The structure and expression of a *Dictyostelium discoideum* gene encoding a prespore-specific protein

A thesis presented by Anne E.Early

for the degree of Ph.D.

at the University of London

Imperial Cancer Research Fund

and

Imperial College

October 1988

ACKNOWLEDGEMENTS

I wish to thank first and foremost my supervisor, Jeff Williams, for his patient teaching and enthusiastic encouragement. I'd also like to thank all the other people in the lab, each of whom contributed to making my four years so stimulating and enjoyable. I'd especially like to thank Phil Mason, Donna Driscoll and Adriano Ceccarelli for everything they have taught me, and Jane Kirk, Hiro Mahbubani and Keith Jermyn, without whom everything would have been so much more difficult. Thanks are also due to Adrian Harwood for critical reading of the manuscript, Liz Johnson for proofreading, John Nicholson for photographic work and David Glover, my university supervisor.

ABSTRACT

The D19 cDNA clone is derived from a mRNA species which is highly enriched in prespore cells during development of the cellular slime mould *Dictyostelium discoideum*. Four independent genomic clones spanning the D19 coding region were isolated. Comparison of the nucleotide sequence with partial peptide sequence obtained elsewhere demonstrated that the D19 gene encodes PsA, a cell-surface glycoprotein of prespore cells. The structure of the extreme C terminus, together with significant homology in this region to contact sites A, a *Dictyostelium* cell adhesion molecule, predicts attachment of PsA to the cell surface via a glycosyl-phosphatidylinositol (GPI) link.

Using a mutant defective in DIF accumulation, it was shown that the concentration of D19 mRNA, and several other prespore messages, decreases in the presence of exogenous DIF. In vitro nuclear transcription and mRNA half-life analyses demonstrated that DIF acts primarily at the transcriptional level to repress the accumulation of D19 mRNA.

The transformation system was used to analyse the control of D19 expression. The gene was marked by insertion of an oligonucleotide encoding an epitope of the human c-myc gene recognised by a monoclonal antibody. With 995 nucleotides of sequence flanking the cap site, both the mRNA and protein were efficiently expressed and correctly localised. Progressive deletion from the 5' end gradually reduced expression, but temporal and spatial specificity was retained

with 338 nucleotides. A 3' to 5' deletion series was created in which the D19 promoter was coupled to the bacterial CAT gene via the TATA box, cap site and 5' non-coding of the actin 15 gene. D19 sequences spanning -122 to -995 directed correctly regulated expression. Together with the 5' deletion series this defines a 216 nucleotide minimal promoter region, and also indicates that transcriptional control is sufficient for the timing and prespore-specificity of expression.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....2

ABSTRACT.....3

CHAPTER 1
INTRODUCTION.....13

I The development and differentiation of *Dictyostelium discoideum*.....14
1) *Dictyostelium discoideum* as a model system.....14
2) The asexual life-cycle.....15
3) Genome organisation and expression.....20
a) Gene organisation and structure.....20
b) Messenger RNA structure and processing.....23
c) Patterns of gene expression during development..24
4) Cellular differentiation and pattern formation.....27
5) The control of cell-type-specific gene expression..32
a) Patterns of cell-type-specific gene expression..32
b) Cyclic AMP and cell-type-specific gene expression 34
c) Other factors regulating cell-type-specific gene expression.....39
6) Regulatory sequences upstream of *D.discoideum* genes 43

II The control of cell-type-specific gene expression in other eukaryotes.....46
1) The structure of eukaryotic promoters.....47
a) The TATA box.....48
b) Proximal promoter elements.....49
i) The CCAAT box.....50
ii) Sp1 binding sites.....51
iii) Other promoter elements.....52
c) Enhancers.....53
2) Transcriptional control of cell-type-specific gene expression.....58
a) Cell-type-specific enhancers.....58
i) *Drosophila* yolk protein gene expression.....58
ii) Immunoglobulin gene expression.....59
iii) Alpha-fetoprotein gene expression.....63
iv) Other cell-type-specific enhancers.....64
b) Negatively acting mechanisms.....65
c) Transcription factors and tissue specificity....67
i) Co-ordinate control by shared transcription factors.....67
ii) Post-translational regulation of transcription factors.....69
3) Higher order control of cell-type-specific gene expression.....70
a) Chromatin structure.....70

b) Regulatory genes and gene networks.....73
c) Extracellular signals.....75
4) Conclusions and implications for the control of cell-type-specific gene expression in *D.discoideum*.....79

CHAPTER 2
MATERIALS AND METHODS.....82

2.I Materials.....82
1) D.discoideum strains.....82
2) Bacterial strains.....82
3) Isotopes.....82
4) Enzymes.....83
5) Chemicals and miscellaneous materials.....83
6) Solutions.....84
7) Antibodies.....84
8) Abbreviations.....85

2.II Methods.....87
1) Bacterial techniques.....87
a) Growth and long-term storage.....87
b) Preparation of transformation competent DH5 cells 87
c) Transformation of *E.coli* strain DH5.....89
d) Colony lifts for screening.....90
2) Slime mould techniques.....90
a) Growth and long-term storage of *D.discoideum* cells.....90
b) Development of *D.discoideum* cells.....93
i) Filter development.....93
ii) Production of slugs and fruiting bodies.....93
c) Percoll gradient fractionation of prespore and prestalk cells.....94
d) HM44 stalk cell induction in submerged monolayer culture.....95
e) Preparation of transcriptionally active nuclei..96
f) Transformation of *D.discoideum* cells.....96
g) Preparation of samples for immunohistochemistry 98
3) DNA preparation and analysis.....100
a) Preparation of plasmid DNA.....100
i) Small scale preparation of plasmid DNA by the boiling method.....100
ii) Large scale preparation of DNA.....100
b) Restriction enzyme analysis.....102
c) Gel electrophoresis of DNA.....102
i) Agarose gels.....102
ii) Polyacrylamide gels.....104
d) DNA recovery from gels.....107
i) Electroelution.....107
ii) Onto NA45 paper.....107
e) Southern transfer.....108
4) DNA manipulation.....108
a) End-filling and labelling of DNA.....108
i) End-labelling of a 5' overhang using the Klenow fragment of DNA polymerase I.....109
ii) Phosphatase.....109
iii) Kinase labelling.....110

b) Subcloning of DNA fragments into vectors.....110
i) Ligation.....110
ii) Use of linkers.....111
iii) Insertion of a double-stranded oligonucleotide into a vector.....112
c) Sequencing by the chemical cleavage method.....112
d) Generation of a directed deletion series using exonuclease III.....113
e) Creation of a deletion series with Bal31 nuclease 116
5) RNA analysis.....117
a) Extraction of nucleic acid from *D.discoideum*...117
i) Phenol extraction.....117
ii) Extraction of RNA with GTME.....118
b) Formaldehyde gels and Northern transfer.....118
c) Primer extension.....119
d) In vitro nuclear transcription.....120
6) Hybridisation.....121
a) Preparation of probes.....121
i) Nick-translation of DNA.....121
ii) Labelling using a mixed oligonucleotide primer.....122
iii) Kinasing of oligonucleotides.....122
b) Hybridisation conditions.....123
i) Probes prepared by nick translation or mixed primer labelling.....123
ii) Oligonucleotide probes.....124
iii) Labelled RNA generated by in vitro nuclear transcription.....125
c) Autoradiography and density scanning.....126
7) Protein techniques.....126
a) Western blotting.....126
i) Extraction of total cellular protein from *D.discoideum*.....126
ii) Running and transfer of one-dimensional SDS polyacrylamide gels.....126
iii) *D.discoideum* colony blotting.....129
iv) Antibody binding to Western and colony blots 129
b) Assaying transformed cells for chloramphenicol acetyl transferase activity.....130
c) Immunostaining.....132
i) Immunostaining and light microscopy of fixed samples.....132
ii) Immunoelectron microscopy.....132
8) Miscellaneous.....133
a) Ethanol precipitation of nucleic acid.....133
b) Sephadex G50 columns.....133
c) Equilibration of Dowex.....134
d) Deionization of formamide.....134
e) Preparation of dialysis tubing.....135

CHAPTER 3
Isolation and structural analysis of genomic clones corresponding to the D19 gene.....136

3.1 Southern transfer analysis of D19 gene organisation.....136

3.2 Structure of the D19 cDNA clone.....139
3.3 Isolation and mapping of D19 genomic clones.....141
3.4 Sequence analysis of the D19 genomic clone.....145
3.5 Identification of the 5' and 3' termini of the D19 mRNA.....147
3.6 Structure of the polypeptide encoded by the D19 gene.....151
3.7 Identity of the D19 gene product with PsA, and homologies to other proteins.....154
3.8 Attempts to produce a PsA-deficient mutant.....159
3.9 Summary.....162

CHAPTER 4
The mechanism of repression of the D19 gene by DIF....164

4.1 The HM44 in vitro assay system.....165
4.2 DIF repression of D19 and two other prespore genes in vitro.....167
4.3 A D19-related sequence induced by DIF during stalk cell induction.....170
4.4 DIF induction of 56 and 63 gene transcription...173
4.5 The repression of D19 by DIF at the level of gene transcription.....174
4.6 The effect of DIF addition on the stability of D19 mRNA.....179
4.7 DIF repression of D19 at low cAMP concentrations. 181
4.8 D19 protein levels during stalk cell induction..184
4.9 Summary.....186

CHAPTER 5
Analysis of D19 gene expression using the transformation system I
Construction of *D.discoideum* transformation vectors and transformation of a marked copy of the D19 gene.....188

5.1 Construction and characterisation of pB10T.....189
5.2 Construction of the pB10TP vector series by multilinker replacement.....193
5.3 Marking of the D19 gene with an epitope of the human c-myc protein.....198
5.4 Expression of the marked gene after DNA-mediated transformation of AX-2 cells.....201
5.5 Localisation of the D19 protein by indirect immunofluorescence.....207
5.6 Summary.....211

CHAPTER 6
Analysis of D19 gene expression using the transformation system II
Characterisation of 5' flanking sequences.....213
6.1 Construction of an external 5' flanking deletion series using exonuclease III.....213
6.2 Nucleotide sequence upstream of the D19 gene....217
6.3 Transformation of deletion derivatives to identify the 5' boundary of the functional promoter.....219
6.4 Generation of 3' flanking deletions and construction of heterologous gene fusions.....227

6.5 Identification of the 3' boundary of the D19 promoter sequences required for regulated expression. 235
6.6 Sequence characteristics of the minimal promoter region, and attempts to reconstitute D19 promoter activity with synthetic oligonucleotides.....247
6.7 Summary.....253

CHAPTER 7
Expression of D19 late in development.....254

7.1 The distribution of D19 mRNA in mature fruiting bodies.....255
7.2 The expression of a transformed copy of the D19 gene in mature culminants.....258
7.3 Expression of other prespore genes during culmination.....262
7.4 Localisation of D19 protein in culminants.....264
7.5 Summary.....267

CHAPTER 8
DISCUSSION.....269

8.1 The structure and function of the encoded polypeptide.....270
8.2 DIF and the control of cell-type-specific gene expression.....283
8.3 The expression of D19 in stalk cells.....293
8.4 Transcriptional control of D19 expression.....301
8.5 Conclusions and future prospects.....314

REFERENCES.....317

PUBLISHED WORKS.....339

FIGURES

3.1 Southern blotting analysis of Dictyostelium genomic DNA 137
3.2 Restriction map of the D19 cDNA clone.....140
3.3 Use of positive clones from the library screen as probes in Northern blot analysis.....142
3.4 Restriction maps of the D19 genomic inserts.....144
3.5 Strategy used to sequence the genomic clone 4.3.....146
3.6 Sequence of the D19 mRNA.....148
3.7 Identification of the 5' terminus of the D19 mRNA by primer extension analysis.....150
3.8 Plot of the hydropathic index of the predicted translation product of the D19 gene.....153
3.9a Predicted PsA leader peptide sequence
3.9b Regions of homology between PsA, CsA and N-CAM.....156
3.10 Colony blot of AX-2 transformants probed with the monoclonal antibody MUD-1.....160

4.1 The HM44 monolayer assay system.....166
4.2 The kinetics of repression of the D19 mRNA by DIF...168
4.3 Comparison of the repression of D19 with that of two other prespore genes.....169
4.4 Analysis of the cross-hybridising mRNA species.....171
4.5 The induction of pDd63 and pDd56 gene transcription by DIF.....175
4.6 Determination of the rate of D19 gene transcription during stalk cell induction.....176
4.7 Northern transfer analysis of cytoplasmic RNA extracted in parallel with nuclei preparations.....177
4.8 The half-life of D19 and PL3 mRNA in the presence and absence of DIF.....180
4.9 Repression of D19 mRNA in the presence of low cAMP levels.....182
4.10 The effect of DIF on D19 protein levels in HM44 cells 185

5.1 Scheme of construction of pB10TP1.....190
5.2 Termination of neomycin resistance gene transcription by CP1 sequences.....191
5.3 Nucleotide sequence of the multilinker of pB10TP1 and pB10TP2.....194
5.4 Restriction maps of the pB10TP vectors.....196
5.5 Construction of vectors containing alternative versions of the marked D19 gene.....200
5.6 Southern transfer analysis of cells transformed with pB10B1 and pB10N6.....202
5.7 Temporal and spatial patterns of expression of the marked gene during development.....203
5.8 Expression of the D19 gene in wild-type and transformed cells.....205
5.9 Detection of expression of the NdeI-marked gene.....206
5.10 Immunostaining of transformants with the anti-c-myc monoclonal, 9E10.....208
5.11 Prespore localisation of staining detected with the c-myc monoclonal antibody.....210

6.1 Endpoints of the 5' flanking deletions generated with exonuclease III....216
6.2 Complete nucleotide sequence upstream of the D19 coding region in the genomic clone 4.3....218
6.3 Southern blotting analysis of 5' deletion constructs after transfomation....221
6.4 Expression of the -442 and -268 constructs during development....222
6.5 Expression of the -364, -338 and -291 deletion constructs....224
6.6 Relative expression levels of the 5' deletion constructs....226
6.7 G-rich sequences identified upstream of several D.discoideum genes....228
6.8 Restriction map of D19 ΔBam.CAT....231
6.9 Relative arrangement of D19 and actin 15 promoter elements in the intact and fusion genes....232
6.10 Sequence across the D19/actin 15 fusion....234
6.11 Expression of the -52 D19/actin fusion construct detected by CAT activity....236
6.12 Expression of the -122 and -163 fusion constructs...237
6.13 Relative expression levels of the 3' deletion constructs....239
6.14 Identification of the start-site of transcription of the D19/actin fusion constructs....241
6.15 Expression of the -52 construct at the RNA level....243
6.16 CAT activity generated by the -52 construct early in development....244
6.17 Cell-type-specificity of the expression of the -122 construct....246
6.18 Minimal promoter region defined by deletion analysis 248
6.19 Relative arrangement of the promoter elements in the oligonucleotide substitution constructs....251
6.20 Expression of the oligonucleotide substitution constructs....252

7.1 Expression of the D19 gene in stalks detected by Northern transfer....256
7.2 Determination of the 5' terminus of the D19 stalk message....259
7.3 Expression of the marked gene in culminants at the RNA level....261
7.4 Expression of other prespore genes in mature culminants 263
7.5 Visualisation of marked PsA in culminants by immunoelectron microscopy....266

8.1 Common structural features of GPI anchors....274

TABLES

2.1 Media used for growth and transformation of *E.coli*...88
2.2 Media used for growth and transformation of *D.discoideum*.....91
2.3 Commonly used enzyme buffers.....103
2.4 Acrylamide gel recipes.....106
2.5 Sequencing by the chemical cleavage method.....114
2.6 Solutions for SDS-PAGE and Western blotting.....128

CHAPTER 1: INTRODUCTION

Understanding the molecular mechanisms responsible for the emergence of distinct cell types from a common progenitor is a major goal of developmental biology. The cellular slime mould *Dictyostelium discoideum* is a unicellular organism which, in certain environmental conditions, undergoes a developmental phase to create a multicellular structure. Hence it is a simple model of the developmental processes occurring in more complex organisms and is used to investigate this problem. Growing cells exist as a homogeneous population of amoebae, but when the food supply is exhausted, development is initiated. The cells aggregate and two major cell types can then be distinguished; prestalk and prespore cells. These cells form a distinct pattern and are the precursors of the mature stalk and spore cells of the culminant. Many developmental markers have been isolated, displaying both stage- and cell-type-specificity, and are utilised to study the control of cellular differentiation and gene expression. Several substances identified as possible morphogens have been shown to influence both patterning and the expression of individual genes.

This thesis describes the structural characterisation of a gene, D19, expressed exclusively in prespore cells, and hence a widely used marker of the prespore differentiation pathway. The mechanism by which a known morphogen, DIF, represses the gene is established, and by reintroducing

altered copies of the gene into *Dictyostelium* using the transformation system, regions important for cell-type-specific expression are localised. The implications of these results for the control of cell-type-specific gene expression and cellular differentiation in *D.discoideum* are discussed.

I The development and differentiation of *Dictyostelium discoideum*

1) *Dictyostelium discoideum* as a model system

D.discoideum is used for the study of many diverse processes which occur in higher organisms, including chemotaxis, cell adhesion, co-ordinate gene expression and pattern formation (for a review, see Loomis, 1982). The major attraction of this organism is its simplicity. During development only two cell types arise from a single precursor, representing a simple choice between differentiation pathways, and reducing the likely number of signals and cellular interactions involved in comparison to more complex organisms. The cell types can be both easily distinguished and physically separated, facilitating studies of pattern formation and cell-type-specific gene expression.

The organism is haploid during its asexual life cycle, allowing mutants to be readily isolated, and genetic analysis using the parasexual system has been successfully applied (reviewed by Newell, 1982). As growth can proceed without development, clonal populations of mutants blocked in development can be propagated. *Dictyostelium* is relatively

easy to cultivate; certain strains can grow axenically in shaken suspension, and development of a population can be initiated synchronously by starvation, enabling purification of large numbers of cells at the same stage of development.

D.discoideum is classified as a lower eukaryote, but its genome organisation, gene structure and RNA processing patterns largely resemble those of other eukaryotes (see section I.3). Therefore the mechanisms controlling gene expression are likely to be shared with higher cells. However, as it has a small genome, the techniques of molecular biology are relatively easy to apply. A transformation system exists, allowing the reintroduction of cloned genes and hence the identification of sequences important for controlling their expression (Nellen et al., 1984). Recently, homologous gene disruption techniques have been developed which will enable the function of specific gene products to be assessed (De Lozanne and Spudich, 1987; Witke et al., 1987).

Unlike other systems, the growth and developmental phases are largely distinct, and differentiation can occur in the complete absence of cell division. Overt differentiation to form prestalk and prespore cells precedes irreversible committment to a pathway of differentiation. Stalk cells finally undergo terminal differentiation and die, but spores remain pluripotent.

2) The asexual life-cycle

D.discoideum has two alternative life cycles, sexual and asexual. During the sexual life cycle a thick-walled macrocyst is formed, containing a zygote nucleus produced by the fusion of starving amoebae of opposite mating types. However, it is the asexual cycle which is commonly used for experimental studies. Upon starvation, individual amoebae aggregate in response to the depletion of certain amino acids (Marin, 1976). After about 24 hours a mature fruiting body is produced, consisting of two distinct cell types; a globule of spores supported by a cellulose-covered stalk composed of highly vacuolated, dead cells (reviewed by Bonner, 1982).

cAMP is the chemotatic signal responsible for aggregation (reviewed by Devreotes, 1982). Individual cells in the population, known as signalling centres, begin to emit pulses of cAMP, and form aggregation territories of 10^5-10^6 cells. The concentric waves of cAMP originating from the signalling centre are propagated outwards because the surrounding amoebae detect external cAMP and synthesize and secrete cAMP in response. Adenylate cyclase is activated by a change in occupancy of the surface cAMP receptors (Gerisch and Wick, 1975; Roos et al., 1975; Dinauer et al., 1980a), and cAMP is secreted at a rate directly proportional to the transient increase in intracellular concentration (Dinauer et al., 1980a). The pulses of cAMP produced have a half-width of about 2 minutes, and are accompanied by oscillations in intracellular cGMP and extracellular H^+ concentrations, and in the exchange of intracellular and extracellular Ca^{2+} and K^+ (Aeckerle et al., 1985; Bumann et al., 1986). To overcome an adaptation effect mediated by phosphorylation of the

receptor (Dinauer et al., 1980b; Devreotes and Sherring, 1985; Klein et al., 1985a+b), extracellular cAMP is removed by phosphodiesterase activity (Malchow et al., 1972), allowing a response to the next pulse to be generated.

Cells develop the ability to respond to cAMP early during development. The number of cell surface cAMP receptors increases about 10-fold to approximately 10^5 per cell, the basal activity of adenylate cyclase increases 20 to 40-fold, and the activity of cell surface and extracellular cAMP phosphodiesterase increases about 50-fold (Gerisch et al., 1985). The expression of cell surface cAMP receptors and adenylate cyclase is enhanced by stimulation of developing cells with pulses of cAMP (Gerisch et al., 1985). Two classes of cell surface cAMP receptor have been identified; 96% fast dissociating (A), likely to mediate the activation of adenylate cyclase, and 4% slowly dissociating, probably responsible for cGMP production (B) (Van Haastert and De Wit, 1984). Both classes can exist in high or low affinity forms, depending on whether cAMP is bound (Van Haastert and De Wit, 1984; Van Haastert et al., 1986). After prolonged exposure to cAMP the cell surface receptors are downregulated (Van Haastert and De Wit, 1984; Van Haastert, 1985; Van Haastert et al., 1986).

An inward movement of amoebae occurs as a response to the cAMP gradient at the leading edge of the wave (Tomachik and Devreotes, 1981). The cells first make end to end contacts, resulting in streams of cells radiating from the aggregation centre. A few hours later, side by side contacts are formed and loose mounds of cells become tight aggregates (reviewed

by Loomis et al., 1987). At this stage a tip is formed at the top of the cell mound, apparently as a direct consequence of the chemotactic attraction of cells to an autonomous oscillator (Schaap et al., 1984). The tip has properties similar to a classical embryonic organiser; for example, if new tips are grafted on to a slug, secondary axes will be formed (Raper, 1940; Farnsworth, 1973; Rubin and Robertson, 1975). The tip apparently controls the movement of surrounding cells; slug tips can attract streams of aggregation competent cells (Bonner, 1949), and if the tip of a migrating slug is removed, movement ceases until a new tip is formed (Raper, 1940).

Cyclic AMP signalling originating from the tip may be important for regulating the size and pattern of aggregates (Schaap and Wang, 1986; reviewed by Schaap, 1986). There is evidence that oscillatory cAMP signalling continues after aggregation has been completed. Cell movement in periodic waves has been observed (Durston et al., 1976; Durston and Vork, 1979), and a chemotactic response can be observed by exposing post-aggregative cells to a cAMP gradient (Maeda, 1977). Aggregate size is dependent on the number of tips, in turn governed by the ability of a tip to inhibit the formation of competing tips. There is evidence that a diffusible substance produced by the tip, probably adenosine, is responsible for this inhibition (Kopachik, 1982; Schaap and Wang, 1986). The enzymes which convert cAMP to adenosine, cAMP phosphodiesterase and 5'nucleotidase, are preferentially localized in prestalk tissue (Armant et al., 1980; Tsang and Bradbury, 1981). The mechanism by which adenosine acts to

prevent the generation of competing oscillators is unclear, but it is known that adenosine interferes with cAMP relay by inhibiting the binding of cAMP to its receptor (Newell and Ross, 1982).

Comparing different slime mould species, where there is an early onset of cAMP signalling (e.g. *D.discoideum*) the size of the prespore zone at the tip is relatively large (Schaap and Wang, 1984; Schaap et al., 1985). This may be the result of two distinct mechanisms; inhibition of prespore differentiation by adenosine (Schaap and Wang, 1986; see section 1.5c), and a reduced ability of prespore cells to respond to chemotactic signals produced by the tip (Matsukuma and Durston, 1979).

The tip elongates to form the first finger which then enlarges to become a cylinder of cells. This structure may culminate immediately to form a mature fruiting body or alternatively pass through a migratory stage. In appropriate environmental conditions the aggregate collapses on to the substratum to give a pseudoplasmodium, or slug, which moves in response to heat and light (Bonner et al., 1950). At this stage, two cell types arranged in a distinct pattern can be discerned (reviewed by Morrissey, 1982). The anterior 20% of the slug consists entirely of prestalk cells. The posterior zone is largely composed of prespore cells, but also contains scattered cells resembling prestalk cells, termed anterior-like cells (Sternfield and Davis, 1982). The aggregate is covered in a thin slime sheath, left behind as a collapsed tube as the slug migrates. It acts as a physical barrier against the loss or gain of cells, and may also have a

function in homeostasis and slug traction (Vardy et al., 1986; Odell and Bonner, 1986).

At the beginning of culmination the slug tip ceases to move foward and points away from the substratum slightly. The posterior zone continues to move, with the result that all the cells are gathered under the tip. Stalk cells begin to form just below the tip by progressive vacuolisation and formation of cellulose cell walls (Bonner, 1944; Raper and Fennell, 1952). New stalk cells are formed from prestalk cells on reaching the stalk tip. The reverse fountain movement of prestalk cells also serves to push the stalk downwards (Raper, 1940). The cells originating from the very posterior end of the migrating slug vacuolate and form the basal disc, and when the stalk makes contact with the substratum the surrounding cells are lifted into the air. While the stalk continues to elongate, the prespore cells differentiate to form ellipsoid spores surrounded by a tough cellulose-containing spore coat. Rising culminants are repelled from one another by the production of ammonia (Bonner and Dodd, 1962). Germination of the spores in the spore head is inhibited by an adenine derivative, discadenine, synthesized during late development (Ihara et al., 1986).

3) Genome organisation and expression

a) Gene organisation and structure

D.discoideum has a small haploid genome 50x 10^3 kb in size (Sussman and Rayner, 1971; Firtel and Bonner, 1972), consisting of 7 chromosomes (reviewed by Newell, 1982). The complexity of the single-copy portion, which constitutes 70-75% of the total, is only 7 times that of *E.coli*, or approximately 2% that of mammalian genomes. An unusual aspect of *Dictyostelium* is the extreme A-T richness of its genome. Overall, the G-C content is 22% (Sussman and Rayner, 1971; Firtel and Bonner, 1972). In protein coding regions this value is 30-40%, while in non-coding regions the G-C content is only 5-20%. Unlike higher eukaryotes, no substantial DNA methylation has been detected (Kindle and Firtel, 1978; Kimmel and Firtel, 1979; Kimmel and Firtel, 1982).

As a substantial number of genes have been cloned and sequenced, comprising both closely related members of gene families and those present as single copies, some general conclusions can be made about gene structure and organisation. Unsurprisingly for an organism with a small genome, there are several examples of closely linked genes, involving genes with no apparent homology as well as members of the same gene family. An example of the former situation is the linkage between the coregulated genes cysteine proteinase 1 (CP1) (Williams et al., 1985) and pDG17 (Driscoll and Williams, 1987). The genes are divergently transcribed and the intergenic region of 910 nucleotides is able to direct bidirectional transcription (Driscoll and Williams, 1987). Similarly, the 3' end of CP1 is separated from a convergently transcribed gene by 271 nucleotides, and the intergenic region contains signals sufficient for

termination of CP1 trancription (see Chapter 5). In contrast, the M4-1 and M4-4 genes, members of the same family and separated by 1.7kb of sequence, are divergently transcribed but not coregulated (Kimmel and Carlisle, 1986). Assuming transcriptional control mechanisms are operating, this indicates that 1.7kb can contain two independent sets of promoter signals.

Reflecting the overall high A-T content, there are very marked codon prefences, and the third base position is very likely to be an A or a T. A minority of genes analysed contain introns, but where they do exist, they are usually short (in the range of 80-180 nucleotides) and extremely A+T-rich (>95%). In spite of this unusual structure compared to other eukaryotes, the junction sequences are very similar, and in particular the consensus GT..AG is invariantly found (Breathnach et al, 1978).

The sequence flanking the 5' end of coding regions is 95% A-T rich. Like most other eukaryotic genes there is a TATA box consensus sequence, TATAAA$^{T}/_{A}$A, located 30-40 nucleotides upstream of the ~~ATG~~ cap site, but unlike other systems the TATA box is immediately followed by an oligo(dT) stretch of varying length and homogeneity. The two regions may combine to provide a recognition site for RNA polymerase II. The actin 2-sub2 pseudogene lacks a TATA sequence, while the actin 3 pseudogene lacks the oligo(dT) stretch (McKeown and Firtel., 1981).

The 3' flanking sequences are extremely A-T rich, and one or more copies of the sequence AATAAA, which have been shown to act as a signal for 3' end processing and poly(A) addition

in other systems (reviewed by Bernstiel et al., 1985), are usually found. This sequence is generally located 10-20 nucleotides 5' to the site of poly(A) addition, but in *Dictyostelium* is found as much as 60 nucleotides upstream. Sequences important for 3' end processing have not been determined for *Dictyostelium*, and it could be that the high A-T content of non-coding DNA means that other signals are of greater importance.

b) Messenger RNA structure and processing

Pulse-chase experiments have shown that 75% of hnRNA is converted to mature cytoplasmic RNA (Firtel et al., 1976), in contrast to mammalian cells, where the figure is typically only about 10% (Darnell, 1979). The 25% lost is partly accounted for by processing to remove introns (Kimmel and Firtel, 1980), but fewer repeated sequences are represented in cytoplasmic RNA than hnRNA (Firtel and Jacobson, 1977). Messenger RNA processing and transport is very rapid in growing cells (Lodish et al., 1974; Firtel et al., 1976), but during development the rate of nuclear RNA processing is reduced, and the time taken for individual transcripts to appear in the cytoplasm can vary (Mangiarotti et al., 1983a).

As in other eukaryotes, 95% of hnRNA and cytoplasmic RNA has a cap structure (Dottin et al., 1976). The length of the 5' non-coding region varies widely between individual messages, but a common feature is the 95% or greater A-T composition. Virtually all newly synthesized hnRNA and mRNA contains poly(A)+ (Palatnik et al., 1979). The poly(A)-

fraction primarily encodes histone proteins, as in other systems, but other messages exist in both a poly(A)- and poly(A)+ state. The poly(A) tail length may be involved in translational control (Palatnik et al., 1984), but there is no evidence that it affects mRNA stability (Palatnik et al., 1980).

Messenger RNA in vegetative cells decays with biphasic kinetics; one population of messages has a half-life of 2 hours, and another 10 hours (Palatnik et al., 1980). In contrast, in developing cells most mRNA species seem to decay with a half-life of 3-4 hours (Margolskee and Lodish, 1980). Individual exceptions have been reported; after cessation of nuclear transcription, discoidin 1 mRNA disappears with a half-life of 30 minutes (Williams et al., 1979), and mRNAs induced during spore germination also have short half-lives (Kelly et al., 1985). In these cases it is likely that an intrinsically short half-life allows control mediated at the level of transcription to act rapidly. However, for some developmentally regulated messages accumulated after aggregate formation, there is evidence that changes in mRNA stability could be an important control mechanism (Mangiarotti et al., 1982; Mangiarotti et al., 1985; see section I.5b).

c) Patterns of gene expression during development

Measurements of enzyme activity provided the first evidence that programmed changes in gene expression occur during development of *D.discoideum*. Changes in the activity

of at least two enzymes (UDP glucose pyrophosphorylase and glycogen phosphorylase) could be specifically correlated with increased rates of polypeptide synthesis (Franke and Sussman, 1973; Thomas and Wright, 1976). Later, developmental changes in patterns of protein synthesis, examined by 2D-polyacrylamide gel electrophoresis (Alton and Lodish, 1977; Cardelli et al., 1985), suggested that co-ordinate changes occur at distinct developmental stages. Of the 80 developmentally regulated proteins detected by Alton and Lodish (1977), 40 were induced during late aggregation as tight cell-cell contacts were being formed, and a second, smaller induction occurred during culmination. Cardelli et al. (1985) detected changes in the synthesis of 189 proteins and identified two additional stages at which major changes occur, early development and early aggregation. However, only proteins able to be resolved and encoded by abundant and moderately abundant mRNA species (10-20 copies/cell) are detected. Therefore, the number of developmentally regulated proteins are underestimated, and accumulation patterns may simply be a reflection of the function of the more abundant proteins.

Patterns of RNA expression during development have been examined using hybridisation kinetics (Firtel, 1972; Firtel and Jacobson, 1977; Blumberg and Lodish, 1980a+b, 1981; Jaquet et al., 1981). Varying estimates were obtained by different groups, but broadly similar expression patterns were observed. Unlike protein synthesis measurements, a low abundance class can be detected, composed of a large number of distinct sequences.

Vegetative cells contain 3000-5000 distinct mRNA sequences, and during development it has been estimated that up to 3000 additional mRNA species are expressed. However, it seems that most of the newly synthesized RNA is the product of 700-900 genes (Jaquet et al., 1981). In agreement with the protein synthesis data, co-ordinate expression at distinct stages seems to occur. Not all the results are consistent; for instance, in pre-aggregation cells there is no significant change in the number of mRNA species, and cross-hybridisation experiments suggest that the population is very similar. However, it is clear from the results of gene cloning that novel genes are expressed during this period. The major change in gene expression seems to occur during late aggregation. Overall, it was estimated that about 80% of the single-copy genome is asymmetrically transcribed during growth and development (Lodish et al., 1982).

Studies of developmental mutants altered in timing of gene expression has reinforced the idea of shared control mechanisms (Saxe and Firtel, 1986). Although groups of genes are expressed at different relative times in different mutants, within each group co-ordinate expression still occurs.

The induction of proteins during development is largely correlated with the appearance of translatable mRNA in the cytoplasm, suggesting control at the level of RNA synthesis or stability (Alton and Lodish, 1977). Most of the aggregation-specific mRNA species are absent from the nuclear RNA of growing or preaggregating cells, and so it seems likely that they are under transcriptional control (Lodish et

al., 1982). The transcription of individual cloned DNAs, examined by in vitro nuclear transcription, has largely supported this (Williams et al., 1979; Landfear et al., 1982). However, as the rate of mRNA processing and transport is variable in developing cells, it may be a control mechanism for some genes (Mangiarotti et al., 1983a). Also, there is evidence that the expression of some aggregation-stage genes is primarily controlled by mRNA stabilisation. Some genes tested are transcribed during growth or early development, but accumulation of their mRNAs is only detected after aggregation (Mangiarotti et al, 1985; A.Ceccarelli, personal communication;). This may be related to the observation that on disruption of aggregates, transcripts derived from some aggregation stage genes are specifically destabilized, as in both cases many of the genes controlled in this manner are expressed predominantly in prespore cells (see section I.5b).

4) Cellular differentiation and pattern formation

Using histological criteria, two distinct populations of cells in the aggregate can first be distinguished just prior to tip formation. Later, a very sharp boundary is seen in the migrating slug between prestalk cells in the anterior fifth and prespore cells in the posterior four-fifths. Prespore cells become smaller and more electron dense due to water loss, and contain vacuoles for storage of spore coat proteins (Hohl and Hamamoto, 1969; Ikeda and Takeuchi, 1971; Schaap et al., 1982). Autophagic vacuoles are selectively lost from

prespore cells, but remain in prestalk cells, allowing retention of vital dyes such as neutral red (Bonner, 1952; Yamamoto et al., 1981). On the whole, prestalk cells are less overtly differentiated than prespore cells and resemble pre-aggregative cells.

Several cell-type-specific antibodies have been generated and used as markers in studies of cellular differentiation and pattern formation (see below). A polyspecific antiserum directed against spores detects an acidic mucopolysaccharide component of prespore vesicles (anti-PSV; Takeuchi, 1963). Several prespore-specific monoclonal antibodies have also been produced, including MUD-1 which recognises a cell-surface antigen (Gregg et al., 1982; Tasaka et al., 1983; Grant et al., 1985). Identification of prestalk-specific antigens has proved more difficult; most of those identified to date are also synthesized in the majority of cells during early aggregation and selectively lost from prespore cells (Tasaka et al., 1983; Barclay and Smith, 1986).

The models which have been formulated to explain how a pattern is created in *D.discoideum* can be divided into two opposing categories; those which propose position-dependent differentiation (e.g. MacWilliams and Bonner, 1979), and those which invoke position-independent differentiation followed by cell sorting (Meinhardt, 1983; MacWilliams et al., 1984; Blaschke et al., 1986). Experimental data exist in support of both arguments.

Cell sorting was first demonstrated by Bonner (1952), who found that transplanted vitally stained cell masses returned to their original positions. This result has since been

repeated with a variety of techniques (Takeuchi, 1969; Leach et al., 1973; Feinberg et al., 1979). Prior to establishment of the pattern, the occurrence of cell sorting has been suggested because of the apparently random appearance of some prespore characteristics (e.g. prespore vesicles) within the aggregate (Forman and Garrod, 1977; Tasaka and Takeuchi, 1981), although the tip always remains prespore-free (Muller and Hohl, 1973). Aggregates formed in suspension culture differentiate into prespore and prestalk cells, initially with a random distribution, but eventually the prestalk cells surround the prespore cells (Sternfield and Bonner, 1977; Forman and Garrod, 1977; Tasaka and Takeuchi, 1981). This demonstrates that positional information is not essential for differentiation, and has been cited as evidence of the differential adhesion hypothesis (Steinberg, 1970). However, any cell sorting which occurs is likely to be at least partly mediated by differing chemotatic properties of prestalk and prespore cells. Prespore cells have lower levels of cell-surface cAMP receptors and cAMP phosphodiesterase (Schaap and Spek, 1984). This reduced signalling capacity is correlated with a reduced ability to respond to chemotactic signals produced by the tip (Matsukuma and Durston, 1979).

Preaggregative cells also show variations in their sorting ability, correlated with their differentiation fate (McDonald and Durston, 1984; Weijer et al., 1984). It has been suggested this difference in chemotactic ability determines the eventual position of cells in the aggregate, forming a prepattern before overt differentiation is seen (Wang et al., 1988a). This was proposed to explain the

observations that the cell cycle phase at the time of growth arrest or conditions of growth and starvation influence the differentiation tendency of preaggregative cells (Leach et al., 1973; Tasaka and Takeuchi, 1981; McDonald, 1984; McDonald and Durston, 1984; Weijer et al., 1984). Cells starved early in the cell cycle (S or G2) exhibit cAMP binding activity, cell-surface cAMP phosphodiesterase activity and the ability to relay cAMP signals earlier than cells which begin development in late G2 (Wang et al., 1988a). They would therefore be expected to have a greater ability to generate and detect chemotactic signals. Consistent with this, they are more likely to initiate aggregation centres (McDonald, 1986) and have a strong tendency to sort to the anterior prestalk region of the slug. Cells which start development late in the cell cycle tend to sort to the prespore region. Modifying conditions of growth and starvation probably also acts by altering the timing of the acquisition of cAMP responsiveness.

The correlation between cell cycle phase and differentiation tendency has been used to support a wholly deterministic model of pattern formation. In isolated cells, the cell cycle phase is closely correlated with the expression of prespore or prestalk markers (Gomer and Firtel, 1987). However, there is disagreement as to whether the prestalk marker used is significantly cell-type-enriched (Pears and Williams, 1987), and expression of the prespore marker may simply reflect the ability of an isolated cell to proceed to a certain stage of development (Williams, 1988). A cell cycle imposed choice of differentiation pathway is also

indicated by experiments in which populations of cells arrested early in the cell cycle formed slugs with a higher than average proportion of prestalk cells (53%), while slugs formed with cells arrested late in the cell cycle had a higher than average proportion of prespore cells (88%) (Weijer et al., 1984). Alternatively, it could be argued that the initially skewed population in some way modified the subsequent formation of a morphogenetic field.

It is clear from regulation experiments that cells are not fully determined, and are subjected to supracellular regulatory mechanisms. Prestalk and prespore cells isolated from slugs regulate to almost normal proportions (Bonner et al., 1955, Sakai, 1973). In prespore isolates regulation involves random differentiation followed by cell sorting, but in prestalk isolates prespore cells appear first at the posterior margin (Bonner, 1952; Sakai, 1973). In related slime mould species there is no apparent predetermination; all slug cells are prespore and the stalk is formed by local redifferentiation of prespore into stalk cells at the stalk-prespore boundary (Gregg and Davis, 1982; Schaap et al., 1985).

The action of a localized signal is evident during culmination of *D.discoideum*, as prespore cells at the tip of the stalk rapidly redifferentiate to form stalk cells (Schaap, 1986). Earlier during development, the appearance of the prespore-specific antigen, MUD-1, is localized at the centre and base of the aggregate, supporting a model of position-dependent differentiation and the existence of a morphogenetic field (Krefft et al., 1984).

Observed cellular behaviour and the properties of molecules known to affect cellular differentiation cannot be completely accommodated by any single existing model of pattern formation. It may be that that cells with a weak differentiation tendency sort to an appropriate position in the aggregate and then receive positional information to promote a particular pathway of differentiation. At present, the distribution of potential morphogens within the multicellular structure can only be predicted, as it is not technically feasible to distinguish between total levels and the proportion which has biological activity.

5) The control of cell-type-specific gene expression

At least four substances are known to influence the choice of cellular differentiation pathways in aggregates; cAMP, adenosine, DIF and ammonia. The effect of diffusible molecules on cellular differentiation, judged by histological and immunological criteria, is parallelled by their effect on cell-type specific gene expression. Therefore, understanding the mechanisms by which these substances modify the expression of individual genes will help to determine how they act to form a differentiated structure.

a) Patterns of cell-type-specific gene expression

Developmentally regulated genes in *D.discoideum* have been isolated by differential screening of cDNA libraries, or by the inclusion of an excess of cold competitor RNA from

growing cells (Williams et al., 1979; Rowekamp and Firtel, 1980; Mangiarotti et al., 1981; Barklis and Lodish, 1983; Mehdy et al., 1983; Jermyn et al., 1987). A high proportion of aggregation stage clones have been found to encode cell-type-enriched transcripts (Barklis and Lodish, 1983; Mehdy et al., 1983).

The cDNA clones encoding prestalk-enriched mRNA sequences originally isolated were divided into two classes, I and II, on the basis of timing of expression and the existence of at least some features of coordinate control (Barklis and Lodish, 1983). However, objections can be raised to this simple classification; for instance, D14 was found to differ from other class I genes, being unresponsive to cAMP in NC4 cells (Oyama and Blumberg, 1986a). More importantly, the degree of prestalk-enrichment of these sequences is significantly lower than for the more recently cloned, DIF-inducible mRNA sequences, used as markers of the stalk cell pathway of differentiation (Jermyn et al., 1987). Unlike the DIF-induced genes, the prestalk class I and II genes are expressed prior to overt cellular differentiation, and therefore may be initially expressed in all cells before selective loss in prespore cells.

A similar classification has been proposed for prespore-enriched mRNA sequences. The class I cDNA clones D18 and D2 encode mRNAs which accumulate during aggregation (Chisholm et al., 1984). Like the class I prestalk genes, they are expressed prior to cellular differentiation; in bacterially-grown cells, expression of these genes reaches a peak in the loose mound and declines rapidly just as prespore cells can

be detected (Oyama and Blumberg, 1986a). Therefore a similar objection to a cell-type-specific classification may apply. However, most of the prespore-specific genes isolated have a similar profile of mRNA accumulation during development, typified by the mRNA sequence encoded by the class II cDNA clone D19 (Chisholm et al., 1984). Expression is first detectable in tight aggregates just prior to tip formation, and hence is coincident with the appearance of overt prespore differentiation.

b) Cyclic AMP and cell-type-specific gene expression

Although cAMP performs multiple functions during development, several observations have suggested that it is the major inducer of prespore differentiation. Sporogenous mutants of the strain V12M2 require only the addition of cAMP to form spores when incubated at low cell density (Kay, 1982). The addition of high levels of cAMP to aggregates of Percoll gradient purified prespore cells prevents normal regulation to produce prestalk cells (Weijer and Durston, 1985). In intact slugs, cAMP can induce the conversion of prestalk to prespore cells. If slugs are allowed to migrate on agar containing 1mM cAMP, the total proportion of prespore-staining tissue increases by 5%, reducing the amount of non-staining tissue adjacent to the substratum and at the extreme rear (Schaap and Wang, 1986). Recently it has been shown that reducing extracellular cAMP levels in the slug by implantation of cAMP-phosphodiesterase-containing beads induces local dedifferentiation of prespore cells (Wang et

al., 1988b). A general reduction in cAMP, achieved by submerging slugs in a cAMP phosphodiesterase-containing solution, causes a significant reduction in prespore staining cells (Wang et al., 1988b).

Cyclic AMP is essential for the control of gene expression throughout the developmental cycle. The mechanism employed seems to be dependent on the developmental stage. Early during development, when cAMP signalling begins, a set of genes expressed during growth are repressed by pulses of cAMP (Kimmel and Carlisle, 1986). Genes induced during early development such as discoidin 1 are unaffected by pulses of cAMP, but are switched off when exposed to a high level of cAMP during aggregate formation (Williams et al., 1980). In contrast, other early genes are induced by pulses of cAMP; some of these are inhibited by the continuous presence of extracellular cAMP (Mann and Firtel, 1987). Still other genes can be prematurely induced by addition of a high level of extracellular cAMP (Williams et al., 1980; Haribabu et al., 1986). Expression of most post-aggregative genes has been found to require the presence of extracellular cAMP (Chung et al., 1981; Barklis and Lodish, 1983; Mehdy et al., 1983).

There is evidence for coordinate control of the expression of groups of prespore and prestalk genes by cAMP (Barklis and Lodish, 1983; Mehdy et al., 1983; Chisholm et al., 1984; Oyama and Blumberg, 1986a+b). Differences in the dependence of specific genes on cAMP, both for induction and maintenance of expression, can be correlated with their cell-type-specificity. A major difference is the response to dissociation of aggregates into single cells. Disaggregation

causes a specific loss of transcripts derived from many aggregation-stage genes. The transcript level can be restored by addition of cAMP to dissociated cells, and in many cases the original loss can be prevented by inclusion of high levels of cAMP during disaggregation (Chung et al., 1981; Barklis and Lodish, 1983; Mangiarotti et al., 1983a). On the whole, prespore-specific messages, especially class II, are lost much more rapidly than prestalk messages, and require higher levels of cAMP to prevent loss or to restore levels; often only a partial restoration is achieved by addition of cAMP.

There is evidence that the primary cause of this loss is mRNA destabilisation. The mRNA half-life decreases from 3-4 hours to 20-30 minutes (Mangiarotti et al., 1982), while the rate of nuclear transcription decreases 5-10 fold (Landfear et al., 1982). It has been suggested that during normal development, cAMP acts to stabilise these messages (Mangiarotti et al., 1983b). The transcripts shown to be synthesised early in development but not accumulated until after aggregation (Mangiarotti et al., 1985; see section I.3c) also seem to be derived from prespore-enriched genes where analysed (A.Ceccarelli, personal communication). However, the relatedness of mechanisms acting early and late in development is not yet clear.

If adenylate cyclase activity, and hence cAMP relay, is inhibited with caffeine or conditions of high osmolarity, post-aggregative gene expression can be restored by adding extracellular cAMP (Gomer et al., 1986; Oyama and Blumberg, 1986b). However, some genes normally repressed by cAMP

continue to be expressed (Kimmel, 1987). This suggests that cAMP is acting via at least two distinct mechanisms. Several possible mechanisms by which extracellular cAMP could act to modify gene expression in *D.discoideum* could be envisaged. Entering cells directly, it could interact with the regulatory subunit of cAMP-dependent protein kinase, so releasing the catalytic subunit. Subsequent phosphorylation events could be of importance, or the regulatory subunit could be involved directly. There is evidence that the R subunit reversibly translocates to the nucleus (Woffendin et al., 1986). Alternatively, on entering cells, cAMP might bind to a distinct protein to mediate its effect on gene expression; novel cAMP binding proteins have been identified (Tsang and Tasaka, 1986). A third possibility is that cAMP acts through the cell surface receptor. Binding to the receptor could either stimulate adenylate cyclase to produce intracellular cAMP, again activating protein kinases, or generate alternative second messengers such as calcium or cGMP.

There is strong evidence that cAMP induction of postaggregative gene expression is mediated by the cell surface receptor (Schaap and Van Driel, 1985; Haribabu and Dottin, 1986; Oyama and Blumberg, 1986b; Gomer et al., 1986). The ability of cAMP analogues to induce postaggregative gene expression can be directly correlated with their ability to bind to the cell surface receptor. In contrast, there is no correlation with ability to permeate the cell or to activate cAMP-dependent protein kinase. This implies that extracellular cAMP does not act by entering the cell

directly, although it does not rule out an involvement of intracellular cAMP.

By inhibiting extracellular phosphodiesterase activity, the extracellular cAMP concentration required for gene induction can be assessed (Gomer et al., 1986; Oyama and Blumberg, 1986b). 1-2μM is required for half-maximal prespore mRNA accumulation, but only 320-520nm is necessary for half-maximal mRNA accumulation of other post-aggregative genes (Oyama and Blumberg, 1986b). The latter value is in the range of the K_D of the low affinity B form of the receptor. Therefore prespore gene induction may require binding of cAMP to a still lower affinity form, perhaps created by a change in receptor environment in aggregates, or possibly a distinct protein. There is some evidence for the existence of such a receptor during the slug stage (Schaap and Spek, 1984). Alternatively, prespore gene expression may simply require increased receptor occupancy.

The second messenger(s) involved in cAMP-mediated induction of postaggregative gene expression has not been identified. Studies of a mutant lacking intracellular cGMP phosphodiesterase argue against the involvement of elevated cGMP levels, as cAMP requirements are not reduced (Schaap et al., 1986). As previously mentioned, inhibition of adenylate cyclase and hence intracellular cAMP production by caffeine can be bypassed (Gomer et al., 1986; Oyama and Blumberg, 1986b). In a mutant, N7 (= synag7), in which adenylate cyclase cannot be activated by cAMP and hence relay is blocked, addition of cAMP can still induce post-aggregative gene expression (Schaap et al., 1986). There is preliminary

evidence from drug experiments that postaggregative gene expression requires the mobilisation of intracellular calcium (Schaap et al., 1986). In amoebae, cAMP has been shown to stimulate the accumulation of inositol trisphosphate, known to release calcium from non-mitochondrial stores (Europe-Finner and Newell, 1986).

The requirement for higher levels of extracellular cAMP for prespore gene expression than for other post-aggregative genes predicts that the cAMP concentration is highest in the prespore zone. Unexpectedly, in slugs the total cAMP concentration is about 50% higher in the prestalk zone (Brenner, 1977). However, such measurements cannot distinguish between intracellular and extracellular levels. A model predicting the distribution of extracellular cAMP in whole slugs has been proposed in which the concentration increases from tip to rear (Wang et al., 1988b). Due to continuing cAMP signalling in the aggregate, concentrations of 1μM or higher could be accumulated around cells. This is in a range suitable to bind to a very low affinity receptor. In combination with the preferential conversion of cAMP to adenosine at the tip and lowered phosphodiesterase activity in the rear, a gradient of cAMP would be created. The higher average level of cAMP in the posterior zone would then reduce cAMP signalling due to adaptation. Quenching of pulsing would result in a more constant level of cAMP than in the anterior zone.

c) Other factors regulating cell-type-specific gene expression

The model described above also predicts a gradient of adenosine which decreases from tip to rear (Wang et al., 1988b). Exposure of slugs to high adenosine levels does not increase the total proportion of non-prespore staining cells, but does affect patterning (Schaap and Wang, 1986). When slugs migrate on adenosine-containing agar, a redistribution of non-prespore cells from the anterior to the base occurs, while prespore staining is generally weaker. In contrast, reducing adenosine levels by treatment with adenosine deaminase results in the appearance of prespore cells within the prestalk zone. In aggregation-competent cells in shaken suspension, which have the ability to express certain postaggregative genes on addition of cAMP, adenosine acts as a competitive inhibitor of cAMP-induced prespore gene expression (Schaap and Wang, 1986). This is thought to be due to inhibition of cAMP binding to the A form of the receptor (Theibert and Devreotes, 1984), so reducing adenylate cyclase activity and overall cAMP levels (Schaap and Wang, 1986). Therefore adenosine may be acting indirectly by interfering with the cAMP-mediated induction of prespore gene expression (see Chapter 8).

DIF, or differentiation inducing factor, was first identified as a low molecular weight substance secreted by cells and able to induce stalk cell formation when added to cells incubated at low density (Town et al., 1976). Using HPLC in combination with the stalk cell assay, five peaks of activity were detected (Kay et al., 1983). The structure of DIF-1, derived from the most biologically active peak, has

been determined (Morris et al., 1987). It has an unusual structure, consisting of a dichloro-dihydroxy-methoxy substituted aromatic ring with an alkyl side chain. It is both water and lipid soluble. A mutant strain, HM44, has been isolated which is deficient in DIF production but can respond to added DIF (Kopachik et al., 1983). It is normally arrested at the loose mound stage of development, but in submerged monolayer culture, there is a greater than 90% efficiency of stalk cell formation on addition of DIF. The requirement of DIF for stalk cell formation is reflected at the level of gene expression. Three genes have been isolated encoding mRNAs dependent on DIF for their induction; pDd63, 56 and 26 (Jermyn et al., 1987; Williams et al., 1987; Ceccarelli et al., 1987; McRobbie and Ceccarelli, 1988). Within limits of detection, they are expressed exclusively in prestalk and stalk cells. The mechanism of DIF action is unknown, but for cells to develop responsiveness to DIF, pretreatment with cAMP is required (Town et al., 1976; Sobolewski et al., 1983).

As well as being required for stalk cell formation, DIF can divert cells from the prespore to the prestalk pathway. Sporogenous mutants will form stalk cells instead of spore cells if DIF is added, and DIF induces isolated prespore cells to redifferentiate into stalk cells (Kay and Jermyn, 1983). Exposure of slugs to DIF causes a large reduction in the proportion of prespore staining tissue (D.Traynor, personal communication). Unexpectedly, the level of DIF is at least two-fold higher in the prestalk than the prespore zone

(Brookman et al., 1987), but it is not known whether it is all in a biologically active form.

Ammonia, produced by the degradation of proteins during development, is suggested to act as an antagonist of DIF (Gross et al., 1983). It seems to inhibit the differentiation of prestalk into stalk cells. Migrating slugs can be induced to culminate by enzymatically depleting them of ammonia; loss of ammonia may be the signal for fruiting body formation during normal development (Schindler and Sussman, 1977). Ammonia also seems to facilitate prespore differentiation; it is required for spore cell formation in large aggregates in submerged culture (Sternfield and David, 1979), while addition of ammonium salts to a sporogenous mutant increases the proportion of cells forming spores (Gross et al., 1983). When differentiation competent cells are shaken rapidly in a medium which inhibits cAMP relay, cAMP and ammonium salts are both required to optimise and maintain the expression of prespore-specific genes (Oyama and Blumberg, 1986c).

Other factors have been detected which apparently stimulate prespore differentiation. However, SPIF (spore inducing factor, Wilkinson et al., 1985), now known to be methionine (D.Ramji, personal communication), and PIF (prespore inducing factor, Kumagai and Okamoto, 1986), which is incompletely characterised, do not seem to be absolutely required, as sporogenous mutants can form spores in the presence of cAMP alone. The ability of isolated cells to form spores also argues against a requirement for cell-cell contact for prespore gene expression. It has been reported that sustained cell association is required for the

expression of prespore genes, as some prespore genes cannot be induced by cAMP until the time of aggregate formation (Mehdy et al., 1983; Chisholm et al., 1984). The requirement for multicellularity can also be circumvented by attachment to plastic petri dishes, which suggests that specific cell-cell interactions are not necessary (Mehdy and Firtel, 1985).

6) Regulatory sequences upstream of *D.discoideum* genes

In *Dictyostelium*, regions important for regulated gene expression have often been proposed simply on the basis of sequence examination. In some cases, comparison of coexpressed genes and members of the same gene family has been possible (Poole and Firtel, 1984; Romans and Firtel, 1985). However, structural similarities within gene families may simply reflect evolutionary conservation. It is often assumed that structural features such as G-C rich motifs or dyad symmetry elements are likely to be important for transcriptional regulation. Hypersensitive site mapping has provided more solid evidence of functional significance (Pavlovic et al., 1986; Ayres et al., 1987), which can now also be tested more directly using the transformation system (Nellen et al., 1984).

The *D.discoideum* actin gene family consists of about 20 members, most of which are expressed (Romans and Firtel, 1985). The 5' flanking sequences of 15 genes were compared, and on this basis they were divided into four groups, reflecting to some extent related expression patterns. Short

GC-rich repeats were identified, located between 100 and 300 nucleotides from the initiation codon, some of which form part of dyad symmetry elements.

The expression of deleted versions of two genes, actins 6 and 15, was analysed by DNA-mediated transformation (Cohen et al., 1986; Nellen et al., 1986). For actin 15, 182 nucleotides of 5' non-coding sequence, including two G-C rich regions, is sufficient for normally regulated transcription (Cohen et al., 1986). A highly conserved copy of the sequence AAAAATGGGATTTTTT, designated a box 4 homology (Romans and Firtel, 1985), is found upstream of actin 6 and 15. For both genes, removal of part of this sequence abolishes expression. In addition, a GC-rich region approximately 600 nucleotides upstream of actin 6 was identified which increases transcription by a factor of 10 (Nellen et al., 1986).

A G-rich sequence element is required for efficient expression of a developmentally regulated gene inducible by cAMP, cysteine proteinase 2 (CP2; Pears and Williams, 1987; Datta and Firtel., 1987). Within 921 nucleotides upstream of the start-site of transcription, the only region with a relatively GC-rich composition is centered at position -235 (Pears and Williams, 1987). This region contains two adjacent homologous G-rich elements, 9 nucleotides in length. The distal element forms part of a region of weak dyad symmetry. The possible importance of these G-rich repeats was suggested by a comparison with the 5' flanking sequences of a coregulated gene, cysteine proteinase 1 (CP1; Driscoll and Williams, 1987). Two direct repeats 80 nucleotides in length are found 220 and 320 nucleotides upstream of the cap site.

Each contain a G-rich core, identical save for one nucleotide substitution. They are similar to the direct repeats upstream of CP2, although no strict consensus exists. The G-rich blocks in CP1 lie close to the mapped positions of two developmentally regulated DNAse I hypersensitive sites (Pavlovic et al., 1986).

Deletion of both CP2 elements or deletion of the distal element alone reduces expression by 50-fold on cAMP induction, while a single point mutation in the distal element causes a 20-fold reduction (Pears and Williams, 1987). As a low level of cAMP inducibility was retained, it suggests either that developmental regulation is conferred solely by a second element, or that a second, weakly acting element in conjunction with the G-rich region confers cAMP inducibility.

The CP1 and CP2 elements are functionally equivalent (Pears and Williams, 1988). Insertion of a 68 base-pair oligonucleotide, spanning most of one of the CP1 direct repeats, upstream of the CP2 gene lacking both G boxes restores efficient cAMP-inducibility. Two copies of the core element also restore full expression, while one copy is partially effective. However it is not clear whether this G-rich element acts as a promoter, upstream activating sequence or enhancer (C.Pears, Ph.D. thesis, Univ. of London).

Similar G-rich sequences are found upstream of several other developmentally regulated genes (Barklis et al., 1985; Haribabu et al., 1986; Ayres et al., 1987; Datta and Firtel., 1987; see Figure 6.7). However, these genes are not coregulated. In the case of SC253 (Ayres et al., 1987;

=pDd56) a developmentally-regulated DNAse I hypersensitive site has been mapped to approximately the same position as the G box. For the other genes, evidence of functional importance must await transformation studies or identification of trans-acting factors.

II The control of cell-type-specific gene expression in other eukaryotes

There are a variety of levels at which genes expressed in a cell-type-specific manner could be regulated; initiation of transcription; mRNA processing, transport or stability; or protein synthesis (reviewed by Darnell, 1982). The best characterised cell-type-specific mechanisms are those involving cis-acting regulatory DNA sequences, which are targets for binding proteins conferring selective transcription by RNA polymerase II (reviewed by Maniatis et al., 1987). Other regulatory mechanisms operating in a cell-type-specific manner have been described; for instance, the increased accumulation of immunoglobulin heavy chain mRNA during the differentiation of B lymphocytes can be largely accounted for by post-transcriptional mechanisms (Gerster et al., 1986). Post-transcriptional mechanisms may regulate the level of several tissue-specific mRNAs during hepatocyte differentiation in mice (Powell et al., 1984), but the appearance of most liver-specific mRNAs is under transcriptional control. Post-transcriptional control of many other genes has been reported, particularly modulation of

mRNA stability. However, they are often "housekeeping" genes known to be cell-cycle regulated (Ayusawa et al., 1986; Graves et al., 1987; Seiler-Tuyns and Paterson, 1987; Coppock and Pardee, 1987), or are required to respond rapidly to an inducing signal such as a growth factor (Dani et al., 1985; Greenberg et al., 1986). Hence they do not show marked cell-type-specificity. Exceptions to this are some tissue-specific mRNAs under the control of steroid hormones, such as vitellogenin (Brock and Shapiro, 1983). Little evidence has been presented to suggest that translational control is of major importance for tissue specificity. During myeloid differentiation of HL-60 cells to form macrophages and neutrophils, the majority of gene products seem to be regulated at the translational level, but this is not the case for other terminal differentiation steps examined (Salser et al., 1987). Thus it appears that control of transcription initiation is the mechanism utilised most widely to regulate cell-type-specific gene expression.

1) The structure of eukaryotic promoters

Although the mechanism by which RNA polymerase II transcription is controlled is not fully characterised, it is seemingly highly conserved among eukaryotes (reviewed by Dynan and Tjian, 1985; McKnight and Tjian, 1986). This is reflected by similarities in promoter organisation and in the function of trans-acting factors. Three components of eukaryotic promoters have been identified: the TATA box, which controls the accuracy of transcription initiation, and

promoter elements and enhancers, which modify the rate of transcription. For most eukaryotic genes analysed, multiple upstream elements have been identified (Serfling et al., 1985a). Enhancers differ from promoter elements in their ability to act in an orientation-independent manner and at great distances, although the distinction between these two classes is becoming increasingly blurred. Multimerization of some promoter elements can confer enhancer-like properties (Searle et al., 1985; Bienz and Pelham, 1986), while in yeast there is only one major class of upstream promoter element (upstream activating sequences), which resemble mammalian enhancer sequences in many respects (Struhl, 1987; Guarente, 1988).

a) The TATA box

The consensus sequence TATAAA is located 25-30 bp from the site of mRNA initiation of most higher eukaryotic genes and is required for accurate initiation of transcription by RNA polymerase II (Breathnach and Chambon, 1981). In yeast, the spacing is apparently less constrained (the TATA box is 40-120 bp from the cap site), but an initiator element exists close to the start-site (Struhl, 1987). The mechanism of transcription initiation is little understood, and factors binding to the TATA box have not been extensively characterised. Chromatographic fractionation and biochemical analysis of mammalian cell extracts have identified four separate activities required to reconstitute accurate transcription in vitro (Matsui et al., 1980; Samuels et al., 1982). One of these factors, TFIID, can be substituted by a

protein purified from yeast (Buratowski et al., 1988), demonstrating the high degree of evolutionary conservation.

Some promoters analysed lack a TATA box but instead have G-C rich sequences close to the start site of transcription, and so a distinct mechanism of transcription initiation may exist. These promoters frequently direct transcription of "housekeeping" genes (reviewed by Dynan, 1986).

b) Proximal promoter elements

Upstream promoter elements, are typically located within about 150 bp of the TATA box and are 8-12 bp in length (reviewed by Dynan and Tjian, 1985; Sassone-Corsi and Borelli, 1986; McKnight and Tjian, 1986). They generally function to increase transcriptional efficiency, but some negatively acting elements have been identified (Osborne et al., 1985; Bouvagnet et al., 1987). Mutagenesis studies have suggested that promoter strength is determined by the number and type of upstream elements, which implies that they can act in a combinatorial manner to promote transcription. Some upstream promoter elements are widely distributed, and are also a common component of viral promoters (reviewed by Jones et al., 1988). This presumably enables the virus to utilise the transcription apparatus of the host. Other promoter elements are apparently gene-specific or confined to a set of genes with a shared response to an inducing signal. Although there is evidence that interactions occur between proteins binding to the TATA box and upstream elements (Takahashi et al., 1986), paradoxically some promoter elements can function independently of orientation (Everett et al., 1983). In many

cases proteins binding specifically to these sequences have been extensively purified.

i) The CCAAT box

The CCAAT box was originally identified as an important RNA polymerase II promoter element from studies of vertebrate ß-globin genes (Efstratiadis et al., 1980), but it is also found 60-80 bp upstream of a large number of other eukaryotic genes (Benoist et al., 1980; McKnight and Tjian, 1986). The same element can serve as as a control sequence in viral origins of DNA replication (Nagata et al., 1983; Rawlins et al., 1984), and it appears that a highly related binding protein or set of proteins is involved (Jones et al., 1987). The factor which activates transcription was designated CTF (CCAAT box transcription factor; Jones et al., 1985), while NF-1 (nuclear factor 1) binds to the adenovirus origin of replication (Rawlins et al., 1984; Leegwater et al., 1985). It was suggested that as CTF and NF-1 were biochemically and functionally indistinguishable, they could be the same protein(s). Purification of human CCAAT-binding factors reveals a series of polypeptides ranging from 52K to 66K in size, and identical heterogeneity is obtained with NF-1 (Jones et al., 1987). Isolation of cDNA clones indicates that at least some of the heterogeneity seen in preparations of the two factors is a result of differential splicing of a single protein (Santoro et al., 1988). However, distinct CCAAT box binding proteins with differing binding specificities have been identified (Dorn et al., 1987; Raymondjean et al., 1988; Chodosh et al., 1988a). Although the functional significance of these multiple factors is not

yet understood, each protein is apparently composed of two different subunits, so increasing the potential for regulation (Chodosh et al., 1988a). Two *S.cerevisiae* gene products, HAP2 and HAP3, also recognise CCAAT-containing transcription elements as a heterologous complex (Keng and Guarente, 1987). Moreover, subunits of a mammalian binding protein, CP1, are functionally interchangeable with those of HAP2/HAP3 (Chodosh et al., 1988b). The CCAAT box has no enhancer-like properties; multiple copies cannot stimulate transcription at a distance (Bienz and Pelham., 1987).

ii) Sp1 binding sites

The hexanucleotide GGGCGG has been defined as the binding site of an extensively characterised transcription factor, Sp1 (Myers et al., 1981; Fromm and Berg, 1982; Briggs et al., 1986). Sp1 was originally identified as a requirement for SV40 transcription (Dynan and Tjian, 1983). The deletion of three 21 bp G-C rich direct repeats, each containing two copies of the hexanucleotide 5'-GGGCGG-3' and located about 50-100 bp from the start site of early viral transcription, severely curtails early promoter function (Myers et al., 1981; Fromm and Berg, 1982; Everett et al., 1983). Interestingly, although this element is obviously asymmetric, the 21 bp repeats retain activity even if inverted (Everett et al., 1983; Gidoni et al., 1985). Sp1 binds the hexanucleotide in a sequence-specific manner (Dynan and Tjian, 1985). It is now clear that Sp1 binding sites are associated with many cellular as well as viral genes, and in some cases Sp1 binding has been demonstrated (Dynan et al, 1986). Comparison of different genes indicates that the

affinity of Sp1 binding is dependent on the sequences which flank the hexanucleotide core (McKnight and Tjian, 1986). A common regulatory feature is not evident among the diverse promoters containing actual or potential Sp1 binding sites, and it is unclear why other genes lack Sp1 sites. A class of genes which commonly contain potential Sp1 binding sites are housekeeping genes, such as hypoxanthine phosphoribosyl transferase (HPRT) and dihydrofolate reductase (DHFR) (Dynan, 1986). A cDNA clone encoding Sp1 has now been isolated (Kadonaga et al., 1987).

iii) Other promoter elements

Other sequences close to the TATA box of individual genes have been identified which seem to be required for basal promoter activity; if they are deleted or mutated, transcription is reduced or abolished (Myers et al., 1986; Quinn et al., 1988). However, some of these sequences are also responsive to inducing agents. For example, deletion analysis suggests that the cAMP-response element (CRE), an 8 bp palindrome highly conserved upstream of set of genes rapidly activated by cAMP, is also required as a basal promoter element for some genes (Short et al., 1986; reviewed by Roesler et al., 1988). The sole function of other promoter elements is to activate transcription in response to inducing signals, such as heat shock. The consensus heat shock element (HSE) is C--GAA--TTC--G (Bienz, 1985; Pelham, 1985). In the Drosophila hsp70 promoter, which has been studied most extensively, a copy of this element is found 15 bp upstream of the TATA box (Pelham, 1982). In the Xenopus hsp70 gene, which contains three HSEs further upstream but lacks the

proximal element, the HSE-TATA box distance can be varied without loss of activity (Bienz and Pelham, 1986). The Xenopus HSEs can also confer heat-inducibility to a heterologous gene in either orientation, whereas a fragment of the Drosophila promoter spanning its two most proximal HSEs has only very weak activity unless multiple copies are present. Thus, studies of the heat shock regulatory demonstrate that the same sequence can function either as a proximal promoter element or as an inducible enhancer.

Binding proteins have been identified and purified for both the CRE (CREB) and HSE (HSTF) (Montminy et al., 1987; Parker and Topol, 1984; Kingston et al., 1987). There is evidence that dimers of HSTF bind to DNA co-operatively, resulting in a conformational change of the DNA backbone (Shuey and Parker, 1986a+b). CREB also activates transcription as a dimer; dimerisation is favoured by phosphorylation of the protein (Yamamoto et al., 1988; see section II.2c(ii)).

c) Enhancers

Enhancers are cis-acting regulatory elements that function independently of distance and position relative to the start site of transcription (Serfling et al., 1985a). They were first identified within the the regulatory region of the DNA tumour virus SV40 (Banerji et al., 1981; Moreau et al., 1981) and are characterised by the presence of multiple, short sequence motifs 8-10 bp in length. The complete SV40 enhancer spans ~ 100bp and genetic evidence suggests the presence of at least three discrete 15-22 bp elements that

co-operate to enhance transcription (Herr and Clarke, 1986; Clarke and Herr, 1987). These elements overlap shorter motifs, such as the "core" motif identified by comparison with many other enhancers (Laimins et al., 1982; Sassone-Corsi and Borrelli, 1986), and the sph motif. Mutagenesis has indicated that these and other sequences are functionally important (Zenke et al., 1986), and many have been shown to represent binding sites for trans-acting factors (Davidson et al., 1986; Wildeman et al., 1986; reviewed by Jones et al., 1988).Interaction with trans-acting factors has been proposed as the possible mechanism of enhancer activity (Serfling et al., 1985). The absence of any spacing restrictions between enhancers and proximal promoter elements may be due to the ability of enhancer-protein complexes to associate with the transcriptional apparatus by DNA looping (Ptashne, 1986).

The modular arrangement of the SV40 enhancer is supported by experiments demonstrating that individual sequence elements can function independently when present as multiple copies (Schirm et al., 1987). Interestingly, although there is no obvious distinction between constitutive and cell-type-specific elements, different regions are active in different cell-types. This suggests that the distribution of trans-acting factors may have distinct cell-type specificity. A model has been proposed, based on the results of manipulating sequence motifs, in which several levels of organisation exist (Ondek et al., 1988). According to this model, the SV40 enhancer is composed of multiple enhancer elements, each with a bipartite structure of two "enhansons". The spacing between enhansons is constrained due to the requirement for

association of two binding factors to give an active enhancer element. This may explain the presence of "common" motifs in enhancers mediating specific activation. Alternatively, some binding factors may contain fully functional activation domains, obviating the need for a bipartite structure (Chiu et al., 1987; Ondek et al., 1988).

In addition to the core sequence, which is not a universal feature of all enhancers, other SV40 motifs are present upstream of various viral and cellular genes. The presence of some of these motifs has been linked to responsiveness to specific inducing agents, such as cAMP and phorbol esters. A trans-acting factor, activator protein 1 (AP-1), has been purified which binds to a 9 bp sequence that can confer TPA inducibility to a heterologous gene (Angel et al., 1987; Lee et al., 1987). AP-1 was therefore judged to be an enhancer binding protein. Affinity-purified preparations of AP-1 consist of a series of related polypeptides ranging in size from 40 to 47 kD, and so AP-1 may constitute a family of closely related DNA binding proteins (Lee et al., 1987; Bohmann et al., 1987). One of the polypeptides is probably the product of the proto-oncogene c-jun (Bohmann et al., 1987; Angel et al., 1988), which was recently shown to be identical to the fos-associated protein p39 (Rauscher et al., 1988). As predicted, interaction of the c-fos protein with AP-1 stimulates transcription of genes which possess the AP-1 binding site (Chiu et al., 1988; Sassone-Corsi et al., 1988). The enhancer binding protein AP-2, which recognises multiple sites in the SV40 control region and also binds upstream of many other genes, confers responsiveness to both TPA and cAMP

(Mitchell et al., 1987; Imagawa et al., 1987). Thus the AP-2 binding site represents a second distinct cAMP response element. The distribution of AP-1 appears to be more ubiquitous than that of AP-2 (Imagawa et al., 1987).

The prototype inducible enhancer is the hormone-dependent element present in the long terminal repeat of the murine mammary tumour virus (MMTV). When inserted in either orientation at various distances from a heterologous gene, enhancer activity is dependent on the presence of an active glucocorticoid hormone-receptor complex (Chandler et al., 1983). Similar binding sites are present upstream of steroid-inducible cellular genes; for instance, two copies of a glucocorticoid response element (GRE) are located 2.5 kb upstream of the transcription initiation site of the tyrosine aminotransferase gene (Jantzen et al., 1987). It has been established that a minimal GRE 15 bp in length is sufficient to confer hormone sensitivity to a heterologous promoter (Strahle et al., 1987). The sequence can be varied, remaining functional with either perfect or only partial symmetry, and is also recognised by the progesterone-receptor complex. The minimal oestrogen response element is a related but distinct 13-15 bp palindrome, implying mechanisms of induction are very similar (Klock et al., 1987; Martinez et al., 1987). Other inducible enhancers include those responsive to such factors as heavy metals (Serfling et al., 1985b), viral infection (Goodbourn et al.,1986) and growth factors (Treisman, 1985). Most seem to be regulated by positive trans-acting factors, but negative mechanisms may be utilised, e.g. the human ß-interferon (ß-IFN) gene. The ß-IFN

gene regulatory element consists of two adjacent domains; one positively acting and one negative (Goodbourn et al., 1986). The negatively acting element can repress transcription from a heterologous promoter, while the positive element acts as a constitutive enhancer.

The similarity between the function of the upstream activating sequences of yeast and mammalian enhancers is reflected by the functional equivalence of proteins binding to them, and suggests at least one mechanism of transcriptional activation has been evolutionarily conserved in higher and lower eukaryotes. GCN4 induces the co-ordinate expression of genes required for amino acid biosynthesis (Hope and Struhl, 1985; Arndt and Fink, 1986), binding as a dimer to a 9 bp dyad sequence (Hill et al., 1986). The C-terminus of v-jun has 44% identity with the DNA binding domain of GCN4 (Vogt et al., 1987), and the core of the consensus sequence recognised by GCN4 is identical to the binding site of AP-1, the cellular homologue of v-jun (Struhl, 1987). The activation domain of GCN4 is distinct from the DNA binding domain (Hope and Struhl, 1986, 1987), and the JUN protein is capable of substituting for both functions in yeast (Struhl, 1988). GAL4 induces transcription of glucose metabolizing genes by binding to a 17 bp region of imperfect dyad symmetry (West et al., 1984; Yokum et al., 1984; Bram et al., 1985; Giniger et al., 1985). Although no obvious primary sequence homology exists with the GCN4 protein, GAL4 also has distinct DNA binding and acidic transcription activation domains (Brent and Patashne, 1985; Keegan et al., 1986). GAL4, when expressed in *Drosophila*

larvae, stimulates transcription of a *Drosophila* promoter linked to GAL4 binding sites (Fischer et al., 1988). This result predicts that the GAL4 protein also has a close homologue in higher eukaryotes. In addition it suggests that the apparent complexity of some mammalian enhancers may not be necessary for their function as transcriptional activators, but instead to achieve complex patterns of gene expression with a limited number of regulatory proteins.

2) Transcriptional control of cell-type-specific gene expression

A variety of mechanisms are used to regulate tissue-specific transcription. This is partly because patterns of expression vary; some genes are expressed in one cell type only, while others are subject to more complex developmental and tissue-specific regulation. Tissue-specific enhancers have been frequently identified, and often act in conjunction with cell-type-specific promoters. There is evidence for co-ordinate control of genes expressed in the same cell type by shared transcription factors, but many genes are controlled by multiple regulatory elements, allowing overlapping but distinct expression patterns. Some genes are regulated by positively acting factors, while the cell-type distribution of others is limited by the activity of negative factors.

a) Cell-type-specific enhancers

i) *Drosophila* yolk protein gene expression

The divergently transcribed yolk protein genes yp1 and yp2 are expressed in the fat bodies and ovaries of adult females as a result of two distinct cis-acting elements acting on both genes (Garabedian et al., 1985). Thus it appears that different factors in different tissues regulate specific genes. The fat body element, which has been localised to a 125 bp DNA fragment located 196 bp upstream of the yp1 cap site, can confer fat body-specific activity to a heterologous promoter (Garabedian et al., 1986). The element is still active when inverted and when placed at varying distances from the heterologous promoter. Therefore the stage-, sex- and tissue-specific expression of the yp1 gene is controlled by a positively acting enhancer.

ii) Immunoglobulin gene expression

The restricted expression of immunoglobulin heavy and light chain genes to the B cell lineage is perhaps the best characterised example of tissue-specificity, but it is still incompletely understood. For the murine heavy chain gene, three distinct sequence elements have been identified which contribute to tissue specificity; an enhancer, a promoter and a region of intragenic DNA. The function of the enhancer was first investigated by deletion analysis, and a minimal region of 140 bps was identified as necessary to activate B cell-specific expression (Gillies et al., 1983; Banerji et al., 1983). However, in the absence of the enhancer, cell-type-specificity was retained by fusion of promoter sequences to the SV40 enhancer (Mason et al., 1985). Intragenic sequences involved in the control of cell-type-specificity may be required for post-trancriptional control, as during B cell

differentiation, enhancer activity and transcription can be detected before mRNA accumulation (Grosschedl and Baltimore, 1985; Gerster et al., 1986). It has been suggested that the enhancer is important for the induction of cell-type-specific expression (Wang and Calame, 1986), while the promoter is primarily involved in the maintenance of cell type specificity (Mason et al., 1985).

The kappa light chain gene is co-expressed with the mu heavy chain gene. The only conserved promoter element, although present in the opposite orientation, is an octamer located at an analogous position, 70 bp upstream of the transcriptional start site (Parslow et al., 1984; Falkner and Zachau, 1984). For both genes, the octamer ATTTGCAT is essential for transcriptional activity, and as this motif is also found upstream of other heavy and light chain genes, an important B-cell-specific function is suggested. The heavy chain enhancer also contains an octamer sequence (Falkner and Zachau, 1984), which may explain the synergistic activity of the promoter and enhancer in controlling immunoglobulin gene expression (Garcia et al., 1986). Moreover, a synthetic promoter consisting of an octamer upstream of a TATA box is sufficient for lymphoid-specific promoter activity (Wirth et al., 1987).

The heavy chain gene enhancer is located within the intron separating the V region J segment and the constant region (Gillies et al., 1983; Banerji et al., 1983; Neuberger, 1983). Initially, in vivo footprinting experiments identified four discrete binding sites within the enhancer in addition to the octamer sequence, observed only in lymphoid

cells (Ephrussi et al., 1985). These were designated E1 to E4, and each contain a conserved sequence, CAGGTGGC. It was therefore assumed that all four sites would be recognised by the same B cell-specific factor. However, more recent in vitro studies have suggested that of each these sites is recognised by a different factor (Augereau et al.,1986; Peterson et al., 1986; Schlokat et al., 1986; Weinberger et al., 1986). Although binding is specific to lymphoid cells, factors with identical binding specificities have been detected in cells of a non-lyphocytic lineage. This has led to the suggestion that the heavy chain gene enhancer is in an inactive conformation in other cells and hence inaccessible for binding (Weinberger et al., 1986; Sen and Baltimore, 1986). Therefore the function of tissue-specific binding proteins may be to activate the DNA, resulting in an open chromatin configuration.

All five sites contribute to the activity of the enhancer, but the octamer element seems to be the most critical for B cell specificity and overall activity (Gerster et al., 1987; Lenardo et al., 1987). More than one factor can bind to the octamer. A protein designated NF-A, isolated from human B cells, was reported to bind in vitro to the octamer motifs of the mu heavy chain and kappa light chain promoters, and to the heavy chain enhancer (Singh et al., 1986). However, a factor with similar binding specificity was isolated from HeLa cells. The existence of a non-specific factor that binds to the octamer sequence might account for the use of this site for regulating transcription of other genes, such as the U1 and U2 RNA genes, and the histone H2B

gene (Sive and Roeder, 1986). Nuclear extracts from distinct cell types differ in their ability to interact with octamer-containing fragments (Landolfi et al., 1986), and a lymphoid-specific protein was described which specifically binds to the octamer of the heavy chain gene promoter (Staudt et al., 1986). This protein was named NF-A2, and the previously identified ubiquitous protein NF-A1. No differences could be detected in the binding sites for the two factors. As a functional B-cell specific promoter can be created with an octamer upstream of a TATA motif, NF-A2 may be able to interact productively with a TATA-binding factor (Wirth et al., 1987), while NF-A1 may function only with additional transcription factors, since in non-tissue-specific promoters the octamer motif is always associated with other elements (Ciliberto et al., 1985; Sive and Roeder, 1986). A cDNA clone encoding NF-A2 has been isolated, and the expression of the gene, oct-2, seems to be tissue-specific (Staudt et al., 1988). Hence, cell-type-specific expression of heavy chain genes may be controlled by the lymphoid-specific expression of the oct-2 gene. A testis-specific octamer binding factor has been identified in the sea urchin (Barberis et al., 1987).

Ubiquitous and cell-type-specific factors interacting with the kappa light chain gene octamer have also been described. OTF-1 and OTF-2 have identical binding properties, but OTF-1 activates the H2B promoter, while OTF-2 activates the kappa promoter (Fletcher et al., 1987; Scheidereit et al., 1987). As for the heavy chain gene factors, it is not

clear how different proteins can distinguish between the octamer elements of different genes.

Like the heavy chain enhancer, a function in establishment rather than maintenance of gene activation has been suggested for the kappa gene enhancer (Atchison and Perry, 1987). The enhancer of the kappa light chain gene has a less complex organisation, as only two binding sites have been identified (Sen and Baltimore, 1986). kE3 appears to bind the same factor which binds to E3 of the heavy chain enhancer, while the B site binds NF-kB. NF-kB binding activity is restricted to cells that express k chains, and so it is assumed that the B site contributes to cell-type-specificity. As both the octamer and B site are also present in the SV40 enhancer, it may be that these sites are preferentially utilised when SV40 is activated in B cells.

