershiftable by mAb1 and mAb2 (HeLa data in figure V.3 and NIH-3T3 data not shown). The increased degree of supershift by mAb1 seen with the HeLa relative to the Rat-1 complex appears to be due to a higher titre.
preparation of antibody (see also data in Chapter 6). Thus while a complex with the mobility of MAC II is detectable in extracts from a variety of cell types expressing high levels of myc protein, it is only the Rat-1 v-myc extracts which have a reduced MAC II level. There does not appear to be a simple correlation between myc protein and MAC II levels in the other extracts, although there is a general tendency towards increased MAC II levels in c-myc overexpressing cells. If the complex actually contains myc protein, it is possible that v-myc, but not c-myc, protein is excluded from it. If this is happening, then the observed signal intensities would approximate c-myc, but not total c- and v-myc levels. While the amount of B complex signal within one cell type is independent of myc protein levels, it can vary between cell types.

5.2.3 Peculiar probe characteristics

DNA probes, such as the -66/+49 fragment, which display MAC II binding activity were found to have very unusual properties. For example, two preparations of fragment (prep 1 and prep 2), both isolated from the same batch of pUCSVter-66/+49IFN plasmid, purified by gel electrophoresis, and labelled to the same specific activity, display very different affinities for MAC II, while the B complex and the other, minor complexes all bind with constant affinities (Figure V.4, lanes 1 and 3). Additionally, the affinity of probe prep 1 for MAC II is reduced following gel repurification of the labelled fragment (lanes 1 and 2). This reduction in affinity does not appear to be due to the gel purification procedure, as the same fragment, recovered from an independent plasmid preparation, end labelled and then gel purified (prep 3) has shown stronger affinity for MAC II than the probe derived from the second preparation (lanes 6 and 7). There does not appear to be an easily separable contaminant inhibiting MAC II binding, as repurification of the probe from prep 2 on a NACS 52 ionic exchange resin column does not alter the difference in affinities.
between it and the probe from prep 1, which binds more MAC II (lanes 1, 3 and 5). The ability of these probes to bind to the B complex is not affected by any of these manipulations, as lanes 1 through 5 are all taken from the same exposure of the same gel and display almost identical levels of B complex binding, contrasting markedly with MAC II binding.

Mixing experiments were performed to further test the presence of a putative inhibitor acting in trans. Two probe preparations, labelled to the same specific activity, one of which binds MAC II well (prep 2) and one of which does not (prep 4), were combined in varying proportions and used in a gel retardation assay (lanes 8-13). The resultant probe mixtures display MAC II binding activity in direct proportion to the amount of prep 2 probe. In contrast the B complex signal is constant through all of the samples, indicating that both probes bind it with similar affinities. Thus any MAC II inhibitory or stimulatory activity would be stoichiometric with probe concentration. More likely, this result suggests that binding is not affected by external factors in the probe preparations, but that an intrinsic and variable feature other than sequence is determining the binding of MAC II.

Synthetic oligonucleotides containing the exact sequence of the -66/+49 probe, including the flanking restriction endonuclease sites were prepared and either individually end-labelled with polynucleotide kinase and subsequently annealed or annealed and then labelled with the Klenow fragment of E. coli DNA polymerase. When these probes were used in the gel retardation assay, only a smear, running the length of the gel track, and no specific complexes, were detected, regardless of the labelling/annealing procedure used (Figure V.4, lane 14). Following gel purification of the labelled, annealed oligonucleotides, no MAC II signal is detectable, while specific binding of the B complex, and many minor complexes, is readily seen (lane 15). To control for the possibility that the oligonucleotides were
not annealing completely, a 12 base oligonucleotide complementary to the 3' end of one strand was end labelled and primer extended to generate full length, double stranded DNA. This probe also bound complex B with high affinity, but did not detect MAC II (lane 16). In fact, all attempts to detect MAC II with oligonucleotide probes have thus far been unsuccessful.

5.2.4 Intrinsic and extrinsic factors

These results suggest that the plasmid derived probe has some property, the level of which varies amongst preparations, which enables it to bind MAC II. Possible factors which are intrinsic to the DNA may include DNA secondary structure, base modifications such as methylation or the tight, possibly covalent, attachment of small molecules. Possible DNA extrinsic factors include protein, organic, heavy metal or salt contaminants. A number of experiments have been performed to assess these possibilities.

Any extrinsic contaminants associating with the DNA must be resistant to removal by all of the procedures used in the fragment preparation, and be active in a direct proportion to the number of DNA molecules present in the binding reaction. The manipulations undertaken with the DNA include alkaline/SDS bacterial lysis, double CsCl/EtBr banding, phenol and phenol/chloroform extraction, ethanol and isopropanol precipitation, fragment isolation by native agarose or polyacrylamide gel electrophoresis, and size exclusion (Sephadex G-50 medium) and ion exchange (NACS-52 resin) chromatography. The labelled fragments which subsequently bound MAC II did not appear to migrate anomalously on native PAGE gels (data not shown), probably excluding large contaminating molecules, which would be expected to alter fragment mobility.
In addition, the probe which was capable of binding MAC II was insensitive to digestion with the RNAases H, T1 and A, as well as the single strand DNA specific nuclease S1 (data not shown). These results would appear to eliminate the possibility that MAC II recognises an RNA contaminant, RNA:DNA hybrids or single stranded DNA.

Various reagents which might have contaminated the DNA or altered its structure were tested for their ability to alter MAC II formation with a probe preparation which displayed weak binding activity. Ethidium bromide (5×10⁻⁶-5.0 mg/ml), chloroquine (0.75 μm-7.5 mM), spermidine (5 mM), formamide (5-10%), MgCl₂ (10 mM), ZnSO₄ (10⁻⁵-2.5×10⁻³ M) and all 4 NTP and dNTPs in combination (1mM each) were found to have minimal effects on MAC II formation, with 7.5 mM chloroquine, 1-5 mg/ml EtBr and 1-2.5 mM ZnSO₄ being generally inhibitory to any complex formation (data not shown). Optimal NaCl concentration from 50 to 250 mM NaCl (100 mM for MAC II; 50-100 mM for B) and pH values from 7.5 to 8.6 (7.5-8.0 for both complexes) for the binding reaction were also determined, and are very similar to the conditions which had been used initially (50 mM NaCl, pH8.0; data not shown). Thus it was not possible to greatly increase or selectively decrease the affinity of the -66/+49 fragment probe for MAC II simply by adding any of a variety of reagents, some of which are plausible contaminants, to the bandshift reaction.

5.2.5 Thermolability of MAC II binding capacity of probe

The temperature stability of the MAC II binding activity of the probe was also assessed. -66/+49 probe which can bind MAC II was heated to 95°C and then either immediately placed on ice or allowed to reanneal by cooling to room temperature over several hours, then stored at 4°C. When used in gel retardation assays (Figure V.5), the rapidly cooled probe detects a smear of bands with moderate mobilities, and one distinct band
of greatly reduced mobility relative to free probe, which migrates significantly ahead of the control MAC II band. The partially reannealed probe detects all of these bands as well as one which comigrates with complex B. In neither of the lanes is a complex which comigrates with MAC II visible. The free probe in the rapidly cooled sample migrates at a position consistent with single strand DNA, while that of the slowly cooled sample is a combination of both double and single strand DNA mobilities, demonstrating that the slowly cooled probe contains reannealed as well as single strand components, while the rapidly cooled probe is predominantly single strand DNA.

Complex B, but not MAC II, binding can thus be reconstituted following probe reannealing. This finding is reminiscent of the results seen using annealed synthetic oligonucleotides, which also bound B but not MAC II. The activity which allows MAC II binding is thus unstable at elevated temperatures, or when double strand DNA is removed. This implies that stable modifications such as methylation are probably not important, but does not exclude either heat labile contaminants or secondary structure conformations which are lost upon DNA melting.

The thermal stability of the binding activity was assessed more rigorously by heating probe in TE for 5 min through a range of temperatures between 45°C and 75°C, followed by rapid cooling on ice before incubation with cellular extracts (Figure V.6, left panel). Binding activity is unaffected by heating at 53°C, reduced at 57°C and 61°C, and almost eliminated at 65°C. In contrast complex B binding is unaffected by the heat treatment, except in the 72°C and 75°C treatments, where the probe is completely melted. When 50 mM NaCl was added to the probe and the identical heating regimen followed, MAC II binding activity was dramatically stabilised (Figure V.6, right panel). MAC II binding activity under these conditions is readily detectable through the 69°C treatment,
with mild and sharp reductions at 72°C and 75°C, respectively. Thus the addition of 50 mM NaCl to the probe in TE increases the transition temperature (Tt) by approximately 14°C. The double strand nature of the DNA is also stabilised by salt addition, as even at 75°C all of the probe remains double stranded and, as expected, binds the B complex. The melting temperature (Tm) of the probe in these conditions is approximately 83°C. Addition of formamide (40% v/v final concentration) to probe in TE prior to heating results in a Tt of 49°C and Tm of 57°C, while the further addition of 50 mM NaCl increases these values to a Tt of 53°C and Tm of 61-64°C (data not shown). The MAC II binding capacity of the probe is thus less thermostable than its double strandedness, and the difference between the Tm of the probe and the Tt for MAC II binding is approximately 8-10°C under several different conditions. This maintenance of a constant differential between the transition and melting temperatures of the probe suggests that this feature is stabilised via hydrogen bonded interactions, in a similar manner to the stabilisation of double stranded DNA structures.

The most likely explanation of this feature is that some form of stable secondary structure is imparted to the DNA while the fragment is present in the supercoiled plasmid. The presence of an additional component besides the DNA fragment is highly unlikely. If there is, it must be very small, as no mobility difference is observed between probes which do or do not bind MAC II. Even a small protein binding to the DNA is unlikely as numerous treatments which should have disrupted DNA:protein interactions (high salt buffers, detergents, extraction from organic solvents) failed to disrupt binding capability, while the addition of low concentrations of NaCl would not be expected to dramatically increase the thermal stability of a DNA:protein interaction. Small RNA molecules are probably not associated with the probe, as enzymes specific for DNA:RNA
hybrids have not altered binding activity. Finally, any non-nucleic acid molecule associated with the DNA would have to do so with very high affinity and sequence specificity, but could not be any of the ones tested to date. Thus a highly stable secondary structure seems the most likely, if unexpected, possibility for imparting MAC II binding capacity to a DNA fragment.

5.2.6 Deletion mapping of MAC II binding site

A series of 5' and 3' deletions starting from the -66/+49 probe were used in an attempt to map the minimal sequences required to bind MAC II. Each probe was used in conjunction with both mAb1 and peptide blocked mAb1 on Rat-1 control extracts. Initial data displayed in Chapter 4 showed that both the -66/+49 and -63/+49 probes can bind a complex of similar mobility. The -63/+49 complex was also found to be specifically supershiftable with mAb1 (Figure V.7). The additional 5' deletions to -51, -44, -22 and -13 do not yield a complex of similar mobility (Figures V.7 and V.8). However on very long autoradiographic exposures, there appear to be weak complexes which bind to the -51 and -44 deletions that may be supershifted by the antibody (Figure V.6, lower panel). This result is quite tentative as the signal:noise ratio on this exposure is quite poor. None of the 3' deletion probes, with endpoints at +22, -21 and -49 bind complexes which are supershiftable. Even extended exposures of the +22 deletion, retaining the greatest amount of 3' sequence, does not appear to bind a supershiftable complex (Figure V.8, lower panel, and data not shown).

Thus in order to bind a complex which is unambiguously identifiable as MAC II on the basis of both its mobility and supershiftability by mAb1, almost 120 bp of DNA appear to be required. Truncation of only 12 bp from the 5' end to position -51 eliminates almost all binding activity, while the removal of ~20 bp from the 3' end completely abolishes all detectable binding. However, given that there appears to be some
component in addition to its sequence required of the probe in order to bind MAC II, it is quite difficult to interpret these data. At this stage one can only conclude which sequences can bind MAC II, but not those which can not.

5.2.7 Competition mapping of MAC II binding site

Competition analyses were performed to attempt to map the sequence requirements for MAC II binding, much as had been done for the E2F complex (see Chapter 4). The endlabelled -66/+49 P2 fragment was used as probe, and 5-200 fold molar excesses of various unlabelled fragments or annealed complementary oligonucleotides were used as competitors (Figure V.9). Efficient competition of MAC II is observed with the -66/+49, -66/+22 and -13/+49 sequences. -71/-21 (the ls1 oligonucleotide) does not compete at all, even at 200 fold molar excess. The intersection of the successfully competing fragments delineates the region from -13 to +22, which has not yet been tested. This sequence does not contain any previously identified protein binding sites, and even excludes the TATAA box, although it contains the mRNA cap site. Fragment preparations which were incapable of binding MAC II but contained the entire -66/+49 sequence also efficiently competed the -66/+49 probe for MAC II binding (data not shown). This result suggests that there is a DNA sequence requirement in addition to the putative structural requirement for MAC II binding.

Complex B is as well competed by the -66/+22 and -71/-21 sequences as it is by the full length probe fragment, while the -13/+49 fragment does not compete efficiently. These results are consistent with the competition data of Chapter 4, in which complex B was assigned to sequences within the -57/-40 region.
5.2.8 Thermostability of MAC II in extracts

The capability of the cell extracts to maintain MAC II activity following heat treatment was also assessed. This experiment was performed in two ways: either extracts were heated and cooled prior to addition of the probe, or the complexes were preformed on the DNA and the entire mixture heated and then cooled. From the experiments performed on the thermostability of the probe, it could be anticipated that at 100 mM, the NaCl concentration of the assay, the probe would maintain MAC II binding activity well above 68°C. The results of this experiment are presented in Figure V.10. Both sets of conditions display unchanged binding of MAC II from 4°C to 45°C. Binding is completely abolished following treatment at 52°C, regardless of the presence of probe. Interestingly, the B complex is highly heat sensitive, being undetectable in extracts warmed even to 37°C. This complex is destabilised by treatment of extracts at temperatures as low as 22°C (data not shown). A more precise series of heat treatments of the extracts without probe indicates that between 49.5°C and 51.7°C MAC II binding activity is effectively eliminated (Figure V.10). One minor high mobility complex is more heat resistant than MAC II, displaying no reduction in binding activity following 53°C treatment, but disappearing after incubation at 68°C (data not shown). Rat-1 extracts demonstrated a similar thermal stability profile to that of the HeLa extracts shown in Figure V.10 (data not shown). The ability of MAC II to form following heat treatment of extracts at temperatures which completely abolish most other specific binding complexes implies that none of these higher mobility complexes is a necessary precursor of MAC II. Whether the heat resistant minor complex is related to MAC II is less clear, though its binding is independent of MAC II binding when MAC II incompetent probes are used (see eg mixing experiments, Figure V.4).
It is intriguing that the MAC II complex is destabilised following incubation at 50°C. The MAC II binding site appears to encompass the P2 promoter TATA box. One of the major TATA box binding activities which is required for in vitro transcription is the TFIID fraction. TFIID activity in cell extracts is lost following a fifteen minute incubation at 47°C (Nakajima, et al, 1988). The similar thermal instability of MAC II and TFIID suggests that MAC II may contain TFIID protein; it will be interesting to learn if this is so.

5.3 Discussion and Conclusions

An initial characterisation of the MAC II protein:DNA complex has been presented in this chapter. The complex has been demonstrated to be immunologically related to the c-myc protein, and its protein components have been shown to interact with DNA sequences from the human c-myc P2 promoter. In addition to its primary sequence, the DNA binding site of this complex seems to require an additional, metastable DNA structure.

5.3.1 Presence of myc protein in MAC II

It has been reproducibly demonstrated using extracts from several cell lines and two different monoclonal antibodies that MAC II is immunologically related to human c-myc protein. It seems likely that the antigenic determinants are due to the presence of myc protein in the complex. Several pieces of data are consistent with this interpretation, while others may argue against it.

The immunological data is the strongest evidence in favor of the presence of myc protein in the complex. The two monoclonal antibodies, mAb1 and mAb2, were raised against short peptides derived from the human c-myc protein sequence. These antibodies have been demonstrated to recognise v- and c-myc protein, and no single other protein appears to cross react with both of them (Evan, et al, 1985, Evan, et al, 1988, and G Evan, D Hancock and T Littlewood, personal...
communication). Similarly, in searches of databases containing translations of approximately thirty thousand cloned sequences (OWL protein database v. 9), only myc family proteins show significant homologies with both the mAb1 and mAb2 cognate peptides (data not shown). It is possible that each antibody is cross reacting with a different non-myc protein, as there are some proteins which can cross react with each antibody in immunoprecipitation or immunoblotting reactions. However, their identities are currently unknown, and of the proteins detected in the database search, none are known or suspected sequence specific DNA binding proteins (G Evan, D Hancock and T Littlewood, personal communication; and data not shown).

There are myc family members in addition to v- and c-myc which may be recognised by the mAb1 and mAb2 antibodies. mAb1 was raised to a peptide which contains one of the most highly conserved sequences amongst the myc protein family. It should be capable of recognising c-myc, N-myc, L-myc, B-myc and S-myc proteins if they are present in the extracts. The mAb2 peptide is derived from one of the least conserved regions of the myc family, being present and conserved primarily in the c-myc proteins. There is some potential cross reactivity with the N- and B- myc proteins, although neither of these proteins has been detected with the mAb2 antibody (G Evan, personal communication). Thus the results are most consistent with c-myc protein being the immunoreactive species in MAC II, although other possibilities can not at present be completely eliminated.

The distribution of MAC II activity in extracts from different cell lines seems consistent with the presence of c-myc in MAC II, although there are some potential inconsistencies in the data. MAC II has been detected in extracts from derivatives of three cell lines, Rat-1, NIH-3T3 and HeLa, and the intensity of the MAC II band is proportional to the amount of c-myc
protein expressed in each of these cell lines or their derivatives. Extracts of HeLa cells, which express more c-myc protein than rodent fibroblasts (Moore, et al, 1987), contain more MAC II binding activity than the fibroblast extracts. Extracts from a Rat-1 cell line expressing very high levels of exogenous human c-myc protein also contain elevated levels of activity. In contrast, the extracts of v-myc expressing Rat-1 cell lines contain lower amounts of MAC II binding activity than the other Rat-1 cell lines, while the extracts from NIH-3T3 cells which express v-myc contain a similar level of binding activity as the control NIH-3T3 extracts (Figure V.2). Since the c-myc gene is down-regulated in the Rat-1, but not the NIH-3T3, cells which express v-myc, the MAC II activity would appear to correlate with c-myc, but not v-myc levels. If c-myc protein is indeed present in the complex, this suggests that v-myc protein is excluded from it.

There are a number of possible explanations for such a phenomenon. The simplest would be that v-myc protein is incapable of entering into the complex which forms on the DNA. Alternatively v-myc protein may be present in the complex in solution, but fails to resolve on the polyacrylamide gel. This is a distinct possibility since, as the v-myc protein is substantially larger than the c-myc protein and contains a large region of retroviral gag protein. It therefore may behave differently in native gel systems. Alternatively, a v-myc complex may be less stable than a c-myc containing complex under the gel electrophoretic conditions, and thus be undetectable. Consistent with this interpretation, it has been repeatedly observed that prolonged electrophoresis (four hours instead of two) of the MAC II complex results in a much weaker and more diffuse signal, while that of complex B is unaffected (data not shown). Additionally, a precedent exists in the example of the TATA box binding protein, TFIID, for a species dependence in the ability to resolve DNA:protein complexes.
using gel-retardation assays. Mammalian TFIID isolated from cell extracts can be detected by DNAase I footprinting assays but not gel retardation assays. Yeast TFIID, which can substitute for the mammalian protein in some functional assays, is detectable in gel retardation assays using yeast protein substituted into mammalian extracts (Buratowski, et al, 1989, Horikoshi, et al, 1989).

MAC II mobility from extracts of Rat-1 cells expressing several different size mutants of the human c-myc protein is unaltered from that of cells expressing wild type c-myc protein. For example, the MAC II mobility in extracts from Rat-1 cells expressing the Δ145-262 human c-myc protein deletion mutant, which migrates as a 42 kd species on denaturing polyacrylamide gels as compared with the 62-65 kd species seen for the wild type protein (Penn, et al, 1990b), is the same as that seen for extracts containing wild type c-myc protein. Such a lack of a relationship between the size of the myc protein present in an extract and the mobility of MAC II would suggest that c-myc protein may not be present in the complex. Although the mobility of protein complexes under non-denaturing conditions is determined by a combination of the charge, molecular weight and three dimensional structure of the proteins, there is a general tendency for related DNA:protein complexes to migrate in proportion to their sizes (Hope and Struhl, 1986). However, the mobility of MAC II is quite low, suggesting that the protein component(s) may be large, and that such size differences in the c-myc protein may not significantly alter the mobility of the complex. In support of this interpretation, the mobility shift observed in the presence of mAb2, imparted by an antibody of approximately 150 kd, is only approximately 2 mm in a gel in which MAC II had migrated 16 mm and free probe 130 mm (a ~12% decrease in mobility). Thus the data concerning a myc protein size-dependency of MAC II are inconclusive.
In conclusion, there is currently no formal proof of the direct presence of c-myc protein in MAC II, and many of the experiments which address this issue can be subjected to contradictory interpretations. Further experimentation is therefore required in order to answer this important question.

5.3.2 Nature of the MAC II binding structure

The DNA to which MAC II binds appears to contain an unusual property which is distinct from its primary structure. This is most evident from the observation that multiple fragment preparations, even if derived from the same plasmid, do not all possess an equivalent capacity to be involved in MAC II complex formation. Furthermore, the binding capacity is stable for long periods when the linear fragment is stored at low temperatures, yet is completely lost following brief incubation at temperatures approximately 8°C below the melting temperature of the DNA fragment.