In summary, studies of immunoglobulin gene expression have demonstrated that distinct factors can interact with identical sequence motifs. Although it is not known how specific binding is controlled, it seems that expression is regulated by the interaction of cell-type-specific factors with a limited number of sequence motifs.

iii) Alpha-fetoprotein gene expression

The examples above concern genes with relatively simple expression patterns. In contrast, the alpha-fetoprotein (AFP) gene exhibits a complex pattern of tissue-specificity and temporal regulation during development. In the mouse it is activated in three independent cell lineages; the visceral endoderm of the yolk sac, the foetal liver and the foetal gastrointestinal tract (Tilghman, 1985). However, its mRNA is

maintained at different steady state levels in each tissue (Andrews et al., 1982; Belayew and Tilghman, 1982). The 5' flanking region of the AFP gene contains a tissue-specific promoter and three upstream regulatory elements with the properties of enhancers (Scott and Tilghman, 1983; Godbout et al., 1986). The enhancers displayed equivalent activities in a transient expression assay in a hepatoma cell line (Godbout et al., 1986). The in vivo activity of the enhancer elements were tested separately, linked to the AFP promoter, by introduction into transgenic mice (Hammer et al., 1987). These experiments demonstrated that the presence of at least one enhancer element was necessary to activate transcription, but function in maintenance of expression was unclear. It was also shown that the elements do not have equivalent biological properties, as each had a different ability to direct expression in the liver and gut. It was suggested that multiple enhancers upstream of the AFP gene act in a combinatorial manner to generate complex patterns of gene expression (Hammer et al., 1987). The diversity cannot simply be due to different concentrations of limiting trans-acting factors in each tissue, because the appropriate construct could direct high levels of expression in all three tissues. Therefore it is likely that accessibility to trans-acting factors is of importance.

iv) Other cell-type-specific enhancers

As described above, both the AFP gene and immunoglobulin genes are regulated by cell-type-specific enhancer elements in conjunction with a cell-type-specific promoter. This form of regulation is shared by other genes that have been

examined, including the rat insulin gene (Walker et al., 1987). An enhancer was identified which is active only in insulin-producing cells, but sequences closer to the start site of transcription, although lacking enhancer activity, were able to function independently to direct tissue-specific expression. Similarly, elements within 150 bp of the transcription start site are sufficient to activate cell-type-specific transcription of the rat albumin gene (Heard et al., 1987). However, an enhancer located 10kb upstream is required for the strict control of liver-specific expression in transgenic mice (Pinkert et al., 1987).

b) Negatively acting mechanisms

Studies of the control of albumin gene expression have suggested another possible regulatory mechanism; repression in non-expressing cells (Cereghini et al., 1987). In general, somatic cell genetic studies are consistent with both positive and negative regulatory mechanisms, as diffusible factors are apparently able to block or induce the expression of specific genes on the fusion of cells of different histotypic origin (reviewed by Weiss, 1982; Baron and Maniatis, 1986). At least four distinct factors present in hepatocytes bind within 150 bp of the initiation site of the albumin gene, while extracts from other tissues give a different pattern of DNA-protein interactions, shown to be dominant after mixing experiments. The action of dominant negative factors has been implicated in the regulation of other genes such as the human retinol binding protein (RBP) gene (Colantuoni et al., 1987). Three regulatory elements

were identified within 334 bp of the initiation site of the RBP gene; a non-tissue-specific enhancer, a negative element and a promoter element. Liver-specific expression required the presence of the negative element and the promoter element.

The cell-type-specificity of the embryonic skeletal myosin heavy-chain gene is controlled by the concerted action of multiple positive and negative regulatory elements (Bouvagnet et al., 1987). This emphasizes the possibility that cell-type-specific expression patterns are achieved by the additive effect of multiple transcription factors, which therefore need not be individually cell-type-specific.

The expression of steroid-regulated genes varies between genes and between different tissue types, and cannot be correlated simply with the presence or absence of steroid hormone receptors. For instance, both the chicken ovalbumin and vitellogenin genes are induced by oestrogen, but vitellogenin is expressed in the liver, while ovalbumin is expressed in oviduct tubular gland cells. Unlike other oestrogen-stimulated genes, for which oestrogen response elements have the properties of inducible enhancers, it seems that the ovalbumin gene is subjected to negative regulation (Gaub et al., 1987). The ovalbumin promoter is cell-specifically repressed in oviduct cells, and this repression is relieved by the addition of steroid hormones. A region of the promoter can confer repression and hormone-induced derepression on a heterologous promoter. In some cells, such as adult hepatocytes, lack of expression is probably a result of the absence of cell-specific factors. However, the non-

expression of ovalbumin in embryonic liver seems to be a result of DNA methylation and/or a closed chromatin configuration, as it is bypassed by microinjection of recombinant plasmids (Dierich et al., 1987). In contrast, the high expression of the Xenopus vitellogenin gene in liver cells may be due to a positive cell-specific activator distinct from the oestrogen response element (Dobbeling et al., 1988).

c) Transcription factors and tissue specificity

i) Co-ordinate control by shared transcription factors

As discussed in section II.2a(ii), different immunoglobulin genes seem to share similar enhancer binding factors. The recognition of binding sites upstream of the co-ordinately expressed genes α-1-antitrypsin (α-1-AT) and transthyretin (TTR) by a common transcription factor, revealed by competition assays, demonstrates that a single factor can activate unrelated genes (Grayson et al., 1988). Both genes are expressed mainly in hepatocytes, and upstream enhancer sequences as well as proximal promoter sequences are required (Costa et al., 1988). Three regions are necessary for activity of both enhancers; one of which interacts with a common factor, and the other two with the shared liver-specific factor (Grayson et al., 1988). A consensus sequence has been proposed as part of a liver-specific enhancer binding site. A liver-specific promoter binding factor, HP1, has been proposed on the basis of a conserved promoter element found upstream of several hepatocyte-specific genes (Kugler et al., 1988). Competition experiments in conjunction

with gel retardation assays have demonstrated that the albumin, α-1-antitrypsin and ß-fibrinogen promoters bind an identical factor, but the HP1-related sequence in the transferrin gene has different binding properties. It seems likely that HP-1 is the same as or related to other liver-specific factors identified; α-1TFB, which binds to the promoter region of the α-1-antitrypsin gene (De Simone et al., 1987), HNF-1, which binds to the α- and ß-fibrinogen genes (Courtois et al., 1987) and APF, which interacts with a single element of the albumin promoter (Cereghini et al., 1988). Thus, despite the obvious lack of structural homology among liver-specific genes, it is probable that they use common factors and elements to achieve their cell-type specificity.

A positively-acting factor specific to the anterior pituitary gland is required for transcription of both the prolactin and growth hormone genes (Nelson et al., 1988); Pit-1, also known as PUF-1 (Cao et al., 1987) or GC2 (West et al., 1987). The two genes are expressed in phenotypically distinct cell types within the anterior pituitary, but at one stage during development are co-expressed. Sequences within 140 bp of the transcription start site, spanning two Pit-1 binding sites, are sufficient for activation of the growth hormone gene, while multiple Pit-1 binding sites spread over 2kb act synergistically to promote prolactin gene expression (Nelson et al., 1988). This may be a reflection of the differences in the expression patterns, and it is likely that additional transcription factors, either positive or negative, are involved.

ii) Post-translational regulation of transcription factors

It has been shown that at least one cell-type-specific transcription factor is activated by a post-translational mechanism. As previously described, NF-kB binds to the B site of the kappa light chain gene enhancer, and was at first thought to be restricted to lymphoid cells (Sen and Baltimore, 1986). In cells that do not express light chain genes, NF-kB is not detectable in either cytoplasmic or nuclear fractions. However, by denaturation, size fractionation and renaturation, NFkB activity can be revealed in cytosolic fractions, suggesting that it is present in a wide variety of cells but inhibited in its binding activity (Baeuerle et al., 1988). If pre-B cells are stimulated with phorbol esters or lipopolysaccharide, which both lead to activation of protein kinase C, in addition to a release of inhibition, the protein translocates to the nucleus. It is considered unlikely that this is a result of covalent modification of NF-kB, as simple renaturation can generate binding activity. The most likely explanation is that the treatments release NF-kB from an inhibitor, allowing translocation to the nucleus and DNA binding, for example due to phosphorylation of the inhibitor by protein kinase C. Although the translocation of pre-existing transcription factors to the nucleus may be widely used to control inducible DNA binding proteins, it is not a general mechanism. The subcellular distribution of AP-1, which binds upstream of many TPA inducible genes, was not significantly affected by TPA stimulation of cells.

As many enhancer binding factors seem to be responsive to TPA, this is a possible indication that post-translational activation of trans-acting factors may commonly occur, possibly by a protein kinase C-dependent phosphorylation event (reviewed by Jones et al., 1988). Recently it was shown that CREB, which binds to a conserved cAMP response element, is phosphorylated in response to cAMP (Yamamoto et al., 1988). This phosphorylation, required for transcriptional activation, shifts the equilibrium towards the dimeric form of the protein, which seems to be a requirement for binding. Again, protein kinase C rather than A kinase seems to be responsible for the phosphorylation. Sequence-specific binding of the heat shock activator protein can be induced without protein synthesis, and one explanation possible is a reversible post-translational modification event (Zimarino and Wu, 1987).

3) Higher order control of cell-type-specific gene expression

a) Chromatin structure

The structure of chromatin has frequently been implicated in the control of tissue-specific gene expression. For instance, differences in chromatin structure could explain the exclusion of the oestrogen receptor and the COUP transcription factor from the apoVLDLII promoter in oviduct but not in liver (Wjinholds et al., 1988), and may be responsible for the cell-type-specific interaction of ubiquitous factors with the tyrosine aminotransferase gene

(Becker et al., 1987). As previously discussed, it seems likely that the cell-type-specificity of the ovalbumin, AFP and immunoglobulin genes is restricted by chromatin configuration. Recently, evidence for major changes in the chromatin structure of the AFP proximal promoter region during development has been reported, suggesting that these sequences are primarily responsible for repression (Godbout and Tilghman, 1988).

The expression of many cell-type-specific genes is known to be associated with tissue-specific hypersensitive sites (reviewed by Gross and Garrad, 1988). Hypersensitive sites are associated with genes that are being actively transcribed, or with the potential for activation, and they represent nucleosome-free regions in an "open" chromatin configuration. They are generally a minor portion of the genome (~1%) with about two orders of magnitude greater sensitivity to nuclease cleavage or chemical modification than other regions of bulk chromatin.

It is not known how hypersensitive sites are generated, but it is generally assumed that the nucleosomes are replaced with transcription factors. Binding of transcription factors can be detected due to protection against nucleases or modifying agents. The induction of a hypersensitive site does not require protein synthesis, and once formed, it can be propagated for at least 20 generations (Groudine and Weintraub, 1982). In *Dictyostelium*, hypersensitive sites associated with developmentally regulated gene expression have been detected (Pavlovic et al., 1986; Ayres et al, 1987).

The position of hypersensitive sites frequently coincides with that of enhancers (Parslow et al., 1982; Mills et al., 1983; Zarat and Yamamoto, 1984; Theisen et al., 1986; Emerson et al., 1987; Georgopoulos et al., 1988). The generation and/or maintenance of an open region may require the initial association of specific or general trans-acting factors, rendering the region accessible to additional but different trans-acting factors which actually activate transcription. For example, around the immunoglogulin heavy chain gene enhancer, chromatin configuration seems to be correlated with availability for expression (Ford et al., 1988). Prior to lineage committment, both T and B lineage lymphocytes have an open, DNAase I hypersensitive IgH enhancer region irrespective of their rearrangement status and transcriptional activity of their immunoglobulin genes. Following adoption of the myeloid lineage, a closed chromatin configuration results.

Higher order chromatin structure may be involved in the regulation of gene expression. Flanking the human ß-globin locus, 50 kb 5' and 20 kb 3', are two erythroid-specific DNAase I superhypersensitive sites (Tuan et al., 1985; Forrester et al., 1986). When linked to a ß-globin minilocus, the ß-globin gene is correctly regulated and expressed at high levels in transgenic mice, independent of copy number and free from position effects (Grosveld et al., 1987). The control regions are dominant, even after fusion to non-erythroid cells, and it is likely that they control the accessibility of the entire ß-globin locus to trans-acting factors. The border regions flanking the DNAase I sensitive

domain described for the ovalbumin gene may perform a similar function (Lawson et al., 1982), although such candidate sites are not associated with all genes. A 24 kb nuclease-sensitive domain spans the chicken lysozyme gene in oviduct nuclei. However, of 10 hypersensitive sites, none is localised near the borders of the nuclease-sensitive region (Gross and Garrad, 1988).

To act as a dominant control region in erythroid cells, one possibility is that the superhypersensitive sites act as nuclear matrix binding sites, interacting with matrix-bound proteins such as topoisomerases. Alternatively they might contain enhancer-like sequences that exert their effects over large distances after the interaction of enhancer-binding proteins. Such enhancers and matrix binding sites could be associated or co-localised (Cockerhill and Garrad, 1986; Gasser and Laemmli, 1986). The mechanism of generation of such a site probably does not require DNA replication (Grosveld et al., 1987), and is presumably in response to a specific physiological signal.

b) Regulatory genes and gene networks

The purification of proteins that bind to DNA sequences responsible for tissue-specific expression is currently a widely used strategy for cloning genes with a regulatory role in cellular differentiation. It seems likely that changes in the expression of numerous structural genes associated with lineage commitment are controlled by a relatively small number of regulatory genes. Genetic analysis of *Drosophila* has provided an insight into the regulatory pathways

controlling insect morphogenesis, and has suggested that many of the genes involved probably do act directly to modify transcription (reviewed by Akam, 1987; Scott and Carroll, 1987).

Several classes of genes control pattern formation in *Drosophila*, expressed in a precise temporal order and spatial pattern; first, maternally active genes, and then zygotic genes, subdivided into segmentation genes and homeotic genes (Scott and Carroll, 1987). There are interactions between classes and within classes of genes. A substantial proportion of these genes contain a 180 bp region known as the homeodomain or homeobox (McGinnis et al., 1984; Scott and Weiner, 1984). This sequence is common to at least six homeotic genes, at least five segmentation genes and at least five maternally active genes (Scott and Carroll, 1987; Gehring, 1987). The function of the homeodomain is suggested by a similarity with sequences in bacterial DNA binding proteins (Laughon and Scott, 1984) and transcriptional regulators encoded by the MAT locus in yeast (Shepherd et al., 1984). A role in regulating gene expression is consistent with the nuclear localization of several homeobox-containing gene products (White and Wilcox, 1984; Beachy et al., 1985; Carroll et al., 1985; Di Nardo et al., 1985) and with reported in vitro DNA binding properties (Desplan et al., 1985). A putative DNA binding "zinc finger" motif has been identified in some gap genes and genes of unknown function (Rosenberg et al., 1986; Schuh et al., 1986; Tautz et al., 1987).

Recently it was shown that a purified recombinant protein encoded by *even-skipped* (*eve*), a segmentation gene containing a homeodomain, binds with high affinity to specific sequences near the 5' ends of the *eve* and *engrailed* (*en*) genes (Hoey and Levine, 1988). This is consistent with its proposed function in regulating the expression of *en*, another segmentation gene, and autoregulating its own expression (Harding et al., 1986; Macdonald et al., 1986). There is no obvious homology between the binding sites, and other homeo box gene products tested also recognised both classes of binding sites. This supports a model in which regulatory interactions among homeobox proteins involve competition for similar cis regulatory sequences.

Genes containing homeodomains have also been identified in organisms as diverse as yeast, nematodes, frogs, mice and humans. Most of the *Drosophila* genes are apparently required for segment formation or segment identity, but some are also expressed later in specific cell types or sets of cells (reviewed by Doe and Scott, 1988; Barad et al., 1988). At least one, *rough*, has been identified which has no apparent involvement in early embryogenesis; instead it is implicated solely with pattern formation in the developing eye (Saint et al., 1988). Therefore it seems to be involved in recruitment or differentiation of individual photoreceptor cells, rather than in specifying regional identity.

Simpler eukaryotes such as nematodes follow a developmental programme based strictly on lineage for many cell types, although other cells may be regulated by those surrounding them (Sternberg and Horvitz, 1986). The *mec-3*

gene is essential for the proper differentiation of touch receptor neurones in *C.elegans* (Way and Chalfie, 1988). The coding sequence contains a homeodomain, and so the encoded protein presumably binds to DNA and regulates the transcription of a set of genes that specify the differentiated phenotype. The activation of *mec-3* in a particular lineage may be in response to cell-autonomous signals, or extracellular signals may be involved.

In vertebrates, there is some evidence that homeobox-containing genes could be involved in both development and differentiation. About 18 homeobox genes have been identified in the mouse; nearly all are expressed at one stage or another of embryogenesis, often within limited anatomical regions (reviewed by Holland and Hogan, 1988). Nearly every mouse or human homeobox gene is expressed in some adult tissue as well as the embryo (Schofield, 1987). Analysis of the expression of mouse and human genes during haematopoesis showed that while certain homeobox transcripts were almost ubiquitous, others were restricted to certain lineages or particular cell types (Kongsuwan et al., 1988). Expression patterns were altered in myeloid and erythroid cell lines induced to undergo terminal differentiation, and in lines switched from lymphoid to myeloid phenotypes, suggesting that homeobox genes may influence developmental decisions.

Using an alternative approach, two genes directly implicated in the control of vertebrate differentiation have been identified (Davis et al., 1987; Pinney et al., 1988). Transfection of a cloned multipotential mouse embryo cell line with human DNA has suggested the existence of two

regulatory loci, Myd and MyoD1, mediating myogenic determination and differentiation. The two genes seem to be sequentially expressed, as transfection of Myd results in the accumulation of MyoD1 transcripts. The action of these two genes produces a myogenic stem cell which can be induced to differentiate under appropriate conditions, and they are probably part of a larger regulatory network.

c) Extracellular signals

Hormones are probably the best characterised extracellular signals known to modify the expression of a set of genes in specific tissues. Hormones act by entering cells directly and binding to a specific receptor, which then translocates to the nucleus, binds to DNA and activates transcription (reviewed by Yamamoto, 1985; Picard and Yamamoto, 1987). There is evidence that binding of glucocorticoid induces the active form of its receptor due to dissociation of a complex with hsp90 (Denis et al., 1988), further evidence for modification of pre-existing transcription factors (see section II.2c(ii)). Different hormone receptors have a closely related structure, in which the hormone-binding, DNA-binding and transcription activation domains are distinct (Green and Chambon, 1987).

The mechanisms by which other extracellular signals known to influence differentiation fate modify the expression of individual cell-type-specific genes is less well understood. Retinoic acid, which in the developing chick limb bud may control pattern formation (Tickle et al., 1982; Wilde et al., 1987; Thaller and Eichele, 1987) can also induce

differentiation of several different cell types in culture (Chytil, 1984). In F9 mouse teratocarcinoma stem cells, which differentiate into primitive endoderm cells on retinoic acid addition, five genes have been identified encoding mRNAs which increase in abundance in response to retinoic acid (Wang et al., 1985). For at least two of these, regulation is at a transcriptional level. A retinoic acid receptor has been cloned which has structural similarities to the receptors for steroid hormones, thyroid hormones and vitamin D3 (Petkovich et al., 1987; Giguere et al., 1987). However, unlike steroid hormones, additional binding proteins exist which may act as shuttles between the cytoplasm and nucleus (Ong et al., 1982; Chytil, 1984).

Several genes, for which loss of function is correlated with alterations in cell identity, encode products which resemble growth factors, cell-surface receptors or components of signal transduction pathways; these include *Notch* and *Decapentaplegic* in *Drosophila* (Lehmann et al., 1983; Padgett et al., 1987), *sevenless* in *Drosophila* (Tomlinson and Ready, 1986) and *lin-12* in *C.elegans* (Greenwald et al., 1983). Factors homologous to mammalian growth factors induce expression of an early marker of the mesoderm lineage in *Xenopus* embryos (Slack et al., 1987). It was subsequently shown that basic fibroblast growth factor (basic FGF) and transforming growth factor-ß (TGF-ß) synergistically induce production of cardiac actin mRNA, encoding the earliest expressed form of muscle actin (Kimelman and Kirschner, 1987). Moreover, *Xenopus* mRNAs have been identified which

encode proteins with homology to these factors (Kimelman and Kirschner, 1987; Weeks and Melton, 1987).

4) Conclusions and implications for the control of cell-type-specific gene expression in *D.discoideum*

Tissue specific gene expression is likely to be controlled by a variety of mechanisms. There is evidence for the existence of transcription factors which activate the expression of unrelated genes in the same tissue, but additional mechanisms are required to account for more subtle differences in expression patterns. The identification and functional properties of multiple binding sites upstream of many genes, such as the AFP gene, suggest that combinational mechanisms may enable diverse patterns of gene expression to be generated by a limited number of common regulatory factors. For other genes, such as Drosophila yp1, tissue specificity is mediated by a single cis-acting element, suggesting the existence of unique transcription factors. In some cases, tissue specificity seems to be conferred by post-translational modification of trans-acting factors with a ubiquitous distribution. It is possible that other transcription factors with an apparently tissue-specific distribution are present in an inactive state in other cells. Differences in chromatin structure between tissues may also allow non-specific factors to activate transcription in a particular cell type, but specific factors may be required for the generation of an open chromatin structure.

During *Dictyostelium* development, only two major cell types are formed. Therefore, in contrast to more complex organisms, relatively simple signals may be sufficient to specify cell-type identity and to regulate the expression of cell-type-specific genes. However, variations do exist in the timing of expression and degree of cell-type enrichment of individual genes. The process of cellular differentiation does not closely resemble that of higher organisms; cells are not irreversibly committed until the latter stages of terminal differentiation (reviewed by Bonner, 1982). Thus a cell-autonomous mechanism is unlikely, especially since cellular differentiation is known to be dependent on the action of extracellular signals.

The ability to undergo fairly rapid dedifferentiation and redifferentiation and the overall structural simplicity of the organism also make it unlikely that complex hierarchies of regulatory genes exist. In support of this, Southern blotting has provided no evidence for the existence of homeobox-containing genes (McGinnis, 1985). Another difference compared to higher organisms is that development proceeds without cell division, and the genome is not detectably methylated. Therefore cell-type-specific gene expression in *Dictyostelium* must be controlled without the involvement of the cell cycle and changes in methylation state. The likely simplicity of regulatory mechanisms utilised by *Dictyostelium* reinforces the contention that it is a particularly tractable system in which to investigate the control of cell-type-specific gene expression, and

ultimately to establish the pathways by which extracellular signals act to control cellular differentiation.

CHAPTER 2 : MATERIALS AND METHODS

2.I MATERIALS

1) *D.discoideum* strains

The wild-type strain V12M2 was obtained from G.Gerisch and is described in Town et al., 1976.

The wild-type strain NC4 is described in Bonner and Slifkin, 1949.

AX-2 is an axenic derivative of NC4, and is described in Watts and Ashworth, 1970.

HM44 is a mutant deficient in DIF production and was isolated from HM27, a derivative of V12M2, as described in Kopachik et al., 1983.

HL31 is a "stalky" mutant derived from NC4 and is described in Morrissey et al., 1981.

2) Bacterial strains

DH1 is an *E.coli* strain giving high transformation efficiencies, described in Hanahan, 1983. Its genotype is F-, *endA1*, *hsdR17* (r_K^-, m_K^+), *supE44*, *thi-1*, λ^{-1}, *recA1*, *gyrA96*, *relA1*.

DH5 is a derivative of DH1, selected to give higher frequencies of transformation (see Hanahan 1985).

3) Isotopes

All isotopes and radiolabelled compounds were purchased from Amersham International, with the exception of α-^{32}P dCTP,

α-^{32}P dATP and α-^{32}P ATP which were also obtained from New England Nuclear-Dupont (NEN).

α-^{32}P dATP = deoxyadenosine 5'-[α-^{32}P] triphosphate

α-^{32}P dCTP = deoxycytidine 5'-[α-^{32}P] triphosphate

both at 3000Ci/mmole.

γ-^{32}P ATP = adenosine 5'-[γ-^{32}P] triphosphate at 5000 Ci/mmole.

α-^{32}P UTP = uridine 5'-[α-^{32}P] triphosphate at Ci/mmole.

^{125}I-Protein A = Affinity purified protein A ^{125}I-labelled with Bolton and Hunter reagent at >30mCi/mg.

^{14}C-Chloramphenicol = D-threo-[dichloroacetyl-1-^{14}C]chloramphenicol at 25μCi/ml.

4) Enzymes

Restriction endonucleases were purchased from Boehringer Mannheim, New England Biolabs (NEB) or Northumbria Biologicals Ltd. (NBL). The Klenow fragment of DNA polymerase I was purchased from Amersham International. *E.coli* polymerase I, calf intestinal phosphatase, Bal 31 nuclease and S1 nuclease were obtained from Boeringer Mannheim. T4 DNA ligase and T4 DNA polymerase were supplied by Bethesda Research Laboratories (BRL), polynucleotide kinase by Pharmacia, super reverse transcriptase by Anglian Biotechnology Ltd., and exonuclease III by NEB.

5) Chemicals and miscellaneous materials

Unless otherwise specified, chemicals were purchased from BDH Chemicals Ltd. or Fisons plc. Analytical grade reagents

were used wherever available. Caesium chloride was supplied by Koch-Light Ltd., and urea (ultra pure) by Schwartz/Mann Biotech. Ethidium bromide, actinomycin D, tetracycline and cAMP were purchased from Sigma. G418 (geneticin sulphate) was obtained from Gibco BRL and ampicillin from Penbritin, while nogalamycin was a gift of the Upjohn Co.. Nucleotides were purchased from PL Biochemicals.

Sephadex G50, sephacryl S-300 and Percoll were obtained from Pharmacia, and agarose (ultra pure) from BRL. Nitrocellulose filters were purchased from both Sartorius and Schleicher and Schuell, while Schleicher and Schuell alone were the suppliers of NA45 paper. Filters used for slime mould experiments were purchased from Millipore. 3mm and 17mm chomatography paper was obtained from Whatman.

6) Solutions

All solutions were made with deionised distilled water produced by the MilliQ system for reagent grade water, and were sterilized either by autoclaving at 10 lbs of pressure for 25 minutes or by filtering through a Nalgene filter. Solutions to be used with RNA were DEP treated by the addition of diethylpyrocarbonate (DEP) to 0.1% just prior to autoclaving. Certain compounds were added after autoclaving, e.g. BSA and thio-containing compounds (DTT and ß-mercaptoethanol). 10% SDS was not autoclaved.

7) Antibodies

Rhodamine-conjugated and non-conjugated rabbit anti-mouse polyclonal antisera (DAKOPATTS, 10g/l) were immunoabsorbed

against methanol-fixed V12M2 slugs and AX-2 vegetative cells before use. 9E10 (Evan et al., 1985) and MUD-1 (Gregg et al., 1982) are monoclonal antibodies derived from mouse hybridomas. The protein concentration of the ascites fluid used as a stock solution was about 0.6mg/ml and 0.2mg/ml respectively.

8) Abbreviations

Common abbreviations used throughout are as follows:

A	adenosine
C	cytidine
G	guanosine
T	thymidine
ATP	adenosine 5' triphosphate
UTP	uridine 5' triphosphate
dATP	deoxyadenosine 5' triphosphate
dGTP	deoxyguanosine 5' triphosphate
dTTP	deoxythymidine 5' triphosphate
dCTP	deoxycytidine 5' triphosphate
cAMP	cyclic adenosine 3'-5' monophosphate
BSA	bovine serum albumin
SDS	sodium dodecyl sulphate
DTT	dithiothrietol
DEP	diethylpyrocarbonate
EDTA	diaminotetra-acetic acid (disodium salt)
HEPES	N-2-hydroxyethylpperazine-N'-2-ethane sulphonic acid
MOPS	3-(N-morpholino) propane sulphonic acid
PIPES	piperazine-NN'-bis-2-ethane sulphonic acid

MES	2-(N-morpholino) ethane sulphonic acid
TEMED	NNN'N'-tetramethylethylenediamine
Tris	hydroxymethyl methyl-amine
OAc	acetate ion
PO_4	phosphate ion
Hz	hydrazine
DMS	dimethylsulphate
DNA	deoxyribonucleic acid
RNA	ribonucleic acid
DNAase	deoxyribonuclease
RNAase	ribonuclease
kb	kilobase
bp	base pairs
1K	1000 revolutions per minute
cpm	counts per minute
Tm	melting temperature of a nucleic acid hybrid
C	degrees centigrade
M	molar
V	volts
Ig	immunoglobulin

2.II METHODS

1) Bacterial techniques

a) Growth and long-term storage

E.coli were grown in L-broth or on L-agar (Table 2.1) at 37°C. When required, the appropriate antibiotic was added to the medium to select for the growth of transformed cells; 100µg/ml for both ampicillin and tetracycline. Frozen stocks of cells harbouring plasmids were prepared by growth to saturation in antibiotic-containing L broth followed by dilution 1:1 with sterile glycerol. The stocks were stored in 1.5ml Nunc tubes either at -20°C (short-term) or at -70°C (long-term). Bacterial cell stocks required for the preparation of competent cells were either stored as glycerols at -70°C or, in the short-term, on Psi agar plates at room temperature (Hanahan, 1985).

b) Preparation of transformation competent DH5 cells (adapted from Hanahan, 1985)

The protocol used includes modifications introduced by V. Simanis. Bacterial stocks were allowed to thaw only sufficiently to permit cells to be scraped from the surface. After streaking onto a Psi agar plate the cells were left to grow at 37°C overnight. A single large colony was used to inoculate a 0.5 ml culture in Psi broth (Table 2.1), shaken at 37°C. When the culture was growing exponentially it was diluted into 5ml of prewarmed broth. After a further two hours' growth it was added to 100mls of prewarmed broth. When the optical density of the culture had reached approximately 0.22 at a wavelength of 550nm on the Cecil spectrophotometer,

Table 2.1

Media used for the growth and transformation of *E.coli*

L broth

10g	Difco bacto tryptone
5g	Difco bacto yeast extract
5g	NaCl

made up to 1 litre with water and adjusted to pH 7.2 with NaOH

L agar

L broth + 15g/litre of Difco agar

Psi broth

5g	Difco bacto yeast extract
20g	Difco bacto tryptone
5g	$MgSO_4$
0.76g	KCl

made up to 1 litre with water and adjusted to pH 7.6 with KOH

Psi agar

Psi broth + 15g/litre of Difco agar

All solutions were autoclaved at 10 lbs of pressure for 25 minutes.

the cells were transferred to chilled tubes and allowed to cool on ice. After pelleting by spinning at 3K for 5 minutes in an HB4 rotor, the supernatant was removed with a pipette, and the cells were resuspended in 40mls of TfbI (30mM KOAc, 10mM $CaCl_2$, 50mM $MnCl_2$, 15% v/v glycerol, pH adjusted to 5.8 with 0.2M acetic acid) using a precooled pipette. The cells were allowed to stand for 5 minutes on ice, and then pelleted by spinning at 2.5K in a Sorvall HB4 rotor at 4°C. They were gently resuspended in 4mls of TfbII (10mM PIPES, 75mM $CaCl_2$, 10mM $RbCl_2$, 15% v/v glycerol, pH adjusted to 7.0 with KOH). The cells were left for 15 minutes on ice before dispensing as 210µl aliquots into prechilled Eppendorf tubes, and snap freezing in liquid nitrogen. The cells were stored at -70°C. Usually a transformation efficiency of $1\text{-}5x10^7$ colonies/µg of DNA was achieved.

c) Transformation of *E.coli* strain DH5
(Adapted from Hanahan, 1985)

Transformation competent cells were allowed to thaw slowly on ice. 10-100pg of supercoiled DNA or a maximum of 25ng of a ligation mix, ethanol precipitated if necessary, was added to 25µl of cells in a maximum volume of about 5µl. After 30-45 minutes on ice, the cells were heat shocked by incubation at 42°C for 90 seconds, and then chilled on ice for 1-2 minutes. At the end of this time, 75µl of Psi broth, prewarmed to 37°C, was added, and the cells were incubated at 37°C for 60 minutes before plating on to L agar containing antibiotic. Colonies were obtained after overnight incubation at 37°C.

d) Colony lifts for screening

(Grunstein and Hogness, 1975)

Transformed colonies were transferred to nitrocellulose by carefully laying a filter on to the surface of the agar plate and leaving for 1-2 minutes, ensuring no air bubbles remained. After addition of orientation marks using a syringe needle, the filter was removed and placed, colony side up, onto pads of 3mm paper soaked in a series of buffers for 5 minutes each. Firstly, the colonies were lysed using 1.5M NaCl, 0.5M NaOH, and then treated with 1.5M NaCl, 0.5M Tris-HCl pH 8.0 followed by 2x SSPE (0.36M NaCl, 20mM NaH_2PO_4 pH 7.4, 2mM EDTA) for 5 minutes each. The filters were air-dried and baked for 2 hours at 80°C. Bacterial debris was removed by treatment of the filters with 50mM NaOH at room temperature for 30 minutes followed by three 10 minute water washes. The filters were then probed as described in section 2.6b(i). The colonies on the plates were allowed to regrow by further incubation at 37°C for 7-8 hours.

2) Slime mould techniques

The methods used for the growth and development of *D.discoideum* cells are based on those described in Sussman, 1966. All cell culture was performed at 22°C.

a) Growth and long-term storage of *D.discoideum* cells

Strains NC4, V12M2, HM44 and HL31 were grown in association with *Klebsiella aerogenes* on SM agar (see Table 2.2). The stocks were maintained by growth across preformed

Table 2.2 **Media used for the growth and development of *D.discoideum*.**

Axenic medium

14.3g	Oxoid bacterial peptone
7.2g	Oxoid yeast extract powder
15.4g	D-glucose
0.51g	Na_2PO_4
0.485g	KH_2PO_4

made up to 1 litre with water and adjusted to pH 6.9 with NaOH.

SM agar

10g	D-glucose
10g	Difco bacto peptone
1g	$MgSO_4$
2.2g	KH_2PO_4
1g	Na_2HPO_4
1g	Difco yeast extract

made up to 1 litre with water and 18g Difco agar added.

KK2

2.55g	KH_2PO_4
0.65g	K_2HPO_4

made up to 1 litre with water. Sterile $MgSO_4$ added to 10^{-3}M after autoclaving, but before use.

KK2 agar

KK_2 + 20g/litre Difco agar

MES-HL5

5g	Oxoid yeast extract
10g	Oxoid yeast peptone
10g	glucose
1.3g	MES

made up to 1 litre with water and pH adjusted to 7.1 using 3M KOH.

All solutions were autoclaved at 10 lbs of pressure for 25 minutes.

bacterial lawns, and cells picked from the growing edge were used to seed new plates. At regular intervals (every 3 months) fresh cultures were recovered from silica gel stocks (see below), except for HM44 which was hatched out freshly just prior to an experiment. To generate large numbers of cells, 0.5-2.0x 10^5 *Dictyostelium* amoebae per plate were mixed with bacteria and spread on SM agar, giving "clearing" lawns after about 48 hours.

AX-2 cells (Watts and Ashworth, 1970) can grow without a bacterial food supply in nutrient broth, i.e. axenically. They were grown in shaken suspension at 180 rpm on a rotary shaker in axenic medium (Table 2.2) containing 15µg/ml tetracycline. Cells were inoculated to give an initial density of about 2-5x 10^5/ml, and the doubling time under these conditions was 9-10 hours. Stationary phase was reached at 1-2x 10^7/ml. Viable axenic cultures could be stored for several months as standing cultures at a density of not more than 5x 10^5/ml. Recovery was achieved by at least ten-fold dilution into fresh medium.

Sporulating strains (i.e. all strains used except HM44) were stored for longer periods as spores on silica gel at 4°C. Spores from 1-2 plates of culminants (1-4x 10^8 cells) were resuspended in 0.5mls of horse serum before addition to silica gel crystals in 2ml universal bottles, accompanied by vigorous shaking on ice. Alternatively, exponentially growing cells were resuspended at about 2x10^7/ml in 95% horse serum, 5% DMSO and stored in 1ml Nunc tubes in liquid nitrogen after gradual cooling (2 hours at -20°C, 2 hours at -70°C). Amoebae were recovered by placing either a silica crystal or the

rapidly thawed liquid nitrogen stock on to an SM plate pre-spread with *Klebsiella aerogenes*.

b) Development of *D.discoideum* cells

i) Filter development

Cells to be developed were harvested during exponential growth, either from plates when the bacterial lawn was visibly clearing, or from suspension cultures at between 2-5x 10^6/ml. The cells were pelleted by centrifugation in 50ml Falcon tubes in an MSE bench centrifuge at 2K for 2 minutes. The cells were washed twice with KK_2 (see Table 2.2), and resuspended in the same buffer at 5x 10^7/ml. 0.5ml aliquots were spread on to washed 2cm white Millipore filters placed on prefilters soaked in 1.5mls lower pad solution (0.022M KCl, 0.0025M $MgCl_2$, 500µg/ml streptomycin sulphate, 0.04M KH_2PO_4, adjusted to pH6.5 with KOH). The cells were allowed to settle before removing excess liquid. The plates were kept in a moist environment for the duration of the experiment.

ii) Production of slugs and fruiting bodies

Exponentially growing cells were harvested and washed with KK_2 as before. The pellet was allowed to drain, forming a slurry of approximately 10^9 cells/ml. A thin streak of approximately 100µl placed across the centre of a 2% water agar plate was allowed to dry slightly before enclosure of plates, interleaved with black discs, inside a plastic bag additionally containing a moist tissue to provide humidity. The plates were then placed inside a specially designed tin, painted black on the inside and with a narrow slit, positioned at right angles to the cell streak, to provide a

unidirectional light source. The slugs were harvested after 24-48 hours (a maximum of 36 hours for AX-2 cells).

Fruiting bodies were obtained by plating 1-2x 10^8 washed cells on to 2% KK_2 agar plates, and allowing development to proceed for at least 24 hours.

c) Percoll gradient fractionation of prespore and prestalk cells (adapted from Ratner and Borth, 1983)

To separate prespore and prestalk cells derived from axenic slugs, preformed 70% Percoll gradients in 20mM EDTA, 5mM MES pH7.0 were used. To 70mls of Percoll, 20.5mls of 10/3 buffer pH7.0 and 9.5 mls of 10/3 buffer pH4.0 were added (i.e. 66.7mM EDTA, 16.7mM MES). 12ml aliquots of this solution were dispensed into 15ml Corex tubes, which were then spun at 15K for 40 minutes in the SS34 rotor. To harvest slugs, the cell streak was cut away prior to flooding the plate with Na^+/K^+ buffer (20mM Na_2KPO_4 pH7.0) and scraping the agar surface with a glass rod. The cells were washed once with room temperature 20/20 buffer (Na^+/K^+ buffer with 20mM EDTA) and resuspended in a reduced volume of the same solution (about 6mls for 20x 9cm plates). The cell clumps were dissociated by successively vortexing, passaging through syringe needles and addition of BALpronase (=12.5mM 2,3-dimercaptopropanol and 1mg/ml pronase type V, both from Sigma) rapidly followed by dilution with cold 20/20. The cell suspension was filtered through Nitex to remove any remaining aggregates, and the eluate was counted, concentrated by centrifugation and resuspended at 2-4x 10^8/ml in 20/20. After layering 250µl on each Percoll gradient, they were spun at

12K for 3 minutes in the HB4 rotor. Prespore cells are denser and band at a lower position. The heavy and light fractions were collected using a Pasteur pipette and were washed free of Percoll by 5-fold dilution in 20/20 and centrifugation. The fractions were respun on separate gradients and harvested as before.

d) HM44 stalk cell induction in submerged monolayer culture (Kopachik et al., 1985)

HM44 cells grown on a lawn of *Klebsiella aerogenes* were harvested as the plates were clearing, and washed at least twice with KK_2 to remove bacteria before resuspending in DIF induction buffer at a density of 2.5x 10^6/ml (=10mM MES, 2mM NaCl, 10mM KCl, 1mM $CaCl_2$, 5mM cAMP, 200µg/ml streptomycin sulphate and 15µg/ml tetracycline, adjusted to pH 6.2 by the addition of KOH). 15ml aliquots of cells (1.125x 10^8) were plated in 14cm diameter Sterilin tissue culture dishes and incubated for 9-11 hours before the addition of DIF. The time at which the medium was changed was designated t_0. The original medium was carefully aspirated off and replaced with 15mls of fresh medium, either with or without 3000 U/ml^{-1} of synthetic DIF-1 per plate. Timepoints were harvested after careful removal of the medium by addition of a few mls of cold KK_2 and scraping the surface of the plate. The cells were pipetted into a 50ml Falcon tube and pelleted by centrifugation. A 2cm plate containing 2mls of cells was also set up to score the efficiency of the induction by counting the percentage of stalk cell formation 24 hours after the addition of DIF. One 14cm plate of cells was used for RNA

extraction, while 3-10 plates of cells per timepoint were harvested for extraction of nuclei.

e) Preparation of transcriptionally active nuclei (adapted from Jacobson, 1976)

Cells to be used for nuclei preparation were pelleted by centrifugation and washed once with cold 0.2% NaCl. All subsequent steps were performed on ice, and centrifugation was at 4°C. The cells were resuspended at 1x 10^8/ml in nuclear lysis buffer (0.05M Hepes pH7.5, 0.005M MgOAc, 10% sucrose, 2% cemusol). After spinning at 6K for 10 minutes in the SS34 rotor, the supernatant was discarded, or retained for cytoplasmic RNA purification. The nuclei were resuspended and spun twice more in lysis buffer, and then taken up in nuclear storage buffer (0.04M Tris pH7.9, 0.01M $MgCl_2$, 0.1mM EDTA, 1mM DTT, 50% glycerol) at 2x 10^8/ml. 200µl aliquots were snap frozen in ethanol/dry ice and stored at -70°C.

f) Transformation of *D.discoideum* cells (adapted from Nellen et al., 1984; Nellen and Firtel, 1985)

AX-2 cells were grown in axenic medium in shaken suspension until a density of 1-3x 10^6/ml was reached. They were then plated in 10ml aliquots on to 9cm Falcon Optilux petri dishes. The cells were allowed to settle out from the suspension and attach firmly to the dish; they were left for at least 3 hours before removing the remaining cells and axenic medium and replacing with 10mls of MES-HL5 pH7.0-7.1

(Table 2.2) containing 100µg/ml ampicillin. The cells were left to stand overnight or for at least 4 hours.

The DNA was added to the cells as a calcium phosphate coprecipitate, with duplicate plates for each sample. Two plates were treated in an identical manner except for the omission of DNA to monitor the level of spontaneous resistance to the drug. Very few colonies ever formed on these control plates, and those that did disappeared quickly as the selection was continued. The DNA precipitate was prepared by adding 25µg of supercoiled plasmid per plate to 0.6ml 1xHBS (2x HBS stock = 274mM NaCl, 10mM KCl, 1.4mM Na_2HPO_4, 12mM dextrose, 42mM HEPES, pH adjusted to 7.05 using NaOH, filter sterilized and stored at -20°C), followed by addition of calcium chloride to 125mM while vortexing (38µl of a 2M stock, filter sterilized and stored at -20°C). The precipitate was allowed to stand for 25 minutes to develop fully. The medium was removed from the cells and the DNA added dropwise at the centre. No medium was added back for 25 minutes while the plates were rocked gently. 10mls of MES-HL5 was then added and left for a further 4-5 hours. The cells were then subjected to a glycerol shock in which the medium was replaced with 2mls of 15% sterile glycerol in 1xHBS for 10 minutes. After thorough removal of the glycerol, 10mls of axenic medium was added for a minimum of 4 hours or at a maximum, overnight, to allow expression of the resistance gene on the transforming plasmid to begin. Fresh axenic medium was then substituted additionally containing 20µg/ml geneticin sulphate (G418) from a stock at 4mg/ml stored in small aliquots at -70°C. After three days' selection the

medium was replaced with axenic medium with a G418 concentration of 10μg/ml.

A recent modification in our laboratory is that the cells are no longer mixed with autoclaved *Klebsiella aerogenes* and plated on to black Millipore filters at this stage. Instead, after a further three days' incubation in tissue culture dishes, resistant colonies become visible. The number of colonies counted at this stage give a reliable indication of the transformation efficiency. The colonies are allowed to increase in size and become a confluent monolayer, helped by frequent medium changes to remove dead cells, resuspension by pipetting to seed new colonies and, if necessary, initial transfer to a 2cm dish. When a monolayer of cells cover a 9cm dish, they are resuspended, diluted to 20-30mls and transferred to a shaken suspension culture in a 50ml flask. After exponential growth has been successfully established the cells can be transferred to larger scale cultures. The cells on plates are subcultured every 3-4 days by resuspension and transfer of 0.5ml to 10mls of fresh medium. After the initial selection step, all transformants are grown in axenic medium containing 10μg/ml G418.

g) Preparation of samples for immunohistochemistry (Gregg, 1982)

Slugs were disaggregated by vortexing after resuspension in KK_2 at a concentration of about 10^7/ml followed by passaging through needles of successively decreasing bore-size, attempting to avoid the introduction of air bubbles. This resulted in a mixture of single cells and clumps

containing about 2-10 cells which were fixed by the addition of ice-cold methanol to a final concentration of 60% (v/v). The cell suspension was dispensed as 5µl drops on multiwell slides which had been coated with 100µg/ml poly L-lysine, rinsed with water and dried. After the cell sample had dried, an additional 10µl of methanol was added and allowed to evoporate. The slides were used directly for antibody staining (see section 2.7c(i)) or stored at 4°C.

Slugs were fixed prior to embedding in Tissue Tek O.C.T. compound (Miles Scientific, BDH) by flooding plates with 4% paraformaldehyde (v/v) dissolved in phosphate buffered saline (PBSA =0.14M NaCl, 0.003M KCl, 0.015M Na_2HPO_4, 0.015M KH_2PO_4, pH7.2). The fixed slugs were transferred using an inoculating loop to a multiwell plate and washed with several changes of PBSA, one wash containing 1% methylene blue. The slug was then moved to a drop of embedding medium on a coverslip which was inverted on to a prefrozen drop on a square of Whatman 17mm paper. The frozen block was stored at -70°C before sectioning. Sections 5 or 10µM thick were cut using a Bright 5030 cryostat and picked up on to slides pretreated with agarose/chrome alum (Krefft et al., 1984). Slides were dipped at 60°C in a Millipore filtered solution containing 0.5g of gelatin in 250mls of water and 0.5ml of 10% w/v $KG(SO_4)_2.12H_2O$. The embedding medium was washed from the slides with PBSA prior to staining, and the slides were stored at -20°C.

To prepare slime trails, slugs were allowed to continue migrating across water-agar plates for 36-48 hours. Blocks of agar bearing slime trails were cut from the plates and

inverted onto poly-L-lysine coated slides. After removal of the blocks, the trails remained attached. Hence the underside of trails were stained.

3) DNA preparation and analysis

a) Preparation of plasmid DNA

i) Small scale preparation of plasmid DNA by the boiling method

(Holmes and Quigley, 1981)

A stationary culture of *E.coli*, transformed with plasmid DNA, was prepared by inoculating 2mls of L broth containing the appropriate antibiotic and growing at 37°C overnight in a rotary shaker. 1.5mls of cells were harvested by spinning in a microfuge for 1 minute, and the supernatant was aspirated off using a drawn out Pasteur pipette. The pellet was resuspended in 200µl of STET (5% Triton X-100, 50mM Tris-HCl pH8.0, 50mM EDTA pH8.0, 8% sucrose) followed by the addition of 20µl of a freshly prepared lysozyme solution (40mg/ml). After vortexing, the tube was placed in a boiling water bath for 45 seconds and then spun for 10 minutes in a microfuge. The pellet of cell debris was removed with a toothpick and 200µl of isopropanol added. After spinning on dry ice, the nucleic acid was pelleted by spinning for 10 minutes. The miniprep pellet was rinsed with 80% ethanol and resuspended in 50µl of TE buffer (10mM Tris-HCl pH8.0, 1mM EDTA).

ii) Large scale preparation of DNA

(adapted from Ish-Horowicz and Burke, 1983)

A 500ml culture of bacteria, containing the required plasmid, was spun at 7K for 5 minutes in 2x 250ml buckets in a GSA rotor. Each pellet was resuspended in 10mls of 50mM glucose, 25mM Tris-HCl pH8.0, 10mM EDTA, and the combined cell solution was lysed by the addition of 40mls of 0.2M NaOH, 1% SDS. The lysate was chilled on ice for 5 minutes before the addition of 20mls of cold 5M KOAc pH4.8. After a further 15 minutes on ice, the cell debris was spun out at 8K for 5 minutes. The nucleic acid present in the supernatant was precipitated by adding 0.6 volumes of isopropanol and spinning at 8K for 10 minutes. The pellet was rinsed with ethanol and drained well before resuspension in 4.5mls of 50mM Tris-HCl pH 8.0, 10mM EDTA. Exactly 4.5mls of the DNA solution was added to 4.85g of caesium chloride, followed by 0.5mls of ethidium bromide (10mg/ml). The combined solution was spun in a 15ml disposable plastic tube in the bench IEC CENTRA-4X centifuge for 10 minutes at maximum speed (about 4K). The supernatant was transferred to fill a Beckman 5ml polyallomer Quick-seal centrifuge tube, which was heat sealed and spun for 16-18 hours in a Beckman VTi65 rotor at 45K and 20°C. The supercoiled plasmid DNA, banding lower in the gradient than any nicked plasmid or residual chromosomal DNA, was harvested with a syringe and re-banded after transferring to a new tube. The tube was filled with rebanding solution (12.5mls of 1g/ml CsCl solution combined with 1ml of EtBr stock solution) and spun as before. The ethidium bromide was removed from the harvested band by passage over a Dowex AG50 column after at least two-fold dilution with TE. The column, in a Pasteur pipette plugged with polymer wool, was first

washed repeatedly with TE. The eluate was collected in a 15ml Corex tube and finally the Dowex was rinsed with one column volume of TE. The DNA was recovered by ethanol precipitation and resuspended in TE.

b) Restriction enzyme analysis

Double-stranded DNA, prepared by either of the two methods above, was subjected to restriction digestion in one of three buffers, depending on the salt requirements of the enzyme used (see Table 2.3). All buffers were stored as a 10x stock prepared freshly every 4-5 weeks. In addition bovine serum albumin, also stored as a 10x stock, was added to every digest to a final concentration of 100µg/ml. In theory only 1 unit of restriction enzyme per µg of DNA was required, but in practice a 2-3 fold excess was added provided that the volume of enzyme, stored in glycerol, did not exceed 10%. The amount of DNA added was dependent on whether the digest was for preparative or analytical purposes, but in general 2µl of miniprep DNA was added. All digests were carried out for at least one hour at 37°C, except for those with TaqI (65°C) or AccI (55°C). After completion of incubation, miniprep digests were subjected to RNAase treatment for 10 minutes at room temperature by addition of 1µl of a 100µg/ml stock.

c) Gel electrophoresis of DNA

(Maniatis et al., 1982)

i) Agarose gels

Table 2.3 **Commonly used enzyme buffers**

10x RE buffer

0.1M Tris HCl pH 7.5
0.1M $MgCl_2$
0.1M ß-mercaptoethanol

10x RE+S

0.1M Tris HCl pH 7.5
0.1M $MgCl_2$
0.1M ß mercaptoethanol
0.5M NaCl

10x RE+SS

0.1M Tris HCl pH 7.5
0.1M $MgCl_2$
0.1M ß-mercaptoethanol
1M NaCl

10x NT buffer

0.5M Tris-HCl pH 7.8
0.05M $MgCl_2$
0.1M ß-mercaptoethanol

Stored as incomplete buffers lacking ß-mercaptoethanol. For complete buffer, 14µl ß-mercaptoethanol added to 2mls.

10x CAP buffer

0.1M Tris-HCl pH 9.5
0.01M spermidine
0.001M EDTA

Stored as incomplete buffer lacking spermidine. For complete buffer, 40µl 0.5M spermidine added to 2mls.

10x CAP kinase buffer

0.5M Tris-HCl pH 7.5
0.1 $MgCl_2$
0.05M DTT

Stored as incomplete lacking DTT. For complete buffer, 100µl 1M DTT added to 1.9mls.

5x ligase buffer

0.25M Tris-HCl pH 7.5
0.05M $MgCl_2$
0.05M DTT
0.0025M ATP

Stored as incomplete buffer lacking DTT and ATP. For complete buffer, 0.1ml of 1M DTT and 10µl of 0.5M ATP added to 1.9mls.

Two sizes of flat-bed horizontal agarose gels were used for the resolution of DNA fragments. Rapid analysis was carried out on small 50ml gels, while larger 150ml gels were used for more detailed mapping and fragment preparation. The gel percentage was varied between 0.6% and 1.5% depending on the sizes of the fragments to be resolved. The appropriate amount of agarose was dissolved by heating in either 50 or 150mls of 0.5x TBE (2.5x stock =0.225M Tris-HCl, 0.225M boric acid, 0.005M EDTA), and the gels were run in this buffer. After the agarose solution had cooled but was still molten, ethidium bromide was added to 1µg/ml and the gels were poured and allowed to solidify. Gels were either poured in preformed trays or on to flat glass plates sealed with tape. Loading buffer was added to the DNA samples (5x loading buffer =15% Ficoll, 0.025% bromophenol blue, 2.5x TBE); restriction digests were loaded directly. 50ml gels were run at 100V for 20 minutes and large gels were run either at 100V for 4 hours or 30V overnight. The DNA was visualized by viewing it under long wavelength (302nm) UV light, either on a transilluminator or, if the DNA was to be used for cloning, provided by a hand-held source. The size of DNA fragments was determined by comparison with marker DNAs; either a combination of pAT153 digested with HinfI and lambda DNA digested with HindIII, or the 1kb DNA ladder from BRL.

ii) Polyacrylamide gels

Native gels were used for both preparative and analytical purposes involving double-stranded DNA, e.g. mapping or purification of fragments for cloning. Denaturing gels, containing 8M urea, were used for preparation and analysis of

single-stranded DNA or double-stranded DNA which had been denatured, e.g. sequencing gels, or purification of oligonucleotides, single-stranded probes or fragments for Maxam and Gilbert sequencing. The percentage of gel used was varied between 5 and 20% (for gel recipes, see Table 2.4). All gels except gradient gels (see below) contained 0.5x TBE and were run in this buffer, and were set by the addition of 1.5% v/v of ammonium persulphate and 0.1% v/v TEMED just prior to pouring between two glass plates. The plates were separated at a fixed distance by spacers made of Teflon or plasticard of varying thicknesses, and were clamped with bulldog clips and sealed at the lower corners with Cowgum. The gel comb used was of the same thickness as the spacers. The gel thickness required was dependent on the amount of DNA to be loaded and the resolution required. In general, 1-3mm spacers were used for preparative gels and native analytical gels, and 0.5mm spacers for sequencing gels. Gradient gels contained a gradient of TBE which altered the resolution properties of the gel and hence increased the amount of readable sequence.

Native gels were run at no more than 400V and the loading buffer used was the same as for agarose gels. For denaturing gels, DNA was ethanol precipitated and resuspended in formamide dyes (80% deionised formamide, 0.5x TBE, 0.1% xylene cyanol). Immediately prior to loading the DNA was denatured by placing at 95°C for 1 minute. Denaturing gels were run at 1000-1500V. The DNA could be visualized in three ways; by autoradiography if radioactively labelled, by staining with ethidium bromide (1µg/ml) or, with sufficient

Table 2.4 Acrylamide gel recipes

For denaturing gels

40% acrylamide stock

190g acrylamide (BioRad)
10g N,N'-bismethylene acrylamide
Made up to 500mls with distilled water

For native gels

40% acrylamide stock

200g acrylamide (BDH)
Made up to 500mls with distilled water.

2% bis stock

4g N,N'-bismethylene acrylamide
Made up to 200mls with distilled water.

All filtered and stored at 4°C.

100mls denaturing gel solution (5-20%)

40g urea
12.5-50mls 40% acrylamide
20mls 2.5x TBE
1.5mls 5% APS
Distilled water to 100mls
100µl TEMED

40mls 5% native gel solution

5mls 40% acrylamide
2.5mls 2% bis
8mls 2.5x TBE
1ml 5% APS
23.5mls distilled water
40µl TEMED

Gradient gel (narrow)

Top gel

32.55g urea
10mls 40% acrylamide
3.75mls 10x TBE
36.75mls water
375µl 10% APS

75µl TEMED

Bottom gel

1.25g sucrose
5.75g urea
1.875mls 40% acrylamide
3.125mls 10xTBE
2.125mls water
62.5µl 10% APS
0.25mls 0.4% BPB
12.5µl TEMED

amounts, by UV shadowing on a thin layer chromatography plate precoated with silica gel (Merck).

d) DNA recovery from gels

i) Electroelution

Slices of agarose or acrylamide gels containing the DNA to be purified were excised with a razor blade, and placed inside dialysis tubing sealed with clips and containing approximately 300µl of 0.5x TBE. The tubing was placed inside a flat-bed gel tank also containing 0.5x TBE, with the slice closest to the negative electrode. 150V were applied for a minimum of 1 hour, followed by 45 seconds during which the polarity was reversed. The buffer was transferred to a tube, and fresh buffer was added to rinse the bag and slice. The DNA was recovered by ethanol precipitation. Electroelution worked poorly for fragments of about 5kb or over.

ii) Onto NA45 paper

This technique was used for the recovery of fragments of up to about 7kb from agarose gels suitable for cloning. After staining and visualisation of the fragment with a hand-held UV light, a razor blade was used to cut slits both in front of and behind the fragment. Strips of NA45 paper wetted with gel buffer were inserted into the slits, and the gel was returned to the electrophoresis tank for re-application of the voltage until the DNA had been seen to run on to the front strip using the hand-held UV light. The DNA was eluted from the paper by immersing in 300µl of 50mM arginine containing 1M NaCl and incubating at 70°C for 1 hour. The DNA was then recovered by ethanol precipitation.

e) Southern transfer

(Southern, 1975)

Large agarose gels (150mls) were immersed in 500mls of 0.25M HCl for 30 minutes at room temperature with gentle rocking to depurinate high molecular weight DNA, followed by 500mls of denaturation solution (0.2M NaOH, 0.6M NaCl). After a further 30 minutes, the gel was neutralised in 1.5M NaCl, 0.5M Tris-HCl pH7.5, again for 30 minutes. The gel was placed on to a wick of Whatman 3mm paper soaked in 10x SSC (20x SSC =3M NaCl, 0.3M trisodium citrate) in contact with a resevoir of 10x SSC. The edges of the gel were sealed with Parafilm in order to ensure all buffer passed through the filter. The nitrocellulose filter was prewetted with water and then 10x SSC before carefully placing on top of the gel and removing all air bubbles. Two sheets of 3mm paper the same size as the gel were placed between the filter and a 5cm layer of tissues covered with a weighted glass plate. The transfer was allowed to proceed overnight, after which the blot was dismantled and the filter rinsed in 5x SSC prior to air drying and baking at 80°C for 2 hours.

4) DNA manipulation

(adapted from Maniatis, 1982)

a) End-filling and labelling of DNA

i) End-labelling of a 5' overhang using the Klenow fragment of DNA polymerase I

This method was used to label restriction fragments to be sequenced by the chemical cleavage method. A mixture of three cold triphosphates to a final concentration of 100μM were added to the digested DNA (5-10μg in 1x restriction buffer), followed by 30μCi of α-32P dATP or dCTP, depending on the sequence of the overhang left by the restriction enzyme. For every 10μl of added volume, 1μl of 10 x nick-translation buffer was included (see Table 2.3) in addition to 1 unit of Klenow per μg of DNA. The reaction was allowed to proceed at room temperature for 15-20 minutes. If the labelled triphosphate was not added at the terminal position, the reaction was chased for about 5 minutes with an excess of cold triphosphate (100μM). The reaction was terminated by phenol extraction, and the labelled DNA was separated from nucleotides and phenol by spinning through a G50 column equilibrated in TE (see section 2.8b).

An excess of cold triphosphates in the absence of labelled triphosphate was substituted when a 5' overhang was required to be blunted for cloning purposes.

ii) Phosphatase

This enzyme removes 5' phosphate groups from DNA termimi. Digested DNA was subjected to the action of phosphatase for two purposes; to prevent recircularisation of cleaved vectors during ligation to target molecules, and to enable the substitution of a labelled phosphate catalysed by T4 polynucleotide kinase. The reaction was carried out with DNA, generally contained in restriction enzyme buffer, diluted to 1μg/15μl in 1x CAP buffer (see Table 2.3). 1μl of calf intestinal alkaline phosphatase (40 units) was added per μg

of DNA, and incubated for 30-60 minutes at 37°C (adding fresh enzyme at 30 minutes). The reaction was stopped by extracting with an equal volume of phenol and the DNA was recovered by passing through a spun G50 column (see section 2.8b) followed by ethanol precipitation.

iii) Kinase labelling

5-10 µg of DNA fragment which had previously been dephosphorylated with phosphatase was placed in a final concentration of 1x kinase buffer (10x stock, see Table 2.3) after the addition of 50µCi of γ-^{32}P ATP. To this mixture 10 units of T4 polynucleotide kinase was added and the reaction was continued for 1 hour. After phenol extraction to prevent further enzyme activity, the unincorporated label was removed using a spun G50 column equilibrated in TE.

b) Subcloning of DNA fragments into vectors

i) Ligation

The vector was prepared by appropriate restriction digestion followed by end-filling if necessary and then phosphatase treatment of the cleaved ends as described in sections 2.3b and 2.4a(ii). The fragments to be cloned were purified as in sections 2.3d(i) and (ii). If required they were also blunted prior to gel purification. A combination of blunt and cohesive ends could be generated by filling one site prior to cleavage of another. Vector and insert were generally combined to give a 2-3 fold molar excess of insert; typically the vector was at a concentration of 20-50 ng/µl in the final ligation mix, also containing 1x ligase buffer (5x stock, see Table 2.3). After the addition of 1µl of T4 DNA

ligase per 10µl reaction volume (2.5 units), the ligation was allowed to continue for at least 12 hours at 14°C. A portion of the ligation mix was used directly to transform competent bacteria (see section 2.1c).

ii) Use of linkers

If the ends of the DNA molecules produced by digestion were not compatible, or if a new restriction site was required at a junction of vector and target or upon recircularization, then the ends were joined via linkers. The linkers (0.5µg) were phosphorylated using 2 units of T4 polynucleotide kinase in 1x kinase buffer (Table 2.3) at 37°C for 1 hour after the addition of ATP to 1mM. The degree of completion of subsequent reactions could be followed by labelling a small proportion of the linkers. This was achieved by incubating the reaction mix, prior to the addition of cold ATP, for 15 minutes in the presence of 10µCi of γ-^{32}P ATP. The kinased linkers were ligated to the blunt-ended DNA overnight at 18°C (0.5µg linkers/0.5µg DNA) in 1x ligase buffer. The reaction was terminated by phenol extraction and unligated linkers, free nucleotides and molecules consisting of up to four self-ligated linkers were removed by passage over a Sephacryl-300 (S-300) spun column. The column was made by plugging a 1ml syringe with polymer wool on to which a base of Sephadex G25 was layered. The column was then filled with S-300, placed in a 15ml Falcon tube and spun at 2K for 2 minutes. The column was washed twice with TE before spinning the phenol-extracted mix. Each DNA terminus now had several linkers attached; all but part of one were removed by digesting with the appropriate

restriction enzyme, e.g. EcoR1 for EcoR1 linkers. To ensure completion of the reaction, about 50 units of enzyme were used and the incubation was continued for several hours. The free linkers were removed with a second S-300 column. The success of the ligation and subsequent digestion was monitored by the removal of small aliquots at various stages. On a high percentage native acylamide gel the self-ligated linkers could be seen as a ladder which reverted to a single low molecular weight band after digestion. The DNA with linkers added was then used for further cloning steps.

iii) Insertion of a double-stranded oligonucleotide into a vector

Insertion of a single copy of an oligonucleotide was ensured by ligation of a molar excess of non-phosphorylated oligonucleotide to cleaved phosphorylated vector. Two complementary oligonucleotides (100ng of each strand) were mixed in 20µl of 1x ligase buffer and annealed by incubating at 60°C for 1 hour. The vector was cleaved with the appropriate enzyme, generating compatible ends, and 20-30ng was ligated to a 50-100 molar excess of oligonucleotide in the same buffer. After transformation, insertion of the oligonucleotide could be detected by labelling one of the strands to screen colony lifts (see section 2.1d).

c) Sequencing by the chemical cleavage method (Maxam and Gilbert, 1980)

10-20µg of DNA to be sequenced was cleaved with a suitable restiction enzyme, i.e. generating ends suitable for kinase-labelling or end-filling with Klenow (see sections

2.4a(i) and (iii)), and sufficiently close to the region of interest. Blunt or 3' recessed ends could be labelled using T4 polymerase. The DNA was then cleaved with a second enzyme to generate fragments labelled at one end only which could be separated on a denaturing acrylamide gel. After visualisation by autoradiography, the appropriate bands were excised and the DNA recovered by electroelution and two rounds of ethanol precipitation. The fragments were resuspended in 30µl of distilled water and split between four tubes, labelled G, G+A, T+C and C, in the ratio 1:2:2:1. The reactions were carried out as decribed in Table 2.5, using the solutions listed. On piperidine treatment, the four sets of reactions led to partial cleavage at G residues alone, at As and Gs, at Ts and Cs or at Cs alone, on average about once every 100 bases. The products were resolved by denaturing polyacrylamide gel electrophoresis as described in section 2.3c(ii).

d) Generation of a directed deletion series using exonuclease III

(adapted from Henikoff, 1984)

In order for the digestion to proceed in one direction only, 5µg of plasmid DNA was cleaved at two restriction sites, leaving a 3' recessed end adjacent to the sequences required to be protected. A 3' overhang is not a substrate for exonuclease III (alternatively digestion may be blocked by filling a 5' overhang with α-thio triphosphates). The digestion producing the 3' recessed end was carried out first to enable completion of the reaction to be checked. The

Table 2.5 **Sequencing by the chemical cleavage method**

G	G + A	T + C	C
5µl DNA	10µl DNA	10µl DNA	5µl DNA
200µl DMS buffer	-	10µl water	15µl 5M NaCl
Chill to 0°C	Chill to 0°C	Chill to 0°C	Chill to 0°C
1µl DMS	25µl 88% formic acid	30µl hydrazine	30µl hydrazine
0°C 30 seconds	20°C 10 minutes	20°C 10 minutes	20°C 10 minutes

To all:-

1. Chill on dry ice, then ethanol precipitate by spinning in a microfuge for 5 minutes.

2. Rinse with 95% ethanol and resuspend in 0.3M NaOAc

3.Ethanol precipitate again, rinse with 95% ethanol and dry.

4. Resuspend in 100µl of 1M piperidine (from a 10.0M stock, stored at 4°C) and incubate at 90°C for 30 minutes.

5. Lyophilise until dry (usually overnight).

6. Resuspend in 20µl of water and lyophilise.

7. Resuspend in 10µl of water and lyophilise.

8. Resuspend in 10µl of water and lyophilise.

9. Resuspend in formamide dyes and denature by heating at 90°C for 3 minutes before loading on to a gel.

Solutions used

DMS stop

1.5M NaOAc
1M 2-mercaptoethanol
100µg/ml yeast tRNA

Hydrazine stop

0.3M NaOAc pH 7.0
0.1mM EDTA pH 8.0
25µg/ml yeast tRNA

DMS buffer

50mM sodium cacodylate pH 8.0
10mM $MgCl_2$
1mM EDTA pH 8.0

completed digest was spun through a G50 column equilibrated with 10mM Tris-HCl pH8.0, 0.1mM EDTA, in a volume of 20µl. Approximately 2µg of DNA (8µl) was added to a reaction mix in a final volume of 30µl of 1x exonucleaseIII buffer (50mM Tris-HCl pH8.0, 1mM MgCl2, 0.5mM DTT) and prewarmed to 30°C. The conditions used were designed to remove approximately 200 bases per minute. The reaction was started by adding 2µl of exonuclease III (60 units) and 1µl aliquots were removed at various timepoints into 5µl of stop buffer on ice, containing Zn^{+} to inhibit the enzyme (20mM $ZnSO_4$, 50mM Tris-HCl pH8.0, 200mM NaCl, 10mM $MgCl_2$, 0.1mM $CoCl_2$). The exonuclease III was then heat inactivated by incubating at 70°C for 10 minutes. After the samples had cooled, S1 nuclease was employed to trim the single-stranded ends by addition of 50µl of a mix containing 1x S1 buffer (33mM NaAc pH4.5, 50mM NaCl, 0.03mM $ZnSO_4$) and 0.5 units of enzyme per tube. The samples were incubated at 18°C for 30 minutes and then the DNA was recovered by ethanol precipitation. The pellet was resuspended in 10µl TE, of which 3µl were set aside for a 0.6% agarose gel to monitor the activity of the exonuclease III. The ends were repaired by adding 5µl of a mix containing 0.25 units of Klenow polymerase together with 40mM dNTPs and 100µg/ml BSA. After 10 minutes at room temperature, 45µl of 5mM EDTA was added and the Klenow destroyed by incubating at 70°C for 10 minutes. The timepoints were incubated overnight at 14°C after the addition of 15µl of a mix containing 1x ligase buffer and 2.5 units of T4 ligase per sample. 10µl of the total of 25µl for each timepoint was used directly to transform 25-30µl of competent DH5 cells.

e) Creation of a deletion series with Bal31 nuclease (Legersky et al., 1978)

Bal 31 nuclease has both a single-stranded endonuclease and a double-stranded exonuclease activity, and the latter property was utilised to create a 3' flanking deletion series (see section 6.4). First, the most suitable conditions (incubation temperature and time, and level of enzyme) to produce deletions over the region of interest were established. To a final reaction volume of 20µl in 1x Bal31 buffer (5x = 0.1M Tris pH7.2, 3M NaCl, 62.5mM $MgCl_2$, 62.5mM $CaCl_2$, 5mM EDTA) 2.5µg of linearised DNA and 3µl of a 1/10 dilution of enzyme (0.6 units) were added. The mix was incubated at 30°C, and every minute for 5 minutes an aliquot containing 0.5µg of DNA was removed and added to 4µl of 40mM EDTA on ice (final concentration 20mM). These conditions led to the removal of approximately 100 base-pairs from each end per minute. The timepoints were phenol-extracted and spun over a G50 column equilibrated with TE. They were then ethanol precipitated and resuspended in diluted 5x ligase buffer (see Table 2.3), giving a final concentration of 1x in a 20µl mix additionally containing 4µl of phosphorylated linkers (30ng/µl final concentration). 1µl (2.5 units) of T4 ligase was added to each and the tubes incubated at 18°C overnight. After a phenol extraction and G50 column to remove ligase activity, the timepoints were digested to remove excess linkers and passed over S-300 columns (see section 2.4b(ii)). The DNA in the eluates was precipitated with ethanol and religated before transformation into DH5 cells.

Before subcloning, the deletions were subjected to restriction enzyme analysis and the endpoints precisely determined by sequencing from the linker site by the chemical cleavage method.

5) RNA analysis

a) Extraction of nucleic acid from *D.discoideum*

i) Phenol extraction

(adapted from Jacobson, 1976)

The cells were harvested by spinning at 2K for 3 minutes in an IEC CENTRA-4X bench centrifuge in 50ml disposible Falcon tubes. At this stage the pellet could be snap frozen on dry ice and stored indefinitely at -70°C. The cells were resuspended in 1ml of HMK buffer (50mM HEPES pH7.5, 40mM $Mg(OAc)_2$, 20mM KCl, DEP treated), to which 0.2M NaCl, 0.5% SDS (v/v) and 0.5% DEP (v/v) had been added. After vortexing, the lysate was successively extracted with equal volumes of saturated phenol and chloroform/isoamyl alcohol (50:1). The extractions were repeated until the interface obtained with phenol was clear (usually twice). The nucleic acid contained in the aqueous phase was transferred to a baked Corex tube and precipitated by addition of DEP treated NaAc pH5.5 to 0.3M and three volumes of ethanol. After chilling on dry ice for 30 minutes, the nucleic acid was pelleted by centrifugation at 10K for 30 minutes at 4°C in an HB4 rotor. The pellet was dissolved in water and reprecipitated in an eppendorf tube. After rinsing with 80% ethanol and drying, the nucleic acid was resuspended in DEP treated water.

ii) Extraction of RNA with GTME

(Maniatis et al., 1982)

Total RNA was extracted with GTME lysis buffer (4M guanidium isothiocyanate, 0.1M 2-mercaptoethanol pH5.0, filter sterilized and stored at -20°C). 5x 10^7 cells were resuspended in 2mls of GTME, and snap frozen and thawed to ensure complete lysis if not stored at 20°C. The lysate was layered on to a 1.1ml CsCl cushion (5.7M CsCl, 0.1M EDTA pH5.0, filter sterilized) in an SW50.1 tube. The tube was filled with paraffin and spun for 16 hours at 33K and 15°C. A Pasteur pipette was used to remove the paraffin and then the caesium solution containing DNA and protein, banding just below the interface. The tube walls were rinsed with 3mls of sterile DEP treated water and inverted. A razor blade was used to cut away the base of the tube containing the RNA pellet, which was resuspended in 100µl of water and transferred to an eppendorf tube. The RNA was recovered by ethanol precipitation after extracting once with phenol and with chloroform, and again resuspended in DEP treated water.

b) Formaldehyde gels and Northern transfer (adapted from Maniatis et al., 1982)

A 1.5% gel was prepared by melting 2.25g of agarose in 108mls of water. After cooling to 60°C, 15mls of 10x gel buffer (0.2M MOPS pH7.0, 0.05M NaOAc, 0.01M EDTA pH8.0) and 27mls of formaldehyde (final concentration 2.2M) were added. After the gel had solidified it was presoaked in 1x running buffer (1x MOPS) for 1 hour before loading. 5µg of RNA, lyophilised if necessary, was resuspended in 4.5µl of DEP treated water. To each sample, 2µl of sterile 10x MOPS

buffer, 3.5µl of formaldehyde and 10.0µl of deionised formamide were added, and the RNA was denatured by incubation at 55°C for 15 minutes. 2µl of 10x gel loading buffer was added to each sample (50% glycerol, 1mM EDTA, 0.4% xylene cyanol) immediately prior to loading. The gel was run at 100V for 4 hours or 30V overnight. The RNA was transferred to nitrocellulose with the same procedures used for Southern blotting (see section 2.3e) except that 20x SSC was used in place of 10x. After transferring overnight, the filter was not rinsed prior to baking for 2 hours at 80°C. The filter was probed as described in section 2.6b.

c) Primer extension

(Williams and Mason, 1985)

Primer extension was performed using a kinase-labelled synthetic oligonucleotide (see section 2.6a(iii)). 100ng of primer was labelled and the product of G50 spun column chromatography was split between 6-10 annealing reactions, each containing 20µg of total RNA. Thus a molar excess of primer was always employed (10fmoles =80pg of a 25mer; 20µg of total RNA =0.25-0.5 fmoles of a message with an abundance of 0.1-0.2%). The 10µl annealing mix also contained 1µl of 10x primer extension buffer (100mM PIPES pH6.4, 4M NaCl), and was sealed in a 25µl sterile glass capillary tube. The sealed tubes were incubated in a water bath overnight. A prediction of the Tm, dependent on the length and base composition of the oligonucleotide, could only be made approximately and so a range of annealing temperatures were selected, typically 10-20°C below the estimated value. The contents of the capillary tube were expelled into 90µl of extension mix

containing 50mM Tris-HCl pH8.2, 10mM DTT, 6mM $MgCl_2$, 250µg/ml actinomycin D, 0.5mM of each of the four dNTPs and 1 unit of super reverse transcriptase. After incubation at 42°C for 1 hour the nucleic acid was recovered by ethanol precipitation and the extension products were resolved by denaturing polyacrylamide gel electrophoresis. The gel was fixed in 10% methanol, 10% acetic acid and dried, and the bands were visualised by autoradiography.

d) In vitro nuclear transcription
(adapted from Jacobson, 1976)

Nuclei were incubated under conditions in which preinitiated transcripts were elongated, so incorporating ^{32}P-labelled UTP. To a 200µl aliquot of nuclei thawed slowly on ice (for details of nuclei preparation, see section 2.2e), 200µl of a 2x transcription buffer was added (80mM Tris-HCl pH7.9, 400mM NaCl, 20mM $MgCl_2$, 2mM DTT, 10% glycerol, 0.5M each of ATP, CTP and GTP, 5µM UTP). The reaction was begun by the addition of α-^{32}P UTP to 200nM (Amersham SP6 grade 3000Ci/mmole, 2.5µl per reaction) and incubated at room temperature for 30 minutes. To stop the reaction, 65µl of stop buffer (8.3% SDS, 1.67mM EDTA pH7.4, 0.167M Tris pH7.4) was added, followed by 20µl of 4M NaCl and 40µl of 1mg/ml tRNA in TE (to enable recovery of elongated transcripts by ethanol precipitation) and 100µl of TE (to prevent phase inversion during phenol extraction). The mix was extracted with an equal volume of phenol followed by chloroform twice, and each time the upper, aqueous phase was transferred to a clean tube. After the last chloroform extraction, the aqueous

phase was split between two eppendorf tubes and four volumes of ammonium acetate were added. The contents of the tubes were vortexed, chilled on dry ice and spun for 10 minutes. The supernatant, containing the unincorporated nucleotides and most of the counts, was discarded. The pellets were resuspended in 100µl DEP treated TE buffer and the precipitation step was repeated until few counts remained in the supernatant (usually twice). At this stage the pellet was washed in 95% ethanol and allowed to dry before resuspending in TE. After counting the labelled RNA could be used directly as a probe.

6) Hybridisation

a) Preparation of probes

i) Nick-translation of DNA

(Rigby et al., 1977)

Double-stranded DNA molecules were labelled in both strands by the translation of single-stranded nicks created by DNAase using *E.coli* DNA polymerase I. A 25µl reaction mix was assembled containing 250ng of DNA, 1µl of a nucleotide mix lacking the dNTP employed for labelling (3.3mM each), 1µl 1mg/ml BSA, 2.5µl 10x nick-translation buffer (see Table 2.3) and 30µCi of either α-^{32}P dCTP or dATP. An appropriate amount of pancreatic DNAase, stored as aliquots at -70°C, was added to give an average probe size of 500bp (~1ng). Two minutes later, 1 unit of DNA polymerase I (Kornberg) was added, and the reaction was allowed to proceed for 2 hours. The nick-translation was halted by the addition of 25µl of SDS buffer and the probe was separated from the unincorporated

nucleotides by passing over a spun G50 column, equilibrated in SDS buffer. The probe could then be stored at -20°C, but immediately before use it was denatured by the addition of NaOH to 0.3M (3M stock), followed by 0.5M Tris-HCl pH7.5 and 0.3M HCl. The specific activity obtained was generally 0.5-1.0x10^8 cpm/µg.

ii) Labelling using a mixed oligonucleotide primer (adapted from Feinberg and Vogelstein, 1983)

This method was also used to label double-stranded plasmids or fragments to high specific activity. A mixed primer of random hexanucleotides (PL Biochemicals) was suspended in TE buffer to give 90 units/ml (1 unit = optical density of 1 at a wavelength of 260nm). 100ng of DNA, made up to a volume of 5µl in water, was denatured by boiling for 3 minutes and rapidly cooled in an ice/water bath. Other components of the reaction mix were then added; 2.5µl of 1M HEPES pH6.6, 2.5µl of 10x nick-translation buffer (see Table 2.3), 2.5µl of a nucleotide mix containing 3.3mM of each of the three nucleotides not used for labelling, 1µl of a 1 in 7 dilution of the oligonucleotide mix, 1µl of 10mg/ml BSA and 7µl of water. To initiate the reaction 25µCi of α-^{32}P dCTP or dATP and 2.5µl of Klenow polymerase were added. After 2.5 hours at room temperature the reaction was stopped and the probe prepared for use as before (see section 2.6a(i)).

iii) Kinasing of oligonucleotides

To prepare an oligonucleotide probe, 100ng of synthetic oligonucleotide was labelled with γ-^{32}P ATP using T4 polynucleotide kinase (see section 2.4a(iii)) without the necessity of prior dephosphorylation or gel purification.

b) Hybridisation conditions

(adapted from Maniatis et al., 1982)

i) Probes prepared by nick translation or mixed primer labelling

The same hybridisation solution was used for both Northern and Southern blots and for colony filters, containing 50% formamide, 5x SSC, 0.1% SDS, 100µg/ml salmon sperm DNA and 10x Denhardts solution. The salmon sperm DNA was denatured by sonication and heated twice to 100°C for 5 minutes before storage at -20°C. Denhardts solution was stored at 20°C as a 100x stock (2% w/v each of BSA, Ficoll and polyvinylpyrrolidone. The filters were sealed in a Jencons bag containing a minimum of 1ml of hybridisation solution per 10cm^2 of nitrocellulose. The filters were prehybridised at 37°C for 4 hours before addition of the denatured probe to the same buffer. Care was taken not to let undiluted probe touch the filter, and before resealing the bag as many of the air bubbles as possible were removed. The probe, at a concentration of about 2-4x10^6 cpm/ml, was mixed thoroughly and incubated at 37°C overnight.

The filters were washed at low stringency in buffer containing 2x SSC, 0.1% SDS at 65°C. Initially the filters were rinsed with buffer at room temperature before four washes of 30 minutes each (4x 15 minutes for Northerns) with prewarmed buffer in a 65°C bath or oven. As during the hybridisation, the filters were gently shaken. Washing at higher stringency was achieved by reducing the salt concentration e.g. from 2x SSC to 1x, and then 0.5x and 0.1x

if required. The filters were sealed between sheets of thin plastic for autoradiography to prevent drying out, and hence to enable subsequent removal of the probe. If Southern or colony blots were to be reprobed, the original probe was removed by immersion in 50mM NaOH for 30 minutes at room temperature, followed by three 10 minute water washes. Removal of probes from Northern blots by placing in boiling water for 2 minutes followed by gradual cooling often resulted in loss of RNA from the filter, and so reprobing was sometimes delayed until the first probe had decayed. Probes could be re-used after storage in hybridisation buffer for up to 1 week. Re-used probes were denatured by incubating at 75°C for 15 minutes.

ii) Oligonucleotide probes

The hybridisation temperature and washing conditions employed were dependent on the length and base composition of the oligonucleotide. The same buffer was used for hybridisation and prehybridisation, containing 6x SSC,0.1% SDS, 10x Denhardts solution and 100µg/ml salmon sperm DNA. For short oligonucleotides (<25-30 nucleotides) the melting temperature of the hybrid formed in the buffer was calculated by the equation $Tm = [4n(G+C) + 2n(A+T)]°C$. The hybridisation was carried out at about 20°C below the calculated Tm, and the washes at 10°C below, initally with 2x 5 minute washes with 6x SSC, 0.1% SDS. If necessary to reduce background, the filter was washed further with conditions of reduced salt (e.g. 5x SSC), increased SDS concentration and increased temperature. Using the c-myc oligonucleotide as a probe (see

Chapters 5 and 6), the hybridisation was performed at 32°C and the washes at 42°C.

iii) Labelled RNA generated by in vitro nuclear transcription

^{32}P-labelled RNA produced in this way was used to probe Southern blots of digested plasmid DNAs. The Southern blots were run and transferred in the usual way, except that at least 2µg of DNA was loaded per track (compared to ~0.5µg for normal analytical purposes). Duplicated digests were run to provide a separate filter for each timepoint, and even loading, running and transfer were checked by visualising the gel with UV light after staining with ethidium bromide. The hybridisation buffer (50% deionised formamide, 5x SSC, 200µg/ml BSA, 300µg/ml Ficoll, 0.2% SDS, 100µg/ml wheatgerm tRNA, 500µg/ml polyrA) was treated by addition of DEP to 0.2% and incubation at 37°C for longer than 2 hours prior to the prehybridisation step. The filters (~5cm x ~8cm) were sealed with 5mls buffer each and prehybridised for 4 hours. Hybridisation was continued for 48 hours. After this time, the filters were washed in 1x SSC (for D19) or 0.5x SSC (for pDd63 and 56), 50% formamide, 0.1% SDS at 37°C with three changes of buffer for 15 minutes each. The stringency employed for D19 was necessary to prevent binding of a related RNA species to the filters, and that for pDd63 and 56 was required to overcome cross-hybridisation of these two related messages (see Chapter 4).

c) Autoradiography and density scanning

All autoradiography except that involving ^{125}I (to reduce background) was carried out using pre-flashed film (Laskey and Mills, 1975) For ^{32}P and ^{125}I a tungstate intensifying screen was used and the cassette stored at -70°C. The film used was either Fuji RX X-ray film or Kodak X-Omat AR film and was developed using a Fuji X-ray film processor (RGII). The resulting autoradiograms were scanned, if required, using an LKB ultroscan XL laser densitometer (2222-010).

7) Protein techniques

a) Western blotting

i) Extraction of total cellular protein from *D.discoideum*

For each timepoint, protein was extracted from approximately 5x 10^6 cells, recovered by centrifugation. The cell pellet was washed twice with KK_2 and transferred to an eppendorf tube, where it was resuspended in 50µl of 1x sample buffer (see Table 2.6) to which bromophenol blue had been added. The buffer was filtered and stored in frozen aliquots at -20°C. The samples were briefly sonicated if required to resuspend the protein. A protein estimation was carried out on 1µl of the sample diluted into 2mls of Biorad protein assay reagent diluted 1 in 5 (stock =55% phosphoric acid, 15% methanol). The optical density was read at 595nm, and compared to a series of standards. 5x 10^6 cells yielded approximately 100µg of protein, and 30µg was loaded per lane of a one-dimensional SDS polyacrylamide gel.

ii) Running and transfer of one-dimensional SDS polyacrylamide gels

(Laemmli, 1970)

Gels were run using the BioRad mini-gel apparatus and consisted of a 10% separating gel overlaid by a 4.5% stacking gel. The gel mould was constructed according to the BioRad manual using ethanol-cleaned plates. The components of the stacking gel and the separating gel are listed in Table 2.6, The gels were degassed before pouring and induced to polymerise by the addition of ammonium persulphate and TEMED. The separating gel was overlaid with 0.5cm of water-saturated isopropanol and allowed to polymerise for 30-60 minutes. The isopropanol:water was then removed, blotting the last traces with 3mm paper, before inserting the gel comb and pouring the stacking gel. The protein preparations, already containing sample buffer, were vortexed, heated at 80°C for 10 minutes and finally spun to pellet cellular debris before removal of the sample to be loaded with a Hamilton syringe. The gels were run at 200V until the dye front had reached the bottom of the gel (about 25-30 minutes). After dismantling the gel apparatus the stacking gel was discarded.

For Western blotting, a nitrocellulose filter cut to size was soaked in distilled water for 15-30 minutes before assembling a cassette containing the gel and filter flanked by 2 sheets of 3mm chromatography paper and 2 fibre pads. The cassette was placed in a small BioRad transfer tank containing 1 litre of transfer buffer (4 parts SDS-PAGE running buffer, 1 part methanol) with the gel closest to the negative electrode. The transfer was carried out at 30V and

Table 2.6 Solutions for SDS-PAGE and Western blotting

Acrylamide stock

30% acrylamide
0.8% methylene-bis-acrylamide
Filtered and stored at 4°C

Lower gel buffer

1.5M Tris-HCl pH 6.8
0.4% SDS

Stacking gel buffer

0.5M Tris-HCl pH 6.8
0.4% SDS

5xSDS-PAGE running buffer

0.125M Trizma
0.5% SDS
1M glycine
pH = 8.3

Sample buffer

1/8 dilution of stacking gel buffer
1.95% SDS
10% glycerol
10mM DTT

10% separating gel

6.7mls acrylamide solution
8.3mls water
5.3mls lower gel buffer
100µl 10% APS
9.3µl TEMED

4.5% stacking gel

0.75mls acrylamide solution
3.0mls water
1.25mls stacking gel buffer

Transfer buffer

80% SDS-PAGE running buffer
20% methanol

Blocking buffer

13.6% low fat milk powder (Marvel)
0.01M NaCl
2.5mM Trizma base
0.2% sodium azide
Adjusted to pH 7.4 with NaOH
Stored at 4°C

Low salt wash buffer

150mM NaCl
50mM Tris
Adjusted to pH 7.4 with HCl

High salt wash buffer

400mM NaCl
50mM Tris
Adjusted to pH 7.4 with HCl

room temperature overnight, or for 1 hour at 100V with ice cooling. After removal from the cassette, the filter could be used directly for antibody binding.

iii) *D.discoideum* colony blotting (Witke et al, 1987)

This procedure was carried out to try to identify insertional mutants following transformation of a cloned gene (see section 3.8). Either individual transformed colonies were picked on to SM agar plates spread with a lawn of Klebsiella aerogenes, or the cells were pooled and plated at low density to give colonies derived from single cells. As D19 is expressed only during development, the colonies were allowed to grow to a sufficient size to produce a central core of starved, developing cells. The plates were overlaid with a dry filter (Schleicher and Scheull BA85) and left for 2 minutes before transferring the filter, colony side up, on to a copper block frozen on a bed of dry ice. After 1 minute the filters were placed at -70°C for 30 minutes. They were then thawed before washing several times in a low salt buffer (see Table 2.6) until the cellular debris had been removed. The plates were stored at 4°C until required.

iv) Antibody binding to Western and colony blots

The procedure employed for Western blots and colony blots was identical. The excess capacity of the nitrocellulose was first blocked by incubation in blocking buffer for 2 hours at room temperature or overnight at 4°C. The blot was then incubated with the test antibody at the required dilution in

blocking buffer (1/100 for MUD-1) for at least 2 hours at room temperature. To reduce the volume required at this stage the filters were placed in a sealed bag (with at least 1ml/10cm^2). The antibody solutions could be re-used several times. The blots were washed twice in low salt buffer, then twice in high salt buffer, and finally again twice in low salt buffer, all for 15 minutes (see Table 2.6). Antibody binding was detected using ^{125}I-protein A. Unless the test antibody was known to be of the correct Ig subclass to bind protein A directly, the blots were first incubated for 1 hour at room temperature with DAKO rabbit anti-mouse Ig at a dilution of 1/200 in blocking buffer. The Ig had been extensively immunoabsorbed to prevent non-specific binding. The blots were washed as before and placed in a box with a sufficient volume of blocking buffer, containing affinity-purified ^{125}I-protein A (10^6 cpm/ml; generally 20mls), to cover them. The washing procedure was repeated and the filters rinsed with water before drying at 37°C for 1 hour and subjecting to autoradiography. A piece of foil was inserted between the blot and the film to prevent sticking.

b) Assaying transformed cells for chloramphenicol acetyl transferase activity

(Gorman et al., 1982)

Cells to be assayed (between 1-5x 10^7) were harvested by centrifugation and washed twice with KK_2. They were resuspended in 1ml of buffer containing 40mM Tris pH7.4, 1mM EDTA and 150mM NaCl per 10^7 cells and, after transferring to an eppendorf tube, were pelleted by spinning in a microfuge.

The cells were resuspended in 100µl ice-cold 0.25M Tris pH7.8 and the tube was vortexed and placed in dry ice for 5 minutes. The sample was transferred directly to 65°C and heated for 10 minutes. Completion of lysis was checked by microscopic examination, and the lysate was then centrifuged at 4°C to pellet cellular debris. The supernatant was transferred to a new tube and stored at -20°C. The protein concentration was determined by reading the absorbance of 10µl in 2mls of Biorad assay reagent (diluted 1 in 5) at 595nm.

The lysate was assayed for CAT activity by diluting a known amount of protein (usually 2µg) into 60µl of water and adding to a tube containing 1µCi of ^{14}C-chloramphenicol (5µl, Amersham). To this mix 70µl of 0.25M Tris-HCl pH7.8 was added and, after a 5 minute preincubation at 37°C, 20µl of 4mM acetyl CoA. The assay reaction was incubated at 37°C for a further 30 minutes and stopped by the addition of 0.5mls of cold ethyl acetate. Each tube was vortexed and spun for 30 seconds. The upper, organic phase was transferred to a new tube and allowed to evaporate in a fume hood. The acetylated products were separated from chloramphenicol by thin layer chromatography on a silica gel plate in a tank containing 95:5 chloroform/methanol (v/v). The samples were resuspended in 15µl of ethyl acetate and spotted on to the plate as 2µl aliquots, allowing to dry in between. After the solvent front had reached within ~1cm of the top the plate it was air dried and subjected to chromatography.

c) Immunostaining

i) Immunostaining and light microscopy of fixed samples

When attempting to detect the c-myc epitope, all antibody incubations were carried out for 30 minutes at room temperature. 10µl of the c-myc monoclonal, 9E10, (1 in 100 dilution in PBSA) was carefully applied per field to the slides in a humidified chamber. The slides, placed in a slide rack, were then washed by immersion in 3 changes of PBSA buffer for 5 minutes each, and allowed to air dry before addition of the second antibody (rhodamine-conjugated rabbit anti-mouse, 1 in 100 dilution = 100µg/ml. After the incubation time the slides were washed as before and allowed to dry. The slides were mounted in Gelvatol (Monsanto) and examined under fluorescent light (600nm).

ii) Immunoelectron microscopy
(Roth, 1983; Coggi et al., 1986)

This procedure was undertaken by R.Tilly and K.Blight at the Electron Microscopy Unit, L.I.F.. Fruiting bodies or slugs were fixed for 2 hours at room temperature in 1% monomeric glutaraldehyde in Sorensen's phosphate buffer, pH7.2. They were then washed in phosphate buffer, soaked for 4 hours in 0.5M ammonium chloride also in phosphate buffer, and finally washed overnight in fresh buffer. Dehydration was begun at room temperature for 30 minutes in 70% ethanol followed by 2 hours in 70% ethanol at -20°C, 4 hours in 100% ethanol at -30°C and then at least three changes of 100% ethanol. Specimens were embedded in Lowicryl resin (Chemishe Werke Lowi) and sections were cut on a Reichert OMU4 and

picked up on coated nickel grids. After incubating sections from both transformants and AX-2 controls with the appropriate antisera, immunolabelling was detected using Auroprobe anti-mouse immunoglobulins conjugated to 15nm colloidal gold (Janssen Life Sciences).

8) Miscellaneous

a) Ethanol precipitation of nucleic acid

Nucleic acid was precipitated by the addition of one-tenth volume of 3M sodium acetate (pH5.5 and DEP treated for RNA, pH7.0 for DNA) and 2.5 volumes of ethanol. The samples were then chilled on dry ice for 10-30 minutes, or overnight at -20°C, before spinning for 10-30 minutes. Large volumes were spun in Corex tubes in an HB4 rotor at 10K and 4°C, while smaller volumes were spun in a microfuge. Pellets were rinsed in chilled 80% ethanol and air dried before resuspension. If two consecutive precipitations were performed the first pellet was taken up directly in 0.3M sodium acetate. If the nucleic acid was present at a low concentration, a carrier of yeast tRNA (1mg/ml stock in TE, stored at -20°C) was added to 50µg/ml.

b) Sephadex G50 columns

Sephadex G50 was equilibrated by immersion in at least 2.5 volumes of water at room temperature overnight. The supernatant was decanted off and replaced with the same volume of water containing 0.02% sodium azide. The swollen G50 was stored at room temperature.

To make a column for spun column chromatography a 1ml syringe with its plunger removed was used. A plug of polymer wool was inserted and the syringe filled with G50 and placed in a 15ml Falcon tube. The column was spun at 3000K for 5 minutes in the IEC CENTRA-4X bench centrifuge. The eluate was discarded and the column washed by topping up with either SDS buffer (10mM Tris-HCl pH7.4, 100mM NaCl, 0.5% SDS, 1mM EDTA) for nick-translations or TE (10mM Tris-HCl pH8.0, 1mM EDTA) for all other purposes. The column was spun as before and the washes repeated twice more. Before use, the recovery of the column was checked with 50µl of buffer.

c) Equilibration of Dowex

100g of Dowex AG 50W-X8 (Biorad) was washed in 400mls of 1M NaOH for 5 minutes, which was then decanted and replaced with 400mls of 1M HCl, again for 5 minutes. After 2 washes with distilled water the Dowex was washed with 1M NaCl. It was stored at 4°C in 1M NaCl, 10mM EDTA, 100mM Tris-HCl, pH7.5 (adjusted after addition with 4M NaOH). Sodium azide was added to 0.02%.

d) Deionization of formamide

The formamide used for the denaturation of RNA for Northern transfer, hybridisation buffer used with RNA probes, and for denaturing gel loading buffer (formamide dyes) was deionised. 500mls of AR formamide was mixed with 10g of Amberlite Monobed MB1 or MB3 resin (BDH), stirring for 30 minutes. The formamide was decanted to a further 10g of resin and stirred for another 30 minutes. The pH was checked to

ensure that it was 7.0 and the formamide stored in small aliquots at -20°C.

e) Preparation of dialysis tubing

A large quantity of the required size of dialysis tubing was placed in a 2 litre beaker containing 1 litre of 1mM EDTA pH8.0. The tubing was weighted down with a conical flask half-filled with water. After boiling for 10 minutes, fresh buffer was substituted and the process repeated twice more. The tubing was then washed by boiling three times in distilled water for 10 minutes. After cooling the tubing was stored in 5mM EDTA pH8.0 at 4°C.

CHAPTER 3 : Isolation and structural analysis of genomic clones corresponding to the D19 gene.

cDNA clones corresponding to *Dictyostelium* mRNA species which change in their abundance during development have been isolated by differential screening of cDNA libraries (Williams et al., 1979; Rowekamp and Firtel, 1980; Mehdy et al., 1983; Jermyn et al., 1987). Many of the clones derived from aggregation stage mRNAs have been shown to hybridise preferentially to RNA isolated from prespore or prestalk cells (Barklis and Lodish, 1983; Mehdy et al., 1983) The cDNA clone D19 was initially isolated from a library prepared using polyA(+) RNA from cells at fifteen hours of development (Landfear et al., 1982; Mangiarotti et al., 1982). Using the labelled clone to probe a Northern blot of RNA from separated prespore and prestalk cells, D19 was found to hybridise to a highly prespore-enriched mRNA species (Barklis and Lodish, 1983).

3.1 Southern transfer analysis of D19 gene organisation.

Samples of CsCl gradient-purified *Dictyostelium* nuclear DNA were digested singly with NsiI, NdeI, HinfI and PvuII, separated by electrophoresis through a 0.8% agarose gel and transferred onto nitrocellulose. The restriction enzymes used were known to cleave within the cDNA insert (see section 3.3). HinfI cuts the *Dictyostelium* genome more frequently

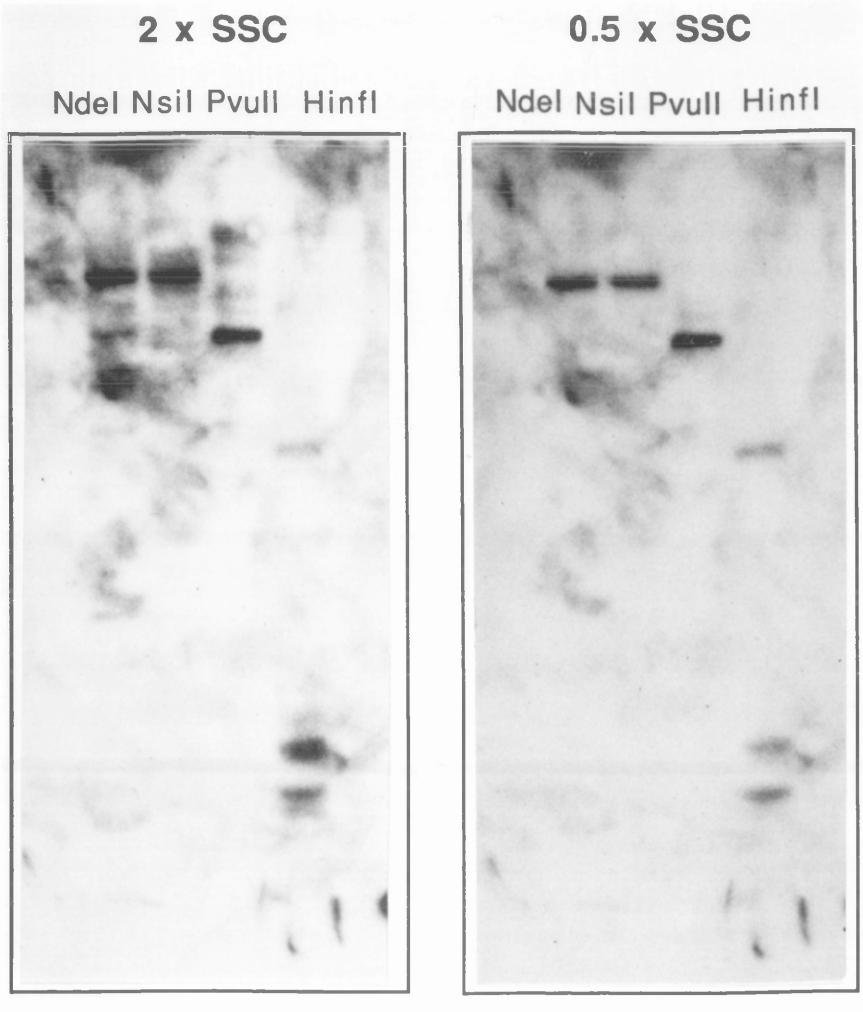


Figure 3.1 Southern blotting analysis of *Dictyostelium* genomic DNA

The two panels show a single Southern blot of *Dictyostelium* nuclear DNA, washed at the two different stringencies indicated. 2µg of CsCl gradient purified DNA, digested with the appropriate enzyme, was loaded per lane. The blot was probed with a nick-translated HaeIII fragment purified from the D19 genomic clone, spanning most of the D19 coding region.

than the other three enzymes because it has a 5 nucleotide recognition sequence while NsiI, NdeI and PvuII are 6-cutters. The Southern blot was probed with a 2.0kb HaeIII fragment purified from the genomic clone 4.3 (see section 3.3), labelled by nick-translation, and the filter was washed at different stringencies (Figure 3.1). Originally a fragment from the cDNA clone containing a short length of vector-derived sequences had been used, but cross-hybridisation problems resulted (data not shown). The HaeIII fragment spans a region extending from just 3' to the ATG, to a point about 1.3kb downstream of the termination codon. At low stringency (2x SSC, 0.1% SDS, 65°C), multiple bands were seen in each lane, although some bands were more intense than others. Washing at 1x SSC did not significantly reduce the hybridisation to any of the bands (data not shown), but at 0.5x SSC most of the signal was concentrated in a single band for NsiI, NdeI and PvuII, and three smaller bands for HinfI. The sites were chosen because they were each known to occur once within the cDNA insert (see section 3.2), but the pattern obtained with HinfI suggests that within the region spanned by the probe, multiple sites exist. The absence of two hybridising bands with the other three enzymes could be either because the sites lie close to one end of the region from which the probe is derived, or because a doublet was produced.

However, the presence of single bands implies strongly that D19 is a single copy gene. Supporting this, mapping of four independent genomic isolates has revealed no differences in genomic structure (see section 3.3). The survival of

multiple bands at 1x SSC suggests that genes with a related structure exist, either containing sequences of overall similarity, or a shorter, more homologous region. Survival of a hybrid at 1x SSC, 65°C for DNA with a GC content of approximately 33% (the value subsequently determined for the coding region of D19) requires homology of greater than about 82% using the formula $Tm = 16.4\log[Na^+]+0.41(GC)+81.5°C$.

3.2 Structure of the D19 cDNA clone.

The bank from which the D19 clone was originally isolated was constructed by insertion of cDNA into the BamHI site of pBR322 (Mangiarotti et al., 1982). An initial sample of DNA was provided by Dr.S.Cohen. The clone was first analysed by restriction enzyme mapping. The insert could not be excised directly as the BamHI site had been destroyed by cloning, but the insert length was calculated to be approximately 720 nucleotides. As the cDNA clone hybridises to a single messenger RNA species 700-750 nucleotides long, it seemed likely that the cDNA insert was virtually full-length. Of 19 restriction enzymes tried with a 6 nucleotide recognition sequence, only PvuII, NdeI and NsiI were found to cleave within the insert, each apparently at a unique site. In addition, a single HinfI site and multiple RsaI sites were identified. The map of the cDNA insert and its orientation with respect to pBR322 are shown in Figure 3.2.

The sequence of part of the cDNA insert was determined. By kinase labelling at the NdeI site, and sequencing by the chemical cleavage method (Maxam and Gilbert, 1980), 210 nucleotides of sequence were read. A potential uninterrupted

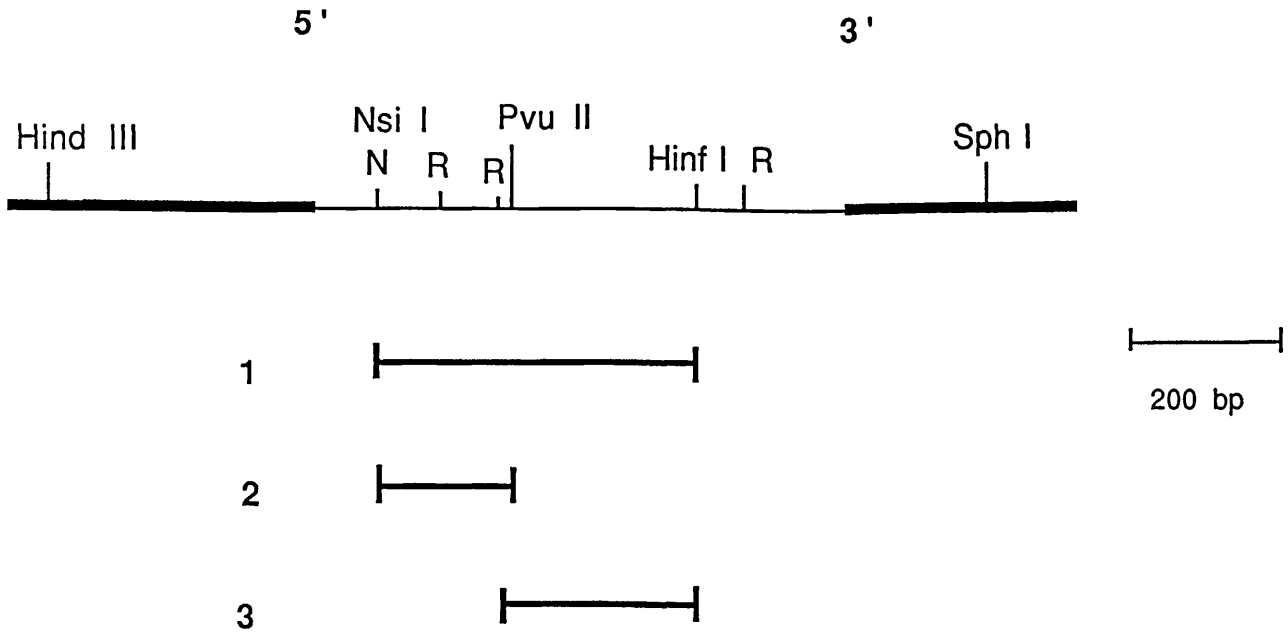


Figure 3.2 Restriction map of the D19 cDNA clone

The cDNA insert is represented by the thin line, and pBR32. sequences by the solid black lines. The bars drawn underneath indicate regions of the insert excised for use as hybridisatio: probes (see text for details).

open reading frame existed in one strand only, indicating the probable orientation of the message (see section 3.4).

3.3 Isolation and mapping of D19 genomic clones.

To isolate genomic clones corresponding to the D19 cDNA clone, a genomic library was screened with a fragment from the cDNA insert. The library was constructed by insertion of partially Sau3A digested, size-selected *Dictyostelium* nuclear DNA into the BamHI site of pAT153, to yield inserts of approximately 5kb (Twigg and Sherratt, 1980; J.Williams, personal communication). The 430bp NsiI-HinfI fragment, containing cDNA sequences only (see Figure 3.2), was gel purified, labelled by nick-translation and used as a probe. Twenty-one positives were picked from one filter of approximately 50,000 clones and, on secondary screening, eight of these survived washing at 0.1x SSC, 0.1% SDS at 60°C. This represented a degree of homology of 94% or greater, using the formula $Tm = 16.4\log[Na^+]+0.41(GC)+81.5°C$. Minipreps of the positive clones were digested with NsiI and HindIII, and Southern blotted. The blots were hybridised with the original probe, and again washed at high and low stringency. Those which continued to hybridise strongly at 0.1x SSC and which were cleaved within the insert were selected for further analysis, a total of four independent clones.

To verify that the genomic clones were homologous to the D19 gene, they were used in parallel with the D19 cDNA clone to probe Northern blots of *Dictyostelium* RNA (Figure 3.3). The RNA analysed was isolated from HM44 cells after 22 hours

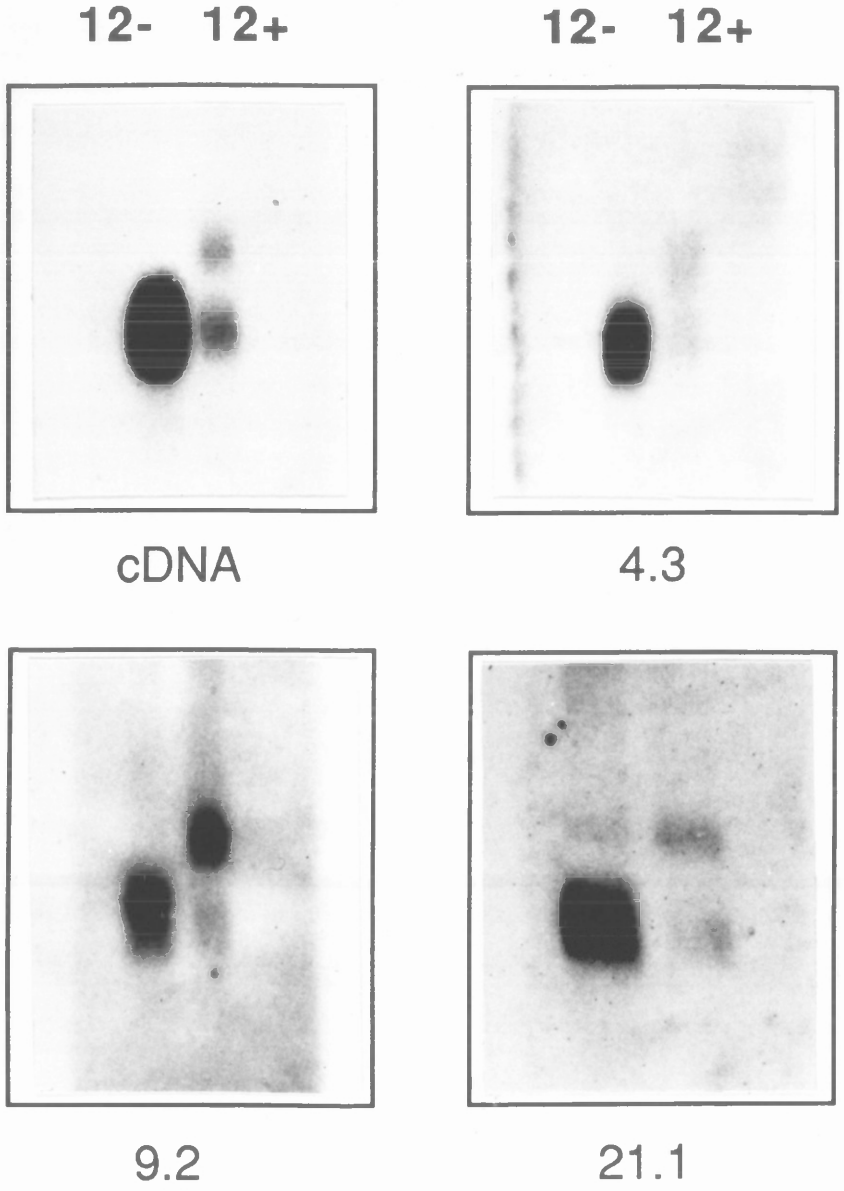


Figure 3.3 Use of positive clones from the library screen as probes in Northern blot analysis

5µg samples of total RNA, isolated from HM44 cells after monolayer culture development with cAMP and DIF or DIF alone, were electrophoresed through a 1% formaldehyde gel and transferred to nitrocellulose. Duplicate filters were probed with the D19 cDNA clone and the genomic clones, 4.3, 9.2 and 21.1, labelled by nick-translation. The numbers indicate the time of exposure to DIF where added, after the 10 hour cAMP preincubation.

of monolayer culture with cAMP, when D19 mRNA is known to be expressed at high levels, and from HM44 cells additionally treated with DIF, where two bands are observed at low stringency (K.Jermyn, personal communication; see section 4.2). Use of the D19 cDNA and all four genomic clones as probes gave an identical hybridisation pattern in each case. Hybridisation to the same sized band in the two RNA samples could not be removed by washing with 50% formamide, 0.1x SSC, 0.1% SDS at 37°C (data not shown). This strongly suggested that each genomic clone contains sequences homologous to the D19 gene.

The structure of the clones was analysed by restriction mapping, in combination with Southern blotting with two probes derived from different regions of the cDNA insert (shown in Figure 3.2). The PvuII-NsiI fragment was derived from the 5' end of the coding region, while the RsaI-HinfI fragment contained sequences closer to the 3' end. The unique PvuII in the centre of the cDNA insert was utilised to determine both the orientation of the genomic inserts with respect to pAT153, and the approximate amount of flanking sequence each clone contained. This was achieved by digesting with PvuII/SphI and PvuII/HindIII. SphI and HindIII are unique in pAT153 and positioned on opposite sides of the insert. In addition, the same probes were used with blots of the clones digested with HaeIII, RsaI, DdeI and HhaI. The results of both the original mapping and subsequent Southern blotting was consistent with all four clones having an identical structure over their shared regions (Figure 3.4). The map of the cDNA clone could also be superimposed.

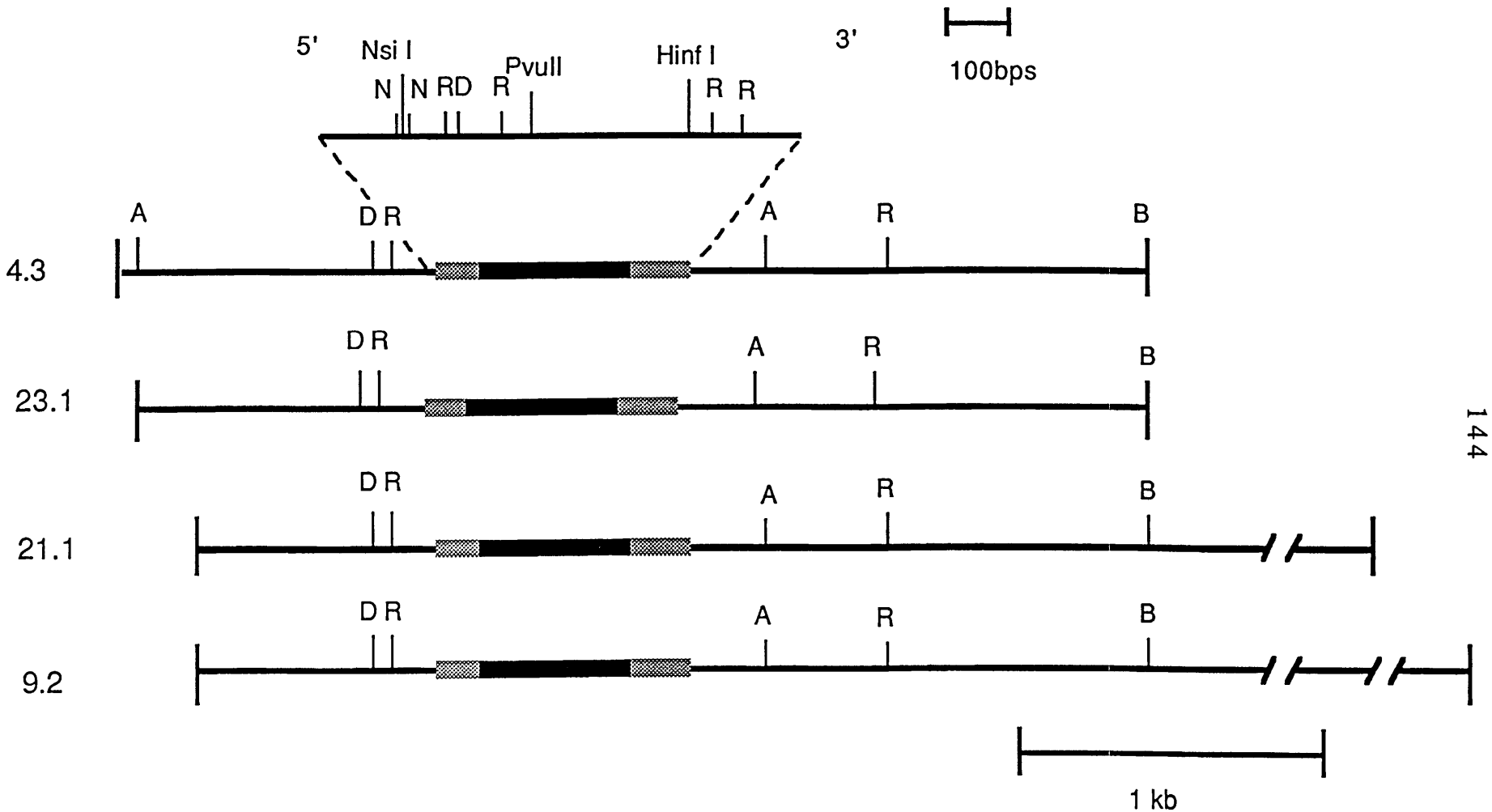


Figure 3.4 Restriction maps of the D19 genomic inserts

The transcribed region (▒▒▒) has been expanded two-fold relative to the intergenic DNA.
A =AvaII, B =BamHI, D =DdeI, N =NdeI, R =RsaI
intergenic DNA ▒▒▒ translated region ▒▒▒ 5' and 3' untranslated DNA

The genomic clones 4.3, 23.1, 21.1 and 9.2 have insert sizes of 3.37 kb, 3.30 kb, 5.24 kb and 13.3 kb respectively. Because of the initial size-selection step, the 13.3kb insert is likely to be due to the cloning of multiple Sau3A fragments. 4.3 was selected for further study as it has the greatest amount of sequence on either side of the PvuII site. The unique NsiI site at the 5' end of the gene is oriented towards the HindIII site of pAT153, while at the 3' end, a BamHI site has been regenerated.

3.4 Sequence analysis of the D19 genomic clone.

The approximate position of the coding region of the D19 gene within the genomic insert of 4.3 was deduced by comparison with the map of the cDNA clone. The identity in structure over this region indicated that the gene contained no introns. As the message size is relatively small (700-750 nucleotides), and the genomic map had been fairly well defined, the chemical cleavage method was used to determine the gene sequence spanning the coding region (Maxam and Gilbert, 1980). Details of the strategy employed are shown in Figure 3.5. As part of an initial strategy to mark the cloned D19 gene, a linker creating an XbaI site was inserted into the PvuII site. (For a description of the approach subsequently used to mark the gene, see section 5.3.) Sequence across the PvuII site was generated from the NdeI-HinfI fragment.

927 nucleotides of contiguous sequence was established, starting 283 nucleotides upstream of the putative initiation

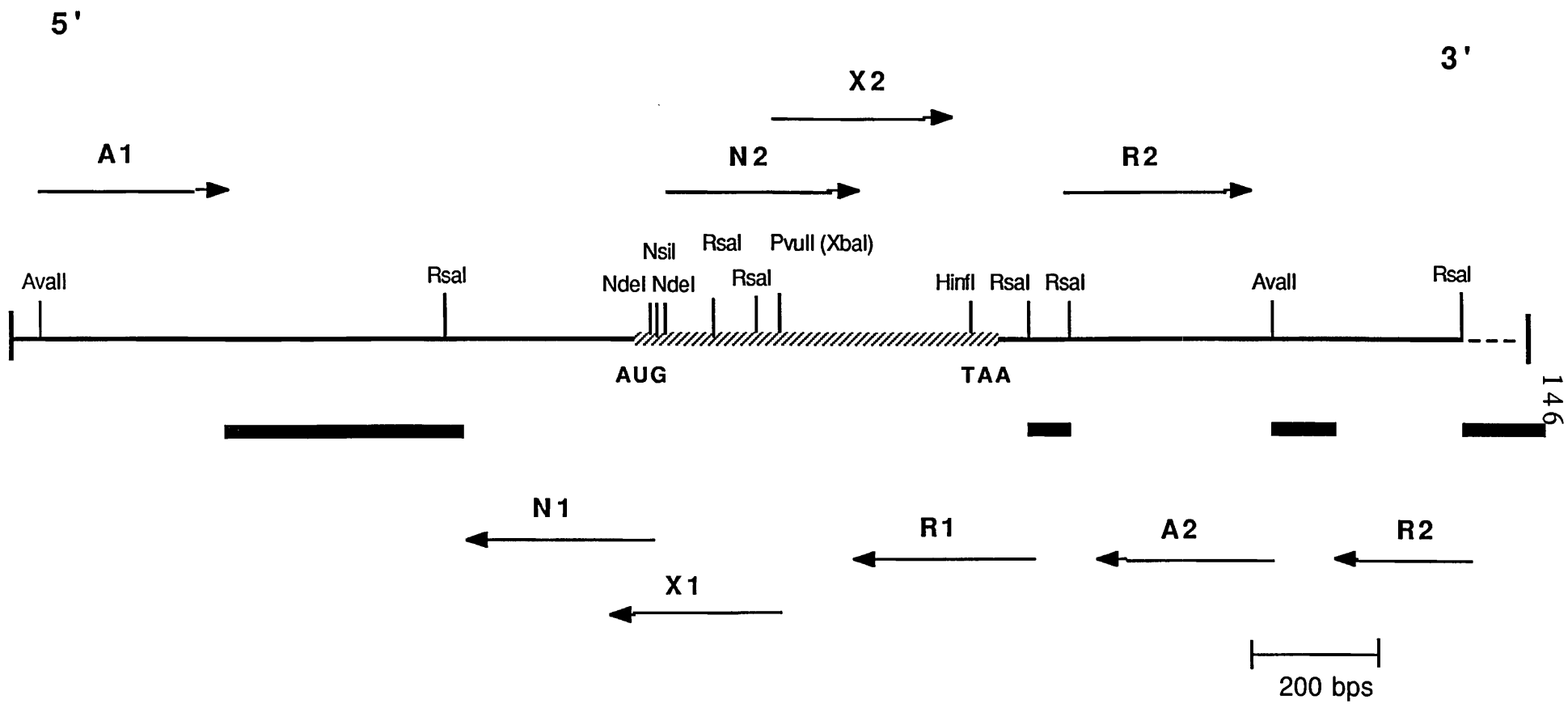


Figure 3.5 Strategy used to sequence the genomic clone 4.3

All fragments were kinased-labelled, and hence all sequence was read in a 5' to 3' direction The arrows indicate the length and direction of sequence read from a particular site. The letters denote the site labelled; A =AvaII, N =NdeI, R =RsaI, X =XbaI.

////////////// coding region ▬▬▬ region not sequenced

codon, and extending 137 nucleotides downstream of the putative termination codon (Figure 3.6). Between these regions is an open reading frame of 504 nucleotides containing no introns. All the restriction sites identified by mapping of the cDNA clone were also predicted by the genomic sequence, and in addition a 210 nucleotide block of genomic sequence was identical to that established from the cDNA clone. Considered in conjunction with the Northern blot analysis, it seems fair to conclude that the D19 cDNA clone and the genomic clone 4.3 are derived from the same gene.

Within the 4.3 insert, the D19 coding region is flanked by approximately 1 kb of 5' sequence and 1.8 kb of 3' sequence. The sequence of 380 nucleotides at the extreme 5' end of the insert, and of two unlinked regions each of about 300 nucleotides at the opposite end, was also established by the chemical cleavage method. The most distant sequence established downstream of the gene was approximately 900 nucleotides from the TAA. In all cases there was no evidence of any coding sequence, which is easily distinguished in *D.discoideum* from the very A-T rich flanking DNA. Therefore it is unlikely that another gene exists within 1 kb on either side of the D19 coding region, and hence another complete gene cannot be contained in the 4.3 insert.

3.5 Identification of the 5' and 3' termini of the D19 mRNA.

The length of the 3' untranslated region of the D19 mRNA was determined by comparison of genomic and cDNA sequence. The site of polyadenylation was judged from the position of

```
  1                                                    ATTTTATTAATATAATTCTTAATAATTATAAAAAAAAAAAAAAAA

 46 AAACAACCAAAATATTATATTATTTTTTGTAAATATAAATTAAAAAA ATG AAA TTC CAA CAT ACA
                                                       MET Lys Phe Gln His Thr   6

111 TTT ATT GCA TTA TTA TCA CTA TTA ACA TAT GCA AAT GCA TAT GAT TAT TTC ACA
    Phe Ile Ala Leu Leu Ser Leu Leu Thr Tyr Ala Asn Ala Tyr Asp Tyr Phe Thr  24

165 ACA ACA TTG GCC AAT CAA AAT CCA GTT TGT GCT TCA GTA GAT GTC ATA CAA AAT
    Thr Thr Leu Ala Asn Gln Asn Pro Val Cys Ala Ser Val Asp Val Ile Gln Asn  42

219 GTT TGT ACT GAG GTT TGT GGT AGA TTT GTC CGT TAC ATT CCT GAT GCT ACC AAT
    Val Cys Thr Glu Val Cys Gly Arg Phe Val Arg Tyr Ile Pro Asp Ala Thr Asn  60

273 ACA AAT CAA TTC ACT TTT GCC GAA TAT ACT ACA AAC CAA TGT ACT GTT CAA GTA
    Thr Asn Gln Phe Thr Phe Ala Glu Tyr Thr Thr Asn Gln Cys Thr Val Gln Val  78

327 ACA CCA GCT GTA ACA AAT ACA TTT ACT TGC GCT GAT CAA ACT TCT TCA CAT GCA
    Thr Pro Ala Val Thr Asn Thr Phe Thr Cys Ala Asp Gln Thr Ser Ser His Ala  96

381 CTC GGT TCT GAT TGG AGT GGA GTT TGT AAA ATC ACT GCC ACA CCC GCA CCA ACA
    Leu Gly Ser Asp Trp Ser Gly Val Cys Lys Ile Thr Ala Thr Pro Ala Pro Thr 114

435 GTT ACA CCA ACA GTT ACA CCA ACA GTT ACA CCA ACA GTA ACA CCA ACA CCA ACA
    Val Thr Pro Thr Val Thr Pro Thr Val Thr Pro Thr Val Thr Pro Thr Pro Thr 132

489 AAT ACA CCA AAT CCA ACC CCG TCC CAA ACA TCA ACA ACA ACT GGT TCT GCT TCA
    Asn Thr Pro Asn Pro Thr Pro Ser Gln Thr Ser Thr Thr Thr Gly Ser Ala Ser 150
                    •   •   •
543 ACT GTT GTT GCA AGT TTA TCA CTC ATT ATT TTC TCA ATG ATT CTT TCC CTT TGT
    Thr Val Val Ala Ser Leu Ser Leu Ile Ile Phe Ser MET Ile Leu Ser Leu Cys 168

597 TAATTTTAAAAATAACAAATGGTAAAAATTATTTGGCATTTTCAAGTACAATTTTAAATTTTTCAAAAAAA

668 AATAAAATTAAATAAATAATTAAATAAATTAATTAATAAAAAAAATAAAATTAAAACTTTTTTAATTTA
```

Figure 3.6 Sequence of the D19 mRNA

The numbers on the left denote nucleotides, and the numbers on the right amino acids. The symbols are as follows; •••, predicted N-glycosylation site, —— ,position of the four tetra-amino acid repeats.

the oligo dT stretch in the cDNA clone. Instead of the original pBR322 cDNA clone, a clone constructed in pUC8 was analysed. Since multilinker sites were available for labelling, sequencing of the ends of the insert by the chemical cleavage method was facilitated. Three cDNA clones, all confirmed by sequencing to be derived from the D19 mRNA, were isolated by screening a DIF-induced library with a fragment of the original cDNA clone (see section 4.4). Although none of the inserts was full-length, each had an oligo dT run at the same position, strongly suggesting that this represented the true 3' end of the message. Hence the 3' untranslated region of the D19 mRNA is 140 nucleotides in length, and is extremely A-T rich, having a total of only 11 G and C residues (see Figure 3.6).

The position of the start site of transcription was established by primer extension. A 23 nucleotide primer, 5'-CCAATGTTGTTGTGAAATAATCA-3', complementary to positions +68 to +81 downstream of the translation start site, was synthesized. This region was chosen because of its relative G-C richness, giving a predicted DNA-RNA hybrid with a Tm of approximately 60°C using the formula Tm = [4n(G+C)+2n(A+T)]°C. The kinase-labelled primer was annealed at 45°C to RNA isolated either from V12M2 prespore cells or from AX-2 slugs. Following extension with reverse transcriptase, the products were analysed by denaturing acrylamide gel electrophoresis. A single, identically-sized band was obtained with both RNA samples (Figure 3.7). This demonstrates that transcription of the D19 gene initiates at a single site approximately 92 nucleotides from the cap site.

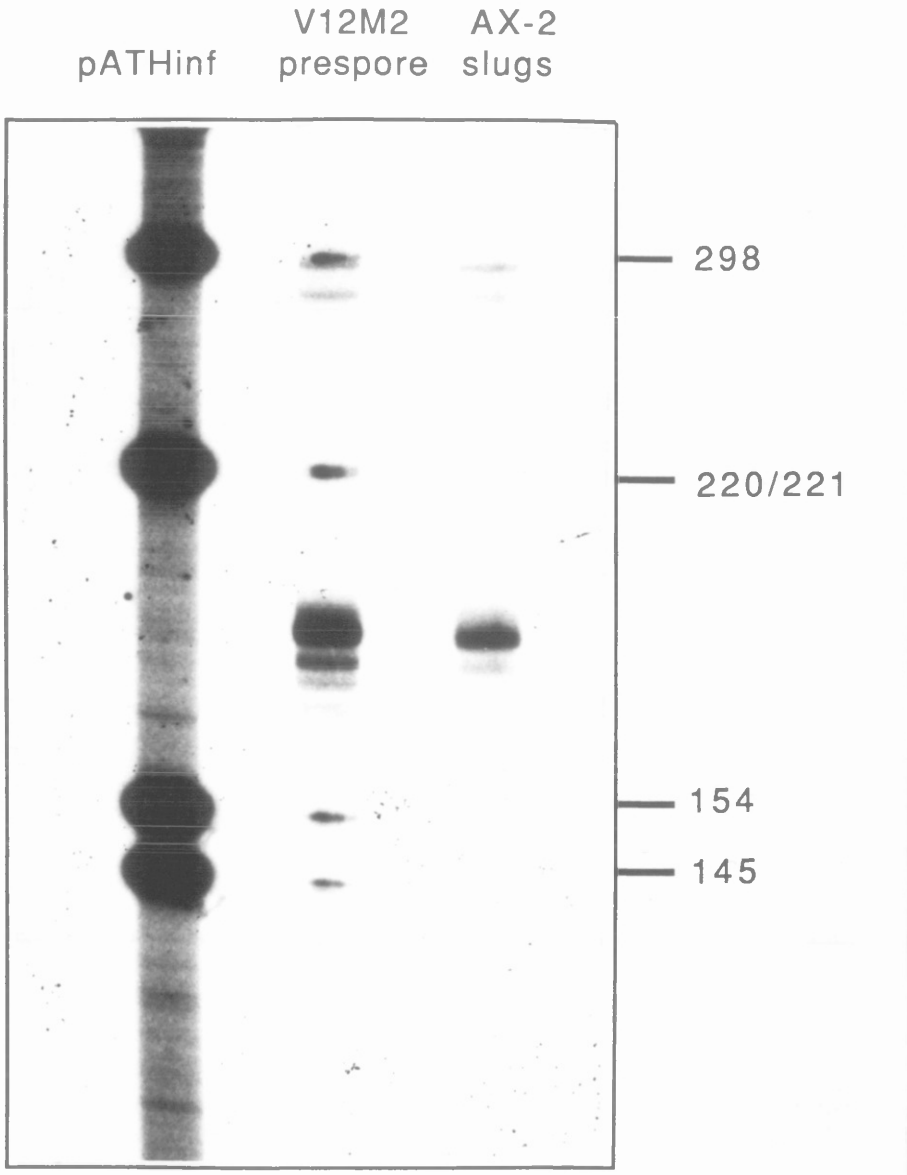


Figure 3.7 Identification of the 5' terminus of the D19 mRNA by primer extension analysis

5ng of primer, labelled with -32P ATP using T4 polynucleotide kinase, was annealed at 45°C to 5g samples of total RNA, isolated either from V12M2 prespore cells or from AX-2 slugs. The extension products were resolved by electrophoresis through a 7.5% denaturing acrylamide gel until the xylene cyanol had reached 20cm. The size of the extended product was estimated by comparison to the size markers (pATHinfI), and the start site of transcription was calculated relative to the known position of the sequence complementary to the primer.

The 5' noncoding region is also very A-T rich, containing only 5 non-contiguous C and G residues.

The estimated total length of the mRNA is 736 nucleotides, which is in good agreement with the results of gel electrophoresis (K.Jermyn, personal communication). The overall structure is similar to that predicted for many *D.discoideum* messages (see Chapter 1). The only other prespore message to be characterised to date, EB4, was reported to have a radically different 5' end structure (Barklis et al., 1985). A 350 nucleotide long G-C rich 5' untranslated region was predicted by S1 mapping, and it was suggested that these sequences could be involved in the control of message stability. As the half-life of the D19 message is apparently subject to the same control mechanisms as EB4 (Barklis and Lodish, 1983), the direct involvement of 5' noncoding sequences was not clear. Recently however, evidence has strongly suggested that the position of the promoter was mistakenly assigned, and the sequence presented is in fact derived from an intron (W. Nellen, personal communication).

3.6 Structure of the polypeptide encoded by the D19 gene.

The length of the D19 coding region is 168 amino acids (Figure 3.6). Analysis of the primary protein sequence was carried out using the programme PCGENE. The encoded protein has a predicted molecular weight of 17,879 daltons and a predicted pI of 4.62. It is composed of a higher than average proportion of nonpolar residues and a lower than average

proportion of potentially charged residues. Unusually it contains over 21% threonine residues, which are potential targets for O-linked glycosylation. There is one predicted N-linked glycosylation site within the C-terminal part of the protein at codon 136. At the N-terminus are a series of residues with the characteristics of a signal peptide (Briggs and Gierasch, 1986; Figure 3.8a). A hydrophobic "core" of fifteen amino acids with a predicted α-helical structure follows a short hydrophilic region adjacent to the methionine. Nineteen residues are found on the N-terminal side of the predicted cleavage site i.e. following Ala-Asn-Ala according to the -1-3 rule (Briggs and Gierasch, 1986). The extreme C terminus also contains a largely hydrophobic stretch of approximately 18 amino acids, but these residues are not followed by polar residues which could form a cytoplasmic anchor domain (Figure 3.9). This region is interspersed with hydrophilic residues, most commonly serine. Immediately before the hydrophobic tail are four copies of a perfectly conserved four amino acid repeat, Pro Thr Val Thr (underlined in Figure 3.6). The repeats form a proline-rich region, containing 10 of the total of 13 proline residues in the protein. Within this 31 amino acid stretch are 9 short regions where proline alternates with a hydroxyamino acid, normally a threonine. This pattern is interspersed with valine, asparagine or alanine residues, present singly. Overall, the primary amino acid sequence predicts a secreted polypeptide with an extended structure, possibly heavily glycosylated.

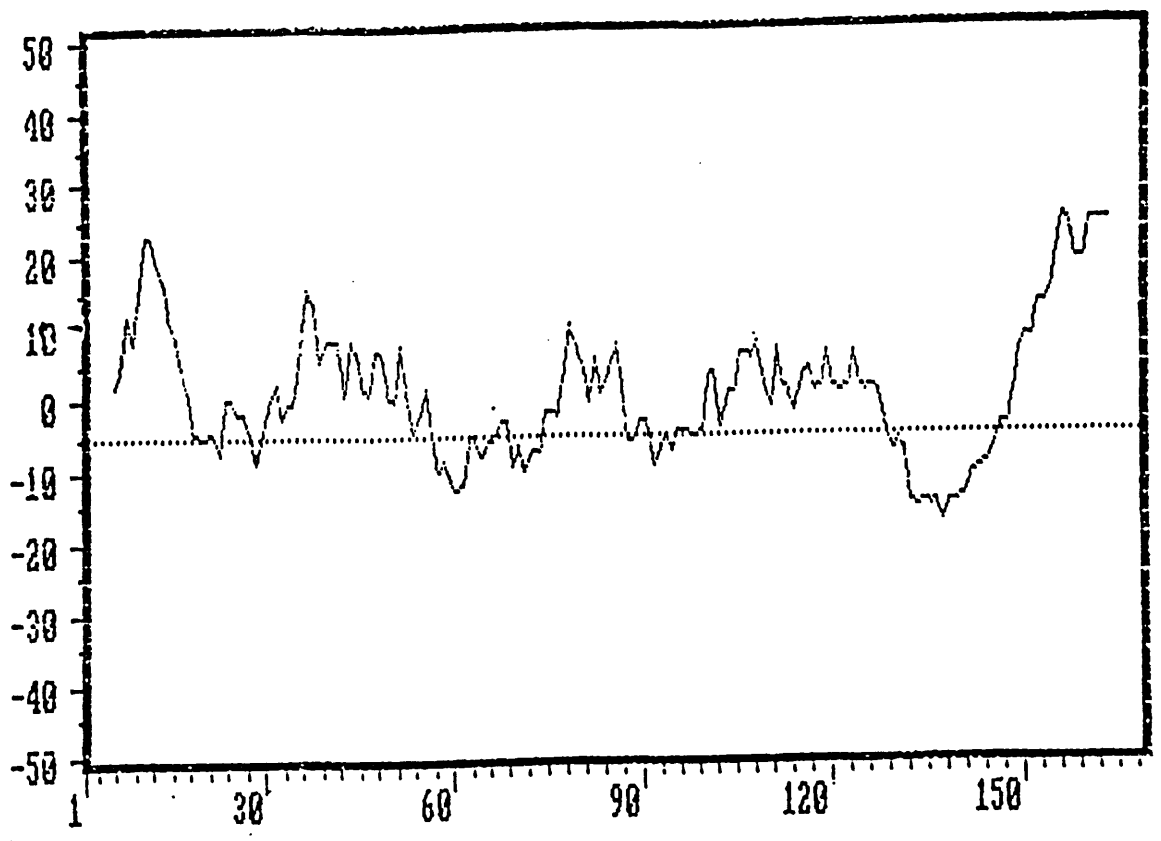


Figure 3.8 Plot of the hydropathic index of the predicted translation product of the D19 gene

Tha plot was computed according to the method of Kyte and Doolittle (1982) with a window size of 9 amino acids. The 168 amino acids of the predicted D19 coding region are represented along the x axis. On the y axis, negative values indicate hydrophobicity, while negative values indicate hydrophilicity.

3.7 Identity of the D19 gene product with PsA, and homologies to other proteins.

PsA is a prespore-specific cell surface glycoprotein, identified by Dr.K.Williams and his coworkers (Grant et al., 1985). It is known to be specifically recognised by the monoclonal antibody MUD-1 (Gregg et al., 1982), and to be identical to the prespore-specific protein sp29 (West and Loomis, 1985). Recently, milligramme quantities of PsA have been purified by affinity chromatography, and the N-terminal region has been subjected to peptide sequencing. The sequence of the known 81 amino acids was found to be identical to the D19 sequence immediately downstream of the predicted signal peptide (K.Williams, personal communication). As the D19 gene is unique in the genome (see section 3.1) and as it has been demonstrated that PsA is encoded by a single locus, pspA (Grant et al., 1985), the D19 gene must encode PsA.

Other predicted characteristics of the D19 protein match the observed characteristics of PsA. The experimentally determined pI of PsA was reported to be between 4 and 5, and the prediction of a signal peptide is consistent with the localisation of PsA on the surface of prespore cells and in the slime sheath, demonstrated with MUD-1 (Williams et al., 1984; Krefft et al., 1985). However, there is a discrepancy between the predicted and actual molecular weights, as PsA has been judged to be 30kD (Krefft et al., 1983). This difference could be accounted for by post-translational modifications. PsA is heavily glycosylated, and in a glycosylation-defective mutant (modB, lacking type II carbohydrate) the molecular weight, 23kD, is much closer to

that predicted from the primary sequence (West and Loomis, 1985; Hohmann et al., 1987; K.Williams, personal communication). In addition, the high hydroxyamino acid content (21% threonine, 8% serine) may give rise to anomalous migration through SDS gels (Both, 1987).

PsA is a polymorphic protein. A separately occurring larger and smaller species have been detected in several wild isolates (Grant et al., 1985). The variant forms have regular ± 2kD differences in molecular weight, best explained by addition or deletion of identical 2kD units from the wild-type 30kD protein. The primary sequence data from the D19 genomic clone suggests a possible mechanism for the generation of this polymorphism. Unequal crossing-over of the repeats would result in the loss of up to three copies of the repeats from one chromosome and insertion of up to three extra copies into the other; at least one copy of the repeat would always be retained. In this way, exact insertion or deletion of up to 12 amino acids might occur. The repeat contains two threonine residues, and if they were subject to O-linked glycosylation, differences of 12 amino acids could account for the molecular weight differences of 2kD. Peptide sequencing of the variant forms of the protein is in progress (K.Williams, personal communication).

There is homology between PsA and the *D.discoideum* protein contact sites A (CsA) over a 61 amino acid region (Figure 3.8b). CsA is known to be involved in establishing EDTA-resistant contacts early in development (Muller and Gerisch, 1978). The sequence of a contact sites cDNA clone has been established (Noegel et al., 1986). The homology to

Figure 3.9a Predicted PsA leader peptide sequence

MET lys phe gln his thr phe ile ala leu leu ser leu leu thr tyr ala asn ala tyr asp

putative cleavage site

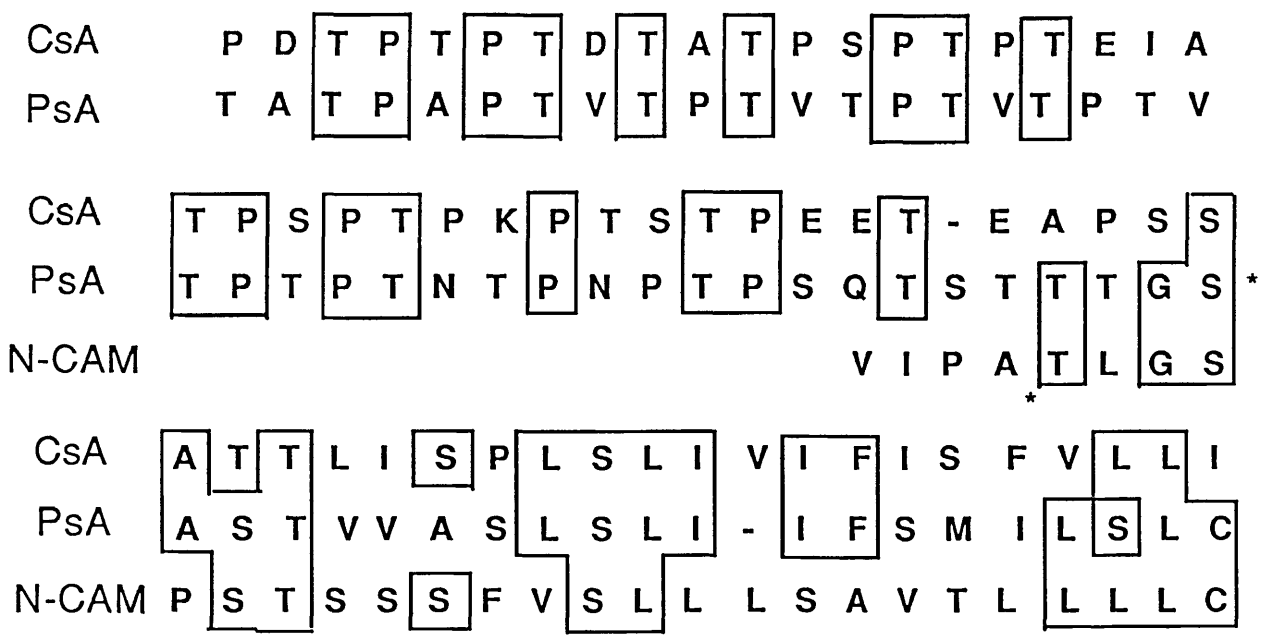


Figure 3.9b Regions of homology between PsA, CsA and N-CAM

The C-terminal 60 amino acids of PsA and CsA (Noegel et al., 1986) and the C-terminal 29 amino acids of chicken N-CAM are compared (Hemperley et al., 1986). The asterisk denotes the position of cleavage prior to the addition of the GPI group, where established. The 25 amino acids following the site of cleavage in N-CAM are unique to the GPI-linked form.

PsA is at the extreme C terminus and in the proline-rich repeats, which in both proteins are found immediately upstream of the C terminus. The repeat sequence in CsA is 32 amino acids long and contains three clusters of proline residues which alternate with serine or threonine. In PsA, the comparable region is 31 amino acids in length and contains 9 shorter interspersed groups of alternating threonines and prolines. For CsA there has been speculation that such a primary sequence is responsible for forming an extended chain in proline-proline conformation, which could project the protein into the extracellular space (Noegel et al., 1986). The homology between CsA and PsA in this region implies that PsA has a similar functional requirement for such a stalk-like structure.

The homology at the extreme C terminus is shared by another known cell-adhesion molecule, namely the glycosyl-phosphatidylinositol-linked (GPI) form of the chicken neural cell adhesion molecule (N-CAM) (Hemperley et al., 1986). This form of membrane-attachment was recently also shown for CsA, using ^{3}H-inositol (G.Gerisch, personal communication). This constitutes strong but indirect evidence that PsA is covalently attached to the plasma membrane via an inositol-containing lipid. This type of linkage was originally identified for the variant surface glycoproteins of trypanosomes, but has since been identified for a wide variety of membrane proteins (Low et al., 1986; Cross, 1987). All cell surface proteins known to be attached in this manner lack putative transmembrane domains, like PsA, and have characteristically hydrophobic C termini containing

interspersed hydrophilic residues. In all cases so far where the polypeptide sequence predicted from the cloned gene has been compared to that of the mature protein (including chicken N-CAM) it has been shown that the C terminus is cleaved off before covalent attachment of the new terminal amino acid to a sugar residue of the glycolipid (Ferguson et al., 1985; Tse et al., 1985). Therefore it is a strong prediction that both CsA and PsA will be processed in the same manner.

Recently the peptide sequence of PsA was completed, and as predicted, the terminal twenty amino acids were missing from the protein (K.Williams, personal communication), although direct experiments are still required to prove a PI link exists. The site of cleavage is between a serine and an alanine residue. No consensus sequence has yet been identified for protein cleavage and addition of PI, and if the situation reflects that encountered for signal peptides, a direct primary sequence homology may not exist. Hence the C terminal homology between CsA and PsA (10/22) and between PsA and chicken N-CAM (11/25) is intruiging, and at best may be suggestive that the three genes belong to the same family. However, direct evidence is clearly needed to indicate a role for PsA in cell adhesion (see Chapter 8). Recently it has been suggested that there is a structural similarity between the C terminus of CsA and a region of the 160,000 dalton chicken N-CAM (Matsunaga and Mori, 1987). However, as the comparison is made to the large intracellular domain of the chicken N-CAM, the functional significance of this similarity is not apparent.

3.8 Attempts to produce a PsA-deficient mutant.

In *D.discoideum*, homologous recombination has recently been utilised to generate mutants lacking specific gene products (De Lozanne and Spudich, 1987; Witke et al., 1987). This approach is possible since the organism is haploid. Transformed incomplete copies of two cloned genes have been found to recombine with their chromosomal homologues; at an estimated frequency of 30% for myosin (De Lozanne and Spudich, 1987), and 5% for α-actinin (W.Witke, personal communication). For myosin, inactivation of the endogenous gene was achieved by insertion of the transformation vector containing a 5' gene fragment into 5' flanking sequences. However, for α-actinin, no vector insertion or gross gene rearrangements were detected, although absence of the protein was correlated with a reproducible alteration of the mRNA. Again, a fragment spanning the amino-terminal coding sequences was used for transformation.

The myosin mutants were originally identified by the alteration in phenotype of a subset of the population, but transformants deficient in α-actinin were isolated directly using a colony blot technique in conjunction with monoclonal antibody probes (Witke et al., 1987). As a specific monoclonal antibody (MUD-1) to the PsA protein exists, it is theoretically possible to isolate PsA-deficient mutants using the same approach.

Initially two plasmids were cotransformed (see Chapter 5); the extrachromosomal vector pCIBBKSP, to encode G418 resistance (Ahern et al., 1988), and a pUC8 based D19 cDNA

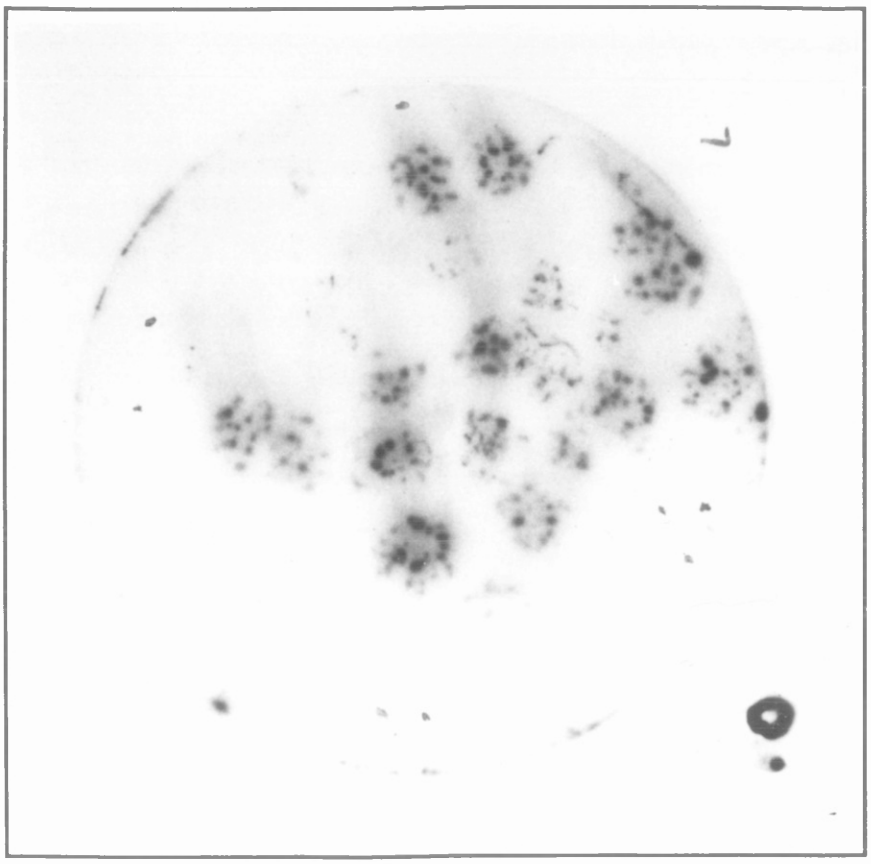


Figure 3.10 Colony blot of AX-2 transformants probed with the monoclonal antibody MUD-1

Transformed clones, plated at low density on a lawn of *Klebsiella aerogenes* after plating at low density to obtain individual colonies, were lifted onto nitrocellulose. The colonies had been allowed to grow until developing cells formed as a result of depletion of the bacterial food source. The cells on the filters were lysed by freeze-thawing and the debris removed by washing in low salt buffer before probing with MUD-1 at a dilution of 1 in 100. ^{125}I-labelled protein A was used to detect antibody binding.

clone containing approximately 600 base-pairs of sequence derived from the 3' end of the gene (see section 4.3). The transformants were pooled and diluted before plating onto bateria to give clonal growth. The ultimate goal was to isolate a mutant, regain G418 sensitivity by loss of the extrachromosomal vector, and then to rescue any resulting mutant phenotype by transformation of an intact copy of the gene using G418 selection. Of 180 colonies screened with MUD-1, all were still expressing PsA (Figure 3.10). As the mechanism of cotransformation is not yet well defined in *Dictyostelium* (Nellen and Firtel, 1985; Firtel et al., 1985), it was uncertain how many of the clones contained both plasmids. If only a subset of the clones contained the pUC8 plasmid, the absence of mutants could be explained.

A second strategy was attempted, using a deleted version of the D19 genomic clone subcloned into the transformation vector pB10TP2. Deletion of the 5' end of the insert (see section 5.6) had left 127 nucleotides upstream of the cap site, and the construct was known to contain insufficient promoter sequences for expression. However, 125 colonies screened again all continued to produce PsA. Therefore the frequency of mutant production is obviously less than the 5% observed for other genes. Several possible reasons exist for this reduced frequency; for myosin and α-actinin the transformed sequences may contain "hotspots" of recombination absent from the D19 gene; consistent with this, reproducible gene rearrangements were detected using Southern and Northern blotting (De Lozanne and Spudich, 1987; Witke et al., 1987). 3' sequences may be less efficient than 5' sequences and, in

the second approach, the transformed gene may have been too large a target to obtain a specific mutant perhaps necessary to abolish PsA synthesis. An alternative but very unlikely possibility is that although PsA has been mapped to a single linkage group (Grant et al., 1985) and Southern blotting suggests that D19 is a single copy gene (section 3.1), there are in fact two highly conserved linked copies. Further attempts to obtain an insertional mutant have been discontinued.

3.9 Summary

1. The D19 gene is present as a single copy in the genome, but a closely related gene family exists.

2. Four independently isolated genomic clones have an identical structure over their shared regions and sequence data from one of the clones has shown the gene to be devoid of introns.

3. Initiation of transcription occurs at a single site. The 5' and 3' noncoding sequences are A-T rich, and span about 90 and 140 nucleotides of the 736 nucleotide-long message repectively.

4. The D19 gene encodes PsA, a prespore-specific cell-surface glycoprotein recognised by the monoclonal antibody MUD-1. Size polymorphisms seen for PsA may be due to unequal crossing-over at the repeat region near the C-terminus.

5. The structure of the C-terminus of the PsA protein, and its homology to contact sites A and a form of chicken N-CAM, strongly suggest that PsA is linked to the cell membrane via phosphatidylinositol. However, evidence as to the function of PsA must await the successful generation of a mutant.

CHAPTER 4: The mechanism of repression of the D19 gene by DIF.

A subset of genes, expressed predominantly in the prespore zone of the slug after aggregation, seems to be dependent on high levels of cAMP both for induction and maintenance of expression (see Chapter 1; Barklis and Lodish, 1983; Oyama and Blumberg, 1986a). The response of D19 to the reduction in cAMP levels on disruption of aggregates, and to the addition of exogenous cAMP, places it in this class (Chisholm et al., 1984). Removal of cAMP produces a rapid decline in D19 mRNA levels by a combination of a sharply reduced transcription rate and the specific destabilisation of the message (Landfear et al., 1982; Barklis and Lodish, 1983). In contrast, two substances have been identified which can reduce the expression of prespore genes, namely DIF and adenosine. Adenosine, when added to aggregation-competent cells shaken in suspension, inhibits the cAMP-mediated induction of prespore gene expression (Weijer and Durston, 1985; Wang et al., 1986). DIF (Differentiation Inducing Factor) is a low molecular-weight chlorinated compound of known structure (Morris et al., 1987) which induces stalk cell differentiation (Town et al., 1976). If slugs are disaggregated and incubated in the presence of DIF, prespore gene expression is repressed (Kay and Jermyn, 1983). The effect of DIF on the expression of individual genes is most conveniently studied in vitro using the HM44 mutant, which

does not produce its own DIF but remains responsive to exogenously added DIF (Kopachik et al., 1983).

4.1 The HM44 in vitro assay system.

HM44 was isolated after MNNG mutagenesis of its parental strain, HM27, a derivative of V12M2 (Kopachik et al., 1983). As a consequence of its failure to accumulate DIF, it is blocked at the loose mound stage of development. This is at a time immediately following aggregation when HM27 is first forming a tip. The level of endogenous DIF production is sufficiently low to study the effect of exogenous DIF on HM44 cells incubated *in vitro* at high density. During development in tissue culture dishes under a buffered salts solution, the cells remain as a monolayer, and hence all are equally exposed to added DIF.

When DIF is added to HM44 cells which have been incubated in cAMP-containing medium for ten hours, 94% of the amoebae differentiate to form fully vacuolated stalk cells after a further 24 hours of development (Kopachik et al., 1983). In the absence of DIF, stalk cells constitute fewer than 1% of the total (Figure 4.1). A comparison was made of the proteins produced by HM44 cells induced to form stalk cells by DIF, with those present after incubation for the same length of time but without DIF (Kopachik et al., 1985). With DIF, 21 out of 23 proteins known to be enriched in stalk cells were expressed at greatly increased levels but in contrast, four major spore-specific proteins were repressed. This suggests that DIF acts to divert cells from the prespore to the stalk cell pathway of differentiation (see Chapter 7).

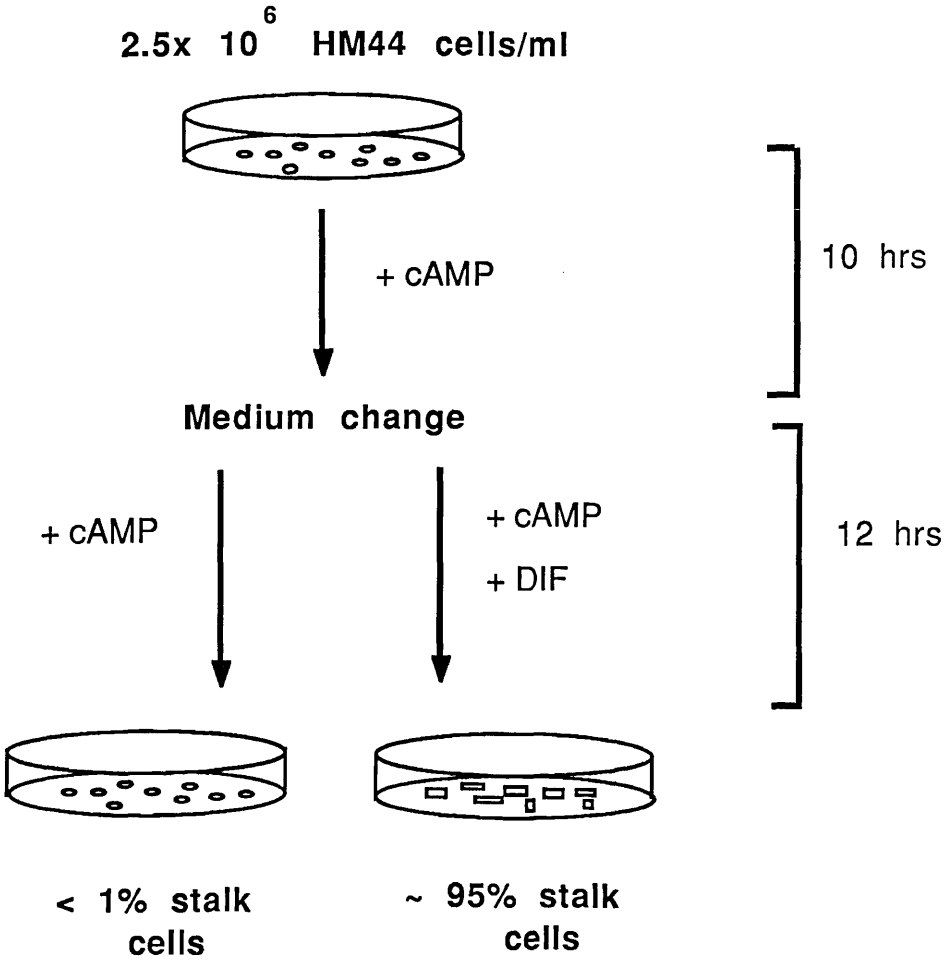


Figure 4.1 The HM44 monolayer assay system

HM44 cells grown on Klebsiella aerogenes on SM plates were washed free of bacteria and incubated in medium containing 10mM MES pH 6.2, 10mM KCl, 2mM NaCl, 1mM $CaCl_2$, 15µg/ml tetracycline 100µg/ml ampicillin, 5mM cAMP, and 3000 U DIF/ml where adde (15mls per 14cm tissue culture plate). When harvesting timepoints as described in Chapter 4, the medium change after 10 hours of preincubation was designated t_0.

4.2 DIF repression of D19 and two other prespore genes in vitro.

The HM44 assay system was used to examine the effect of DIF on prespore-specific gene expression at the RNA level. A stalk cell induction was performed by incubating cells in medium containing 5mM cAMP for ten hours before substitution of medium additionally containing 3000U of DIF per ml to half of the plates (Kopachik et al., 1985; see Chapter 2). RNA was isolated from timepoints harvested in parallel during development, and the levels of mRNAs corresponding to specific cloned genes were compared in the presence and absence of DIF. Using the labelled D19 cDNA clone as a probe, it was found that D19 mRNA levels induced during the cAMP preincubation remain constant for at least 24 hours without DIF, but when DIF is added, message levels decline rapidly (K.Jermyn, personal communication; Figure 4.2). After 4 hours, only 10% of the former level remained. As no further reduction was seen, the residual message may derive from those amoebae which never mature into stalk cells. It takes about 90 minutes for the amount of D19 mRNA to be halved.

Two other prespore mRNA sequences, corresponding to the cDNA clones PL3 and D7 (Barklis and Lodish, 1983; Oyama and Blumberg, 1986c), are repressed by DIF with identical kinetics (Figure 4.3). The labelled clones were again used to probe Northern blots of RNA isolated from HM44 cells. PL3 and D7, like D19, have been assigned to prespore class II on the basis of their time of appearance and sensitivity to extracellular cAMP (Chisholm et al., 1984). Therefore it seems likely that they are repressed by a similar mechanism.

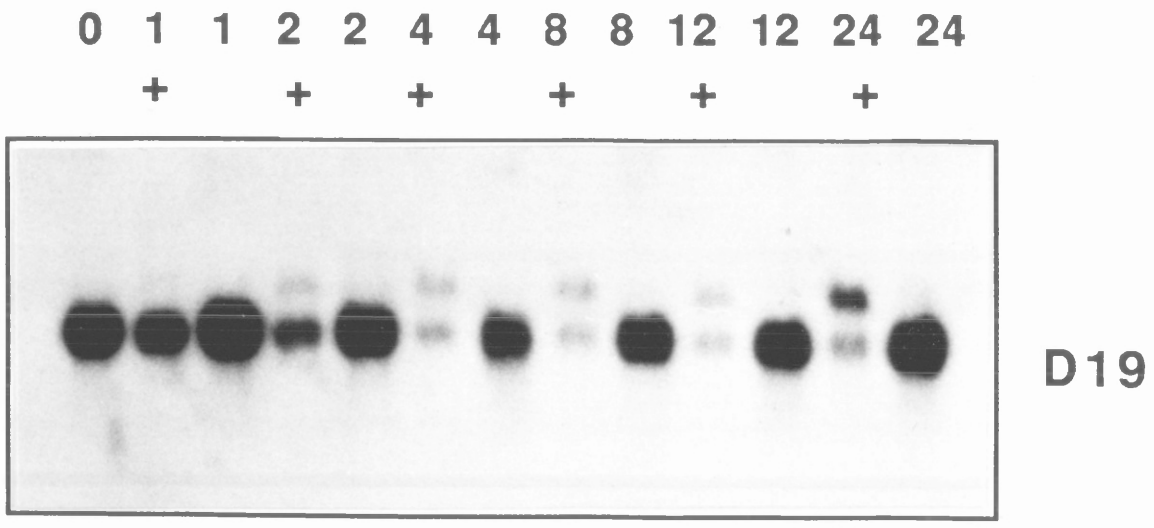


Figure 4.2 The kinetics of repression of the D19 mRNA by DIF

The symbol 0 indicates the time at which DIF, to 3000U/ml, was added to HM44 cells in monolayer culture after 10 hours' preincubation with medium containing 5mM cAMP. The incubation time after DIF addition is shown. Total nucleic acid was purified and 5µg samples analysed by Northern transfer after electrophoresis through a 1% formaldehyde gel, using the D19 cDNA clone as a probe. The filter was washed for 4x 15' with 2x SSC, 0.1% SDS at 65°C.

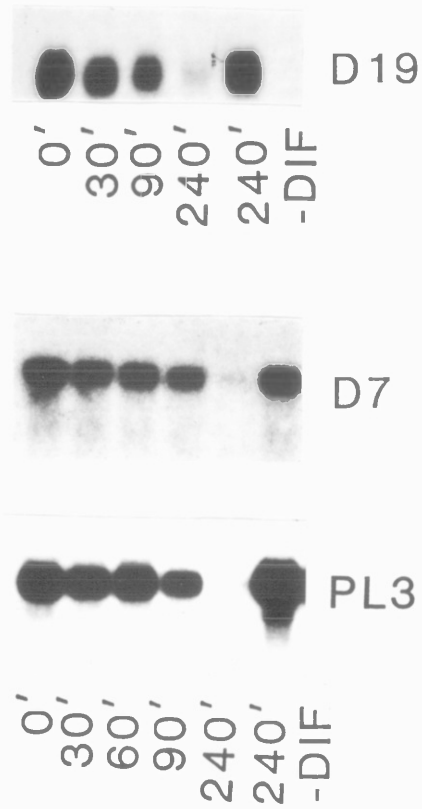


Figure 4.3 Comparison of the repression of D19 with that of two other prespore genes

HM44 cells were incubated for the times shown in the presence or absence of DIF, after 10 hours' incubation in medium containing DIF alone. Total nucleic acid was extracted and 5µg samples analysed by Northern transfer after electrophoretic separation on a 1% formaldehyde gel. A mixture of the D7 and the PL3 cDNA clones was used to probe a single filter. The blot probed with the D19 cDNA clone derives from a separate experiment, but in both cases greater than 90% of amoebae formed stalk cells after 12 hours' incubation in the presence of DIF. The filters were washed with buffer containing 2x SSC as previously described.

4.3 A D19-related sequence induced by DIF during stalk cell induction.

It was observed that as levels of D19 mRNA declined, a higher molecular-weight band appeared which cross-hybridised to the cDNA probe (K.Jermyn, personal communication, Figure 4.2). The band, approximately 100 nucleotides larger than D19 mRNA, and reproducibly seen in RNA samples from different experiments, is first detectable two hours after DIF addition. It reaches peak levels after 24 hours of induction. Hybridisation to the band can be completely removed by washing with 50% formamide, 0.5x SSC, 0.1% SDS at 37°C, but survives at 2x SSC (Figure 4.4a). Using the formula Tm $=16.4\log[Na^+] +0.41(GC) +81.5°C$, the approximate degree of homology of the hybrid between the RNA species and the DNA probe was calculated. Assuming a 33% GC content (that of D19 coding sequence), the homology is between about 65% and 75%.