The possibility that this binding capacity is due to the positive action of an additional molecule interacting with the DNA seems unlikely due to several observations. The mobility of probe which can bind MAC II is indistinguishable from that of probe which cannot bind MAC II on both native polyacrylamide or agarose gels. DNA fragments which differ in length by three bp have been resolved on these gels, implying that if there is an associated molecule, it alters the mobility of the DNA fragment by significantly less than three bp does. Furthermore, the MAC II capacity can be retained following recovery of probe purified by gel electrophoresis. It has also not been possible to dissociate the binding activity from the probe by a number of treatments which would be expected to remove most molecules which would interact electrostatically with the DNA. This would appear to eliminate most protein:DNA interactions. The instability of the binding capacity at elevated temperatures suggests that the binding
capacity is not due to covalent modification of the DNA, and the constant
temperature differential between the MAC II binding transition point and
the probe fragment melting point suggest that the MAC II binding capacity
is stabilised by the same kind of interaction which stabilises the double
stranded DNA, ie hydrogen bonding. Lastly, addition of several possible
contaminating molecules has not been found to increase the affinity of
MAC II for the probe. MAC II binding capacity may thus be due to an
intrinsic, metastable property of the probe DNA.

Hints as to the nature of this property come from the origins of DNA
fragments which have MAC II binding capacity. The only probes which
have been found capable of binding MAC II are all fragments isolated from
plasmid DNA. The most consistent fragment preparations with respect to
MAC II binding have been derived from the same plasmid preparations,
although there is additional variability in the MAC II binding ability
amongst fragment preparations derived from the same plasmid
preparation. Furthermore, synthetic oligonucleotides, whether made
double stranded by annealing or by primer extension from one template
strand, were incapable of binding MAC II. These data suggest that the
DNA assumes a stable MAC II binding structure while present in the
plasmid, and that such a structure is maintained in the linear fragment.
However, it may not be possible to impart this structure directly to linear
DNA.

DNA has been documented to enter many different non-B DNA
conformations. These include right hand A or left hand Z forms, and
unusual structures such as purine-pyrimidine rich triple strand H DNA,
cruciforms, stem-loops, slipped base pairings, anisomorphic and bent
instances these structures are stabilised in plasmids by the presence of
negative supercoiling, and are only transiently present in linear DNA.
The requirement that the MAC II binding structure be very stable in linear DNA appears to rule out the possibility of Z DNA, cruciforms, stem-loop structures or anisomorphic DNA (Rich, et al, 1984, Wells, et al, 1988). Structures which have been reported to be moderately stable in linear DNA include slipped base pairs (Hentschel, 1982), triple strand DNA (Sklenár and Feigon, 1990) and some S1 nuclease hypersensitive sites (Wells, et al, 1988). A slipped base pair structure appears to be inconsistent with the MAC II binding site, as MAC II binding is maintained at NaCl concentrations in excess of 200 mM, while slipped base pairs are undetectable at NaCl concentrations of 120 mM NaCl or greater (Hentschel, 1982). Triple stranded H-DNA is usually found in polypurine-polypyrimidine tracts of supercoiled plasmids, and often requires an acidic pH. Formation of this structure also requires an inverted repeat sequence, in order that one DNA strand can fold back into the major groove of the double stranded component of the other half of the repeat, and base pair with its complementary sequence (Htun and Dahlberg, 1989). There are also reports of such structures forming at neutral pH using linear oligonucleotides (Sklenár and Feigon, 1990), and it has been suggested that a triple stranded region can form between a single stranded oligonucleotide and double stranded DNA upstream of the human c-myc P1 promoter (Cooney, et al, 1988). Such triple stranded structures have always been reported to result in the introduction of nuclease hypersensitive sites at the position of the bend in the DNA, where single stranded DNA becomes exposed (Htun and Dahlberg, 1989, Wells, et al, 1988). While there are polypurine stretches within the MAC II binding sequences, they do not contain inverted repeats, and no nuclease sensitivity which correlates with the ability to bind MAC II has been found, suggesting that this type of structure is not responsible for MAC II binding. Bends in double strand DNA are often an intrinsic
property of specific DNA sequences (Travers, 1989) and can probably be eliminated from consideration, as their stability and formation should be independent of transient heat treatment or the source of the DNA fragment, as well as requiring sequence motifs which are not found within the MAC II binding probe. It thus appears that there are no obvious explanations of the MAC II binding property of the probe based on previously described DNA structures.

While it seems that the MAC II binding structure is distinct from many other known DNA conformations, there is a recent report on differences in the kinetics of interaction between BamHI restriction endonuclease and two supercoiled pBR322 plasmid DNA preparations, in which the behaviour of one of the plasmid preparations parallels that of the MAC II competent probe. The conclusions of this paper are striking enough to merit partial reproduction here:

...It is evident that the two different preparations of plasmid DNA are cleaved differently. This, in spite of the fact that the assay conditions were identical, the bacterial strain that the plasmid was grown in was identical, and the growth and plasmid preparation protocols were identical. ... When the same preparation of DNA was prewarmed [to 65°C] prior to assay, the rate constants returned to the value of the first preparation. Hence, although identical in sequence to the plasmid in preparation 1 and prepared in an identical way, the plasmid in preparation 2 was in a different (but apparently kinetically stable) conformational state. This study suggests that kinetically stable topological differences may be introduced in the handling of the DNA (Nardone, et al, 1990).

Thus, although this study was performed on supercoiled plasmid DNA, the variations in the behaviour of the plasmids appears very similar to those seen with DNA probes which have MAC II binding capacity, and their nature remains unexplained.

Several previous studies have examined the c-myc P2 promoter proximal sequences for evidence of unusual DNA structures both in vivo and in vitro. These studies have relied on nuclease to detect sequences which might be present in an altered structure which would thus be either
hypersensitive or particularly resistant to cleavage. DNAase I hypersensitive sites in vivo have been detected on both the human and murine c-myc promoters at positions within approximately 70 bp of the P2 cap site (Mango, et al, 1989, Siebenlist, et al, 1988, Siebenlist, et al, 1984). The sequences to which these sites map include the binding domains of the E2F, Me1a1 and B/β complexes as well as part of the putative MAC II binding domain. As these analyses are performed in the presence of cellular factors, it is not possible to distinguish between structures which may preexist before factor binding and structures which are induced upon association of proteins with the DNA. An in vitro study of S1 nuclease sensitive sequences using supercoiled plasmids which include regions of the human c-myc gene have not identified any regions of hypersensitivity near the P2 promoter (Boles and Hogan, 1987). A site was identified upstream of the P1 promoter which is dependent on supercoiling (Boles and Hogan, 1987), and may exist as a triple stranded structure (Boles and Hogan, 1987, Cooney, et al, 1988). The formation of this structure was found to be dependent on the inclusion of sequences several kb away, presumably due to an effect they have on the manner in which the energy of supercoiling is distributed (Boles and Hogan, 1987). It is possible that a MAC II binding, nuclease hypersensitive, structure might similarly be present in plasmid DNA under restricted sequence contexts, even though such a site is not detectable in the linear DNA fragment although it is difficult to imagine what structure might be retained in the probe fragment.

5.3.3 MAC II binding site

Presently the minimal MAC II binding site requires the human c-myc gene sequences from -63 to +49 relative to the P2 promoter mRNA cap site (see Table V.1). This assignment remains preliminary, as it is clear that a property in addition to primary DNA sequence is required for MAC II
binding, and it has not yet been determined how to reproducibly impart this property to the DNA probes. Until this has been determined, negative binding and competition results must be treated with caution.

Competition experiments have revealed that specificity for a primary sequence independent of any other properties of the DNA must exist, as full length -66/+49 fragments which are incapable of binding MAC II when used as probes efficiently compete complex formation. In addition some smaller fragments which are incapable of binding MAC II can compete its formation. For example, the -13 to +49 and -66 to +22 fragments both fall into this category. It is not clear whether they compete the sequence or non-sequence recognition properties of the MAC II proteins, as these fragments were not heat treated to remove any potential non-sequence binding component prior to the competition assays. The inability of the lsl oligonucleotide to compete MAC II implies that the sequence requirements contain elements outside of the -71 to -22 region, as this competitor presumably lacks the putative structural portion of the binding site. Further definition of the sequence dependent component of the binding site should be possible using finer heat inactivated deletion and point mutated constructs.

5.3.4 Relationship to negative regulation of myc transcription

In Chapter 3 it was found that both the -71 to -22 and -51 to +49 human c-myc P2 promoter regions were sufficient to impart responsiveness to myc protein levels to a heterologous indicator. In this Chapter a complex has been described which may contain c-myc protein and will form on human c-myc P2 promoter sequences which extend from -63 to +49. These sequences contain extensive regions of overlap, suggesting that there may be a role for the MAC II complex in regulating transcription from the myc promoter. However, as a more rigorous genetic relationship between the cis acting regulatory elements and the MAC II binding site is not yet
available, the data are consistent with too many interpretations to either confirm or exclude a role for MAC II in the regulation of myc transcription. Further refinement of their relationship will thus require higher resolution determinations of both the myc responsive elements within the promoter and the MAC II DNA binding site.
Figure V.1: Supershifting Rat-1 complexes with anti-myc antibodies

Gel retardation assays using 10 μg Rat-1 control extracts, 1.5 μg poly d(IC)•poly d(IC) and the -66/+49 human c-myc P2 promoter fragment as probe, incubated with monoclonal antibodies. Extracts were incubated with antibody before (~10 min) or after (post) combining with probe. The indicated antibodies were preincubated with 1 μg of peptide (+pep) before adding to the complex formation mix.

IgG preparations from ascitic fluid of two αmyc monoclonal antibodies are indicated (mAb1, mAb2). Cognate peptide for mAb1 is pep1, and for mAb2 is pep2 (see text). An αLFA-1 antibody ascitic IgG preparation was also used (control mAb).

Complex A from previous figures (eg IV.11) has been renamed MACII. Complex B is also indicated (B).

The positions on the human c-myc protein of the antigenic peptides are indicated at the bottom of the figure.
Figure V.2: Cell type distribution and response to myc levels of MAC II

Extracts from a variety of cell populations expressing different amounts and mutants of c-myc protein were used in assays with the -66/+49 human c-myc fragment probe. Extracts were prepared from proliferating Rat-1, NIH-3T3 and HeLa cells. Pools of Rat-1 cells expressing v-myc, wild type human c-myc protein and human c-myc protein mutants Δ41-53 and Δ145-262 and NIH-3T3 cells expressing v-myc were also used for extract preparation (Penn, et al, 1990b).

10 μg of each extract and 1.5 μg poly d(IC)•poly d(IC) were used in each assay.

The major complexes (MACII, B) are indicated.

F: Free probe
<table>
<thead>
<tr>
<th>cell type:</th>
<th>Rat-1</th>
<th>NIH-3T3</th>
<th>HeLa</th>
</tr>
</thead>
<tbody>
<tr>
<td>exogenous myc:</td>
<td>control $\Delta 41-53$ $\Delta 145-262$ WT v-myc</td>
<td>control v-myc</td>
<td>control</td>
</tr>
</tbody>
</table>

![Image of gel electrophoresis](image-url)
Figure V.3: Supershifting HeLa complexes with anti-myc antibodies

HeLa cell extract was used in gel retardation assays containing anti-myc monoclonal antibodies. Each reaction contained 5 µg extract, 1.5 µg poly d(IC)•poly d(IC) and -66/+49 human c-myc P2 promoter fragment probe.

Extracts were incubated with antibody before (~10 min) combining with probe. The indicated antibody preparations were preincubated with 1 µg of peptide (+pep) before adding to gel retardation mix.

IgG preparations from ascites fluids of two αmyc monoclonal antibodies are indicated (mAb1, mAb2). Cognate peptide for mAb1 is pep1, and for mAb2 is pep2.

The major complexes (MACII, B) are indicated. Gel retardations with Rat-1 control extracts are included for reference.

The positions of the antigenic peptides on the human c-myc protein are indicated at the bottom of the figure.

F: Free probe
Figure V.4: Variations in MAC II binding capacity of different probe preparations

Multiple preparations of human c-myc P2 promoter -66/+49 probe were compared for their capacity to bind MAC II. Binding activity of four independent preparations of the BamHI/PstI fragment isolated from pUC-66/+49IFN (+97/+208 relative to P1) or pUCSVter-66/+49IFN are shown (1st, 2nd, 3rd and 4th prep). Synthetic oligonucleotides which contain the exact same sequence as the restriction fragment were annealed and used as probe (oligos).

Lanes 1-7 show binding capacity of preps 1 to 3 following repurification of the labelled probe variously by PAGE (gel purified) or ion exchange chromatography (NACS purified).

Lanes 8-13 show the binding capacity of a mixture of two labelled preparations, 4 and 2, combined in varying proportions (Mixing). Both probes were labelled to the same specific activity.

Lanes 14 and 15 show the binding capacity of annealed, labelled oligonucleotides before and after gel purification of the annealed DNA. The oligonucleotides had been purified by PAGE prior to labelling.

Lane 16 shows the binding capacity of an antisense strand oligonucleotide annealed to an end labelled 12 nt primer which was then extended by Sequenase 2.0 modified T7 DNA polymerase.

Reactions were performed with 5 μg extract, 1.5 μg poly d(IC)•poly d(IC) and the indicated probe; lanes 1-7, 14-16 used Rat-1 LJ extract, lanes 8-13 HeLa extract.

The positions of the major complexes (MAC II, B) are indicated; F: Free probe.
Figure V.5: Loss of binding upon melting MAC II probes

The ability of plasmid derived -66/+49 probe to bind MAC II was assessed following heating of the probe to 95°C for 2 min and then cooling to 0°C immediately (95°C) or over several hours (reannealed). Binding reactions contained 10 μg Rat-1 control extract and 1.5 μg poly d(IC)*poly d(IC).

The positions of the major complexes (MAC II, B) are indicated.

The two sets of 95°C and reannealed lanes are exposures of different durations of the same gel.

F I: Free probe, single strand.

F II: Free probe, double strand.
Figure V.6: Thermal stability of MAC II binding capacity of probes

Plasmid derived -66/+49 probe was heated to the indicated temperature for 5 min and immediately cooled on ice prior to use in gel retardation assays.

Probes were maintained in the buffered salt solution indicated (TE or TE + 50 mM NaCl). Binding reactions contained 5 µg HeLa extract and 1.5 µg poly d(IC)•poly d(IC) and were performed at the same final salt concentration.

The positions of the major complexes (MAC II, B) are indicated.

F: Free probe.
buffer:  

<table>
<thead>
<tr>
<th>TE</th>
<th>TE + 50 mM NaCl</th>
</tr>
</thead>
</table>

temp:  

<table>
<thead>
<tr>
<th>4°</th>
<th>45°</th>
<th>49°</th>
<th>53°</th>
<th>57°</th>
<th>61°</th>
<th>65°</th>
<th>69°</th>
<th>72°</th>
<th>75°</th>
</tr>
</thead>
</table>

MAC II  }

B  }

F  
Figure V.7: Deletion analysis of MAC II binding (1)

Various fragments of the human c-myc P2 promoter were used in gel retardation assays. MAC II binding capability of each probe was ascertained by the ability of mAb1 to specifically supershift a complex of the appropriate mobility. The probe used and whether the reaction contained mAb1 or mAb1 and pep1 are indicated.

The left hand panel shows a 5' deletion series, with 5' endpoints of the probe fragments at positions -66, -63 and -22. These three probes have 3' ends at +49. The right hand panel shows a 3' deletion series, with 3' endpoints at positions +49, +22, -21 and -49 and a shared 5' end of -66.

The positions of the major complexes (MAC II, B, β) are indicated.

Reactions were performed with 10 μg Rat-1 control extract, 1.5 μg poly d(IC)•poly d(IC) and the indicated probe.

F: Free probe.
Figure V.8: Deletion analysis of MAC II binding (2)

Various fragments of the human c-myc P2 promoter were used in gel retardation assays. MAC II binding capability of each probe was ascertained by the ability of mAb1 to specifically supershift a complex of the appropriate mobility. The probe used and the whether reaction contained mAb1 or mAb1 and pep1 are indicated.

The panel shows deletions with 5' endpoints at positions -66, -51, -44 and -13 and a shared 3' endpoint at +49 and a 3' deletion of -66/+22. The upper and lower panels show short and long exposures of the same gel.

The positions of the major complexes (MAC II, B) are indicated.

Reactions were performed with 5 μg Rat-1 control extract, 1.5 μg poly d(IC)•poly d(IC) and the indicated probe.

F: Free probe.
Figure V.9: Competition mapping of MAC II binding

Unlabelled DNA probes of various portions of the human c-myc P2 promoter region from spanning the sequences -71:+49 were used as competitors for binding of labelled -66/+49 probe. Competitors include fragment preparations containing the sequences -66/+49 (A), -66/+22(B) and -13/+49(D). The -71/-22(C) competitor was prepared from annealed 1s1 oligonucleotides.

Each reaction contained 10 µg Rat-1 extract, 1.5 µg poly d(IC)•poly d(IC) and labelled -66/+49 probe. Molar excesses of each unlabelled competitor are indicated. The major complexes (MACII, B) are indicated, as are schematic diagrams of the competitors, with positions indicated relative to the P2 mRNA cap site.

F: Free probe.
Figure V.10: Thermal stability of MAC II in extracts with or without probe

Cell extracts (Extract incubation) or preformed complexes, containing probe as well as extract, (Complex incubation) were heated to the indicated temperature for 5 min followed by immediate cooling on ice.

10 μg HeLa extract, 1.5 μg poly d(IC)•poly d(IC) and -66/+49 plasmid derived probe were used in all reactions.

The positions of the major complexes (MAC II, B/β) are indicated.

F: Free probe.
Extract treatment: incubation

MAC II ►

B/β ►

F
Table V.1: MAC II binding and competition: Summary

The sequence of all probes used in gel retardation, supershifting and competition experiments with MAC II is presented. Capital letters indicate sequences which are identical with wild type human c-myc, lower case those which differ.

The Shift column indicates whether a supershift was detected when using mAb1 in the binding reaction. Comp indicates the ability of a given sequence to compete MAC II bound to -66/+49 probe.

+: a positive signal
[-]: probably negative but with a possibility of binding MAC II
--: negative

A blank at any position indicates the experiment was not performed.

WTMYC: wild type human c-myc sequence
WTCON: Minimal MAC II binding site from the supershiftable complex data is underlined

The binding sites of the E2F and 'Me1a1' complexes as well as the human c-myc P2 TATA box and mRNA cap site are indicated.
### MAC II

#### DOUBLE STRAND BINDING AND COMPETITION

<table>
<thead>
<tr>
<th></th>
<th>Shift</th>
<th>Compete</th>
</tr>
</thead>
<tbody>
<tr>
<td>wtwc</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-66/+49 S'...qate cggGccGGg aAAGACGGG AGGGGCCGAT CCGCCTGACG ATAAAAGGGG GGTTCGCGG CTTTATCTAA CGCGCTCTAG TAGTCCGCG CAGAGGCCA GGGAGCGCC GGGGGGCGC CGTATG</td>
<td>3'</td>
<td>+</td>
</tr>
<tr>
<td>-63/+49 S'...q atccGccGGg aAAGACGGG AGGGGCCGAT CCGCCTGACG ATAAAAGGGG GGTTCGCGG CTTTATCTAA CGCGCTCTAG TAGTCCGCG CAGAGGCCA GGGAGCGCC GGGGGGCGC CGTATG</td>
<td>3'</td>
<td>+</td>
</tr>
<tr>
<td>-51/+49 S'...qate cggGccGGg aAAGACGGG AGGGGCCGAT CCGCCTGACG ATAAAAGGGG GGTTCGCGG CTTTATCTAA CGCGCTCTAG TAGTCCGCG CAGAGGCCA GGGAGCGCC GGGGGGCGC CGTATG</td>
<td>3'</td>
<td>[-]</td>
</tr>
<tr>
<td>-44/+49 S'...qate cggGccGGg aAAGACGGG AGGGGCCGAT CCGCCTGACG ATAAAAGGGG GGTTCGCGG CTTTATCTAA CGCGCTCTAG TAGTCCGCG CAGAGGCCA GGGAGCGCC GGGGGGCGC CGTATG</td>
<td>3'</td>
<td>[-]</td>
</tr>
<tr>
<td>-21/+49 S'...qate cggGccGGg aAAGACGGG AGGGGCCGAT CCGCCTGACG ATAAAAGGGG GGTTCGCGG CTTTATCTAA CGCGCTCTAG TAGTCCGCG CAGAGGCCA GGGAGCGCC GGGGGGCGC CGTATG</td>
<td>3'</td>
<td>[-]</td>
</tr>
<tr>
<td>-13/+49 S'...qate cggGccGGg aAAGACGGG AGGGGCCGAT CCGCCTGACG ATAAAAGGGG GGTTCGCGG CTTTATCTAA CGCGCTCTAG TAGTCCGCG CAGAGGCCA GGGAGCGCC GGGGGGCGC CGTATG</td>
<td>3'</td>
<td>[-]</td>
</tr>
<tr>
<td>-66/-42 S'...qate cggGccGGg aAAGACGGG AGGGGCCGAT CCGCCTGACG ATAAAAGGGG GGTTCGCGG CTTTATCTAA CGCGCTCTAG TAGTCCGCG CAGAGGCCA GGGAGCGCC GGGGGGCGC CGTATG</td>
<td>3'</td>
<td>-</td>
</tr>
<tr>
<td>-64/-43 S'...qate cggGccGGg aAAGACGGG AGGGGCCGAT CCGCCTGACG ATAAAAGGGG GGTTCGCGG CTTTATCTAA CGCGCTCTAG TAGTCCGCG CAGAGGCCA GGGAGCGCC GGGGGGCGC CGTATG</td>
<td>3'</td>
<td>-</td>
</tr>
<tr>
<td>wtcw</td>
<td></td>
<td></td>
</tr>
<tr>
<td>S'...qate cggGccGGg aAAGACGGG AGGGGCCGAT CCGCCTGACG ATAAAAGGGG GGTTCGCGG CTTTATCTAA CGCGCTCTAG TAGTCCGCG CAGAGGCCA GGGAGCGCC GGGGGGCGC CGTATG</td>
<td>3'</td>
<td>-</td>
</tr>
</tbody>
</table>
Chapter 6 The MAC I complex
6.1 Introduction

The unusual properties of the probes which bind MAC II, raise the possibility that some form of secondary structure, including perhaps a single strand component, is required for binding activity. Denatured -66/+49 probe detects a band with a mobility intermediate between the B complex and MAC II, as well as a smear of bands which migrate with mobilities ranging from slightly faster than the B complex through to the position of the free probe (See Figure V.8). The mobilities of MAC II and the uppermost of these complexes are similar, prompting the question of whether this complex was a single strand binding equivalent of MAC II.