This higher molecular weight species has not been detected in timecourses from the wild-type strains V12M2 and NC4 (K.Jermyn, personal communication). Nevertheless, it is of interest to understand how DIF can lead to the generation of an mRNA highly related to a prespore-specific sequence normally repressed by DIF. Either the message is derived from a separate gene with a promoter which is inactive during normal development, or during stalk cell differentiation in vitro, transcription of the D19 gene leads to mRNA with an altered structure. As the D19 gene contains no introns, it is impossible for the upper band to be the product of differential splicing. Additional sequences at the 5' or 3'

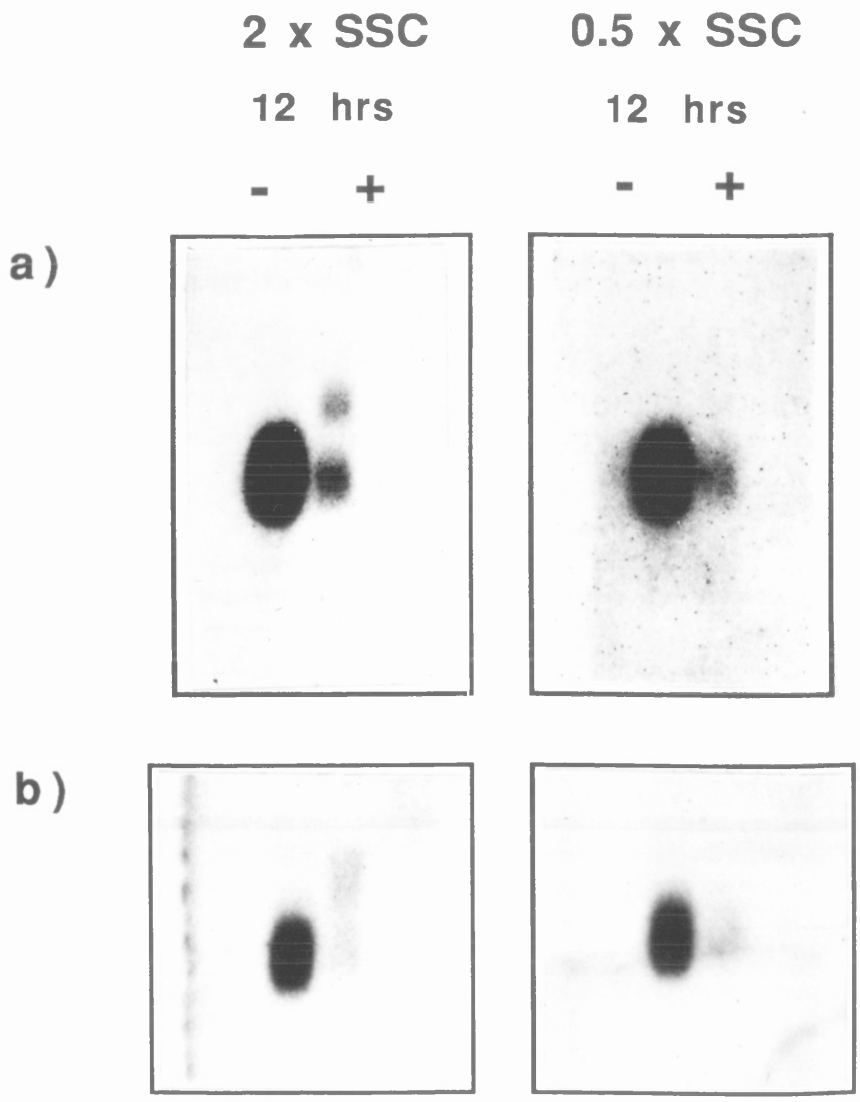


Figure 4.4 Analysis of the higher molecular weight cross-hybridising mRNA species

5µg samples of RNA, extracted from HM44 cells incubated in monolayer culture either in the presence or absence of DIF, were run side by side on a 1% formaldehyde gel and after trnsferring to nitrocellulose the blots were probed with a) the D19 cDNA clone and b) the D19 genomic clone 4.3. The filters were washed initially with 2x SSC, 0.1% SDS at 65°C, followed by 0.5x SSC, 0.1% SDS at 65°C.

ends, resulting from the use of alternative promoters or polyadenylation signals, would create insufficient structural differences to cause differential hybridisation of the cDNA probe. For the same reasons it is very unlikely that DIF blocks processing of a D19 nuclear precursor, and this was confirmed by the use of the genomic clone 4.3 as a probe. Although its structure should be identical to that of any precursor, hybridisation to the upper band was again removed at 0.5x SSC (Figure 4.4b). Therefore the higher molecular weight message must be derived from a separate gene.

Attempts were made to isolate both cDNA and genomic clones containing the D19-related DIF-induced sequence. An HM44 cDNA bank was constructed by inserting G-tailed cDNA, synthesized from poly(A+)RNA isolated after 12 hours' incubation with DIF, into the C-tailed PstI site of pUC8 (K.Jermyn, personal communication). The bank was screened with an internal fragment from the D19 cDNA insert (see section 3.3) and washed at low stringency. Three positive clones were obtained, but sequence from the ends of the inserts (using the chemical cleavage method after labelling at sites in the multilinker) revealed complete identity with the established D19 sequence (data not shown). As part of the sequence spans the 3' untranslated region, it is extremely unlikely that a related gene would have no mismatches.

As described in section 3.3, thirteen of the genomic clones originally isolated do not hybridise to the cDNA clone at high stringency. Five of these clones were used to probe Northern blots of RNA isolated from HM44 cells after 12 hours' incubation with DIF. All hybridised to the upper band

at low stringency, but failed to hybridise at high stringency (0.1x SSC, data not shown). Therefore, none of the clones tested has complete identity with the higher molecular weight message. Due to the failure to isolate a cDNA clone, further analysis of existing genomic clones was considered too complex.

4.4 DIF induction of 56 and 63 gene transcription.

pDd63 and pDd56, isolated by differential screening of the DIF-induced cDNA bank described in section 4.3, are derived from mRNA sequences which are highly prestalk-enriched during normal development, and accumulate only in response to DIF during stalk cell induction in HM44 cells (Jermyn et al., 1987). Nuclear run-on assays were used to determine the level at which DIF acts to induce accumulation of these mRNAs (Figure 4.5). Transcriptionally active nuclei were harvested from HM44 cells at various times during a stalk cell induction. The cells were incubated in tissue culture dishes for ten hours with cAMP before DIF addition. Nuclei were harvested just prior to the medium change (0') and at 15, 30, 90 and 240 minutes after DIF addition. Nuclei were also isolated from cells incubated for 240 minutes in medium with no DIF.

The nuclei were incubated under conditions in which pre-initiated transcripts are elongated (Jacobson, 1976), and the labelled RNA was hybridised to Southern blots of plasmid DNAs, digested to purify a fragment of the insert from the remainder of the vector. Thus the vector band acts as a control for background hybridisation. The cDNA clone IG7 was

also included. IG7 derives from an mRNA which is expressed constitutively during development, and which is equally abundant in prestalk and prespore cells (K.Jermyn, personal communication), and so was used to allow comparison of hybridisation between timepoints.

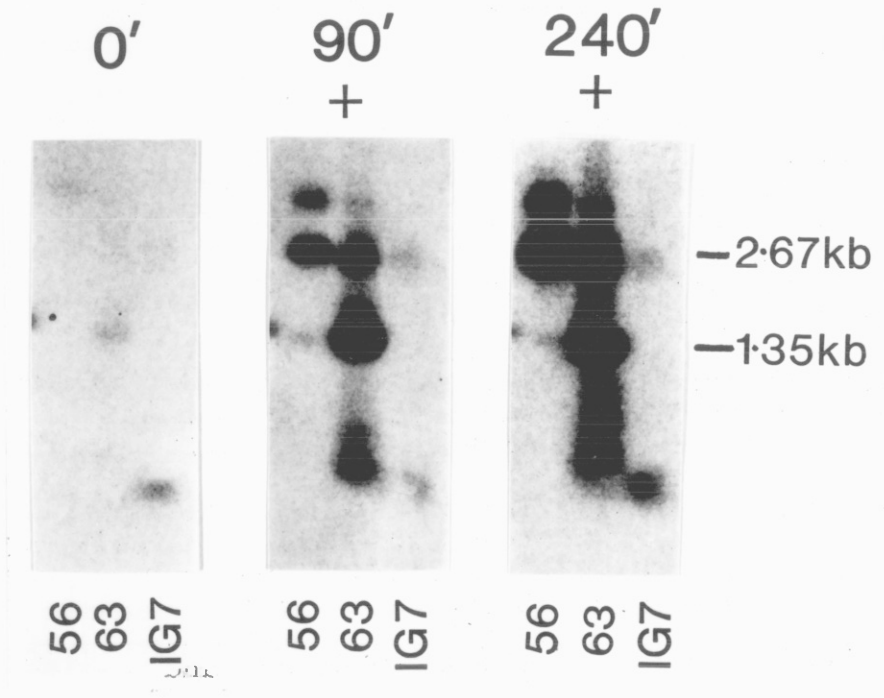
There is a very low level of 56 and 63 gene transcription in control nuclei (0' and 240' without DIF), but 15 minutes after DIF addition there is a detectable increase in transcription of 63. The major rise in transcription of 56 is not seen until between 90 minutes and four hours after DIF addition, but for both genes the timing of the increases in transcription is closely correlated with the appearance of cytoplasmic mRNA. pDd63 mRNA accumulates linearly, while pDd56 mRNA displays sigmoidal kinetics (Jermyn et al., 1987). Four hours after DIF addition, there is approximately a 50-fold elevation in the rate of transcription of both genes.

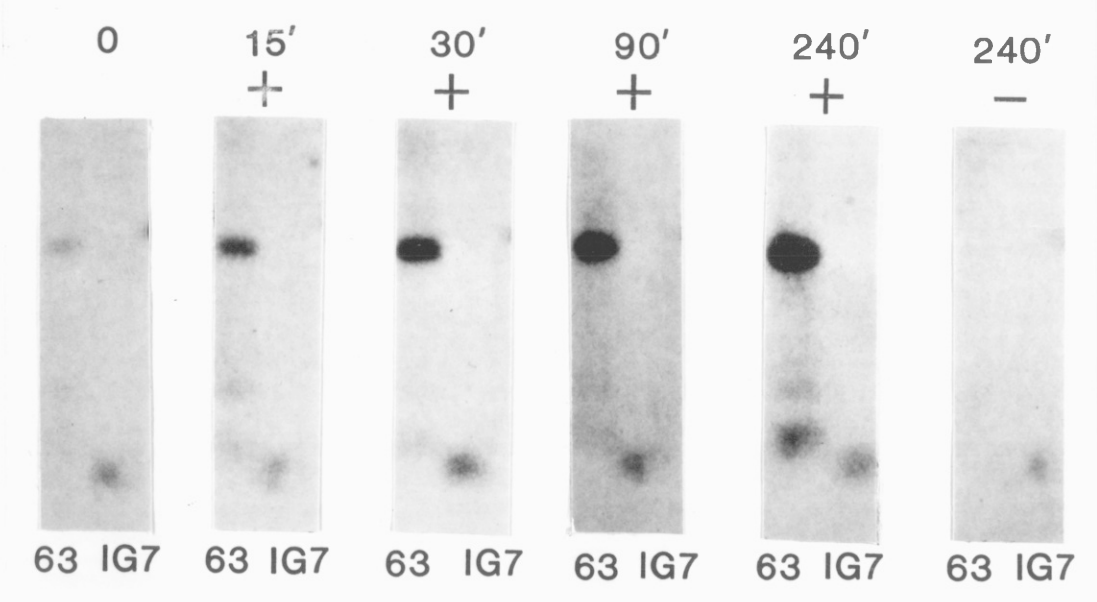
4.5 The repression of D19 by DIF at the level of gene transcription.

The experiments described above were also used to examine the effect of DIF on the transcription of D19 (Figure 4.6). Labelled RNA was used to probe Southern blots of the D19 and IG7 cDNA clones. Again, the insert fragments were separated from their respective vectors by digestion and electrophoresis through a 0.8% agarose gel. The timepoints used were chosen to span the period when the greatest reduction in D19 mRNA levels is seen. For each experiment, a Northern blot of RNA isolated from the cytoplasmic fractions generated during nuclei preparation was probed with a mixture

Figure 4.5 The induction of pDd63 and pDd56 gene transcription by DIF

Transcriptionally active nuclei were isolated from HM44 cells in monolayer culture at the times shown after the addition of DIF. Nuclear run-on transcription was performed by adding ^{32}P-UTP, among other components, to 200µl aliquots of nuclei at a concentration of 2x 10^8/ml. After 30 minutes' incubation, the extended labelled transcripts were purified and hybridised to Southern blots of digested plasmid DNA. The electrophoretic separation of insert and vector fragments enables non-specific hybridisation to be eliminated, and the vector band acts as a control for background hybridisation. The pDd63 cDNA clone was digested with PvuII and ClaI, yielding an internal fragment of 1.35kb. pDd2052, a genomic clone of 56, was digested with PvuII and ClaI to give an internal fragment of 2.67kb. IG7, a cDNA clone expressed constitutively during development, was included to allow comparison of signal intensity between timepoints. In all cases, 2µg of DNA was loaded per track. Hybridisation was continued for 48 hours at 37°C in buffer containing 50% formamide, and the filters were washed for 4x 15' at 37°C in 50% formamide, 0.5x SSC, 0.1% SDS.


0′
90′
+
240′
+
2·67kb
1·35kb
56
63
IG7
56
63
IG7
56
63
IG7


0
15′
+
30′
+
90′
+
240′
+
240′
−
63 IG7
63 IG7
63 IG7
63 IG7
63 IG7
63 IG7

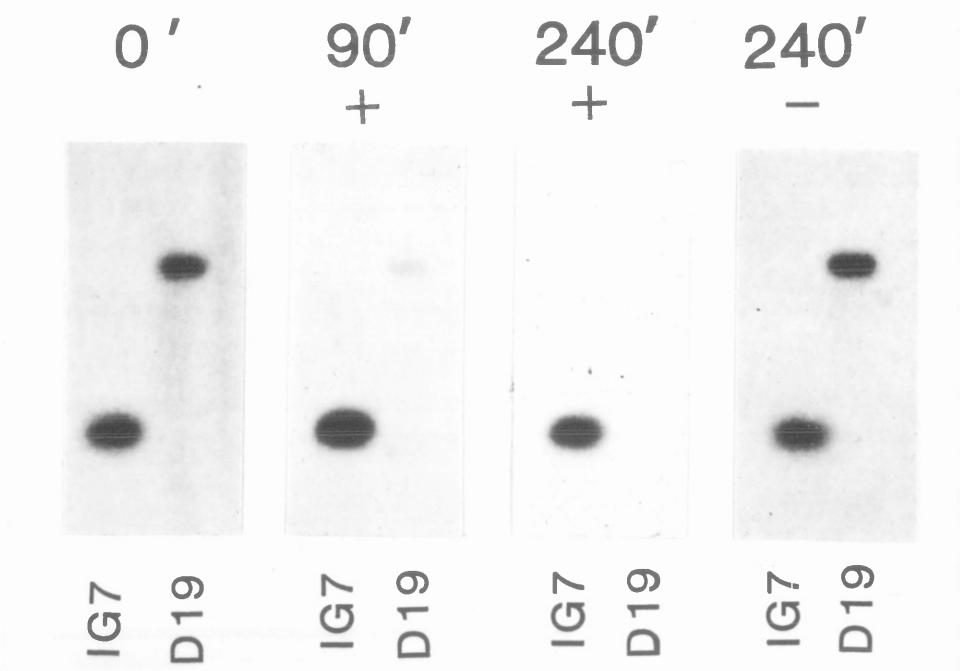
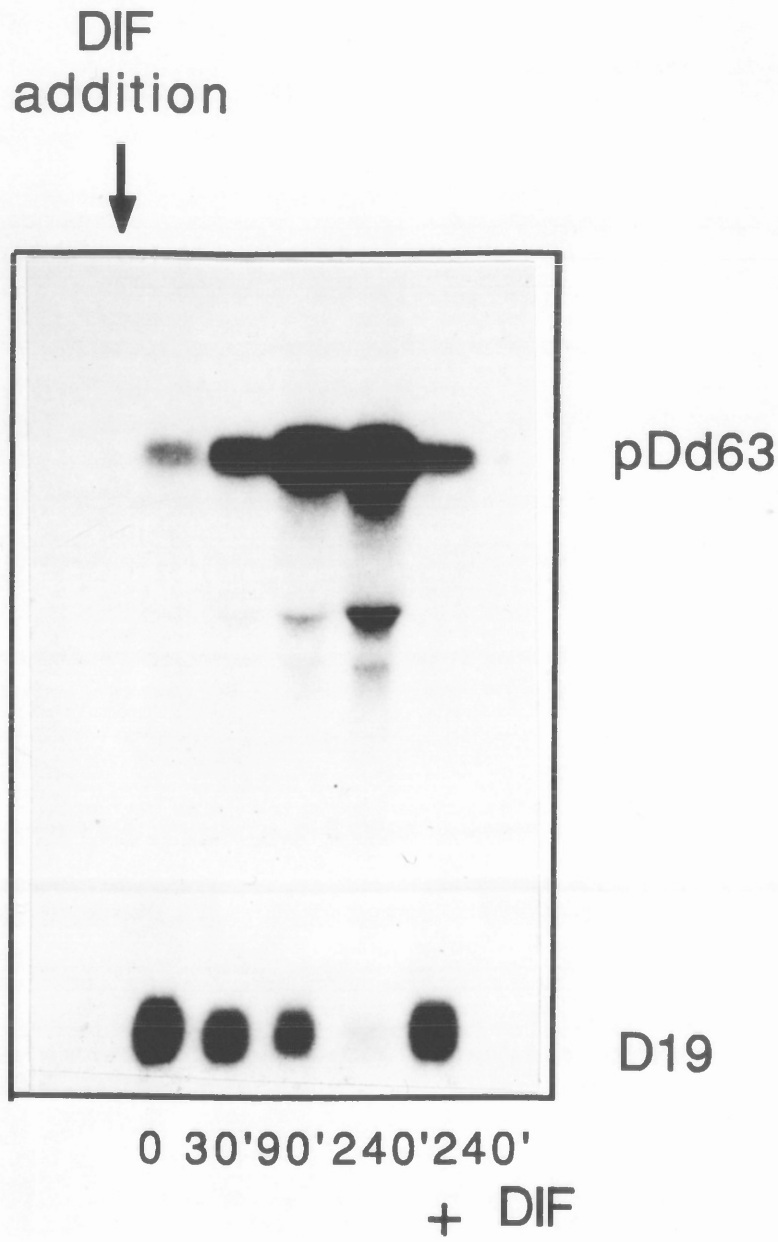


Figure 4.6 Determination of the rate of D19 gene transcription during stalk cell induction

Transcriptionally active nuclei were prepared from HM44 cells incubated in monolayer culture in the presence or absence of DIF for the times shown, following a 10 hour preincubation with cAMP alone. Nuclear run-on assays were performed as described previously, and the resulting ^{32}P-labelled RNA was used to probe Southern blots of the IG7 and D19 cDNA clones. IG7, a cDNA cloned into the SmaI site of pUC8, was digested with EcoRI and HindIII, giving a 500bp insert fragment. D19, cloned into the BamHI site of pBR322, was digested with HindIII and SalI, giving a 1.3kb insert fragment. In both cases, 2µg of DNA was loaded per track. The hybridisation was performed as described previously, except that the filters were washed with 50% formamide, 1x SSC for 4x 15' at 37°C.



4.7 Northern transfer analysis of cytoplasmic RNA extracted in parallel with nuclei preparations

Cytoplasmic nucleic acid was isolated from the supernatant generated during nuclei preparation. 5µg samples, corresponding to the times shown after the addition of DIF, were run on a 1% formaldehyde gel and the resultant Northern blot was probed with a mixture of the pDd63 and D19 cDNA clones. The filter was washed with 2x SSC, 0.1% SDS for 4x 15' at 65°C.

of the pDd63 and D19 cDNA clones. This was to confirm that DIF induction of 63 and repression of D19 had been efficient, although an initial indication was provided by scoring the level of stalk cell induction of a sample plate, 24 hours after DIF addition. In all experiments the rate of stalk cell formation was at least 90%.

A Northern blot of cytoplasmic RNA from the same experiment (Figure 4.7) was probed with the D19 and pDd63 cDNA clones. Hence a direct comparison of nuclear and cytoplasmic events is possible. Within 90 minutes of DIF addition, the signal obtained for D19 in the run-on transcription assay decreased approximately ten-fold. After 4 hours' incubation with DIF there was a further decrease in D19 transcription such that only a faint signal was seen after a long exposure of the autoradiogram. In contrast, where DIF was not included in the medium, the signal intensity remained constant over the four hour period. Comparison with the rate of decline in cytoplasmic mRNA levels suggests that a major part of the repression of D19 can be explained by a reduction in nuclear transcription. The speed with which DIF affects D19 transcription, unlike the induction of 56 and 63, cannot be accurately judged as it is intrinsically difficult to determine the precise onset of repression when an mRNA has any significant degree of stability.

4.6 The effect of DIF addition on the stability of D19 mRNA.

The half-life of all individual mRNAs analysed in developing *Dictyostelium* aggregates has been estimated to be between three and four hours (Mangiarotti et al., 1982). If the D19 mRNA turns over with this half-life in HM44 cells, a 50% reduction in mRNA levels 90 minutes after DIF addition would not be possible even if transcription ceased immediately. In addition, as the D19, D7 and PL3 genes are known to be regulated at a transcriptional and post-transcriptional level by cAMP (Landfear et al., 1982; Barklis and Lodish, 1983), experiments were performed to determine whether or not the repression by DIF also has a post-transcriptional component. These experiments utilised the RNA synthesis inhibitor nogalamycin, which has been successfully used to determine mRNA half-lives in germinating spores (Ennis, 1981; Kelly et al., 1985). A stalk cell induction was again performed using HM44 cells. After 10 hours' incubation with cAMP, the medium was changed and supplemented with nogalamycin at 300µg/ml. Five minutes later, DIF was added to half of the plates and cells were harvested at various times after drug addition, up to a total of four hours. Cells were also harvested just prior to drug addition, and a terminal timepoint from cells incubated with DIF but without nogalamycin was included to check the extent of pDd63 induction.

A Northern blot of total RNA extracted from the cells was probed with a mixture of the pDd63 and D19 cDNA clones (Figure 4.8). The low level of 63 mRNA synthesized in the presence of nogalamycin, compared to that induced in its absence, demonstrates the effectiveness of the drug. At four

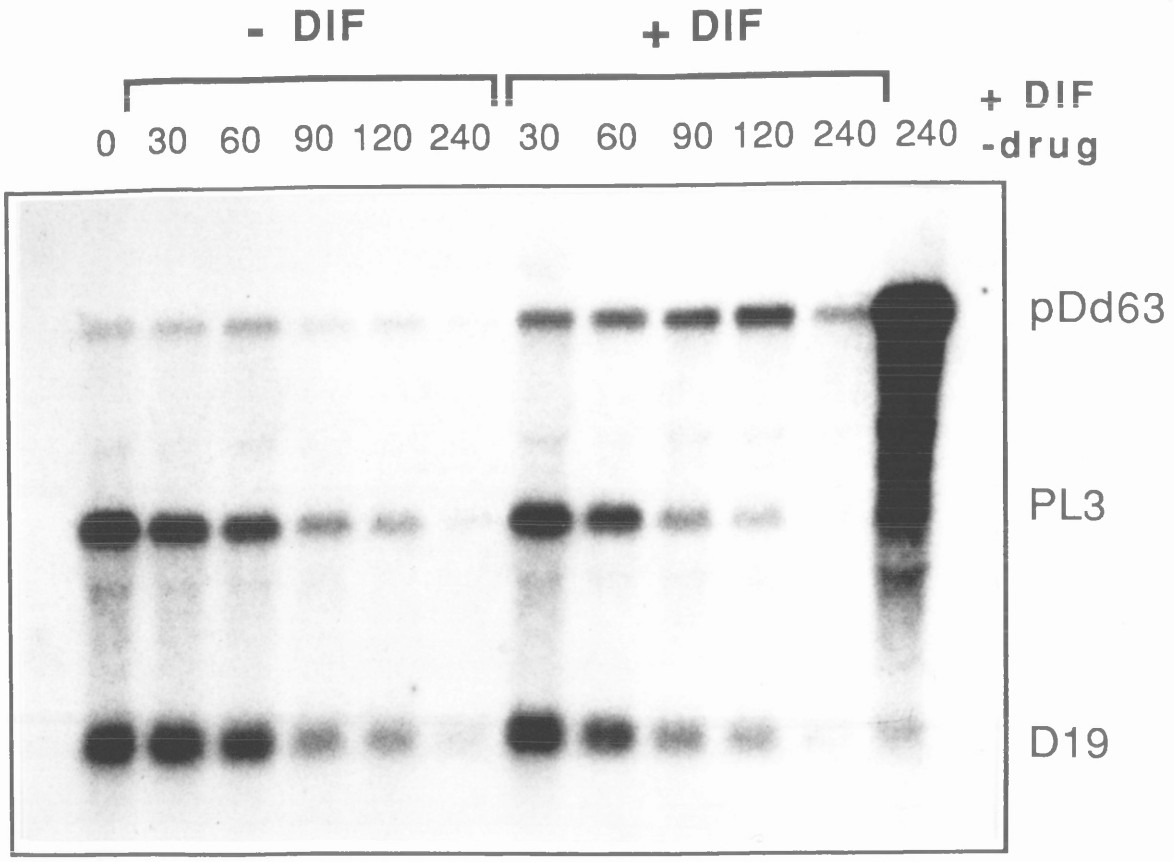


Figure 4.8 The half-life of D19 and PL3 mRNA in the presence and absence of DIF

HM44 cells were incubated in medium containing 300µg/ml nogalamycin (Ennis,after a 10 hour preincubation in the presence of cAMP. DIF was added to half the plates after 5 minutes' exposure to the drug. Timepoints were taken at the times shown after drug addition. A timepoint harvested 4 hours after DIF addition, in the absence of nogalamycin, was also included. Total nucleic acid was extracted and analysed by Northern transfer as previously described. The blot was probed with a mixture of the pDd63 and D19 cDNA clones, followed by the PL3 cDNA clone. The filter was washed with 2x SSC, 0.1% SDS for 4x 15' at 65°C.

hours with DIF alone, D19 mRNA levels were much reduced compared to the zero timepoint. In the presence of the drug, the rate of decay of the D19 mRNA was not altered by the addition of DIF. Reprobing the filter with PL3, another prespore clone (see section 4.2), gave the same result; the half-life of the mRNA was unchanged (Figure 4.8). In an experiment where DIF was added 30 minutes prior to nogalamycin, the subsequent rate of decay of the mRNA was not significantly altered (data not shown). Hence, it is unlikely that nogalamycin interferes with DIF action by preventing the synthesis of a DIF-inducible component.

The half-life of the D19 mRNA is about 90 minutes in both the presence and absence of DIF. This value, taken in conjunction with the result of the nuclear run-on assays, is consistent with the effect of DIF being mediated primarily at the transcriptional level.

4.7 DIF repression of D19 at low cAMP concentrations.

DIF has been proposed to act by interfering with the cAMP-mediated induction of prespore gene expression. In aggregation competent NC4 cells, incubated in monolayer culture, DIF and cAMP seem to act antagonistically (Wang et al., 1986). DIF represses cAMP-receptor interactions at low cAMP concentrations (0.1mM) and reduces the cAMP relay response. However, the experiments described here were performed using a saturating, 5mM concentration, where DIF does not prevent cAMP binding to the receptor in NC4 cells.

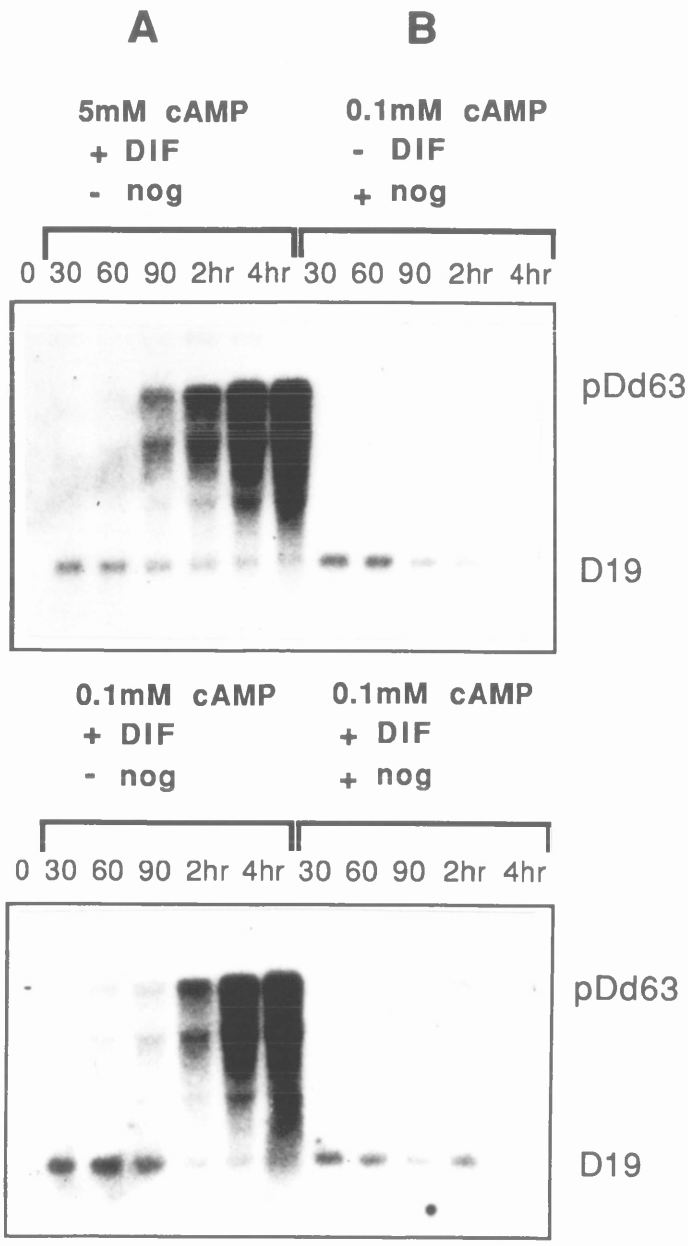


Figure 4.9 Repression of D19 mRNA by DIF at low cAMP concentration

All cells were first incubated for 10 hours in 5mM cAMP before substituting buffer containing the components listed above. Timepoints were harvested at the times shown after the buffer change, and total RNA was extracted and analysed by Northern transfer as previously described. The blot was probed with a mixture of the D19 and pDd63 cDNA clones, and washed with 2x SSC, 0.1% SDS for 4x 15, at 65°C.

In order to investigate the possible antagonism between the action of DIF and cAMP in regulating prespore gene expression in HM44 cells, stalk cell inductions were performed with reduced levels of cAMP in the DIF-containing medium. The 10 hours of preincubation were carried out as usual with 5mM cAMP. A comparison was made between the rate of DIF-induced decline in D19 mRNA levels with 5mm cAMP and 0.1mM cAMP (Figure 4.9). In addition, the mRNA half-life in the presence and absence of DIF at 0.1mM cAMP was measured using nogalamycin. The efficiency of both DIF induction and nogalamycin action was monitored by probing the blot with pDd63 as well as the D19 cDNA clone. There was no change in the kinetics of decay of D19 message levels as a consequence of reducing the extracellular cAMP concentration, and DIF again does not alter the half-life of D19 mRNA at low cAMP concentrations. Overall, there was no evidence of mutual antagonism between cAMP and DIF.

These results are in apparent contradiction to those reported for NC4 cells (Wang et al., 1986). However, the DIF concentration used for all HM44 experiments (3000U/ml), which consistently resulted in decreased D19 mRNA levels, would have been insufficient to repress prespore gene expression in NC4 cells incubated with 5mM cAMP. At this cAMP concentration, 10,000U DIF/ml had little effect on prespore antigen levels in NC4 cells. Therefore it may be that with the assay conditions used, HM44 cells are more sensitive to DIF than NC4 cells, or expressed conversely, the action of DIF in NC4 cells is more effectively counteracted by cAMP.

This interpretation is supported by recent evidence that efficient stalk cell formation in low density monolayer cultures of NC4 cells requires removal of cAMP (Berks and Kay, 1988). Prespore repression in HM44 cells might show a greater dependence on cAMP levels if lower levels of DIF were used. As the repression of D19 in HM44 cells is apparently cAMP independent at the DIF concentration used, DIF may be counteracting cAMP induction at a post-receptor, intracellular stage (see Chapter 8).

4.8 D19 protein levels during stalk cell induction.

As a specific monoclonal antibody which recognises the D19 protein is available (see section 3.7), it was utilised to determine whether DIF affects the rate of protein as well as RNA disappearance. It has previously been shown that adding DIF to HM44 cells prevents the appearance of four major prespore-specific proteins, detected by metabolic labelling with ^{35}S-methionine (Kopachik et al., 1985), but the effect of DIF on the turnover of specific pre-existing proteins is not known. During a stalk cell induction performed in the standard manner, cells were harvested 1.5, 4, 8, 12, and 24 hours after the addition of DIF. Timepoints were also taken just prior to DIF addition, and from a parallel set of incubations without DIF. Samples of protein from each extract were analysed by one-dimensional SDS-polyacrylamide gel electrophoresis, and protein from V12M2 cells at 12 hours of development was included as a positive control of antibody binding. A Western blot of the gel was

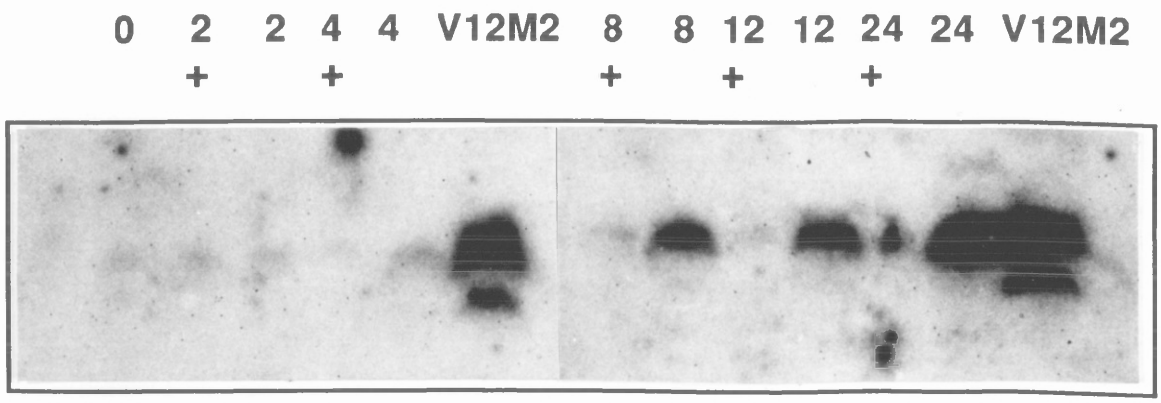


Figure 4.10 The effect of DIF on D19 protein levels in HM44 cells

Total cellular protein was extracted from HM44 cells in monolayer culture at the times shown after addition of 3000U DIF/ml to half the plates, following 10 hours' preincubation with cAMP. A sample from V12M2 cells was included as a positive control. 30µg samples of protein were subjected to one dimensional polyacrylamide gel electrophoresis with a combination of a 4.5% stacking gel and a 10% separating gel. A Western blot of the gel was incubated with MUD-1 at a dilution of 1 in 100, and binding was detected with ^{125}I-protein A.

incubated with MUD-1, and binding was detected using ^{125}I-protein A (Figure 4.10).

The level of D19 protein at the time of DIF addition is very low compared to that detected 24 hours later in the absence of DIF. An increase in signal can be seen after two hours, the major rise occurring between two and four hours after the zero timepoint. However, where DIF was included, the initial level of D19 protein remained constant throughout the remainder of the induction. Although this apparent stability may be due to a continuing low level of protein synthesis, it is clear that DIF does not induce a rapid loss of pre-existing D19 protein. This result emphasizes the primary role of transcriptional mechanisms in mediating DIF repression of the D19 gene.

4.9 Summary

1. DIF represses D19 and two other prespore class II genes, PL3 and D7, at the RNA level. A D19-related sequence induced by DIF during stalk-cell induction is almost certainly derived from an independent gene.

2. DIF induces the prestalk-specific genes 56 and 63 at the transcriptional level.

3. The repression of D19 by DIF, demonstrated by nuclear run-on assays, is also mediated primarily at the transcriptional level, as the stability of D19 mRNA is unaffected by DIF addition.

4. In HM44 cells, at the concentrations of DIF used, antagonism with cAMP cannot be demonstrated for prespore repression.

5. DIF does not cause a reduction in levels of pre-existing D19 protein.

CHAPTER 5: Analysis of D19 gene expression using the transformation system I: Construction of *D.discoideum* transformation vectors and transformation of a marked copy of the D19 gene.

The potential for analysis of gene expression in *Dictyostelium* has recently been expanded by the development of a transformation system. By reintroducing copies of cloned genes into slime moulds, it is possible to study both the function of specific gene products and the DNA sequence elements which control their expression (Reymond et al., 1985; Pears et al., 1987; Knecht and Loomis, 1987; Witke et al., 1987). The first transformation vector to be constructed, B10, is based on pBR322, and contains the neomycin resistance gene from Tn5 fused to the promoter of the actin 6 gene (Nellen et al., 1984). After introduction as a calcium phosphate coprecipitate, the vector confers resistance to the amino-glycoside G418. Selection with G418 produces stable transformants. Using B10 as a starting point, alternative vectors were contructed in this laboratory in order to facilitate manipulation of *Dictyostelium* DNA prior to transformation.

The vector pB10 (J.Williams, personal communication; Figure 5.1) was constructed by insertion of the EcoRI-BamHI neomycin resistance (neo) gene-containing fragment from B10 into EcoRI-BamHI-cleaved pEMBL18+ (Dente et al., 1983). The phage f1 origin of replication remains intact in pB10, and so

the vector can produce single-stranded DNA upon super-infection of bacterial cells containing the vector with the phage. The single-stranded DNA is suitable for sequence analysis by the chain-termination method (Sanger et al., 1977), and can also be used for other purposes, e.g. the production of single-stranded probes for use in S1 mapping and for performing site-directed mutagenesis. The latter method is of importance since *Dictyostelium* DNA, especially in the A-T rich regions flanking genes, has few useful restriction enzyme cleavage sites. In addition, the pEMBL sequence increases the yield of plasmid DNA relative to that given by B10. This is because sequences flanking the plasmid origin of replication have been mutated, preventing the strict control of copy number which occurs in ColE1-based plasmids. This feature is useful for *Dictyostelium* transformation, as comparatively large amounts of DNA are required (see Chapter 2).

5.1 Construction and characterisation of pB10T.

In pB10 the actin 6/neomycin resistance gene fusion lacks a *Dictyostelium* terminator, and a heterodisperse series of transcripts of high molecular weight is produced (Nellen and Firtel, 1985). As readthrough transcription could affect transcription of genes located elsewhere in the vector, a genomic fragment was inserted to provide *Dictyostelium* polyadenylation and termination sequences. In pB10 (Figure 5.1), the 3' region of the cysteine proteinase 1 gene, located on a 1.4kb Sau3A fragment derived from the genomic clone pDG10 (Pears et al., 1985), has been inserted adjacent

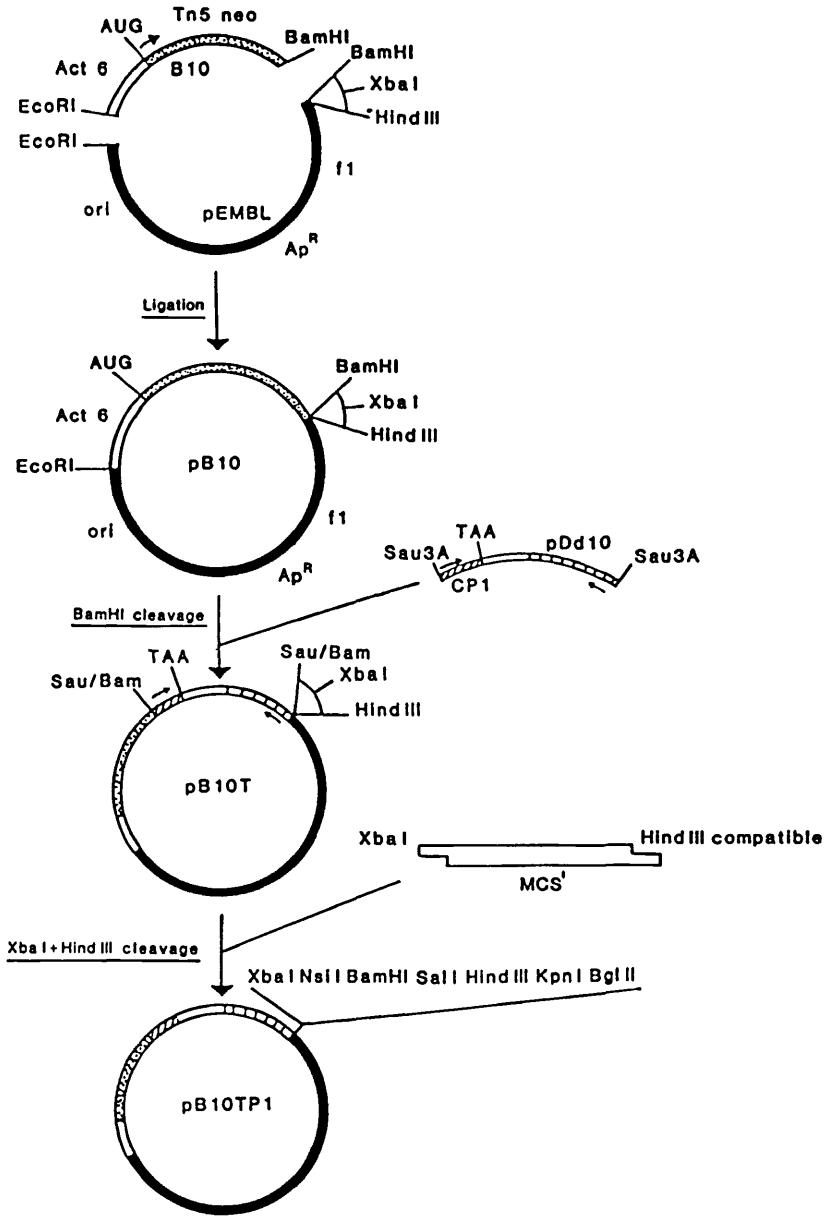


Figure 5.1 Scheme of construction of pB10TP1

Arrows are used to indicate the direction of transcription of the neomycin resistance gene, the neomycin/Tn5 fusion and Dd10. The synthetic oligonucleotide inserted to provide a novel multilinker sequence is denoted by MCS (multiple cloning site).

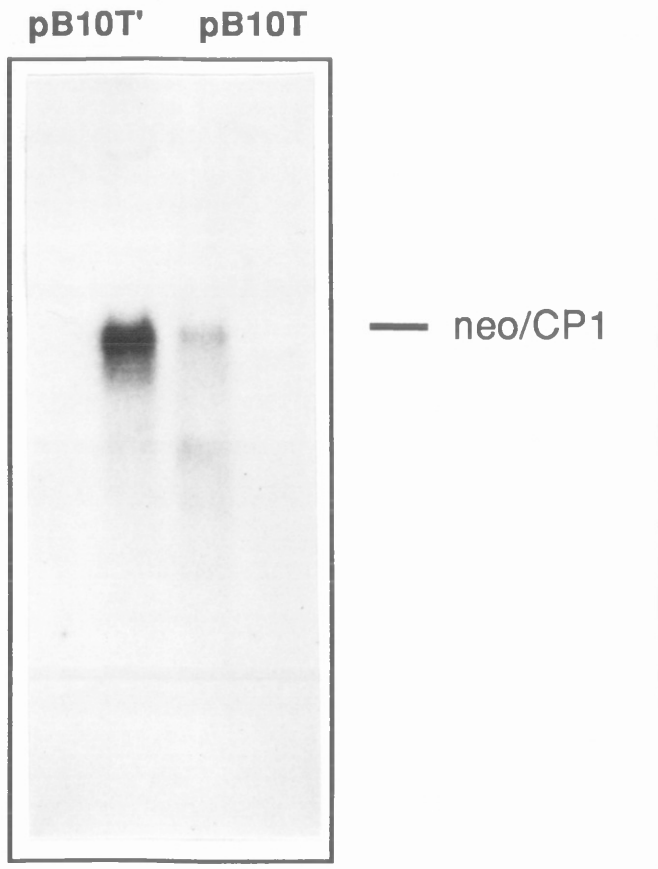


Figure 5.2 Termination of neomycin resistance gene transcription by CP1 sequences

Total RNA was extracted with phenol from growing cells transformed either with pB10T or pB10T', and 5µg samples were analysed by Northern transfer. The blot was probed with pB10, ^{32}-P labelled by nick-translation, which hybridises to the neomycin resistance gene transcript. The filter was washed for 4x 15' with 2x SSC, 0.1% SDS at 65°C.

to Tn5 neo after cleavage at the pB10 BamHI site and ligation of the complementary termini. A fusion gene has therefore been created; the neomycin resistance gene plus 268 nucleotides of CP1, spanning the last intron of the gene. The unknown nucleotide sequence downstream of CP1 was established by the chemical cleavage method (Maxam and Gilbert, 1980). By labelling at the AccI site at position 6351, the PvuII site at 6422 (using T4 polymerase) and the XbaI site in the multilinker (see Figure 5.4a), it was possible to obtain 740 nucleotides of contiguous sequence, which overlapped previously established sequence downstream of CP1 (J.Williams, personal communication). This sequence analysis revealed the presence of a convergently transcribed gene, termed Dd10. The termination codons of the two genes are separated by 271 nucleotides. Assuming there to be no overlapping transcription, functional termination signals for both genes must be contained in this region.

The 5' proximal sequences from CP1 in pB10T function efficiently to produce a discrete fusion transcript. Figure 5.2 shows a Northern blot of RNA extracted from growing cells, transformed with either pB10T or a related vector containing a shorter fragment from pDG10 (see section 5.2). The blot was probed with pB10, which hybridises to the discrete neo/CP1 fusion transcript. The size of this transcript indicates that 3' end processing occurs approximately 150 nucleotides downstream of the CP1 termination codon.

In pB10T there are 650 nucleotides of coding sequence derived from the Dd10 gene. The primary amino acid sequence

of this region has no obvious homologue in the currently available databases. A fragment from Dd10, used as a probe in Northern transfer analysis, hybridises to a constitutively expressed mRNA species 1.1 kb in length (D.Driscoll, personal communication; data not shown). Fusion of the 3' region of Dd10 to the 5' proximal regions of other *Dictyostelium* genes inserted in the multilinker is a potential means of analysing the function of transcriptional control elements. Fusing a 5' fragment from the 56 gene to Dd10 in this way resulted in a discrete transcript of the length expected for an mRNA terminating between CP1 and Dd10 (K.Jermyn, personal communication). However, the same sequences from Dd10 did not function effectively to terminate the CP2 gene (C.Pears, Ph.D. thesis). Therefore termination may require interaction with more proximal sequences or recognition of some feature of overall secondary structure.

5.2 Construction of the pB10TP vector series by multilinker replacement.

To increase the usefulness of pB10T, the multilinker of pEMBL18+ was replaced with a new synthetic oligomer (Figure 5.3). The reading frame of the Dd10 coding region with respect to sites in the multilinker is shown, since it is preferable to fuse two coding regions in frame in order to prevent premature cessation of translation and hence an increased susceptibility of the mRNA to degradation. The multilinker was inserted in two stages; an XbaI-HindIII linker was first cloned using the corresponding original multilinker sites, followed by addition of a HindIII-KpnI-

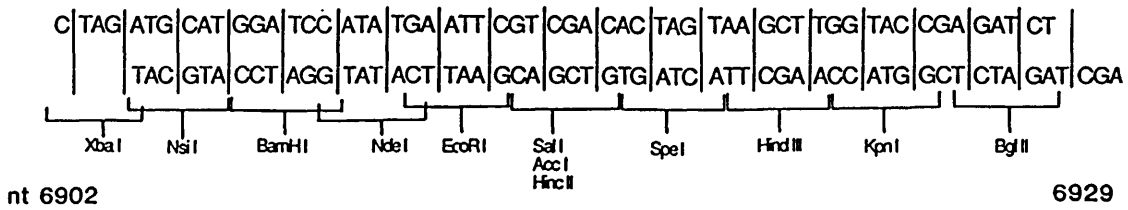


Figure 5.3 Nucleotide sequence of the multilinker of pB10TP1 and pB10TP2

The vertical bars indicate the frame of the multilinker relative to the coding region of Dd10.

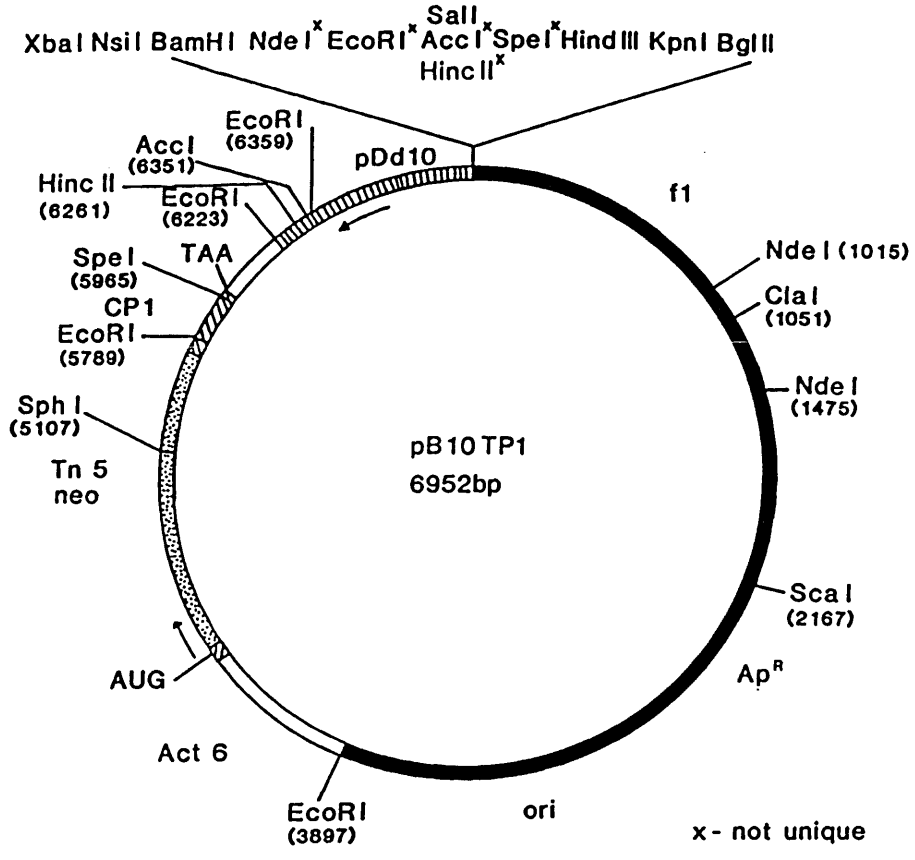
BglII linker into the HindIII site. As the second linker could be inserted in either orientation, two versions of the multilinker resulted in addition to the shorter intermediate. The new vectors were designated pB10TP1, illustrated in Figures 5.3 and 5.4a, and pB10TP3, in which the order of the HindIII and BglII sites is reversed.

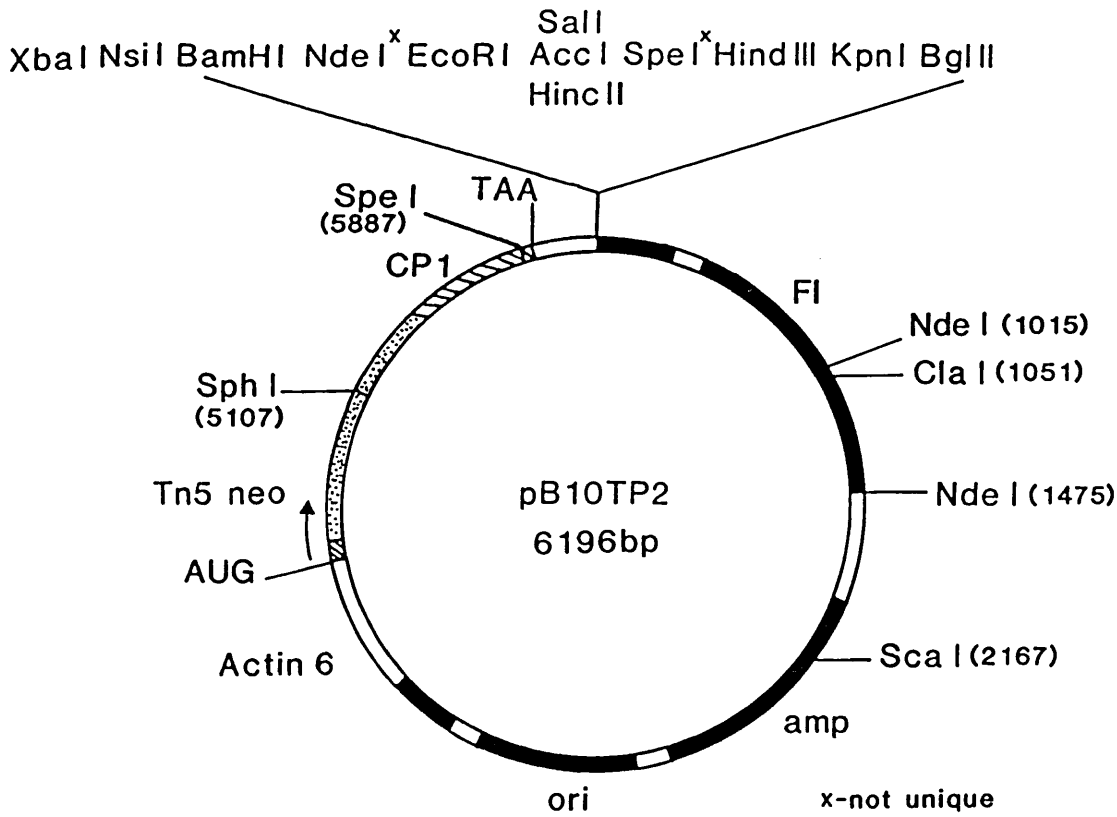
The new multilinker contains sites suitable for initial cloning, sequence analysis and mutation prior to transformation. It contains several sites with a high A-T content (NsiI, NdeI and SpeI) which facilitate subcloning as they occur frequently in the *Dictyostelium* genome. The NsiI recognition sequence is unique in the vector and is doubly useful becuase it generates a PstI-compatible cohesive end. The NdeI and SpeI sites are not unique; there are two closely-spaced NdeI sites in the pEMBL sequence and SpeI is present in the CP1 coding region. Cleavage with SpeI deletes the termination signals and cleavage with NdeI eliminates the ability to produce single-stranded DNA, but in both cases transformation is still possible. The HincII site is also not unique but digestion with this enzyme only removes a portion of Dd10 and it is therefore a potentially useful site. Also within the linker is a unique BamHI site and a unique BglII. Either of these could be used to clone Sau3A partial digests in order to generate a *Dictyostelium* genomic clone bank for use in complementing specific mutants.

As exonuclease III is unable to degrade a 3' protruding terminus, the KpnI site was included in the linker to protect pEMBL sequences from exonucleolytic digestion when creating a deletion series from a cloned fragment (Henikoff, 1984). The

Figure 5.4 Restriction maps of the pB10TP vectors

The position of the sites shown were precisely deduced from the total nucleotide sequence of the two vectors. This was in part newly determined and in part derived from other studies (Beck et al., 1982; Dente et al., 1983; Romans and Firtel, 1985). The sequence of previously uncharacterised DNA from Dd10 was obtained using the method of Maxam and Gilbert (1980) (see text for details). Restriction sites marked with the x superscript are not unique.


Xba I Nsi I BamH I Nde I x EcoR I x Sal I Acc I x Spe I x Hind III Kpn I Bgl II
Hinc II x
EcoR I (6359)
Acc I (6351)
Hinc II (6261)
EcoR I (6223)
pDd10
f1
Spe I (5965)
TAA
CP1
EcoR I (5789)
Sph I (5107)
Tn 5 neo
pB10TP1
6952bp
Nde I (1015)
Cla I (1051)
Nde I (1475)
Sca I (2167)
Ap R
AUG
Act 6
EcoR I (3897)
ori
x- not unique


Xba I Nsi I BamH I Nde I x EcoR I Sal I Acc I Spe I x Hind III Kpn I Bgl II
Hinc II
Spe I (5887)
TAA
CP1
Sph I (5107)
Tn5 neo
AUG
Actin 6
pB10TP2
6196bp
FI
Nde I (1015)
Cla I (1051)
Nde I (1475)
Sca I (2167)
amp
ori
x-not unique

presence of a site downstream from KpnI allows sequencing of the deletion derivatives by a choice of strategy. As the KpnI site is close to the binding site for the standard M13 sequencing primer, single-stranded DNA can be sequenced by the chain termination method (Sanger et al., 1977). Alternatively, by labelling at the terminal multilinker site, sequence can be determined by the chemical cleavage method (Maxam and Gilbert, 1980). This is sometimes important as it is difficult to sequence very A-T rich DNA using the dideoxy procedure.

pB10TP1 is 6962 base-pairs in length and has been successfully used to clone *Dictyostelium* DNA up to 4.75 kb with no apparent effect on the frequency of recombinants or the yield of plasmid DNA. However, a related vector was constructed to maximize the number of unique restriction sites within the multilinker (J.Williams, personal communication). By destroying the EcoRI site in pB10 by cleavage and end-filling, an EcoRI fragment containing the 3' proximal region of CP1, again isolated from pDG10, could be inserted. The 520 nucleotide fragment was end-filled and cloned into the filled BamHI site of the vector to give pB10T'. The distal terminus of the fragment lies 22 nucleotides downstream of the Dd10 gene, but a discrete neo/CP1 transcript still results (Figure 5.2). The multilinker was replaced as before, giving pB10TP2 (HindIII KpnI BglII, Figure 5.4b) and TP4 (BglII KpnI HindIII). Due to the absence of Dd10 sequences and the method of construction, EcoRI, AccI and HincII are additionally unique, offering a greater flexibility in subcloning, but there is no direct

means of "marking" cloned promoters. As pB10TP4 was used in analysis of the expression of the D19 genomic clone, an alternative method of marking the gene was required.

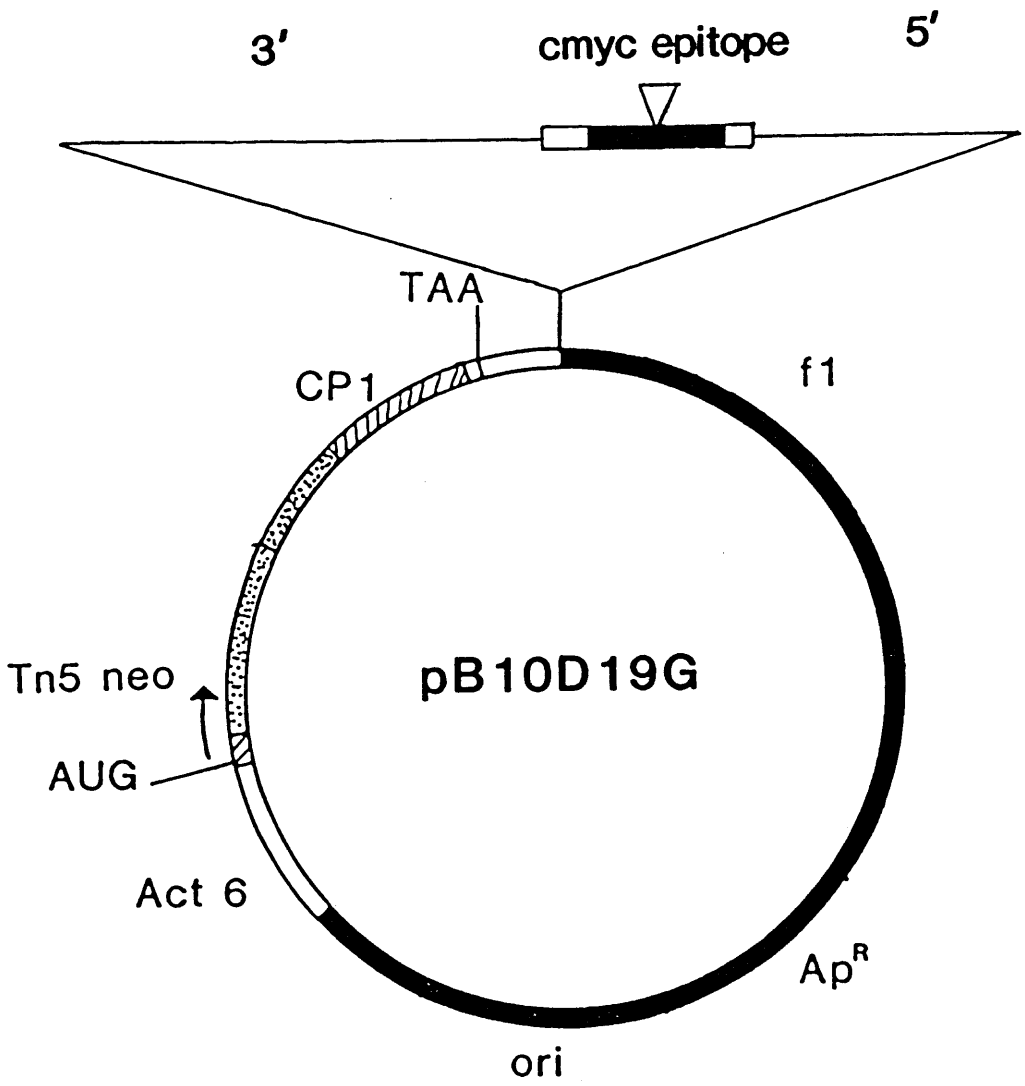
5.3 Marking of the D19 gene with an epitope of the human c-myc protein.

The PvuII site within the coding region of D19, unique in the original pAT153-based genomic clone, was utilised to clone a 33 nucleotide-long sequence encoding an epitope of the human c-myc protein. This method of tagging proteins was first used to detect and localise an hsp70-like protein in rat liver cells (Munro and Pelham, 1986). The epitope Glu-Gln-Lys-Leu-Leu-Ser-Glu-Glu-Asp-Leu-Asn is specifically recognised by the anti-myc monoclonal 9E10 (Evan et al., 1985). Also, transcripts from the transformed gene could be distinguished from those of the endogenous gene using the non-coding strand of the oligonucleotide as a probe.

The nucleotide sequence synthesized was designed firstly to incorporate preferred *Dictyostelium* codon usage, and secondly to create a BglII site at the 3' end, while regenerating the PvuII site at the other. The two strands of the synthetic oligonucleotide were annealed before ligation to the linearized vector; a 50-fold molar excess of non-phosphorylated oligonucleotide over phosphorylated vector ensured efficient insertion of a single copy of the epitope. The epitope was inserted in frame within codon 81. Owing to an error in the oligonucleotide sequence synthesized, glutamate was substituted for glutamine in the encoded peptide.

The BglII site was utilised for sequencing by the chemical cleavage method (Maxam and Gilbert, 1980) to check the orientation of the epitope. The correctly marked construct was designated B1. The c-myc-marked gene was excised from pAT153 by cleaving at the EcoRI site in the vector and the BamHI site at the 3' boundary of the insert, and cloned into the corresponding multilinker sites of pB10TP4 (Figure 5.5). The resulting construct, pB10B1, contains the entire D19 coding region with 995 nucleotides of sequence upstream of the cap site and about 1.9kb of downstream sequence.

The accessibility of the pB10B1 c-myc epitope for antibody binding was unknown as the oligonucleotide is positioned close to the centre of the coding region. An alternative version of the marked gene was constructed in parallel to increase the chances of detecting D19 expression at both the RNA and protein level. An oligonucleotide with NdeI-compatible termini (Figure 5.5) was inserted between the two closely spaced NdeI sites at the N-terminus of the D19 coding region, again while in pAT153. The epitope itself was unaltered. The gene was subcloned into pB10TP4 as before to give pB10N6. In pB10N6, the epitope is theoretically more accessible as it positioned close to one end of the protein, but the cloning site is within the putative signal sequence just 5' to the cleavage point. The position of cleavage was recently confirmed by comparison with peptide sequence of the mature protein (see Chapter 3). If processing has not been disrupted by the additional sequences, the epitope will be


3'
cmyc epitope
5'
TAA
CP1
f1
Tn5 neo
pB10D19G
AUG
Act 6
Ap^R
ori

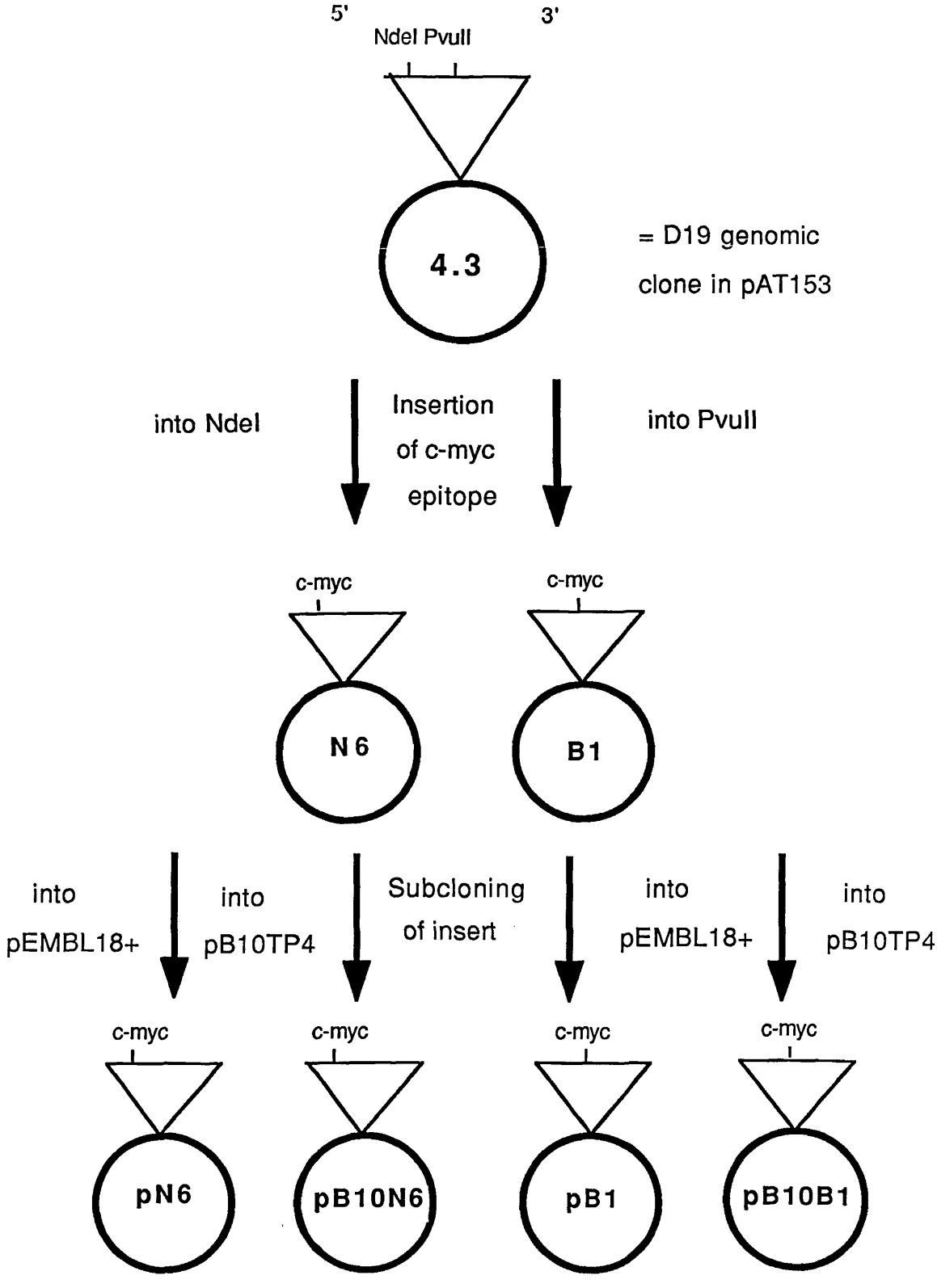


Figure 5.5 Construction of vectors containing alternative versions of the marked D19 gene

Facing page, more detailed map of pB10B1 (= pB10D19G).

removed. No other suitable unique sites for insertion of the epitope exist within the D19 gene.

5.4 Expression of the marked gene after DNA-mediated transformation of AX-2 cells.

AX-2 cells were transformed with pB10B1 and pB10N6 using a protocol involving calcium phosphate coprecipitation and glycerol shock (see Chapter 2). Stable transformants were selected with 10µg/ml G418, and expression of the gene was analysed in pooled populations of at least 100 clones for each construct. Southern blotting analysis indicated that the vector copy number within the population was on average approximately 10 per cell for pB10B1, and 5-10 for pB10N6 (Figure 5.6).

The expression of the marked D19 gene was analysed by Northern transfer. Total nucleic acid was extracted from cells harvested during development on filters. A development of AX-2 cells was set up in parallel to that of pB10B1-transformed cells. Timepoints were taken at defined morphological stages, and growing cells (0 timepoint) were harvested directly from shaken suspension culture.

Expression of the transformed gene was detected using the non-coding strand of the oligonucleotide, ^{32}P-labelled with T4 polynucleotide kinase, as a probe. In pB10B1-transformed cells, marked D19 mRNA accumulated with the kinetics expected for the transcript of the endogenous D19 gene (Barklis and Lodish, 1983; Oyama and Blumberg, 1986a; Figure 5.7a). The mRNA was first faintly detectable at the tipped mound stage, and the level then sharply increased to

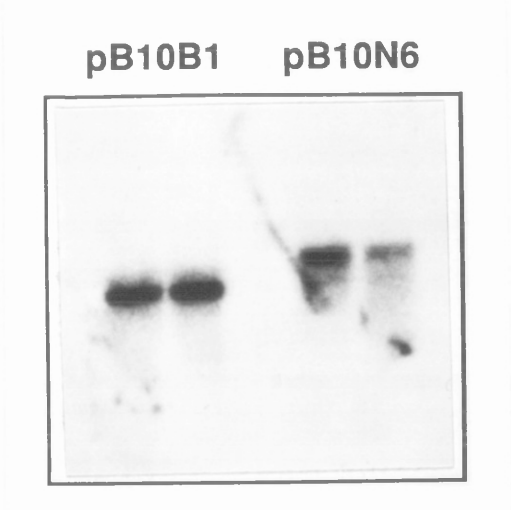
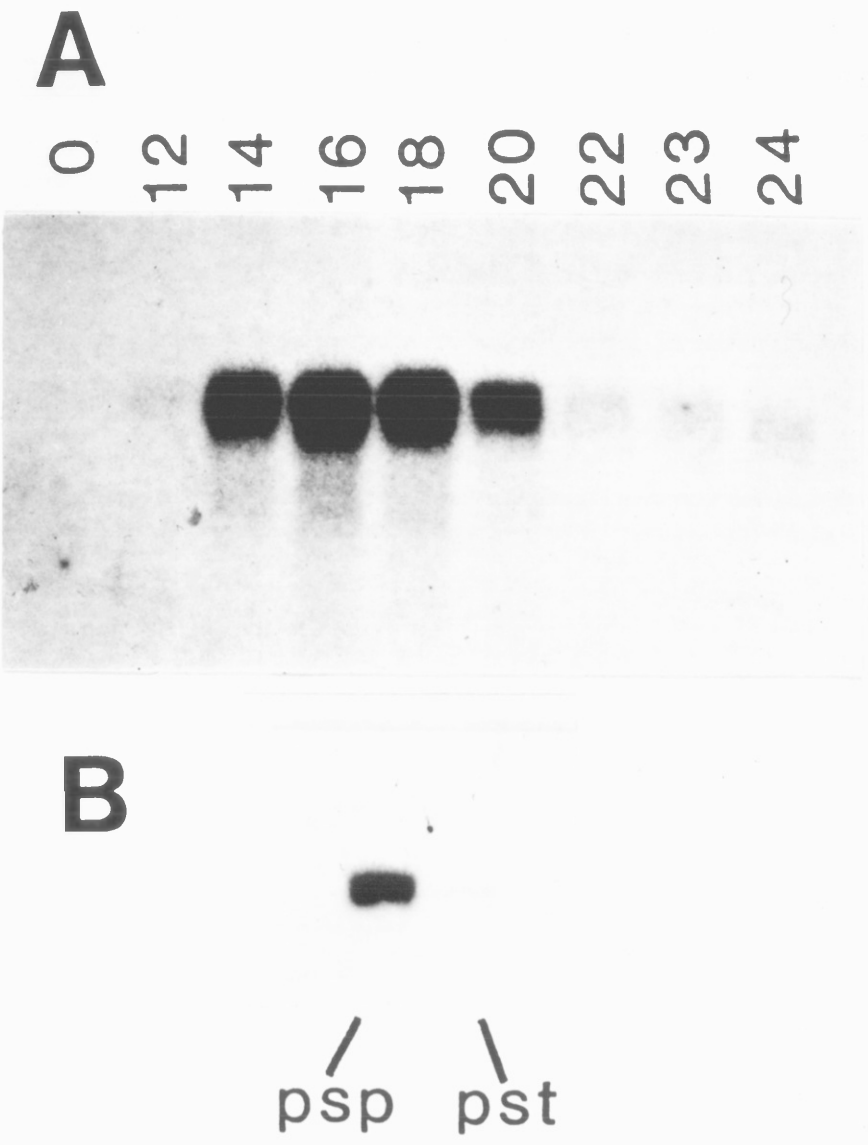


Figure 5.6 Southern transfer analysis of cells transformed with pB10B1 and pB10N6

Total nucleic acid was phenol-extracted from growing cells, and 50µg samples were RNAase-treated and run on a 0.8% agarose gel. After denaturation of the DNA and transfer to nitrocellulose, the filter was probed with ^{32}P labelled pB10TP2. Known masses of pB10TP2 were also included to allow quantitation of the hybridisation signal subsequently obtained (7.5ng, 2.5ng and 0.5ng). The filter was washed for 4x 30' with 2x SSC, 0.1% SDS at 65°C. The copy number for each transformant was calculated with the knowledge that pB10B1 and N6 are about 9kb in size, and using the following assumptions; 50µg of total nucleic acid contains 1µg of DNA and the size of the *Dictyostelium* genome is 50,000kb. Knowing the fraction of the DNA which corresponds to the vector, the proportionate mass per cell and hence the copy number can be calculated.

Figure 5.7 Temporal and spatial patterns of expression of the marked gene during development

Pooled populations of AX-2 cells transformed with pB10B1 were analysed. Total cellular RNA was extracted with phenol and 5µg samples run on a 1% formaldehyde gel. After transfer to nitrocellulose, the blot was probed with the non-coding strand of the c-myc oligonucleotide, end-labelled with ^{32}-P using T4 polynucleotide kinase, which specifically detects mRNA from the transformed gene. Hybridisation was performed at 32°C in buffer consisting of 6x SSC, 0.1% SDS , and for the washes (5x SSC, 0.1% SDS, 4x 15') the temperature was increased to 42°C.
(A) Times shown indicate the intervals after initiation of development on filters. The corresponding morphological stages are as follows: 12hrs, loose mounds; 14 hours, tipped mounds; 16 hours, first finger; 18hrs, standing slug; 20hrs, late preculminates; and 24hrs, culminates.
(B) Disaggregated slug cells were separated by two rounds of Percoll density gradient centrifugation (Ratner and Borth, 1983), and RNA was isolated from separated prespore and prestalk cells as before.


A
0
12
14
16
18
20
22
23
24
B
psp
pst

reach a peak of abundance at the first finger and standing slug stages. The concentration fell in preculminants, and a low level was maintained during culmination. The nick-translated D19 cDNA clone detects the product of both the transformed and endogenous genes as a single band. Using this probe, no changes in the accumulation pattern could be discerned with the same RNA samples (Figure 5.8). That the accumulation of endogenous D19 mRNA has the same pattern is also demonstrated by timepoints harvested from AX-2 cells at equivalent stages of development (Figure 5.8). It is clear that the relative level of D19 mRNA in transformed cells has been substantially increased by transcription of the marked gene. There is approximately five times as much mRNA in the transformed cells compared with an equivalent number of AX-2 cells. The expression level of the endogenous and introduced genes must be comparable since the copy number of the transformed gene is only approximately ten-fold greater.

The D19 mRNA in cells transformed with pB10B1 is correctly spatially as well as temporally regulated. Slugs were disaggregated and the cells separated by two rounds of Percoll density gradient centrifugation (Ratner and Borth, 1983). Prespore and prestalk cells can be separated because they are of differing densities (Tsang and Bradbury, 1981). mRNA isolated from the light (prestalk) and heavy (prespore) fractions was again subjected to Northern transfer analysis using the c-myc oligonucleotide probe (Figure 5.7b). There is strong enrichment of the c-myc marked mRNA in the prespore fraction, estimated to be 10-fold by scanning densitometry. As the prestalk fraction is invariably contaminated to some

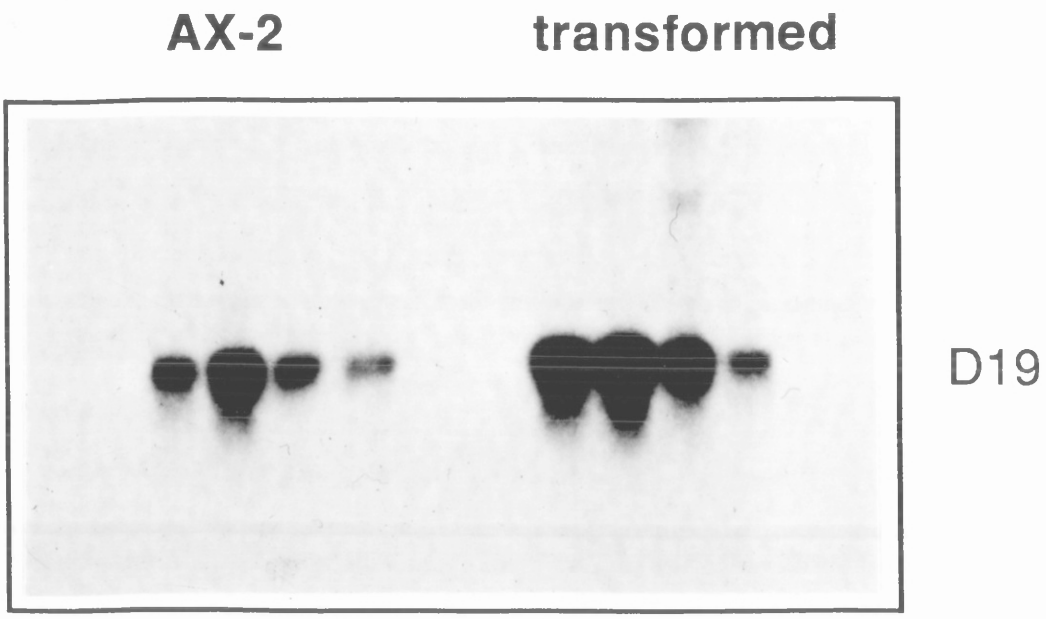


Figure 5.8 Expression of the D19 gene in wild-type and transformed cells

Cells were harvested at equivalent morphological stages during a filter development of either AX-2 cells or cells transformed with pB10B1. Total cellular RNA was extracted with phenol and 5µg samples analysed by Northern transfer as previously described, except that the blot was probed with the D19 cDNA clone, labelled with ^{32}P, which detects the product of both the endogenous and transformed gene.

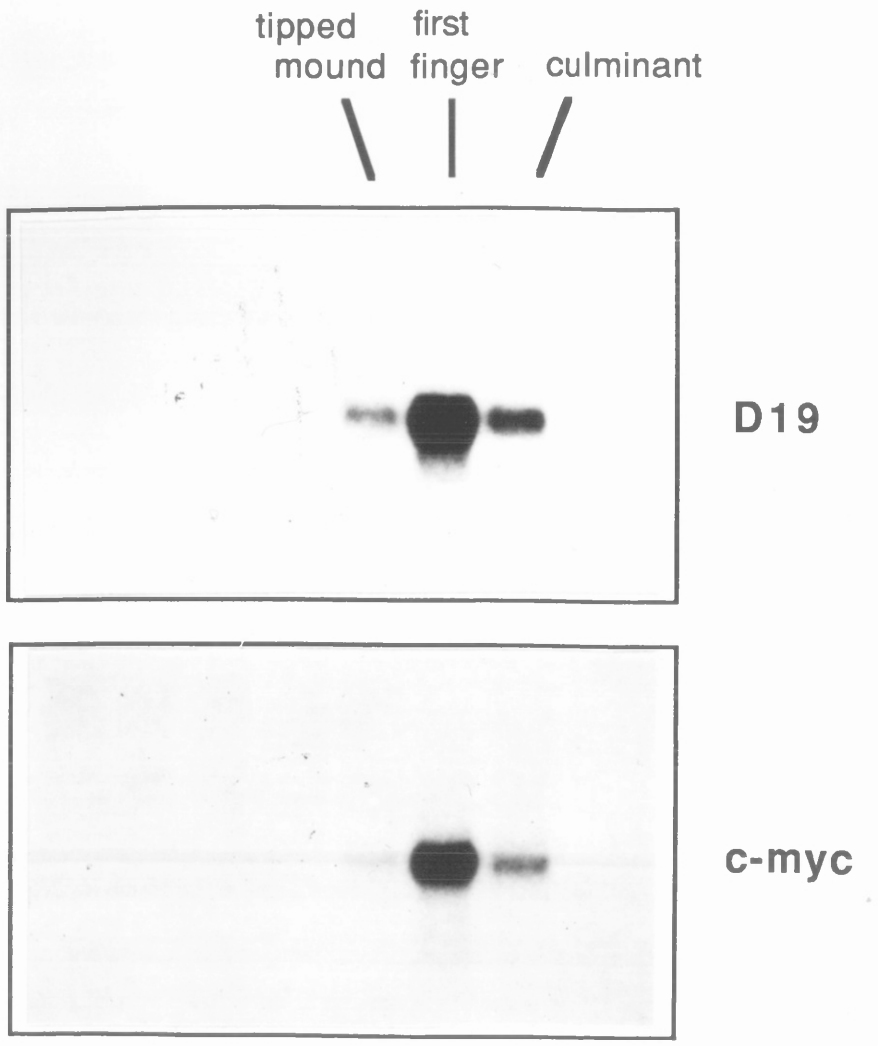


Figure 5.9 Detection of expression of the NdeI-marked gene

Total nucleic acid was extracted with phenol from timepoints harvested during a development on filters and analysed by Northern transfer as described previously. Duplicate blots were probed either with the non-coding strand of the c-myc oligonucleotide, which detects the product of the marked gene alone, or with the D19 cDNA, which hybridises to the product of both the endogenous and transformed genes. The hybridisation and washing conditions used were dependent on the probe.

extent with prespore cells using this separation technique (Tsang and Bradbury, 1981; Ratner and Borth, 1983), the true value must be greater.

Trancripts from the pB10N6 construct, detected with the c-myc oligonucleotide in Northern transfer analysis, also show correct developmental accumulation (Figure 5.9). The mRNA was first detectable in tipped aggregates. Message levels had increased dramatically in abundance by the first finger stage, and were reduced at culmination. Again, no differences were seen when probing the same timecourse with the D19 cDNA clone. The transformed gene is apparently expressed as efficiently as the pB10B1 version.

5.5 Localisation of the D19 protein by indirect immunofluorescence.

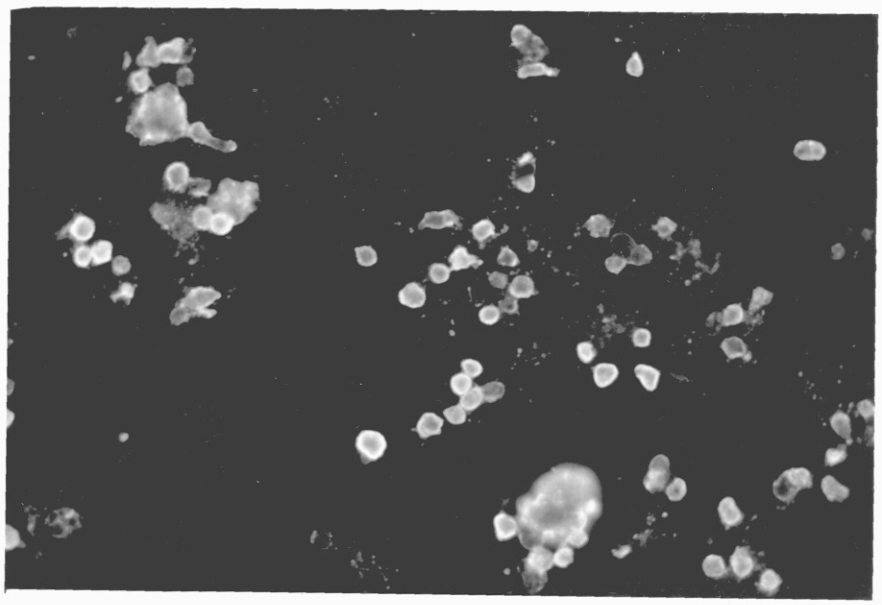
Pooled populations of transformants containing either pB10B1 or pB10N6 were screened for expression of the marked gene at the protein level by indirect immunofluorescence. Migrating slugs were disaggregated, fixed in methanol and stained with 9E10 (Evan et al., 1985), followed by a rhodamine-conjugated second antibody. As a control, cells from AX-2 slugs were treated in the same way. Staining was observed for cells transformed with pB10B1 only. Bright cell surface staining of a high proportion of these cells (Figure 5.10a) was in contrast to an almost complete absence of fluorescence for AX-2 and pB10N6-transformed cells (data not shown). This demonstrates that the position of the oligonucleotide within pB10B1 is such that the epitope is accessible for binding, and that the substitution of

Figure 5.10 Immunostaining of transformants with the anti-c-myc monoclonal antibody, 9E10

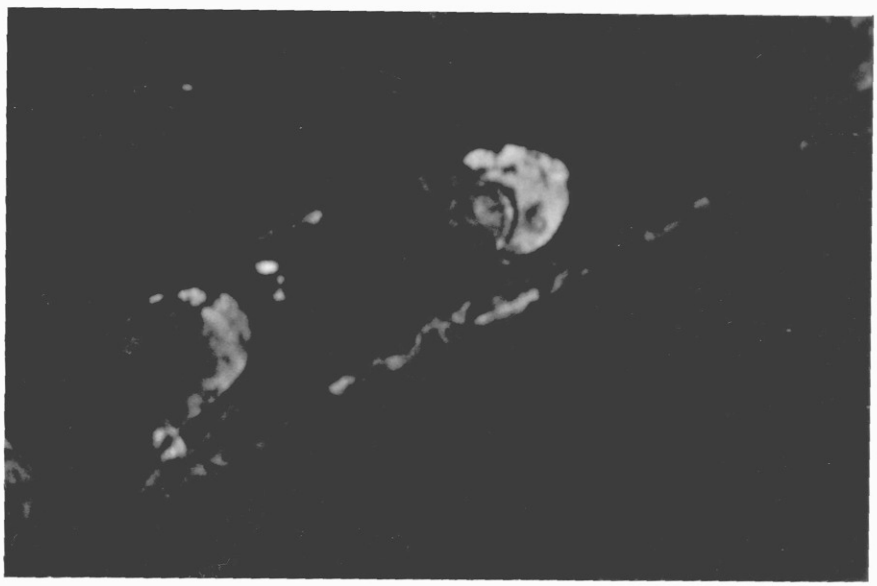
A) Cells expressing the c-myc-marked D19 gene transformed with pB10B1 were induced to form slugs by plating as a streak on water agar and exposing to a unidirectional light source. Disaggregated slug cells were fixed in 60% methanol at a density of 5x 10^6/ml, and dispensed as 5µl drops on to a poly-L-lysine coated multiwell slide. The cells were stained by incubation with a 1 in 100 dilution of 9E10 for 30', followed by a rhodamine-conjugated second antibody for the same length of time. Magnification, x 400.