Data presented in this chapter demonstrate that this complex, now termed MAC I, also contains a myc-related antigen. An initial characterisation of its binding properties, cell type distribution, relationship to myc protein levels and stoichiometry on DNA are presented.

6.2 Results

6.2.1 Effect of antibodies

The mAb1 monoclonal antibody, alone or blocked with its cognate peptide, was added to a gel retardation assay using denatured -66/+49 probe and Rat-1 extracts (Figure VI.1). Addition of the antibody results in the quantitative supershifting of the MAC I complex. This activity of the antibody is inhibited by preincubation with the antigenic peptide. Repeated attempts at supershifting this complex with mAb2 have invariably been unsuccessful (data not shown). MAC I thus shares at least one antigenic epitope with the myc protein, which maps in human c-myc to amino acids 41-55, and is within one of the most highly conserved sequences amongst the myc family of proteins. While there is one common antigenic determinant detectable by mAb1 in the two complexes,
the inability to supershift MAC I with mAb2 suggests that the complexes have different structures.

6.2.2 Strand specificity

Synthetic oligonucleotides corresponding to the two strands of the -66/+49 fragment were labelled, gel purified, and used in binding reactions (Figure VI.1). Only the antisense strand probe binds MAC I with high affinity, while the sense probe detects only a weak smear. This result demonstrates that MAC I binds a single strand component of the probe preparation, and that it displays some sequence specificity.

Antisense, sense, denatured or double strand unlabelled -66/+49 DNA were used as competitors against labelled antisense probe (Figure VI.2). The antisense strand competes itself very efficiently, almost completely blocking complex formation at 5 fold molar excess, and the denatured double strand DNA competes almost as well. Sense strand competitor blocks binding marginally at 5 fold, and moderately well at 50 fold, excess. Double strand DNA, in contrast, does not compete at all. When the antisense strand is present in the competition reactions, a nonspecific band which migrates ahead of MAC I increases in intensity with increasing competitor concentrations, a result which is reproducibly seen. In the sense strand competition reactions, however, the intensity of this band is generally unaffected or marginally reduced. In addition, free probe which migrates with the mobility of double strand DNA appears in proportion to the concentration of sense strand competitor. This suggests that the sense strand is not directly competing for MAC I binding, but is rather annealing to the probe, preventing MAC I complex from forming on the antisense DNA. The denatured competitor is probably exhibiting both actual competition through its antisense strand component and also annealing to free probe through its sense strand component, thus displaying both double strand free probe and an increase in the non-specific binding
complex. The inability of the double strand DNA to compete MAC I binding provides further evidence that MAC I and MAC II are distinct complexes.

6.2.3 Antibody titration

The amount of mAb1 required to produce a detectable MAC I supershift was determined by titrating a 30 fold range of antibody concentrations against a constant amount of extract (Figure VI.3). The magnitude of the supershift increases in proportion to the amount of antibody added from 0.1 to 2 μl. At 2 μl the supershifted band becomes a doublet, with the higher mobility component the less intense of the two. This ratio becomes inverted with 3 μl Ab, where the higher mobility signal is very strong, and the lower mobility band is notably weaker. The sum of the intensities of these two bands is much greater than that seen without any antibody added, suggesting that the antibody may facilitate complex formation or enhance its stability. All of the complex is not shifted at every antibody concentration: a greater proportion of complex becomes shifted with increasing antibody concentration, until it is quantitatively shifted by 1 μl antibody.

The finding that different amounts of antibody induce differing magnitudes of mobility shift has interesting implications for the valency of the complex. If each complex has only one epitope, one would expect two positions for a mobility shift: one position would reflect one complex bound to one antibody, and the other, slower migrating form would reflect two complexes bound by one bivalent antibody. At limiting antibody concentrations the higher mobility form should predominate, with a shift to low mobility complex at equivalent Ab/Ag concentrations, culminating in a return to higher mobility complex when in antibody excess. More than two supershifted positions should not be detectable under monovalent antigen conditions. However at least four or five different
supershifted mobilities are seen in the titration experiment, suggesting that each MAC I complex contains at least two and possibly more mAb1 epitopes. The variations in mobility may be due either to the formation of Ab:Ag lattices of different sizes, or to the binding of multiple antibodies to each individual protein:DNA complex.

6.2.4 Effects of different carriers

While it is possible to resolve MAC I from the high mobility background smear, it would be preferable if these non-specific bands could be completely eliminated. Towards this end, the ability to detect MAC I binding activity using different non-specific carriers, while reducing unwanted signal, was tested (Figure VI.4). Poly d(IC)•poly d(IC) and single strand Bluescript plasmid DNA both allow detection of supershiftable complex, while double strand salmon sperm DNA does not. Additional bands, which have proved to be non-specific, as well as the smear, are detected with the single strand Bluescript carrier. A variety of other synthetic single strand polymers, (including poly dA, poly dT, poly dI, and poly dC) or DNA preparations have been used as carriers, but none has proved as efficacious as poly d(IC)•poly d(IC) (data not shown). All experiments have thus been performed using poly d(IC)•poly d(IC) carrier. Interestingly, the blockage of MAC I activity by double strand salmon sperm DNA is a property shared by MAC II.

6.2.5 Relationship to myc protein levels and cell type specificity

The presence of a myc antigenic site in MAC I raises the obvious question of whether myc protein is the antigenic component. The same possible correlations between MAC I activity and myc protein levels exist for this complex as were described for the MAC II complex (Chapter V). Extracts from Rat-1 cells expressing different myc constructs were therefore assayed for both MAC I levels and their ability to be supershifted by mAb1 (Figure VI.5). Rat-1 cells which express the human c-myc protein deletion
mutants Δ41-53 (deleted for the mAb1 epitope) and Δ145-262 (deleted for
the mAb2 epitope), as well as v-myc and wild type human c-myc, should
have down-regulated expression of the endogenous c-myc gene (Penn, et
al, 1990b). The total amount of myc protein in each of these extracts
should be greater than that present in the control extracts. All of these
extracts display a binding activity which has a similar mobility and
quantity to the MAC I activity of the control extract. This suggests that if
myc protein is present in MAC I, it is not the rate limiting component for
complex formation.

mAb1 also supershifts an equivalent proportion of the complex, and by
the same extent, in each extract. Since the Δ41-53 protein is deleted for the
mAb1 epitope, one might expect the antibody to be ineffectual if MAC I
contains c-myc protein. The maintenance of a supershiftable phenotype in
the Δ41-53 extract suggests that either myc is not present in the complex, or
each complex contains both mutant and wild type protein. RNAase
protection analysis of RNA isolated from the Δ41-53 cell pool in parallel
with the extract preparation reveals that while the endogenous c-myc gene
expression is reduced in these cells, it is not eliminated (Figure VI.4).
Endogenous c-myc protein is thus likely to also be present in significant
quantities. Since there is a reasonable probability that each complex
contains multiple antigenic proteins, (see above), it might be expected that
a large proportion of the complexes formed from this extract would be
supershiftable. This experiment must thus be considered inconclusive.

In addition to the Rat-1 cell lines, NIH-3T3 cells, which do not
autoregulate their endogenous c-myc expression, and HeLa cells were
assayed for MAC I binding activity. Both control and v-myc expressing
NIH-3T3 cells contain similar amounts of a complex which comigrates
with MAC I and is also supershiftable. While less complex is seen in the
NIH-3T3 than the Rat-1 extracts, it is difficult to normalise results from
different cell types, and thus assess their significance. The HeLa extracts, in contrast, do not appear to contain significant quantities of MAC I binding activity. This result is striking, in light of the strong MAC II signal seen with HeLa extracts, and argues quite forcefully against MAC I and MAC II containing identical protein constituents. Since HeLa extracts should contain myc protein, then either additional proteins, or a specific modification(s) to the myc protein must be required to form MAC I if it does contain myc protein.

6.2.6 Competition mapping of binding site

An extensive series of single strand oligonucleotides and denatured double strand fragments were used in competition assays (0, 5 and 50 fold molar excess) in an attempt to map a minimal sequence required to compete, and hopefully bind, MAC I. Data for antisense or denatured competitors are presented in Figure VI.6; sense strand and unrelated competitions in Figure VI.7. The data are summarised in Table VI.1. All antisense sequences are referred to using the 5' to 3' coordinates of their complementary sense strand. Thus "-66/+49 antisense" is actually 5':+49/-66:3'. Upstream refers to positions closer to P1, while downstream refers to positions closer to the poly A site.

Deletions with upstream endpoints at -66, -44 and -13 and a downstream endpoint at +49 all compete successfully for MAC I binding, though the -13 fragment competes noticeably less well, displaying little competition at 5 fold excess. Deletion to position +22 from the downstream end (-66/+22) does not impair competitive function, while truncation to -22 (-71/-22: ls1) completely abolishes competition. Ls2 and Ls3 antisense linker scanning oligonucleotides (see Figure III. 26), have also been used, and are equally deficient in competitive function (data not shown). The -22/+22 oligonucleotide, defined as a potential minimal binding element by these downstream deletions, is also an efficient

333
competitor. Division of this 45 nt segment into -22/-1 and +1/+22 halves eliminates all competitive ability. Two segments predominantly outside of the -22/+22 region (-39/-19, which contains a point mutation in the TATAA box, and a 12-mer which overlaps the myc sequence from +44/+49) fail to compete MAC I binding. Antisense oligonucleotide competitors have thus defined a 45 nt sequence centered around the P2 cap site, which is capable of competing MAC I, neither half of which competes alone. The -13/+49 deletion displays partial competition activity, and may thus move the upstream boundary to -13, though this has yet to be demonstrated.

Sense strand competitions have been tested with far fewer sequences (Figure VI.7). Full length -66/+49 competes, as was shown previously. The Is1 sense strand (-71/-22) and the -21/-1 sequence both display minimal competitive function, and only at 50 fold excess. Both of these sequences compete better than the +1/+22 sense strand, which does not noticeably compete. Two oligonucleotides, 7070 and 7074, which contain sequences from the human c-myc third exon, show a small degree of competition at 50 fold excess. The competitive abilities of the unrelated and -21/-1 sequences are similar and marginal, and suggest that the -21/-1 competition may be non-specific. The weak level of these sense strand competitions suggests that if they are specific, they are competing by forming double strand structures which somehow interfere with protein: single strand DNA interactions.

6.2.7 Deletion mapping of binding site

The competition data have defined a potential binding site in the region from -22 to +22 on the antisense strand. A comparable series of experiments to those presented above were done using deletions of the -66/+49 fragment as probes (instead of competitors) in gel retardation assays. Either single strand oligonucleotides or denatured DNA fragments
were used as probes, in the presence or absence of mAb1 to positively identify MAC I. The results of these experiments are summarised in Table VI.1, and representative autoradiograms presented in Figure VI.8. Probes displayed in the table encompass 5' and 3' deletions similar to those used in the competition assays; also indicated are whether a probe bound a complex of the appropriate mobility to be MAC I, and the ability of that complex to be supershifted by mAb1. The reactions displayed in the figure include antibody combined with either blocking peptide 1 or control peptide 2.

The minimal supershiftable upstream deletion maps to position -21 in the -21/+49 probe (probe C Figure VI.8), while sequences downstream have been eliminated through +31 (-44/+31, data not shown). These endpoints have been mapped to within 8 nt on the upstream side, and 9 nt on the other. Whether the intersection of these probes, -21/+31, will bind supershiftable complex has not yet been determined. Significantly, several probes bound complexes which roughly comigrate with MAC I, but which are not supershiftable. For example, the -66/+22 and -22/+22 (probe D in Figure VI.8) probes, which both compete MAC I, bind complexes which are insensitive to mAb1. The low mobility -22/+22 complex is competable both by itself and antisense -66/+49 (data not shown). There is thus a paradoxical relationship between the ability of a sequence to compete MAC I and to bind it in a mobility shift assay.

The affinity of each probe for MAC I appears to correlate with the amount of upstream sequence it contains. For instance -66/+49 displays a much stronger MAC I signal than -21/+49 (Figure VI.8, compare probes B and C). This is a reproducible phenomenon, and has been seen with each of the upstream deletions tested. The relative affinities of the downstream deletions has not been assessed. On extended exposures it is apparent that the sense strand of probe -66/+49 also has a weak affinity for a
supershiftable MAC I complex, which binds marginally less well than the -21/+49 probe (Figure VI.9, probes A and C), though a more refined definition of this binding activity has yet to be attempted. In stark contrast to the variable binding capabilities of double strand probes used for binding MAC II, every preparation of single strand probe of the same sequence has been found to bind with similar efficiency (data not shown).

6.2.8 Supershift competitions

The combination of competition and deletion data demonstrated that probes such as -22/+22 were capable of both competing the MAC I band detected by -66/+49 and detecting a complex which was not supershiftable yet is competeable by -66/+49. This suggests that the unshifted MAC I band may contain a mixture of both the shiftable and unshiftable complexes, all of which have the potential to be competed by the same DNA fragment. It was therefore important to demonstrate that the competition data were valid against antigenic MAC I, as opposed to the non-shiftable component. This issue was addressed by performing competitions against supershifted MAC I with several of the probes used previously. Extracts were combined with mAb1 antibody followed by unlabelled competitor. After a brief incubation radioactive -66/+49 antisense probe was added and complexes resolved as usual (Figure VI.10). Of the sequences tested, all of those which previously blocked unshifted complex formation, i.e. -66/+49, -21/+49, -66/+22 and -21/+22, prevented the formation of a supershifted complex when in 50 fold molar excess. Similarly, the -71/-22 oligonucleotide which did not compete previously, was again unable to prevent supershifted MAC I formation. The combination of these competition data and the demonstration that at sufficiently high Ab concentrations all of the detectable MAC I band can be supershifted (Figure VI.4), imply that the non-shiftable complex seen with the -21/22 antisense probe is not normally detected by the -66/+49 antisense probe. However,
more direct characterisation of these complexes is required to confirm this conclusion.

6.2.9 Thermostability of MAC I in extracts

The dramatic effects of heat treatments on MAC II in extracts raised the issue of how MAC I would behave in response to similar heat treatments. The response of extracts to heat treatment in the presence or absence of probe was addressed (Figure VI.11). In contrast to MAC II, MAC I binding activity is reduced following a 37°C incubation of the extract, and is completely abolished at 45°C. The non-specific binding complexes are much more stable, as the low mobility complex is resistant up to 52°C, and the high mobility complexes of the smear are only marginally affected even by 68°C treatment. There is, in addition, a dramatic alteration in MAC I thermosensitivity when the complex is preformed on the DNA before heating. In the presence of probe DNA all of the complexes, including MAC I, are totally resistant to treatment even at 68°C. Exactly how complex formation leads to thermostabilisation is not yet clear, although it may be significant that the non-specific complexes are stabilised as well as MAC I. The heat response of MAC I is thus distinct from that of MAC II, providing an additional degree of disparity between them.

6.3 Discussion and Conclusions

In this Chapter the single strand DNA binding complex called MAC I, which possesses both sequence- and strand-specificity for binding the antisense strand of the human c-myc P2 promoter, has been described. This complex has been demonstrated to be antigenically related to the myc protein family, and most likely contains multiple antigenic determinants per complex. A DNA sequence on which MAC I can form has been mapped, and includes the sequences just downstream of the human c-myc P2 TATA box.
6.3.1 Does MAC I contain myc protein?

Data exist both in favour of and against the notion that myc protein is present in MAC I. While no interpretation is currently irrefutable, the data seem most consistent with the interpretation that MAC I contains a different member of the myc protein family than c- (or v-) myc. As was seen with the MAC II complex in Chapter 5, the strongest evidence for the presence of myc protein in MAC I is the recognition of the complex by an anti-myc monoclonal antibody (mAb1), resulting in the supershifting of the complex. mAb1 was raised to a peptide which represents the most highly conserved sequence of the myc family of proteins, amino acids 44-55 of the human c-myc protein, APSEDIWKKFEL. This sequence, or a close variant of it, is present in all of the myc family proteins known, and the antibody has been demonstrated to recognise v-, c-, N- and L-myc proteins. mAb1 can cross react with some other cellular proteins in immunoprecipitation assays, although their identities are unknown (G Evan, D Hancock and T Littlewood, personal communication). The mAb2 antibody has been repeatedly found to have no effect on the MAC I complex. This antibody was raised to a peptide which is not well conserved outside of the c- and v-myc proteins, and it has been found to react only with these proteins (Evan, et al, 1985, and G Evan, D Hancock and T Littlewood, personal communication). The inability of this antibody to recognise the MAC I complex, even if combined with the cell extracts before the addition of probe, suggests that its epitope is either obscured even in the absence of the DNA probe, or is not present in the complex. When used against the MAC II complex, mAb 2 has induces a supershift even when added to the reaction following complex formation on the DNA. Taken together, these results imply that while a myc family protein is probably present in MAC I, it is distinct from c-myc, which is itself likely to be present in MAC II.
This interpretation is bolstered by several other data. The amount of MAC I present in extracts of Rat-1 cells expressing a number of different myc protein constructs was found to be invariant. Similarly the cell type specificity of MAC I formation does not appear to correlate with the amount of c-myc protein expressed. For instance, HeLa cells, which express relatively high levels of c-myc protein (Evan, et al, 1988), have practically undetectable levels of MAC I activity, while MAC I is readily detected in extracts from Rat-1 and NIH-3T3 cells. More importantly, MAC I in extracts from Rat-1 cells expressing a human c-myc deletion mutant which does not contain the mAb1 epitope was supershifted as quantitatively as the complex in control extracts. While there are extenuating circumstances which can be used to rationalise all of these results in favor of c- (or v-) myc protein being an integral component of MAC I, it is more straightforward to postulate that a different member of the myc protein family is present in this complex.

The other known members of the family, N-, L-, S- and B-myc all contain sequences which may be recognised by mAb 1. It is not known whether any of these proteins are expressed in the fibroblast cell lines which have been demonstrated to contain MAC I activity. Additional members of the family may exist, and could also be involved in the formation of this complex.

### 6.3.2 Stoichiometry of MAC I

Regardless of which myc protein family member(s) are involved in MAC I formation, the dose response curve of the antibody-dependent supershift has very interesting implications. As discussed in Section 6.2.3, it is apparent that there are multiple mAb1 antigenic epitopes present in MAC I. The differences in supershift mobilities can be caused by the establishment of Ab:Ag networks, in which a lattice of antigenic reactions forms, or it can be due to the interaction of multiple antibodies with a
single complex. The observation that distinct supershifted bands are seen at each antibody concentration argues more for lattice formation than independent Ab:Ag complexes, which would be expected to display a distribution of bands corresponding to the number of antibodies bound per complex. While lattice formation is formally possible with two epitopes, three or more are probably required. This is because the bivalency of the antibody leads to a strong preference for second site occupancy by the same Ab molecule, as cooperative binding will occur due to the essentially unimolecular kinetics following the first binding interaction. Similarly, even if the first antibody molecule interacts with a second protein:DNA complex, a second antibody molecule might be expected to occupy the remaining site on each of the molecules, thus preventing formation of an extended Ab:Ag network. Thus it would appear that three or more epitopes should be present in the complex if a lattice is actually to form. Another observation lending credence to the lattice model is that at the highest antibody concentrations used, the magnitude of the supershift was seen to be reduced. The simplest explanation for this observation is that as antigen concentrations are limiting, excess antibody saturates the available epitopes, leading to the formation of smaller Ab:Ag networks. Taken together these data suggest that the supershifted complexes are actually formed into lattice-like networks, and that each complex thus contains three or more exposed antigenic epitopes.

If the antigenic protein has the same general structure as the myc family members, it would be expected to have both helix-loop-helix and leucine zipper domains at its carboxyl terminus. Each of these domains is theoretically capable of dimerising, and bacterial c-myc protein has been shown to form homotetramers through the interaction of these domains (Dang, et al, 1989a). Thus a general model for the MAC I structure might
include up to four molecules of the antigenic protein interacting with the DNA.

6.3.3 MAC I binding site

The combined use of deletions and competition studies has allowed a partial description of the DNA element(s) involved in MAC I formation (see Table VI.1). Probe deletions have defined an upstream limit to a minimal MAC I binding fragment between positions -22 and -13, and a downstream sequence limit between positions +21 and +31, relative to the human c-myc P2 promoter cap site. Assuming that there is no redundancy in the binding site, this would define a minimal element from positions -22 to +31. Competition experiments have further demonstrated that MAC I binding is sequence specific. Using deletions of the binding sequences, it has been possible to map elements which can prevent MAC I formation. It has been found that all sequences which are capable of binding MAC I are capable of competing for MAC I binding as well. In addition there exists a subset of oligonucleotides which can compete for MAC I binding but are incapable of binding MAC I themselves. Such a result was seen with the -21 to +22 antisense fragment, which not only competes for MAC I binding, but also binds a complex which has a similar mobility to, but is distinct from, MAC I, in that it is not supershifted by mAb1 (Figure VI.10). These results suggest that the minimal MAC I binding site is composed of multiple elements, each of which is insufficient for MAC I binding.