B) Slime trails, resulting from the migration of pB10B1-transformed slugs, were inverted on to poly-L-lysine-coated slides. The underside of the trails were stained with 9E10 in the manner described above. Magnification, x 255.

A



B



glutamate for glutamine does not prevent antibody-antigen interaction. The absence of any staining for the pB10N6 transformants is probably due to removal of the epitope from the mature protein by processing, since the mRNA is abundantly expressed (Figure 5.8). It may be that the additional 33 nucleotides do not prevent cleavage at the normal position, or cleavage occurs within the epitope, abolishing antibody binding. Because of the inability to detect the D19 gene product in cells transformed with pB10N6, the pB10B1 construct was used as the starting point for subsequent manipulations.

The localisation of the protein on the cell surface of the pB10B1 transformants is as expected, since the D19 gene encodes PsA (see Chapter 3), a known cell-surface glycoprotein (Grant et al., 1985). It is assumed that the proportion of cells which do show staining are prespore cells, but double staining with anti-PSV (prespore vesicles; Tasaka et al., 1983) using a fluoroscein-conjugated second antibody was not clear enough to allow unambiguous comparisons (data not shown). An attempt was made to stain sections of pB10B1-transformed slugs with the 9E10 monoclonal. Cryostat sections were stained using the same protocol employed for disaggregated cells (Figure 5.11). Although a higher background resulted, staining is clearly more intense in the posterior, prespore zone of the slug, which should represent about four-fifths of the total volume. The staining in this region strongly suggests that the protein as well as the RNA is being correctly synthesized in prespore cells.

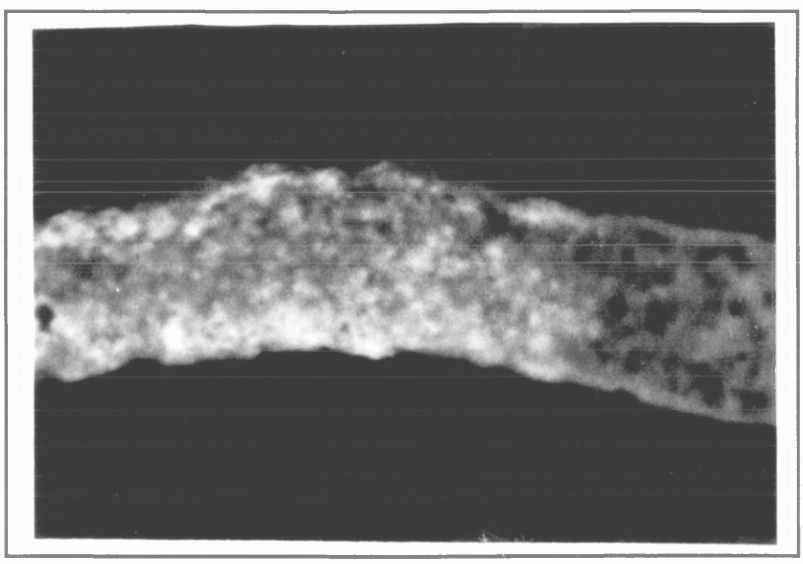


Figure 5.11 Prespore localisation of staining detected with the c-myc monoclonal antibody

Slugs composed of cells transformed with pB10B1 were fixed with paraformaldehyde, placed in a block of embedding medium and frozen at -20°C. 5μm thick longitudinal sections were cut from the block using a cryotat, and the sections were picked up on slides and stained with 9E10 as previously described. Magnification, x 385.

The anti-c-myc monoclonal antibody was also used to stain the underside of slime trails. A distinct horseshoe-shaped pattern is seen (Figure 5.10b). These imprints indicate regions where the slugs have made a close contact with the agar surface during locomotion. Instead of the staining of individual discarded cells, the pattern seen is apparently more diffuse. It suggests that PsA is released into the slime sheath, either by a specific mechanism or simply as a result of cell lysis during slug movement. PsA is very unlikely to be an integral component of the extracellular matrix, unlike 56 and 63, predicted to be stalk tube proteins (McRobbie et al., 1987). Monoclonal antibodies directed against epitopes of these proteins do not give rise to strong cell-surface staining.

5.6 Summary

1) The pB10TP vectors make it possible to clone and manipulate *Dictyostelium* DNA prior to direct use in transformation. pB10TP1 and 3 provide a means of marking cloned genes, while pB10TP2 and 4 allow a wider choice of cloning strategies.

2) Marking the D19 gene with an oligonucleotide encoding an epitope of the human c-myc protein, recognised by a monoclonal antibody, has enabled specific detection of the mRNA and protein products of the transformed gene.

3) Transformation of a marked copy of the D19 gene into *Dictyostelium* demonstrates that 995 base pairs upstream of the cap site are sufficient for correct temporal and spatial D19 mRNA accumulation.

4) The transformed protein, visualised by indirect immunofluorescence, is correctly localised on the surface of cells. Insertion of the epitope within the N-terminal signal sequence, instead of towards the middle of the coding region, led to an absence of detectable protein, presumably due to removal of the epitope by processing.

CHAPTER 6: Analysis of D19 gene expression using the transformation system II: Characterisation of 5' flanking sequences.

Having established that the 5' flanking region contained in the genomic clone 4.3 is able to direct correctly regulated expression, a more detailed analysis of the sequences required for this regulation was undertaken using the transformation system. An important question is whether transcriptional and/or post-transcriptional mechanisms are responsible for cell-type-specific expression, as there is evidence that the expression of individual prespore genes is controlled by cAMP at the level of transcription and mRNA stability (Landfear et al., 1982; Barklis and Lodish, 1983; see Chapter 1).

6.1 Construction of an external 5' flanking deletion series using exonuclease III.

Within the 4.3 genomic insert, the D19 gene is flanked by approximately 1kb upstream of the start site of transcription. The sequence of 190 nucleotides immediately 5' to the cap site was established as a result of the strategy employed to sequence the D19 coding region (see Chapter 3). Only one restriction site, AvaII, was identified upstream of this region. 350 nucleotides of sequence close to the 5' end of the insert derived after labelling at this site did not overlap with any previously established. As *Dictyostelium*

intergenic DNA is extremely A-T rich, it was unlikely that sufficient additional sites suitable for spanning the unknown sequence existed. Therefore a 5' flanking external deletion series was constructed with two aims; to establish the unknown sequence upstream of the gene and to identify the regions important in the control of gene expression using the *Dictyostelium* transformation system.

The deletions were created using a pEMBL derivative of the c-myc-marked genomic clone B1 (see Chapter 5). A filled HindIII-BamHI fragment, spanning the entire insert, was cloned into the filled XbaI-BamHI sites of pEMBL18+ (Dente et al., 1983) to give pB1. In this construct the 5' end of the gene is oriented towards the f1 primer binding site of pEMBL. To create a directed deletion series with exonuclease III (Henikoff, 1984), a means of blocking the activity of the enzyme is required. This is achieved either by filling with thiophosphates or by cutting with a restriction enzyme which generates a 3' overhang. pB1 was cleaved at the PstI and SalI sites in the multilinker, both located on the 5' side of the insert, before exonuclease III digestion. SalI cleaves closer to the insert and leaves a 5' overhang, while PstI generates a 3' overhang which protects the vector sequences from digestion. The cleaved starting material was incubated with exonuclease III for various lengths of time. Timepoints were taken from 30" to 10' in an attempt to span 1kb, with the assumption that exonuclease III digests approximately 200 bases per minute.

A pEMBL18+ derivative was chosen because of the restriction sites available. Suitable sites were required for

initial cloning, cleavage prior to exonuclease III treatment, labelling for sequencing by the chemical cleavage method (Maxam and Gilbert, 1980) and finally excision for subcloning. Additionally, the deletions could be sequenced by the chain termination method (Sanger et al., 1977). The multilinker of pB10TP also contains suitable restriction sites (see Chapter 5), and the need for subcloning would have been obviated. However, experience has suggested that the presence of the additional slime mould DNA contained in this vector can lead to the generation of a high proportion of large deletions, presumably due to nicking at A-T rich regions by S1 nuclease (A.Ceccarelli, personal communication). An unsuccessful attempt was originally made to construct a 3' flanking deletion series in pB10TP4 to sequence across the D19 coding region (data not shown).

S1 nuclease is employed to trim the the single-stranded ends left by exonuclease III. The enzyme level and incubation temperature giving efficient digestion of single-stranded DNA without significant nicking of double-stranded DNA was determined for the pEMBL subclone. After S1 nuclease treatment, the DNA was repaired with Klenow, ligated and transformed into *E.coli*. Minipreps were mapped by digestion with HindIII, a unique multilinker site, and BglII, which cleaves 270 nucleotides downstream of the ATG, within the c-myc oligonucleotide.

The end points of the deletions were determined precisely by sequencing using the chemical cleavage method (Maxam and Gilbert, 1980) after labelling at the HindIII site located 5' to the insert. In total 84 colonies were screened, resulting

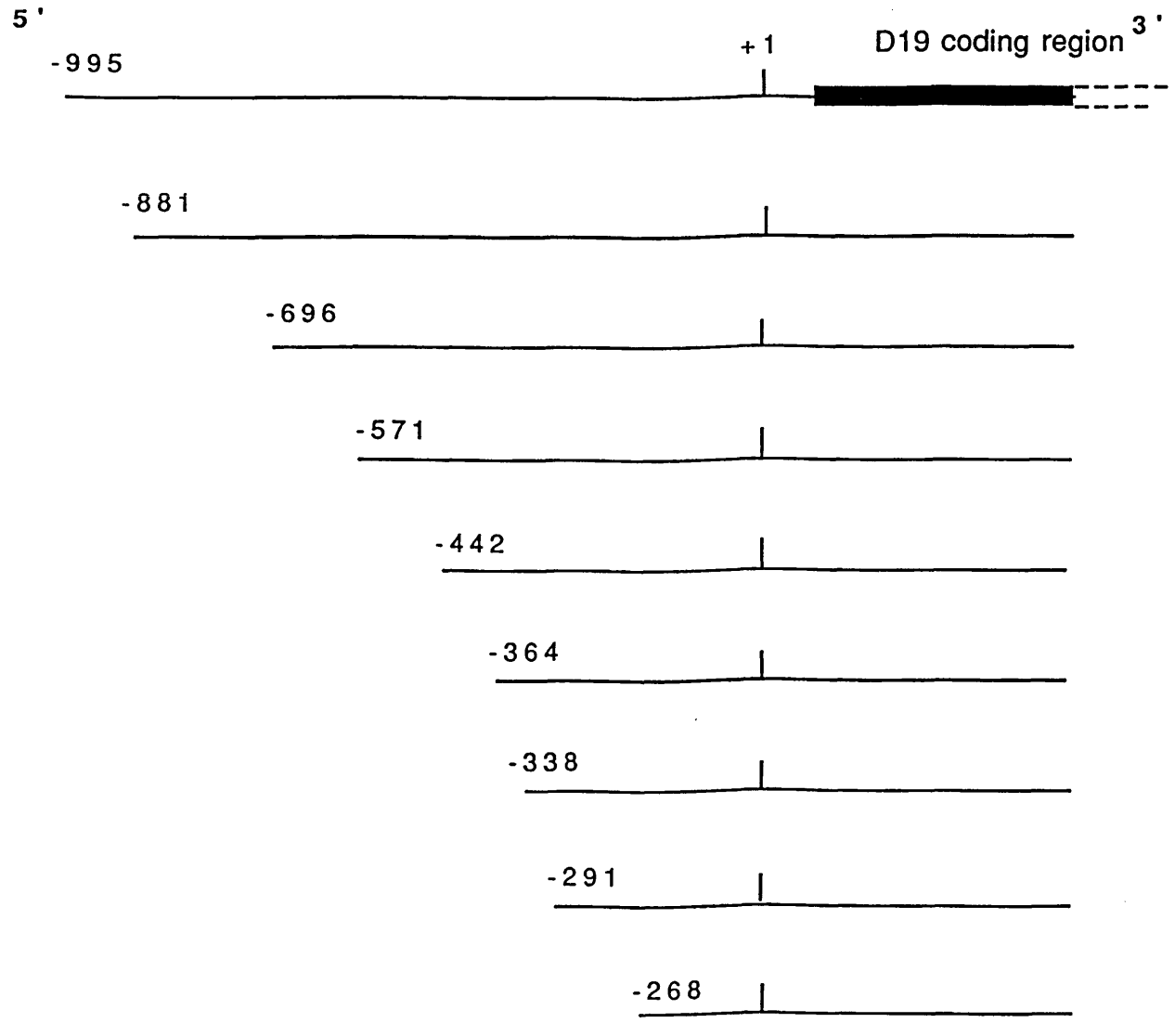


Figure 6.1 Endpoints of the 5' flanking deletion generated with exonuclease III

The numbers shown indicate the length of D19 sequence remaining upstream of the start-site of transcription (+1).

in 25 potentially useful deletions. However, as many were clustered, only 8 were used for further analysis (Figure 6.1).

6.2 Nucleotide sequence upstream of the D19 gene.

The range of deletion sizes generated allowed the complete sequence upstream of the D19 coding region within 4.3 to be determined (Figure 6.2). Sequence from six deletions was required to span the 804 nucleotides of previously unknown sequence. The deletion endpoints correlated well with the predictions of restriction mapping, and there were no mismatches where overlaps occurred. Overall, the sequence composition is typical of *Dictyostelium* intergenic DNA (Kimmel and Firtel, 1982). The 994 nucleotides from the cap site to the end of the insert are composed of 86% A+T and 14% G+C. Because of the extreme A-T richness of *Dictyostelium* DNA, clusters of Gs and Cs are thought to represent potential regulatory signals (Cohen et al., 1986; Driscoll and Williams, 1987; Pears and Williams, 1987). The distribution of Gs and Cs within the upstream sequence was analysed with the programme PCGENE. Interestingly, an asymmetric distribution exists. With respect to the coding strand, the Gs are found mainly within the more distal region, while most of the Cs are concentrated within a 140 base pair region centered 200 nucleotides upstream of the cap site. The Gs and Cs themselves are not evenly distributed. From -330 to -770 six major G-rich blocks, on average 10 nucleotides in length, and one slightly longer C-rich block can be identified (designated I to VIII, Figure 6.2). The



Figure 6.2 Complete nucleotide sequence upstream of the D19 coding region in the genomic clone 4.3

The sequence from the 5' boundary of the insert to the initiation codon is shown. The numbering is with respect to the cap site, denoted by * (see section 3.5). The two potential TATA box sequences upstream of the oligo dT run are indicated. G-C rich sequences are underlined, and numbered where appropriate (see text for details).

more proximal C-rich region contains a higher concentration of G and C residues (37 in 140 nucleotides), generally more evenly distributed, although extremely C-rich motifs exist (underlined in Figure 6.2).

6.3 Transformation of deletion derivatives to identify the 5' boundary of the functional promoter.

The deletions were subcloned from pEMBL18+ into pB10TP4 as a HindIII-EcoRI fragment encompassing the entire insert. The resulting contructs used in transformation differ from pB10B1 only in amount of 5' sequence flanking the gene.

Six subcloned deletions were originally transformed into *Dictyostelium*, terminating 880, 696, 571, 442, 268, and 127 nucleotides upstream of the cap site. Stable transformants were selected with 10µg/ml G418. However, due to difficulties experienced with the transformation procedure at this time, the population sizes obtained were very low. The number of surviving colonies ranged from one to four, and no viable transformants were obtained with the -127 construct. Growth in 2cm tissue culture dishes was necessary initially to establish sufficient numbers of actively dividing cells. The transformants were first screened for expression of the marked gene by indirect immunofluorescence with the c-myc monoclonal antibody 9E10 (see Chapter 5). Slugs derived from each transformant were disaggregated and fixed in methanol prior to staining. Cells transformed with the full-length construct, pB10B1, and untransformed AX-2 cells were used as positive and negative controls respectively.

Expression of the D19 protein was detected until all but 268 nucleotides upstream of the start-site of transcription had been removed (data not shown). Bright fluorescence was seen with all other deletions, including the next largest leaving 442 nucleotides 5' to the cap site. The intensity of staining was apparently little diminished relative to that obtained with the full-length construct. This suggested a sharp cut-off point between a fully active and inactive promoter. However, the sudden absence of staining may be due to the threshhold of sensitivity of the detection method.

Southern blotting data showed that the transformants each contained an approximately equal vector copy number of between 5 and 15 per cell (Figure 6.3). Probing with pB10B1 also demonstrated that no gene rearrangements or further deletions had occurred. The absence of staining for -268 is therefore unlikely to be caused by an event during transformation, despite the survival of a single colony only. However, as the possibility existed that expression of the gene had been affected by the site of integration, both the -442 and -268 constructs were retransformed. Two and eight colonies were obtained respectively. The result was unchanged as only slug cells from the -442 transformant expressed levels of D19 protein detectable with 9E10.

Expression of mRNA from the two transformed genes was then analysed by Northern transfer. Transformants were developed on filters, and total nucleic acid was extracted from timepoints harvested between 12.5 hours and 22 hours after starvation. Nucleic acid was also extracted from growing cells. Probing a Northern blot with the anti-coding

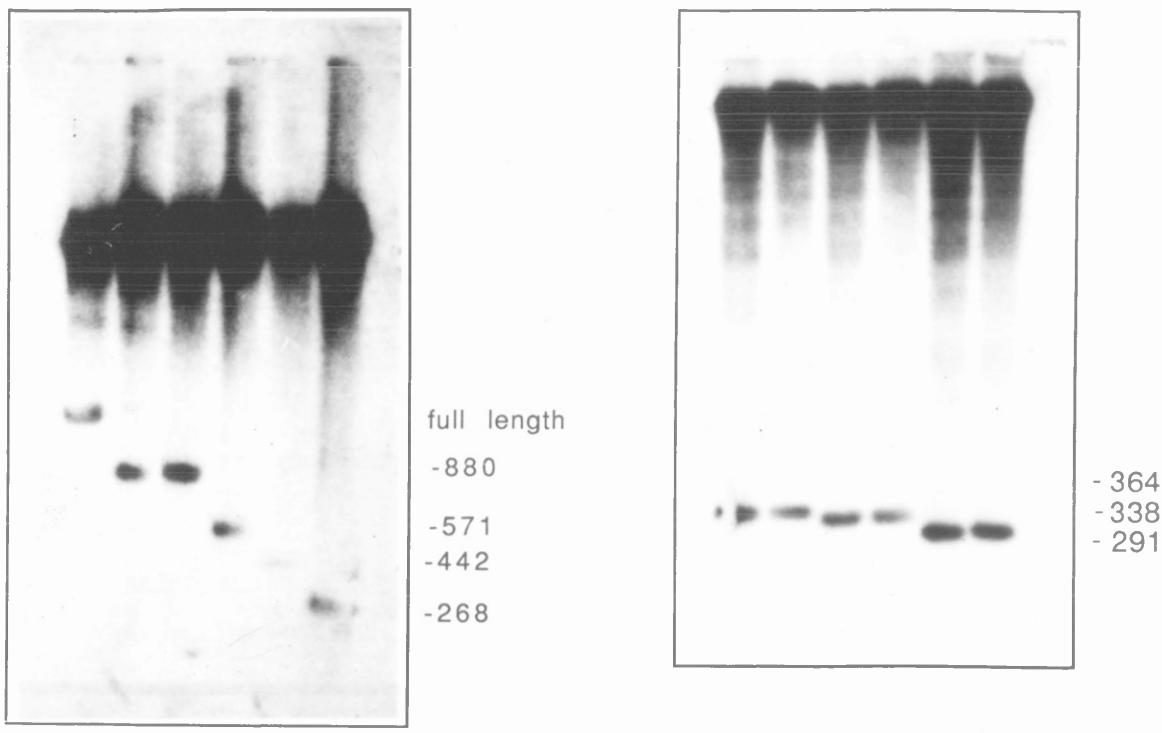


Figure 6.3 Southern blotting analysis of 5' deletion constructs after transformation

50µg of total nucleic acid, extracted with phenol from pooled populations of growing transformants, was digested with HindIII and EcoRI, which cleave the insert from the vector. The size of the insert fragment gave an indication that no large deletions or gross rearrangements had occurred during transformation. After treatment with RNAase and electrophoretic separation on a 0.8% agarose gel, the DNA samples were denatured and transferred onto nitrocellulose. The filter was probed with the full-length construct, pB10B1, and washed in 2x SSC, 0.1% SDS at 65°C for 4x 15'. Known amounts of vector were run on the same gels to allow quantitation of copy number as described in the legend to Figure 5.6.

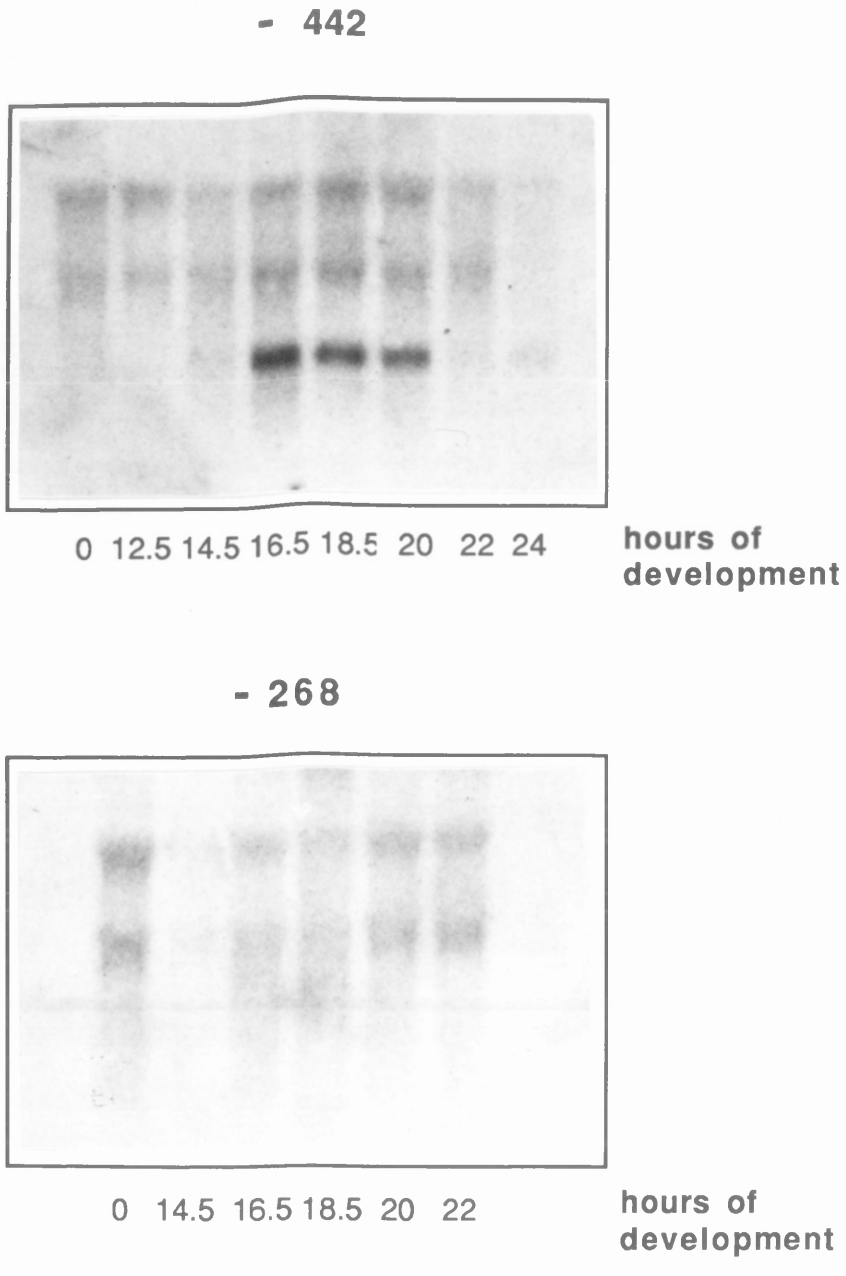


Figure 6.4 Expression of the -442 and -268 deletion constructs during development

Total nucleic acid was extracted with from transformed cells harvested during a development on filters. 5µg samples were run on a 1% formaldehyde gel and transferred to nitrocellulose. The blots were probed with the anti-coding strand of the c-myc oligonucleotide as described in Chapter 5, which hybridises to the product of the marked gene but not that of endogenous D19. The developmental timepoints correspond to the following morphological stages; 0 hrs, vegetative cells; 12.5 hrs, loose mounds; 14.5 hrs, tight mounds; 16.5 hrs, tipped mounds; 18.5 hrs, first finger/standing slug; 20 hrs, mexican hat (early preculminants); 22 hrs, late preculminants; and 24 hrs, culminants.

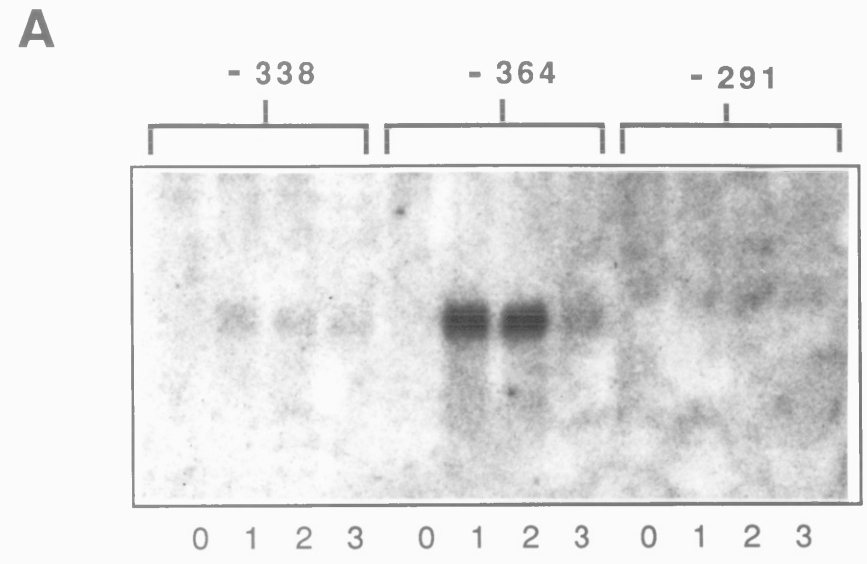
strand of the c-myc oligonucleotide, expression of the transformed gene could be detected for the -442 construct only (Figure 6.4). The signal intensity of the band corresponding to the marked transcript was extremely weak; the two upper bands are background hybridisation to ribosomal RNA and give an indication of the length of exposure time required. No D19 mRNA was detected with the -268 timecourse even after prolonged exposure of the autoradiograph. The low expression of the -442 construct at the RNA level was not predicted by immunofluorescence, and illustrates the difficulty of making meaningful quantitative comparisons with this type of data.

Although of low abundance, the expression of D19 mRNA is correctly temporally regulated by 442 nucleotides of upstream sequence. Detectable levels of marked transcript are only present in tipped aggregates, standing slugs and preculminants. In order to localise this boundary more closely, additional deletions were screened by mapping and sequencing to identify intermediate points. Three new deletions were obtained, terminating 364, 338 and 291 nucleotides upstream of the cap site (Figure 6.1). After subcloning and transformation as before, population sizes of approximately ten clones resulted, and expression of the marked gene was again monitored by staining of disaggregated slug cells. Staining was seen for the -338 and -364 transformants only. This result was reflected at the RNA level (Figure 6.5a), as very faint, correctly developmentally regulated expression was seen with -338, with slightly stronger expression for -364. However, no expression was

Figure 6.5 Expression of the -364, -338 and -291 deletion constructs

(A) Comparison of the expression of the -364, -338 and -291 deletion constructs at equivalent stages of development. Cells transformed with each of the three constructs were developed in parallel on filters, and harvested as tipped mounds (1), on formation of first fingers (2), or during culmination. Vegetative cells were also harvested (0). Total RNA, extracted with phenol, was analysed by Northern transfer as previously described, using the non-coding strand of the c-myc oligonucleotide as a probe.

(B) Prespore enrichment of the product of the marked gene flanked by 338 nucleotides of 5' sequence. Disaggregated slug cells, transformed with the -338 construct, were separated into prespore and prestalk fractions by Percoll density gradient centrifugation (Ratner and Borth, 1983) Phenol-extracted total RNA was then analysed by Northern transfer as previously described.


A
- 338
- 364
- 291
0 1 2 3 0 1 2 3 0 1 2 3
B
- 338
prespore
prestalk

detectable with -291. Southern blotting indicated roughly equivalent vector copy numbers of 10 per cell (Figure 6.3).

The construct containing 338 nucleotides of upstream sequence also directs prespore-specific expression (Figure 6.5b). Transformed slugs were disaggregated, and prespore and prestalk cells were separated on Percoll density gradients (Ratner and Borth, 1983). A Northern blot of RNA extracted from each fraction was probed with the c-myc oligonucleotide. There is approximately a 5-fold enrichment of the marked transcript in prespore cells. Therefore the 5' boundary of the functional D19 promoter must lie somewhere between 291 and 338 nucleotides upstream of the cap site.

A comparison was made of the expression of the various deleted genes with that of the full-length construct at the RNA level. Equivalent amounts of RNA from each transformant, extracted at the time when maximal expression of the marked gene was occurring, were blotted and hybridised to the c-myc oligonucleotide probe (Figure 6.6). A large decrease in expression levels is correlated with reduced amounts of 5' flanking sequence. From the full-length construct to -338 the decline is estimated to be about 40-fold by scanning densitometry. Differences exist between independent populations transformed with the same construct, which may be due to small population sizes. However, the overall pattern is unchanged. The reduction in relative expression levels suggest that additional elements required for efficient expression may be located upstream of sequences directing correct temporal and spatial regulation. An alternative

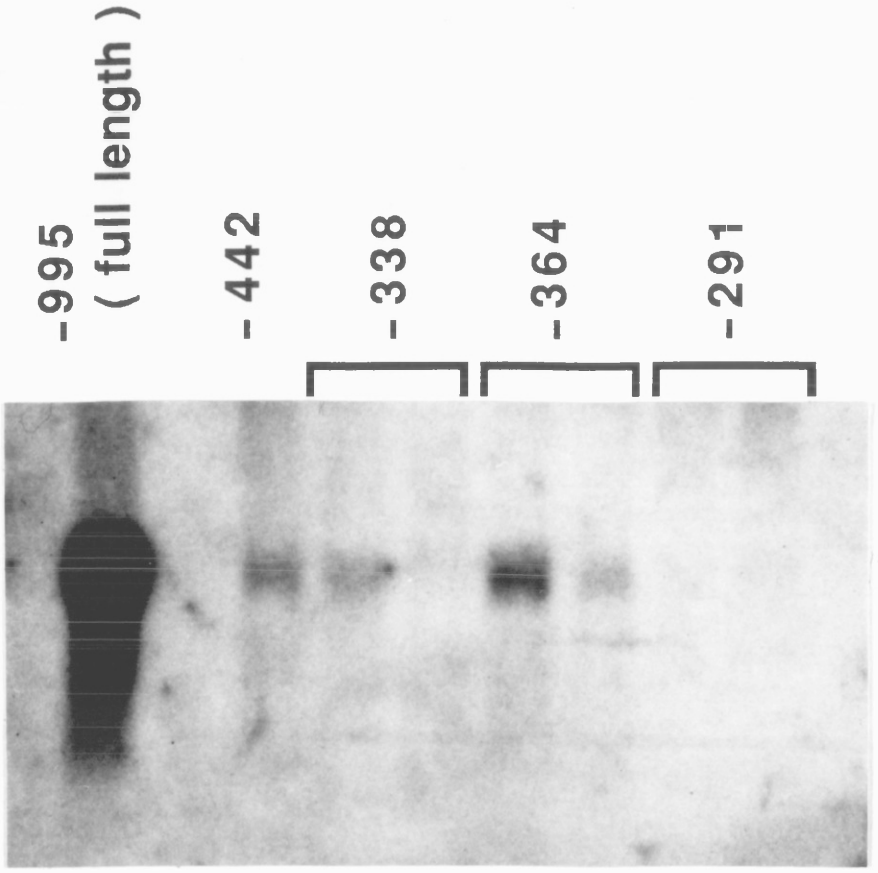


Figure 6.6 Relative expression levels of the 5' deletion constructs

Equal amounts of total nucleic acid (5µg), harvested from the time during development when peak expression of the transformed gene was known to be occurring, were analysed in parallel by Northern transfer. The blot was probed with the c-myc oligonucleotide as previously described. A single pooled population was analysed for the full length construct and the -442 construct, but two independent populations were compared for the -364, -338 and -291 transformants.

explanation is that the increaing proximity of flanking G-C rich vector sequences is inhibiting expression.

Possible candidates for the additional upstream elements are the G-rich regions described in section 6.2 (also see Figure 6.2). These sequences show similarities to G-rich boxes identified upstream of other *Dictyostelium* genes (Figure 6.7), primarily cysteine proteinase 1 and 2 (Driscoll and Williams, 1987; Pears and Williams, 1987); D11, a prestalk-enriched gene (Barklis et al., 1985); discoidin 1A, 1B and 1C (Poole and Firtel, 1984); and SC253 (Ayres et al., 1987; = pDd56). For cysteine proteinase 2, mutation or deletion of these elements results in greatly reduced efficiency of expression (about 50-fold; Pears and Williams, 1987). D19 regions III and VII (Figure 6.2) are the most G-rich and hence more closely resemble sequences upstream of other genes. All but region VIII were removed in the last deletion to promote expression (-338). The boundary of this deletion was within VIII, downstream of GCTA leaving TGATGG. The deletion with 442 nucleotides upstream of the cap site was expressed at a consistently higher level (see Figure 6.6) and additionally contained regions VI and VII.

6.4 Generation of 3' flanking deletions and construction of heterologous gene fusions.

The closest 5' flanking deletion which still expresses D19 retains 338 nucleotides upstream of the start-site of transcription. To define the sequences required for correctly regulated expression more closely, a second deletion series was constructed in which DNA was progressively deleted from

Gene		Orientation relative to gene	Position relative to cap site
CP1	GGGGGTATGGTGG	→	- 765
	AAAGCATAGGTGG	←	- 320
	AAAGGAATGGGGA	→	- 290
	AAAGCATTGGTGG	←	- 243
	AAAGGAATGGGGG	→	- 213
CP2	TTAGGGGTGGGTT	→	- 217
	CAGGGGGTGTGTT	→	- 233
D 11	GGGGGAGTGTTGG	→	- 156
Discoidin 1 A	TTAGGGGTGGTGT	→	- 222
	TGGTGAGTGGTTT	→	- 165
1 B	TGAAAAGTGGGGT	→	- 176
1 C	AAATTATTGGGGG	←	- 215
pDd 56	AGGGGGTTGGGGG	→	- 560
D 19	GAGCAACAAGGCA	→	- 427
	ATAGGTTTGGGAA	→	- 385
	ATGCTATGATGGA	→	-336

Figure 6.7 G-rich sequences identified upstream of several *D.discoideum* genes

The three sequences shown for D19 correspond to regions VI, VII and VIII underlined in Figure 6.2. The G-rich sequence centered at -336 is interrupted by the -338 deletion, while all three are retained in the -442 construct.

the 3' end. The deletions were generated with Bal31, starting at a point in the coding region, as no suitable restriction sites existed further upstream. Therefore, elements essential for expression, such as the TATA box, were removed before the sequences of interest were reached. To overcome this problem, the deleted 5' flanking regions of D19 were subcloned upstream of a heterologous gene, actin 15 CAT, in the vector ΔBamCAT (C.Pears, Ph.D thesis).

The promoter of the actin 15 gene is one of the best characterised in *D.discoideum* (Romans and Firtel, 1985; Cohen et al., 1986; Nellen et al., 1986). Actin 15 is expressed at low levels in vegetative cells, but the mRNA abundance increases about ten-fold after two hours of development (Knecht et al., 1986). 270 basepairs of 5' flanking DNA are sufficient to confer this pattern of accumulation on a heterologous gene (Cohen et al., 1986). Deletion analysis has localised the sequences involved to either one or both of the GC-rich areas at -138 and -157 relative to the ATG (designated boxes 3 and 4). The box 4 sequence is found upstream of a coregulated actin gene (Nellen et al., 1986). Deletion of two other G-C rich regions (boxes 1 and 2) had no effect. However, expression was correlated with the appearance of two DNaseI hypersensitive sites mapping to boxes 2 and 4 (-138 and -190).

A non-expressing version of the actin 15 gene was constructed in the transformation vector. The actin 15 coding sequences from the genomic clone SC79, with 290 bp of 5' flanking and 500 bp of 3' flanking DNA, were cloned into pB10TP1. Two BamHI sites were created simultaneously at

nucleotide -190 and -138 by site-directed mutagenesis (C.Pears, Ph.D thesis). By cleaving with BamHI and recircularising, the G-C rich areas at -138, 157 and -190 were deleted and replaced with a cloning site. It was hoped that insertion of upstream sequences from other genes would enable identification of promoter elements by reconstitution of expression.

As a more rapid method for detecting gene expression, the chloramphenicol acetyltransferase (CAT) gene was inserted in frame into the actin 15 coding region. The level of expression can thus be assayed by measuring levels of CAT activity in crude cell extracts, rather than by analysing extracted RNA (Gorman et al., 1982). A 787 bp HindIII fragment from pYEJ001 (Pharmacia), encompassing most of the CAT coding sequences, was inserted close to the 5'end of the actin 15 coding region to give A15CAT (K.Duffy, unpublished results). The deleted version of the A15 promoter was then included by substitution of an XbaI-BglII fragment (C.Pears, Ph.D. thesis). The resulting constuct was designated ΔBam.CAT (Figure 6.8).

When transformed into AX-2 cells, pooled populations, with an average copy number of 40-50 per cell, show no detectable CAT activity above the level in non-transformed cells, either during growth or after four hours of development in shaken suspension (C.Pears, Ph.D. thesis). In contrast, with an intact actin 15 promoter, enzyme activity is detectable in cell extracts harvested during both growth and development.

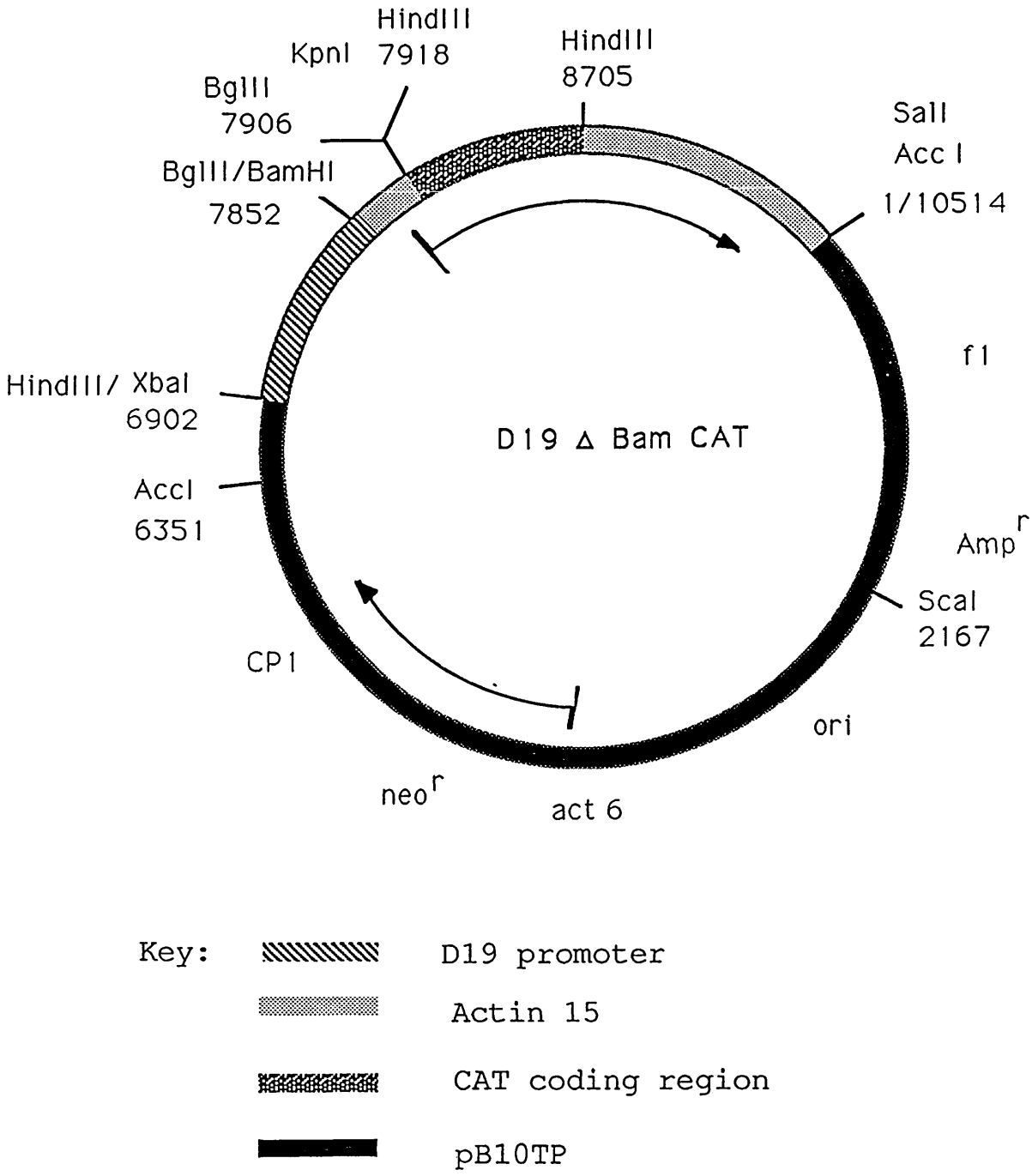


Figure 6.8 Restriction map of D19 ΔBam.CAT

D19 ΔBam.CAT was constructed by directed insertion of D19 promoter sequences purified from the 3' deletion constructs as a BglII-HindIII fragment into XbaI-BamHI of ΔBam.CAT. The arrows indicate the direction of transcription of the Tn5 neo/CP1 and the actin 15/CAT fusion.

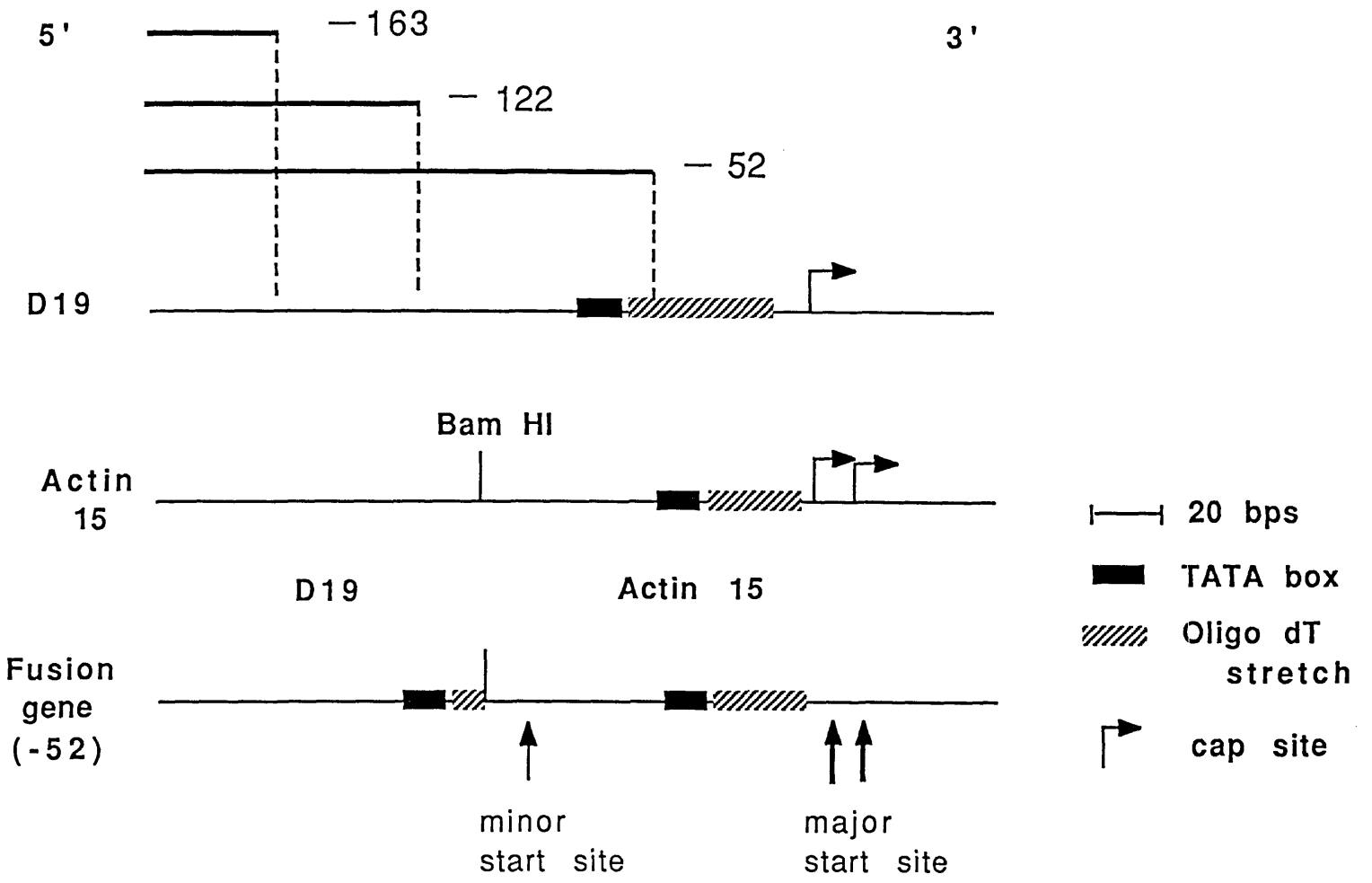


Figure 6.9 Relative arrangement of D19 and actin 15 promoter elements in the intact and fusion genes

pB1 (pEMBL18+ containing the D19 genomic insert with 994 nucleotides of 5' flanking sequence; see section 6.1) was used as the starting material for the 3' flanking deletion series. This construct enabled sensitive detection of the point at which expression was lost, as initial expression levels were known to be high. The vector, cleaved at the unique NsiI site within the coding region, was digested with Bal31. Bal31 digestion is not unidirectional (Legersky et al., 1978), but was not required to be since the deletions were then subcloned into ΔBamCAT. Conditions were established which removed approximately 100 bases from each end/minute, and digestion was continued for up to 5 minutes. After addition of BglII linkers and recircularisation, the DNA was transformed into *E.coli*. The sizes of the deletions were mapped by digestion within the linker with BglII, and with HindIII, which cleaves to the 5' side of the insert. The deletion endpoints were precisely localised by Maxam and Gilbert sequencing after labelling at the BglII site. Three deletions were selected for further analysis, terminating 52, 122 and 163 nucleotides upstream of the cap site. The upstream sequences remaining were inserted into the filled XbaI-BamHI sites of ΔBam.CAT as a filled HindIII-BglII fragment. Thus, all actin 15 flanking sequences upstream of the BamHI site were additionally removed. Figure 6.9 and 6.10 show the relative arrangement of D19 and actin 15 promoter elements in the deletion constructs and the sequence across the site of fusion respectively.

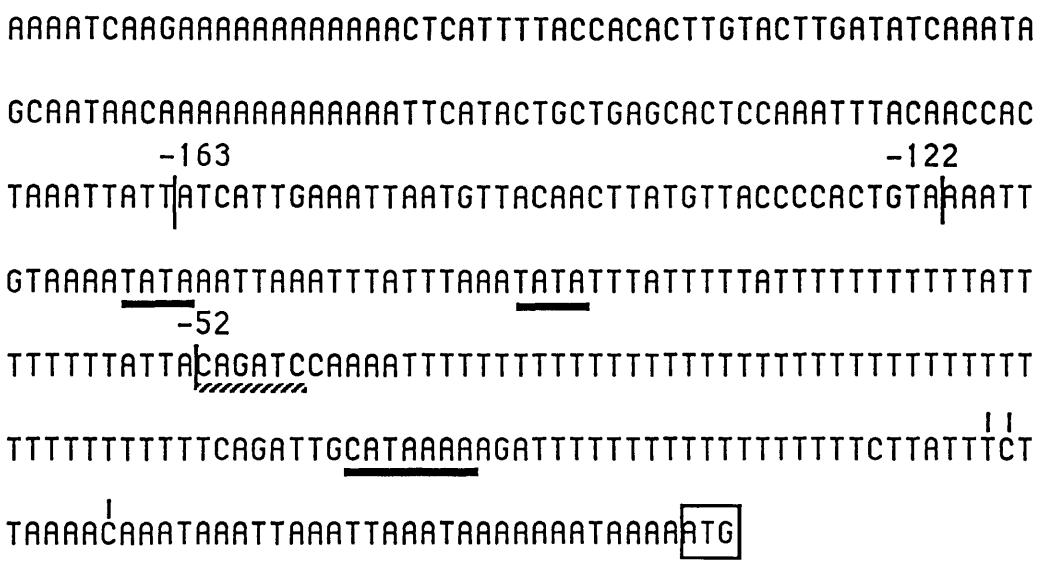


Figure 6.10 Sequence across the D19/actin 15 fusion

The D19 sequence shown is that of the shortest 3' flanking deletion (-52), and the boundaries of the other two deletion are marked. All constructs contained an additional 713 nucleotides of D19 upstream sequence not shown. The extra sequences introduced at the junction between D19 and actin 15 by cloning are underlined(~~~~~). The putative D19 and actin 15 TATA boxes (——) and actin 15 start sites of transcription (|) are also indicated.

6.5 Identification of the 3' boundary of the D19 promoter sequences required for regulated expression.

The three fusion gene constructs and A15CAT, which is known to express CAT at high levels, were transformed into AX-2 cells. Stable transformants were selected with 10µg/ml G418, and pooled populations of 500-1000 clones were analysed. Gene expression was initially detected by means of CAT assays. Crude extracts were harvested from growing cells in shaken suspension and from cells developed on filters. A15CAT provides a positive control for the assay, and extracts from AX-2 cells were also included to assess the background level of CAT activity. Known amounts of extracted protein were incubated with cold acetylCoA and ^{14}C-chloramphenicol. Acetylated products were separated from the starting material by thin layer chromatography.

CAT activity was detected in extracts from cells transformed with two of the three deletion constructs only; those deleted to 52 and 122 nucleotides upstream of the cap site (Figures 6.11 and 6.12a). When 163 nucleotides are removed, the levels of acetylated chloramphenicol are not above background (Figure 6.12b). Where expression does occur, it is correctly developmentally regulated. For both the -52 and -122 constructs, the enzyme activity profile reflects that the accumulation pattern expected for D19 mRNA. As protein levels are being measured instead of RNA, the major increase in expression occurs slightly later and the peak level decreases little during culmination. Actvity above background is first observed at tipped mound/first finger and reaches a peak in early preculminants. Therefore the 3'

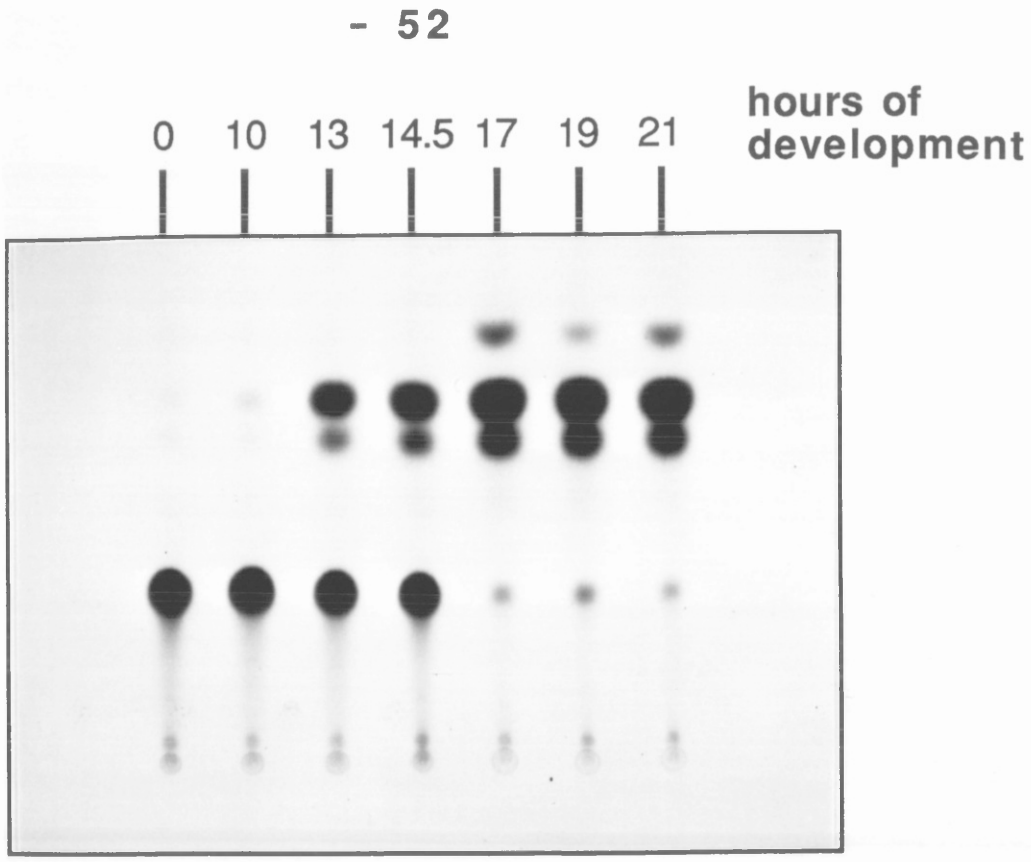
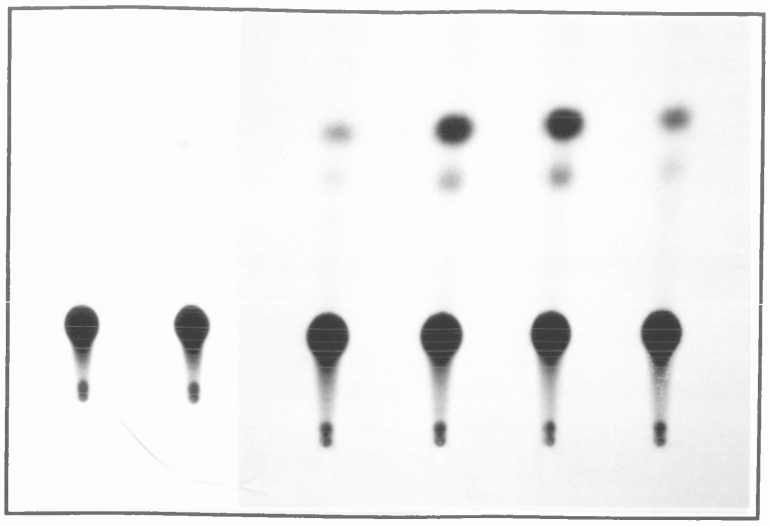


Figure 6.11 Expression of the -52 D19/actin fusion construct detected by CAT activity

Cellular extracts were harvested from a pooled population of transformants, either growing in shaken suspension (0 timepoint) or developed on filters. 2µg of extract from each point was incubated for 30' at 37°C with acetyl CoA and ^{14}C-chlomamphenicol, and the acetylated products were extracted with ethyl acetate and spotted on to a thin layer chromatography (TLC) plate. The separation was continued until the solvent front had reached within ~1cm of the top of the plate. The lowest spot represents unmodified chloramphenicol, while in order of increasing mobility the acetylated products are 1-acetate chloramphenicol, 3-acetate chloramphenicol and 1,3-acetate chloramphenicol. After drying, the plate was subjected to autoradiography. The times shown correspond to the following morphological stages;0, vegetative cells, 10hrs, loose mounds, 13hrs, tipped mounds, 14.5hrs, first finger, 17hrs, slugs, 19hrs, early preculminants, and 21hrs, late preculminants.

- 122



-163

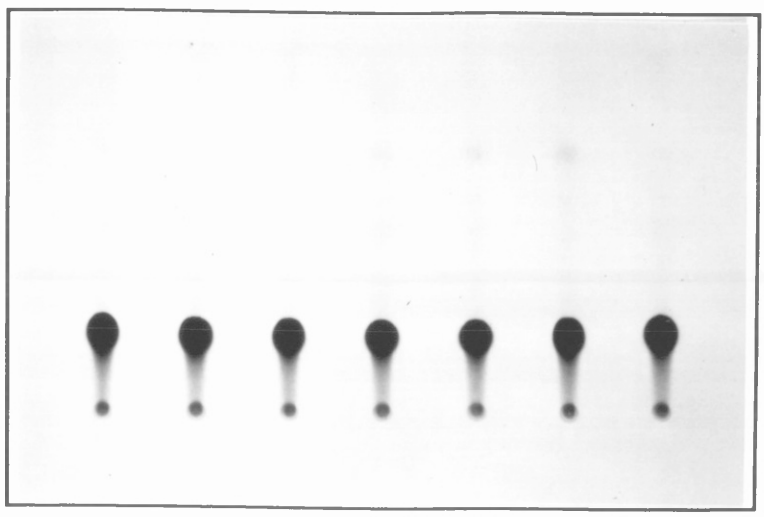


Figure 6.12 Expression of the -122 and -163 fusion constructs

Extracts were harvested during filter developments of pooled populations transformed with the respective constructs, as well as from growing cells. Equal amounts of extract (2µg) were assayed for each point as before, and the acetylated products separated from the unmodified chloramphenicol on a TLC plate. From left to right, the samples analysed for the -122 transformant are; 0hrs, 13hrs, 18hrs, 20hrs, 22hrs and mature culminants (harvested after 36hrs). No background was seen with extract from growing AX-2 cells or after omitting extract altogether (data not shown). Also from left to right, the -163 set of reactions shown are AX-2 0; -163 0, 12.5hrs, 15hrs, 17.5hrs and 20hrs; and no extract.

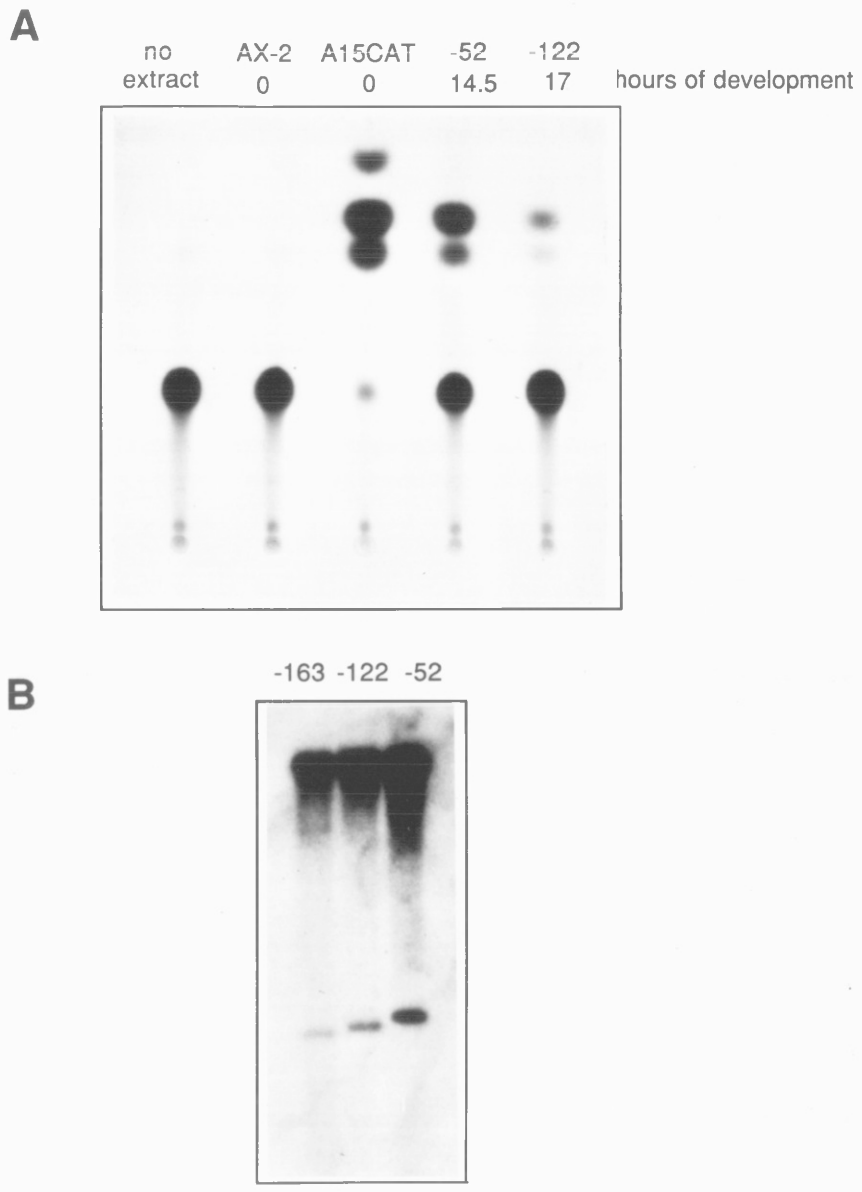
boundary of the D19 promoter lies between -122 and -163 nucleotides upstream of the cap site, a 41 bp GC-rich region (see section 6.6).

However, there is a large difference in the relative CAT activity levels of the two expressing constructs. A comparison was made between extracts harvested from AX-2 and A15CAT-transformed growing cells, and from cells transformed with the D19 constructs maximally expressing the fusion protein (Figure 6.13a). Ten times as much extract (10μg) from the -122 transformant was added. The A15CAT extract acetylated almost all of the starting material, producing a third, upper spot which is diacetylated chloramphenicol. With the -52 construct, the peak enzyme activity is at about 50% of this level. However, the amount of extract used from cells transformed with the -122 construct has at least 5-fold less activity. Thus the true expression level is reduced at least 50-fold relative to the -52 construct. Southern blotting indicated approximately equal vector copy numbers of between 10 and 25 for cells transformed with each of the three deletion constructs (Figure 6.13b).

The reason for such a significant reduction in efficiency with the removal of 70 nucleotides is not known. This region is composed almost exclusively of As and Ts, with just one G residue close to the 5' end (Figure 6.18). The -52 construct interrupts the beginning of the oligo-dT run upstream of the start site of transcription, while the -122 deletion terminates just 3' to a C-rich block, and it may be that a few extra nucleotides downstream, including the G, are required for full activity of this potential sequence motif.

Figure 6.13 Relative expression levels of the 3' deletion constructs

A) Samples from the two 3' deletion constructs to show expression, harvested at the time of peak expression (14.5hrs for -52 and 17hrs for -122), were compared to a sample from growing cells transformed with the actin 15/CAT construct. 2µg of protein were included for actin 15/CAT and for the -52 construct, and 20µg for the -122 construct.The incubation and separation by thin layer chromatography was carried out as previously described. Samples from non-transformed Ax-2 growing cells and incubations performed with no cellular extract were also compared to check the levels of background CAT activity.
B) Total nucleic acid, extracted with phenol from the stable transformants containing the 3' deletion constructs, was analysed by Southern transfer. 50µg of nucleic acid was digested with BglII and XbaI, which flank the 3' and 5' boundaries of the actin15/D19 upstream sequences, and electrophoresed through a 0.8% agarose gel. After denaturation and transfer, the blot was probed with pB10B1, labelled with ^{32}P. The filter was washed with 2x SSC, 0.1% SDS for 4x 30' at 65°C.


A
no extract
AX-2 0
A15CAT 0
-52 14.5
-122 17
hours of development
B
-163 -122 -52

Alternatively, the As and Ts removed may be part of an independent element required for efficient expression. A third possibility is that the spacing between the D19 and actin transcriptional signals is critical for efficient expression. For example, if transcription initiated at a different position in the -122 construct the resulting message might be more unstable. Alternatively, as the -52 construct, unlike the -122 construct, retains the D19 TATA box, precise interactions between proteins binding here and those further upstream may be required for maximal expression.

Primer extension indicated that the start sites of transcription for both of the fusion genes were unchanged compared to A15CAT (Figure 6.14). A 20-nucleotide oligomer, homologous to the anticoding strand of the CAT gene 70 nucleotides downstream of the site of fusion, was used as a probe. An almost identical pattern of multiple bands were seen with each RNA sample. Multiple start sites have previously been identified for actin 15 and three major transcripts exist (Tsang et al., 1982). The upper three bands map very closely to the reported start sites, and the smaller bands are likely to be drop-off products. The fact that the same bands are seen with RNA from all three transformants suggests that the actin 15 TATA box is being utilized in each. A faint upper band is also present with the -52 construct. The increase in message size almost exactly corresponds to the distance between the D19 and actin 15 TATA boxes in this construct (~40bp). The additional start site maps just upstream of the actin 15 TATA box, which suggests

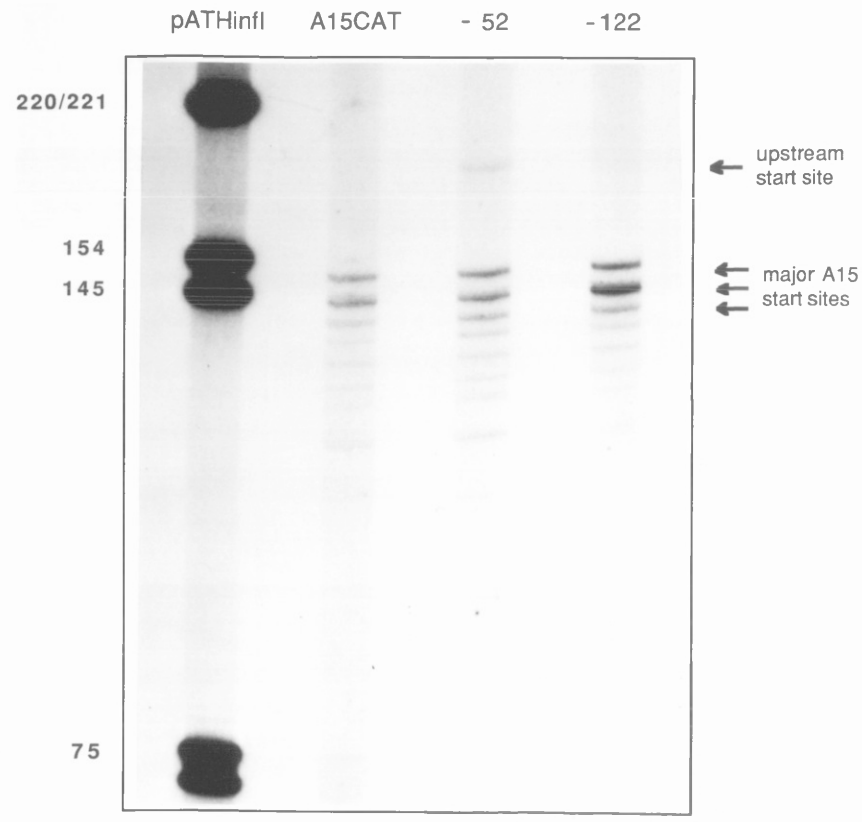


Figure 6.14 Identification of the start-site of transcription of the D19/actin 15 fusion constructs

2µg of total RNA from the A15CAT 0 and -52 16.5hrs samples, and 10µg of total RNA from purified prespore cells of the -122 transformant, was annealed to an excess of primer (~5ng), labelled with T4 polynucleotide kinase, at 45°C. The sequence of the primer was 5'-CAACGGTGGTATATCCAGTG-3', and hence the annealing temperature used was 15°C below the Tm. Primer extension was performed by incubation of the annealed template and primer with units of reverse transcriptase for 60' at 42°C. The nucleic acid, purified by ethanol precipitation, was analysed on a 7.5% denaturing gel. The size of the extended product was determined by comparison with pATHinfI. The position(s) of the start site of transcription could then be mapped relative to the sequence complementary to the primer.

that the D19 TATA box is being used. However, as this constitutes only a minor part of the total expression, it cannot explain the increased efficiency compared to the -122 construct.

Expression of the -52 construct was also examined by Northern blot analysis (Figure 6.15). RNA was extracted from time-points harvested during development on filters. A Northern blot was hybridized to a fragment of the CAT gene, which does not cross-hybridize with any endogenous *Dictyostelium* messages. A sample of RNA from A15CAT transformed cells was included as a positive control. A single discrete band was seen of identical size for both transformants. The pattern of mRNA accumulation during development was very similar to that observed for D19, and the mRNA abundance was comparable to that of the A15CAT transcript in growing cells, which correlates closely with the results of the CAT assays.

In both the -52 construct and the -122 construct, 5' non-coding sequences have been removed, and hence transcriptional signals alone remain. Despite the absence of post-transcriptional control mechanisms, the developmental regulation of D19 expression appears to be normal. Thus, these constructs was used for further analysis of the transcriptional control of D19 expression.

Firstly, CAT assays were performed with cells harvested earlier during development, i.e. from 0 to 12 hours (Figure 6.16). The level of CAT activity was not above background until the time at which D19 mRNA is usually detected. This indicates that the promoter is inactive during the initial

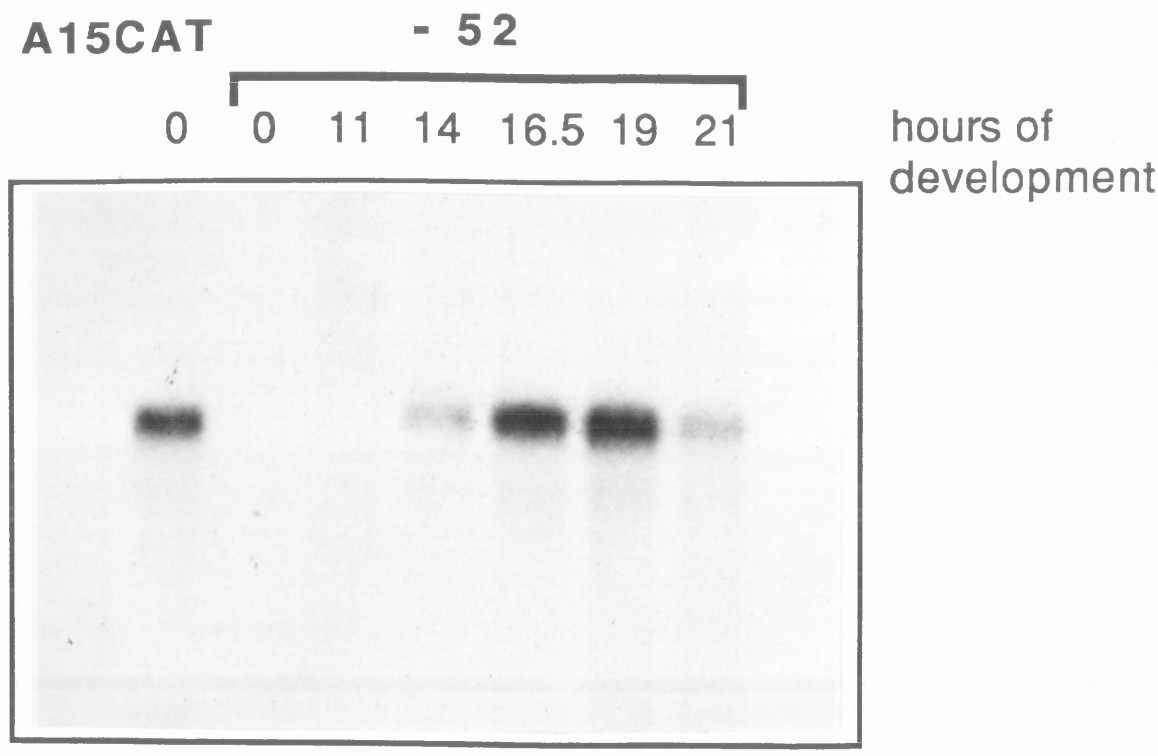


Figure 6.15 Expression of the -52 construct at the RNA level

A pooled population of cells transformed with the -52 D19/actin fusion construct was analysed by Northern transfer. RNA was extracted with GTME from timepoints harvested during a filter development, and 5µg samples were run on a 1% formaldehyde gel. A sample from growing cells transformed with A15CAT was included as a positive control. The blot was probed with a 787bp HindIII fragment from the CAT gene purified from pYEJ001, labelled with ^{32}P. The filter was washed with 2x SSC,0.1% SDS for 4x 15' at 65°C. The times shown correspond to the following developmental stages;0, growing cells, 11hrs, loose mounds, 14hrs, tipped mounds, 16.5hrs, slugs, 19hrs, early preculminants, 21hrs, late preculminants.

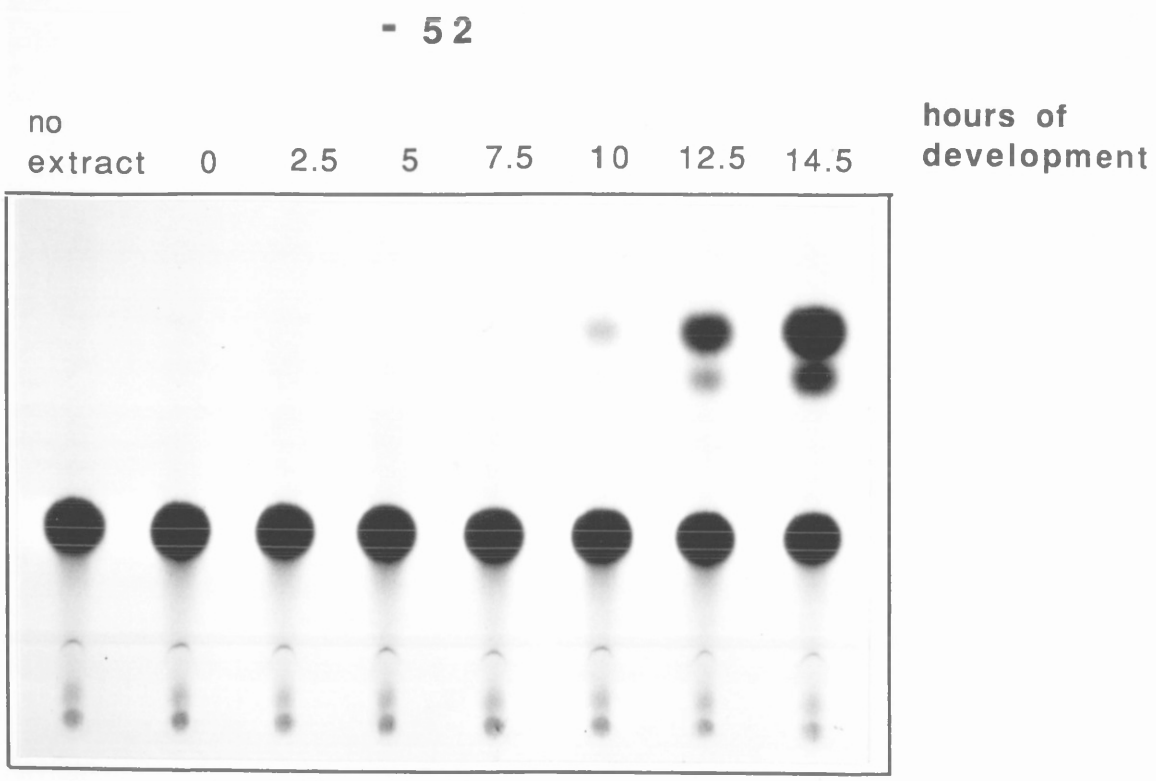


Figure 6.16 CAT activity generated by the -52 construct early in development

Cellular extracts were harvested at the times shown during the early stages of a development on filters. 2µg samples of protein were analysed by performing CAT assays as previously described. Again, a sample from growing cells transformed with the actin 15/CAT construct was included as a positive control, and a reaction omitting extract as a negative control.

stages of development, and confirms that the accumulation of the D19 gene product is largely controlled at the transcriptional level. For some prespore clones, transcription has been shown to occur several hours before cytoplasmic mRNA is detected (Mangiarotti et al., 1985). It has been suggested either that the delay is caused by a lag in transport from the nucleus, or that accumulation of mRNA is triggered by an increase in stability (Mangiarotti et al, 1983a+b). Hence for D19 there is no evidence for either of these mechanisms (see Chapter 8).

The -122 construct has provided information about the control of cell-type specific expression. Dissociated slug cells were separated into prespore and prestalk fractions, again by Percoll density gradient centrifugation, and nucleic acid was extracted and analysed by Northern transfer (Figure 6.17). A fragment from the CAT gene fragment, used as a probe, detected at least 10-fold higher levels of fusion transcript in prespore over prestalk cells. A very similar degree of prespore enrichment resulted using the D19 cDNA clone to probe the same RNA samples (data not shown). In a repeated experiment, where a comparable result was obtained with the two original probes, additional use of the IG7 cDNA clone, which hybridises to a non-cell-type-specific mRNA species, indicated that equal amounts of intact RNA had been loaded in each lane (data not shown). Overall these experiments show that transcriptional control is sufficient for prespore-specific expression of the D19 gene.

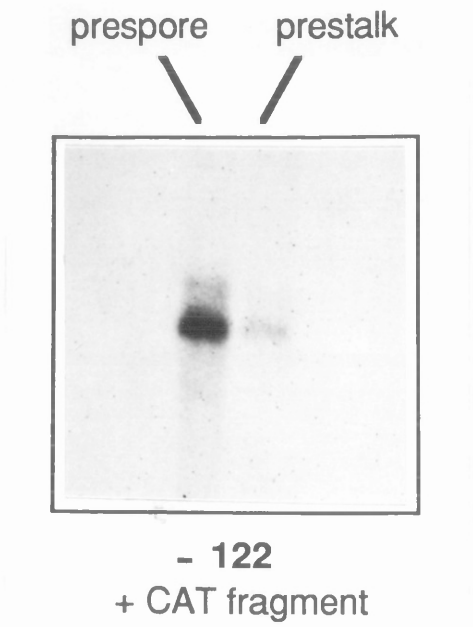


Figure 6.17 Cell-type-specificity of the expression of the -122 construct

A pooled population of cells transformed with the -122 D19/actin fusion construct were induced to form slugs by development on water agar plates exposed to a unidirectional light source. Disaggregated slug cells were separated into prespore and prestalk fractions by Percoll density gradient centrifugation, and total nucleic acid was extracted with phenol. 5µg samples were analysed by Northern transfer as previously described, using the CAT gene fragment as a probe.

6.6 Sequence characteristics of the minimal promoter region, and attempts to reconstitute D19 promoter activity with synthetic oligonucleotides.

Deletion analyses have shown that 216 nucleotides contain the transcriptional signals necessary for correctly regulated, if inefficient, expression (Figure 6.18). This region contains no sequences showing similarity to any previously identified motifs, nor any perfect dyad symmetry elements. A few short direct repeats can be identified; two separated copies each of ACCAC and CCACT, one set of which overlap, two adjacent copies of ACTTG, and two closely linked copies of ATGTTAC. In addition there is an inverted repeat overlapping one of the direct repeats. Close to the 3' end is a highly C-rich motif, CCCCAC, which shows similarity to an inverted copy of a G box (see section 6.3). Both this and the inverted repeat are removed by the -163 deletion. Slightly further upstream, centered at -195, is a longer GC-rich block, CATACTGCTGAGCACTCC. Comparison with upstream sequences from another prespore gene, 914 (A.Ceccarelli, personal communication) has shown partial homology of D19 sequences to a dyad symmetry element, TACCACTATGGTA. At the analogous position upstream of the D19 cap site (centered at -236) the sequence AACCACACTGTA is present, with 10/14 matches. At -174 ACCACTA is found, homologous to the distal half of the element. Further upstream, outside the minimal region defined by deletion analysis, two copies of TGGTA exist at -597 and -625.

The effect of removing sequences important for developmental regulation by creating small internal deletions

Figure 6.18 Minimal promoter region defined by deletion analysis

At both the 5' and 3' ends, the last deletions to show expression and the first to abolish expression are shown. Short additional sequences were included at the 5' ends of the synthetic oligonucleotides to create BamHI compatible ends and a 5' BamHI site after cloning, i.e. upper strand, GATCC, lower strand, GATC.


-338
TGATGGATAATTTTTTTATTTTTTTACAAATAATAA
-291
AAATAATTTAATTTTATTGTAAAATCAAGAAAAAAA
AAAAACTCATTTTACCACACTTGTACTTGATATCAA
ATAGCAATAACAAAAAAAAAAAAATTCATACTGCTG
-163
AGCACTCCAAATTTACAACCACTAAATTATTATCAT
-122
TGAAATTAATGTTACAACTTATGTTACCCCACTGTA
AAATTGTAA
Key
Partial homology to 914 element
Inverted repeat
Direct repeat
Oligonucleotide I
Oligonucleotide II

may be impossible to judge if expression is simultaneously abolished. Therefore an alternative approach was adopted to identify sequences responsible for regulated expression; insertion of synthetic oligonucleotides into a modified version of ΔBam.CAT (C.Pears, Ph.D thesis). 68(+)CAT contains a 68 nucleotide stretch of the cysteine proteinase 1 (CP1) promoter including a 9 nucleotide G-rich region. Two almost perfect repeats of a slightly longer 80 bp region containing just one mismatch are found upstream of CP1, and two sequence elements similar to the G-rich core, located upstream of the coregulated gene CP2, are required for efficient expression (Figure 6.7). The CP1 and CP2 elements seem to be functionally equivalent because CP1 sequences can restore expression of a deleted version of the CP2 gene (C.Pears and J.Williams, 1988). When the CP1 68-mer was inserted, by means of a synthetic oligonucleotide, upstream of the A15CAT gene fusion in ΔBam.CAT, expression was not restored (C.Pears, Ph.D. thesis). However, it was hoped that if regions of the D19 promoter were additionally inserted, the combination of elements present would be sufficient to induce expression. A possible explanation for the position of the 5' boundary of the D19 promoter, as defined by deletion analysis, is that it lies just within a G-rich element. The next deletion downstream, which fails to express, does not contain this sequence. Therefore, the CP1 68-mer may be able to substitute for this putative element, allowing the role of D19 sequences further downstream to be investigated, for example to determine if they confer prespore-specificity to the fusion construct.