As MAC I forms on single strand DNA, it is conceivable that the binding site may be composed of elements which are distinguished by their secondary structure(s), as well as primary sequences. The affinity of MAC I for different probes correlates with the amount of sequence present upstream of position -22, which may be due to a stabilising effect of the upstream sequences on a secondary structure. MAC I was also found to
have a weak affinity for a probe composed of the sense strand -66 to +49 sequence. As the two strands of the probe do not share the same primary sequence, but should have the potential to assume mirror images of a similar secondary structure, it is possible that the complex is recognising such a structure with low affinity on the sense strand.

An RNA secondary structure prediction program (Devereux, et al, 1984, Freier, et al, 1986, Zuker and Stiegler, 1981) has been used to identify any potential structural motifs which are specific for the MAC I binding sequences. The energetically most favorable predicted secondary structures for several oligonucleotides tested for MAC I binding are presented in Figure VI.12. The most notable feature of these structures is a predicted stem-loop forming just downstream (to the left in the Figure) of the P2 cap site. The sequence 5' CTCGCTG 3' is present as an inverted repeat in the regions +18 to +24 and +2 to +8, separated by nine nt. Another short inverted repeat of three nt is present three and four nt outside of the major repeat. The combined sequences would lead to a stem of 13 to 14 bp containing 10 paired positions, with a nine nt loop and a small bubble in the stem. The sequences required to form this stem-loop are present in all of the probes which have been seen to bind MAC I.

If this structure is indeed recognised by MAC I, it alone can not be sufficient for binding. Indeed, the presence of sequences through position -13 are insufficient for MAC I formation, even though the upstream end of the stem is at -4. Thus it may be that this upstream sequence is recognised by the complex, in addition to the downstream stem-loop.

The major inverted repeat contained in the MAC I recognition sequence is present in the c-myc genes of humans, mice and rats (see Table III.2). In contrast, the feline gene does not contain a potential stem loop in this region. This observation implies that either the stem loop per se is not important for MAC I binding, or that the complex will not form on
the feline gene. It will be interesting to learn whether these structures are involved in MAC I formation.

6.3.4 Other sequence specific single strand binding proteins

While the vast majority of sequence specific DNA binding proteins which have been identified recognise double strand DNA sequences, there are several recent reports describing sequence specific single strand binding proteins. Such proteins have been found to bind sequences which are present upstream of the promoters of several eukaryotic and viral genes. These include the human and rat growth hormone genes, the mouse adipocyte specific adipsin gene, the chicken vitellogenin gene, the rat prolactin gene, and the SV40 early genes (Feavers, et al, 1989, Gaillard, et al, 1988, Lannigan and Notides, 1989, Pan, et al, 1990, Peritz, et al, 1988, Wilkison, et al, 1990). In most of these cases the protein detected displays specificity for one strand of the DNA in preference to its complement, although the adipsin promoter binding protein will bind a sequence which is present on both strands due to the presence of an inverted repeat of the binding site. The proteins themselves are in general poorly characterised, except in the case of the prolactin gene binding protein, in which instance the estrogen receptor has been found to display a higher affinity for a single strand as opposed to a double strand estrogen receptor binding site (Lannigan and Notides, 1989). As these proteins have usually been identified in the course of investigations of the control of RNA polymerase II-dependent transcription initiation, their binding sites usually overlap those of double strand binding proteins within sequence elements defined as important for the regulation of these genes. It is not known whether the DNA sequences which these proteins recognise are single stranded in vivo, and it is thus difficult to assess the significance of the in vitro interactions. In at least two situations, involving the factors which interact with the human growth hormone and murine adipsin
genes, however, the levels of single strand binding activity correlate with changes in expression of each gene, suggesting that they may be functionally relevant (Peritz, et al, 1988, Wilkison, et al, 1990). Furthermore, many studies have identified regions of chromatin which are hypersensitive to nucleases, results which are often interpreted as being indicative of regions of local single strandedness (Gross and Garrard, 1988, for review).

MAC I shares some properties with these other complexes, in that it also has a preference for one strand of the c-myc promoter, and its binding site is located at a position which is presumably involved in interactions with RNA polymerase II, as it spans the mRNA cap site. Since in each of these situations the functional significance of single strand DNA:protein complex formation has not been definitively demonstrated, this remains the most important issue which must be addressed.

6.3.5 Relationship of MAC I and myc autosuppression

As a myc response can be detected using constructs which do not contain the MAC I binding sequences (eg Figure III.29), the down-regulation of c-myc transcription probably does not require MAC I. However, because of the apparent redundancy of the response elements, it is still possible that this complex may be involved in attenuation of transcription from the P2 promoter, as the MAC I binding site spans the mRNA cap site. Thus it is important to further characterise this complex. In particular, definition of the requisite contact points between the protein component of MAC I and the DNA is critical, as such information would allow the function of sequences which are incapable of binding MAC I, due to minimal mutations, to be rigorously assessed.
Figure VI.1: Supershifting single strand binding complex with anti-
myc antibody

The ability of α-myc mAb1 to affect complexes which bind single
strand DNA was assessed using heat denatured (95°C) -66/+49 human
c-myc P2 promoter probe. 10 µg Rat-1 control extract were incubated
with mAb1 (95°C + mAb1) or mAb1 preblocked with its cognate peptide
(95°C + mAb1 + pep1) before adding to the binding reaction. 1.5 µg poly
d(IC)*poly d(IC) were used as carrier in each reaction.

Synthetic oligonucleotides corresponding to the sense (sense strand)
or antisense strand (antisense strand) sequences of the -66/+49 probe
were also used as probes.

A reaction containing double strand -66/+49 fragment probe is
included for reference (-).

The major specific complexes are indicated (MAC II, B and MAC I).

F I: free single strand probe.

F II: free double strand probe.
Figure VI.2: Self competition of MAC I binding: single and double strand competitors

Competition for MAC I binding to antisense -66/+49 probe by unlabelled DNA was assessed for either strand (Antisense; Sense) of the -66/+49 sequence, both -66/+49 strands denatured (Both, denatured) or double strand -66/+49 DNA (Double). Denatured double strand competitor was prepared by a 5 min 95°C incubation and immediate cooling on ice. The indicated molar excess of competitor was incubated with 4 µg Rat-1 control extract before addition of the labelled probe.

**Free I**: free single strand probe

**Free II**: free double strand probe
<table>
<thead>
<tr>
<th>Competitor Strand:</th>
<th>Antisense</th>
<th>Sense</th>
<th>Both, denatured</th>
<th>Double</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fold Excess:</td>
<td>0 5 50</td>
<td>0 5 50</td>
<td>0 5 50</td>
<td>0 5 50</td>
</tr>
</tbody>
</table>

- **MAC I**
- **Free I**
- **Free II**
Figure VI.3: Titration of antibody against MAC I

An IgG preparation of mAb1 ascitic fluid was titrated for MAC I supershifting capacity. The indicated volume of antibody was preincubated with 4 µg Rat-1 control extract. Antisense -66/+49 probe and poly d(IC)•poly d(IC) carrier were used in each reaction.

The positions of MAC I (MAC I) and free probe (F) are indicated
Figure VI.4: Carrier specificity of MAC I binding

The formation of MAC I on antisense 66/+49 single strand probe using different non-specific carriers was assessed. The use of 1.5 μg poly d(IC)•poly d(IC) (pdIC•pdIC), 1 μg double strand salmon sperm DNA (Salmon Sperm DNA), or ~1 μg single strand M13 based Bluescript vector (Single strand BS) DNA in the binding reactions is indicated.

The presence of MAC I was determined by the ability of a complex to be supershifted following preincubation of the extract with mAb1 αmyc antibody (+ mAb1).

The MAC I (MAC I) and supershifted MAC I (Shift) complexes are indicated.

F: Free probe.
carrier: pdIC·pdIC, Salmon Sperm DNA, Single strand BS + mAb 1

MAC I

Shift

F
Figure VI.5: Presence of MAC I in different cell types and in response to myc levels

Left panel: Extracts from a variety of cell populations expressing different amounts and mutants of c-myc protein were used in assays with the antisense -66/+49 human c-myc P2 probe. Extracts were prepared from proliferating Rat-1, NIH-3T3 and HeLa cells. Pools of Rat-1 cells expressing v-myc, wild type human c-myc protein and human c-myc protein mutants Δ41-53 and Δ145-262 and NIH-3T3 cells expressing v-myc were each used for extract preparation (Penn, et al, 1990b).

4 µg of each extract and 1.5 µg poly d(IC)•poly d(IC) were used in each assay. The indicates extracts (+) were preincubated with mAb1 αmyc antibody.

The MAC I and supershifted MAC I complexes (MAC I, Shift) are indicated.

F: Free probe

Right: The endogenous c-myc expression (P2) level in10 µg RNA prepared from the control or Δ41-53 Rat-1 cell pools used for the extract preparations was analysed by RNAase protection. Endogenous c-myc gene expression was monitored with the rat myc exon I probe and GAPDH expression (GAPDH) with the Rat GAPDH probe.
Figure VI.6: Competition mapping of MAC I binding: antisense strand competitors

Sequences within the -71/+49 human c-myc promoter, antisense strand were used as competitors for MAC I binding to antisense -66/+49 probe. The molar excess of each competitor used is indicated. The -66/+22 competition was performed with denatured (95°C, 5 min) double strand DNA, all others with synthetic oligonucleotides. The -39/-19 oligonucleotide contains a T->A point mutation within the 'TTTTATA' box at position -29 (pm-29).

All binding reactions were performed in the presence of 4 μg Rat-1 control extract and 1.5 μg poly d(IC)•poly d(IC).

The MAC I (MAC I) and free probe (F) positions are indicated.
Figure VI.7: Competition mapping of MAC I binding: sense strand and unrelated competitors

Sequences within the -71/+49 human c-myc promoter, sense strand were used as competitors for MAC I binding to antisense -66/+49 probe. All competitors were synthetic oligonucleotides, and the molar excess used in each reaction is indicated. All binding reactions were performed in the presence of 4 μg Rat-1 control extract and 1.5 μg poly d(IC)•poly d(IC).

The synthetic oligonucleotides 7070 and 7074 contain sequences from the 3rd exon of the human c-myc gene, and are not homologous to the MAC I binding sequences. Their sequences are:

7070: 5' A GTA ATA TTG AAA AAA GCC ACA GCA TAC ATC TTA AGC GTC CAA GCA GAG AGG CAA AAG CT 3'

7074: 5' AAA CGA AGC TTT TTT GCC CTT CGC GAC CAG GGC CCG GAA CTC GAG 3'

The MAC I (MAC I) and single and double strand free probe (Free I and Free II) positions are indicated.
<table>
<thead>
<tr>
<th>Probe</th>
<th>-66/+49 Antisense</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strand</td>
<td>Sense</td>
</tr>
<tr>
<td>Competitor</td>
<td>-66/+49</td>
</tr>
<tr>
<td>Fold Excess</td>
<td>0</td>
</tr>
<tr>
<td>MAC I</td>
<td></td>
</tr>
<tr>
<td>Free I</td>
<td></td>
</tr>
<tr>
<td>Free II</td>
<td></td>
</tr>
</tbody>
</table>
Figure VI.8: Deletion mapping of MAC I binding via supershifting (short exposure)

Single strand oligonucleotide probes were used to map the binding region of supershiftable MAC I. Synthetic probes containing the sense or antisense strand of the human c-myc P2 promoter -66/+49 region were used in binding reactions in the presence of mAb1, mAb1 and its cognate peptide, pep1 or mAb1 and a control peptide, pep2. Each reaction contained 4 µg Rat-1 control extract and 1.5 µg poly d(IC)•poly d(IC) plus antibody and peptide as indicated. The probes are presented schematically at the bottom of the figure.

A: -66/+49 sense strand
B: -66/+49 antisense strand
C: -21/+49 antisense strand
D: -21/+22 antisense strand
Figure VI.9: Deletion mapping of MAC I binding via supershifting
(long exposure)

This is the same gel as in the previous figure, exposed for a longer time.
<table>
<thead>
<tr>
<th>probe:</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>mAb 1:</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>peptide 1:</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>peptide 2:</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
</tr>
</tbody>
</table>

-66: sense
-21: antisense
+22: antisense
+49: antisense
Figure VI.10: Mapping of MAC I binding: Competing supershifted complexes

The ability of various sequences to compete the binding of the protein(s) which form supershifted MAC I on -66/+49 antisense probe was assessed. mAb1 was preincubated with 4 µg Rat-1 control extracts which were subsequently incubated with competitor and then probe. All binding reactions used 1.5 µg poly d(IC)•poly d(IC) as carrier.

The presence of mAb1 and molar excess of specific competitors are indicated along with the positions of MAC I (MAC I) in its normal and supershifted (Shift) forms.

F: Free probe
<table>
<thead>
<tr>
<th>competitor:</th>
<th>-66/+49 anti</th>
<th>-71/-22 anti</th>
<th>-21/+49 anti</th>
<th>-66/+22 boiled</th>
<th>-21/+22 anti</th>
</tr>
</thead>
<tbody>
<tr>
<td>fold excess:</td>
<td>-</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
<td>-0</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>50</td>
<td>50</td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td>mAb 1:</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

---

[Image of gel electrophoresis with bands labeled 'Shift' and 'MAC 1']

---

[Arrow indicating 'Shift' and 'MAC 1']
Table VI.1: Deletion and competition mapping of MAC I binding:

Summary

The sequence of all probes used in gel retardation, supershifting and competition experiments with MAC I is presented. Capital letters indicate sequences which are identical with wild type human c-myc, lower case those which differ.

The band column indicates whether a complex with a mobility similar to MAC I was detected and the +mAb1 column whether a supershift was detected when using mAb1 in the reaction. Comp indicates the ability of a given sequence to compete MAC I bound to -66/+49 antisense strand probe, and Comp + mAb1 whether this sequence competes the supershifted form of MAC I.

+: a strong positive signal
+/-: clearly positive, but not a strong as -66/+49
-/+: weakly positive
--/+: very weakly positive, e.g. sense strand probe

These ratings are arbitrarily defined, and relative within the same type of assay (e.g. competition vs competition)

A blank at any position indicates the experiment was not performed.

WTMYC: wild type human c-myc sequence

WTCON: Consensus MAC I binding site from the supershiftable complex data is underlined

The binding sites of the E2F and 'Me1a1' complexes as well as the human c-myc P2 TATA box and mRNA cap site are indicated.
### SINGLE STRAND BINDING AND COMPETITION

#### ANTISENSE STRAND

<table>
<thead>
<tr>
<th>Position</th>
<th>Nucleotide Sequence</th>
<th>Bandshift</th>
<th>Comp</th>
<th>Comp + mAbl</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRTC</td>
<td><code>GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-66/49</td>
<td><code>...&lt;gcggccct TTTTCCGC TCCCTGCA GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>-63/46</td>
<td><code>...&lt;gcggccct TTTTCCGC TCCCTGCA GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td>+</td>
<td>+</td>
<td>-/+</td>
</tr>
<tr>
<td>-60/43</td>
<td><code>...&lt;gcggccct TTTTCCGC TCCCTGCA GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td>+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>-57/40</td>
<td><code>...&lt;gcggccct TTTTCCGC TCCCTGCA GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td>+</td>
<td>+</td>
<td>+/-</td>
</tr>
<tr>
<td>-54/37</td>
<td><code>...&lt;gcggccct TTTTCCGC TCCCTGCA GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>-51/34</td>
<td><code>...&lt;gcggccct TTTTCCGC TCCCTGCA GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>-48/31</td>
<td><code>...&lt;gcggccct TTTTCCGC TCCCTGCA GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>-45/28</td>
<td><code>...&lt;gcggccct TTTTCCGC TCCCTGCA GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>-42/25</td>
<td><code>...&lt;gcggccct TTTTCCGC TCCCTGCA GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td>+</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

#### SENSE STRAND

<table>
<thead>
<tr>
<th>Position</th>
<th>Nucleotide Sequence</th>
<th>Bandshift</th>
<th>Comp</th>
<th>Comp + mAbl</th>
</tr>
</thead>
<tbody>
<tr>
<td>WRTC</td>
<td><code>GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-66/49</td>
<td><code>...&lt;gtccggccct TTTTCCGC TCCCTGCA GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td>-/+</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>-71/22</td>
<td><code>...&lt;gtccggccct TTTTCCGC TCCCTGCA GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>-39/19</td>
<td><code>...&lt;gtccggccct TTTTCCGC TCCCTGCA GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
<tr>
<td>-1/2</td>
<td><code>...&lt;gtccggccct TTTTCCGC TCCCTGCA GACGCTGC AAAAACGCA AGGGAGAT</code></td>
<td>-</td>
<td>+</td>
<td></td>
</tr>
</tbody>
</table>
Figure VI.11: Thermal stability of MAC I binding activity with or without probe

The effect of temperature on MAC I in extracts, or preformed on the -66/49 antisense probe was investigated. Rat-1 control extracts were incubated at the indicated temperature for 5 min and then cooled on ice before probe addition (Extract incubation). Alternatively Rat-1 control extracts were combined with probe prior to incubation at the indicated temperature (Complex incubation). All reactions were performed using 5 μg extract and 1.5 μg poly d(IC)•poly d(IC) as carrier.

The mobilities of MAC I (MAC I) and free probe (F) are indicated.
<table>
<thead>
<tr>
<th>treatment:</th>
<th>Extract incubation</th>
<th>Complex incubation</th>
</tr>
</thead>
<tbody>
<tr>
<td>temp:</td>
<td>4°C 37°C 45°C 52°C 68°C</td>
<td>4°C 37°C 45°C 52°C 68°C</td>
</tr>
</tbody>
</table>

**MAC I**

**F**
Figure VI.12: Secondary structure predictions of MAC I binding site

Each of the single strand sequences which correspond to the antisense human c-myc P2 promoter fragment indicated was analysed using an RNA secondary structure analysis program (Devereux, et al, 1984, Freier, et al, 1986, Zuker and Stiegler, 1981). (See Table VI.1 for literal representations of each sequence.) The most energetically favorable structure is presented in two formats for each sequence.

The left hand graphic indicates complementary nucleotides by a curving line above the graphically represented horizontal sequence. Note that each sequence is presented in the 5' to 3' orientation, such that upstream sequences within the promoter are to the left. The right hand graphic represents the same structure in standard base pair format. The right-hand most column indicates the ability of each sequence to bind MAC I when used as a probe (see Table VI.1).

The sequences involved in the core stem-loop structure are indicated at the bottom of the figure.
-66/+49

-22/+49

-13/+49

-21/+22

-21/+31

predicted +

5' TCTGCCCTGCTGGAATTACTACAGCGAGUAGATAAAGCCCCGAAAACCG 3'
Chapter 7 Perspectives
7.1 Summary of conclusions

The initial aims of this project were two-fold. The first was to investigate the cis-acting elements and trans-acting factors through which the negative autoregulatory control of c-myc transcription is mediated in order to increase the understanding of the regulatory mechanisms governing c-myc expression. The second was to approach the same phenomenon as a function of c-myc protein, in order to gain insight into the molecular mechanisms by which c-myc protein may exert its effects on the control of cell proliferation and differentiation.

Using a combination of molecular genetic and biochemical approaches it has been possible to determine sequences which are sufficient to impart responsiveness to elevated levels of myc protein to a heterologous indicator, and to partially characterise some of the protein:DNA complexes which can form on these DNA sequences in vitro (see Figure VII.1). It appears likely that within the region extending from 71 bp upstream and 49 bp downstream of the human c-myc P2 promoter mRNA cap site there exist multiple redundant elements which can mediate the down regulatory response. Several protein complexes have been identified, some of which (E2F, B/β) appear to be similar to the previously described complexes E2F and Me1a1, and others of which have not been described previously. These include two complexes, MAC I and MAC II, which are antigenically related to the c-myc protein. MAC I binds specifically to single stranded DNA and shows a strong preference for sequences in the antisense strand of the human c-myc gene surrounding the P2 cap site. Each MAC I protein:DNA complex contains multiple copies of a myc related epitope, which is most likely present on a protein of the myc family that is distinct from the c-myc protein (see Chapter 6). MAC II is also antigenically related to c-myc protein, and is likely to contain c-myc as one of its protein components (see Chapter 5). This complex forms on
DNA fragments which possess a unique metastable property that appears to be imparted by prior incorporation of the DNA fragments in bacterially derived plasmids. The minimal sequences on which this complex has been observed to form extend from 63 bp upstream through 49 bp downstream of the human c-myc P2 promoter mRNA cap site.

The region of the P2 promoter to which the myc response has been assigned (-71 to +49) has been implicated in at least four other regulatory aspects of c-myc expression (Figure VII.1). This region is involved in the maintenance of basal levels of myc transcription and in the correct positioning of the RNA polymerase, as the sequence spans the P2 TATA box and mRNA cap site, as well as including elements which have been implicated in the control of basal c-myc P2 promoter mediated transcription in a number of in vivo and in vitro systems (Asselin, et al, 1989, Cooney, et al, 1988, Hall, 1990, Hay, et al, 1987, Lipp, et al, 1987, Nishikura, 1986, Spencer, et al, 1990). Transcription modulatory controls can also be mediated through this region, as the -66 to -57 sequences have been implicated in the response of the c-myc gene to stimulation by serum growth factors and the adenovirus E1a protein, and contain the E2F binding site (Mudryj, et al, 1990, Thalmeier, et al, 1989). The upstream sequences between nucleotides -54 and -38, which are highly conserved between the human and mouse c-myc genes, and which include the Me1a1 protein binding site, have been shown to be necessary for the premature termination of transcription at 3' end of murine c-myc exon I in stable transfection assays (Miller, et al, 1989). In addition the sequences from -45 to -38 have also been implicated in the regulation of the serum and E1a responses (Hiebert, et al, 1989, Mudryj, et al, 1990). It has also been found that transcripts which initiate from the c-myc P1 promoter can prematurely terminate at sequences on either side of the P2 promoter TATA box (S. Roberts and D. Bentley, personal communication). Thus
there are numerous controls on c-myc expression which map to these overlapping sequences, and it seems highly probable that the DNA:protein complexes which have been identified in the present study are involved in some of these or other unidentified regulatory mechanisms, and not necessarily in the c-myc feedback response.