Two oligonucleotides were synthesized; a 75-mer corresponding to the sequences between -112 and -182 (I), and an 84-mer corresponding to the sequence immediately upstream (II) (underlined in Figure 6.18). All of the C-rich region is spanned by the two oligomers. The CP1 oligonucleotide was originally inserted in both orientations in ΔBam.CAT. To resemble the proposed arrangement of the D19 promoter, the D19 oligonucleotides were cloned into the construct with a downstream BamHI site. Additionally, in this construct the orientation of the CP1 oligonucleotide is that found in the intact gene. The D19 oligonucleotides were designed to retain a single upstream BamHI site after cloning. Insertion in the correct orientation was confirmed by mapping. I and II were inserted singly, and then II was inserted upstream of I, the same relative arrangement as the native promoter (Figure 6.19).

On transformation into AX-2 cells and selection in the usual manner, pooled populations of about 50 clones resulted. The vector copy number ranged between 10 and 60 per cell for the three constructs (Figure 6.20b). Expression was measured by performing CAT assays with extracts harvested from growing cells, cells at eight hours of development (when CP1 is maximally expressed) and slugs (when D19 is maximally expressed). As the expression of the -122 deletion had been relatively inefficient, increased amounts of extract (10μg cf 2μg) were used to try to increase the chances of detecting the likely low expression levels. No expression was detected in any of the stable transformants (Figure 6.20a). Therefore, the addition of 159 nucleotides from the D19 promoter to an

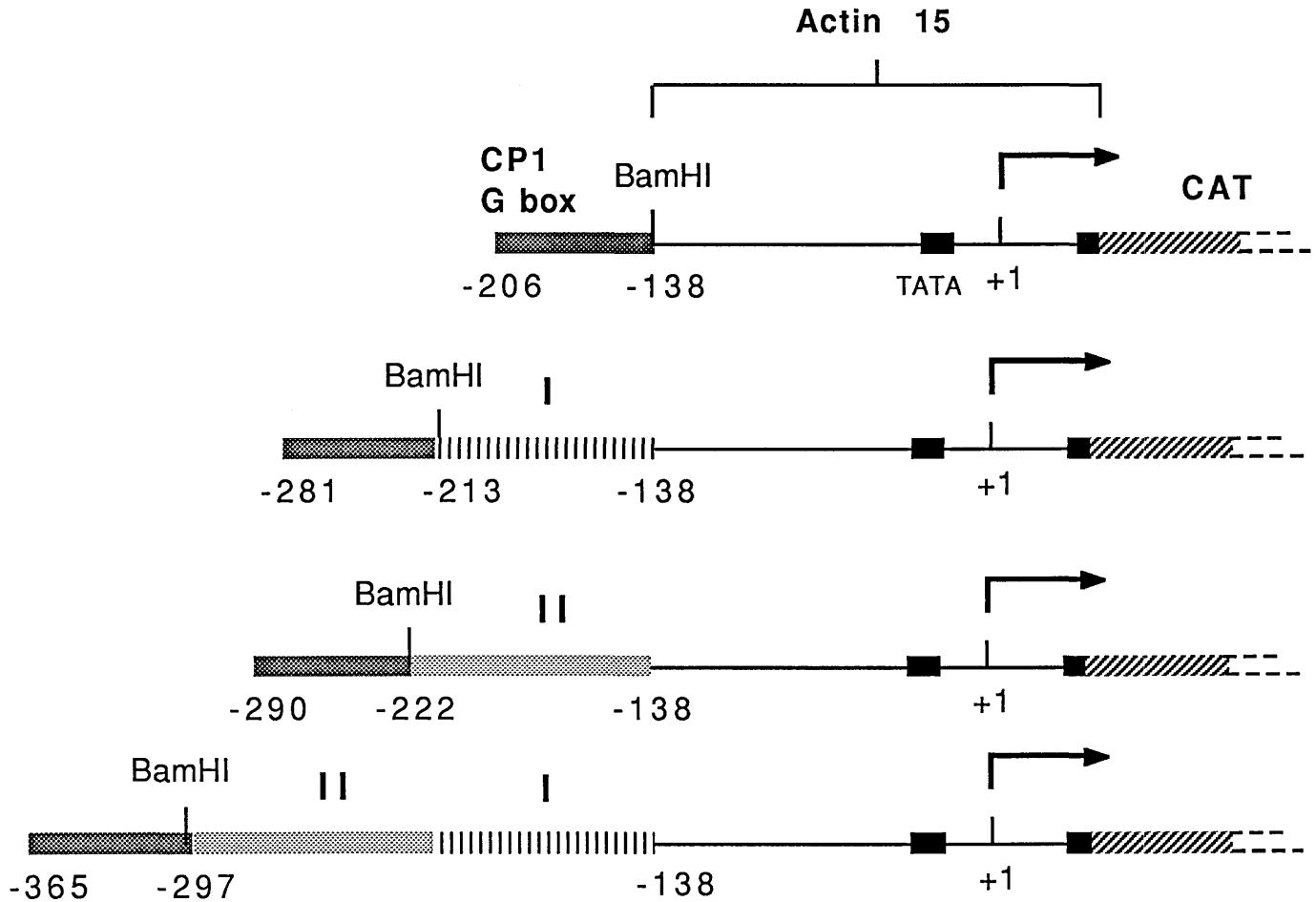
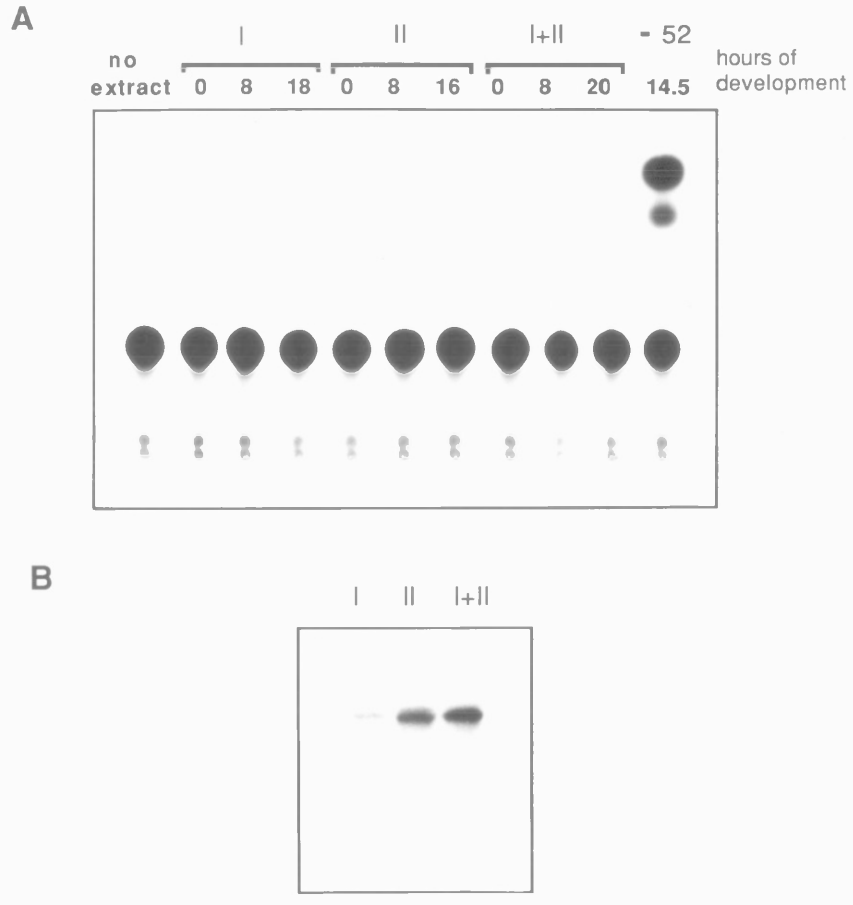


Figure 6.19 Relative arrangement of promoter elememnts in the oligonucleotide substitution constructs

The D19 synthetic oligonucleotides (I+II) were inserted into the BamHI site site 138 nucleotides upstream of the actin 15 cap site (+1).

Figure 6.20 Expression of the oligonucleotide substitution constructs

A) CAT assays were performed with 2µg of protein extracted from pooled populations of cells transformed with each of the three oligonucleotide substitution constructs; containing the downstream sequence alone (I), the upstream sequence alone (II), or both oligonucleotides (I+II).
The extracts were harvested from growing cells, from cells 8 hours after the initiation of development on filters, and from slugs. An equal amount of extract from growing cells transformed with the actin 15/CAT construct was assayed in parallel as a positive control.
B) 50µg samples of total nucleic acid extracted with phenol from the same transformants examined above were analysed by Southern transfer. The DNA was linearised with XbaI, and after electrophoresis and blotting as previously described, the filter was probed with pB10TP2 and washed as before.


A
I
II
I+II
- 52
no extract
0 8 18 0 8 16 0 8 20 14.5
hours of development
B
I II I+II

68bp region of CP1 was insufficient to reconstitute promoter activity, showing either the expression pattern of D19 or CP1. It may be that any induction which occurred was below the level of sensitivity of detection of the assay. Alternatively it is possible that although the G-rich boxes of CP1 and CP2 are functionally similar, one to another, they cannot work in combination with a gene with a divergent pattern of mRNA accumulation (see Chapter 8).

6.7 Summary.

1) Deletion from the 5' end of the intergenic sequences upstream of the D19 gene results in a reduced level of expression, but 338 nucleotides of flanking DNA direct correctly regulated expression.

2) The 122 nucleotides immediately upstream of the cap site can be removed without abolishing regulated expression, albeit at a reduced level.

3) Construction of fusion genes in which the D19 promoter is placed upstream of a heterologous gene show that transcriptional mechanisms are sufficient for correct timing and cell-type specificity of expression.

4) D19 upstream sequences can work in conjunction with the TATA box of a heterologous gene. However, a 159 bp region of D19 and a 68 bp region of the CP1 promoter containing a G-rich element do not combine to form an active promoter.

CHAPTER 7 : Expression of D19 late in development.

The expression of the D19 gene is at a peak during the migrating slug stage of development, and declines as culmination proceeds (Oyama and Blumberg, 1986a; K.Jermyn, personal communication; see Chapter 5). In mature fruiting bodies, the amount of D19 mRNA detected is about 10-20% of maximal levels. Although the D19 transcript is known to be highly enriched in prespore cells (Barklis and Lodish, 1983), the localisation of D19 mRNA in culminants has not been previously investigated.

The localisation of PsA, the D19 gene product, has been established with the monoclonal antibody MUD-1. MUD-1 (Gregg et al., 1982) recognizes an epitope specific to PsA (Krefft et al., 1983; Grant et al., 1985), and has been utilised for immunoblotting, staining of frozen sections and fluorescence-activated cell sorting (FACS) (Krefft et al., 1983, 1984 1985). The MUD-1 antigen is undetectable in vegetative amoebae (Krefft et al., 1983). It is first faintly observed at tip formation, which correlates well with the timing of cytoplasmic mRNA appearance (Barklis and Lodish, 1983, Oyama and Blumberg, 1986a). The antigen appears within cells 1-2 hours before surface expression, after which the number of molecules of antigen/prespore cell increases until culmination, when the cell surface concentration declines rapidly (Krefft et al., 1984). There is a major loss of cell surface labelling in the transition from prespore to spore

cells, and PsA was found to be in a membrane-bound state within the spore coat (Krefft et al., 1984). Some preliminary data that surprisingly PsA is present in stalk cells was also reported (Williams et al., 1984).

The time of appearance and distribution of the protein sp29, now known to be PsA, has also been examined (Morrissey et al., 1984). By ^{35}S-methionine incorporation, followed by silver staining and autofluorography of 2D-gels, the synthesis of PsA could first be detected after tip formation at about 14 hours of development. The peak of PsA accumulation occurs at about 20 hours, and the protein is only faintly detectable in preculminants and culminants. Of 20 prespore proteins detected, PsA was considered to be one of 8 which does not persist into spores. In these experiments, no PsA protein was detected in isolated stalks. However, it is likely that these detection techniques were less sensitive than the use of MUD-1, since PsA was detected earlier in development with antibody staining, at tip formation (Krefft et al., 1984).

7.1 The distribution of D19 mRNA in mature fruiting bodies.

The localisation of D19 mRNA in AX-2 culminants was determined by Northern transfer analysis of total nucleic acid extracted from separated spore and stalk cells (K.Jermyn, personal communication; Figure 7.1). Spores are easily purified since they are more resistant to lysis with the anionic detergent NP40 than other cell types. After detergent pretreatment, frozen spores were lysed by grinding

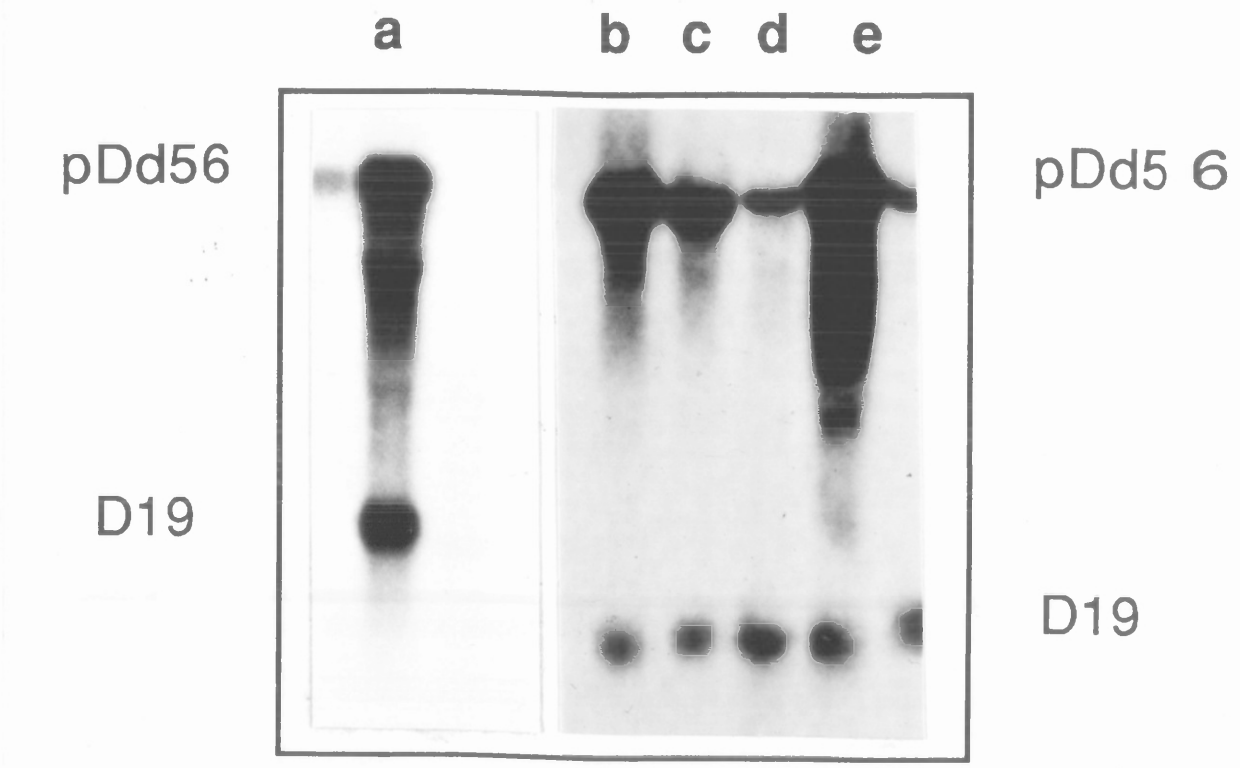


Figure 7.1 Expression of the D19 gene in stalks detected by Northern transfer

Two Northern blots are shown, both probed with a mixture of the D19 and pDd56 cDNA clones and washed at 0.1x SSC, 0.1% SDS for 4x 15' at 65°C. The 5µg total RNA samples are derived from the following sources; a) AX-2 stalks (left) and spores (right); b) V12M2 monolayer stalks, c) NC4 "plucked" stalks, d) HL31 stalks after 24 hours of development and e) HL31 stalks after 48 hours. The samples were analysed on a 1% formaldehyde gel.

in a pestle and mortar. Stalks were prepared by extensive filtering of harvested culminants through Nitex. The stalks are retained while the spores and additional immature cells pass through; any remaining spores do not lyse. Surprisingly, the D19 cDNA clone used as a probe detected hybridising transcripts in RNA from stalk cells but not from spore cells. As this was an unexpected result for a gene encoding highly prespore-enriched mRNA, further RNA samples were then analysed (K.Jermyn, personal communication). The RNA tested was extracted from the following sources; V12M2 cells in low density monolayer culture after completion of stalk cell formation, NC4 plucked stalks, and HL31 stalks. V12M2 and NC4 are wild-type strains (Town et al., 1976; Bonner and Slifkin, 1949), while HL31 is a "stalky" mutant (Morrissey et al., 1981). The mature NC4 fruiting bodies were individually plucked fron the surface of the KK_2 agar plates to avoid harvesting any immature aggregates, and close to 100% stalk cell formation occurs when V12M2 cells are incubated in monolayer culture (K.Jermyn, personal communication). Thus, all three samples were chosen in an attempt to maximise the purity of the stalk cell preparation. In each case a single band of identical molecular weight resulted from use of the D19 cDNA clone as a probe of a Northern blot (Figure 7.1). The signal was retained upon washing at high stringency (0.1x SSC, 65°C). If the hybridising band is derived from a distinct gene it must be closely related to D19 (>98% using the formula $Tm = [16.4\log(Na^+) + 0.41GC + 81.5]^\circ C$).

The idea of a related but distinct stalk-specific message was not supported by primer extension analysis. Stalk and

spore mRNA from AX-2 cells was annealed to the primer originally used to establish the 5' end of authentic D19 mRNA (see section 3.5; Figure 7.2). A single extended product resulted with the stalk RNA sample alone. No bands were obtained with the spore RNA sample, consistent with the results of Northern transfer analysis. The size of the extended product was identical for both the AX-2 stalk and V12M2 prespore samples. For a distinct message it is more likely either that the primer would fail to anneal or that the relative position of the start-site would be altered. If the mRNA species in stalk cells is derived from the D19 gene, the presence of a single band indicates that transcription is not occurring from a separate, stalk-specific, promoter.

7.2 The expression of a transformed copy of the D19 gene in mature culminants.

The expression of an introduced copy of the D19 gene in mature fruiting bodies was investigated. Because the gene was marked with an oligonucleotide encoding an epitope of the c-myc protein, its expression can be distinguished from that of the endogenous gene (see Chapter 5). Marked trancripts are specifically detected by Northern blot analysis using the anti-coding strand of the oligonucleotide as a probe. The construct used, pB10B1, contains 995 nucleotides of sequence upstream of the start-site of transcription, which is sufficient for correctly regulated expression earlier in development (see Chapter 5).

RNA was isolated from separated stalk and spore cells at 24 hours of development, and from stalks 12 hours later.

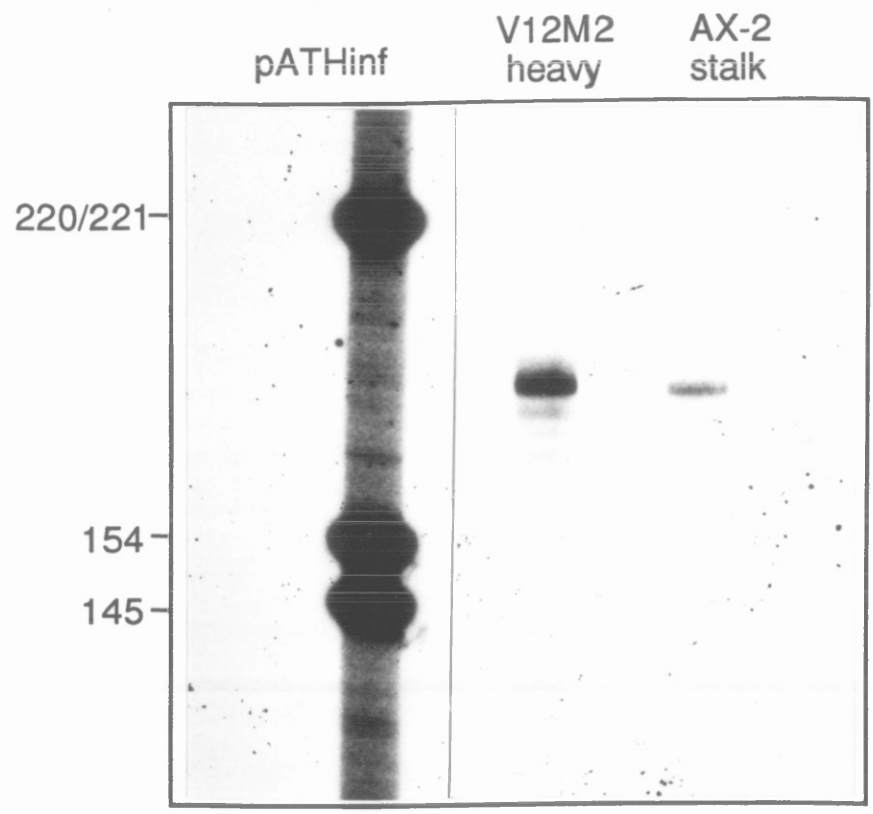


Figure 7.2 Determination of the 5' terminus of the D19 stalk message

Primer extension analysis was carried out as described in the legend to Figure 3.7, except that the RNA samples compared were extracted from stalks of AX-2 cells or prespore cells from V12M2 slugs (heavy fraction of a Percoll density gradient).

Probing a Northern blot with the c-myc oligonucleotide revealed the same distribution of D19 expression previously seen with non-transformed cells (Figure 7.3). The expression in spores was so much reduced compared to both 24 hour and 36 hour culminants that a meaningful value for the stalk-spore ratio could not be determined by scanning densitometry. These results strongly suggest that the transcript previously identified is derived from the D19 gene.

A problem with RNA analysis using this approach is that the yield of RNA from stalk cells seems to be lower than from cells earlier in development and from spores themselves. This is expected since stalks are composed of terminally differentiated, predominantly dead cells (reviewed by Bonner, 1982) in contrast to spore cells, which have a potential for rapid germination (Giri and Ennis, 1978). As equal masses of RNA were loaded onto the gel, the relative mRNA levels detected do not reflect the ratio per cell. Therefore, any finite level of stalk expression will have been enriched.

Also, because of this comparatively low RNA content, stalk preparations are highly susceptible to contamination with RNA from cells at earlier stage of development. The second stalk preparation harvested at 36 hours was included to try to minimise the extent of this contamination. However, the purity of the stalk preparations could not be ascertained, as the proportion of immature cells remaining was unknown. This caveat also applies to the experiments described in section 7.1.

In spite of these criticisms, the hybridisation signals obtained with stalk RNA still seem comparatively intense. An

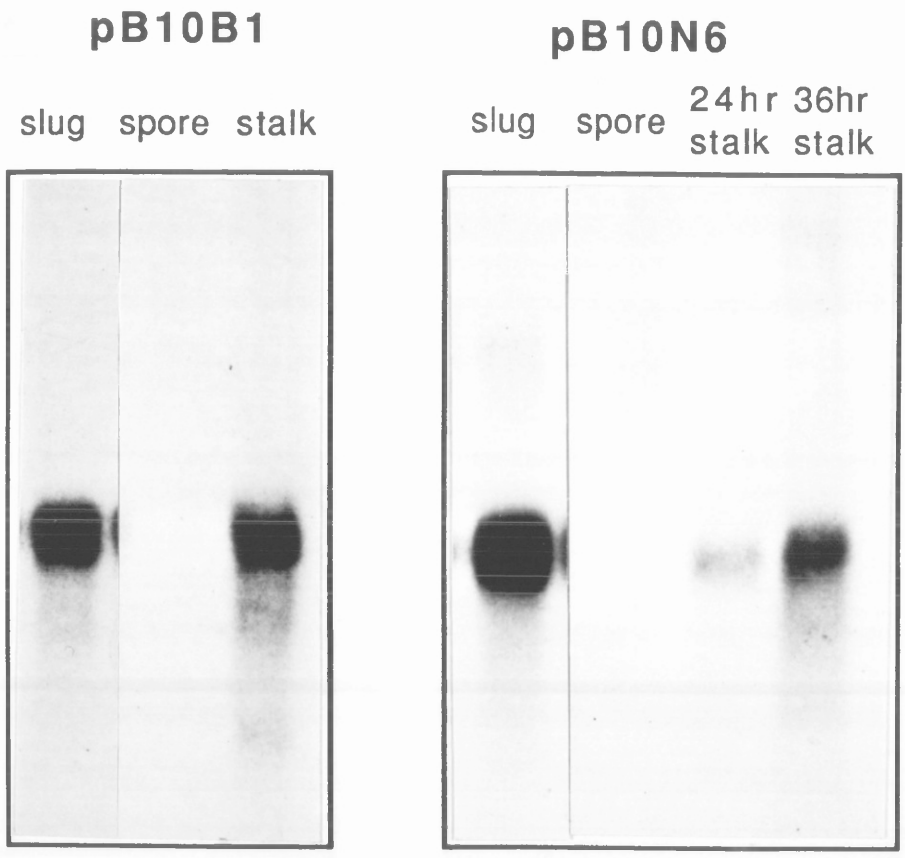


Figure 7.3 Expression of the marked D19 gene in culminants at the RNA level

Total nucleic acid was extracted from transformants, expressing either the B1 or N6 version of the c-myc-marked, D19 gene, at the developmental stages shown. 5µg of each sample was analysed by Northern transfer as previously described, except that the filter was probed with the non-coding strand of the c-myc oligonucleotide at 32°C in hybridisation buffer containing 6x SSC. The filter was washed in 5x SSC, 0.1% SDS at 42°C for 4x 5'.

indication of this signal strength is provided by side-by-side comparison with the developmental timepoint when maximal D19 expression is occurring (Figure 7.4). An equal amount of RNA from the same timecourse was used. However, a more definitive conclusion is that as prespore cells differentiate into spore cells, expression of the D19 gene ceases. As PsA levels are known to be very highly reduced in in culminants over slugs (Morrissey et al., 1984), and staining of spores with MUD-1 is not detected (Krefft et al., 1984), it is consistent that the level of D19 mRNA in spores should be low. Unless translational control is operating, a decline in mRNA levels should precede that of protein levels.

7.3 Expression of other prespore genes during culmination.

Three other highly prespore-enriched genes were tested to determine whether apparent expression in stalks is a unique property of D19. D7 and PL3 are cDNA clones that have an expression pattern similar to D19 during development and also show a similar response to cAMP levels (Oyama and Blumberg 1986a+c). SP96 is a cDNA clone derived from a major spore coat protein (Dowds and Loomis, 1984). The SP96 protein is reported to be present in spore cells and absent from stalk cells (Devine et al., 1982). The cDNA clones were used to probe Northern blots of RNA extracted from separated spore and stalk cells as before. As SP96 is a known spore coat protein, it was initially used as a control to test the integrity of the spore RNA. However, for all three genes there was an absence of detectable spore transcripts, and the

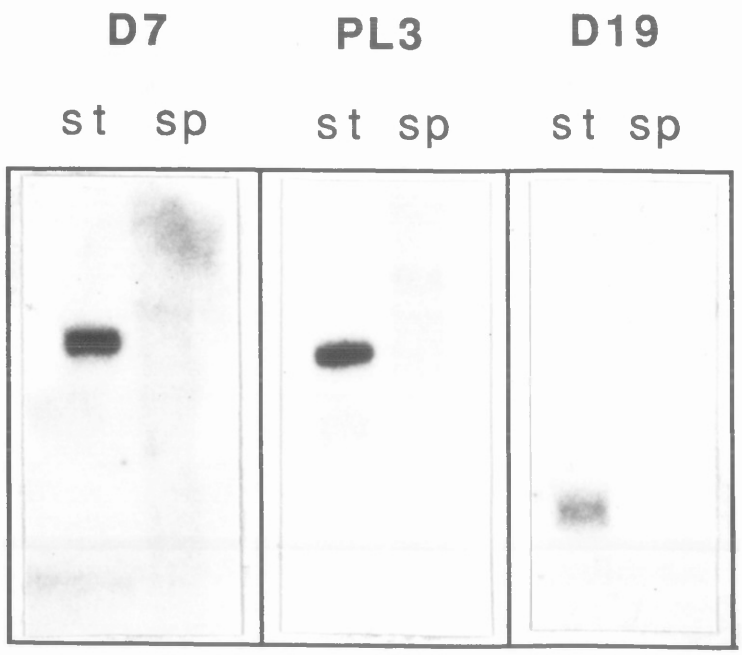


Figure 7.4 Expression of other prespore genes in mature culminants

Duplicate 5µg samples of total nucleic acid from Ax-2 stalks and spores were analysed by Northern transfer, and were first probed separately with the D7, PL3 and D19 cDNA clones using the conditions previously described. The three filters were then reprobed with the SP96 cDNA clone. The hybridisation pattern which resulted was indistinguishable from that given by PL3 (data not shown); on the PL3 filter, the two bands were superimposed.

presence of a single hybridising transcript in stalk RNA. The SP96 and PL3 transcripts comigrate and may be derived from the same gene.

Thus the paradoxical result that mRNA derived from prespore-enriched genes is present in stalk cells and not in spore cells is strongly suggested. The absence from spores is more easily rationalised, since at the protein level many prespore-specific products are absent from spores (Borth and Ratner, 1983; Morrissey et al., 1984). It is possible that even genes encoding spore cell components are not required to be transcribed after spore cell formation. The absence of any residual level of SP96 transcript is unexpected however, as SP96 protein is known to be at significant levels in spores (Dowds and Loomis, 1984). In order to check the integrity of the spore RNA, a gene expressed in germinating spores would be more suitable as a probe. This has not been carried out. The possible presence of prespore-specific transcripts in stalk cells requires investigation using an alternative approach, owing to the uncertainty about the purity of stalk RNA preparations.

7.4 Localisation of D19 protein in culminants

Fruiting bodies composed of cells transformed with the pB10B1 construct were used to investgate the distribution of the protein product of the marked gene. It was hoped that this would give a clearer idea of which cells were expressing D19, i.e. whether the protein was associated with mature, vacuolated stalk cells or other, immature cells. Frozen sections through culminants were stained with 9E10, the

monoclonal antibody which recognises the c-myc epitope (Evan et al., 1985). Sections of transformed culminants were analysed in parallel with samples of AX-2 fruiting bodies to rule out the possibility either that the staining was due to non-specific background or a result of binding to a related epitope on a distinct protein. The samples were then treated with gold-conjugated second antibodies, and examined by electron microscopy. An attempt to use a fluorescent second antibody for light microscopy was unsuccessful because of non-specific background.

With the AX-2 samples, no staining at all could be discerned, either with stalks or spores (Figure 7.5). In contrast, strong staining was detected for some of the transformant sections. The intensity of the staining varied, partly dependent on the nature of the material analysed. There was little or no staining of spores (data not shown), while the most intense staining was seen in the extracellular matrix surrounding stalk cells (Figure 7.5), although this staining was not uniform. No staining was seen within stalk cells, although most the intracellular space was filled with a large vacuole, possibly rendering detection of cytoplasmic components difficult.

As PsA, the D19 gene product, is a known cell surface protein (Grant et al., 1985), it is not surprising that little or none should be retained within cells. It is not known why the protein detected was not more closely associated with the cell membrane. It is not possible to conclude that the protein detected was actually produced by the adjoining stalk cells. The protein could have been

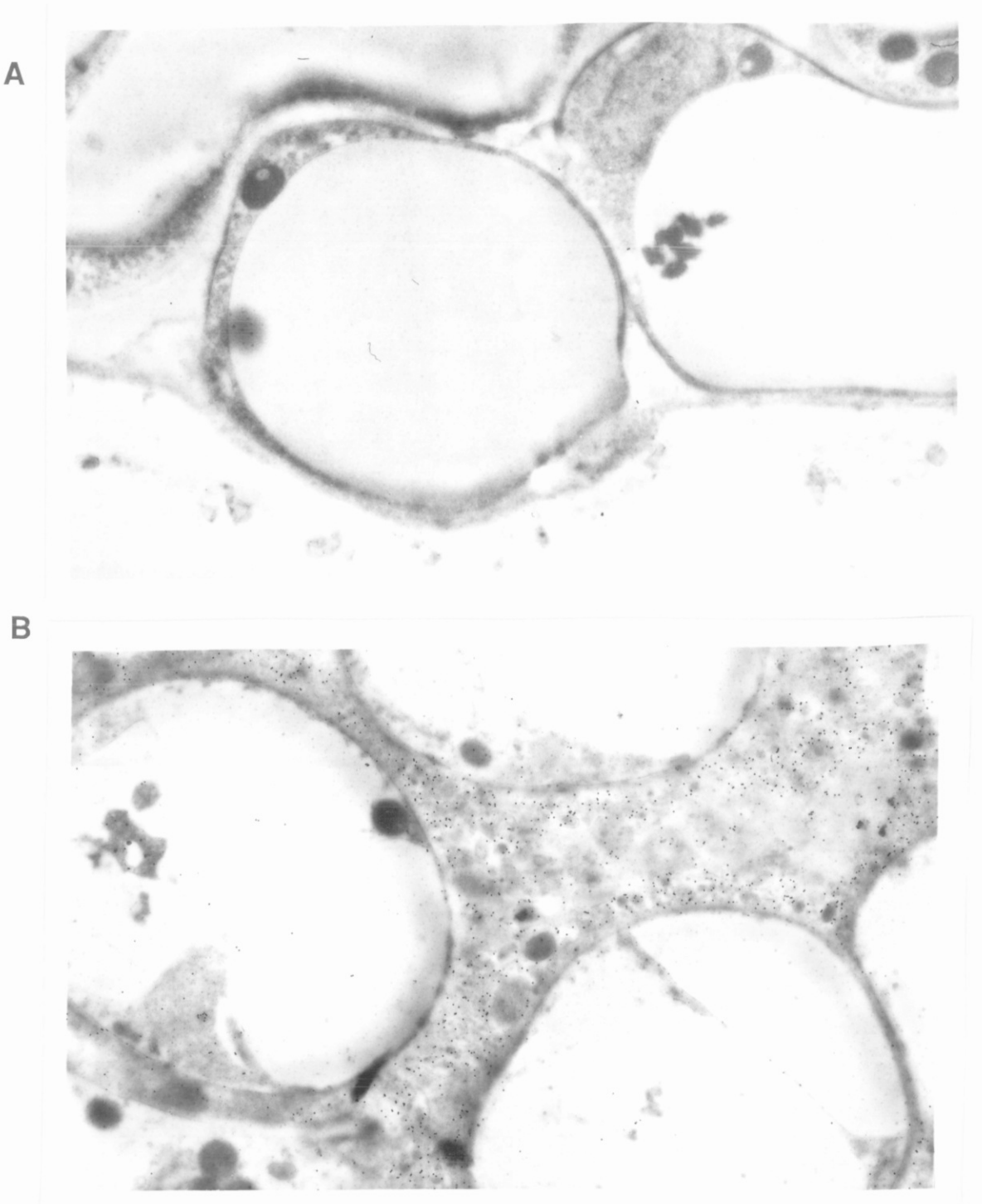


Figure 7.5 Visualisation of marked PsA in culminants by immunoelectron microscopy

Fruiting bodies of (A) AX-2 cells and (B) cells transformed with pB10B1 were analysed. Sections through mature culminants were probed with the anti-c-myc monoclonal 9E10 at a dilution of 1 in 100. Magnification, x 5930.

released into a part of the sheath at an earlier stage of development, which then became associated with stalk cells during culmination. Alternatively, if the marked protein had been synthesized by the cells with which it was associated, specific release from the cell surface may have occurred as the mature stalk cells were formed.

In order to be certain that D19 is expressed either in stalk cells, or in cells which later become stalk cells, the D19 promoter could be used to drive the expression of a gene encoding a protein retained by the cell. This approach was not attempted. However, even using such an approach, it would not be easy to distinguish between the possibilities that stalk cell expression is a result of specific reinduction of D19, or transdifferentiation of prespore cells previously expressing D19.

7.5 Summary

1) The residual level of D19 mRNA detected in culminants seems to be present in stalks rather than spores.

2) Analysis of cells transformed with a marked copy of the D19 gene shows that the stalk-specific mRNA is not the product of a related gene.

3) The expression of three other prespore genes tested is detected in stalk cells but not in spore cells at the RNA level.

4) The protein product of the transformed D19 gene is localised in the extracellular matrix surrounding stalks.

CHAPTER 8: DISCUSSION

The isolation and characterisation of cell-type-specific genes is of crucial importance in understanding the molecular mechanisms underlying cellular differentiation. These markers can be used not only to allow identification of cell types and to follow the generation and maintenance of pattern, but also to investigate how the signals responsible for differentiation act to modify the expression of individual genes. Ultimately by fully defining the chain of events from the production of a morphogenetic signal to changes in gene expression, it will be understood how distinct cell types are produced from a single progenitor.

The cellular slime mould *Dictyostelium discoideum* is ideally suited to studying this problem. Only two major cell types arise during development, reducing the number of signals and cellular interactions likely to be involved in comparison to species with a more complex organisation. Unlike other systems, the signals responsible for cellular differentiation have been well characterised. Several potential morphogens with a known structure have been identified. The two major cell types can be separated, and this has enabled the isolation of genes expressed predominantly on one or other pathway of differentiation. Evidence is now accumulating that substances known to promote cellular differentiation act directly to modify the expression of cell-type-specific genes.

This thesis describes the characterisation of a gene, D19, expressed exclusively in prespore cells, and therefore a marker of the prespore pathway of differentiation. The mechanism by which a known morphogen acts to modify the expression of the gene has been determined, and the transformation system has been used to localise the regulatory signals within the gene required for cell-type-specific expression.

8.1 The structure and function of the encoded polypeptide

Comparison of the predicted amino-acid sequence of the D19 polypeptide with the results of peptide sequencing has established beyond doubt that the D19 gene encodes the prespore-specific cell-surface glycoprotein PsA (Grant et al., 1985), also known as sp29 (West and Loomis, 1985). PsA was first identified as a antigenic determinant present on the surface of prespore cells and in the slime sheath, detected by the monoclonal antibody MUD-1 (Grant and Williams, 1983). The epitope recognised by MUD-1 is unique to PsA, but the nature of the determinant is not known (Krefft et al., 1983). Although there are variations in the size of the protein detected by MUD-1 in different strains, there is genetic evidence that this represents polymorphism at a single locus. Segregant haploids, derived from diploids constructed between strains containing different sized proteins, contain only one form of the protein (Grant et al., 1985). The locus was designated pspA, and assigned to linkage

group 1. This classification was confirmed by mitotic recombination experiments.

A closely related non-polymorphic protein, present only in the sheath, was also identified (ShA; Grant et al., 1985). Both PsA and ShA are recognised by the monoclonal antibody MUD-50, which recognises a carbohydrate determinant, and hence they are probably processed similarly, although ShA additionally carries other unique determinants. It was suggested that as PsA and ShA are also both developmentally regulated and have similar molecular masses and pI values, they could be structurally related.

The above data can be correlated with the results of Southern blotting analysis, using the D19 gene as a probe. D19 is clearly a single copy gene; at high stringency, hybridisation to one predominant band only is generally seen, depending on the restriction enzyme employed. However, at a lower stringency, multiple bands are obtained, evidence of the existence of genes with a related structure. One of these bands may be the result of cross-hybridisation to ShA. As PsA and ShA, which may have arisen by duplication of a common ancestor, are coregulated, it would be of interest to compare 5' flanking sequences. For other genes this has been used as an approach to identify important, conserved regulatory sequences (Poole and Firtel, 1984; Romans and Firtel, 1985, Pears and Williams, 1988). However, the gene encoding ShA has not been isolated to date.

Determination of the nucleotide sequence spanning the complete coding region of D19 has provided information about the structure and processing of the PsA protein. Firstly, it

has suggested a possible mechanism for the generation of the polymorphism. Two variant forms of the protein were identified in 6 out of 21 different wild isolates tested, having a regular increase or decrease of about 2kD in apparent molecular mass (Grant et al., 1985). Towards the C terminus of the protein encoded by D19 are four copies of a four amino acid repeat, pro thr val thr. Unequal crossing-over at the repeats would lead to the loss or gain of up to three copies, which could account for 2kD differences if this region of the protein was subject to post-translational modification. There is biochemical evidence that PsA is modified by O-linked glycan, but the extent of acylation, phosphorylation or sulphation has not been assessed (E.Smith and K.Williams, personal communication). The O-glycosylation is thought to be mostly linked to threonines in the repeat region for two reasons; the failure to detect many of these residues by peptide sequencing, implying some form of modification exists, and the susceptibility of modB mutants, which lack O-linked glycan due to a glycosylation defect (Murray et al., 1984), to loss of PsA from the cell surface. As the repeat region of PsA is predicted to have an extended structure, it could be imagined how the absence of bulky carbohydrate groups could enhance proteolysis by reducing steric protection.

The prediction of a hydrophobic leader peptide is consistent with the known cell-surface localisation of PsA (Gregg et al., 1982; Krefft et al., 1985). However, the amino acid sequence at the C terminus does not indicate the presence of a membrane-spanning domain, followed by a

cytoplasmic polypeptide anchor. Instead, the characteristics of the C terminus resemble those of a set of proteins known to be linked to the plasma membrane via a glycolipid known as a GPI link or anchor (reviewed by Cross, 1987; Ferguson and Williams, 1988). Analysis of different GPI anchors has revealed that they share common structural features, including linkage to the carboxyl group of the terminal amino acid via ethanolamine phosphate, the presence of phosphatidylinositol and a glycan bridging these two components (Figure 8.1). The proteins attached also have shared features: the lack of a cytoplasmic domain and the presence of a hydrophobic tail containing hydrophilic serine and threonine residues at the extreme C terminus. Where the sequence of the mature protein has been compared to that of its cognate gene, it is apparent that the hydrophobic portion is cleaved off prior to the covalent attachment of the GPI anchor to the new terminal amino acid (Boothroyd et al., 1980; Tse et al., 1985; Hemperley et al., 1986). Peptide sequence of the C terminus of the mature PsA protein has shown that 21 amino acids are missing (A.Gooley and K.Williams, personal communication). The cleavage position lies between a serine and an alanine residue, resulting in removal of most of the hydrophobic domain. For other proteins, this region seems to act as a processing signal and has no apparent primary sequence conservation, analogous to the signal peptide. Fusion of the terminal 37 amino acids of a GPI-linked protein, decay accelerating factor, to the C terminus of a secreted protein was sufficient to direct attachment of a GPI-anchor (Caras et al., 1987).

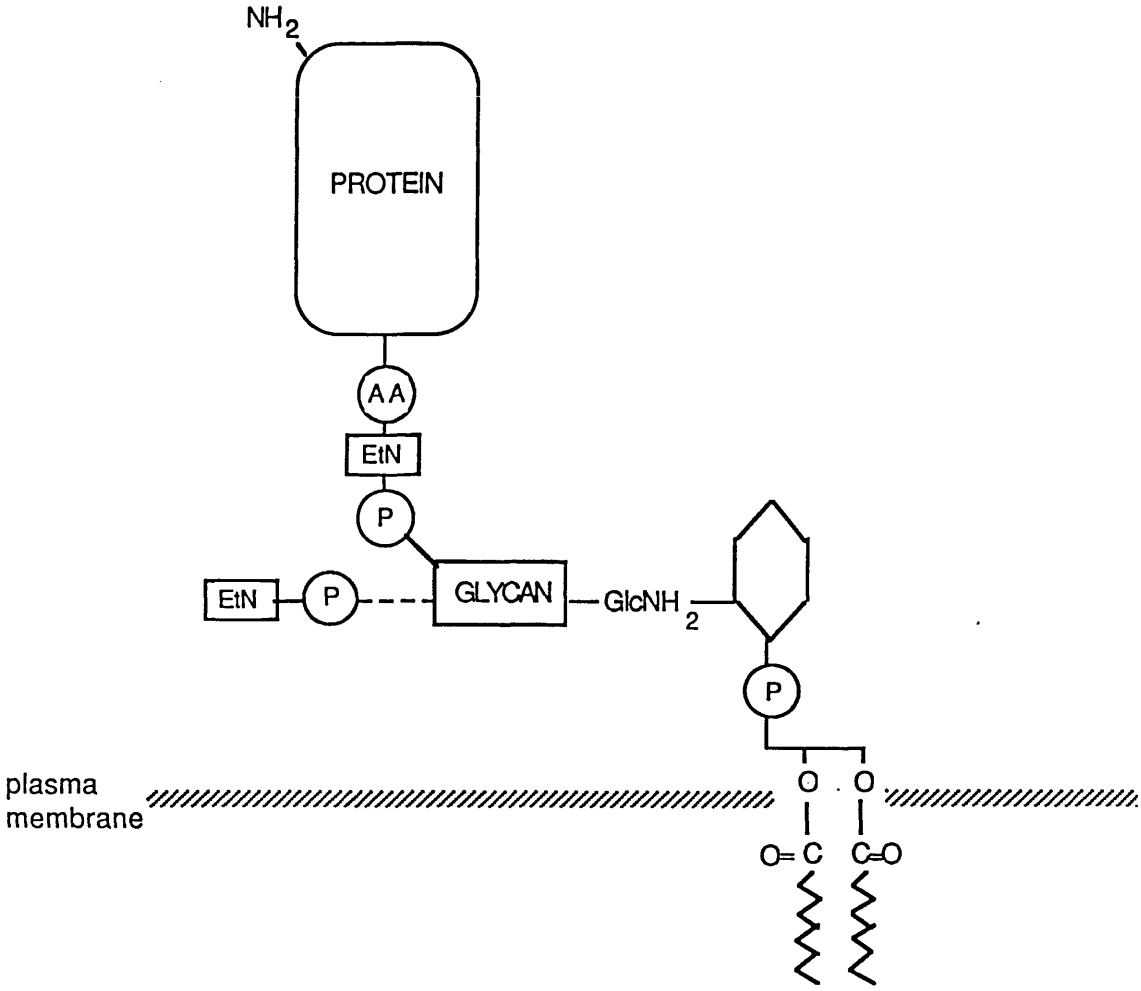


Figure 8.1 Common structural features of GPI-anchors

AA identifies the terminal amino acid; EtN, ethanolamine and P, phosphate. The ethanolamine side-chain indicated by a dotted connecting line is not found in all structures.

Adapted from Ferguson and Williams, 1988.

There has been much speculation about the implications of protein attachment via GPI anchors. It was originally thought that it may be a primitive form of attachment, analogous to the linkage of bacterial proteins to membrane lipids, and retained for a few eukaryotic proteins (Tse et al., 1985). However, it has become clear that it is a widespread mechanism in eukaryotes, and hence may serve different functions depending on the individual protein.

Many of the proteins with GPI anchors so far identified can be grouped into a few major functional classes. Several GPI-linked proteins may have a role in transmembrane signalling, such as the T cell proteins Thy-1 and TAP, involved in T cell activation (Low and Kincade, 1986; Reiser et al., 1986). Defects in the production of GPI-linked proteins lead to deficiencies in T cell activation which can be overcome with calcium ionophores and phorbol esters, but a direct involvement of the GPI linkage itself is not proven. The GPI anchor could be a source of a number of biologically active mediators, such as diacylglycerol, phosphatidic acid or IP3. Recently it has been shown that T cells contain a structurally related set of GPI molecules, at least one of which is specifically degraded in response to insulin and may function as a second messenger of hormone action (Gaulton et al., 1988).

If plasma membrane proteins are free from direct interactions with the cytoskeleton, lateral mobility could be facilitated, which could be important in allowing the rapid cross-linking implicated in signal transduction. Measurements on Thy-1, alkaline phosphatase and decay accelerating factor

have shown that these proteins are freely diffusing at rates comparable to those of lipid probes (Ishihara et al., 1987; Noda et al., 1987; Thomas et al., 1987). However, at least one GPI-anchored protein, PH-20, a sperm membrane protein, has a lateral diffusion rate more than a thousand times slower than lipid (Phelps et al., 1988).

Lateral mobility is likely to be irrelevant to the densely-packed trypanosome variant surface glycoproteins (VSGs), the first examples of GPI-linked proteins to be characterised. Instead it was suggested that the cleavage of the GPI link by enzymic action may permit specific protein release from the plasma membrane during the life cycle of the parasite, or during antigenic variation (Low et al., 1986). There is as yet no evidence for this, or for the cleavage of the GPI moiety of potential signalling molecules to generate second messengers. However there is evidence that a developmentally-regulated endogenous proteolytic activity controls the release of a GPI-linked membrane protein from Plasmodium falciparum. Specific release can be monitored because the soluble form only has proteolytic activity, also induced by addition of PI-specific phospholipase C (Braun-Breton et al., 1988).

Another class of proteins commonly linked to the plasma membrane via a GPI group are cell adhesion molecules. Two well-characterised examples are the chicken neural cell adhesion molecule (N-CAM; Cummingham et al., 1987) and LF3-A, a cell-surface glycoprotein which interacts with the T lymphocyte CD2 membrane glycoprotein (Dustin et al., 1987; Seed, 1987). Both proteins exist in distinct forms, attached

either by a GPI-linkage or by a polypeptide membrane spanning domain. For N-CAM, alternative forms are generated by differential splicing, but the two forms of LF3-A are derived from distinct biosynthetic precursors. The existence of different forms of attachment for the same protein suggests the GPI anchor has a specific function. For instance, increased lateral mobility could facilitate diffusion to sites of adhesion. An asymmetric cell surface distribution of cell adhesion molecules is often observed (Edelman, 1986). Specific release of GPI-linked proteins could cause rapid breakage of cell-cell bonds, and in the extracellular milieu the released protein could act as a competitive inhibitor of cell-cell interactions (He et al., 1986).

Although there is as yet no biochemical evidence for PsA attachment via a GPI link, it has recently been proven for two other *Dictyostelium* cell-surface proteins, contact sites A (CsA; G.Gerisch, personal communication) and antigen 117 (Sadeghi et al., 1988). As bacterial phospholipase C, used as a diagnostic test in other systems, did not release either protein from the cell surface, evidence was provided by the use of radiolabelled compounds known to be components of other GPI linkages. Both CsA and antigen 117 are thought to be cell adhesion molecules. Antigen 117 is suggested to have a function in cell-cell adhesion during the aggregation stage (Brodie et al., 1983), while CsA is among the best characterised cell adhesion molecules in any biological system (reviewed by Loomis et al., 1987).

Cells begin to have cohesive properties two hours after the initiation of development. During early development

dissociated cells will clump together only in the absence of EDTA, while at later stages EDTA does not inhibit clumping (Gerisch, 1961). The membrane components involved in early adhesion have not been unequivocally identified, although there is evidence that a 24kD glycoprotein is involved (Knecht et al., 1987). The EDTA-resistant adhesion mechanism, which appears after 8 hours of development, is mediated by CsA. It was shown that a surface glycoprotein of 80kD could neutralize antisera which block adhesion of developed cells (Muller and Gerisch, 1978). A mutant strain which did not form EDTA stable contacts specifically lacked the CsA glycoprotein (Noegel et al., 1985). Type 1 glycosylation, which is N-linked and cotranslational, rather than type 2 glycosylation, which is post-translational, is involved in the formation of EDTA-stable contacts (Ochai et al., 1982; Hohmann et al., 1987).

The gene encoding CsA has been isolated and sequenced (Noegel et al., 1986). Comparing the predicted amino sequence of the extreme C termini of CsA and PsA, a very similar structure is found, both in terms of overall chemical composition and primary sequence. The extreme C terminus of CsA is also hydrophobic with interspersed hydrophilic residues, and of the last 22 amino acids, 10 are identical to those present in PsA. This is not only good supportive evidence for the attachment of PsA by a GPI anchor, but also suggests that the two proteins may be evolutionarily related (see below).

As it is not known whether any *Dictyostelium* proteins are attached to the membrane by both polypeptide and glycolipid

linkages, there is no compelling evidence that GPI anchors have a specific function. The attachment of CsA via a GPI group obviously cannot provide flexibility as it is postulated to do for N-CAM (Cunningham et al., 1987). However, for antigen 117, accumulation of released protein can be detected in the extracellular medium during development (Sadeghi et al., 1988). This release seems to involve the removal of attached fatty acid, and occurs in the presence of intact cells. Thus it has been suggested that a specific phospholipase activity is responsible. It is tempting to speculate that the potential for rapid release of PsA could be utilised during development; for instance the conversion of prespore cells to prestalk cells during regulation, to stalk cells during transdifferentiation, to spore cells during normal development or to preaggregative cells during disruption of multicellular aggregates. During development cells are known to have a capacity for rapid dedifferentiation (Soll and Finney, 1986), and modulation of prespore mRNA levels at least occurs in response to disaggregation (Barklis and Lodish, 1983; Chisholm et al., 1984). DIF is a likely signal for the conversion of prespore to prestalk or stalk cells (Kay and Jermyn, 1983). However, addition of DIF to HM44 cells does not reduce the pre-existing level of cell-associated PsA, although mRNA levels are significantly reduced (see Chapter 4). The rapid loss of labelling from the surface of prespore cells is reported to be accompanied by the appearance of staining intracellular vesicles.

The homology between CsA and PsA at the hydrophobic C terminus is also shared by the GPI-linked form of the neural cell adhesion molecule (Hemperley et al., 1986). As a consensus sequence cannot be identified for GPI-linked proteins, such direct homology is unlikely to be necessary for processing. The homology between PsA and CsA extends upstream of the extreme C terminus and hence upstream of the PsA cleavage site, spanning a proline-threonine-rich repeat region. For CsA, this region is proposed to form a "stalk" to project the protein into the extracellular space (Noegel et al., 1986). The analogous location in PsA suggests an analogous function, extremely unlikely to be related to GPI attachment. Thus the primary sequence homology between N-CAM, CsA and PsA may be evidence for a functional and/or evolutionary relationship. Although there is no significant homology of other region of PsA to CsA or N-CAM, a distinct structure is necessary property of a protein mediating a specific interaction.

Adhesion mechanisms during the multicellular stage of *Dictyostelium* may be cell-type-specific. Molecules involved in intercellular adhesion in other systems often have a distribution restricted to certain tissue types, and may be involved in the regulation of pattern formation (reviewed by Edelman, 1986). It has been suggested that differential adhesion may be responsible for cell sorting during pattern formation in *D.discoideum*. Aggregates formed in suspension culture differentiate into prespore and prestalk cells, initially with a random distribution, but then sort out with the result that the prestalk cells surround the prespore

cells (Sternfield and Bonner, 1977; Forman and Garrod, 1977; Tasaka and Takeuchi, 1981). Attempts have been made to measure the cohesive properties of separated prestalk and prespore cells, and some differences were found in the effects of EDTA and the blocking ability of an antibody raised against a cell-surface glycoprotein, gp150 (Lam et al., 1981). However, there is good evidence that any cell sorting which occurs is at least partly mediated by differing chemotactic properties of prestalk and prespore cells (Wang et al., 1988a).

Post-aggregative adhesion mechanisms are little understood but are distinct from those at earlier stages of development. Cell-cell adhesion in migrating slugs and culminants is resistant to antisera obtained using membrane preparations from pre-aggregative cells, whereas antisera derived from cells at equivalent developmental stages is active (Steinemann and Parish, 1980; Wilcox and Sussman, 1981). Although a fraction enriched in a 95kD protein can neutralise the blocking activity of this sera, a direct role in cell-cell adhesion has not been proved. The third adhesion mechanism is also resistant to EDTA.

There is no experimental data to suggest that PsA is involved in cell adhesion. PsA is a small protein (30kD), while in general the well characterised cell adhesion molecules are large cell surface glycoproteins (Edelman, 1986). However, as previously described, a 24kD protein has been implicated in the *Dictyostelium* early cell adhesion mechanism (Knecht et al,. 1987). Although an antisera against

PsA exists, any ability to block cell-cell interactions has not been reported.

Whatever function PsA has, it must be specific to prespore cells and confined to the developmental stages following aggregation and preceding culmination. Using MUD-1, production of PsA is not detected until tipped aggregates have formed, and as culmination proceeds there is a rapid loss of cell surface labelling (Krefft et al., 1984). The level of D19 mRNA reaches a peak in the migrating slug, but declines sharply at culmination and is absent from spores. As detected by 2D SDS-PAGE, the protein is also absent from spores (Morrissey et al., 1984). It is thus extremely unlikely that PsA is a spore coat protein, particularly since earlier during development such proteins are stored in intracellular vesicles (Devine et al., 1983).

PsA is present in the extracellular matrix, as judged by staining of the slime sheath. This results in horseshoe-shaped prints which indicate regions of attachment of the slug to the substratum. The significance of these patterns is questionable, since they are obtained with antisera directed against many different *Dictyostelium* proteins, and they may simply be caused by deposition of lysed cells as the slug moves (McRobbie et al., in press). For some proteins such as ShA, staining of the sheath alone is seen, indicating an extracellular function (Williams et al., 1984). However, there is a possibility that strong staining of the sheath is obtained for PsA because of increased or specific release of protein, mediated by cleavage of the GPI-link. It may be that the membrane-bound and extracellular forms of PsA could have

distinct functions. No specific function has yet been assigned to any individual component of the sheath.

The generation of a mutant lacking PsA by use of gene disruption techniques is still the approach most likely to establish the function of the protein. If further attempts were to be made, vectors used successfully for this purpose with other genes could perhaps be utilized (De Lozanne and Spudich, 1987; Witke et al., 1987). If a mutant lacking PsA had been successfully obtained, it would have been interesting to have examined the size, pattern and migratory ability of slugs, all of which could be affected by a reduction in cohesivity or the absence of specific cell-cell interactions.

8.2 DIF and the control of cell-type-specific gene expression

The mechanism by which DIF modifies the expression of D19 and several other genes was analysed using the mutant strain, HM44 (Kopachik et al., 1983). This system was employed because HM44 cells are unable to synthesize DIF, but retain the capacity to respond to added DIF. Therefore by choosing to include or to omit DIF to cells in submerged monolayer culture, the specific effect on the expression of individual genes can be judged. One of the disadvantages of this system is that the precise gene defect in HM44 has not been identified. The DIF biosynthetic pathway has not been characterised, and it is possible that a pleiotropic gene has

been mutated. Alternatively, there could be a more than one mutation present. However, if either of these suggestions is correct, the effect on development must be a subtle one. Addition of DIF alone is required for apparently normal development to proceed, as judged by morphological and biochemical criteria (Kopachik et al., 1983). Wild-type cells, incubated at low density, do not have any additional requirements for stalk cell formation (Town et al., 1976).

Another potential problem is the difference between the situation in vivo and the culture conditions used. When HM44 cells are incubated in submerged monolayer culture in tissue culture dishes at a density of 2.5x 10^6/ml, some cAMP signalling obviously occurs as the cells stream and form small clumps. However, the conditions do not exactly reproduce those of multicellular aggregates, as each cell is only making limited contacts with other cells. The cells experience high levels of added cAMP, which may exceed the physiological range. Inclusion of high cAMP levels is necessary due to the action of extracellular phosphodiesterases during the ten hour preincubation required for the cells to develop DIF responsiveness. Thus there is a possibility that the culture conditions may modify or bypass normal control mechanisms. However, it is unlikely that the basic underlying mechanisms will be affected.

The only anomaly encountered in the experiments described here was the induction of an mRNA species on DIF addition which cross-hybridises to D19 at low stringency. As discussed in Chapter 4, it is almost certainly the product of a distinct gene. However, this messenger RNA cannot be detected

during normal development in wild-type strains. Efforts were made to isolate the gene responsible to enable comparison with the sequence of D19, which has a related structure but is regulated differently by DIF. These attempts were not successful. It is not known why this gene is not expressed in other strains, but it could be normally required at a different stage of the life-cycle or under environmental conditions not reproduced in the laboratory. If the level of DIF added exceeded physiological concentrations, normal repression could have been counteracted or bypassed.

In general, gene expression in HM44 reflects the known or proposed effects of DIF on cellular differentiation. Addition of DIF to wild-type cells incubated at low density induces stalk cell formation (Town et al., 1976). By differential screening of cDNA libraries, constructed with RNA isolated from DIF-induced HM44 cells, three distinct clones were isolated. They were subsequently found to encode prestalk and stalk-specific components of wild-type cells (Jermyn et al., 1987; McRobbie et al., 1988). During a stalk cell induction in HM44 cells, mRNAs derived from two of these genes, pDd63 and pDd56, accumulate rapidly in response to DIF.

DIF can also divert cells from the prespore to the prestalk pathway. In HM44 cells, addition of DIF prevents the synthesis of known prespore-specific proteins (Kopachik et al., 1985). The results presented here show that DIF causes a specific reduction in the levels of three prespore-specific messenger RNAs in HM44 cells. The three genes analysed, D19, PL3 and D7, have been reported to exhibit a comparable dependence on the presence of extracellular cAMP for their

expression, and an identical response to disruption of aggregates, i.e. a rapid loss of their cognate mRNAs. In their response to DIF they also apparently exhibit co-ordinate behavior, since levels of all three mRNA species decline with very similar kinetics.

The mechanism of this repression was determined for D19. The results of in vitro nuclear transcription and mRNA half-life experiments strongly suggest that DIF acts primarily at the level of transcription to mediate repression of D19. There is approximately a 50% reduction in the rate of nuclear transcription 90 minutes after the addition of DIF. The exact timing of the onset of the response is difficult to judge as the mRNA has a half-life of about 90 minutes.

Analogous nuclear transcription experiments indicate that the induction of pDd56 and 63 expression in HM44 cells is also controlled at the level of transcription. For pDd63, induction of transcription can be detected 15 minutes after DIF addition. Therefore DIF must be acting almost directly to induce transcription, for example causing post-translational modification of a pre-existing component. Even more directly, DIF may act by analogy with the mechanism of steroid hormone action, i.e. interaction with an intracellular receptor, which then binds to DNA (see below; reviewed by Yamamoto, 1985). As the timing of the onset of repression of D19 is less certain, the number of steps involved cannot be predicted as easily.

There is evidence that DIF and cAMP act antagonistically in NC4 cells to modify prespore gene expression. In aggregation competent NC4 cells incubated in monolayer

culture, DIF represses cAMP-receptor interactions at low cAMP concentrations (Wang et al., 1986). Hence the cAMP relay response, propagated by the binding of cAMP to the high affinity A form of the receptor, is reduced. However, there is no effect on the interaction of higher concentrations of cAMP with low affinity forms of the receptor. Measurements, made under conditions in which the activity of the extracellular phosphodiesterase is competitively inhibited, have shown that comparatively high levels of cAMP are required to induce prespore gene expression (Oyama and Blumberg, 1986b; Gomer et al., 1986). It has been suggested that either prespore gene expression requires high receptor occupancy, or involves interaction with a very low affinity form of the receptor. In the latter case this binding would be unaltered by DIF, since it is ineffective at high cAMP concentrations. Thus in multicellular aggregates, DIF would indirectly prevent the induction and maintenance of prespore gene expression by reducing cAMP relay and lowering the overall concentration of extracellular cAMP.

For most of the HM44 experiments described here, the cells were incubated in medium containing a saturating concentration of extracellular cAMP. As under these conditions DIF should not be interfering in the binding of cAMP to its receptor (Wang et al., 1986), DIF may have been acting at a post-receptor, intracellular stage to repress D19. Alternative mechanisms could be envisaged; DIF may interact with a step of the second messenger pathway leading from the cAMP receptor, or could act independently on the transcriptional machinery.

At lower cAMP concentrations, the failure to demonstrate any antagonism between DIF and cAMP in HM44 cells may be due to differences in sensitivity of HM44 and NC4 cells to DIF and/or cAMP. High levels of cAMP apparently counteract prespore repression by DIF in NC4 cells, and only by reducing cAMP levels does DIF become active (Wang et al., 1986). If HM44 cells are inherently more sensitive to DIF, maximal repression might occur (via a receptor-independent mechanism) even in the presence of high cAMP levels. Thus by reducing cAMP levels the extent of repression would not be potentiated. This model is supported by recent evidence that efficient stalk cell formation in low density monolayer cultures of NC4 cells requires removal of cAMP (Berks and Kay, 1988). It predicts that prespore repression in HM44 cells might show a greater dependence on cAMP if lower levels of DIF were used. It has also been observed that V12M2 cells are apparently less sensitive to DIF than HM44 (G.Weeks, personal communication).

A further argument supporting the idea that DIF is acting at a post cAMP receptor stage in HM44 cells is the lack of modulation of mRNA stability. The half-life of post-aggregative mRNA species in AX-3 aggregates has been estimated to be 3-4 hours, but on disaggregation it is turned over very rapidly; the accompanying 5-10 fold reduction in transcription rate cannot account for the rate of loss (Mangiarotti et al., 1982; Landfear et al., 1982). If high levels of extracellular cAMP are included, transcription is restored and the mRNA is stabilised. It may be that bifurcating second messenger pathways control these two

processes, and DIF interacts only with the branch concerned with transcriptional control. However, as the half-life of D19 mRNA in HM44 cells is apparently lower than the value expected in multicellular aggregates of AX-3 cells, the lack of cell contact may affect the control of stabilisation. Thus any effect that DIF may have on mRNA stability in vivo would not be detected in vitro.

If the mechanism of DIF action does resemble that of steroid hormones as suggested above, DIF could bind either to a single receptor or to distinct receptors to mediate both induction of prestalk genes and repression of prespore genes. Thyroid hormone, an effector which binds to a steroid-related receptor to regulate gene transcription, resembles DIF in positively and negatively regulating transcription of specific genes (Glass et al., 1988). The mechanism of repression is not known, but it has been shown that the receptor can bind to a slightly diverged response element upstream of genes activated by oestrogen, failing to activate transcription. It is proposed that competitive inhibition is responsible for repression. The situation is not strictly analogous in *Dictyostelium*, as cAMP rather than a DIF-related compound is the major prespore inducer (see Chapter 1). However, the arrangement of the promoter elements may be such that binding of the putative DIF receptor to the D19 promoter blocks cAMP induction. It would be of interest to transform HM44 cells with constructs containing modified versions of the D19 promoter to determine whether the sequences required for DIF repression and cAMP induction are exactly co-localised, overlapping or completely distinct. The most

likely explanation for the first of the three possibilities is that DIF interacts with a stage of the cAMP induction pathway. However, if independent upstream sequences are required for repression, the action of DIF probably involves a separate trans-acting factor.

The expression of PsA in tipped aggregates, detected with MUD-1, seems to be in response to a localised signal. Staining cells first appear in the base and central portion of the aggregate, and are never seen in the tip (Krefft et al., 1985). This data supports a position-dependent model of cellular differentiation, in which prespore cells arise in response to a gradient of an extracellular signal. The pattern may be partly a result of a gradient of DIF, decreasing from tip to base, repressing prespore gene expression in the tip, in addition to an inducing signal. Consistent with this, the DIF-induced gene products are localised in the tip (K.Duffy, K.Jermyn and J.Williams, personal communication). At present, nothing is known about the distribution of the biologically active form of DIF in the multicellular aggregate. Thus there is no evidence to distinguish the possibilities of a gradient of DIF or a gradient of DIF responsiveness.

A morphogen may be defined as a substance that induces different cellular fates at different concentrations (Slack, 1987). In theory, the action of DIF to induce prestalk gene expression and to repress prespore gene expression is directed against the same set of cells, and so this may occur when a single threshold value is reached. However, recent data has shown that the DIF-inducible genes are only

expressed in a subset of the non-prespore staining tissue (K.Duffy, K.Jermyn and J.Williams, personal communication). Although this may be a reflection of interaction with other morphogens, it could mean that lower concentrations of DIF are required for repression than induction. Opposing this, preliminary evidence suggests that for the HM44 stalk cell induction at least, lower DIF concentrations are required for pDd63 induction than for D19 repression (K.Jermyn, personal communication).

The size of the prespore-free zone in different slime mould species is directly related to the onset of cAMP signalling by the tip (Schaap et al., 1985). Where the onset is early, such as *D.discoideum*, the prespore-free zone is relatively large. In species where it is late, almost all cells are initially prespore. This suggests that the production of the substance responsible for prespore repression is co-incident with cAMP signalling. Adenosine is produced by the hydrolysis of cAMP, and the enzymes required for this conversion are preferentially localized in prestalk tissue, where signalling originates (Armant et al., 1980; Tsang and Bradbury, 1981). In addition, adenosine is a competitive inhibitor of prespore gene expression in aggregation competent cells in shaken suspension, which have the ability to express certain post-aggregative genes on addition of cAMP (Schaap and Van Driel, 1985). Like DIF, this may be an indirect effect due to reduction in cAMP relay (and hence overall cAMP levels) as a result of inhibition of cAMP binding to the A form of the receptor.

Unlike DIF, exposure of slugs to high adenosine levels does not increase the total proportion of non-prespore staining cells, but does affect patterning (Schaap and Wang, 1986). Assessing the relative roles of DIF and adenosine in achieving prespore repression in vivo is difficult because of the lack of knowledge about physiological concentrations and distribution. An added complexity is the interaction with the inducer of prespore gene expression, cAMP. It has been predicted that due to continued cAMP signalling in aggregates, in conjunction with the preferential conversion of cAMP to adenosine in the tip, a gradient of extracellular cAMP is created, increasing from tip to rear (Wang et al., 1988b). If adenosine exerts all its effects indirectly by modulating extracellular cAMP levels, it may be more appropriate to consider the balance between DIF and cAMP since the data presented here suggests that DIF may be additionally acting at a post-receptor stage. Therefore there could be competition at an intracellular level between induction by cAMP and repression by DIF.

DIF and adenosine may act primarily at different stages to repress the appearance of prespore cells. Adenosine production may be involved in setting up the cAMP gradient when cAMP signalling is first initiated by the tip, resulting in sufficient accumulation of cAMP for prespore induction only at the base of the aggregate as well as repressing prespore expression at the tip. The presence of prespore-negative cells at this stage which also fail to express the DIF-induced genes (K.Duffy, K.Jermyn and J.Williams, personal communication) suggests that DIF is unlikely to be

responsible for prespore repression in these cells. As higher DIF levels are required for repression of D19 than for induction of pDd63 in HM44 cells (K.Jermyn, personal communication), it would not be expected that levels insufficient for DIF induction could mediate DIF repression.

8.3 The expression of D19 in stalk cells

The results described in Chapter 7 suggest, but do not prove, that D19 is expressed in stalk cells but not in spore cells. This conclusion is surprising for two reasons; earlier during development D19 is expressed specifically in prespore cells, and DIF, which induces stalk cell formation, also represses D19 expression. No other examples of prespore-specific products present in stalk cells have been reported to date. It is known from fate mapping that spore cells are derived from prespore cells and prestalk cells become stalk cells (Raper, 1940). Two-dimensional gel electrophoresis has been used to examine the patterns of protein synthesis in prespore, prestalk, spore and stalk cells (Borth and Ratner, 1983; Morrissey et al., 1984; Cardelli et al., 1985). In general, it appears that there is a greater continuity of expression between prespore and spore cells than between prestalk cells and stalk cells. This has led to the conclusion that many of the gene products synthesized by prespore cells are required for spore formation, while prestalk differentiation is largely unrelated to stalk cell formation. The synthesis of stalk-specific proteins is not detected until the onset of culmination.

Although all of the spore-specific proteins detected are synthesized earlier during development in prespore cells, there are many examples of prespore-specific proteins absent from spores. Eight out of twenty prespore proteins identified by Morrissey et al. (1984) are in this category, including sp29, now known to be PsA, the D19 gene product. A lower proportion of prespore proteins identified by Borth and Ratner (1983), two out of twelve, did not persist into spores. Therefore it is to be expected that D19 mRNA should be absent from spores, because unless translational controls are operating, a reduction in mRNA levels should precede a loss of protein. Also, as discussed in section 8.1, it is very unlikely that PsA is a structural component of the spore coat. It is likely that a low level of expression would be detected more easily at the RNA level, which may explain the lack of detection of any prespore protein in stalk cells to date.

The evidence for expression of D19 mRNA in stalks was provided by Northern transfer analysis. The RNA analysed was extracted from isolated spore and stalk cells. Spores are easily purified because they are more resistant to lysis by nonionic detergents. After repeatedly washing and spinning cells in detergent-containing buffer, only the spore cells remain, which can then be lysed by freeze-thawing. As no signal was detected in the spore preparations using the D19 cDNA clone as a probe, there is obviously no contamination with other expressing cells.

In contrast, it is more difficult to prepare uncontaminated samples of mature stalk cells. They are

purified by repeated cycles of vigorous agitation and filtering of culminants to remove spores. This process is continued until very few spores can be detected in the eluate, but it is unlikely that all the spores are removed owing to their adhesive properties. In spite of this they probably do not constitute the major contamination problem, since they contain little or no D19 mRNA. Also, phenol-extraction of spores without prior lysis is very inefficient. Most contamination of stalk RNA is likely to be caused by the presence of immature prespore cells within the culminant. The proportion of such cells is not known. As culmination proceeds, the total amount of RNA per cell is reduced. Therefore, the yield of RNA from a prespore cell may constitute a disproportionately large part of the total. Most of the reduction in yield is due to a decline in ribosomal RNA synthesis, although the abundance of most messages is reduced to some extent (reviewed by Lodish et al., 1982), and so the proportion derived from the single copy fraction of the genome increases. This means that analysing equal amounts of RNA for each sample, any finite level of D19 mRNA in stalk cells will be enriched compared to earlier timepoints. No attempt was made to adjust the amounts of RNA compared.

A possible explanation for the apparent stalk expression is that the mRNA species is the product of a related but separate gene, specifically induced in stalk cells. The mRNA species is identical in size to D19 and so cannot be the novel message induced in HM44 cells by DIF (see Chapter 4). Washing at high stringency failed to diminish the intensity of the hybridisation signal. Therefore the message is either

the product of D19 or a very closely related gene. Conclusive proof that the transcript detected in the stalk fraction was the product of the D19 gene was provided by extracting RNA from a transformant containing a marked copy of the gene (see Chapter 5), enabling its specific detection.

In an attempt to reduce the possible contamination of stalk cells with prespore cells, RNA was extracted from four sources; 1) NC4 plucked stalks, to reduce contamination with immature aggregates; 2) V12M2 cells after stalk cell formation in monolayer culture, conditions in which virtual 100% stalk cell formation results; 3) stalks of HL31, a "stalky" mutant and 4) stalks of Ax-2 transformants 12 hours after culmination. The fact that D19 mRNA could be detected in each of these samples is strong evidence that the gene is expressed in stalk cells. However, criticisms can be levelled: it is not clear whether further maturation of any remaining immature cells continues after the completion of culmination; the expression of genes in monolayer culture may not accurately represent the in vivo situation; and the developmental programme of HL31 prespore cells does not switch to stalk cell differentiation until mid- to late culmination (Morrissey et al., 1981, 1984), and so prespore repression may be incomplete.

The localisation of marked PsA in culminants detected with the monoclonal antibody 9E10 (Evan et al., 1985) is further evidence for the expression of D19 in stalks, as PsA is apparently associated with stalks but not spores. The results of the electron microscopy were clear-cut; no background was seen in AX-2 cells subjected to the same

procedures. Again, use of the transformant ensured that the results obtained were not due to cross-reaction with a related but stalk-specific species.

However it was not possible to determine in which cells PsA was originally synthesized, as the staining was localised in the extracellular matrix. The association of PsA staining with stalk cells could merely reflect the presence of sheath material, and thus a consequence of loss of PsA from the surface of prespore cells earlier in development. Alternatively, the lack of cell surface staining could be caused by the specific release of PsA on maturation of stalk cells, for example due to cleavage of the GPI link. The cells in which PsA was actually produced could be stained if the protein was retained by the cells. Perhaps the simplest way to achieve this would be to fuse the D19 promoter to a heterologous gene, encoding a non-secreted protein again bearing a unique epitope. If the coding region of D19 itself was mutated to try to prevent export of the protein, the protein and/or mRNA may have reduced stability, and the efficiency of translation could be affected. The most obvious heterologous gene available is the actin 15/CAT fusion (see Chapter 6), as the D19 promoter has been inserted upstream and is known to direct correctly regulated expression. In addition, monoclonal antibodies exist which recognise the CAT protein. A less straightfoward approach which might ultimately yield more informative results is in situ hybridisation. By this means the D19 mRNA could be localised. The RNA content of a cell is a more direct indication of the genes being expressed at that time than the proteins present.

Also, it may be that a heterologous protein has unnatural stability properties, and so expression levels would not be accurately reflected.

The supposed expression of D19 mRNA in stalks is not a unique property of the gene, as all prespore genes tested hybridised to mRNA species in stalk RNA but not in spore RNA. Unlike D19 it was impossible to be certain in each case that the transcript was the product of the same gene, as there was no knowledge about the existence of cross-hybridising messages. The absence of any hybridisation to spore samples brought the integrity of the RNA into question, as at least one of the genes is known to encode a spore coat protein (Devine et al., 1982). However, only RNA corresponding to proteins required during germination would be required to be present after spore maturation.

If expression of prespore genes in stalk cells is a generalised phenomenon, it seems unlikely to be due to reinduction of transcription during late development. While it may be possible that the products of one or two prespore genes might have an additional function in stalks, it is not credible for all prespore proteins. A more likely explanation is that transdifferentiation of a subset of prespore cells occurs during late culmination. In other slime mould species which lack a prespore-free zone entirely, the stalk is formed by redifferentiation of prespore cells (Schaap et al., 1985). The capacity for a rapid prespore to stalk switch may have been retained during evolution. In support of this, redifferentiation of prespore cells in *D.discoideum* has been observed at the prespore/stalk boundary at the top of the

stalk, i.e. the last part of the stalk to be formed (Schaap, 1986). Therefore a prediction is that any residual prespore gene products are located at the top of the stalk in the mature culminant. The examination of stalk sections by electron microscopy did not enable the region staining most intensely to be identified, and visualisation at lower power with an alternative staining strategy was not undertaken.

It is known that isolated prespore tissue regulates more rapidly than prestalk tissue (Raper, 1940; Bonner and Slifkin, 1949). It is possible that prestalk cells can only become prespore in character after a considerable lag because the greater number of spore than stalk-specific components, many of which are highly abundant (Borth and Ratner, 1983). Analysis of protein synthesis has suggested that the majority of stalk-specific products are not synthesized until late in culmination, considerably after the induction of the synthesis of proteins later detected in spores (Borth and Ratner, 1983; Morrissey et al., 1984; Cardelli et al., 1985). As there is no apparent requirement for continuity with prestalk protein synthesis, this implies that a relatively short time is required for a cell to accumulate the necessary proteins to assume stalk character. Thus it is possible that transdifferentiation of prespore cells to stalk cells could occur rapidly at the end of culmination. Such transdifferentiation is witnessed on a larger scale in some mutant strains such as HL31, where the prespore/prestalk ratio is normal during the slug stage, but the culminant is predominantly composed of stalk cells (Morrissey et al., 1981). The presence of prespore-specific products need not

adversely affect stalk cells, since they have a purely structural role.

As dedifferentiation following disaggregation seems to involve the rapid loss of prespore-specific mRNA, it seems surprising that relatively high levels of prespore message should be retained in stalk cells after the transdifferentiation of prespore cells. However, this apparent inconsistency may be due to the existence of distinct regulatory mechanisms. When aggregates are disrupted, the signal for mRNA destabilisation is a drop in cAMP levels (Chung et al., 1981; Mangiarotti et al., 1983b), whereas the signal for stalk cell formation is DIF (Town et al., 1976). The data presented here has established that DIF represses prespore gene expression at the transcriptional level, but has no apparent effect on message stability (see Chapter 4 and section 8.2). The overall rate of loss of prespore mRNA on DIF addition in vitro is relatively slow; levels of D19 mRNA decline with a half-life of 90 minutes. As overt stalk cell differentiation in vivo is not detectable until about 18 hours of development, it is not surprising that a significant proportion of the prespore RNA should be retained in mature fruit.

Assuming transcription is not reduced in the prespore cells which form spores prior to the time it is blocked (presumably by DIF) in prespore cells proposed to form stalks, the complete absence of D19 mRNA in spores requires the existence of an additional, spore-specific regulatory mechanism, such as mRNA destabilization. It seems logical that mRNA levels should be subject to more careful modulation

in spores, which are required for the next stage of the life cycle, than in stalks, which die on terminal differentiation. Without a reduction in half-life in spores, differences in transcription between prespore cells with different eventual fates would be required before the onset of culmination. This is inconsistent with rapid transdifferentiation occurring during the latter stages of culmination; conversely, the prediction of post-transcriptional control is dependent on the validity of the transdifferentiation model.

8.4 Transcriptional control of D19 expression

Analysis of the sequences important for the regulated expression of D19 have served to emphasise the primary importance of transcriptional control mechanisms. Previous studies had suggested that post-transcriptional regulation may be involved in the control of prespore-specific gene expression. For several genes shown to be transcribed early in development, accumulation of messenger RNA cannot be detected until after aggregation (Mangiarotti et al., 1985). This has been cited as possible evidence for control by specific stabilisation. Later in development the mRNA, where tested, is localised in prespore cells (A.Ceccarelli, personal communication). The half-life of post-aggregative mRNA species is estimated to be approximately 3 to 4 hours (Chung et al., 1981). However, if the aggregates are disrupted, many developmentally regulated messages are rapidly lost with a much shorter half-life, while other constitutive messages are maintained (Mangiarotti et al.,

1982). The mRNA species which seem to exhibit this behaviour most strongly are derived from prespore genes (Barklis and Lodish, 1983; Chisholm et al., 1984). In most cases the message level can be restored by the addition of high levels of extracellular cAMP (Barklis and Lodish, 1983; Mangiarotti et al., 1983b). As well as acting to stabilise the messages, addition of cAMP increases the transcription rate of some of these genes by 5-10 fold (Landfear et al., 1982). However, without a major post-transcriptional component the rapid changes in prespore mRNA levels observed would not be possible.

It is not known whether message stabilisation/ destabilisation mechanisms early and late in development are related. A major rise in extracellular cAMP occurs on aggregate formation, the time when expression of many prespore genes is first detected. It could be envisaged that the rise in cAMP at this time acts to stabilise messages derived from prespore genes transcribed early during development, while disaggregation reverses this process due to a drop in the extracellular cAMP concentration. Such a mechanism would be a simple means of co-ordinating the expression of a class of cell-type-specific genes.

D19 mRNA is rapidly lost on disruption of aggregates due to specific destabilisation (Barklis and Lodish, 1983). However, data presented here has shown that transcription of D19 is not detected before the accumulation of D19 mRNA in the cytoplasm. This was judged by determining the expression pattern of a fusion gene used in transformation which contained only upstream, non-transcribed D19 sequences. The

target for post-transcriptional control mechanisms by definition must lie within the structure of the mRNA, i.e. the transcribed region of the gene. Thus the constructs in which putative D19 promoter sequences were placed upstream of the actin 15 CAT gene were a means of examining transcriptional regulation of D19 expression in the absence of any specific post-transcriptional control.

The presence of sequences encoding the bacterial chloramphenicol acetyltransferase gene allowed expression to be monitored by measuring enzyme activity (Gorman et al., 1982). The absence of CAT activity early in development strongly implies that the induction of D19 expression is regulated at the level of transcription. Hence D19 is not regulated by post-transcriptional mechanisms, at least early during development. It could be argued that as these assays are not measuring trancription directly, CAT activity, which requires protein synthesis, might occur only after a delay. Also, the time of induction of CAT activity and the appearance of D19 mRNA was not compared within the same experiment. This would have been difficult to interpret because the inherent sensitivity of the techniques differ and different aspects of gene expression are being compared. However, CAT assays are considered to be highly sensitive, and so there is unlikely to be a significant delay between D19 mRNA production and CAT activity. The onset of transcription of the prespore genes proposed to be subject to early post-transcriptional mechanisms, detected by nuclear run-on assays, is at about 2 hours of development (A.Ceccarelli, personal communication). From the CAT assays

D19 transcription is estimated to begin at about 10 hours, a difference unlikely to be affected by problems of sensitivity. Finally, run-on assays have confirmed that in contrast to some other prespore genes, D19 is not transcribed early in development (A.Ceccarelli, personal communication).