7.2 Mechanisms of myc suppression

The data obtained in the present study have not been resolved to a high enough degree to derive direct conclusions as to the mechanisms by which myc protein down-regulates its own transcription. However it is possible to speculate on the mechanisms which may be functioning. The inability to define a unique sequence through which the suppression is mediated suggests that there exist functionally redundant cis-acting elements. The actions exerted through each of these elements may be mechanistically similar, or alternatively, multiple distinct mechanisms may operate through different sequence and protein components. A number of plausible models can be postulated, each of which takes into account some of the results presented in the previous chapters.

One model is suggested by the observation that the upstream initiated transcripts are almost invariably regulated coordinately with the c-myc promoter P1 and P2 initiated transcripts (eg Figure III.10-12). The coordinated response could occur through a physical blockage to the passage of a transcription complex. This would both prevent the further elongation of the upstream initiated transcripts and the synthesis of appropriately initiated transcripts. Such a physical blockage could be instigated by either the MAC I or MAC II complex, as the binding sites of both complexes effectively span the P2 promoter cap site, which would position them appropriately to occlude access of the RNA polymerase. Such a mechanism might be similar to the binding of a prokaryotic repressor molecule to its cognate operator site, which can either prevent
the binding of RNA polymerase to the promoter or alter its ability to participate in an active transcription complex (Goodbourn, 1990 for review).

Alternatively, the repression may be mediated by other kinds of protein:protein interaction. These interactions could involve the complexes which form on the DNA near the P2 promoter cap site and elements of the RNA polymerase-containing transcription complex. These interactions could inhibit the formation of pre-initiation complexes on the DNA, or they could block any of the intermediary steps leading to the initiation of transcription. Interactions might occur directly between the DNA binding complex and the transcription complex, or indirectly, through bridging protein(s) anchored to the DNA via contacts with the DNA binding protein(s). The proteins involved in these complexes could act by disrupting the functioning of positively acting factors as they interact with the basal transcription complexes, by recruiting negative effectors of transcription, or through more direct interactions with the RNA polymerase. It is conceivable that myc protein could act at any of these levels, as it possesses sequences which may be involved in homo- or heterodimerisation, other protein:protein interactions and protein:DNA interactions (Penn, et al, 1990b). The E2F protein has been demonstrated to exist in various heteromeric forms, making it a plausible candidate for interactions with myc protein (Bagchi, et al, 1990). Similarly, interaction with the other complexes, such as complex B, are also possible. MAC I and MAC II may also be more directly involved in these types of regulatory control. It will only be possible to determine which, if any, of these mechanisms are involved in mediating c-myc autosuppression once the requisite sequence elements as well as the proteins with which they interact have been better defined.
One alternative approach to the general strategy of introducing DNA constructs into cells would be to reconstitute the regulatory events using in vitro assays. Correctly initiated RNA polymerase II-dependent in vitro transcription from the c-myc promoter has been described for both the human and murine genes (Cooney, et al, 1988, Hall, 1990). However, physiological regulation of these genes in vitro has not yet been described. In vitro transcription analyses of a number of other promoters have been developed to the stage where some regulatory events which operate through upstream DNA elements have been mimicked in vitro (Mitchell, et al, 1987). In addition, several of the proteins involved in the basal transcription apparatus, as well as numerous other upstream regulatory factors, have been cloned over the past several years (Kadonaga, 1990, Mitchell and Tjian, 1989). As more general and sequence specific transcription factors are cloned, the reconstitution of regulable transcription complexes on the c-myc promoter using defined and pure components should become a more viable proposition.

7.3 Possible functions for MAC I and MAC II

Further characterisation of MAC I and MAC II will be of particular interest as these complexes contain c-myc or c-myc related proteins, and may provide insight into the biochemical activities of the myc proteins. At present it is not possible to determine whether these complexes are involved in regulating c-myc transcription or whether they mediate cellular functions other than controlling expression of the c-myc locus. They may be involved in mediating other regulatory actions attributed to myc protein such as the control of entry into S phase or monitoring the potential for the maintenance of the proliferative state (see Chapter 1). While such activities may involve the control of transcription, there is no prerequisite for this particular biochemical action.
On a more mechanistic level, it is conceivable that the functions of the MAC II or MAC I complexes involve multiple distinct regulatory stages, which when combined suggest the potential for sophisticated switching controls. These regulatory steps might involve control not only of complex assembly, but also of the binding potential at each binding site. Such possibilities are implied by the sequence independent properties of the MAC I and MAC II binding sites, namely the single strand component required for MAC I binding (Chapter 6), and the poorly understood properties of the DNA probes which are capable of binding MAC II (Chapter 5). While it is possible that the properties which allow MAC II formation are constitutively present in vivo, this seems unlikely as the probes display such variable MAC II binding capacity in vitro. Similarly, the single stranded conditions required for MAC I formation are probably not always present, although the passage of transcription complexes through the c-myc P2 promoter probably leads to the frequent denaturing of the relevant sequences. Thus the necessary conditions for complex formation may be imparted to the genomic DNA only under particular conditions. The ability of the complexes to form may also be regulated by the rate-limiting presence of necessary components such as c-myc protein. These controls may lead to strongly conditional complex formation, which in turn may be important in the synchronisation of certain biological activities such as DNA synthesis and cell division. The possibility that myc protein may operate via such mechanisms to regulate fundamental cellular functions appears worthy of further investigation.

7.4 Future prospects

In order to determine the biological significance of the MAC I and MAC II complexes it will be necessary to characterise them to a much greater degree than has currently been achieved. In particular, it would be very useful to determine the DNA sequences with which the proteins in
these complexes interact to a higher degree of precision. Such information might allow the identification of individual nucleotides which are critical for complex formation. Experiments using fragments of the human c-myc promoter which had been specifically mutated at these residues could then be performed in order to ascertain the necessity of these complexes for correct regulation of the c-myc promoter.

It is also necessary to determine whether c-myc protein is actually present in either MAC I or MAC II. Currently the only data available are of an immunological nature. Confirmation of the identity of the protein components of these complexes by other means would strengthen any conclusions which could be made. Several approaches are evident. One would be to tag introduced exogenous myc protein with a non-myc epitope, and observe whether its cognate antibody would recognise the introduced protein. Alternatively, in vitro translated myc protein might be capable of contributing to complex formation. In a divergent approach, biochemical purification of the proteins involved in these complexes would also lead to their unambiguous identification. A combination of divergent approaches should ultimately lead to the unambiguous resolution of this question.

In summary the major outstanding issues raised in this study are: the relationship of the sequence and protein elements which prevail in the autoregulation of c-myc expression, the nature of the components incorporated in the myc related complexes, the cellular functions of these complexes and their relationship to the normal activities of c-myc protein. Their resolution should lead to greater insight into the ways in which the c-myc oncogene contributes to the control of cell proliferation and differentiation.
Figure VII.1: Regulatory elements and protein factors at the c-myc P2 promoter: Summary

Positions of both regulatory sequence elements and protein factors which have been identified at the c-myc P2 promoter are symbolically represented.

Sequence coordinates are given relative to the human c-myc P2 promoter mRNA cap site. Positions of the TATA box and mRNA cap site are indicated.

Data are presented from the following references:

Serum stimulation: (Mudryj, et al, 1990)
Downstream termination: (Miller, et al, 1989)
P1 dependent termination: (S Roberts and D Bentley, personal communication)
Autosuppression
MAC II binding
MAC I binding
Basal transcription
Serum and E1a stimulation
Downstream termination
P1 dependent termination
References


Alitalo K, JM Bishop, DH Smith, EY Chen, WW Colby and AD Levinson (1983) Nucleotide sequence of the v-myc oncogene of avian retrovirus MC29 Proc Natl Acad Sci USA 80 100-104.


Andres AC, dVM van, CA Schonenberger, F Fluckiger, M LeMeur, P Gerlinger and B Groner (1988) Ha-ras and c-myc oncogene expression interferes with morphological and functional differentiation of mammary epithelial cells in single and double transgenic mice Genes Dev 2 1486-1495.

Ariga H, Y Imamura and SMM Iguchi-Ariga (1989) DNA replication origin and transcriptional enhancer in c-myc gene share the c-myc protein binding sequences EMBO J 8 4273-4279.


Bargmann CI, MC Hung and RA Weinberg (1986) Multiple independent activations of the neu oncogene by a point mutation altering the transmembrane domain of p185 *Cell* 45 649-657.


Bentley DL and M Groudine (1986a) A block to elongation is largely responsible for decreased transcription of c-myc in differentiated HL60 cells *Nature* 321 702-706.


Brown K, M Quintanilla, M Ramsden, I Kerr, S Young and A Balmain (1986) v-ras genes from Harvey and Balb murine sarcoma viruses can act as initiators of two stage mouse skin carcinogenesis Cell 46 447-456.


Cai M and RW Davis (1990) Yeast centromere binding protein CBF1, of the helix-loop-helix protein family, is required for chromosome stability and methionine prototrophy Cell 61 437-446.


Chirgwin JM, AE Przybyla, RJ MacDonald and WJ Rutter (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease Biochemistry 18 5294-5299.


Coughlin SR, WMF Lee, PW Williams, GM Giels and LT Williams (1985) c-myc gene expression is stimulated by agents that activate protein kinase C and does not account for the mitogenic effect of PDGF Cell 43 243-251.


Dalla Favera R, M Bregni, J Erikson, D Patterson, RC Gallo and CM Croce (1982a) Human c-myc onc gene is located on the region of chromosome 8 that is translocated in Burkitt lymphoma cells Proc Natl Acad Sci USA 79 7824-7827.


Dean M, RA Levine, W Ran, MS Kindy, GE Sonenshein and J Campisi (1986) Regulation of c-myc transcription and mRNA abundance by serum growth factors and cell contact *J Biol Chem* 261 9161-9166.


Der CJ, TG Krontiris and GM Cooper (1982) Transforming genes of human bladder and lung carcinoma cell lines are homologous to the ras genes of Harvey and Kirsten sarcoma viruses *Proc Natl Acad Sci USA* 79 3637-3640.


387


Dransfield I and N Hogg (1989) Regulated expression of Mg2+ binding epitope on leukocyte integrin alpha subunits EMBO J 8 3759-3765.


Dunnick W, BE Shell and C Dery (1983) DNA sequences near the site of reciprocal recombination between a c-myc oncogene and an immunoglobulin switch region Proc Natl Acad Sci USA 80 7269-7273.

Eick D and GW Bornkamm (1986) Transcriptional arrest within the first exon is a fast control mechanism in c-myc expression Nuc Acids Res 14 8331-8346.


Endo T and B Nadal-Ginard (1986) Transcriptional and posttranscriptional control of c-myc during myogenesis: its mRNA remains inducible in differentiated cells and does not supress the differentiated phenotype Mol Cell Biol 6 1412-1421.


Falcone G, F Tato and S Alema (1985) Distinctive effects of the viral oncogenes myc, erb, fps, and src on the differentiation program of quail myogenic cells Proc Natl Acad Sci USA 82 426-430.


Freytag SO (1988) Enforced expression of the c-myc oncogene inhibits cell differentiation by precluding entry into a distinct predifferentiation state in G0/G1 Mol Cell Biol 8 1614-1624.


Friend SH, R Bernards, S Rogelj, RA Weinberg, JM Rapaport, DM Albert and TP Dryja (1986) A human DNA segment with properties of the
gene that predisposes to retinoblastoma and osteosarcoma Nature 323 643-646.

Frykberg L, T Graf and B Vennstrom (1987) The transforming activity of the chicken c-myc gene can be potentiated by mutations Oncogene 1 415-421.


Godeau F, H Persson, HE Gray and AB Pardee (1986) c-myc expression is dissociated from DNA synthesis and cell division in Xenopus oocyte and early embryonic development EMBO J 5 3571-3577.


autocrine growth and enhanced tumorigenicity in v-myc-transformed avian macrophages Cell 45 357-364.


Griep AE and HF DeLuca (1986) Decreased c-myc expression is an early event in retinoid acid-induced differentiation of F9 teratocarcinoma cells Proc Natl Acad Sci USA 83 5539-5543.


Hall DJ (1990) Regulation of c-myc transcription in vitro: dependence on the guanine-rich promoter element ME1a1 Oncogene 5 47-54.


Hayflick J, PH Seeburg, R Ohlsson, S Ohlsson-Pfeiffer, D Watson, T Papas and PH Duesberg (1985) Nucleotide sequence of two overlapping myc-related genes in avian carcinoma virus OK10 and their relation to the myc genes of other viruses and the cell *Proc Natl Acad Sci USA* 82 2718-2722.


Hiebert SW, M Lipp and JR Nevins (1989) E1A-dependent trans-activation of the human MYC promoter is mediated by the E2F factor *Proc Natl Acad Sci USA* 86 3594-3598.

Hope IA and K Struhl (1986) GCN4, a eukaryotic transcriptional activator, binds as a dimer to target DNA Cell 46 885-894.


Jones NC, PWJ Rigby and EB Ziff (1988) Trans-acting protein factors and
the regulation of eukaryotic transcription: lessons from studies on
DNA tumor viruses *Genes and Dev* 2 267-281.

Jones TR and MD Cole (1987) Rapid cytoplasmic turnover of c-myc
mRNA: requirement of the 3' untranslated sequences *Mol Cell Biol* 7
4513-4521.

Kaczmarek L, J Hyland, R Watt, M Rosenberg and R Baserga (1985)
Microinjected c-myc as a competence factor *Science* 228 1313-1315.

Activation and repression of mammalian gene expression by the c-
mvc protein *Genes Dev* 1 347-357.

Kadonaga JT (1990) Gene transcription: basal and regulated transcription

Kahn P, A Leutz and T Graf (1986) Individual and combined effects of viral
oncogenes in haematopoietic cells. In *Oncogenes and Growth Control*,
Kahn P and T Graf, eds. (Springer-Verlag, Berlin) pp 312-319.

Kakkis E and K Calame (1987) A plasmacytoma-specific factor binds the c-
mvc promoter region *Proc Natl Acad Sci USA* 84 7031-7035.

is found specifically in plasmacytomas *Curr Top Microbiol Immunol*
141 231-237.


Kan NC, CS Flordellis, GE Mark, PH Duesberg and TS Papas (1984) A
common onc gene sequence transduced by avian carcinoma virus
MH2 and by murine sarcoma virus 3611 *Science* 223 813-816.

Regulation of cell cycle duration by c-myc levels *Oncogene* 4 773-787.

Katagiri F, E Lam and N-H Chua (1989) Two tobacco DNA-binding
proteins with homology to the nuclear factor CREB *Nature* 340 727-
730.

Keath EJ, PG Caimi and MD Cole (1984a) Fibroblast lines expressing
activated c-myc oncogenes are tumourigenic in nude mice and
syngeneic animals *Cell* 39 339-348.

Keath EJ, A Kelekar and MD Cole (1984b) Transcriptional activation of the
translocated c-myc oncogene in mouse plasmacytomas: Similar RNA
levels in tumor and proliferating normal cells *Cell* 37 521.


Lachman HM, KS Hatton, AI Skoultchi and CL Schildkraut (1985) c-myc mRNA levels in the cell cycle change in mouse erythroleukemia cells following inducer treatment Proc Natl Acad Sci USA 82 5323-5327.


Marcu KB, LJ Harris, LW Stanton, J Erikson, R Watt and CM Croce (1983) Transcriptionally active c-myc oncogene is contained within NIARD, a DNA sequence associated with chromosome translocations in B-cell neoplasia *Proc Natl Acad Sci USA* 80 519-523.


Morgenstern JP and H Land (1990b) A series of mammalian expression vectors and characterisation of their expression of a reporter gene in stably and transiently transfected cells NAR 18 1068.


Murre C, PS McCaw, H Vaessin, M Caudy, LY Jan, YN Jan, CV Cabrera, JN Buskin, SD Hauschka, AB Lassar, H Weintraub and D Baltimore (1989b) Interactions between heterologous helix-loop-helix proteins generate complexes that bind specifically to a common DNA sequence Cell 58 537-544.


Nardone G, ME Wastney and P Hensley (1990) DNA structural polymorphism modulates the kinetics of superhelical DNA cleavage by BamHI restriction endonuclease J Biol Chem 265 15308-15315.


Nepveu A and KB Marcu (1986) Intragenic pausing and anti-sense transcription within the murine c-myc locus EMBO J 5 2859-2865.


Penn LJZ, MW Brooks, EM Laufer and H Land (1990a) Negative autoregulation of c-myc transcription EMBO J 9 1113-1121.


expression in developing human embryos Proc Natl Acad Sci USA 82 5050-5054.


Rabbitts PH, A Forster, MA Stinson and TH Rabbitts (1985a) Truncation of exon I from the c-myc gene results in prolonged c-myc mRNA stability EMBO J 4 3737-3733.


Schwab M, K Alitalo, KH Klempnauer, HE Varmus, JM Bishop, F Gilbert, G Brodeur, M Goldstein and J Trent (1983) Amplified DNA with limited homology to myc cellular oncogene is shared by neuroblastoma cell lines and a neuroblastoma tumor Nature 305 245-248.


Shih C, B-Z Shilo, MP Goldfarb, A Dannanberg and RA Weinberg (1979) Passage of phenotypes of chemically transformed cells via transfection of DNA and chromatin Proc Natl Acad Sci USA 76 5714-5718.

Shih C-K, M Linial, MM Goodenow and WS Hayward (1984) Nucleotide sequence 5' of the chicken c-myc coding region: localization of a noncoding exon that is absent from myc transcripts in most avian leukosis virus-induced lymphomas Proc Natl Acad Sci USA 81 4697-4703.


Simon MC, TB Fisch, BJ Benecke, JR Nevins and N Heintz (1988) Definition of multiple, functionally distinct TATA elements, one of which is a target in the hsp70 promoter for E1A regulation Cell 52 723-729.

Simon MC, RJ Rooney, T Fische, N Heintz and JR Nevins (1990) E1a-dependent activation of the c-fos promoter requires the TATAA sequence Proc Natl Acad Sci USA 87 513-517.


SivaRaman L and B Thimmappaya (1987) Two promoter-specific host factors interact with adjacent sequences in an E1a-inducible adenovirus promoter Proc Natl Acad Sci 84 6112-6116.


Stehelin D, HE Varmus, JM Bishop and PK Vogt (1976) DNA related to the transforming gene(s) of avian sarcoma viruses is present in normal avian DNA Nature 260 170-173.


Stewart TA, PK Pattengale and P Leder (1984b) Spontaneous mammary adenocarcinomas in transgenic mice that carry and express MTV/myc fusion genes Cell 38 627-637.


Taylor MV, M Gusse, GI Evan, N Dathan and M Mechali (1986) Xenopus myc proto-oncoprotein expression during development: expression as a stable maternal mRNA uncoupled from cell division EMBO J 5 3563-3570.


Wigler M, A Pellicer, S Silverstein, R Axel, G Urlaub and L Chasin (1979a) DNA-mediated transfer of the adenine phosphoribosyltransferase locus into mammalian cells Proc Natl Acad Sci USA 76 1373-1376.


Wright S and JM Bishop (1989) DNA sequences that mediate attenuation of transcription from the mouse protooncogene myc Proc Natl Acad Sci USA 86 505-509.

Wright WE, DA Sassoon and VK Lin (1989) Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD Cell 56 607-617.


C-MYC: evidence for multiple regulatory functions

Linda J.Z. Penn*, Edward M. Laufer and Hartmut Land

The nuclear c-myc proto-oncogene promotes cell proliferation and can inhibit terminal differentiation as well as induce immortalisation in its oncogenic form. There is increasing evidence that c-myc exerts these biological activities by modulating transcription and by directly affecting the initiation of DNA replication. The regulation of these disparate activities may involve the carboxyl end of the c-myc protein, which is essential for transformation and autosuppression of c-myc transcription. Conserved motifs in this region of the c-myc protein may mediate complex formation and sequence-specific nucleic acid binding.

Key words: autoregulation / DNA replication / proliferation / transcription / transformation

Amongst the known proto-oncogenes the cellular myc (c-myc) gene is one of those most frequently implicated in carcinogenesis. Its oncogenic activation results from constitutive expression of the structurally unaltered c-myc protein (reviewed in refs 1, 2) which suggests that at least some functions of normal and activated c-myc genes are identical. In this review we will focus on the efforts and approaches employed to analyse the biological role and molecular functions of c-myc in proliferation and differentiation; and we will discuss recent protein structure predictions which may further our understanding of c-myc's function in growth control.

The c-myc gene encodes a short-lived nuclear phosphoprotein which is well conserved in vertebrates. It is a member of a family of structurally and functionally closely related genes consisting of three well defined cellular genes, c-myc, N-myc and L-myc, as well as several c-myc-derived oncogenes transduced by acutely transforming retroviruses (v-myc) (reviewed in refs 1, 2). The gene family likely extends to other functional genes termed R-myc, p-myc and B-myc; however at present these genes are only partially characterised. C-myc and v-myc genes have served interchangeably as tools to investigate the biological role of myc oncogenes. Indeed, their functions seem indistinguishable, although the v-myc oncogenes have undergone structural alterations which can enhance their potencies.

c-myc potentiates cellular self-renewal and can inhibit differentiation

The ability to increase cellular self-renewal potential can be regarded as the major biological activity of myc genes. This activity is reflected in the high levels of c-myc expression found in proliferating normal tissues (e.g. refs 7-10) and can be directly demonstrated upon introduction of myc oncogenes into tissue culture cells, reconstituted organs and transgenic mice. A large variety of cell types can be affected, including cells of fibroblastic, epithelial, haematopoietic, myogenic, neuroepithelial and glial origin (M. Noble and H. Land, unpublished). In vivo, increased self-renewal potential represents itself as hyperplasia which can affect specific cell lineages: in vitro it can lead to the manifestation of indefinite cell proliferation, a process also known as immortalisation. In non-continuous cell cultures as well as in animal models c-myc oncogenes alone are insufficient to induce malignant transformation. However, they contribute to this process in cooperation with other oncogenic lesions which include ras and myb oncogenes.

The notion that myc is not only involved in controlling cell division but also affects differentiation stems from observations that increased self-renewal is inseparable from an inhibition of cellular maturation in a number of cell types such as myoblasts, pre-B cells and glial progenitors (M. Noble and H. Land, unpublished). Furthermore, overexpression of c-myc...
blocks the differentiation of several continuous cell lines in vitro. In addition, a significant decrease in expression of the endogenous c-myc gene often precedes the onset of terminal differentiation, and artificial depletion of c-myc by anti-sense RNA can lead to accelerated differentiation. 

Therefore, it seems that expression of c-myc acts to promote proliferation, while its absence is required to allow differentiation to proceed. This is consistent with the idea that the c-myc-induced block in differentiation is primarily a consequence of the increase in self-renewal, although a more direct role of c-myc in the control of differentiation cannot be fully excluded.

Mediator between mitogens and cell cycle progression.

In non-transformed cell cultures c-myc expression is strongly responsive to mitogens and other environmental signals. For example, c-myc expression is transiently induced when quiescent cells are stimulated to reenter the cell cycle by serum or specific growth factors. Moreover, in exponentially proliferating cells c-myc expression is maintained at constant intermediate levels while mitogens are present. However, expression is rapidly reduced when these growth factors are removed or when cells cease to proliferate upon cell contact. This responsiveness to environmental stimuli defines c-myc as an integral component of the cellular signalling mechanism which establishes communication between the environment and the mechanisms controlling cell division.

In fact, there are several lines of evidence suggesting that c-myc itself is a key regulator of cell cycle progression. C-myc can reduce or partially abrogate the requirement for specific growth factors in the induction of DNA synthesis and its overexpression can increase the cellular growth rate possibly due to a shortening of G1. Moreover, in some cells c-myc antisense RNA can delay the progression through this phase of the cell cycle. However, c-myc seems not to function exclusively in G1. The constant synthesis of c-myc protein throughout the cell cycle in exponentially proliferating cells suggests that c-myc may have additional activities. Indeed, c-myc levels may be monitored continuously or at multiple restriction points to prevent entry into growth arrest (G0) following mitosis.

c-myc and DNA replication

The possibility that c-myc may be directly involved in the regulation of DNA replication has been discussed in the literature on several occasions but still remains a controversial issue. Evidence that c-myc might be required in cells for entry into S phase was provided by experiments in which the application of c-myc antisense oligonucleotides to human peripheral blood T lymphocytes blocked the initiation of DNA synthesis. Moreover, cellular DNA sequences acting as putative replication origins, one of which also exhibits enhancer function, have been immunoprecipitated in complexes containing c-myc and c-myc has also been found to promote extrachromosomal replication of a simian virus 40 (SV40) based vector. In contrast, a variety of antibodies against human c-myc protein were unable to interfere with initiation or elongation of SV40 DNA replication in vitro, nor did anti-myc antibody injection into nuclei of cultured fibroblasts prevent these cells from entering S-phase.

The most convincing evidence supporting a direct role for c-myc in the control of DNA replication comes from experiments with developing Xenopus laevis embryos performed in Marcel Mechali's laboratory. In the mature oocyte c-myc is one of the stored maternal proteins and resides in the cytoplasm until fertilisation, at which point it is rapidly translocated into the nucleus. This redistribution of the myc protein coincides with the zygote entering a phase of 12 synchronous cell cleavages which occur within 30-min intervals in the absence of any transcriptional activity. Injection of myc-specific antibodies blocked these rapid cell cleavages while coinjection of the same antibody preparation together with its peptide antigen had no effect. Further studies examining the influence of myc antibodies on in vitro DNA synthesis in extracts from Xenopus oocytes suggest that myc seems to be required for the initiation of replication, but that it is dispensable during elongation (M. Mechali, personal communication).

Taken together, c-myc seems to be required for progression through the G1 phase of the cell cycle and there is increasing evidence in favor of a direct role of c-myc in the initiation of DNA replication. In addition, the action of c-myc may not be restricted to its roles in G1 and S, since myc has also been shown to suppress ras oncogene-induced cell cycle arrest in G2/M. Thus, c-myc may be involved in the regulation of several functions required for cell cycle progression and continuous proliferation.
c-myc as a transcriptional regulator

There have been many attempts to determine whether c-myc acts as a regulator of transcription. This possibility has been particularly attractive because of the putative role for myc in the transmission of mitogenic signals as well as its nuclear localisation and non-specific DNA binding activity (reviewed in ref 2). In addition, there is an apparent functional analogy to the nuclear E1A gene of adenovirus which also induces immortalisation of primary cell cultures in vitro,64 and cooperates with ras oncogenes in transformation.65 These activities of the viral protein have been correlated with its ability to regulate gene transcription,66-70 suggesting the transforming activity of myc may involve a similar mode of action.

Autosuppression

Prompted to search for potential cellular target genes of c-myc, several investigators soon postulated that c-myc might suppress its own expression. This hypothesis was based on observations in Burkitt's Lymphoma showing that the expression of the translocated c-myc allele is constitutive, while the normal gene copy is silent.71,72 The expression of the normal c-myc gene is similarly suppressed in a variety of other spontaneously arising tumours in which one allele has been activated,73-76 and in tumour cells derived from transgenic mice harbouring a constitutively active myc gene.76,77 In addition, the exogenous expression of a myc gene from heterologous promoters also leads to the suppression of the endogenous gene in haematopoietic cells,79-84 and in rodent fibroblasts.85-87 However, it remains unclear whether a negative feedback mechanism is indeed the cause of the observed c-myc suppression outlined above. In fact, it has been suggested that in Burkitt's Lymphoma the normal c-myc gene may have been permanently silenced by other means, such as methylation or alterations in chromatin structure.88-90 Moreover, there has been disagreement as to whether c-myc is indeed a self-regulating gene since in some cell lines expression of exogenous myc is not followed by the suppression of endogenous c-myc expression.25,31,49,91-94

Recent experiments have addressed some of the ambiguities stated above and indicate that human c-myc can indeed repress endogenous c-myc gene expression in a wide variety of non-established and established rodent fibroblasts. This regulation occurs at the level of transcription initiation95 which is similar to the mechanism of c-myc down-regulation observed in rodent fibroblasts expressing v-myc oncogenes.87,96 The transcriptional nature of this suppression is supported by experiments which show that in stably transfected fibroblasts the human c-myc promoter linked to an indicator can be down-regulated in response to exogenous myc expression (E.M. Laufer, unpublished results).

The extent of c-myc suppression is proportional to the concentration of c-myc protein. Indeed, endogenous c-myc RNA is readily detectable in proliferating Rat-1 cells expressing 600-700 molecules of myc protein per cell and is suppressed ≥90% in cells expressing ≥5000 molecules per cell, a concentration which corresponds to c-myc levels 1-2 h post serum-stimulation. C-myc autosuppression is thus observed within the concentration range of c-myc protein found in a normal cell and may function as a homeostatic regulatory mechanism of c-myc expression.95

Interestingly, the c-myc protein is not the only trans-acting factor required for the autosuppression mechanism. This has become apparent from somatic cell fusions between autosuppression responsive and non-responsive cell lines which resulted in hybrids demonstrating the autosuppression positive phenotype.95 Therefore, it is possible that some of the cells which do not exhibit c-myc-induced autosuppression may be deficient in at least one of the functionally required components.

Cellular targets

The results on negative autoregulation offer the intriguing possibility that c-myc may also regulate the expression of other cellular genes, some of which may be involved in the cellular transformation process. One candidate is the αL gene, encoding a subunit of the heterodimERIC LFA-1 cell-adhesion receptor, which was found to be transcriptionally repressed by exogenous expression of c-myc in EBV-immortalised B lymphoblastoid cells (F. Grignani and R. Dalla-Favera, personal communication). This alteration could contribute to B cell malignancy, since it might help these cells to escape from immunosurveillance. In transient cotransfection assays myc has been shown to activate heat shock protein 70 promoters and to suppress the activity of the metallothionein I promoter, although the effects on the endogenous genes remain unclear.96,97 In addition to transcrip-
tional regulation, c-myc can also modulate the expression of cellular genes (mrl and mr2) by a post-transcriptional mechanism. A group of cellular genes whose expression is increased or decreased in cells expressing exogenous c-myc have also been identified, but the mechanisms involved in their regulation have not yet been addressed. They include class I HLA, histone H1 and H1-var.1, as well as the 3C477 and 3CH92 G0/G1 transition genes. N-myc has also been shown to suppress the expression of MHC class I genes, apparently by modulating MHC class I enhancer binding activities. Since the regulation of transcription and posttranscriptional processes seems to represent an integral part of c-myc function, it will be important to investigate the molecular role c-myc plays in these mechanisms.

**Structural and functional domains of the c-myc protein**

The identification of functional domains of the c-myc protein have provided some intriguing clues which should help our understanding of the structural basis for its activities. The transformation and autosuppression functions of c-myc have recently been thorough-

![Figure 1](image-url)
nuclei of cells. In addition, the region responsible for the non-specific DNA binding activity of c-myc has been mapped to amino acids 290-318 (see ref 110; Figure 1). However, this activity is dispensable for cotransformation and therefore its biological relevance remains unclear. Moreover, the significance of c-myc phosphorylation, particularly at the two casein kinase II-specific sites within the acidic and carboxyl-terminal regions of the protein has not yet been fully explored but may be involved in c-myc regulation.

Structural analysis of c-myc has revealed some unique features of the protein molecule. Using a computer algorithm of Taylor and Thornton, the 439 amino acid human c-myc protein has been modelled into an amino-terminal α-helix/β-sheet domain, a predominantly α-helical segment at the carboxyl-terminus, and a less structured 'hinge' region approximately in the centre of the protein. Immediately to the carboxyl-side of the hinge is a region strikingly rich in acidic residues. Large deletions in this area do not significantly affect the activities of c-myc in cotransformation of rat embryo fibroblasts or autosuppression, but similar mutations in v-myc oncogenes abolish their ability to transform chicken macrophages. This loss of c-myc function suggests that the acidic region may be required for transformation in certain cell-types. Finally, comparison of sequences within the myc gene family has revealed stretches of maximal homology which presumably encode functionally important regions conserved throughout evolution.

Indeed, the two domains required for c-myc transformation and autsuppression activities lie within the amino-terminal half and at the carboxyl-end of the protein, which contain such conserved regions. However, not all of the conserved regions are required for the known activities of c-myc, further suggesting additional functions have yet to be identified (Figure 1).

Homologies to 'leucine zipper' and 'helix loop helix' proteins

The carboxyl terminal region of the myc protein has attracted particular interest since it contains three structural motifs which were first identified in several DNA binding proteins. In these proteins, two of these structures, the 'leucine zipper' (LZ), and the 'helix loop helix' (HLH) can promote protein:protein dimerisation (refs 117, 119-121; R. Davis, A. Lassar and H. Weintraub, personal communication), while the third, a positively charged domain, can mediate sequence-specific DNA binding (refs 122-124; R. Davis, A. Lassar and H. Weintraub, personal communication). Most of the proteins which contain either of the dimerisation domains also possess a basic domain, and thus have been referred to as basic 'zipper' or 'basic helix-loop-helix' proteins. Regional alignments of the myc family proteins with the homologous regions of the DNA binding proteins are shown in Figure 2. In human c-myc, the basic region is contained within amino acids 354-369, the 'helix loop helix' motif from residue 370 to 406 and the 'leucine zipper' motif from residue 406 to the carboxyl terminus.

Many 'leucine zipper' containing proteins have now been identified; most are transcription factors, of which the best characterised include the CCAAT/enhancer binding protein (C/EBP) and the proto-oncogenes c-fos and c-jun. (For a review of 'leucine zipper' proteins, see chapter 1 and references therein.) These proteins form highly specific homo- or heterodimeric complexes before binding DNA. The structure responsible for this dimerisation, the 'leucine zipper', is an approximately 30 amino acid helix with leucines at every seventh position. As a consequence of this spacing, hydrophobic side chains are present along one face of the helix. Dimerised helices associate in vitro in a parallel conformation with these hydrophobic faces juxtaposed, in a structure analogous to the coiled coil found in multimeric, filamentous proteins such as keratins or lamins. The specificity and stability of the helix:helix interactions are apparently determined by the leucines in combination with the many charged residues found within the zipper. The myc proteins contain LZ motifs at their carboxyl termini, as they possess four or five appropriately spaced leucines interspersed with many charged residues (Figure 2).

The basic HLH proteins also appear to be sequence-specific DNA binding proteins and have been implicated in the developmental control of gene expression. Representative members of this family include the myogenic control protein MyoD, the immunoglobulin enhancer binding proteins E12 and E47 and their Drosophila homolog daughterless (da), as well as the proteins of the Drosophila achaete-scute complex (AS-C). These proteins can form homo- or heterodimers, and, as with the LZ containing proteins, a strong
selectivity exists for which proteins will pair.\textsuperscript{118,128} The dimerisation domain is characterised by two predicted amphipathic helices containing hydrophobic residues at every third or fourth position. Perhaps significantly, the amino acids at specific positions within these helices are highly conserved (Figure 2) (ref 118; R. Davis, A. Lassar and H. Weintraub, personal communication). The helices are separated by a region which usually contains one or more helix breaking residues, and has been modelled as a loop. This class of structures is ideally shaped like the Greek letter \( \Omega \), with the end residues pinched closely together.\textsuperscript{118,130} The helices within the dimeric complexes may form a structure similar to a coiled coil, but little experimental detail about the nature of this interaction is available. The myc proteins, as shown in Figure 2, contain a region which has all the hallmarks of an HLH motif, having two potential amphipathic helices punctuated by the appropriate hydrophobic amino acids, which are separated by a proline-containing non-helical segment.\textsuperscript{118}

Both the basic HLH and basic 'zipper' proteins can bind to DNA through a stretch of residues immediately amino terminal to their respective dimerisation domains. Contained within a region of 15-20 amino acids in both groups of proteins are two basic clusters which are required for DNA binding (refs 119, 124, 125, 137; R. Davis, A. Lassar and H. Weintraub, personal communication). This domain determines the DNA sequence specificity in the basic LZ proteins, as it presumably does in the basic HLH proteins.\textsuperscript{112,124, 131} The amino acid conservation in this region is highest within the proteins of either the LZ or HLH class, and conservation is also evident between the two groups, particularly in the distribution of positively charged residues.\textsuperscript{138} The myc proteins also contain a basic region with a similar distribution of positively charged residues, showing greater overall homology with the basic domains of the HLH proteins than the LZ proteins. It is situated immediately amino terminal to the HLH domain, and thus is separated by about 50 amino acids from the LZ structure (Figure 2).

Implications for c-myc function

The implications of the homologies to HLH and LZ proteins for potential myc activity are profound. Myc proteins possess what seem to be two dimerisation motifs, a leucine zipper and a helix loop helix, in a region that is necessary for both transformation and autoregulatory activity. It also contains a basic region which shows marked homology with the DNA binding domains of both of these classes of proteins (Figure 1).\textsuperscript{138} It does not seem unreasonable to postulate that myc protein may interact with itself or different proteins in multiple complexes through either the HLH or LZ domains. The myc protein in these complexes may then acquire specific nucleic acid binding potential or other activities which would be dictated by its particular molecular context.

While the circumstantial evidence is consistent with myc complex formation, there is very little direct evidence in support of these notions. Many studies have investigated the disposition of myc protein in cellular extracts and none have reported that it consistently and convincingly coprecipitates with any proteins or exists in multimeric form. In addition, in vitro translated protein has not been found to form either homodimeric complexes or heterodimers with members of either the LZ or HLH protein classes, under conditions in which these other proteins will readily multimerize (refs 119, 128; T. Littlewood and G. Evan, personal communication). In contrast, bacterially expressed protein has been found to exist in homotetrameric form, with the leucine zipper apparently mediating dimerisation, and the HLH domain the dimerisation of these dimers.\textsuperscript{139} However, the crucial factors which allow the bacterial protein to complex have yet to be identified. There are also reports of a protein which can be chemically cross-linked to v-myc\textsuperscript{140} and specific DNA sequences which can be coprecipitated with c-myc,\textsuperscript{55,56} though the nature and biological significance of these interactions awaits further elucidation.

If myc does exist in different homo- or heteromeric complexes in vivo, then, by analogy with the LZ and HLH containing proteins, their formation may be highly regulated in several ways. For instance, the presence of one or more of the constituent proteins may be limiting. Indeed, the missing activity in the myc autosuppression insensitive cells discussed previously may be one such partner. Alternatively, the regulation may be more subtle, requiring, for instance, covalent post-translational modification(s) to enable the protein:protein interactions. The interactions could also conceivably be highly transient or of low affinity. Such possibilities may explain why it has been so difficult to detect myc-containing complexes in vitro.

There is accumulating evidence that c-myc is a multi-functional protein. The ability of c-myc to regulate transcription correlates well with its transforma-
tion activity; and it is possible that at least in part c-myc may stimulate proliferation by suppressing cellular genes which restrict cell growth. In addition, the involvement of c-myc in the control of DNA replication in the absence of transcription may represent a distinct component of c-myc function. Nevertheless, our present knowledge of how c-myc controls all these activities at the molecular level is clearly limited. It will be interesting to discover whether c-myc protein does indeed form complexes and act through specific nucleic acid sequences in order to exert its effects on the control of proliferation and differentiation.

Acknowledgements

We would like to thank D. Bentley, G. Evan, N. Jones, J. Rothbard and M. Sternberg for helpful discussions. We would also like to thank those colleagues who have communicated data prior to publication. LJZP was supported by the Medical Research Council of Canada.

References

12. Palmieri S, Kahn P, Graf T (1983) Quail embryo fibroblasts transformed by four v-myc-containing virus isolates show enhanced proliferation but are non tumorigenic. EMBO J 2:2985-2989
30. Cleveland JL, Jansen HW, Bister K, Fredrickson TN, Morse III HC, Ihle JN, Rapp U (1986) Interaction between raf and...
64. Houweling A, van den Elsen FJ, van der El Aj (1980) Partial transformation of primary rat cells by the left most 4.5% fragment of adenovirus 5 DNA. Virology 105:537-550
transcriptional regulation in transformation. EMBO J 6:2053-2060


129. Nucleic Acids Res 15:6419-6436


158. Braun T, Buschhausen DG, Bober E, Tannich E, Arnold HH (1989) A novel human muscle factor related to but distinct from MyoD1 induces myogenic conversion in 10T1/2 fibroblasts. EMBO J 8:701-709

159. Wright WE, Sassoon DA, Lin VK (1989) Myogenin, a factor regulating myogenesis, has a domain homologous to MyoD. Cell 56:607-617


Domains of Human c-myc Protein Required for Autosuppression and Cooperation with\textit{ ras} Oncogenes Are Overlapping

LINDA J. Z. PENN,† MARY W. BROOKS, EDWARD M. LAUFER, TREvor D. LITTLEWOOD, JAY P. MORGENSTERN, GERRARD I. EVAN, WILLIAM M. F. LEE, and HARTMUT LANDI

Growth Control and Development Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London, WC2A 3PX, and Biochemistry of the Cell Nucleus Laboratory, Imperial Cancer Research Fund, Dominion House, London, EC1A 7BE, United Kingdom, and Hematology/Oncology Section, Department of Medicine, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received 5 February 1990/Accepted 4 June 1990

Amino acids 106 to 143 and 354 to 433 of the human c-myc protein (439 amino acids) were shown to be required for the protein to suppress c-myc gene transcription and were found to exactly overlap with those necessary for c-myc to cooperate with \textit{ras} oncogenes in the transformation of rat embryo fibroblasts. The essential carboxyl-terminal region harbors structural motifs (a basic region, a helix-loop-helix motif, and a "leucine zipper") which, in other proteins, can mediate dimerization and sequence-specific DNA binding.

Recently it has been shown that c-myc down regulates the initiation of its own transcription through a putative homeostatic mechanism. The extent of this suppression is proportional to the concentration of c-myc protein and is observed within the concentration range of c-myc protein found in a normal cell. In fact, it was found that 1,000 to 2,000 molecules must be expressed in Rat-1 cells in order to approach at least half-maximal suppression of endogenous c-myc RNA (16). Moreover, suppression of endogenous c-myc transcription by \textit{v-myc} (3, 16) and down regulation of c-myc expression in response to expression of exogenous N-myc genes have been observed (3, 6, 17; L. J. Z. Penn and M. W. Brooks, unpublished data), suggesting that the ability to repress gene expression may be conserved throughout the \textit{myc} gene family.