The absence of post-transcriptional control early in development but its operation during disaggregation for D19 does not necessarily indicate that distinct mechanisms are involved. It would be of interest to express D19 mRNA early during development by fusing the transcribed region to a promoter known to be active at this stage.

Use of the same constructs described above has also indicated that the D19 promoter is active almost exclusively in prespore cells. Therefore, transcriptional mechanisms are sufficient for prespore specificity. These results strongly suggest that the cell-type-specificity of D19 expression is primarily controlled at the transcriptional level, and predict the existence of transcription factors active in prespore cells only. However, it is impossible to rule out the involvement of post-transcriptional mechanisms completely. The half-life of the actin 15 CAT message was not measured. If D19 mRNA is significantly more stable than actin 15 CAT mRNA, and a low level of transcription does occur in prestalk cells, specific message destabilisation in prestalk cells may be additionally required for the degree of prespore-enrichment observed for the D19 message. Alternatively, if control of message stability is non-cell-type-specific, or is prespore-specific, stabilisation may occur in conjunction with cell-type-specific transcription to

enable prespore-specific message accumulation to occur. The specific destabilisation of D19 mRNA on disaggregation suggests a potential for post-transcriptional control exists, although such a mechanism may operate solely during dedifferentiation.

It it not clear how widely transcriptional regulation of prespore-specific gene expression is utilised as a mechanism of control in *D.discoideum*. The relative contribution of transcriptional and post-transcriptional mechanisms have not yet been assessed for other prespore genes at the multicellular stage. For the prespore genes known to be transcribed early during development, before overt differentiation of prespore and prestalk cells, the simplest model is continued transcription in all cells combined with specific mRNA stabilisation in prespore cells. However, it has not been established whether these genes are transcribed in all cells early in development. It could be that they are only activated in a subset of cells which later become prespore cells. Even if they are transcribed in all cells at early stages, it is still a formal possibility that transcription is later prespore-specific. Sequence data is now being generated for other prespore genes. In order to assess the value of comparison to identify conserved features, whether transcribed or non-transcribed, it will be necessary to characterise the control mechanisms involved more fully.

There are no obvious reasons why different control mechanisms should be employed for genes apparently expressed in the same cell types at the same time. In general, post-

transcriptional control enables more rapid changes in expression. Therefore it may be that for the prespore mRNAs induced by post-transcriptional mechanisms, maximal expression is required to be attained more quickly than for D19. Transcriptional control may have different benefits, such as ensuring absolute cell-type specificity of expression.

Differences in the relative contribution of transcriptional and post-transcriptional mechanisms may simply be required to modulate individual expression patterns precisely. Although prespore genes cannot be divided into simple, perfectly coregulated classes, and individual genes are obviously subject to differing control mechanisms, there is some degree of co-ordination in the timing of expression as well as the cell-type specificity (Chisholm et al., 1984; Morrissey et al., 1984). This is suggested by protein synthesis measurements as well as analysis of the expression of individual clones, although it is possible that insufficient genes have been studied to reveal the true extent of the complexity existing. Thus, it seems consistent that prespore gene expression could be controlled by a common extracellular signal(s). Many experiments have shown that the interaction of cAMP with the cell-surface receptor is required for prespore gene expression (see Chapter 1). This is consistent with the dependence of prespore differentiation on cAMP, but has to be reconciled with the involvement of cAMP in stimulating the expression of many other genes. The observation that higher levels of cyclic AMP are required for prespore-specific gene expression than for other genes (Gomer

at al., 1986; Oyama and Blumberg, 1986b) fits with the proposed cAMP gradient existing in the slug (Wang et al., 1988b), and suggests how cell-type-specific regulation might be achieved. Thus cAMP could interact with a distinct, low-affinity form of the cell surface receptor to activate a separate second messenger pathway, leading to the specific induction of prespore gene expression. Some experiments using agonists such as verapamil and TMB8 have hinted that Ca2+ could be a prespore-specific second messenger (Schaap et al., 1986), although there is as yet no conclusive evidence.

Examination of the sequence upstream of D19 has not revealed any homology to cAMP or Ca2+/phorbol ester response elements identified in other systems (reviewed by Roesler et al., 1988). Deletion analysis indicates that a 216 nucleotide region, beginning 122 nucleotides upstream of the start site of transcription, contains sufficient signals to direct correctly regulated expression, both with regard to cell-type specificity and timing. This is a direct inference from the combined results of deleting from the 5' end and from the 3' end. In each case, expression was lost before there was any change in the pattern of regulation. However, elements required for correct timing and cell type-specificity of expression need not overlap those required to promote efficient expression within the minimal region. It could be that in both sets of deletions, regulated expression was effectively abolished because it fell below the minimum level which could be detected, owing to loss of an efficiency element. Therefore the nature of the signals contained in the regions whose removal resulted in complete loss of expression

cannot be judged, i.e. between -122 and -163, and -338 and -291.

Between -338 and -291, most of the sequence is extremely A-T rich except for the extreme 5' end, where the deletion interrupts a G-rich element, GCTATGATGG. The next deletion upstream contains an additional 26 nucleotides, composed entirely of As and Ts except for the four nucleotides of the G-rich element missing from -338 and a single additional C residue. In two sets of transformants, consisting of pooled populations, the expression of -364 was detectably higher than for -338. It is difficult to quantitate the differences in expression levels because variation occurred between independent populations transformed with the same construct. Southern blotting indicated approximately equal copy numbers for all transformants and so there was no correlation between reduced expression and low copy number. It may be that slight variations in expression occurred because of the small population sizes of some transformants, so increasing the influence of the site of integration on expression.

The sequence GCTATGATGG has 7/10 matches with part of a consensus sequence, identified by comparing G-rich sequences found upstream of several *Dictyostelium* genes (see figure 6.7). The consensus, $^{A}/_{G}$ $^{A}/_{G}$ $^{A}/_{G}$ G G $^{A}/_{G}$ $^{T}/_{G}$ T G G $^{G}/_{T}$ G $^{G}/_{T}$, is very approximate and only a few elements have no mismatches. It may be that G-richness is the important shared feature rather than primary sequence homology, possibly as a consequence of the extreme A-T richness of the surrounding DNA. The genes so far found to have these G-rich motifs do not exhibit the same patterns of expression; for instance D11

is inducible by cAMP (Barklis and Lodish, 1983) while discoidin 1 is repressed by cAMP (Williams et al., 1980). Therefore, if they do have any functional equivalence, they are unlikely to be required for cAMP-mediated induction of expression.

Cysteine proteinase (CP) 1 and 2 are co-regulated and both have similar but non-homologous G-rich repeats (Pears et al., 1985; Driscoll and Williams, 1987; Pears and Williams, 1987). The arrangement of the repeats differ; within about 900bp of the CP1 cap site are five short GC-rich regions, possibly reflecting the presence of a divergently transcribed coregulated gene, while within about 900bp upstream of CP2 are two G-rich blocks only. The two most proximal CP1 repeats form the core of an almost perfectly conserved, larger direct repeat, while in CP2 they are much closer to eachother. However, they are functionally equivalent; two copies of the "core" G box sequence from CP1 can restore efficient cAMP-inducible expression to a deleted version of the CP2 lacking its G-rich repeats (Pears and Williams, 1988). It has also been shown that mutation or deletion of one copy of the CP2 G box does not completely abolish expression; a low level remains that is still cAMP inducible (Pears and Williams, 1987). Although this suggests that the G box is simply functioning to promote efficient expression, it does not rule out the possibility that the low level of cAMP-inducible expression remaining is due to activity of a cryptic element.

Therefore the reduced expression between the -364 and -338 deletions may be the result of interrupting of a G-rich element required to promote efficient expression, and the

absence of expression in the -291 deletion due to complete removal of this sequence. Further upstream, outside the minimal promoter region, seven additional predominantly G-rich motifs are found, interspersed with extremely A-T rich sequence. There is a 40-fold reduction in the level of expression of the full-length construct compared to the -338 construct, as judged by the relative amount of marked transcript at the time of peak D19 expression. The successive removal of the G-rich sequences could be responsible for the gradual reduction in the efficiency of expression observed. For instance, the -442 construct is expressed significantly more strongly than the - 338 construct and contains two additional G-rich elements.

The 3' deletion construct terminating 52 nucleotides upstream of the cap site produces approximately a 20-fold higher level of CAT activity than the construct with 122 nucleotides deleted. The endpoint of the -52 construct lies within the oligo dT run, and hence the D19 TATA box sequences are intact. However, they are deleted by the removal of an additional 70 nucleotides. Primer extension analysis indicates that use of the D19 TATA box is not responsible for the increased expression of the -52 construct, as only a very minor proportion of the transcripts initiate upstream of the expected actin 15 start sites. Therefore there is either a positively acting element between 52 and 122 nucleotides upstream of the cap site, or the relative spacing of the elements in the -52 construct is more favourable. In spite of the reduced level of expression, the -122 construct demonstrates that heterologous promoter elements can work in

combination, as the actin 15 TATA box is used exclusively. Without the inclusion of D19 promoter sequences, the deleted actin 15 CAT construct is not expressed at all.

Deleting to -163 abolishes expression. Between -122 and -163 is a C-rich sequence element, CCCCACTG. This resembles a G box in the opposite strand and could explain the lack of expression. It is also possible that loss of expression on deleting from both the 5' and 3' ends is unrelated to or only partially due to removal of G-C rich sequences; A-T rich regions may also be of importance, although it is more difficult to imagine how they could represent specific signals. Deletion analysis of CP2, also called pst-cath, has shown that an A-T rich sequence located 5' to the G-C rich sequence is required for maximal expression (Datta and Firtel, 1988).

The rationale behind the use of synthetic oligonucleotides to try to localise important sequences more closely was based on two assumptions. Firstly, it was considered likely that the putative G box sequence at the 5' end of the minimal promoter region was required for efficient expression, but not for cell-type specificity. As a consequence of this, it was hoped that the G box from CP1 could be substituted for this sequence, allowing the function of shorter, more C-rich downstream regions to be tested. However, oligonucleotides corresponding to two adjacent segments of the minimal promoter region 84 and 75 nucleotides in length, inserted either singly or together, failed to restore any expression of the deleted actin 15 promoter containing the CP1 G box.

There are several possible reasons for this failure. It could be that the region synthesized did not include sequences absolutely required for expression. If this true, it necessarily implies that certain A-T rich sequences are essential for expression. Another explanation is that any expression which was restored was below background levels. The downstream oligonucleotide only spanned sequences present in the -122 deletion, which gave low levels of expression even though its 5' terminus was located 995 nucleotides upstream of the cap site. Therefore, a single G rich sequence in combination with this 3' terminus may be insufficient for detectable expression. A third possibility is that the elements from CP1 and D19 are unable to combine to promote expression. This may be the result of spacing problems or an inappropriate arrangement, or simply reflect the nature of the elements themselves. The G box from CP1, although able to restore expression of a related gene (Pears and Williams, 1988), may be ineffective with a completely unrelated gene. This could be due to, for example, a requirement for interaction of a G box factor with other factors binding elsewhere. It may also be that although there is an apparent similarity between the G boxes of different genes, they are in fact recognised by distinct factors. Therefore the activity of the CP1 element may be limited to a particular stage of development, before D19 transcription is initiated. It is known that the CP1 G box cannot restore expression of the deleted actin 15 promoter (C.Pears, Ph.D.thesis), and so in order for the addition of D19 sequences to have done so they may have required some inherent promoter activity. If

the sequences inserted contained signals directing cell-type specificity alone, this requirement would not have been fulfilled. Therefore an alternative experiment to test the function of D19 sequences could be to try to modify the expression of a gene already active, i.e. to change the partitioning between prestalk and prespore cells. However, for a strongly expressed gene such as actin, any change may be difficult to detect.

In light of the inability to define the elements required for regulated expression by manipulating the sequence, alternative approaches may be more effective in identifying important regions. By eye, no obvious features are observed within the minimal region, either with homology to any previously defined consensus sequence or with a potentially important structure, such as a region of dyad symmetry. There are a few short direct repeats, some adjacent and some separated. However, comparison with the upstream sequence of another prespore gene, AC914, has revealed partial homology to a dyad symmetry element at about the same position relative to the cap site (A.Ceccarelli, personal communication). No significant similarity to any other potential *Dictyostelium* promoter element exists. An oligonucleotide spanning the AC914 sequence is currently being used to screen a lambda gt11 library using the method of Singh et al. (1988), to try to directly isolate a gene possibly encoding a prespore-specific trans-acting factor (J.Williams, personal communication).

Other, less direct means of identifying trans-acting factors are possible, and such experiments will benefit from

the information provided by the deletion analysis. These include DNAase I hypersensitivity site mapping, gel retardation measurements and footprinting studies using DNAaseI or DMS. A potential problem is that D19 is not expressed in all cells. The 20% of cells which never express D19 may mask differences between the timepoints in which maximal transcription is occurring and those in which there is no transcription.

8.5 Conclusions and future prospects

The gene D19, encoding the cell surface protein PsA, is now the best characterised marker of the prespore pathway of differentiation in *Dictyostelium discoideum*. It is expressed, within limits of detection, exclusively in prespore cells. The induction of its expression is concomitant with the onset of prespore differentiation, as judged by histological and immunological criteria. Therefore, the mechanisms which control the expression of this gene are relevant to an understanding of the control of cell-type-specific differentiation.

Both the D19 cDNA clone and the monoclonal antibody MUD-1 are frequently used as prespore-specific markers. With the knowledge that the protein recognised by MUD-1 is encoded by D19, analysis of the control of D19 gene expression can be directly related to studies of pattern formation utilising MUD-1.

Initially, the limited number of prespore genes isolated were divided into two major classes on the basis of

accumulation profile and requirement for cAMP. However, with the subsequent analysis of additional prespore-enriched genes, it is becoming clear that there is a great deal of heterogeneity in patterns of expression. In particular, prespore genes seem to show individual variation in the extent to which they are regulated post-transcriptionally. The expression of D19 seems to be regulated predominantly by transcriptional mechanisms, while the induction of other genes seems to be controlled by specific mRNA stabilization. However, D19 may be representative of a subclass of prespore genes; for instance, the cDNA clone PL3 hybridises to a transcript which is coregulated with D19 mRNA in all experiments where both genes have been analysed.

Even if the precise mechanisms by which D19 is regulated are not shared by a large number of other genes, it is still of interest to determine how substances known to control cellular differentiation, such as cAMP and DIF, act to modify the expression of individual genes. The data presented here have shown that transcriptional mechanisms are sufficient to mediate cell-type-specific expression, and that DIF acts at the transcriptional level to repress D19. Deletion analysis has defined a 216 nucleotide upstream region which contains the signals required for correctly regulated expression. This information will be of assistance in tracing the pathway from the receipt of a morphogenetic signal by the cell, possibly at the cell surface, to an alteration in the expression of a specific gene within the nucleus. From the gene working backwards, an obvious starting point is the isolation of

trans-acting factors required for cell-type-specific transcription.

REFERENCES

Aekerle,S., Wurster,B. and Malchow,D (1985) EMBO J. 4 39-43

Ahern,K.G., Howard,P.K. and Firtel, R.A. (1988) Nucleic Acids Res. 16 6825-6838

Akam,M. (1987) Development 101 1-22

Alton,T.H. and Lodish,H.F. (1977) Dev.Biol. 60 207-216

Andrews,G.K., Jantzen,R.G. and Tamaoki,T (1982) Dev.Biol. 89 111-116

Angel,P., Imagawa,M, Chiu,R., Stein,B, Imbra,R.J., Rahmsdorf, H.J., Jonat,C., Herrlich,P. and Karin,M (1987) Cell 49 729-739

Angel,P., Allegretto,E., Okino,S., Haltori,K., Boyle,W., Hunter,T. and Karin,M. (1988) Nature 322 166-171

Armant,D.R., Stetler,D.A. and Rutherford,C.L. (1980) J.Cell Sci. 45 119-129

Arndt,K. and Fink,G. (1986) PNAS 83 8516-8520

Atchison,M.L. and Perry,R.P. (1987) Cell 48 121-128

Augereau,P and Chambon,P (1986) EMBO J. 5 1791-1797

Ayres,K., Neuman,W., Rowekamp,W.G. and Chung,S. (1987) Mol. Cell.Biol. 7 1823-1829

Ayusawa,D., Shimizu,K., Koyama,H., Kaneda,S., Keiichi,T. and Seno,T. (1986) J.Mol.Biol. 190 559-567

Baeuerle,P.A. and Baltimore,D. (1988) Cell 53 211-217

Banerji,J., Olsen,L. and Schaffner,W. (1983) Cell 33 729-740

Banerji,J., Rusconi,S and Schaffner,W. (1981) Cell 27 299-308

Barad,M., Jack,T., Chadwick,R. and McGinnis,W. (1988) EMBO J. 7 2151-2161

Barberis,A., Superti-Furga,G. and Busslinger,M. (1987) Cell 50 347-359

Barclay,S.L. and Smith,A.M. (1986) Differentiation 33 101-110

Barklis,E. and Lodish,H.F. (1983) Cell 32 1139-1148

Barklis,E., Pontius,B., Barfield,K. and Lodish,H.F. (1985) Mol.Cell.Biol. 5 1465-1472

Baron,M.H. and Maniatis,T. (1986) Cell 46 591-602

Beachy,P.A., Helfand,S.L. and Hogness,D.S. (1985) Nature 313 545-551

Beck,E., Ludwig,G., Auerswald,E.A., Reiss,B. and Schaller,H. (1982) Gene 19 327-336

Becker,P.B., Ruppert,S. and Schutz,G. (1987) Cell 51 435-448

Belayew,A. and Tilghmann,S.M. (1982) Mol.Cell.Biol. 2 1427-

Benoist,C, O'Hare,K., Breathnach,R. and Chambon,P. (1980) Nucleic Acids Res. 8 127-142

Berks,M. and Kay,R.R. (1988) Dev.Biol. 126 108-114

Bernstiel,M.L., Busslinger,M. and Strub,K. (1985) Cell 41 349-359

Bienz,M. (1985) Trends Biochem.Sci. (1985) 10 157-161

Bienz, M. and Pelham, H.R.B. (1986) Cell 45 753-760

Blaschke,A, Weijer,C. and MacWilliams,H. (1986) Differentiation 32 1-9

Blumberg,D.D. and Lodish,H.F. (1980a) Dev.Biol. 78 268-284

Blumberg,D.D. and Lodish,H.F. (1980b) Dev.Biol. 78 285-300

Blumberg,D.D. and Lodish,H.F. (1981) Dev.Biol. 81 74-80

Bohmann,D., Bos,T.J., Admon,A., Nishimura,T., Vogt,P.K. and Tjian,R. (1987) Science 238 1386-1392

Bonner,J.T. (1944) Amer.J.Bot. 31 175-182

Bonner,J.T. (1949) J.Exp.Zool. 106 1-26

Bonner,J.T. (1952) Amer.Nat. 86 79-89

Bonner,J.T. (1982) In: The Development of *Dictyostelium discoideum*, 1-33. W.F.Loomis (ed.) Academic Press

Bonner,J.T., Chiquoine,A.D. and Koldene,M.Q. (1955) J.Exp. Zool. 130 133-157

Bonner,J.T., Clarke,W.W.,Jr., Neely,C.L.,Jr. and Slifkin,M.K. (1950) J. Cell. Comp. Physiol. 36 149-158

Bonner,J.T. and Dodd,M.R. (1962) Dev.Biol. 5 344-361

Bonner,J.T. and Slifkin,M.K. (1949) Amer.J.Bot. 36 727-734

Boothroyd,J.C., Cross,G.A.M., Hoeijmakers,J.H.J. and Borst,P. (1980) Nature 288 624-626

Borth,W. and Ratner,D. (1983) Differentiation 24 213-219

Both,G.R. Replication of the Reoviridae, In: RNA Genetics. P.Ahlquist, E.Domingo and J.Holland (eds.) CRC Press Inc., Boca Raton, Fla., in press.

Bouvagnet,P.F., Strehler,E.E., White,G.E., Strehler-Page,M., Nadal-Ginard,B. and Mahdavi,V. (1987) Mol.Cell.Biol. 7 4377-4389

Bram,R. and Kornberg,R.D. (1985) PNAS 82 43-47

Braun-Breton,C., Rosenberry,T.L. and Pereira da Silva,L. (1988) Nature 332 457-459

Breathnach,R., Benoist,C., O'Hare,K., Gannon,G. and Chambon, P. (1978) PNAS 75 4853-4857

Breathnach,R. and Chambon,P.A. (1981) Ann.Rev.Biochem. 50 349-383

Brenner,M. (1977) J.Biol.Chem. 252 4073-4079

Brent,R. and Ptashne,M. (1985) Cell 43 729-736

Briggs,M.R., Kadonaga,J.T., Bell,S.P. and Tjian,R. (1986) Science 234 47-52

Briggs,M.S. and Gierasch,L.M. (1986) Adv.Protein Chem. 38 109-180

Brock,M.L. and Shapiro,D.J. (1983) Cell 34 207-214

Brodie,C., Klein,C. and Swierkosz,J. (1983) Cell 32 1115-1123

Brookman,J.J., Jermyn,K.A. and Kay,R.R. (1987) Development 100 119-124

Bumann,J., Malchow,D. and Wurster,B. (1986) Differentiation 31 85-91

Buratowski,S., Hahn,S., Sharp,P.A. and Guarente,L. (1988) Nature 334 37-42

Cao,Z., Barron,E.A.,Camillo,A.J. and Sharp,Z.D. (1987) Mol. Cell.Biol. 7 3402-3408

Caras,I.W., Weddell,G.N., Davitz,M.A., Nussenveig,V. and Martin,D.W.Jr. (1987) Science 238 1280-1283

Cardelli,J.A., Knecht,D.A., Wunderlich,R. and Dimond,R.I. (1985) Dev.Biol. 110 147-156

Carroll,S.B. and Scott,M.P. (1985) Cell 43 47-57

Ceccarelli,A., McRobbie,S.J., Jermyn,K.A., Duffy,K., Early, A.E. and Williams,J.G. (1987) Nucleic Acids Res. 15 7463-7476

Cereghini,S., Blumenfeld,M. and Yaniv,M. (1988) Genes Dev. 2 957-974

Cereghini,S.,Raymondjean,M., Garcia Carranca,A., Herbomel,P. and Yaniv,M. (1987) Cell 50 627-638

Chandler,V.L., Maler,B.A. and Yamamoto,K.R. (1983) Cell 33 489-499

Chisholm,R.L., Barklis,E. and Lodish,H.F. (1984) Nature 310 67-69

Chodosh,L.A., Baldwin,A.S., Carthew,R.W. and Sharp,P.A. (1988a) Cell 53 11-24

Chodosh,L.A., Oleson,J., Hahn,S., Baldwin,A.S., Guarente,L.

and Sharp,P.A. (1988b) Cell 53 25-35

Chiu,R., Boyle,W.J., Meek,J., Smeal,T., Hunter,T. and Karin,M. (1988) Cell 54 541-552

Chiu,R.,Imagama,M.,Imbra,M.J., Bockoven,J.R. and Karin,M. (1987) Nature 329 648-651

Chung,S., Landfear,S.M., Blumberg,D.D., Cohen,N.S. and Lodish,H.F. (1981) Cell 24 785-797

Chytil,F. (1984) Pharmacol.Rev. 36 93S-100S.

Ciliberto,G., Buckland,R., Cortese,R. and Philipson,L. (1985) EMBO J. 4 1537-1543

Clarke,J. and Herr,W. (1987) J.Virol 61 3536-3542

Cockerhill,P.N. and Garrad,W.T. (1986) Cell 44 273-282

Coggi,G., Dell'Orto,P. and Viale,G. (1986) In: Immunocytochemistry-Modern Methods and Applications, 54-70. J.M.Polak and S.Van Noorden (eds.) Wright, Bristol U.K.

Cohen,S.M., Knecht,D., Lodish,H.F. and Loomis,W.F. (1986) EMBO J. 5 3361-3366

Colantuoni,V., Pirozzi,A., Blance,C. and Cortese,R. (1987) EMBO J. 6 631-636

Coppeck,D.L. and Pardee,A.B. (1987) Mol.Cell.Biol. 7 2925-2932

Costa,R.H., Lai.E., Grayson,D.R. and Darnell,J.E. (1988) Mol. Cell.Biol. 8 81-90

Courtois,G., Morgan,J.G., Campbell,L.A., Fourel,G. and Crabtree,G.R. (1987) Science 238 688-692

Cross,G.A.M. (1987) Cell 48 179-181

Cunningham,B.A., Hemperley,J.J., Murray,B.A., Prediger,E.A., Brackenbury,R. and Edelman,G.A. (1987) Science 236 799-806

Dani,C., Mechti,N., Piechaczyk,M., Lebleu,B., Jeanteur,P. and Blanchard,J.M. (1985) PNAS 82 4896-4899

Darnell,J.E. (1979) Prog.Nucleic Acids Res.Mol.Biol. 22 337-353

Darnell,J.E. (1982) Nature 297 365-371

Davis,R.L., Weitraub,H. and Lassar,A.B. (1987) Cell 51 987-1000

Datta,S. and Firtel,R.A. (1987) Mol.Cell.Biol. 7 149-159

Datta,S. and Firtel,R.A. (1988) Genes.Dev. 2 294-304

Davidson,I., Fromental,E., Augereau,P., Wildeman,A., Zenke, M. and Chambon, P. (1986) Nature 323 544-548

De Lozanne,A. and Spudich,J.A. (1987) Science 236 1086-1091

Denis,M., Poellinger,L., Wikstom,A-C. and Gustafsson,J-A. (1988) Nature 333 686-688

Dente,L., Cesareni,G. and Cortese,R. (1983) Nucleic Acids Res. 11 1645-1655

De Simone,V., Ciliberto,G., Hardon,E., Paonessa,G., Palla,F., Lundberg,L. and Cortese,R. (1987) EMBO J. 6 2759-2766

Desplan,C., Theis,J. and O'Farrell,P.H. (1985) Nature 318 630-635

Devine,K.M., Morrissey,J.M. and Loomis,W.F. (1982) PNAS 79 7361-7365

Devine,K.M., Bergman,J.E. and Loomis,W.F. (1983) Dev.Biol. 99 437-446

Devreotes, P.N. (1982) In: The Development of *Dictyostelium discoideum*, 117-168. W.F.Loomis (ed.) Academic Press

Devreotes, P.N. and Sherring,J. (1985) J.Biol.Chem. 260 6378-6384

Dierich,A., Gaub,M., LePennec,J., Astinott,P. and Chambon,P. (1987) EMBO J. 6 2305-2312

DiNardo,S., Kuner,J.M., Theis,J. and O'Farrell,P.H. (1985) Cell 43 59-69

Dinauer,M.C, MacKay,S. and Devreotes,P.N. (1980a) J.Cell Biol. 86 537-544

Dinauer,M.C., Steck,T.L. and Devreotes,P.N. (19806) J.Cell Biol. 86 554-561

Dobbeling,U., Ross,K., Klein-Hitpass,L., Morley,C., Wagner,U. and Ryffel,G.U. (1988) EMBO J. 7 2495-2501

Doe,C.Q. and Scott,M.P. (1988) Trends Neurosci. 11 101-106

Dorn,A., Bollekens,J., Staub,A., Benoist,C. and Mathis,D. (1987) Cell 50 863-872

Dottin,R.P., Weiner,A.B. and Lodish,H.F. (1976) Cell 8 233-244

Dowds,B.A.C. and Loomis,W.F. (1984) Mol.Cell.Biol. 4 2273-2278

Driscoll,D.M. and Williams,J.G. (1987) Mol.Cell.Biol. 7 4482-4489

Durston,A.J., Cohen.M.H., Drage,D.J., Patel, M.J., Robertson, A. and Wonio,D. (1976) Dev.Biol. 52 173-180

Durston,A.J. and Vork,F. (1979) J.Cell Sci. 36 261-279

Dustin,M.L., Selvaraj,P., Mattoliano,R.J. and Springer,T.A. (1987) Nature 329 846-848

Dynan,W.S. (1986) Trends Genet. 2 196-197

Dynan,W.S., Sazer,S., Tjian,R. and Schimke,R.T. (1986) Nature 319 246-248

Dynan,W.S. and Tjian,R. (1983) Cell 32 669-680

Dynan,W.S. and Tjian,R. (1985) Cell 35 79-87

Edelman,G. (1986) Ann.Rev.Cell.Biol. 2 81-116

Efstratiadis,A., Poskony,J.W., Maniatis,T., Lawn,R.M., O'Connell,C., Spritz., R.A., De Riel,J.K., Forget,B.G., Weissmann,S.M., Slighton,J.L., Blechl,A.E., Smithies,O., Baralle,F.E., Shoulders,C.C. and Proudfoot,N.J. (1980) Cell 21 653-668

Emerson,B.M., Nickol,J.M., Jackson,P.D. and Felsenfeld,G. (1987) PNAS 84 4786-4790

Ennis,H.L. (1981) Antimicrob.Agents Chemother. 19 657-665

Ephussi,A., Church,G.M., Tonegawa,S. and Gilbert,W. (1985) Science 227 134-140

Europe-Finner,G. and Newell,P.C. (1986) J.Cell Sci. 87 221-229

Evan,G.I., Lewis,G.K., Ramsay,G. and Bishop,J.M. (1985) Mol. Cell.Biol. 5 3610-3616

Everett,R.D., Baty,D. and Chambon,P. (1983) Nucleic Acids Res. 11 2447-2464

Falkner,F.G. and Zachau,H.G. (1984) Nature 310 71-74

Farnsworth,P. (1973) J.Embryol.Exp.Morphol. 29 253-266

Feinberg,A., Springer,W.R. and Barondes,S.H. (1979) PNAS 76 3977-3981

Feinberg,A.P. and Vogelstein,B. (1983) Analytical Biochem. 132 6-13

Ferguson,M.A.J., Low,M.G. and Cross,G.A.M. (1985) J.Biol. Chem. 260 14547-14555

Ferguson,M.A.J. and Williams,A.F. (1988) Ann.Rev.Biochem. 57 285-320

Firtel,R.A. (1972) J.Mol.Biol. 79 295-314

Firtel,R.A. and Bonner,J.T. (1972) J.Mol.Biol. 66 339-361

Firtel,R.A. and Jacobson,A. (1977) In: Biochemistry of Cell Differentiation II 15 377-429 J.Paul (ed.) Univ.Park Press, Baltimore

Firtel,R.A., Kindle,K. and Huxley,M.P. (1976) Fed.Proc. 35 13-22

Firtel,R.A., Silan,C., Ward,T.E., Howard,P., Metz,B.A., Nellen,W. and Jacobson,A. (1985) Mol.Cell.Biol. 5 3241-3250

Fischer,J.A., Giniger,E., Maniatis,T. and Ptashne,M. (1988) Nature 332 853-856

Fletcher,C., Heintz,N. and Roeder,R. (1987) Cell 51 773-781

Ford,A.M., Watt,S.M., Furley,A.J.W., Molgaard,H.V. and Greaves,M.F. (1988) EMBO J. 7 2393-2399

Forman,D. and Garrod,D.R. (1977) J.Embryol.Exp.Morphol. 40 215-228

Forrester,W.C., Thompson,C., Elder,J.T. and Groudine,M. (1986) PNAS 83 1359-1363

Franke,J. and Sussman,M. (1973) J.Mol.Biol. 81 173-185

Fromm,M. and Berg,P. (1982) J.molec.appl.Genet. 111 457-481

Garabedian,M.J., Hung,M.C. and Wensink,P.C. (1985) PNAS 82 1396-1400

Garabedian,M.J., Shepherd,B.M. and Wensink,P.C. (1986) Cell 45 859-867

Garcia,J.V., Stafford,J. and Queen,C. (1986) Nature 322 383-385

Gasser,S.M. and Laemmli,U.K. (1986) Cell 46 521-530

Gaub,M., Dierich,A., Anstinotti,D., Touitou,I. and Chambon,P. (1987) EMBO J. 6 2313-2320

Gaulton,G.N., Kelly,K.L., Pawlowski,J., Mato,J.M. and Jarrett,L. (1988) Cell 53 963-970

Gehring,W.J. (1987) Science 236 1245-1252

Georgopoulos,K., Van den Elsen,P., Bier,E., Maxam,A. and Terhorst,C. (1988) EMBO J. 7 2401-2407

Gerisch,G. (1961) Naturweissenschaften 48 436-437

Gerisch,G. and Wick,U. (1975) Biochem.Biophys.Res.Comm. 65 364-370

Gerisch,G., Hagmann,J., Hirth,P., Rossier,C., Weinhart,U. and Westphal,M. (1985) CSH Symp.Quant.Biol. 50 813-822

Gerster,T., Mathias,P., Thali,M., Jiricny,J. and Schaffner,W. (1987) EMBO J. 6 1323-1330

Gerster,T., Picard,D. and Schaffner,W. (1986) Cell 45 45-52

Gidoni,D., Kadonaga,J.T., Barrera-Saldna,H., Takahashi,K., Chambon,P. and Tjian,R. (1985) Science 230 511-517

Gigiere,V., Ong,E.S., Segui,P. and Evans,R.M. (1987) Nature 330 624-629

Gillies,S., Morrison,S., Oi,V. and Tonegawa,S. (1983) Cell 33 717-728

Giniger,E., Varnum,S.M. and Ptashne,M. (1985) Cell 40 767-774

Giri,J.G. and Ennis,H.L. (1978) Biochem.Biophys.Res.Commun. 77 282-289

Glass,C.K., Holloway,J.M., Devary,O.V. and Rosenfeld,M.G. (1988) Cell 54 313-323

Godbout,R., Ingram, R. and Tilghman, S.M. (1986) Mol.Cell. Biol. 6 477-487

Godbout,R. and Tilghman,S.M. (1988) Genes Dev. 2 949-956

Gomer,R.H., Armstrong,D., Leichtling,B.H. and Firtel,R.A. (1986) PNAS 83 8624-8628

Gomer,R.H. and Firtel,R.A. (1987) Science 237 758-762

Goodbourn,S., Zinn,K. and Maniatis,T. (1985) Cell 41 509-520

Gorman,C.M., Moffat,L.F. and Howard,B.H. (1982) Mol.Cell. Biol. 2 1044-1051

Grant,W.N., Welker,D.L. and Williams,K.L. (1985) Mol.Cell. Biol. 5 2559-2566

Grant,W.N. and Williams,K.L. (1983) EMBO J. 2 935-940

Graves,R.A., Pandey,N.B., Chodchoy,N. and Martzluff,W.F. (1987) Cell 4 615-626

Grayson,D.R.,Costa,R.H., Xanthopoulus,K.G. and Darnell,J.E. (1988) Science 239 786-788

Green,S. and Chambon,P. (1987) Nature 325 75-78

Greenberg,M.E., Hermanowski,A.L. and Ziff,E.B. (1986) Mol. Cell.Biol. 6 1050-1057

Greenwald,I.S., Sternberg,P.W. and Horvitz,H.R. (1983) Cell 34 435-444

Gregg,J.H. and Davis,R.W. (1982) Differentiation 21 200-205

Gregg,J.H., Krefft,M., Haas-Kraus,A. and Williams,K.L. (1982) Exp.Cell.Res. 142 229-233

Gross,D.S. and Garrad,W.T. (1988) Ann.Rev.Biochem. 57 159-197

Gross,J.D. ,Bradbury,J., Kay,R.R. and Peacey,M.J. (1983) Nature 303 244-245

Grosschedl,R. and Baltomore,D. (1985) Cell 41 885-897

Grosveld, F., Blom van Assendelft,G., Greaves,D.R. and Kollias,G. (1987) Cell 51 975-985

Groudine,M. and Weintraub,H. (1982) Cell 30 131-139

Grunstein,M. and Hogness,D. (1975) PNAS 72 3961-3963

Guarente,L. (1988) Cell 52 303-305

Hammer,R.E., Krumlauf,R., Camper,S.A., Brinster,R.L. and Tilghman,S.M. (1987) Science 235 53-58

Hanahan,D. (1983) J.Mol.Biol. 166 557-580

Hanahan,D. (1985) In; DNA Cloning Vol.1; A practical approach, 109-136. D.M.Glover (ed.) IRL Press

Harding,K., Rushlow,C., Doyle,H.J., Hoey,T. and Levine,M. (1986) Science 233 953-959

Haribabu,B. and Dottin,R.P. (1986) Mol.Cell.Biol. 6 2402-2408

Haribabu,B., Rakjovic,A. and Dottin,R.P. (1986) Dev.Biol. 113 436-442

He,H-T., Barbet,J., Chaix, J-C. and Goridis,C. (1986) EMBO J. 5 2489-2494

Heard,J.M., Herbomel,P., Ott,M-O., Mattura-Rollier,A., Weiss,M. and Yaniv,M. (1987) Mol.Cell.Biol. 7 2425-2434

Hemperley,J.J., Edelman,G.M. and Cunningham,B.A. (1986) PNAS 83 9822-9826

Henikoff,S. (1984) Gene 28 351-359

Herr,W. and Clarke,J. (1986) Cell 45 461-470

Hill,D.E., Hope,I.A., Macke,J.P. and Struhl,K. (1986) Science 234 451-457

Hoey,T. and Levine,M. (1988) Nature 332 858-861

Hohl,H.R. and Hamamoto,S.T.(1969) J.Ultrastruct.Res. 26 442-453

Hohmann,H-P., Bozzaro,S., Merkl,R.,Wallraff,E.,Yoshida,M., Weinhart,U. and Gerisch,G. (1987) EMBO J. 6 3663-3671

Holland,P.W.H. and Hogan,B.L.M. (1988) Genes Dev. 2 773-782

Holmes,D.S. and Quigley,M. (1981) Anal.Biochem. 114 193-197

Hope,I.A. and Struhl,K. (1985) Cell 43 177-188

Hope,I.A. and Struhl,K. (1986) Cell 46 885-894

Hope,I.A. and Struhl,K. (1987) EMBO J. 6 2781-2784

Ihara,M., Tanaka,Y., Yanagisawa,K., Taya,Y. and Nishimura,S. (1986) Biochem.Biophys.Acta 881 135-140

Imagawa,M., Chiu,R. and Karin,M. (1987) Cell 51 251-260

Ikeda,T. and Takeuchi,I. (1971) Dev.Growth Diff. 13 221-229

Ish-Horowitz,D. and Burke,J. (1981) Nucleic Acids Res. 9 2989-2992

Ishihara,A., Hou,Y. and Jacobson,K. (1987) PNAS 84 1290-

Jacquet,M., Part,D. and Felenbok,B. (1981) Dev.Biol. 81 155-166

Jacobson,A. (1976) In: Methods in Molecular Biology Vol. 8, 161-209. J.A.Last (ed.) Marcel Dekker,Inc. New York

Jantzen,H.M., Strahle,U., Gloss,B., Stewart,F., Schwid,W., Boshart,M., Miksicek,R. and Schutz,G. (1987) Cell 49 29-38

Jantzen,K. Fritton,H.P. and Igo-Kemenes,T. (1986) Nucleic Acid Res. 14 6085-6099

Jermyn,K.A., Berks,M., Kay,R.R. and Williams,J.G. (1987) Development 100 745-755

Jones,K.A., Kadonaga,J.T., Rosenfeld,P.J., Kelly,T.J. and Tjian,R. (1987) Cell 48 79-89

Jones,K.A., Yamamoto,K.R. and Tjian,R. (1985) Cell 42 559-572

Jones,N.C., Rigby,P.W.J. and Ziff,E.B. (1988) Genes Dev. 2 267-281

Kadonaga,J.T., Camer,K.R., Masiarz,S.R. and Tjian,R. (1987) Cell 51 1079-1090

Kay,R.R. (1982) PNAS 79 3228-3231

Kay,R.R., Dhokia,B. and Jermyn,K.A. (1983) Eur.J.Biochem. 136 51-56

Kay,R.R. and Jermyn,K.A. (1983) Nature 303 242-244

Keegan,L., Gill,G. and Ptashne,M. (1986) Science 231 699-704

Kelly,R., Kelly,L.J. and Ennis,H.L. (1985) Mol.Cell.Biol. 5 133-139

Keng,T. and Guarente,L. (1987) PNAS 84 9113-9117

Kimmel,A.R. (1987) Dev.Biol. 122 163-171

Kimmel,A.R. and Carlisle,B. (1986) Dev.Biol.83 2506-2510

Kimmel,A.R. and Firtel.R.A. (1979) Cell 16 787-796

Kimmel,A.R. and Firtel.R.A. (1980) Nucleic Acids Res. 8 5599-5610

Kimmel,A.R. and Firtel.R.A. (1982) In: The Development of *Dictyostelium discoideum*, 233-324. W.F.Loomis (ed.) Academic Press

Kimelman,D. and Kirschner,M. (1987) Cell 51 869-877

Kingston,R.E., Schuetz,T.J. and Larin,Z. (1987) Mol.Cell.

Biol. 7 1530-1534

Kindle,K.L. and Firtel.R.A. (1978) Cell 15 763-778

Klein,P.A., Fontana,D., Knox,B., Theibert,A. and Devreotes,P.N. (1985a) CSH Symp.Quant.Biol. 50 787-799

Klein,P.A., Theibert,A., Fontana,D. and Devreotes,P.N. (1985b) J.Biol.Chem. 260 1757-1764

Klock,G., Strahle,U. and Schutz,G. Nature 329 734-736

Knecht,D.A., Cohen,S.M., Loomis,W.F. and Lodish,H.F. (1986) Mol.Cell.Biol. 6 3973-3983

Knecht,D.A., Fuller,D. and Loomis,W.F. (1987) Dev.Biol. 121 277-283

Knecht,D.A. and Loomis,W.F. (1987) Science 236 1081-1091

Kongsuwan,K., Webb,E., Housiaux,P. and Adams,J.M. (1988) EMBO J. 7 2131-2138

Kopachik,W. (1982) J.Embryol.Exp.Morphol. 68 23-35

Kopachik,W., Oohata,A., Dhokia,B., Brookman,J.J. and Kay.R.R. (1983) Cell 33 397-403

Kopachik,W., Dhokia,B. and Kay.R.R. (1985) Differentiation 25 209-216

Krefft,M., Voet,L., Mairhofer,H. and Williams,K.L. (1983) Exp.Cell Res. 147 235-239

Krefft,M., Voet,L., Gregg,J.H., Mairhofer,H. and Williams, K.L. (1984) EMBO J. 3 201-206

Krefft,M., Voet,L., Gregg,J.H. and Williams,K.L. (1985) J.Embryol.Exp.Morph. 88 15-24

Krefft,M., Voet,L., Mairhofer,H. and Williams,K.L. (1983) Exp.Cell. Res. 147 235-239

Kugler,W., Wagner,U. and Ryffel,G.U. (1988) Nucleic Acids Res. 16 3165-3174

Kumagai,A. and Okamoto,K. (1986) Differentiation 31 79-84

Kyte,J. and Doolittle,R.F. (1982) J.Mol.Biol. 157 105-132

Laemmli,U.K. (1970) Nature 227 680-685

Laimins,L.A., Khoury,G.,Gorman,C., Howard,B. and Gruss,P. (1982) PNAS 79 6345-6457

Lam,T.Y., Pickering,G., Geltosky,J. and Siu,C.H. (1981) Differentiation 20 22-28

Landfear,S.M., Lefebvre,P., Chung,S. and Lodish,H.F. (1982) Mol.Cell.Biol. 2 1417-1426

Landolfi,N.F., Capra,J.D. and Tucker.P.W. (1986) Nature 323

548-551

Laskey,R.A. and Mills,A.D. (1975) Eur.J.Biochem. 56 335-341

Laughon,A. and Scott,M.P. (1984) Nature 310 25-31

Lawson,G., Knoll,B., March,C., Woo,S., Tsai,M. and O'Malley,B. (1982) J.Biol.Chem 257 1501-1507

Leach,C.K.,Ashworth,J.M. and Garrod,D.R. (1973) J.Embryol. Exp.Morph. 29 647-661

Lee,W., Mitchell,P. and Tjian,R. (1987) Cell 49 741-752

Leegwater,P.A., van Driel,W. and van der Vliet,P.C. (1985) EMBO J. 4 1515-1521

Legerski,R.J., Hodnett,J.L. and Gray,H.B. (1978) Nucleic Acids Res. 5 1445-1464

Lehmann,R., Jiminez,F., Dietrich,U. and Campos-Ortega,J.A. (1983) Roux's Arch.Dev.Biol. 192 62-74

Lenardo,M., Pierce.J.W. and Baltimore,D. (1987) Science 236 1573-1577

Lodish,H.F., Blumberg,D.D., Chisholm,R, Chung,S., Coloma,A., Landfear,S., Barklis, E., Lefevbre,P., Zuker,C and Mangiarotti,G. (1982) In: The Development of *Dictyostelium discoideum*, 325-352. W.F.Loomis (ed.) Academic Press

Lodish,H.F., Jacobson,A., Firtel,R.A., Alton,T. and Tuchman,J. (1974) PNAS 71 5103-5108

Loomis,W.F. (1982) The Development of *Dictyostelium discoideum* Academic Press

Loomis,W.F., Knecht,D.A. and Fuller,D. (1987) In: Molecular Approaches to Developmental Biology, 339-349. Alan Liss Inc. NY

Low,M.G., Ferguson,M.A.J., Futerman,A.H. and Silman,I. (1986) Trends Biochem.Sci. 11 212-215

Low,M.G. and Kincade,P.W. (1985) Nature 318 62-64

Macdonald,P.M., Ingram,P.W. and Struhl,G. (1986) Cell 47 721-734

McDonald,S.A. (1984) Wilhelm Roux Arch.Dev.Biol. (1984) 194 50-52

McDonald,S.A. (1986) Dev.Biol. 117 546-549

McDonald,S.A. and Durston,A.J. (1984) J.Cell Sci. 66 195-204

McGinnis,W. (1985) CSH Symp.Quant.Biol. 50 263-270

McGinnis,W., Hart,C.P., Gehring,W.F. and Ruddle.F.H. (1984)

Cell 38 675-680

McKeown,M. and Firtel,R.A. (1981) Cell 24 799-807

McKnight,S. and Tjian,R. (1986) Cell 46 795-805

McRobbie,S.J. and Ceccarelli,A. (1988) Nucleic Acids Res. 16 4738

McRobbie,S.J., Tilly,R., Blight,K., Ceccarelli,A. and Williams,J.G. (1988) Dev.Biol.

MacWilliams,H.K., Blaschke,A. and Prause,I. (1984) In: Microbial Development, 779-785. R.Losick and L.Shapiro (eds.) CSH Press

MacWilliams,H.K. and Bonner,J.T. (1979) Differentiation 14 1-22

Maeda,Y. (1977) Dev.Growth Differ. 19 201-205

Malchow,D., Nagete,B., Schwarz,H. and Gerisch,G. (1972) Eur J.Biochem. 28 136-142

Mangiarotti,G., Ceccarelli,A. and Lodish,H.F. (1983b) Nature 301 616-618

Mangiarotti,G., Chung,S., Zuker,C. and Lodish,H.F. (1981) Nucleic Acids Res. 9 947-963

Mangiarotti,G., Giorda,R., Ceccarelli,A. and Perlo,C. (1985) PNAS 82 5786-5790

Mangiarotti,G., Lefevbre,P. and Lodish,H.F. (1982) Dev.Biol. 89 82-91

Mangiarotti,G., Zuker,C., Chisholm,R.L. and Lodish,H.F. (1983a) Mol.Cell.Biol. 3 1511-1517

Maniatis,T., Fritsch,E.F. and Sambrook, J. (1982) "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Maniatis,T., Goodbourn,S. and Fischer,J.A. (1987) Science 236 1237-1245

Mann,S.K.O. and Firtel.R.A. (1987) Mol.Cell.Biol. 7 458-469

Margolskee,J.P. and Lodish,H.F. (1980) Dev.Biol. 74 37-49

Marin,F.T. (1976) Dev.Biol. 60 389-395

Mason,J.O., Williams,G.T. and Neuberger,M.S. (1985) Cell 41 479-487

Martinez,E., Givel,F. and Wahli,W. (1987) EMBO J. 6 3719-3727

Matsunaga,T. and Mori,N. (1987) Scand.J.Immunol. 25 485-495

Matskuma,S. and Durston,A.J. (1979) J.Embryol.Exp.Morph. 50 243-251

Matsui,T., Segall,J., Weil,P.A. and Roeder,R.G. (1980) J.Biol.Chem. 255 11992-11996

Maxam,A. and Gilbert,W. (1980) Methods Enzymol. 65 499-560

Mehdy,M.C. and Firtel,R.A. (1985) Mol.Cell.Biol. 5 705-713

Mehdy,M.C., Ratner,D. and Firtel,R.A. (1983) Cell 32 763-771

Meinhardt,H. (1983) Differentiation 24 191-202

Mitchell,P.J., Wang,C. and Tjian,R. (1987) Cell 50 847-861

Mills,F.C., Fisher,L.M., Kuroda,R., Ford,A.M. and Gould,H.J. (1983) Nature 306 809-812

Montminy,M.R. and Bilezikjian,L.M. (1987) Nature 328 175-178

Moreau,P., Hen,R., Wasylyk,B., Everett,R., Gaub,M.P. and Chambon,P. (1981) Nucleic Acids Res. 9 6047-6067

Morris,H.R., Taylor,G.W., Masento,M.S., Jermyn,K.A. and Kay, R.R. (1987) Nature 328 811-814

Morrissey,J.H. (1982) In: The Development of *Dictyostelium discoideum*, 411-449. W.F.Loomis (ed.) Academic Press

Morrissey,J.H., Farnsworth,P.A. and Loomis,W.F. (1981) Dev. Biol. 83 1-8

Morrissey,J.H., Devine,K.M. and Loomis,W.F. (1984) Dev.Biol. 103 414-424

Muller,K. and Gerisch,G. (1978) Nature 274 445-449

Muller,U. and Hohl,H.R. (1973) Differentiation 1 267-276

Munro,S. and Pelham,H.R.B. (1986) Cell 46 291-300

Murray,B.A., Wheeler,S., Jongens,T. and Loomis,W.F. (1984) Mol.Cell.Biol. 4 514-519

Myers,R.M., Rio,D.C., Robbins,A.K. and Tjian,R. (1981) Cell 25 373-384

Myers,R.M., Tilly,K. and Maniatis,T. (1986) Science 232 613-618

Nagata,K., Guggenheimer,R.A. and Hurvitz,J. (1983) PNAS 80 6177-6181

Nellen,W. and Firtel,R.A. (1985) Gene 39 155-163

Nellen,W., Silan,C. and Firtel,R.A. (1984) Mol.Cell.Biol. 4 2890-2898

Nellen,W., Silan,C., Saur,U. and Firtel,R.A. (1986) EMBO J. 5 3367-3372

Nelson,C. Albert,V.R., Elsholtz,H.P., Lu,L.I-W. and Rosenfeld,M.J. (1988) Science 239 1400-1405

Neuberger,M.S. (1983) EMBO J. 2 1373-1378

Newell,P.C. (1982) In: The Development of *Dictyostelium discoideum*, 35-70. W.F.Loomis (ed.) Academic Press

Newell,P.C. and Ross,F.M (1982) J.Gen.Microbiol. 128 2715-2724

Noda,M., Yoon,K., Radan,G.A. and Koppel,D.E. (1987) J.Cell. Biol. 105 1671-1677

Noegel,A., Gerisch,G., Stadler,J. and Westphal,M. (1986) EMBO J. 5 1473-1476

Noegel,A., Harloff,C., Hirth,P., Merkl,R., Modersitzki,M., Stadler,J., Weinhart,U., Westphal,M. and Gerisch,G. (1985) EMBO J. 4 3805-3010

Ochai,H., Stadler,J., Westphal,M., Wagle,G., Merkl,R. and Gerisch,G. (1982) EMBO J. 1 1011-1016

Odell,G.M. and Bonner,J.T. (1986) Phil.Trans.R.Soc.Lond. B312 487-525

Ondek,B., Gloss,L. and Herr,W. (1988) Nature 333 40-45

Ong,D.E., Crow,J.A. and Chytil,F. (1982) J.Biol.Chem. 257 13385-13389

Osborne,T.F. Goldstein,J.L. and Brown,M.S. (1985) Cell 42 203-212

Oyama,M. and Blumberg,D.D. (1986a) Dev.Biol. 117 550-556

Oyama,M. and Blumberg,D.D. (1986b) PNAS 83 4819-4823

Oyama,M. and Blumberg,D.D. (1986c) Dev.Biol. 117 557-566

Padgett,R.W., St.Johnston,R.D. and Gelbart,W.M. (1987) Nature 325 81-84

Palatnik,C.M., Storti,R.V. and Jacobson,A. (1979) J.Mol.Biol. 128 371-395

Palatnik,C.M., Storti,R.V., Capone,A.K. and Jacobson,A. (1980) J.Mol.Biol. 141 99-118

Palatnik,C.M., Wilkins,C. and Jacobson,A. (1984) Cell 36 1017-1025

Parker,C.S. and Topol,J. (1984) Cell 37 273-283

Parslow,T.G., Blair,D.L., Murphy,N.J. and Granner,D.K. (1984) PNAS 81 2650-2654

Parslow,T.G. and Granner,D.K. (1982) Nature 299 449-451

Pavlovic,J. Banz,E. and Parish,R.w. (1986) Nucleic Acids Res. 14 8703-8722

Ptashne,M. (1986) Nature 322 697-701

Pears,C.J. (1987) Ph.D. thesis, University of London.

Pears,C.J., Mahbubani,H.M. and Williams,J.G. (1985) Nucleic Acids Res. 13 8853-8866

Pears,C.J. and Williams,J.G. (1987) EMBO J. 6 195-200

Pears,C.J. and Williams,J.G. (1988) Nucleic Acid Res. 16 8467-8486

Pelham,H.R.B. (1982) Cell 30 517-528

Pelham,H.R.B. (1985) Trends Genet. 1 31-35

Peterson,C.L. Orth,K. and Calame,K.L. (1986) Mol.Cell.Biol. 6 4168-4178

Petkovich,M., Brand,N.J., Krust,A. and Chambon,P. (1987) Nature 330 444-450

Phelps,B.M., Primkoff,P., Koppel,D.E., Low,M.G. and Myles,D.G. (1988) Science 240 1780-1782

Picard,D. and Yamamoto,K.R. (1987) EMBO J. 6 3333-3340

Pinkert,C.A. Ornitz,D.M., Brinster,R.L. and Palmiter,R.D. (1987) Genes Dev. 1 268-276

Pinney,D.F., Pearson-White,S.H., Konieczny,S.F., Latham,K.E. and Emersom,C.P.Jr. (1988) Cell 53 781-793

Poole,S.J. and Firtel,R.A. (1984) J.Mol.Biol. 172 203-220

Powell,D.J., Friedman,J.M., Oulette,A.J., Krauler,K.S. and Darnell J.E. (1984) J.Mol.Biol. 179 21-35

Quinn,P.G., Wong,T.W., Magnuson,M.A., Schabb,J.B. and Granner,D.K. (1988) Mol.Cell.Biol. 8 3467-3475

Raper,K.B. (1940) J.E.Mitchell Sci.Soc. 56 241-282

Raper,K.B. and Fennell,D.I. (1952) Bull.Torrey Bot.Club 79 25-51

Ratner,D. and Borth,W. (1983) Exp.Cell Res. 143 1-13

Rawlins,D.R., Rosenfeld,P.J.,Wides,R.J., Challberg,M.D. and Kelly,T.J. (1984) Cell 37 309-319

Rauscher,F.J.III, Cohen,D.R., Curran,T., Bos,T.J., Vogt,P.K., Bohmann,D., Tjian,R. and Franza,B.R.Jr. (1988) Science 240 1010-1016

Raymondjean,M., Cereghini,S. and Yaniv,M. (1988) PNAS 85 757-761

Reiser,H. Oettgen,H., Yeh,E.T.H., Terhorst,C., Low,M.G., Bernacerraf,B. and Rock,K.L. (1986) Cell 47 365-370

Reymond,C.D, Nellen,W. and Firtel,R.A. (1985) PNAS 82

7005-7009

Rigby,P.W., Dieckman,M., Rhodes,C. and Berg,P. (1977) J.Mol. Biol. 113 237-251

Roesler,W.J., Vandenbank,G.R. and Hanson,R.W. (1988) J.Biol. Chem. 263 9063-9066

Romans,P. and Firtel,R.A. (1985) J.Mol.Biol. 183 311-326

Roos,W., Nanjundiah,U., Malchow,D. and Gerisch,G. (1975) FEBS Lett. 53 139-142

Rosenberg,U.B., Schroder,C., Priess,A., Kienlin,A., Cote,S., Riede,I. and Jackle,H. (1986) Nature 319 336-339

Roth,J. (1983) In: Techniques in Immunocytochemistry, 217-284. G.R.Bullock and P.Petrusz (eds.) Academic Press NY

Rowekamp,W. and Firtel,R.A. (1980) Dev.Biol. 79 409-418

Rubin,J. and Robertson,A. (1975) J.Embryol.Exp.Morphol. 33 227-241

Sadeghi,H., da Silva,A.M. and Klein,C. (1988) PNAS 85 5512-5515

Saint,R., Kaliois,B., Lockett,T.J. and Elizur,A. (1988) Nature 334 151-154

Sakai,Y. (1973) Develop.Growth Differ. 15 11-19

Salser,W.A., Davis,R.C., Chou,C-C, Teraoka,S.N., Thomason,A., Nelson,R. and Slovin,J. (1987) In: Molecular approaches to Developmental Biology, 509-520. Alan R.Liss, Inc. NY

Samuels,N., Fire,A. and Sharp,P.A. (1982) J.Biol.Chem. 257 14419-14427

Sanger,F., Nicklen,S. and Coulsen,A.R. (1977) PNAS 74 5463-5467

Santoro,C., Mermod,N., Andrews,P.C. and Tjian,R. (1988) Nature 324 218-224

Sassone-Corsi,P. and Borelli,E. (1986) Trends Genet. 2 215-219

Sassone-Corsi,P., Lamph,W.W., Kamps,M. and Verma,I.M. (1988) Cell 54 553-560

Saxe,C.L.III and Firtel,R.A. (1986) Dev.Biol. 115 407-414

Schaap,P. (1986) Differentiation 33 1-16

Schaap,P., Pinas,J.E. and Wang,M. (1985) Dev.Biol. 111 51-61

Schaap,P., Konijn,T.M. and Van Haastert,P.J.M. (1984) PNAS 81 2122-2126

Schaap,P. and Spek,W. (1984) Differentiation 27 83-87

Schaap,P., Van Der Molen,L. and Konijn,T.M. (1982) Differentiation 22 1-5

Schaap,P. and Van Driel,R. (1985) Exp.Cell Res. 159 388-398

Schaap,P., Van Lookeren Campagne,M., Van Driel,R., Spek, W., Van Haastert,P.J.M. and Pinas, J. (1986) Dev.Biol. 118 52-63

Schaap,P. and Wang,M. (1984) Dev.Biol. 105 470-478

Schaap,P. and Wang,M. (1986) Cell 45 137-144

Scheidereit,C., Heguy,A. and Roeder,R.G. (1987) Cell 51 783-793

Schindler,J. and Sussman,M. (1977) J.Mol.Biol. 116 161-169

Schirm,S, Jiricny,J. and Schaffner,W. (1987) Genes Dev. 1 65-74

Schlokat,U., Bohmann,D, Scholer,H. and Gruss,P. (1986) EMBO J. 5 3258-3266

Schofield,P.N. (1987) Trends Neurosci 10 3-6

Schuh,R., Aicher,W., Gaul,U., Cte,S., Priess,A., Maier,D., Seifert,E., Nauber,U., Schroder,C., Kemler,R. and Jackle,H. (1986) Cell 47 1025-1032

Scott,M.P. and Carroll,S.B. (1987) Cell 51 689-698

Scott,M.P. and Weiner,A.J. (1984) PNAS 81 4115-4119

Scott,R.W. and Tilghman,S.M. (1983) Mol.Cell.Biol 3 1295-1309

Sen,R. and Baltimore,D. (1986) Cell 46 705-716

Searle,P.F., Stuart,G.W. and Palmiter,R.D. (1985) Mol.Cell. Biol. 5 1480-1489

Seed,B. (1987) Nature 329 840-842

Seiler-Tuyns,A and Paterson,B.M. (1987) Mol.Cell.Biol. 7 1048-1054

Serfling,E., Jasin,M. and Schaffner,W. (1985) Trends Genet. 1 224-230

Serfling,E., Lubbe,A., Dorsch-Hasler,K. and Schaffner,W. (1985) EMBO J 4 3851-3859

Shepherd,J.C.W., McGinnis,W., Carrasco,A.E., De Robertis,E.M. and Gehring,W.J. (1984) Nature 310 70-71

Short,J.M., Wynshaw-Boris,A., Short,H.P. and Hanson,R.W. (1986) J.Biol.Chem. 261 9721-9726

Shuey,D.J. and Parker,C.S. (1986a) J.Biol.Chem. 261 7934-7940

Shuey,D.J. and Parker,C.S. (1986b) Nature 323 459-461

Singh,H., LeBovitz,J.H., Baldwin,A.S.Jr. and Sharp,P.A.

(1988) Cell 52 415-423

Singh,H., Sen,R., Baltimore,D. and Sharp,P.A. (1986) Nature 319 154-158

Sive,H.L. and Roeder,R.G. (1986) PNAS 83 6382-6386

Slack,J.M.W. (1987) Nature 327 553-554

Slack,J.M.W., Darlington,B.G., Heath,J.K. and Godsave,S.F. (1987) Nature 326 197-200

Sobelewski,A., Neave,N. and Weeks,G. (1983) Differentiation 25 93-100

Soll,D.R. and Finney,R. (1986) In: The Genetic Regulation of Development, W.F.Loomis (ed.). 45th Symposium of the Society for Developmental Biology

Southern,E. (1975) J.Mol.Biol. 98 503-517

Staudt,L.M., Clerc,R.G., Singh,H., LeBointz,J.H., Sharp,P.A. and Baltimore,D. (1988) Science 241 577-580

Staudt,L.M., Singh,H., Sen,R., Wirth,T., Sharp,P.A. and Baltimore,D. (1986) Nature 323 640-643

Steinberg,M.S. (1970) J.Exp.Zool. 173 395-433

Steinemann,C. and Parish,R. (1980) Nature 286 621-623

Sternberg,P.W. and Horvitz,H.R. (1986) Cell 44 761-772

Sternfield,J. and Bonner,J.T. (1977) PNAS 74 268-271

Sternfield,J. and David,C.N. (1979) J.Cell Sci. 38 181-191

Sternfield,J. and David,C.N. (1982) Dev.Biol. 93 111-118

Strahle,U., Klock,G. and Schutz,G. (1987) PNAS 84 7871-7875

Struhl,K. (1986) Cell 49 295-297

Struhl,K. (1987) Cell 50 841-846

Struhl,K. (1988) Nature 332 649-650

Sussman,M. (1966) In: Methods in Cell Physiology, 397-415. D.Prescott (ed.) Academic Press, NY

Sussman,R.R. and Rayner,E.P. (1971) Arch.Biochem.Biophys. 144 127-137

Takahashi,K. (1986) Nature 319 121-126

Takeuchi,I. (1963) Dev.Biol. 8 1-26

Takeuchi,I. (1969) In: Nucleic Acid Metabolism, Cell Differentiation and Cancer Growth, 297-304. E.V.Coudry and S.Seno (eds.) Pergammon Press, Oxford

Tasaka,M., Noce,T. and Takeuchi,I. (1983) PNAS 80 5340-5344

Tasaka,M. and Takeuchi,I. (1981) Differentiation 18 191-196

Tautz,D., Lehmann,R., Schnurch,H., Schuh,R., Seifert,E.,

Kienli,A., Jones,K. and Jackle,H. (1987) Nature 327 383-389

Thaller,C. and Eichele,G. (1987) Nature 327 625-628

Theibert,A. and Devreotes,P.N. (1984) Dev.Biol. 106 166-173

Theisen,M., Stief,A. and Sippel,A.E. (1986) EMBO J. 5 719-724

Thomas,D.A. and Wright,B.E. (1976) J.Biol.Chem 251 1258-1263

Thomas,J., Webb,W., Davitz,M.A. and Nussenveig,V. (1987) Biophys.J. 51 522a

Tickle,C., Alberts,B., Wolpert,L. and Lee,J. (1982) Nature 296 564-566

Tilghman,S.M. (1985) In: Oxford Surveys on Eukaryotic Genes, 2 160- N.McLean (ed.) Oxford University Press, NY

Tomachik,K.J. and Devreotes,P.N. (1981) Science 212 443-446

Tomlinson,A. and Ready,D.F. (1986) Science 231 400-402

Town,C.D., Gross,J.D. and Kay,R.R. (1976) Nature 262 717-719

Treisman,R. (1985) Cell 42 889-902

Tsang,A. and Bradbury,J.M. (1981) Exp.Cell Res. 132 433-441

Tsang,A., Mahbubani,H. and Williams,J.G. (1982) Cell 31 375-382

Tsang,A.S. and Tasaka,M. (1986) J.Biol.Chem. 261 10753-10759

Tse,A.G.D., Barclay,A.N., Watts,A. and Williams,A.F. (1985) Science 230 1003-1008

Tuan,D., Solomon,W., Qiliang,L.S. and Irving,M.L. (1985) PNAS 32 6384-6388

Van Haastert,P.J.M. (1985) Biochim.Biophys.Acta 846 324-333

Van Haastert,P.J.M. and De Wit,R.J.W. (1984) J.Biol.Chem. 259 13321-13328

Van Haastert,P.J.M., De Wit,R.J.W., Janssens,P.M.W., Kesbeke, F. and DeGoede,J. (1986) J.Biol.Chem. 261 6904-6911

Vardy,P.H., Fisher,L.R., Smith,E. and Williams,K.L. (1986) Nature 320 526-529

Vogt,P.K., Bos,T.J. and Doolittle,R.F. (1987) PNAS 84 3316-3319

Walker,M.D., Karlsson,O., Edlund,T., Barnett,J., Nir,U. and Rutter,W.J. (1987) In; Molecular Approaches to Developmental Biology, 619-627. (ed.) Alan R.Liss, Inc.

Wang,M., Aerts,R.J., Spek,W. and Schaap,P. (1988a) Dev.Biol. 125 410-416

Wang,M., Schaap,P. and Van Haastert,P.J.M. (1986)

Differentiation 33 24-28

Wang,M., Van Driel,R. and Schaap,P. (1988b) Development 103 611-618

Wang,S-Y, LaRosa,G.J. and Gudas,L.J. (1985) Dev.Biol. 107 75-86

Wang,X.F. and Calame,K. (1986) Cell 47 241-247

Watts,D.J. and Ashworth,J.M. (1970) Biochem.J. 119 171-174

Way,J.C. and Chalfie,M. (1988) Cell 54 5-16

Weeks,D.L. and Melton,D.A. (1987) Cell 51 861-867

Weijer,C.J. and Durston,A.J. (1985) J.Embryol.Exp.Morph. 86 19-37

Weijer,C.J., Duschl,G. and David,C.N. (1984) Differentiation 28 13-23

Weinberger J., Baltimore,D. and Sharp,P.A. (1986) Nature 322 846-848

Weiss,M.C. (1982) In: Somatic Cell Genetics, 169-179. C.T.Casey and D.C.Robbans (eds.) NY Plenum Press

West,B.L., Catanzaro,D.F., Melton,S.H., Cattini,P.A., Baxter, J.D. and Reudelhuber,T.L. (1987) Mol.Cell.Biol. 7 1193-1197

West,C.M. and Loomis,W.F. (1985) J.Biol.Chem. 260 13803-13809

West,R.W.Jr., Yocum,R. and Ptashne,M. (1984) Mol.Cell.Biol. 4 2467-2478

White,R.A.H. and Wilcox,M. (1984) Cell 39 163-

Wijnholds,J., Philipsen,J.N.J. and AB,G. (1988) EMBO J. 7 2757-2763

Wilcox,D. and Sussman,M. (1981) PNAS 78 358-362

Wilde,S.M., Wedden,S.E. and Tickle,C. (1987) Development 100 723-733

Wildeman A.G. (1986) Mol.Cell.Biol. 6 2098-2105

Wilkinson,D.G., Wilson,J. and Hames,B.D. (1985) Dev.Biol. 107 38-46

Williams,J.G. (1988) Development 103 1-16

Williams,J.G., Ceccarelli,A., McRobbie,S., Mahbubani,H., Kay, R.R., Early.A., Berks,M. and Jermyn,K.A. (1987) Cell 49 185-192

Williams,J.G., Lloyd,M.M. and Devine,J.M. (1979) Cell 17 903-913

Williams,J.G. and Mason,P. (1985) In: Nucleic Acid Hybridis-

ation: A practical approach, 139-160. B.D.Hames and S.J. Higgins (eds.) IRL Press

Williams,J.G., North,M.J. and Mahbubani,H. (1985) EMBO J. 4 999-1006

Williams,J.G., Tsang,A.S. and Mahbubani,H. (1980) PNAS 77 7171-7175

Williams,K.L., Grant,W.N., Krefft,M., Voet,L. and Welker,D.L. (1984) In: Molecular Biology of Development, 75-90. ?(ed.) Alan R.Liss, Inc.

Wirth,T., Staudt,L. and Baltimore,D. (1987) Nature 329 174-178

Witke,W., Nellen,W. and Noegel,A. (1987) EMBO J. 6 4143-4148

Woffendin,C., Chambers,T.G., Schaller,K.L., Leichtling,B.H. and Rickenberg,H.V. (1986) Dev.Biol. 115 1-8

Yamamoto,A., Maeda,Y. and Takeuchi,I. (1981) Protoplasma 108 55-70

Yamamoto,K.K., Gonzales,G.A., Biggs,W.H.III and Montminy,M.R. (1988) Nature 334 494-498

Yamamoto,K.R. (1985) Annu.Rev.Genet. 19 209-252

Yokum,R.R., Hanley,S., West,R.W. Jr. and Ptashne,M. (1984) Mol.Cell.Biol 4 1985-1988

Zarat,K.S. and Yamamoto,K.R. (1984) Cell 38 29-38

Zenke,M., Grundtstrom,T., Mattes,H., Wintzereith,M., Schatz,C., Wildeman,A. and Chambon,P (1986) EMBO J. 5 378-397

Zimarino,V. and Wu,C. (1987) Nature 327 727-730

Gene, 59 (1987) 99–106
Elsevier

GEN 02161

Two vectors which facilitate gene manipulation and a simplified transformation procedure for *Dictyostelium discoideum*

(Recombinant DNA; cellular slime mould; transformation; vector; neomycin resistance)

A.E. Early and J.G. Williams

Imperial Cancer Research Fund, South Mimms, Herts. EN6 3LD (U.K.)

Received 26 February 1987
Revised 13 April 1987
Accepted 22 April 1987

SUMMARY

We have constructed and characterized two *Dictyostelium* transformation vectors (pB10TP1 and pB10TP2) designed for the facile sequence determination, mutagenesis and functional analysis of *Dictyostelium* genes. The vectors incorporate the B10 neomycin-resistance (*neo*) gene [Nellen et al., Mol. Cell. Biol. 4 (1984) 2890–2898] and sequences derived pEMBL18+ [Dente et al., Nucl. Acids Res. 11 (1983) 1645–1655], enabling the production of single-stranded template and increasing the yield of double-stranded DNA. A new multiple cloning site (MCS) has been inserted adjacent to the M13 sequence primer binding site so that single-stranded template DNA isolated from recombinants prepared using these vectors is suitable for sequence analysis and site-directed mutagenesis. The linker incorporates restriction sites suitable for the preparation of a directed deletion series and useful in cloning, including some sites with recognition sequences frequent in the extremely A + T-rich *Dictyostelium* genome. A *Dictyostelium* genomic fragment has been included to provide transcription termination signals for the *neo* gene. One of the two vectors (pB10TP1) contains the 3′-proximal portion of a constitutively expressed mRNA of unknown function. It is located downstream from the MCS so that 5′-proximal fragments of genes, cloned into the MCS, generate fusion transcripts which are distinguishable from transcripts of the corresponding endogenous genes. The complete nucleotide sequence of the two vectors has been established and a comprehensive restriction map deduced. We also describe a modification of the published transformation system, which allows it to be applied to the commonly used strain Ax-2, and another generally applicable modification which greatly reduces the time required to obtain stable transformants.

INTRODUCTION

Upon starvation, individual *Dictyostelium* amoebae form multicellular aggregates which undergo differentiation to yield a mass of spores supported by a stalk comprised of dead, cellulose-ensheathed cells (see Loomis, 1982, for review). During development there are major changes in gene expression and a

Correspondence to: Dr. J.G. Williams, Imperial Cancer Research Fund, South Mimms, Herts. EN6 3LD (U.K.) Tel. 0707-44444.

Abbreviations: Ap, ampicillin; bp, base pair(s); cAMP, adenosine 3′:5′-cyclic monophosphate; CP1, cysteine proteinase 1; kb, kilobase pairs of 1000 bp; MCS, multiple cloning site; MCS′, see Fig. 4; MOPS, 3-(*N*-morpholino)propanesulphonic acid; *neo*, neomycin-resistance gene; nt, nucleotide(s); *ori*, origin of DNA replication; ^{R}, resistance; ss, single stranded; Tn, transposon.

0378-1119/87/$03.50 © 1987 Elsevier Science Publishers B.V. (Biomedical Division)

large number of developmentally regulated genes have been isolated (Williams et al., 1979; Roewekamp and Firtel, 1980; Barklis and Lodish, 1983; Mehdy et al., 1983). The expression of these genes has been extensively studied at a cellular level (Williams et al., 1980; Chisholm et al., 1984; Mehdy and Firtel, 1985; Oyama and Blumberg, 1986), but identification of DNA elements controlling their expression at a molecular level requires their transfection into *Dictyostelium*. The transformation vector B10 contains the neomycin resistance gene from Tn*5* fused to the promoter of the actin *6* gene (Nellen et al., 1984). After introduction as a calcium phosphate coprecipitate, the vector confers resistance to the amino-glycoside G418. Upon drug selection, the vector is found as a tandem linear array which co-migrates with high-M_r DNA. The number of copies increases proportionately with the G418 concentration used in the selection. Using B10, very large transcripts of the Tn*5* gene result because the vector does not contain a signal for transcription termination downstream from the fusion gene. The derivative B10S contains a region from the actin *8* gene which, although theoretically in the incorrect orientation to act as a terminator, does function to yield a discrete transcript (Nellen and Firtel, 1985).

We have used the actin *6* Tn*5* *neo* fusion gene from B10 to construct the vectors pB10TP1 and pB10TP2 with the aim of facilitating manipulation of *Dictyostelium* DNA prior to transformation. The *Dictyostelium* genome has an unusual structure in that the non-coding regions almost entirely consist of A and T residues organised in homopolymer runs or alternating dA-dT tracts. These regions are normally devoid of restriction sites and hence conventional approaches to sequence manipulation are difficult to apply. The vectors incorporate several features designed to circumvent this problem, while retaining a transformation efficiency comparable to B10S (Nellen and Firtel, 1985).

We have also modified the transformation procedure in several ways. The original procedure was developed for the Ax-3 strain of *Dictyostelium*. We find that it is not directly applicable to the Ax-2 strain and we have developed conditions which allow its transformation. The original procedure includes a step where cells are grown clonally on bacterial lawns. This is inconvenient and we describe an altered transformation procedure whereby this step is eliminated.

EXPERIMENTAL AND DISCUSSION

(a) Construction of pB10

The vector pB10 (Fig. 1) was constructed by insertion of the *Eco*RI-*Bam*HI fragment from B10 (Nellen et al., 1984) into *Eco*RI + *Bam*HI-cleaved pEMBL18+ (Dente et al., 1983). The pEMBL family of vectors contain the *ori* of the phage f1 and, upon super-infection of cells containing the vector with the phage, vector ss DNA is packaged into phage particles and released from the host (Dente et al., 1983). The f1 *ori* remains intact in pB10 and the vector produces ss DNA suitable for sequence analysis by the chain-termination method (Sanger et al., 1977). The ss DNA can be used for other purposes, including the production of ss probes for use in S1 mapping and for performing site-directed mutagenesis (Zoller and Smith, 1982). The latter method is of great importance in manipulating DNA devoid of useful restriction-enzyme cleavage sites. It has been used in this laboratory to create a novel restriction-enzyme cleavage site in a highly A + T-rich region to perform fine mapping of a putative control element (Pears and Williams, 1987).

In addition, the pEMBL sequence increases the yield of double-stranded DNA obtained relative to that given by B10. This is because sequences flanking the plasmid *ori* have been mutated, preventing the strict control of copy number which occurs in ColE1-based plasmids. On average, 2–4 μg of plasmid DNA is produced per ml of *Escherichia coli* culture without the necessity for chloramphenicol amplification. This feature is very useful, as comparatively large amounts of DNA are used in *Dictyostelium* transformation.

(b) Construction and characterisation of pB10T

In pB10 the *act* 6::*neo* fusion gene lacks a *Dictyostelium* terminator, and a heterodisperse series of transcripts of high M_r is produced. This readthrough transcription could affect transcription of genes located elsewhere in the vector, hence we have inserted a genomic fragment to provide *Dictyostelium* polyadenylation and termination sequences. In pB10T (Fig. 1), the 3′ region of the *CP1* gene, located on a 1.4-kb *Sau*3A fragment derived from the genomic clone pDG10 (Pears et al., 1985),

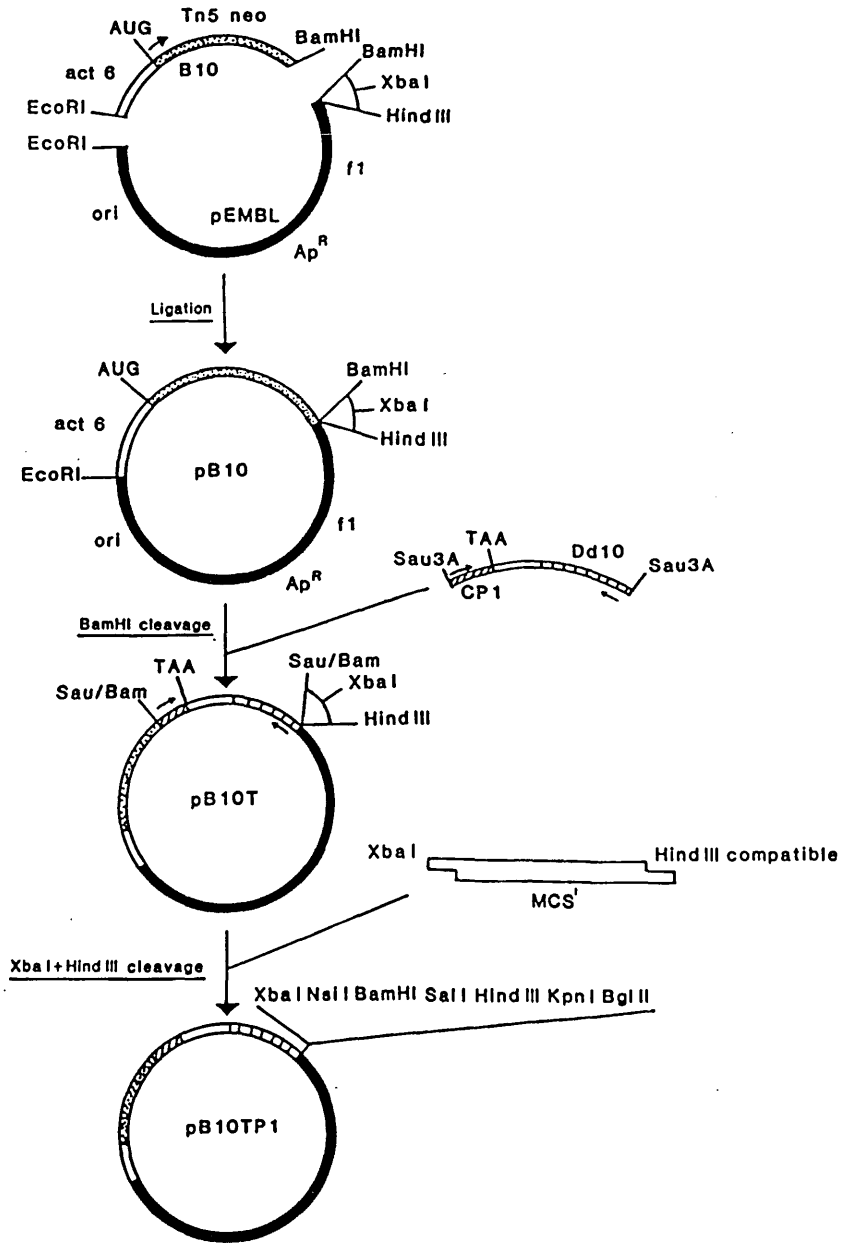


Fig. 1. Scheme of construction of pB10TP1. This scheme illustrates the construction of pB10TP1 as described in sections **a**, **b** and **c**. The identity of sequences included in the vector are denoted as follows: blackened box, pEMBL; stippled box, Tn*5 neo*; hatched box, *CP1*; striped box, *Dd10*; open box, *Dictyostelium* intergenic DNA (*act6* promoter or *CP1/Dd10* terminator). Arrows are used to indicate the direction of transcription of Tn*5 neo*, the Tn*5 neo/CP1* fusion and *pDd10*. The *Eco*RI-*Bam*HI fragment from B10 (Nellen et al., 1984) was purified on a 1% agarose gel and the *Sau*3A fragment from pDG10 (Pears et al., 1985) was purified on a 5% native acrylamide gel. An *Eco*RI + *Bam*HI-cleaved sample of pEMBL 18+ DNA (Dente et al., 1983) and a *Bam*HI-cleaved sample of pB10 DNA were phosphatased prior to ligation. The symbol *Sau/Bam* indicates the position where ligation of pB10 and the pDG10 fragment occurred, resulting in destruction of the *Bam*HI recognition sequence but retention of the *Sau*3AI site. After cleavage with *Xba*I + *Hin*dIII the pB10T vector was not phosphatased prior to insertion of MCS' (see Fig. 4); the two strands of the synthetic oligodeoxynucleotide were annealed at 60°C and then ligated without phosphorylation to ensure insertion of a single copy.