In order to investigate which regions of the human c-myc protein are required for autosuppression, we have mapped the c-myc protein for this function by using a subset of the c-myc in-frame deletion and insertion mutants first described by Stone et al. (19) (Fig. 1A). The mutants were introduced and expressed in the Rat-1 cell line via the replication-incompetent retroviral vectors pDORneo (Fig. 1B) and pMV6 (19), which contain the neomycin resistance gene (see Table 1). The respective c-myc-derived mutants were transfected into \textit{Ψ}-2 cells (13), and the resultant helper-free ecotropic virus particles were used to infect early-passage Rat-1 cells. Drug-resistant Rat-1 cell colonies (150 to 250) were subsequently pooled and harvested as subconfluent proliferating cell populations. These cells were analyzed in parallel for the expression of endogenous c-myc RNA and exogenous c-myc wild-type or mutant protein (for methods, see reference 16).

To determine the capacity of each of the c-myc mutants to induce suppression of endogenous c-myc RNA expression, RNA was prepared from the Rat-1 cells (16) which had been infected with retroviruses carrying the human c-myc gene, an exon II and III-specific c-myc cDNA, mutant c-myc genes (Table 1), or the neomycin resistance gene alone. The level of endogenous c-myc expression was subsequently determined by RNase protection, using probes to detect exon I-specific sequences of endogenous rat c-myc RNA as well as rat glyceraldehyde-3-phosphate dehydrogenase RNA (Fig. 2; for details on methods, see reference 16). c-myc RNA expression from P2, the major site of transcription initiation in rat cells, is shown in Fig. 2, although c-myc RNA transcripts initiating from both start sites in exon I were detectable and were found to be similarly regulated (16; other data not shown). The assays also showed that wild-type (16) or mutant (data not shown) c-myc proteins expressed from either the genomic (exons I, II, and III) or cDNA (exons II and III) constructs suppressed endogenous c-myc expression with similar efficiency (Table 1).

Most of the human c-myc protein mutants remained competent to down regulate the expression of endogenous c-myc RNA. However, mutations in either of two regions of human c-myc rendered the protein inactive. One region was defined by the inactive deletion mutant D106-143 and the two functional insertion mutants, In 105 and In 144, which flank this critical region on either side. The second section required for autosuppression was localized to the carboxyl end of c-myc and included amino acids 354 to 433. All six mutations within this region of c-myc resulted in mutant protein which was inactive for myc autosuppression activity. The left-hand boundary of this domain was marked by the active deletion mutant D265-353. However, the right-hand boundary is less well defined, since the most distal insertion mutant, In 434, demonstrated only 50% wild-type levels of activity and the effect of mutations between amino acids 435 and 439 at the carboxyl terminus have not been tested.

In conclusion, we consider the regions containing amino acids 106 to 143 and 354 to 433 to be essential for human c-myc protein to suppress endogenous c-myc RNA expression in Rat-1 cells (Fig. 2 and Table 1). For complete evaluation of the results described above it was critical to determine that the c-myc proteins expressed from the retroviral promoter corresponded to the expected mutations and that these proteins were expressed at sufficiently high levels to observe suppression of the endogenous c-myc gene. Therefore, the mutant c-myc proteins were
analyzed by immunoblotting (7, 16) whole-cell extracts from polyclonal Rat-1 cell populations expressing each of the mutant proteins. A rabbit pan-myc polyclonal antibody was used (xmyc-1; 8) (Fig. 3). All mutant proteins were detectable and migrated on sodium dodecyl sulfate-polyacrylamide gels according to the expected size for each particular mutation (19). Mutant D41-53 was the only exception, and it was undetectable because it had undergone a deletion of the epitope to which the myc antibody used in these experiments was raised. Together with restriction analyses of the corresponding plasmid DNAs (data not shown), these experiments confirmed that the mutant myc protein products did indeed correspond to the respective mutations of the c-myc gene. In addition, extracts of the different Rat-1 cell populations, each expressing a c-myc protein mutant, were analyzed quantitatively by a myc-specific enzyme-linked immu-

TABLE 1. Properties of human c-myc mutants

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Presence of:</th>
<th>Exogenous c-myc protein expression in:</th>
<th>c-myc activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pMV6 cDNA, exons II and III</td>
<td>pDOR gene, exons I, II, and III</td>
<td>Immunoblot</td>
</tr>
<tr>
<td>Wild type</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>In 6</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>D7-38</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>D41-53</td>
<td>+</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>In 55</td>
<td>+</td>
<td>+</td>
<td>2,000</td>
</tr>
<tr>
<td>D55-92</td>
<td>+</td>
<td>+</td>
<td>5,850</td>
</tr>
<tr>
<td>D93-103</td>
<td>+</td>
<td>+</td>
<td>8,000</td>
</tr>
<tr>
<td>In 105</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>D106-143</td>
<td>+</td>
<td>+</td>
<td>3,050</td>
</tr>
<tr>
<td>In 144</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>In 203</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>In 262</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>D265-317</td>
<td>+</td>
<td>+</td>
<td>3,750</td>
</tr>
<tr>
<td>D265-353</td>
<td>+</td>
<td>+</td>
<td>2,650</td>
</tr>
<tr>
<td>D265-367</td>
<td>+</td>
<td>+</td>
<td>8,900</td>
</tr>
<tr>
<td>In 370</td>
<td>+</td>
<td>+</td>
<td>3,200</td>
</tr>
<tr>
<td>In 373</td>
<td>+</td>
<td>+</td>
<td>2,750</td>
</tr>
<tr>
<td>D371-412</td>
<td>+</td>
<td>+</td>
<td>7,650</td>
</tr>
<tr>
<td>In 413</td>
<td>+</td>
<td>+</td>
<td>3,550</td>
</tr>
<tr>
<td>D414-433</td>
<td>+</td>
<td>+</td>
<td>8,050</td>
</tr>
<tr>
<td>In 434</td>
<td>+</td>
<td>+</td>
<td>3,000</td>
</tr>
</tbody>
</table>

a D, In-frame deletion; In, in-frame linker insertion.
b ND, Not determined because of deletion of epitope to which myc peptide antibody was raised; ELISA, enzyme-linked immunosorbence assay.
c See reference 19.
FIG. 2. Expression of endogenous rat c-myc RNA in response to the expression of wild-type (w.t.) and mutant human c-myc genes. RNAs (10 μg) from Rat-1 cells infected with control retrovirus and with retrovirus carrying wild-type or mutant human c-myc genes were analyzed by RNase protection using single-stranded RNA probes complementary to endogenous rat c-myc exon 1 and glyceraldehyde-3-phosphate dehydrogenase (GADPH)-specific sequences. c-myc RNA expression from P2, the major site of transcription initiation in rat cells, is shown. Expression of exogenous human c-myc RNA was not detectable under the assay conditions. The RNA analyzed was prepared from polyclonal Rat-1 cell populations, with the exception of the cells expressing mutants In 105 and D145-262, from which RNAs isolated from representative clones are shown (see text also). Following electrophoresis through a 6% denaturing polyacrylamide gel, the protected probe was visualized by autoradiography. Each of the mutant c-myc constructs was assayed in at least three independent experiments and consistently yielded identical results.
FIG. 3. Qualitative analysis of human c-myc protein mutants. Extracts from polyclonal Rat-1 cell populations expressing exogenous c-myc genes were resolved by electrophoresis through a 10% sodium dodecyl sulfate-polyacrylamide gel and analyzed by immunoblot using a rabbit pan-myc polyclonal antiserum and 125I-labeled protein A. Bands corresponding to human c-myc proteins were visualized by autoradiography. Minor bands visible in lanes DI06-143, D371-412, and In 413 correspond to proteolytic products of human c-myc protein. The endogenous rat c-myc protein is expressed at levels (16) below the limits of detection of this experiment. Protein size standards are shown in kilodaltons.

FIG. 4. Expression levels of wild-type and mutant c-myc proteins in Rat-1 cells. Extracts from Rat-1 cells expressing exogenous wild-type and mutant c-myc genes were analyzed by a myc-specific enzyme-linked immunosorbence assay. Bars represent the average of duplicate assays in which the number of molecules of c-myc protein per cell was determined by comparison with an internal purified myc protein standard. The results shown were derived from retrovirus-infected polyclonal populations of cells, with the exception of In 105 and DI45-262, for which several clones were analyzed. Results representing one of each of these clones are shown (see text also).

This analysis demonstrated that nearly all of the c-myc mutants were expressed at or above 2,000 molecules of c-myc protein per cell and therefore could readily be scored for their myc autosuppression function. However, the expression levels of two mutants (In 105 and DI45-262), as measured in polyclonal cell populations, fell below this mark, and concomitant down regulation of endogenous c-myc expression was not observed (data not shown). Further analysis of several Rat-1 cell clones (e.g., as in Fig. 4) revealed that in those cells expressing elevated levels of mutant c-myc proteins In 105 and DI45-262, the endogenous c-myc gene was consistently found to be suppressed (e.g., as in Fig. 2).

This analysis has led to the identification of two regions in the human c-myc protein which are essential for negative autoregulation of c-myc expression (Fig. 5). The first region resides in the amino half of the protein between amino acids 106 and 143 and contains sequences between amino acids 129 and 143 which are highly conserved among the members of the myc gene family (12, 18), which suggests that the functions of myc proteins associated with this region have been conserved throughout evolution. Database searches with the peptide sequence of this region did not detect significant homology with any defined structural polypeptide motif or known protein outside the myc family (data not shown).

The second region involved in c-myc autosuppression is also conserved within the myc gene family. It is located in the carboxy-terminal region of the c-myc protein between amino acids 354 and 433 and contains one of the two nuclear localization signals identified in the human c-myc protein (residues 364 to 374) (4). In addition, three motifs, a basic region (residues 355 to 369), a helix-loop-helix structure (residues 370 to 406), and a "leucine zipper" (residues 406 to 439 at the carboxyl terminus) associated with DNA-binding proteins have recently been identified within this region of c-myc (11, 15). The presence of these motifs in the functionally essential carboxyl end of c-myc suggests that myc proteins may form complexes and possibly thereby bind specific DNA sequences. In fact, bacterially derived myc protein can form homotetramers through its carboxyl end in vitro (5), and it has recently been suggested that c-myc can bind directly to a specific DNA sequence (1).

In contrast to our results, an in-frame deletion of 108 amino acids of v-myc which corresponds to the region of c-myc at the exon II-exon III border was shown to be inactive for c-myc down regulatory activity (3). We showed that this region in c-myc, approximately corresponding to amino acids 145 to 262, could be deleted while the c-myc protein remained fully functional for its autosuppression function. This discrepancy may be due to intrinsic differences between v-myc and c-myc, the specific amino acids deleted, or the rodent fibroblast cell lines used in the two studies. Alternatively, the v-myc deletion mutant may have been expressed at subthreshold levels to detect c-myc suppression. Indeed, we found it necessary to identify cell clones expressing high levels of the c-myc mutant protein DI45-262 in order to detect suppression of the endogenous rat c-myc gene.

Regions other than amino acids 106 to 143 and 354 to 439 of the c-myc gene are dispensable for the myc autosuppression function. Nevertheless, some of these regions, such as amino acids 45 to 68 of c-myc, are very highly conserved among the myc gene family, which suggests that this domain may be important for as yet unknown function of the myc protein (Fig. 5). Similarly, the highly conserved acidic region of the myc proteins (Fig. 5) was shown in v-myc to be required for chicken hematopoietic cell transformation but unnecessary for the transformation of chicken embryo fibroblasts (2, 9). These results suggest that myc activities can be cell-type dependent or may rely on additional and possibly tissue-specific cellular factors.
It is intriguing that the functional domains of \( c-myc \) required for autosuppression overlap with those essential for \( H-a-ras \) cotransformation of secondary rat embryo fibroblasts (19). Indeed, the results presented here indicate a correlation between the transforming function of \( c-myc \) and its ability to suppress gene transcription. Both these activities are completely lost in the mutants affecting amino acids 106 to 143 or 354 to 433 (Fig. 5). Therefore, it seems possible that cellular transformation by \( c-myc \) involves the repression of cellular genes which act to restrict cell proliferation.

We thank Shaun Dawkins for technical assistance.

L.J.Z.P. was supported by a postdoctoral fellowship from the Medical Research Council of Canada.

LITERATURE CITED

15. Rosenbaum, H., E. Webb, J. M. Adams, S. Cory, and A. W.

---

FIG. 5. Structural and functional domains of human \( c-myc \) protein. Secondary-structure predictions based on amino acid sequence information suggest that the 439-amino-acid nuclear phosphoprotein may be composed of an \( \alpha \)-helix-\( \beta \)-sheet domain (amino acids 1 to 203) and a predominantly \( \alpha \)-helical domain at the carboxyl end (amino acids 238 to 439) which are separated by a less-structured hinge region (amino acids 204 to 237). Additional structural motifs include a highly acidic domain (\( \text{EH} \)), a basic region (\( \text{EL} \)), a helix-loop-helix domain (\( \text{II/III} \)), and a leucine zipper (\( \text{ENa} \)). Regions most highly conserved among the members of the \( myc \) gene family (\( \text{EH} \)) are also indicated. The functionally essential domains of \( c-myc \) identified to date are represented by the solid black boxes in the specified schematic diagrams and map as follows: autosuppression, residues 106 to 143 and 353 to 433; cooperation with \( ras \) oncogenes in rat embryo fibroblasts, residues 106 to 143 and 353 to 433; nuclear localization, residues 320 to 328 and 364 to 374; and nonspecific DNA binding, residues 290 to 318. II/III, Border of exon II- and exon III-encoded sequences.


Negative autoregulation of c-myc transcription

Linda J.Z. Penn, Mary W. Brooks, Edward M. Lauffer and Hartmut Land

Growth Control and Development Laboratory, Imperial Cancer Research Fund, Lincoln's Inn Fields, London WC2A 3PX, UK

Present address: The Hospital for Sick Children, Department of Microbiology, 555 University Ave., Toronto, Ontario MSG 1X8, Canada

Communicated by R. Treisman

The introduction of activated c-myc and v-myc genes into a variety of non-established and established cells results in the suppression of endogenous c-myc expression. As measured in Rat-1 fibroblasts, the suppression occurs at the level of transcriptional initiation. Moreover, the extent of the down-regulation is proportional to the cellular concentration of c-myc protein, and the critical concentration range in which the endogenous c-myc RNA is effectively suppressed corresponds to that found in non-transformed cells. In addition, the autoregulatory mechanism is not only dependent on c-myc protein, but also requires additional trans-acting factors. These results suggest a negative feedback mechanism can act as a homeostatic regulator of c-myc expression in vivo.

Key words: feedback/somatic cell hybrid/oncogene-suppression/transformation

Introduction

The activation of the c-myc proto-oncogene is implicated in the genesis of a wide range of tumours and can involve such diverse mechanisms as amplification, translocation, promoter insertion or retroviral transduction (reviewed in Cole, 1986; Kelly and Siebenlist, 1986). These alterations lead to constitutive expression of the c-myc gene which is sufficient to cause oncogenic activation in vitro and in vivo (Lee et al., 1985; Stewart et al., 1984; Adams et al., 1985; Land et al., 1986). To understand the role of c-myc in tumorigenesis, its function and the mechanisms which govern its expression have been intensively investigated.

c-myc is a nuclear phosphoprotein with potent biological properties. Its presence seems to be required for proliferation, while its absence may be obligatory for complete withdrawal from the cell cycle. When constitutively expressed, c-myc can extend the self-renewal capacity and block the terminal differentiation of various cell types which ultimately may lead to their immortalization. Moreover, c-myc can cooperate with ras-like oncogenes to fully transform primary rodent embryo cells (reviewed in Penn et al., 1990). However, the molecular basis for these effects remains unknown. Similarly, the exact mechanisms by which c-myc is regulated are still obscure, although it has become clear that expression is controlled by a complex interplay of regulatory mechanisms operating at levels involving transcriptional initiation and elongation, as well as mRNA stability (reviewed in Marcu, 1987).

An approach to explore important functional as well as regulatory aspects of c-myc became apparent when evidence was found that c-myc might be involved in the negative control of its own expression. Such a mechanism was postulated for the first time when it was found in murine plasmacytomomas and Burkitt's lymphomas that the expression of the translocated c-myc allele is constitutive, while the expression of the normal c-myc allele is often undetectable (Adams et al., 1983; Dunnick et al., 1983; Leder et al., 1983). The expression of the normal c-myc gene is similarly suppressed in a variety of other spontaneously arising tumours in which one allele has been activated (e.g. Bernard et al., 1983; Nishikura et al., 1983; Shih et al., 1984; Forrest et al., 1987), and in tumour cells derived from transgenic mice harbouring a constitutively active myc gene (e.g. Adams et al., 1985; Rosenbaum et al., 1989; Dildrop et al., 1989). In addition, the expression of c-myc, N-myc or v-myc genes from heterologous promoters also leads to the suppression of the endogenous c-myc gene in various cell types (e.g. Rapp et al., 1985; Morse et al., 1986; Schwartz et al., 1986; Cory et al., 1987; Dean et al., 1987; Lombardi et al., 1987; Poter et al., 1987; Cleveland et al., 1988; Clynes et al., 1988; Mango et al., 1989). These observations suggest myc negative feedback may be quite common. However, the mechanism of these presumed autoregulatory events has not been thoroughly explored and may not be universal, though initiation of c-myc transcription has been reported to be suppressed by v-myc (Cleveland et al., 1988). Indeed, it has been suggested that in Burkitt's Lymphoma the normal c-myc allele may have been permanently silenced by other means, such as alterations in chromatin structure or methylation (Siebenlist et al., 1984; Dunnick et al., 1985; Nishikura and Murray, 1988). Further evidence questioning the existence of myc autosuppression stems from reports showing that in several cells, the expression of exogenous myc genes does not correlate with the suppression of endogenous c-myc expression (e.g. Keath et al., 1984; Coppola and Cole, 1986; Dmitrovsky et al., 1986; Kelekar and Cole, 1987; Stone et al., 1987; Zerlin et al., 1987).

In order to clarify some of the ambiguities mentioned above we overexpressed c-myc or v-myc genes in a series of non-established and established rodent fibroblasts and found endogenous c-myc expression to be suppressed. Subsequently, we have determined that the suppression of the endogenous c-myc gene occurs at the level of transcriptional initiation and that the degree of this suppression is proportional to physiological concentrations of c-myc protein. Moreover, analysis of a non-responsive cell line indicated...
that additional trans-acting cellular factor(s) are required for negative feedback regulation. These results indicate that c-myc protein can act as a modulator of its own gene transcription and are consistent with the concept that a negative feedback mechanism plays a role in the regulation of the c-myc gene in normal cells.

Results

Expression of exogenous c-myc and v-myc genes leads to the down-regulation of endogenous c-myc expression in rodent fibroblasts

To determine if endogenous c-myc RNA expression was suppressed in cells constitutively expressing c-myc protein, the Rat-1 cell line as well as secondary rat and mouse embryo fibroblasts were infected with a retrovirus (DORhe-myc I, II, III) carrying the three exons of the human c-myc gene together with the neomycin resistance gene (neo') or with a control retrovirus (DORneo) carrying the neo' gene alone. Drug resistant colonies of the Rat-1 cells were pooled while those from the secondary cell cultures were cloned, and then analysed as early-passage, sub-confluent proliferating cells. Analysis of endogenous c-myc gene expression by Rnase protection (for probe details see Figure 1d) showed that those cells expressing an exogenous c-myc gene had an ~5-10-fold suppressed level of endogenous c-myc expression compared with control levels (Figure 1a; compare lanes 2, 3, and 4 with lane 1).
the following experiments.

To ascertain whether intact exogenous c-myc protein was required to trigger the down-regulation of the endogenous gene, the effects of wild-type and mutant c-myc were compared. Rat-1 cells were infected with retroviruses carrying the human c-myc cDNA of coding exons II and III (MV6hc-myc2,3), a c-myc cDNA which had been mutated by the introduction of a frameshift mutation at amino acid 48 (MV6hc-myc2,3fs), or a truncated c-myc gene lacking nearly all exon III derived sequences (DORAhc-myc), or infected with a control retrovirus (MV6bneo). Expression of the human c-myc cDNA in the pooled populations of Rat-1 cells resulted in the suppression of endogenous c-myc RNA, whereas expression of mutant c-myc genes did not alter the endogenous c-myc expression level, showing that intact, c-myc protein is required to down-regulate endogenous c-myc expression (Figure 1b).

To determine if the expression of an exogenous v-myc gene could also repress c-myc, a variety of fibroblast cells were infected with a retrovirus carrying the v-gag-myc gene from the MC29 strain of avian myelocytomatosis virus (DoKymyc) or a control retrovirus (DORneo). The resultant pooled populations of Rat-1, Fischer Rat 3T3, KPREF, A31 BALB/3T3, Swiss 3T3 and C3H10T1/2 cells showed that the introduction of a v-myc gene resulted in the down-regulation of endogenous c-myc expression (Figure 1c).

To ensure the products of the exogenous v-myc gene, the human c-myc gene (exons I, II and III) and the human c-myc cDNA (exons II and III) were expressed, myc protein was detected by immunoblotting whole cell extracts (data not shown). Densitometric scanning of these immunoblots further demonstrated that this exogenous myc protein was expressed at 8- to 12-fold higher levels than endogenous c-myc protein in control Rat-1 cells (for c-myc see Figure 4a).

Similarly, human c-myc protein was shown to be expressed in cell clones derived from secondary rodent fibroblast cultures by a myc specific enzyme linked immunosorbence assay (ELISA; data not shown).

Thus the myc down-regulation mechanism is dependent on the expression of intact c-myc or v-myc protein and is operative in a variety of rodent fibroblasts. The specific activities of the three different myc gene products derived from human c-myc exons I, II and III, or II and III, and v-myc genes, used in the suppression of endogenous c-myc expression, were very similar in Rat-1 cells which allowed these exogenous myc genes to be interchangeably used in the following experiments.

Suppression of transcriptional initiation

To determine if the suppression of c-myc expression was regulated at the level of transcription, the activity of the endogenous c-myc gene was determined in the absence or presence of an exogenous c-myc or v-myc gene by nuclear run-on assays. Nuclei were prepared from Rat-1 cells infected with retroviruses carrying genes encoding either intact or truncated human c-myc proteins or with a control retrovirus carrying only a neo gene. Run-on transcription assays were performed in the presence (Figure 2a; lanes 2, 4 and 6) or absence (Figure 2a; lanes 1, 3 and 5) of α-amanitin, an inhibitor of RNA polymerase II activity. Subsequently, the 32P-labelled RNA transcripts were hybridized with filters containing single-stranded DNA probes (Figure 2c) corresponding to either strand of rat c-myc exon I and double-stranded DNA probes specific for γ-actin and ribosomal RNA. The latter two probes controlled for the efficacy of α-amanitin and the run-on reactions among the different nuclear preparations.

The experiment showed that RNA polymerase II specific transcription of the c-myc encoding sense strand was readily detectable in Rat-1 cells infected with the control neo virus (Figure 2a; lane 1) and that this transcription was clearly suppressed in cells expressing an exogenous c-myc gene (Figure 2a; lane 3). As expected, endogenous c-myc transcription was not affected in Rat-1 cells expressing the truncated human c-myc gene (Figure 2a; lane 5). A similar experiment demonstrated that an exogenous v-myc gene also induced suppression of the endogenous c-myc gene at the transcriptional level (Figure 2b). This transcriptional down-regulation is specific to the endogenous c-myc gene and does not affect transcription of the γ-actin gene (Figure 2a and b). Since the c-myc specific probe, by analogy with the mouse gene, corresponds to sequences upstream of the predicted attenuation site (Nepveu and Marcu, 1986); and overlapped P2, the major site of c-myc transcription initiation in rat cells (Hayashi et al., 1987; see also Figure 1c), it is...
within the concentration range of c-myc protein expressed in a normal cell. To determine if such a dose response relationship existed, the effects of differing levels of c-myc protein on endogenous c-myc gene expression were analysed. Rat-1 cells infected with a retrovirus carrying human c-myc cDNA were randomly cloned and their human c-myc protein concentrations and endogenous rat c-myc RNA levels were subsequently assayed in parallel by the myc ELISA and RNase protection, respectively. The relationship between these two parameters was found to be inversely proportional (Figure 3a and b). The endogenous c-myc RNA levels in control Rat-1 cells not expressing human c-myc protein were arbitrarily defined as 100%. Accordingly, cells expressing ~2500 or >5000 molecules of human myc protein per cell demonstrated ~20% and ~10% expression of endogenous c-myc, respectively. Similarly, a number of independent, pooled populations of Rat-1 cells infected with retrovirus carrying the human c-myc gene, expressed an average of 5000 molecules of human c-myc protein per cell and, as expected, expression of the endogenous c-myc gene was consistently reduced 10-fold (e.g. Figure 1b, lane 2). Thus, the extent of suppression in Rat-1 cells is proportional to myc protein concentration, so that in response to 1000–2000 molecules of c-myc protein per cell, c-myc RNA levels are expected to be reduced by 50%, whereas at ≥5000 molecules per cell ≥90% suppression is evident.

To determine the concentration of rat c-myc protein expressed in the Rat-1 cells, the level of endogenous c-myc protein was initially quantitated in sub-confluent exponentially proliferating cells. These levels could not be directly determined since the detection antibody used in the ELISA did not recognize rat c-myc protein. Instead, this protein could be detected by immunoblotting, using a myc-specific antibody raised to a peptide common to all c-myc proteins. The approximate amount of rat c-myc protein was determined by comparing signal intensities corresponding to cell extracts from Rat-1 cells and Rat-1 cells expressing a known concentration of exogenous human c-myc protein which had been previously quantitated by myc ELISA (Figure 4a). The indicated signal in Figure 4a, lane 2 corresponds to a concentration of 5000 human c-myc protein molecules per cell and densitometric analysis of the indicated human (Figure 4a, lane 2) and rat (Figure 4a, lane 1) c-myc protein signals demonstrated an 8-fold difference in their intensities. Therefore, the concentration of rat c-myc protein in exponentially proliferating cells amounted to ~600–700 molecules, at which expression of endogenous c-myc RNA is clearly demonstrable (e.g. Figure 1b, lane 1). Since the dose response curve (Figure 3a and b) was determined in such proliferating Rat-1 cells the contribution of rat c-myc protein to the total c-myc protein concentration is ~600–700 molecules per cell in control cells and is proportionately less in those cells expressing human c-myc protein.

To determine if endogenous c-myc protein in normal Rat-1 cells can accumulate to a level at which the autosuppression mechanism could show maximal effect, c-myc levels were quantitated before and after serum-stimulation of these cells. The myc protein in nuclear (Figure 4b) or whole cell (data not shown) extracts from quiescent and serum-stimulated Rat-1 cells, as well as varying amounts of COLO 320 cell extracts was immunoblotted. Having directly quantitated the human c-myc protein concentration in the COLO 320 cells by myc ELISA, the intensities of the immunoblot signals

very likely that in cells expressing intact exogenous myc protein, endogenous c-myc gene transcription is suppressed at the level of initiation rather than elongation.

**Autosuppression is concentration dependent**

The results showing that c-myc transcription can be suppressed in cells expressing an exogenous myc gene are consistent with the concept that a negative feedback mechanism plays a role in the normal regulation of c-myc expression. If so, the suppression of c-myc should be dependent on the concentration of its gene product, and the number of c-myc molecules required for the mechanism to function must fall
corresponding to the COLO 320 and Rat-1 cells were compared and the amount of rat c-myc protein determined. From these experiments we can conclude that the expression of c-myc protein in Rat-1 cells is clearly elevated to at least 5000 molecules/cell between 1 and 2 h following serum stimulation (Figure 4b).

Serum treatment of Rat-1 cells induces a rapid increase in c-myc RNA expression which peaks 1 h following serum-stimulation and is then down-regulated to basal level expression within 8—16 h (Figure 4c). It is intriguing that the c-myc protein level in Rat-1 cells is correspondingly elevated 1—2 h post-serum (Figure 4b) to a concentration range which can significantly reduce the amount of c-myc RNA expressed (Figure 3a and b). However, the lower levels of c-myc protein in sub-confluent proliferating cells suggest negative feedback would have relatively less effect during this phase of growth (Figure 4a). These observations show that the quantitative requirements for a negative feedback mechanism to be active in vivo are met and suggest that this mechanism may contribute to the homeostatic regulation of c-myc RNA expression during the different phases of cellular proliferation.

Additional cellular factors are required for negative feedback

Unlike the many other cell lines tested (Figure 1), introduction of exogenous myc into NIH 3T3 cells did not result in the down-regulation of endogenous c-myc RNA expression (Figure 5a) although expression of the exogenous myc protein was clearly detectable by immunoblotting (data not shown). This failure to respond could be due to a mutation in cis such that the c-myc promoter is no longer sensitive to the regulatory factors involved in the feedback mechanism. Alternatively, this cell line may lack a critical trans-acting factor necessary for myc down-regulation.

To distinguish between these possibilities the non-responsive mouse NIH 3T3 cell line was fused with the responsive Rat-1 cell line to generate stable interspecific somatic cell hybrids. By this approach the responsiveness of the mouse c-myc promoter could be tested in the context of the hybrid cell background. The level of rat and mouse c-myc RNA expression was determined by RNase protection analysis of the stable dual-drug resistant hybrids that were generated by fusion of Rat-1 cells infected with a control retrovirus carrying the hygromycin resistance (hygro') gene and NIH 3T3 cells infected with a control retrovirus carrying the neo' gene; as well as Rat-1 cells infected with a retrovirus carrying the v-myc and hygro' genes and NIH 3T3 cells infected with a retrovirus carrying the v-myc and neo' genes. As expected, rat c-myc RNA expression was suppressed in hybrids expressing exogenous v-myc (Figure 5b; compare lanes 9—12 with 5—8). In addition, the expression of mouse c-myc RNA was similarly suppressed in the hybrid cells expressing exogenous v-myc (Figure 5b; compare lanes 9—12 with 5—8). This contrasts with the non-responsiveness of the mouse NIH 3T3 parental cells (Figure 5a) and of control somatic cell hybrids (see below and Figure 5b, lanes 3 and 4). Thus the mouse c-myc promoter in NIH 3T3 cells is capable of responding to autosuppression. This suggests the non-responsiveness of these cells may be due to the lack of functional trans-acting factor(s) required for the negative feedback mechanism.

To control for the fusion procedure itself altering the myc down-regulation potential of the cells, control hybridizations were performed by fusing the parental cells to themselves. For example, Rat-1 cells expressing the neo' or hygro' gene were fused and the hybrids selected for their resistance to both G418 and hygromycin. Similar interspecific fusions were conducted with v-myc expressing Rat-1 cells as well as NIH 3T3 cells which did or did not express v-myc (see
Fig. 5. Autosuppression of c-myc requires additional trans-acting cellular factors. RNA (10 μg) from cells infected with a control retrovirus (—) or with a retrovirus carrying the v-gag–myc gene (+) were analysed by RNase protection for the expression of endogenous c-myc (for probe details see Figure 1d) and GAPDH genes (using a rat GAPDH probe). The protected probe was resolved on denaturing polyacrylamide gels and visualized by autoradiography. (a) Protected RNA from NIH 3T3 cells. (b) Protected RNA from stable somatic cell hybrids generated from intraspecific fusion of Rat-1 cells (RAT-1/RAT-1), intraspecific fusion of NIH 3T3 cells (3T3/3T3) and interspecific fusion of Rat-1 and NIH 3T3 cells (Rat-1/3T3 HYBRIDS). RNA (10 μg) of these hybrids was also probed for the expression of exogenous v-gag–myc using an MC29 gag-specific probe.

Discussion

We have provided evidence that introduction and expression of c-myc or v-myc genes into a variety of rodent fibroblasts results in the suppressed initiation of endogenous c-myc transcription. The extent of this repression is proportional to the concentration of c-myc protein and the range of down-regulating concentrations corresponds to that found in non-transformed cells. Moreover, autoregulation is not only dependent on c-myc protein, but also requires additional trans-acting factor(s).

As shown by nuclear run-on transcription analysis, exogenous myc gene expression specifically induces the repression of RNA polymerase II-directed transcriptional initiation from the endogenous c-myc promoter. These results provide strong evidence that c-myc can function as a transcriptional modulator of cellular gene expression. In support of these conclusions the activity of a stably transfected c-myc promoter fusion construct can be suppressed by exogenous myc expression in Rat-1 cells (Laufer and Land, in preparation). The degree to which transcription attenuation (Bentley and Groudine, 1986; Eick and Bornkamm, 1986) and post-transcriptional mechanisms (Blanchard et al., 1985; Piechaczyk et al., 1985; Swartwout et al., 1987) may contribute to the down-regulation of c-myc by the feedback mechanism has not been directly investigated. However, their involvement is likely to be minimal, since endogenous c-myc gene expression and transcriptional initiation are repressed to a similar extent when activated c-myc genes are introduced into Rat-1 cells.

Quantitative analysis suggests that c-myc autosuppression may play a role in regulating c-myc expression in vivo. The negative feedback response is proportional to c-myc protein concentration, and the non-linear nature of this dose–response relationship shows that small increases of myc protein concentration can significantly reduce the expression of endogenous c-myc. Indeed, endogenous c-myc
The expression of the c-myc gene is regulated by a complex set of mechanisms (reviewed in Marcu, 1987; Piekaczynz et al., 1987). Among these, autosuppression may act as a regulator of c-myc transcriptional initiation. As a homeostatic response, negative feedback would not necessarily exclude the effects of other regulatory mechanisms, but would contribute to the overall level of c-myc expression. Indeed, serum induces endogenous c-myc expression in control cells and cells expressing an exogenous myc gene to the same extent. However, the absolute amount of endogenous c-myc RNA is significantly reduced in cells expressing an exogenous myc gene (Penn and Brooks, unpublished results).

The NIH 3T3 cells in our laboratory, which do not demonstrate suppressed c-myc expression after exposure to elevated levels of myc protein, were used as a tool to further investigate the mechanism of negative feedback. In somatic cell hybrids between these cells and autosuppression sensitive Rat-1 cells, the rat cell factors were able to induce suppression of mouse c-myc. While we cannot determine whether the autoregulatory deficiency in the NIH 3T3 cells is due to a mutation or whether it represents the normal state in this cell type, these experiments clearly showed that the mechanism of myc autosuppression requires not only the myc protein but also at least one additional cellular trans-acting factor. Moreover, the demonstration of this autoregulatory response in a wide range of cell lines, including early passage non-established cultures, combined with the demonstration that the c-myc gene of NIH 3T3 cells could be rendered responsive to elevated myc levels, further suggests that autosuppression is a normal control pathway for c-myc expression in vivo.

The precise roles of the c-myc protein and the other trans-acting factor(s) in the myc autosuppression mechanism are unknown. It remains unclear whether c-myc protein modulates gene transcription by binding directly to specific DNA sequences within the c-myc promoter or by a more indirect mechanism of action. The function of the trans-acting factor also remains obscure. However, there is potential for myc to form a complex with this factor since structurally, myc has many hallmarks of a protein which interacts with other proteins and/or DNA. The myc protein contains a basic domain homologous to DNA-binding regions, as well as leucine zipper and helix-loop-helix motifs through which it may complex with other proteins (Davis et al., 1987; Villares and Cabrera, 1987; Landschulz et al., 1988; Tapscott et al., 1988; Murre et al., 1989a,b; Prendergast and Ziff, 1989). Indeed, Dang et al. (1989) have shown bacterially derived myc protein can form homotetramers in vitro through the carboxyl-region of c-myc which contains these motifs. It is intriguing that these domains, which are conserved among members of the myc gene family, are required for both co-transformation (Stone et al., 1987) and autosuppression (Penn et al., in preparation).

Overexpression of the c-myc gene due to alterations affecting its regulation can lead to cellular immortalization and thereby contribute to the step-wise progression of tumorigenesis. The loss of a functional trans-acting factor required for the negative feedback mechanism could similarly result in such overexpression, and the corresponding mutation of this suppressor of c-myc expression might mark the locus of an 'anti-oncogene'. Indeed, mutations of this kind may play a part in the evolution of those cell lines in which an autoregulatory response is not found. However, the exact contribution of a disabled autosuppression mechanism to uncontrolled growth or the process of tumorigenesis requires further understanding of the factors involved in this pathway of c-myc gene regulation.

Materials and methods

Cell culture and retroviral infection

Secondary rat and mouse embryo fibroblasts were prepared from 13–14 day old fetuses (Land et al., 1983). The cell lines used in this study include: Rat-1, a subclone of the Fischer rat embryo fibroblast line F2408 (Lania et al., 1980); KPREP, a population of spontaneously immortalized rat embryo fibroblasts (L.J. Penn, unpublished results); FR 3T3 (Seif and Cazin, 1977); BALB/3T3 clone A31 (Aaronson and Todaro, 1968); SW 3T3 (Todaro and Green, 1968); C3H10T1/2 (Reznikoff et al., 1973); COLO 320, human colon adenocarcinoma (Quinn et al., 1979); and NIH 3T3 cells (laboratory stock).

All fibroblast cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS), 100 μg/ml kanamycin and 2 μg/ml gentamycin. COLO 320 cells were cultured in 5% FCS/RPMI 1640.

To produce infectious retroviral particles recombinant viral constructs were transfected into ψ-2 cells (Mann et al., 1983) and transiently produced ecotropic retrovirus harvested 24–48 h later. To create stable high titre virus producers, ψ-2 cells, which had been pretreated for 18 h with 0.03 μg/ml tunicamycin were then infected with the transiently produced retroviruses in the presence of polybrene (8 μg/ml), and subsequently selected for expression of either the neo' or hygro' genes carried by the viruses in 1 mg/ml G418 sulphate (Genetecin, Gibco) or 300 μg/ml hygromycin B (Sigma), respectively. Individual colonies were picked and screened for high titre virus production. Secondary embryo rodent cells and cell lines were infected either with retrovirus from the stable ψ-2-producer cell lines or from transient ψ-2 transfections, and selected in either 1 mg/ml G418 or 150 μg/ml hygromycin B. Unless otherwise mentioned, the drug-resistant colonies were pooled and analysed as early passage, sub-confluent cells 24 h after plating. For serum-stimulations, confluent early passage cells were serum-starved for 2 days in 0.5% FCS/DMEM and then stimulated with 10% FCS/DMEM.

Recombinant retroviruses

The control retroviral constructs carrying the neo' gene alone include pMV6 (Kirschmeier et al., 1988; gift of W.Lee) and pDORneo (Oster and al., in preparation). The v-myc retroviral construct, pDoKv-myc (Thompson et al., 1989) is based on Moloney sarcoma virus carrying the v-gag–myc oncogene derived from the avian myelocytomatisos virus MCV29. The pM6hc-myc,23 cDNA plasmid contains Moloney leukemia virus derived sequences and a human c-myc cDNA covering exons II and III from the Thai site upstream of the initiator ATG codon to the Nol site downstream of the TAA stop codon (Stone et al., 1987). The retrovirus carrying the frameshift mutant of human c-myc (MV6hecmyc2,3fs)was generated in the exon II and III cDNA by inserting a BamHI linker (5'CCGATCCG3'; Pharmacia) into the unique EcoRV site in c-myc exon II. This created an out of frame mutation at amino acid 48 in the human c-myc protein. The deletion mutant of the human c-myc gene (DORshcmyc), which is truncated downstream of the unique Clal site, is interrupted in the coding region of exon III at approximately amino acid 300 (Varley et al., 1987). A 3.5 kb Nhel–HindIII fragment, which includes c-myc exon II, and the breakpoint in exon III (a gift of J.Varley) was subcloned into the Moloney murine leukemia virus retroviral plasmid pDORneo. The retroviral construct, pDORhc-myc I, II, III gene, carries the 4.8 kb Xhol–Nhel human c-myc gene (subcloned from pm21; Stone et al., 1987) inserted into pDORneo. The pZipBam/hygro and pZipmyc/hygro constructs were generated from
...pZip/myc (Thompson et al., 1989) by removing the v-myc gene with BglII and ligating in the hygromycin resistance gene, as a BamHI fragment from pHygroBam (a gift of V.von Hoyningen-Huene). pZip-Bam/hygro was then completed by digestion with BamHI to remove the v-Ha-ras gene. pZip/myc/hygro was subsequently constructed by inserting a BglII linker 2.9 kb BstEII-SalI fragment which encoded the v-gag—myc oncogene MC29 into the BamHI site of ZipBam/hygro.

**RNAse protection**

RNA was prepared by the guanidinium isothiocyanate method of Chirgwin et al. (1979). RNAse protection was conducted essentially as described by Zin et al. (1963) except that the RNA probe was first purified by electrophoresis through a 6% denaturing polyacrylamide gel, excised from the gel, eluted for >2 h with agitation at 37°C in elution buffer (0.5 M NH4OAc, 0.1% w/v SDS, 0.1 mM EDTA) and ethanol precipitated. Hybridizations were conducted at 52—54°C.

The probes were generated using T3 RNA polymerase (Stratagene) from linearized Bluescript KS and SK cloning vectors (Stratagene) containing the following DNA fragments: a 930 bp PstI—SalI fragment containing rat c-myc exon 1, a 420 bp HindIII—SacI fragment encoding v-myc oncogene MC29 (Reddy et al., 1983); a 900 bp Psrl fragment containing chicken GAPDH subcloned from gEMI (Dugacizky et al., 1983; gift of D.Bentley); and a 450 bp Apa—TaqI fragment containing rat GAPDH subcloned from PRLC GAP (Tso et al., 1985; gift of J.Tso, X.-H.Sun and R.Wu). The protected probes were resolved by electrophoresis on 6% or 8% denaturing polyacrylamide gels and visualized by autoradiography on X-O MAT film (Genetica). To quantitate signal intensities films were preflushed and exposed at —70°C with an intensifier screen.

**Nuclear run-on transcription**

Nuclear run-on transcription assays were performed by a modified version of previously described procedures (Groudine et al., 1981; Lian et al., 1985). To isolate nuclei, Rat-1 cells containing various integrated proviruses (see legend to Figure 2) were stimulated with serum for 40 min, then washed twice with cold PBS, scraped and pelleted. The cells were resuspended in RSB (10 mM Tris—HCl, pH 7.4, 10 mM NaCl, 5 mM MgCl2); 3 x 107 cells/ml were incubated in freezing buffer (50 mM Tris-HCl, pH 8.0, 5 mM MgCl2, 40% glycerol, 0.5 mM DTT) at a concentration of 2—2.5 x 107 nuclei/ml, frozen and stored at —70°C. For in vitro transcriptions 60 ml of nuclei were combined with 20 ml of 25 mM Tris—HCl, pH 8.0, 5 mM MgCl2, 250 mM KCl, 4 ml of 33 mM ATP, TTP and GTP each, 12.5 ml of [a-32P]UTP at 5000 Ci/mmol and 2.5 ml RNA10 (100 U, Promega). Where indicated α-amamin was added at a final concentration of 2 μg/ml. The reactions were incubated for 30 min at 30°C and were stopped by a 10 min treatment with 400 U of RNase-free DNase (Boehringer) at 30°C followed by Proteinase K digestion in the presence of 5 mM EDTA and 1% SDS for 5 min at 65°C and then 45 min at 45°C. The RNA was extracted with phenol—chloroform and the presence of 5 mM EDTA and 1% SDS for 5 min at 65°C and then separated from unincorporated nucleotides by passage through a Sephadex G50 spin column.