102

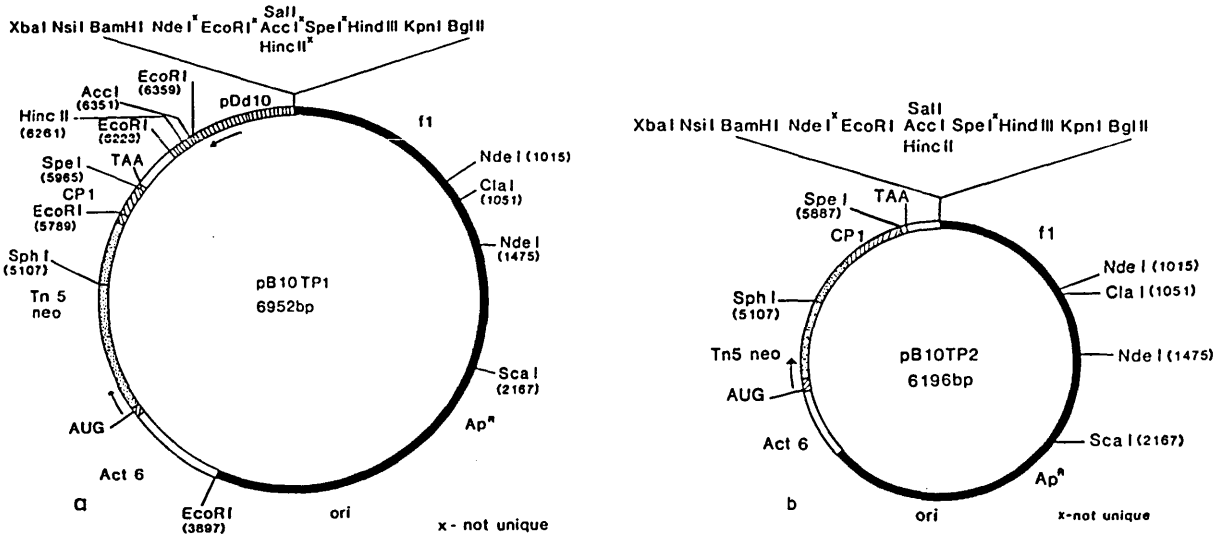


Fig. 2. Restriction maps of the pB10TP vectors. The position of the sites shown was precisely deduced using the total nucleotide sequence of the two vectors. This was in part newly determined and in part derived from other studies (Beck et al., 1982; Dente et al., 1983; Romans and Firtel, 1985). The sequence of previously uncharacterised DNA, from *Dd10*, a gene derived from pDG10 (see section b; Pears et al., 1985), and the upstream region of *act6*, was obtained using the method of Maxam and Gilbert (1980). Correct cloning of the multilinker was verified using the chain-termination method (Sanger et al., 1977). The entire nucleotide sequence will be provided if a formatted disk suitable for an IBM-PC computer is sent to the corresponding author (J.G.W.). (a) pB10TP1; (b) pB10TP2. Restriction sites marked with the x superscripts are not unique. For an explanation of other symbols used and the identity of the sequences included as denoted by shading, see Fig. 1. In addition, a short region of the *act6*-coding region is indicated by a hatched box.

has been inserted adjacent to Tn*5 neo* after cleavage at the pB10 *Bam*HI site and ligation of the complementary termini. Nucleotide sequence analysis downstream from the *CP1* gene fragment revealed the presence of a convergently transcribed gene, which we term *Dd10*. The termination codons of the two genes are separated by 271 nt. Assuming there to be no overlapping transcription, functional termination signals for both genes must be contained in this region. The data presented in Fig. 3 show that the 3′-proximal sequences from *CP1* function efficiently to produce a discrete fusion transcript. A Northern transfer of RNA extracted from pB10T transformants grown in 15 μg G418/ml was hybridized with a labelled fragment from pDd7, a cDNA clone (Williams et al., 1985). As the probe derives from the 3′ end of the CP1 coding region, it hybridises to two transcripts; the *neo*::*CP1* fusion transcript (1), expressed constitutively, and the endogenous *CP1* mRNA (2), which shows normal developmental accumulation in response to cAMP addition. The size of the fusion transcript indicates that 3′ processing occurs approximately 150 nt downstream from the *CP1* termination codon. With RNA extracted from non-transformed Ax-2 cells, the probe hybridises to the endogenous *CP1* mRNA alone.

In pB10T there are 650 nt of coding sequence derived from the *Dd10* gene. The sequence of this region has been determined but there is no obvious homologue in the currently available databases. A fragment from this region, used as a probe in Northern transfer analysis, hybridises to a constitutively expressed mRNA species of 1.1 kb in length (not shown). Fusion of the 3′ region of *Dd10* to the 5′-proximal regions of other *Dictyostelium* genes inserted in the MCS has been used, in another study in this laboratory, as a means of analysing the function of potential transcriptional control elements. A discrete fusion gene transcript was produced with the length expected for an RNA terminating in the region between *CP1* and *Dd10* (K.A. Jermyn, K. Duffy and J.G.W., unpublished results). It is preferable to fuse two coding regions in frame

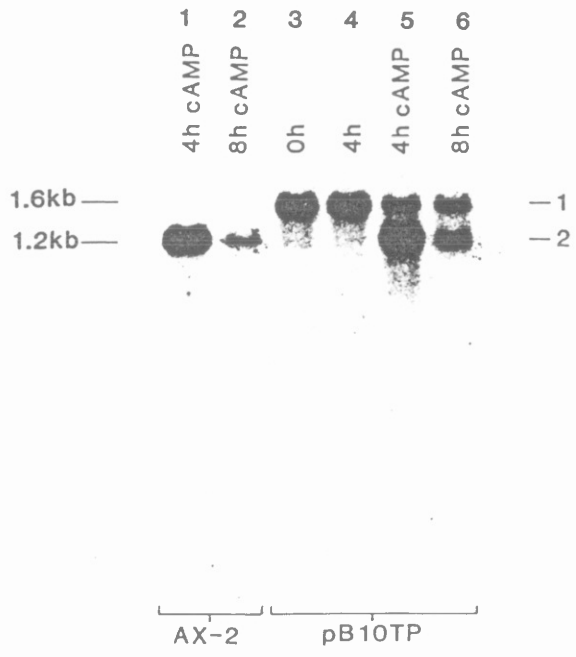


Fig. 3. The CP1 termination signal leads to the formation of a discrete neo transcript. After harvesting from the bacterial lawn, a pB10TP1 transformant of Ax-2 was grown in suspension with stepwise-increasing G418 levels until a peak of 15 μg G418/ml was reached. At this stage development of the cells was initiated by starvation. Addition of cAMP at an early stage of development induces premature expression of *CP1* (Williams et al., 1980). Cells were harvested from the transformant (pB10TP1) and the control (Ax-2) at the times shown after the initiation of development; the presence of extra exogenous cAMP is indicated. Thus: lanes 1 and 2, Ax-2 cells developed with cAMP; lanes 3 and 4, transformant developed without cAMP; lanes 5 and 6, transformant developed with cAMP. Phenol-extracted RNA from cells was denatured and separated on a 1% agarose gel containing 2.2 M formaldehyde. Both the gel and running buffer contained 0.02 M MOPS buffer pH 7.0, 0.05 M sodium acetate, and 0.01 M EDTA pH 8.0. A Northern transfer of the gel on to nitrocellulose was hybridised with a ^{32}P-labelled nick-translated (Rigby et al., 1977) *Bgl*II-*Hin*dIII fragment of 520 bp purified from pDd7 (Williams et al., 1985).

to prevent premature cessation of translation and hence an increased susceptibility of the messenger RNA to degradation. Fig. 4 shows the reading frame of the *Dd10*-coding region with respect to sites in the MCS of pB10TP1. pB10TP1 was generated by replacing the MCS present in pB10TP (see section **c**). However, the reading frame of the *Dd10*-coding region with respect to the *Xba*I site is identical in pB10TP and pB10TP1.

(c) Replacement of the MCS

To increase the usefulness of pB10T, the MCS of pEMBL18 + was replaced with a new synthetic oligomer (MCS^1) to yield pB10TP1 (Figs. 1 and 2a). This MCS^1 contains sites suitable for initial cloning, sequence analysis and mutation prior to transformation. It contains several restriction enzyme cleavage sites with a high AT content. These sites are found very frequently in the *Dictyostelium* genome and hence subcloning is facilitated. The *Nsi*I recognition sequence, which is unique in the vector, has four out of six A or T residues, and is doubly useful because it generates a *Pst*I-compatible cohesive end. The *Nde*I and *Spe*I sites, which also have a high AT content, are not unique. There are two closely spaced *Nde*I sites in the pEMBL sequence and *Spe*I is present in the CP1 coding region. Cleavage with *Spe*I deletes the termination signals and cleavage with *Nde*I eliminates the ability to produce ss DNA. However, in both cases transformation of *Dictyostelium* is still possible. The *Hin*cII site is also not unique in the MCS^1 but cleavage with this enzyme removes only a portion of *Dd10* and it is therefore a potentially useful site. Also within the MCS^1 is a unique *Bam*HI and a unique *Bgl*II site. Either of these sites could be used to clone *Sau*3AI

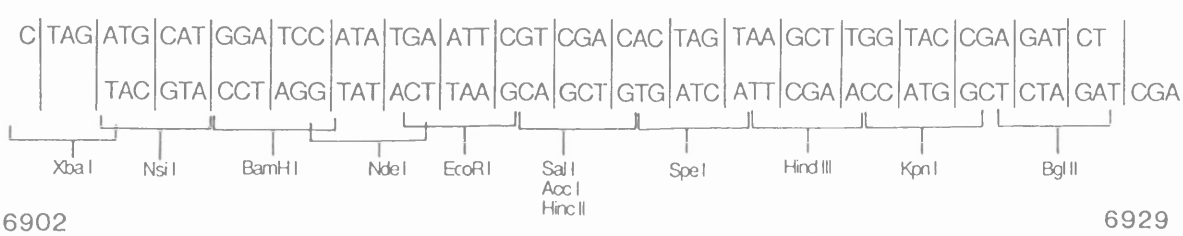


Fig. 4. Nucleotide sequence of the MCS^1 of pB10TP1 and pB10TP2. The vertical bars indicate the frame of the MCS^1 relative to the coding region of *Dd10*.

partial digests in order to generate a *Dictyostelium* genomic clone bank for use in complementing specific mutants.

As exonuclease III is unable to degrade a 3′-protruding terminus, the *Kpn*I site was included in the linker to protect pEMBL sequences from exonucleolytic digestion when creating a deletion series from a cloned fragment (Henikoff et al., 1984). The *Bgl*II site downstream from the *Kpn*I site allows the deletion derivatives to be sequenced by a choice of strategy. As the *Kpn*I site is close to the binding site for the standard M13 sequence primer, ss templates produced after superinfection with phage f1 can be readily sequenced. Alternatively, by labelling at the *Bgl*II site, nucleotide sequences within the MCS′ can be determined by the chemical cleavage method (Maxam and Gilbert, 1980). The availability of the latter strategy is of considerable importance as we have sometimes experienced difficulty in determining the sequence of very A + T-rich segments of DNA using the dideoxy procedure.

pB10TP1 is 6962 bp in length and has successfully been used to clone *Dictyostelium* DNA fragments up to 4.75 kb in length with no apparent effect on the frequency of recombinants nor on the yield of plasmid DNA from them.

(d) Construction of pB10TP2

The vector pB10TP2 was designed to maximize the number of unique restriction sites within the MCS. The *Eco*RI site in pB10 was destroyed by cleavage and end-filling to yield pB10R1. An *Eco*RI fragment of 520 nt, containing the 3′-proximal region of the *CP1* gene, was isolated from pDG10 (Pears et al., 1986) and end-filled. This fragment was inserted into the end-filled *Bam*HI site of pB10R1 to yield pB10RT. The MCS in this vector was replaced with MCS′ as in pB10TP1 to yield pB10TP2 (Fig. 2b).

The *Eco*RI fragment from pDG10 used to construct pB10TP2 contains the 3′ proximal region of *CP1* and most of the intergenic region, with its distal terminus lying 22 nt downstream from the 3′ end of the *Dd10* gene. Thus, aside from the larger number of unique restriction sites present in pB10TP2, the major difference between these two vectors is the complete absence from pB10TP2 of any sequences derived from the *Dd10* gene. However, *neo* transcripts observed using pB10TP2 are the same size as those obtained with pB10TP1 (not shown), and so both vectors must contain the sequences necessary for correct transcription termination. Due to the absence of *Dd10* sequences and the method of construction, *Eco*RI, *Acc*I and *Hin*cII sites are additionally unique in the MCS[1] (Fig. 2b). Thus pB10TP2 offers greater flexibility in subcloning, but does not provide a direct means of 'marking' cloned promoters by fusion to the *Dd10* gene.

(e) Use in transformation; modifications for Ax-2 cells

After manipulation of the cloned DNA, derivatives of the vector can be used directly in transformation to assess the expression of the gene within *Dictyostelium*. Transformation conditions for the B10 vector were established using the strain Ax-3 (Loomis, 1971; Nellen et al., 1985) which differs in several important respects from the independently isolated strain Ax-2 (Watts and Ashworth, 1970). This latter strain is routinely used in this and many other laboratories, and we have established conditions for its transformation. The calcium phosphate co-precipitation procedure remains unaltered, but the glycerol shock treatment was found to lead to a complete loss of viability. Hence the length of the glycerol shock was reduced from between 5 and 9 min to 2 min, and the concentration of glycerol was reduced from 18% to 15%. Prior to transformation cells were grown in HL5 axenic medium rather than MES-HL5 medium (Watts and Ashworth, 1970; Firtel and Lodish, 1973). After five days' preselection in medium into 20 μg G418/ml and plating on *Klebsiella aerogenes* at 40 μg G418/ml, between 200 and 2000 colonies per filter were obtained. Thus the transformation efficiency is comparable to that obtained with Ax-3 cells (Nellen et al., 1984).

After harvesting and axenic growth of the transformants in suspension with increasing concentrations of G418, Southern blotting indicated an amplification of vector copy number (Fig. 5). A peak of approx. 80 copies per cell was reached, but the sharpest increase was seen in the transition from 5 to 10 μg/ml. The pattern of bands is consistent with integration of the plasmid as a tandem linear array. Junction fragments were not observed as the cells analysed were pooled populations.

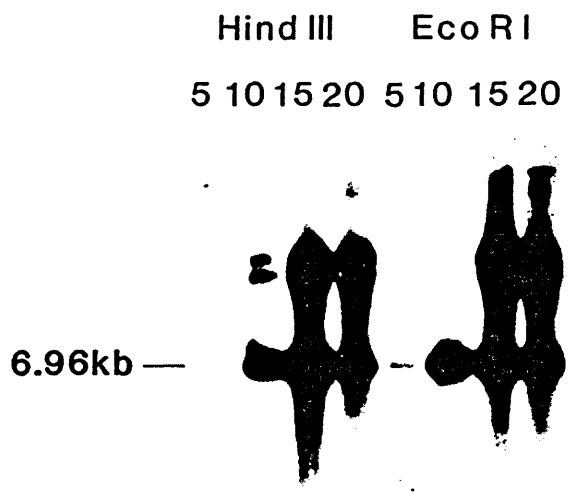


Fig. 5. The relationship between G418 concentration and vector copy number. Total cellular nucleic acid was from pB10TP1 transformants grown in successively increasing G418 concentrations (indicated for each lane in μg/ml) and 50 μg of total nucleic acid was digested with *Eco*RI or *Hin*dIII. The digests were electrophoresed through a 0.8% agarose gel containing 0.5 × TBE buffer (= 0.044 M Tris, 0.04 M borate, 0.001 M EDTA). A Southern transfer of the gel was hybridized with pB10TP1 (see section e) ^{32}P-labelled by nick translation (Rigby et al., 1977). The size of the linearized vector is indicated (6.96 kb); other bands are due to incomplete digestion of the DNA. The copy number was calculated using the assumption that 50 μg of total nucleic acid contains 1 μg of DNA. By comparison of the relative intensity of hybridisation of the probe to nucleic acid from transformed cells, to the signal obtained with a known mass of vector, the proportion of DNA from the transformant which is vector-derived was estimated. From a knowledge of the size of the *Dictyostelium* genome (50000 kb) compared to the size of the vector (6.96 kb) the copy number was then deduced.

(f) Changes in the transformation procedure

The transformation procedure developed by Nellen et al. (1984) is reproducible and yields a reasonable number of transformants per μg of DNA. However, it is impossible to estimate the true transformation efficiency because, before clonal growth on a bacterial lawn, the cells are pre-selected for five to six days in axenic medium containing G418 at 20 μg/ml. The pre-selection is necessary to kill non-transformed cells, which would otherwise prematurely clear the bacterial lawn, but cell division during pre-selection will cause the efficiency to be over-estimated. In addition to being time-consuming, plating on bacteria is the step at which contamination is most likely to occur, because four separate filters must be placed manually into the dishes to support the lawn of *Klebsiella aerogenes*. To simplify the procedure, and also to obtain a more accurate estimate of the transformation efficiency, we have developed a modified protocol which entirely dispenses with growth on bacteria.

The selection procedure is based on the observation that cells will grow in axenic medium in plastic petri dishes (R. Firtel, personal communication), and that the growth appears to be clonal (K. Duffy, personal communication). We performed the normal pre-selection step at 20 μg G418/ml but, after three days, reduced the concentration to 10 μg G418/ml. After a further three days' incubation, micro-colonies became visible to the naked eye. After a further two or three days the colonies were removed by pipetting and either transferred to a fresh plate containing medium with 10 μg G418/ml or inoculated into flasks for growth in suspension. The number of transformants could be readily determined by counting the colonies during a change of medium as they were easily visible when the plate was held obliquely. We have never observed spontaneous mutants under these selection conditions, and we routinely obtained 20 to 50 colonies when transformation is performed with pB10TP or its recombinants. After pooling the transformants and growth in suspension with 10 μg G418/ml, the vector sequences are typically estimated by Southern blot analysis to be at 20–30 copies per cell (not shown). In a set of transformants plated on to bacteria after five days pre-selection in the normal manner, obtained at the same time as an experiment giving the above result, there were 200–300 transformants. This suggests that some cell division does occur during the pre-selection step. The true transformation efficiency could be even lower than 50 per 12 μg per 2 × 10^7 cells if some of the colonies we observed resulted from cells which have attached secondarily. However, we believe this to be unlikely since there is no apparent tendency for the formation of satellite colonies as would be expected if re-attachment were occurring.

With these changes, the time elapsed between addition of DNA to the cells and establishment of suspension cultures of transformants is approximately halved. In addition the procedure is much less labour-intensive and the incidence of microbial contamination is greatly reduced.

(g) Conclusions

(1) The two transformation vectors described here make it possible to clone and manipulate *Dictyostelium* DNA more easily prior to direct use in transformation. The pB10TP1 vector provides a means of marking cloned genes while pB10TP2 allows a very wide choice of cloning strategies. In addition, the vectors produce terminated *neo* transcripts and can be readily propagated in the large amounts required for transformation.

(2) By altering the transformation protocol, it can be applied successfully to the Ax-2 strain and stable transformants growing in suspension can be obtained more quickly and easily.

ACKNOWLEDGEMENTS

We would like to thank Dr. Ellen Henderson for her kind permission to use the Northern transfer data presented in Fig. 3.

REFERENCES

Barklis, E. and Lodish, H.F.: Regulation of *Dictyostelium discoideum* mRNAs specific for prespore or prestalk cells. Cell 32 (1983) 1139–1148.

Beck, E., Ludwig, G., Auerswald, E.A., Reiss, B. and Schaller, H.: Nucleotide sequence and exact localization of the neomycin phosphotransferase gene from transposon Tn*5*. Gene 19 (1982) 327–336.

Chisolm, R.L., Barklis, E. and Lodish, H.F.: Mechanism of sequential induction of cell-type-specific mRNAs in *Dictyostelium* differentiation. Nature 310 (1984) 67–69.

Dente, L., Cesareni, G. and Cortese, R.: pEMBL: a new family of single-stranded plasmids. Nucl. Acids Res. 11 (1983) 1645–1655.

Firtel, R.A. and Lodish, H.F.: A small nuclear precursor of messenger RNA in the cellular slime mould *Dictyostelium discoideum*. J. Mol. Biol. 166 (1973) 295–314.

Henikoff, S.: Unidirectional digestion with exonuclease III creates targeted breakpoints for DNA sequencing. Gene 28 (1984) 351–359.

Loomis, W.F.: Sensitivity of *Dictyostelium discoideum* to nucleic acid analogues. Exp. Cell Res. 64 (1971) 484–486.

Loomis, W.F. (Ed.): The Development of *Dictyostelium discoideum*. Academic Press, New York, 1982.

Maxam, A. and Gilbert, W.: Sequencing end-labelled DNA with base-specific chemical cleavages. Methods Enzymol. 65 (1980) 499–560.

Mehdy, M.C. and Firtel, R.A.: A secreted factor and cAMP jointly regulate cell-type-specific gene expression in *Dictyostelium discoideum*. Mol. Cell Biol. 5 (1985) 705–713.

Mehdy, M.C., Ratner, D. and Firtel, R.A.: Induction and modulation of cell-type-specific gene expression in *Dictyostelium*. Cell 32 (1983) 763–771.

Nellen, W. and Firtel, R.A.: High-copy-number transformants and co-transformation in *Dictyostelium*. Gene 39 (1985) 155–163.

Nellen, W., Silan, C. and Firtel, R.A.: DNA-mediated transformation in *Dictyostelium discoideum*: regulated expression of an actin gene fusion. Mol. Cell Biol. 4 (1984) 2890–2898.

Oyama, M. and Blumberg, D.D.: Changes during differentiation in requirements for cAMP for expression of cell-type-specific mRNAs in the cellular slime mould, *Dictyostelium discoideum*. Develop. Biol. 117 (1986) 550–556.

Pears, C.J. and Williams, J.G.: Identification of a DNA sequence element required for efficient expression of a developmentally regulated and cAMP-inducible gene of *Dictyostelium discoideum*. EMBO J. 6, (1987) 195–200.

Pears, C.J., Mahbubani, H.M. and Williams, J.G.: Characterisation of two highly diverged but developmentally co-regulated cysteine proteinase genes in *Dictyostelium discoideum*. Nucl. Acids Res. 13 (1985) 8853–8866.

Rigby, P.W.J., Dieckmann, M., Rhodes, C. and Berg, P.: Labelling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J. Mol. Biol. 113 (1977) 237–251.

Roewekamp, W. and Firtel, R.A.: Isolation of developmentally regulated genes from *Dictyostelium*. Develop. Biol. 79 (1980) 409–418.

Romans, P. and Firtel, R.A.: Organisation of the *Dictyostelium discoideum* actin multigene family. J. Mol. Biol. 183 (1985) 311–326.

Sanger, F., Nicklen, S. and Coulson, A.R.: DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74 (1977) 5463–5467.

Watts, D.J. and Ashworth, J.M.: Growth of myxamoebae of the cellular slime mould *Dictyostelium discoideum*. Biochem. J. 119 (1970) 171–174.

Williams, J.G., Lloyd, M.M. and Devine, J.M.: Characterisation and transcription analysis of a cloned sequence derived from a major developmentally regulated mRNA of *Dictyostelium discoideum*. Cell 17 (1979) 903–913.

Williams, J.G., Tsang, A.S. and Mahbubani, H.: A change in the rate of transcription of a eukaryotic gene in response to cAMP. Proc. Natl. Acad. Sci. USA 77 (1980) 7171–7175.

Williams, J.G., North, M.J. and Mahbubani, H.: A developmentally regulated cysteine proteinase in *Dictyostelium discoideum*. EMBO J. 4 (1985) 999–1006.

Zoller, M.J. and Smith, M.: Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any fragment of DNA. Nucl. Acids Res., 10 (1982) 6487–6500.

Communicated by K.F. Chater.

Development 103, 519-524 (1988)
Printed in Great Britain © The Company of Biologists Limited 1988

A *Dictyostelium* prespore-specific gene is transcriptionally repressed by DIF *in vitro*

A. E. EARLY and J. G. WILLIAMS*

Imperial Cancer Research Fund, Clare Hall Laboratories, Blanche Lane, South Mimms, Herts, EN6 3LD, UK

* To whom reprint requests should be addressed

Summary

One important role of DIF, the stalk cell-specific inducer of *Dictyostelium*, may be to divert cells from the spore cell pathway of differentiation. The D19 gene encodes an mRNA which is highly enriched in prespore over prestalk cells in the migratory slug. We show, using a mutant defective in DIF accumulation, that the concentration of D19, and several other prespore mRNA sequences, decreases in the presence of exogenous DIF. There is evidence that both transcriptional and post-transcriptional controls operate to regulate expression of these genes. We have performed *in vitro* nuclear transcription and mRNA half-life analyses, and find that DIF acts at the transcriptional level to repress the accumulation of the D19 mRNA.

Key words: *Dictyostelium discoideum*, DIF, prespore gene, transcription.

Introduction

Cells in the anterior one fifth of the *Dictyostelium* slug normally differentiate to form dead, vacuolated stalk cells while the majority of cells in the posterior region differentiates into spores. Cyclic AMP is the chemoattractant responsible for cellular aggregation and cAMP signalling is thought to persist during slug migration (reviewed by Schaap, 1986). In disaggregated cells, shaken rapidly in suspension, the local extracellular cAMP concentration is reduced and several prespore-specific genes have been shown to have a reduced transcription rate and reduced mRNA stability (Landfear *et al.* 1982; Mangiarotti *et al.* 1983). These effects can be overcome by the addition of cAMP (Mangiarotti *et al.* 1983). Thus, prespore gene expression is regulated by cAMP at both a transcriptional and a post-transcriptional level.

Two substances, adenosine and DIF, have been shown to be capable of repressing prespore cell differentiation. In aggregation-competent cells, shaken in suspension, addition of adenosine inhibits cAMP-induced prespore gene expression (Weijer & Durston, 1985; Wang *et al.* 1986). In intact slugs, reduction of the adenosine concentration by enzymic depletion results in the appearance of prespore cells in the prestalk zone (Schaap & Wang, 1986). DIF is a low molecular weight chlorinated compound of known structure (Morris *et al.* 1987), which will induce isolated amoebae to differentiate into stalk cells (Town *et al.* 1976). If slugs are disaggregated and incubated in the presence of DIF, then prespore gene expression is repressed (Kay & Jermyn, 1983). The effect of DIF on the expression of individual genes can best be studied *in vitro* using the HM44 mutant which is blocked in the accumulation of DIF but remains DIF-responsive (Kopachik *et al.* 1983). If cells of the HM44 strain are incubated with cAMP for approximately 10 h, and DIF is then added, at least 90 % of the amoebae form stalk cells. In the absence of DIF only 1 % of the cells differentiates into stalk cells. After incubation with cAMP the cells contain prespore-specific proteins, distinguishable by 2-D gel electrophoresis and the addition of DIF suppresses their expression (Kopachik *et al.* 1985).

The D19 cDNA clone derives from an mRNA sequence that is highly enriched in the prespore region of the slug (Barklis & Lodish, 1983). We have established the sequence of the D19 gene and shown it to encode PsA, a cell surface glycoprotein of prespore cells (Grant *et al.* 1985; Early, Williams, Meyer, Por, Smith, Williams & Gooley, manuscript submitted). PsA is known to be specifically recognized by a monoclonal antibody, MUD-1 (Krefft *et*

al. 1983). Using this gene, we have investigated the mechanism whereby DIF acts to repress specific prespore gene expression.

Materials and methods

Dictyostelium discoideum *growth and development*

The mutant strain HM44 (Kopachik *et al.* 1983) was grown in association with *Klebsiella aerogenes* on SM agar plates (Sussman, 1966). Stalk cell induction was performed in submerged monolayer culture as described by Kopachik *et al.* (1985). Cells were incubated in the presence of 5 mM-cAMP for 10 h and the buffer was replaced with fresh buffer additionally containing 3000 units DIF-1 ml^{-1} (Morris *et al.* 1987). Plates containing nogalamycin, which was added to 300 $\mu g\,ml^{-1}$ 5 min prior to the addition of DIF, were placed in a light-proof box to prevent inactivation of the drug. Nogalamycin was the generous gift of the Upjohn Co., Kalamazoo, MI, USA.

RNA isolation and Northern transfer

One plate of cells, containing approximately 4×10^7 amoebae, was used to prepare RNA for each timepoint of the stalk cell induction. The cells were pelleted and frozen in dry ice. The frozen pellets were directly resuspended into 1 ml of buffer containing 50 mM-Hepes, 40 mM-MgOAc, 20 mM-KCl, 0·5 % SDS, 200 mM-NaCl and 0·5 % diethylpyrocarbonate. Total nucleic acid was prepared by repeated phenol–chloroform extraction. Aliquots (5 μg) of RNA were electrophoresed through a 1 % denaturing formaldehyde gel and transferred to nitrocellulose filters (Maniatis *et al.* 1982). The filters were prehybridized for 4 h at 37°C in 50 % formamide, 5×SSC, 10×Denhardt's solution, 100 $\mu g\,ml^{-1}$ heat-denatured salmon sperm DNA, 0·1 % SDS and hybridized in the same conditions overnight. The filters were probed with cDNA clones labelled with ^{32}P by the method of Feinberg & Vogelstein (1983). The filters were washed four times for 15 min each in 2×SSC, 0·1 % SDS at 65°C before autoradiography.

Isolation of nuclei and in vitro *nuclear transcription*

Synthetically active nuclei were prepared and stored as described by Jacobson (1976) and used for *in vitro* nuclear RNA synthesis as described previously (Williams *et al.* 1987). The labelled RNA was hybridized to Southern transfers bearing restriction digests of the D19 cDNA clone and of a constitutively transcribed sequence, 1G7. The filters were prehybridized at 37°C in a buffer containing 50 % deionized formamide, 5×SSC, 200 $\mu g\,ml^{-1}$ BSA, 0·2 % SDS, 300 $\mu g\,ml^{-1}$ Ficoll, 500 $\mu g\,ml^{-1}$ poly rA and 100 $\mu g\,ml^{-1}$ wheat germ tRNA. After a 4 h prehybridization the ^{32}P-labelled transcripts were added to this buffer and hybridized for two days at 37°C. The filters were washed four times for 15 min each in 50 % formamide, 1×SSC, 0·1 % SDS at 37°C.

Results

(A) *The addition of DIF to HM44 cells decreases the concentration of prespore-specific mRNA sequences*

In a standard stalk cell induction, HM44 cells are incubated in tissue-culture plates with cAMP alone for 10 h and then with cAMP and DIF for an additional 12 h (Kopachik *et al.* 1985). After 10 h in the presence of cAMP, prespore-specific proteins are present and addition of DIF leads to their disappearance. In order to examine this process at an RNA level, the D19 cDNA clone was used to probe a Northern blot of RNA extracted from HM44 cells harvested during a stalk cell induction (Fig. 1). The amount of DIF used was sufficient to transform more than 85 % of amoebae into morphologically recognizable stalk cells after a 12 h incubation period. Total RNA was extracted from cells harvested just before the medium change and at intervals afterwards. After the 10 h cAMP preincubation, the cells contain D19 mRNA and, in the absence of DIF, the level of D19 mRNA remains constant. When DIF is present, the level of D19 mRNA declines such that, approximately 4 h after DIF addition, it is at about 10 % of its former level. This low level is maintained throughout the remaining incubation. The residual D19 mRNA may derive from those amoebae that never mature into stalk cells. It takes about 90 min for the amount of D19 mRNA to be halved. Two other prespore mRNA sequences, those corresponding to the cDNA clones PL3 and D7 (Barklis & Lodish, 1983; Oyama & Blumberg, 1986), were found to be repressed by DIF with identical kinetics (Fig. 2).

(B) *DIF acts to repress D19 gene transcription*

Nuclear run-on assays were used to determine if the repression of D19 by DIF is mediated at the level of gene transcription (Fig. 3). Transcriptionally active nuclei were isolated from HM44 cells during a stalk cell induction. Again the cells were incubated in

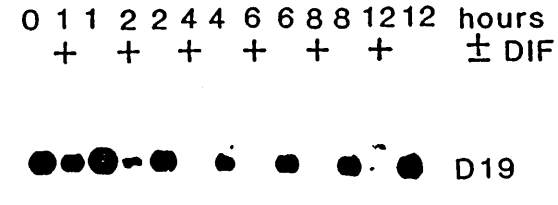


Fig. 1. The kinetics of repression of the D19 mRNA by DIF. The symbol 0 indicates the time at which DIF was added to HM44 cells, after 10 h preincubation with cAMP alone. The time of incubation in the presence or absence of DIF is shown in hours. Total nucleic acid was purified and analysed by Northern transfer using the D19 cDNA clone as a probe, as described in the methods section.

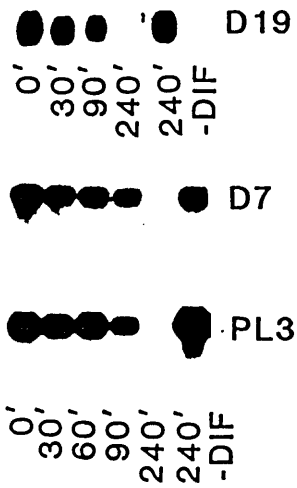


Fig. 2. Comparison of repression of the D19 with the repression of two further prespore genes. HM44 cells were incubated for the times shown in the presence or absence of DIF, after 10 h incubation in medium containing cAMP alone. Total nucleic acid was extracted and analysed by Northern transfer as before. A mixture of the D7 and PL3 cDNA clones was used to probe the lower filter. The Northern transfer probed with the D19 cDNA clone derives from a separate experiment, but in both cases greater than 90 % of amoebae formed stalk cells after 12 h incubation in the presence of DIF.

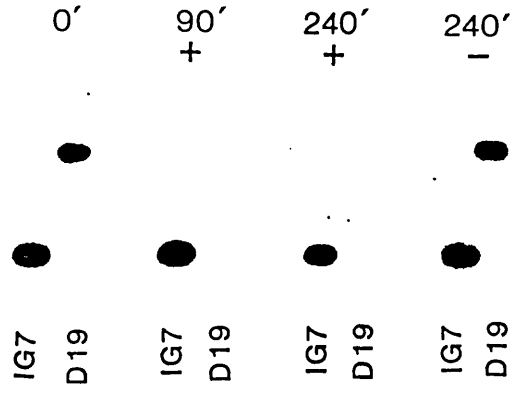


Fig. 3. Determination of the rate of D19 gene transcription during stalk cell induction. Transcriptionally active nuclei were prepared from HM44 cells incubated in the presence or absence of DIF for the times shown, following 10 h preincubation with cAMP. ^{32}P-labelled RNA from the nuclear run-on assays was used to probe Southern blots of the 1G7 and D19 cDNA clones. The restriction enzyme digestion separates the insert from the plasmid vector band, which then acts as a control for background hybridization. 1G7, a cDNA cloned into the *Sma*I site of pUC8, was digested with *Eco*RI and *Hin*dIII, giving a 500 bp insert fragment. D19, a cDNA cloned into the *Bam*HI site of pBR322, was digested with *Hin*dIII and *Sal*I, giving a 1·3 kb insert fragment. In both cases, 2 μg of DNA was used per track.

tissue culture plates for 10 h before DIF addition. Nuclei were harvested just prior to the medium change and at 90 min and 240 min after the addition of DIF. Nuclei were also isolated from cells incubated for 240 min in medium with no DIF. The times were chosen to span the period when the greatest reduction in mRNA levels is observed. The cytoplasmic fractions obtained during the nuclei preparation were used to monitor the stalk cell induction. Total nucleic acid was extracted and a Northern blot was probed with a mixture of the D19 and the pDd63 (Jermyn *et al.* 1987) cDNA clones. pDd63 hybridizes to an mRNA that is rapidly induced at the transcriptional level by DIF (Williams *et al.* 1987).

The nuclei were incubated under conditions in which preinitiated transcripts are elongated and the radiolabelled RNA was hybridized to Southern blots of two plasmid DNAs, D19 and 1G7 (Fig. 3). The 1G7 cDNA clone derives from an mRNA which is expressed constitutively during normal development and which is equally abundant in prestalk and prespore cells (K. A. Jermyn, personal communication). The relative intensities of the signals obtained for D19 were normalized by using exposure times giving approximately equal signal intensities for 1G7 (Fig. 3).

Within 90 min of DIF addition, the signal obtained for D19 decreases approximately tenfold. After 4 h incubation with DIF there is a further decrease in D19 transcription such that only a faint signal is obtained after a very long exposure of the autoradiogram. In contrast, where DIF was not included in the medium, the signal intensity remains constant over the 4 h period. Thus DIF almost completely abolishes transcription of the D19 gene. Comparing the transcription rate with the rate of decline in cytoplasmic mRNA levels in the same experiment (Fig. 2), we conclude that a major part of the repression of D19 observed upon DIF addition can be explained by a reduction in nuclear transcription.

(C) *The stability of D19 mRNA is unaffected by DIF addition*

The average half-life of individual mRNAs in *Dictyostelium* developing aggregates of AX-3 cells has been estimated to be between 3 and 4 h (Mangiarotti *et al.* 1982). If the D19 mRNA turned over in HM44 cells with this half-life, a 50 % reduction in mRNA levels 90 min after DIF addition would not be possible, even if transcription ceased immediately. Furthermore, the D19, PL3 and D7 genes are known to be regulated at both a transcriptional and a post-transcriptional level by cAMP (Landfear *et al.* 1982; Barklis & Lodish, 1983). Hence, experiments were performed to determine whether or not the repres-

sion by DIF also has a post-transcriptional component.

These experiments utilized the RNA synthesis inhibitor nogalamycin, which has been successfully used to determine mRNA half-lives in germinating *Dictyostelium* spores (Ennis, 1981; Kelly *et al.* 1985). A stalk cell induction was again performed using HM44 cells. After a 10 h incubation with cAMP alone the medium was changed and supplemented with nogalamycin at 300 $\mu g\,ml^{-1}$. 5 min later, DIF was added to half of the plates and cells were harvested at various times after nogalamycin addition, up to a total of 4 h. Cells were also harvested just prior to drug addition and a terminal timepoint from cells incubated with DIF, but without nogalamycin was included to check the induction. A Northern blot of total RNA extracted from the cells was probed with a mixture of the pdD63 and the D19 cDNA clones (Fig. 4). The low level of pDd63 mRNA synthesized in the presence of nogalamycin, compared to that induced in its absence, demonstrates the effectiveness of the drug. At 4 h, with DIF alone, D19 mRNA levels are much reduced compared to zero times. In the presence of the drug, the rate of decay of the D19 mRNA is not altered by the addition of DIF. The half-life of the mRNA is about 90 min in both the presence and absence of DIF. This value, taken in conjunction with the result of the nuclear run-on assays, is consistent with the effect of DIF being mediated purely at the transcriptional level.

In an experiment where DIF was added 30 min prior to nogalamycin, the subsequent rate of decay of the mRNA was not significantly different from that observed with the protocol described above (data not shown). Hence, it is unlikely that nogalamycin is interfering with DIF action by preventing the synthesis of a DIF-inducible component necessary for destabilization.

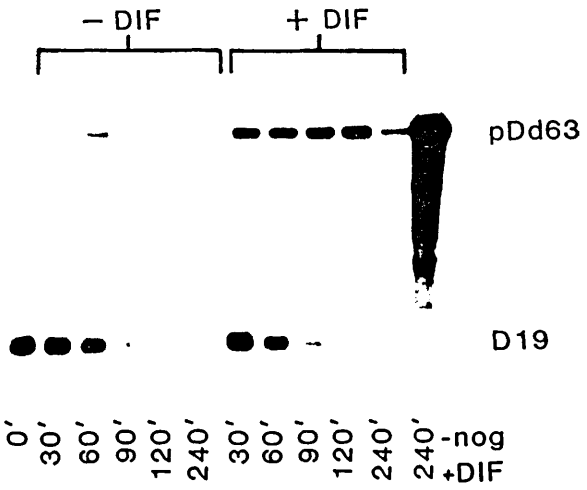


Fig. 4. Determination of the half-life of the D19 mRNA in the presence and absence of DIF. HM44 cells were incubated in medium containing nogalamycin after a 10 h incubation in the presence of cAMP. DIF was added to half the plates after 5 min in the presence of the drug. A timepoint harvested 4 h after DIF addition, in the absence of the drug, was also included. Total nucleic acid was extracted and analysed by Northern transfer with a mixture of the pDd63 and D19 cDNA clones.

Discussion

Previous observations have shown that DIF acts to repress prespore cell differentiation (Kay & Jermyn, 1983; Kopachik *et al.* 1985). We have now shown that DIF represses the expression of the D19, PL3 and D7 prespore genes in HM44 cells. The mRNA species that hybridize to these cDNAs have all been assigned to the prespore II class on the basis of their time of appearance and sensitivity to extracellular cAMP (Chisolm *et al.* 1984). All three mRNA sequences disappear from HM44 cells on exposure to DIF with similar kinetics. It seems likely, therefore, that they are all repressed by DIF through a similar mechanism.

The rate of transcription of the D19 gene, determined from nuclear run-on assays, is reduced in the presence of DIF while the half-life of the mRNA is unaffected. Given the relatively short, 90 min, half-life of the mRNA, the reduction in transcription rate is sufficient for transcriptional regulation primarily to account for the observed rate of mRNA disappearance. The induction by DIF of two prestalk-enriched mRNA sequences, pDd63 and pDd56, has also been shown to be mediated at a transcriptional level (Williams *et al.* 1987; Ceccarelli *et al.* 1987). The response to DIF of the pDd63 gene is very rapid as new transcripts are detected within 15 min of the addition of DIF. We cannot accurately judge the speed with which DIF affects D19 transcription because it is intrinsically difficult to determine the precise onset of repression when an mRNA has any significant degree of stability. We have also used the MUD-1 antibody (Krefft *et al.* 1983) to examine the effect of DIF on the level of D19 protein in HM44 cells (data not shown). If DIF is added after 10 h incubation with cAMP alone, there is no further net accumulation of protein over the next 14 h, but the pre-existing low level of protein does not decline. This implies that DIF does not cause a rapid turnover of prespore protein, and hence its regulation of gene transcription is of primary importance in mediating prespore repression.

DIF has been proposed to act by interfering with the cAMP-mediated induction of prespore gene expression (Wang *et al.* 1986). In aggregation-competent NC4 cells, incubated in monolayer culture,

DIF and cAMP act antagonistically. DIF represses cAMP–receptor interactions at low cAMP concentrations and reduces the cAMP relay response (Wang *et al.* 1986). However, our experiments have been performed using a saturating, 5 mM, concentration of cAMP, where DIF does not prevent cAMP binding to the receptor. Possibly, therefore, DIF is counteracting cAMP induction of D19 expression at some intracellular stage of signal transduction. Such a mechanism would provide greater scope for specificity of repression and, consistent with this, not all genes induced by cAMP are repressed by DIF in HM44 cells (Jermyn *et al.* 1987). It would also explain the absence of mRNA destabilization, since this is apparently dependent upon the disruption of cAMP signalling (Mangiarotti *et al.* 1982, 1983). Adenosine is thought to block cAMP-induced prespore differentiation directly, by inhibition of cAMP binding (Wang *et al.* 1986). Unlike DIF, therefore, adenosine might be expected to exert both a transcriptional and post-transcriptional effect on prespore gene expression.

When sections of developing aggregates are stained with MUD-1, a region of nonstaining cells is observed at the top of the mound (Krefft *et al.* 1984). Possibly, therefore, factors such as DIF and/or adenosine may be repressing prespore differentiation in this region. Adenosine seems a likely candidate for such a role since the enzymes that hydrolyse cAMP to generate it are enriched in the prestalk region (Armant *et al.* 1980; Brown & Rutherford, 1980). Also, enzymic reduction of adenosine levels in the slug leads to prespore differentiation in the prestalk region (Schaap & Wang, 1986). However, exposure to high adenosine levels does not lead to an increase in the total proportion of prestalk tissue in the slug, while the analogous experiment using DIF results in a major reduction in the proportion of prespore staining cells (D. Traynor and R. Kay, personal communication). This does not necessarily indicate a more important role for DIF but may simply reflect differences in the levels and spatial distribution of the exogenously added DIF and adenosine. As DIF and adenosine act at different stages in the cAMP signal transduction pathway, one attractive idea is that they might function synergistically to divert cells from the prespore pathway of differentiation. Further analysis of the precise mechanism of repression should help to elucidate their relative contribution to this process.

References

Armant, D. R., Stetler, D. A. & Rutherford, C. L. (1980). Cell surface localisation of 5′ AMP nucleotidase in prestalk cells of *Dictyostelium discoideum*. *J. Cell Sci.* **45**, 119–129.

Barklis, E. & Lodish, H. F. (1983). Regulation of *Dictyostelium discoideum* mRNAs specific for prespore or prestalk cells. *Cell* **32**, 1139–1148.

Brown, S. S. & Rutherford, C. L. (1980). Localisation of cyclic nucleotide phosphodiesterase in the multicellular stage of *Dictyostelium discoideum*. *Differentiation* **16**, 173–184.

Ceccarelli, A., McRobbie, S. J., Jermyn, K. A., Duffy, K., Early, A. & Williams, J. G. (1987). Structural and functional characterisation of a *Dictyostelium* gene encoding a DIF-inducible, prestalk-enriched mRNA sequence. *Nucl. Acids Res.* **15**, 7463–7476.

Chisolm, R. L., Barklis, E. & Lodish, H. F. (1984). Mechanisms of sequential induction of cell-type-specific mRNAs in *Dictyostelium* differentiation. *Nature, Lond.* **310**, 67–69.

Ennis, H. L. (1981). Nogalamycin inhibits ribonucleic acid synthesis in growing and developing cells of the slime mould *Dictyostelium discoideum*. *Antimicrob. Agents Chemother.* **19**, 657–665.

Feinberg, A. P. & Vogelstein, B. (1983). A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. *Analytical Biochem.* **132**, 6–13.

Grant, W. N., Welker, D. L. & Williams, K. L. (1985). A polymorphic, prespore-specific cell-surface glycoprotein is present in the extracellular matrix of *Dictyostelium discoideum*. *Mol. Cell Biol.* **5**, 2559–2566.

Jacobson, A. (1976). Analysis of mRNA transcription in *Dictyostelium discoideum*. In *Methods in Molecular Biology*, vol. 8 (ed. J. A. Last), pp. 161–209. New York: Marcel Dekker, Inc.

Jermyn, K. A., Berks, M., Kay, R. R. & Williams, J. G. (1987). Two distinct classes of prestalk-enriched mRNA sequences in *Dictyostelium discoideum*. *Development* **100**, 745–755.

Kay, R. R. & Jermyn, K. A. (1983). A possible morphogen controlling differentiation in *Dictyostelium*. *Nature, Lond.* **303**, 242–244.

Kelly, R., Kelly, L. J. & Ennis, H. L. (1985). *Dictyostelium discoideum* mRNAs developmentally regulated during spore germination have short half-lives. *Mol. Cell Biol.* **5**, 133–139.

Kopachik, W., Dhokia, B. & Kay, R. R. (1985). Selective induction of stalk cell-specific proteins in *Dictyostelium*. *Differentiation* **28**, 209–216.

Kopachik, W., Oohata, A., Dhokia, B., Brookman, J. J. & Kay, R. R. (1983). *Dictyostelium* mutants lacking DIF, a putative morphogen. *Cell* **33**, 397–403.

Krefft, M., Voet, L., Gregg, J. H., Mairhofer, H. & Williams, K. L. (1984). Evidence that positional information is used to establish the prestalk-prespore pattern in *Dictyostelium discoideum* aggregates. *EMBO J.* **3**, 201–206.

Krefft, M., Voet, L., Mairhofer, H. & Williams, K. L. (1983). Analysis of proportion regulation in slugs of *Dictyostelium discoideum* using a monoclonal antibody and FACS-IV. *Expl Cell Res.* **147**, 253–256.

Landfear, S. M., Lefebvre, P., Chung, S. & Lodish, H. F. (1982). Transcriptional control of gene expression during development of *Dictyostelium discoideum*. *Mol.*

Cell Biol. **2**, 1417–1426.

MANIATIS, T., FRITSCH, E. F. & SAMBROOK, J. (1982). *Molecular Cloning: A Laboratory Manual. Cold Spring Harbor.* New York: Cold Spring Harbor Laboratory.

MANGIAROTTI, G., CECCARELLI, A. & LODISH, H. F. (1983). Cyclic AMP stabilizes a class of developmentally regulated *Dictyostelium discoideum* mRNAs. *Nature, Lond.* **301**, 616–618.

MANGIAROTTI, G., LEFEBVRE, P. & LODISH, H. F. (1982). Differences in the stability of developmentally regulated mRNAs in aggregated and disaggregated *Dictyostelium discoideum* cells. *Devl Biol.* **89**, 82–91.

MORRIS, H. R., TAYLOR, G. W., MASENTO, M. S., JERMYN, K. A. & KAY, R. R. (1987). Chemical structure of the morphogen differentiation inducing factor from *Dictyostelium discoideum. Nature, Lond.* **328**, 811–814.

OYAMA, M. & BLUMBERG, D. D. (1986). Cyclic AMP and NH_3/NH_4^+ both regulate cell-type-specific mRNA accumulation in the cellular slime mould, *Dictyostelium discoideum. Devl Biol.* **117**, 557–566.

SCHAAP, P. (1986). Regulation of size and pattern in the cellular slime moulds. *Differentiation* **33**, 1–16.

SCHAAP, P. & WANG, M. (1986). Interactions between adenosine and oscillatory cAMP signalling regulate size and pattern in *Dictyostelium. Cell* **45**, 137–144.

SUSSMAN, M. (1966). Biochemical and genetic methods in the study of cellular slime mould development. *Meth. cell. Physiol.* **12**, 397–410.

TOWN, C. D., GROSS, J. D. & KAY, R. R. (1976). Cell differentiation without morphogenesis in *Dictyostelium discoideum. Nature, Lond.* **262**, 717–719.

WANG, M., SCHAAP, P. & VAN HAASTERT, P. J. M. (1986). Multiple effects of differentiation-inducing factor and cyclic-AMP signal transduction in *Dictyostelium. Differentiation* **33**, 24–28.

WEIJER, C. J. & DURSTON, A. J. (1985). Influence of cyclic AMP and hydrolysis products on cell type regulation in *Dictyostelium discoideum. J. Embryol. exp. Morph.* **86**, 19–37.

WILLIAMS, J. G., CECCARELLI, A., MCROBBIE, S. J., MAHBUBANI, H., KAY, R. R., EARLY, A., BERKS, M. & JERMYN, K. A. (1987). Direct induction of *Dictyostelium* prestalk gene expression by DIF provides evidence that DIF is a morphogen. *Cell* **49**, 185–192.

(*Accepted 22 March 1988*)

MOLECULAR AND CELLULAR BIOLOGY, Aug. 1988, p. 3458–3466
0270-7306/88/083458-09$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 8, No. 8

Structural Characterization of *Dictyostelium discoideum* Prespore-Specific Gene D19 and of Its Product, Cell Surface Glycoprotein PsA

A. E. EARLY,[1] J. G. WILLIAMS,[1]* H. E. MEYER,[2] S. B. POR,[3]† E. SMITH,[3] K. L. WILLIAMS,[3] AND A. A. GOOLEY[3]

Imperial Cancer Research Fund, Clare Hall Laboratories, Blanche Lane, South Mimms, Potters Bar, Hertfordshire, EN6 3LD, United Kingdom[1]; Medizinische Fakultat, Institut fur Physiologische Chemie, Ruhr-Universitat Bochum, D-4630 Bochum, Federal Republic of Germany[2]; and School of Biological Sciences, Macquarie University, Sydney, New South Wales 2109, Australia[3]

Received 8 February 1988/Accepted 20 May 1988

The *Dictyostelium discoideum* cell surface antigen PsA is a glycoprotein which first appears in the multicellular stage soon after tip formation and is selectively expressed on prespore cells. The D19 gene encodes an mRNA sequence which is highly enriched in prespore over prestalk cells in the slug stage. We have determined 81 amino acid residues of N-terminal sequence from immunoaffinity-purified PsA protein and shown this sequence to be identical to the predicted sequence of the D19 gene. There are several short repeat elements close to the C terminus, and unequal crossing-over within these is proposed to account for the size polymorphism observed in PsA protein isolated from different *D. discoideum* strains. The repeats are proline rich and show similarity to the C-terminal region of the *D. discoideum* cell adhesion molecule, contact sites A. The extreme C terminus, which is also homologous to contact sites A, is characteristic of proteins attached to the plasma membrane via a glycosyl-phosphatidylinositol link. We have marked the PsA gene by insertion of an oligonucleotide encoding an epitope of the human c-*myc* protein. A construct containing this gene and 990 base pairs of 5′-flanking region directed correct temporal and spatial mRNA accumulation. We found the marked PsA protein, detected with the human c-*myc* antibody, to be correctly localized on the surface of cells.

The identification of cell-type-specific markers is of central importance in the analysis of cellular differentiation and pattern formation. The ideal marker should be exclusively restricted to the cell lineage which it characterizes, and, if it is to be used to analyze the initial events in cellular differentiation, it should be expressed early in the ontogeny of the cell lineage. The anterior one-fifth of the migratory slug of *Dictyostelium discoideum* contains stalk cell precursors, while the posterior four-fifths consists predominantly of prespore cells. Genes, which are expressed in a cell-type-specific manner during development, are readily isolated (1, 17, 27), because prestalk- and prespore-cell-enriched populations can be separated by microdissection or by density gradient centrifugation of disaggregated slugs (33, 42). The D19 cDNA clone derives from a highly prespore-enriched mRNA obtained from density gradient-separated cells (1). The D19 gene has been defined as a class 2 prespore sequence on the basis of the time of appearance of its mRNA during development and its dependence on extracellular cyclic AMP for continued expression (7). It has frequently been used as a prespore-specific marker in studies investigating the requirements for cell-type-specific gene expression and cellular differentiation (17, 32, 36).

An alternative approach to generating cell-type-specific markers has been to raise monoclonal antibodies (MAbs). The MAb MUD1 detects the surface glycoprotein PsA (prespore-specific antigen) first appearing on the prespore cell surface early in the prespore pathway, soon after tip formation (19). The MUD1 MAb has often been used to specifically localize prespore cells and to quantitate prespore gene expression (18, 37). PsA was originally defined as a 30-kilodalton (kDa) prespore-specific cell surface protein carrying a unique epitope which is recognized by MUD1 (19). By using genetic, immunological, and biochemical criteria, it was demonstrated that PsA is encoded by a single locus, *pspA*, and that the gene product is a protein with a molecular weight polymorphism in several wild isolates (12). Another epitope on PsA is recognized by the MAb MUD50 (13), which, in addition to PsA, recognizes a diverse group of prespore proteins (S. Alexander, E. Smith, L. Davis, A. Gooley, S. B. Por, L. Browne, and K. L. Williams, Differentiation, in press). West and Loomis (44) have shown that the PsA (SP29) glycoprotein lacks a glycoconjugate(s) in *modB* mutants, and it has been shown that the MUD50 epitope is, or is part of, this carbohydrate (Alexander et al., in press).

Another desirable feature of a marker is that it be well defined structurally and functionally. PsA is the best characterized cell-type-specific surface molecule of the multicellular stage. Here, we have determined the entire sequence of the D19 gene and have shown that it encodes the PsA protein, which was purified and partially sequenced.

MATERIALS AND METHODS

Strains and culture conditions. PsA was purified from the slug stage of three strains of *D. discoideum*: two wild isolates, WS576 and WS380B (9); and a mutant, HU2440, which carries the *modB501* mutation (44) in the genetic background of wild isolate DD61 (12; A. A. Gooley and K. L. Williams, manuscript in preparation).

Preparation of slug extracts and purification of PsA by immunoaffinity chromatography. Two methods were used to

* Corresponding author.
† Present address: St. Vincent's Hospital, Darlinghurst, Sydney, New South Wales, Australia.

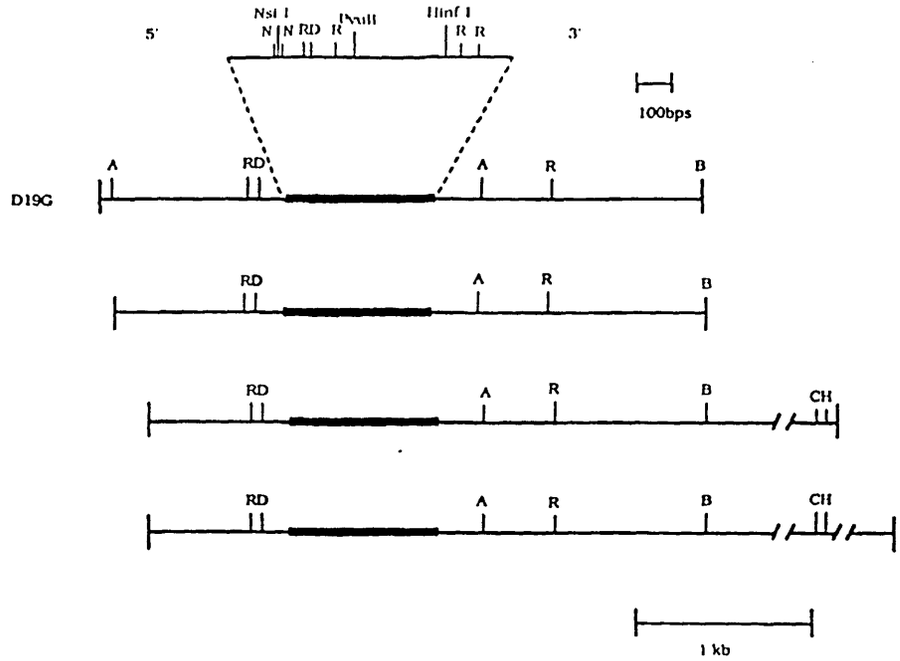


FIG. 1. Structure of the D19 genomic clones. The transcribed region ( ) contains the translated region ( ) and the 5' and 3' untranslated regions ( ). This part of the map has been expanded twofold relative to the intergenic DNA. Abbreviations for restriction sites are as follows: A, *Ava*II; B, *Bam*HI; C, *Cla*I; D, *Dde*I; H, *Hin*dIII; N, *Nde*I; R, *Rsa*I. D19G and three other genomic clones are shown.

prepare detergent extracts of slug cells. First, ~5 × 10^{10} slug cells were lysed in phosphate-buffered saline–0.1% (vol/vol) Nonidet P-40 (NP-40; Sigma Chemical Co.). The detergent extracts were passaged over two guard columns (20 ml of Sepharose 6B followed by 4 ml of CNBr-Sepharose 4B blocked with glycine) before chromatography on a MUD1 affinity column. Increased yields of PsA were obtained with a combination of detergents. Crude membranes were prepared from ~5 × 10^{10} slug cells by the Tween 40 (Sigma) detergent method used to purify Thy-1 from rat thymocytes (35). The membranes were then dissolved in 2% (wt/vol) sodium deoxycholate (BDH), and before binding the extract to the MUD1 affinity column, it was found necessary to add NP-40 to a final concentration of 1% (vol/vol) to restore antibody-antigen interaction. PsA was bound to a column of 1 ml of CNBr-Sepharose (Pharmacia) coupled to the MAb MUD1. Bound PsA was eluted from the column in 1-ml samples of 50 mM diethylamine (Sigma)–NaOH (pH 11.5)–0.1% (vol/vol) NP-40 and immediately adjusted to pH 8.0 with solid glycine. Fractions containing PsA, as detected with MUD1 in a modified dot-immunobinding assay (39), were pooled and lyophilized.

Purified PsA was subjected to sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) by the method of Laemmli (22) with 12.5% (wt/vol) acrylamide in the resolving gel and electroblotted onto nitrocellulose by the method of either Towbin et al. (41) or Kyhse-Anderson (20).

Endoproteinase Glu-C (*Staphylococcus aureus* V8 protease) digestion. Protease digests of reduced and carboxymethylated PsA and peptide separation on high-pressure liquid chromatography were essentially as described by Meyer et al. (28). Cleavage with V8 protease (Miles Laboratories, Inc.) (5% wt/wt) was performed for 2 h (three times) at 37°C in 25 mM Tris hydrochloride–0.1 mM EDTA–0.025% (wt/vol) SDS (pH 7.8). The digestion mixture (4.5 nmol in 3.0 ml) was chromatographed on a 10-μm Vydac pH-stable reversed-phase column (4.6 by 250 mm) with a hexafluoroacetone (E. Merck AG)-containing solvent system. Solvent A contained (vol/vol) hexafluoroacetone-NH_3 (pH 8.6); solvent B contained 64% (vol/vol) acetonitrile–20% (vol/vol) propanol–0.03% (vol/vol) hexafluoroacetone-NH_3 (pH 8.6). A gradient ranging from 0 to 75% solvent B was applied at a flow rate of 1 ml/min.

Sequence analysis. The amino acid sequences of the N terminus and V8 peptides were determined by automated gas phase Edman degradation with on-line identification of the PTH-amino acids (16). Chemicals and solvents for automated sequence analysis were from Applied Biosystems.

Isolation and analysis of the D19G genomic clone. A genomic library, constructed by insertion of a partial *Sau*3A digest of *D. discoideum* DNA from strain AX2 into the *Bam*HI site of the plasmid pAT153, was screened with an internal *Nsi*I-*Hin*fI fragment from the D19 cDNA clone (1). The genomic clone (D19G) was sequenced by the method of Maxam and Gilbert (26) after labeling at the *Nde*I, *Rsa*I, and *Hin*fI sites (Fig. 1). Other cloning methods used were essentially as described by Maniatis et al. (25).

Insertion of a human c-*myc* epitope into D19G. The c-*myc* oligonucleotide is a 33-base-pair sequence which encodes an eight-amino-acid epitope of the human c-*myc* protein recognized by a MAb, 9E10 (10, 29). The oligonucleotide was synthesized with *D. discoideum* codon usage and inserted in frame into the unique *Pvu*II site in the D19G clone (coding strand: 5'-CT GAA GAA AAA TTA ATT TCA GAA GAA GAT CTC T-3') (Fig. 1). The two strands of the oligonucleotide were annealed at 60°C for 1 h in 1× ligase buffer (50 mM Tris [pH 7.5], 10 mM $MgCl_2$, 10 mM dithiothreitol, 5 mM ATP) and then ligated to the linearized vector with the same buffer. A 50-fold molar excess of nonphosphorylated oligonucleotide over vector was used to ensure efficient insertion of a single copy of the epitope. The oligonucleotide

was designed so that a *Pvu*II site was regenerated at one end while at the other, a *Bgl*II site was created. This latter site was used in sequencing to verify correct orientation of the epitope. The c-*myc*-marked gene was excised from pAT153 by using the *Eco*RI site in the vector and a *Bam*HI site which was regenerated at the 3′ boundary of the insert. By using this fragment, the gene was cloned into the corresponding multilinker sites of the *D. discoideum* transformation vector, pB10TP2 (8).

DNA-mediated transformation and analysis of cellular differentiation. The *D. discoideum* strain AX2 (43) was used for DNA-mediated transformation by the modified protocol described by Early and Williams (8), which is based on the procedure of Nellen et al. (30). Pooled populations of transformants were grown in HL5 medium (43) in shaken suspension. Transformants were developed on membrane filters (Millipore Corp.) after washing by centrifugation in 20 mM KH_2PO_4-K_2HPO_4, pH 6.0 (40). To obtain migrating slugs, cells were developed as single streaks of approximately 10^9 cells per ml on 2% (wt/vol) water agar. The plates were kept moist and dark except for a single slit, which provided a unidirectional light source. Prestalk and prespore cells were separated by the method of Ratner and Borth (33).

Analysis of gene expression by Northern (RNA) transfer. Total cellular nucleic acid was extracted by alternating phenol and chloroform extractions. Five micrograms of each nucleic acid sample was electrophoresed through a formaldehyde denaturing gel and transferred onto nitrocellulose as described by Maniatis et al. (25), except that the filters were baked and prehybridized immediately after transfer. The hybridization was carried out in 6× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% SDS at 32°C with the kinase-labeled, noncoding strand of the c-*myc* oligonucleotide as a probe. The filters were washed for 5 min (twice) at 37°C in 5× SSC–0.1% SDS.

Immunohistochemical analysis of transformants. Disaggregated cells in KH_2PO_4-K_2HPO_4 were prepared by passaging slugs through a series of syringes with needles of successively decreasing bore size. The disaggregated cells were washed and suspended at about 5×10^6/ml in KH_2PO_4-K_2HPO_4 before fixing in 60% methanol and dispensing as 5-μl drops onto poly-L-lysine subbed multiwell slides.

Slides were stained with the affinity-purified c-*myc* MAb, 9E10, prepared from culture fluid (10 μl per field) (10). This was diluted 1:50 in phosphate-buffered saline, and the slides were incubated in a humidified chamber for 30 min at room temperature. The specimens were washed extensively with phosphate-buffered saline for 15 min which was followed by the addition of 10 μl of rhodamine-conjugated rabbit anti-mouse immunoglobulin, diluted 1:30 in phosphate-buffered saline, again for 30 min at room temperature. Washing was carried out as before. The slides were mounted in Gelvatol and examined under a fluorescence phase-contrast microscope.

RESULTS

Purification of PsA by MUD1 affinity chromatography. PsA was purified from slug cell extracts of three strains of *D. discoideum* by MAb MUD1 affinity chromatography. Extracts prepared by lysing WS380B cells in NP-40 were first chromatographed over two guard columns. These columns removed nonspecific material, particularly a 37-kDa protein (this contaminant was later identified as the lectin discoidin). Figure 2 shows three stages of the purification procedure; the original NP-40 extract (lane 2); the material which did not bind to the MUD1 column (lane 3); PsA eluted off the column, seen here to migrate at 27 kDa (lane 4); and the same material as in lane 4, except that it was electroblotted onto nitrocellulose and probed with MAb MUD1 (lane 5).

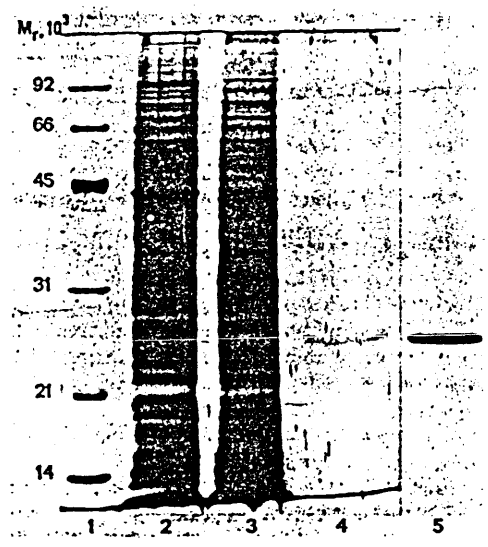


FIG. 2. Purification of PsA from strain WS380B. Protein samples were separated by SDS-PAGE and silver stained. Lanes: 1, molecular weight markers; 2, original 0.1% (vol/vol) NP-40 cell extract; 3, MUD1 affinity column flow through; 4, PsA eluted off the column at pH 11.5; 5, the same material as in lane 4, electroblotted onto nitrocellulose and stained with MAb MUD1 culture supernatant and peroxidase-conjugated second antibody.

PsA purified from this WS380B extract was subjected to automated Edman degradation, and 12 of the first 14 residues were assigned (Table 1). This procedure was repeated for a polymorphic wild strain, WS576 (PsA, 33 kDa), and sequence analysis identified 16 of the first 21 residues, which were identical to the N terminus of WS380B (Table 1). This established that two different-sized versions of PsA had the same N terminus. Typical yields of PsA by this method were 50 to 100 μg of protein per 10^{11} cells. Since insufficient material was recovered to allow further analysis, a purification protocol involving a combination of sodium deoxycholate and NP-40 was devised (see Materials and Methods).

This method resulted in typical yields of 0.5 to 1.5 mg of protein per 10^{11} cells, although the level of purification of PsA was somewhat lower than that achieved by the original method. However, the eluate was much enriched for PsA (Fig. 3, lane 1), seen here to migrate at 22 kDa (strain HU2440, a *modB* mutant lacking a glycoconjugate). The identity of PsA was confirmed by electroblotting the material onto nitrocellulose and probing with MAb (MUD1) (Fig. 3, lane 2). Attempts to further purify PsA resulted in irreversible loss on high-pressure liquid chromatography columns and low recovery from G-252 Sephadex. However, precipitation of the protein in prechilled acetone and reduction and carboxymethylation resulted in either loss of the contaminant or blocking of its N terminus (see below).

N-terminal amino acid sequence of PsA purified from HU2440. PsA purified from HU2440 by MUD1 affinity chromatography was subjected to N-terminal degradation, and 54 residues were identified. Only one sequence could be identified from the preparation (Table 1). Therefore, the contaminating protein was not present in sufficient quantities to interfere with the sequencing of the intact PsA protein.

TABLE 1. Amino acid sequences of PsA purified from three strains

Strain	*pspA* allele[a]	Sequence[b]
WS380B	352	YDYFT--LAN QNPV
WS576[c]	353	YDYFT--LAN QNPV-A-V-V I
HU2440[d]	351	YDYFTTTLAN QNPVCASVDV IQNVCTEVCG RFVRYIPDAT NTNQFTFAEY TTNQ
HU2440[e]	351	YDYFTTTLAN QNPVCASVDV IQNVCTEVCG RFVRYIPDAT NTNQFTFAEY TTNQCTVQVT PAVTNTFTCA DQTSSHALGS D-S

[a] The gene locus which encodes the PsA protein is called *pspA*. Numbers denote alleles for different isolates (12).
[b] Cleavages at the glutamate residues are shown in boldface. Amino acids are designated by the single-letter code. -, Unassigned residue.
[c] WS576 sequence from the 0.1% (vol/vol) NP-40 extract.
[d] Sequence from the material purified by using sodium deoxycholate and NP-40.
[e] Sequences from the V8 digest of PsA purified from HU2440.

V8 peptide map of PsA purified from HU2440 and sequencing of selected peptides. Three major peptides from the V8 digest of PsA purified from HU2440 were separated by reversed-phase high-pressure liquid chromatography (Fig. 4). Peptide I spans from residue 1 to 27, and peptide II spans from residue 28 to 49. The double peak, III, was analyzed to determine whether it represented two separate V8 fragments or whether part of the peak was undigested PsA. Sequence analysis revealed that both peaks had the N terminus Tyr-Thr-Thr-Asn, which is identical to the N-terminal sequence at positions 50 through 53. This suggests that the double peak resulted from a form of microheterogeneity due either to bound SDS, which was present in the V8 digestion buffer (38), or to some form of posttranslational modification (PsA is phosphorylated [Alexander et al., in press]). The V8 fragments IIIA and IIIB were pooled, and the sequence was extended from residue 50 to serine at position 83, with residue 82 unassigned.

A fourth peptide (IV; Fig. 4) from the V8 digest of HU2440 contained at least three sequences, one of which showed 100% homology with a V8 fragment of the lectin discoidin (3). The lack of discoidin in PsA purified by using the guard columns is explained by its removal by binding to these columns.

Isolation and characterization of a D19 genomic clone. Southern blotting of genomic AX2 DNA at high stringency indicated that D19 was a single-copy gene, but at lower stringency, cross-hybridizing restriction fragments were detected (data not shown). A restriction fragment derived from within the D19 cDNA insert was used, at high stringency, to probe a genomic library in the plasmid pAT153. Four independent positive clones were isolated and characterized by restriction mapping and Southern blotting. These four clones have an identical structure in the regions in which they overlap (Fig. 1). The position of the gene was deduced by restriction mapping. The clone with the greatest amount of sequence flanking the D19 coding region was designated D19G and used for all further analyses (Fig. 1).

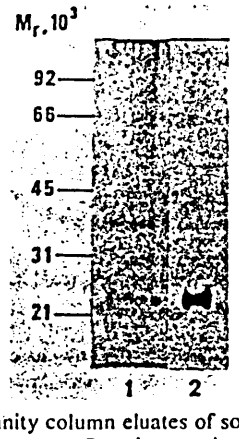


FIG. 3. MUD1 affinity column eluates of sodium deoxycholate–NP-40 HU2440 cell extract. Protein samples were separated by SDS-PAGE. Lane 1, Coomassie blue stain of PsA (*); lane 2, the same sample electroblotted onto nitrocellulose and stained with MAb MUD1 and peroxidase-conjugated second antibody.

The D19G clone contains an insert of 3.37 kilobases (kb). A total of 1.1 kb of nucleotide sequence, spanning the coding and 5′ and 3′ untranslated regions of the gene, was determined. A single start site of transcription was detected by primer extension by using RNA isolated from V12M2 prespore cells or from AX2 slugs (data not shown). The 5′ noncoding region is approximately 90 nucleotides long and has an A+T content of 95%. The length of the 3′ noncoding region was estimated from the cDNA clone to be 150 nucleotides. The coding region is 504 nucleotides long, contains no introns, and encodes a predicted polypeptide of 168 amino acids (Fig. 5). The overall size of the mRNA is estimated to be 736 nucleotides, which is in close agreement with the results of gel electrophoresis.

Structure of the predicted polypeptide. The N terminus of the D19 gene encodes a polypeptide tract which has the characteristics of a signal sequence (6). A hydrophobic "core" of 15 amino acids, with a predicted α-helical structure, follows a short hydrophilic region adjacent to the methionine. Comparison of the sequence on the immediate C-terminal side of the predicted cleavage site (i.e., following Ala-Asn-Ala) shows it to be identical to the 81 N-terminal residues of the PsA protein. Since we have shown the D19 gene to be unique in the genome, it must encode the PsA protein. The protein encoded by the D19 gene has a predicted molecular weight of 17,879 and a predicted pI of 4.62. It is composed of a higher-than-average proportion of nonpolar residues and a lower-than-average proportion of potentially charged residues (Fig. 6). Remarkably, it contains over 21% threonine residues. These are potential sites for O-linked glycosylation. Although there is one potential N-linked glycosylation site (Asn-Pro-Thr; Fig. 5), this is not utilized (4). Bause (4) has shown that no glycosyl transfer was detected on those peptides which contain a proline residue in position X in the Asn-X-Ser/Thr triplet sequence.

The extreme C terminus contains a largely hydrophobic stretch of approximately 18 amino acids, but these are not followed by polar residues which could act as a membrane anchor sequence (Fig. 5 and 6). This region is interspersed with hydrophilic residues, most commonly serine. Immediately before the hydrophobic tail there are four copies of a

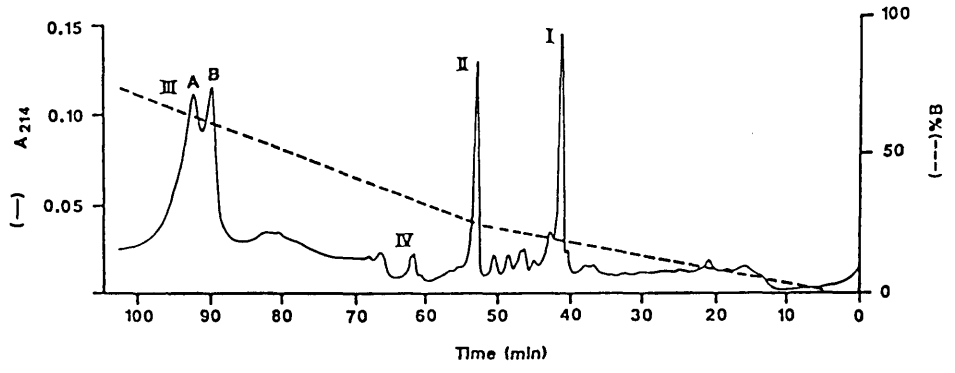


FIG. 4. Separation of peptides from a V8 digest of PsA purified from HU2440. The digestion mixture was separated on a Vydac C8 pH-stable column in an acetonitrile-propanol (solvent B) gradient at a flow rate of 1 ml/min. PsA peptides are designated I, II, and IIIA and B with respect to the time they eluted off the column. Peak IV is a peptide of the lectin discoidin. %B, Percentage of solvent B.

perfectly conserved, four-amino-acid repeat, Pro-Thr-Val-Thr. These repeats are followed by three similar, but imperfect, repeats. The repeats form a proline-rich region, containing 10 of the total 13 proline residues present in the protein. Within this 31-amino-acid stretch there are nine short regions where proline alternates with a hydroxyamino acid, normally a threonine residue. This pattern is interspersed with valine, asparagine, or alanine residues, present as single amino acids. Overall, the primary amino acid sequence predicts a polypeptide with an extended structure.

Expression of a marked copy of the gene after transformation into *D. discoideum* and localization of the protein product. The PsA gene was marked by insertion of an oligonucleotide encoding an eight-amino-acid epitope of the human c-*myc* protein (29), which is recognized by the 9E10 MAb (10). The epitope was inserted in frame within codon 81. The marked gene was subcloned into the *D. discoideum* transformation vector pB10TP2 (8). The final construct, pB10D19G, contained the entire coding region with approximately 1.0 kb of sequence upstream of the cap site and about 1.9 kb of downstream sequence (Fig. 7).

After transformation into AX2 cells and selection of stable transformants by using G418, expression of the gene was analyzed in pooled populations of cells. The vector copy number within the population was approximately 10. The mRNA transcribed from the marked D19 gene was analyzed by Northern transfer hybridization by using the noncoding strand of the c-*myc* oligonucleotide as a probe. This mRNA accumulated with the same kinetics as the transcript of the endogenous PsA gene (Fig. 8A). The mRNA was first

```
1                                                    ATTTTATTAATATAATTCTTAATAATTATAAAAAAAAAAAAAAA

46 AAACAACCAAAATATTATATTATTTTTTGTAAATATAAATTAAAAAA ATG AAA TTC CAA CAT ACA
                                                      MET Lys Phe Gln His Thr   6

111 TTT ATT GCA TTA TTA TCA CTA TTA ACA TAT GCA AAT GCA TAT GAT TAT TTC ACA
    Phe Ile Ala Leu Leu Ser Leu Leu Thr Tyr Ala Asn Ala Tyr Asp Tyr Phe Thr  24

165 ACA ACA TTG GCC AAT CAA AAT CCA GTT TGT GCT TCA GTA GAT GTC ATA CAA AAT
    Thr Thr Leu Ala Asn Gln Asn Pro Val Cys Ala Ser Val Asp Val Ile Gln Asn  42

219 GTT TGT ACT GAG GTT TGT GGT AGA TTT GTC CGT TAC ATT CCT GAT GCT ACC AAT
    Val Cys Thr Glu Val Cys Gly Arg Phe Val Arg Tyr Ile Pro Asp Ala Thr Asn  60

273 ACA AAT CAA TTC ACT TTT GCC GAA TAT ACT ACA AAC CAA TGT ACT GTT CAA GTA
    Thr Asn Gln Phe Thr Phe Ala Glu Tyr Thr Thr Asn Gln Cys Thr Val Gln Val  78

327 ACA CCA GCT GTA ACA AAT ACA TTT ACT TGC GCT GAT CAA ACT TCT TCA CAT GCA
    Thr Pro Ala Val Thr Asn Thr Phe Thr Cys Ala Asp Gln Thr Ser Ser His Ala  96

381 CTC GGT TCT GAT TGG AGT GGA GTT TGT AAA ATC ACT GCC ACA CCC GCA CCA ACA
    Leu Gly Ser Asp Trp Ser Gly Val Cys Lys Ile Thr Ala Thr Pro Ala Pro Thr 114

435 GTT ACA CCA ACA GTT ACA CCA ACA GTT ACA CCA ACA GTA ACA CCA ACA CCA ACA
    Val Thr Pro Thr Val Thr Pro Thr Val Thr Pro Thr Val Thr Pro Thr Pro Thr 132

489 AAT ACA CCA AAT CCA ACC CCG TCC CAA ACA TCA ACA ACA ACT GGT TCT GCT TCA
    Asn Thr Pro Asn Pro Thr Pro Ser Gln Thr Ser Thr Thr Thr Gly Ser Ala Ser 150
                    •   •   •

543 ACT GTT GTT GCA AGT TTA TCA CTC ATT ATT TTC TCA ATG ATT CTT TCC CTT TGT
    Thr Val Val Ala Ser Leu Ser Leu Ile Ile Phe Ser MET Ile Leu Ser Leu Cys 168

597 TAATTTTAAAAATAACAAATGGTAAAAATTATTTGGCATTTTCAAGTACAATTTTAAATTTTTCAAAAAAA

668 AATAAAATTAAATAAATAATTAAATAAATTAATTAATAAAAAAATAAAATTAAAACTTTTTAATTTA
```

FIG. 5. Sequence of the D19 mRNA. The sequence was obtained from the genomic clone D19G by the chemical cleavage method (26). The open reading frame shown is confirmed by comparison with the protein sequence (Table 1). The length of the 5' noncoding region was predicted by primer extension analysis (data not shown), while the length of the 3' noncoding region was deduced by comparison with sequence from the cDNA clone. Symbols: ——, position of the four tetra-amino acid repeats; •••, Asn-Pro-Thr triplet sequence which, although it conforms to the Asn-X-Thr sequence, is not a predicted site for N-linked glycosylation (4). The numbers on the left represent nucleotides, while the numbers on the right represent amino acids.

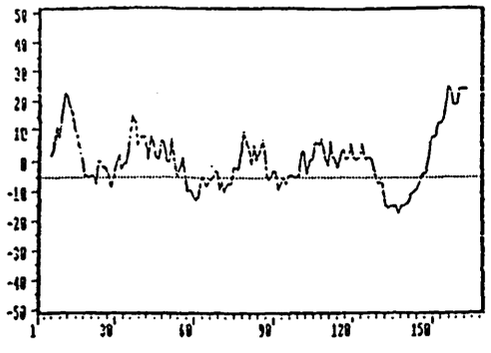


FIG. 6. Plot of the hydropathic index of the predicted primary translation product of the D19 gene. The plot was computed according to the method of Kyte and Doolittle (21) with a window size of nine amino acids. The 168 amino acids of the predicted D19 coding region are represented along the x axis. On the y axis, positive values indicate hydrophobicity, while negative values indicate hydrophilicity.

detectable at the tipped aggregate stage, reached a peak of abundance at the first finger and standing slug stages, and fell in concentration at culmination. The mRNA also showed correct spatial localization. Slugs of the transformed population were disaggregated, and the cells were separated by two rounds of Percoll density gradient centrifugation. Comparison of the light (prestalk) and heavy (prespore) mRNA fractions showed strong enrichment of the mRNA in prespore cells (Fig. 8B). The D19 cDNA insert detected mRNA from both the endogenous and transformed genes as a single band when it was used as a probe in Northern transfer. In a transformed population with approximately 5 to 10 copies of the transforming gene, there was approximately five times as much mRNA at this position in the gel as was found in an equivalent number of AX2 cells harvested at the same stages of development (data not shown). Thus, 990 base pairs of sequence upstream of the cap site contains all the information required for correctly regulated expression at approximately the level of the endogenous gene.

The transformed population was screened for expression of the marked gene by indirect immunofluorescence (Fig. 9). Methanol-fixed, disaggregated slug cells were stained with the anti-c-*myc* MAb and with a rhodamine-conjugated second antibody. Bright, cell surface staining of a high proportion of the transformed cells was seen, but there was no staining of (control) cells harvested from nontransformed slugs (data not shown).

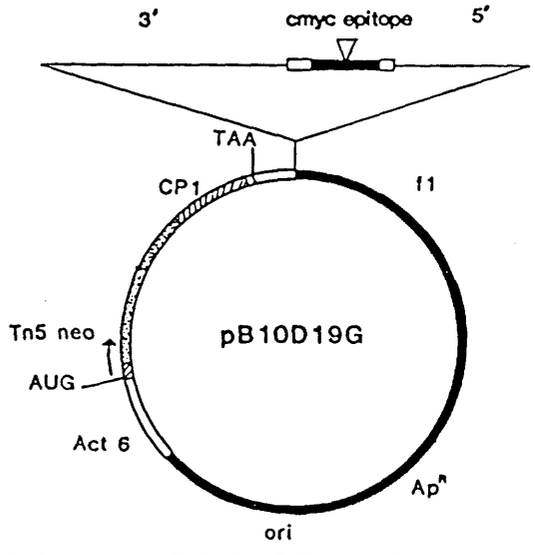


FIG. 7. Structure of pB10D19G. The plasmid was constructed by insertion of the entire D19G genomic fragment into the transformation vector, pB10TP2 (8). The position and orientation of the D19 gene within the insert are shown. The relative position of the c-*myc* epitope within the coding region is also indicated. The solid black line within the vector represents sequences derived from pEMBL. The actin 6 promoter (Act 6; ▭) drives expression of the Tn*5* neomycin resistance gene (Tn5 neo; ▭). This transcript, which enables the selection of *Dictyostelium* cells resistant to G418, is terminated with sequences from the cysteine proteinase 1 gene (CP1; ▨). ori, Origin of replication; f1, phage f1.

DISCUSSION

The predicted amino acid sequence of the coding region of the D19 gene is identical to the first 81 residues of PsA purified by MAb MUD1 affinity chromatography. Other aspects of the DNA sequence correlate well with the physical properties of the PsA protein. The pI of PsA is between 4 and 5 (12), and the pI predicted from the primary sequence of the D19 polypeptide is 4.62. The sequence of the PsA gene indicates that there is a signal peptide and the predicted cleavage site would generate an N terminus identical to that obtained for the protein. However, the molecular weight of the protein is predicted to be 17,879, while that of PsA has been judged to be 30,000 (19). This apparent discrepancy can be accounted for by posttranslational modifications. PsA and *D. discoideum* contact sites A (CsA) are glycoproteins which show major decreases in size in *modB* mutants (11, 44; Alexander et al., in press). The primary defect caused by the *modB* mutation is thought to be in the mechanism responsible for the retardation of the protein until type 2 glycosylation occurs (15).

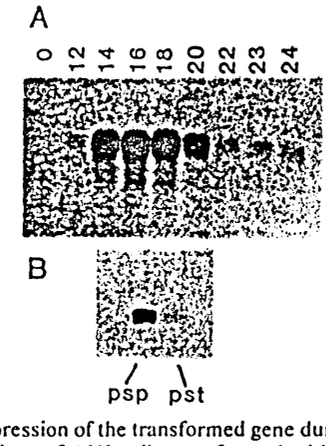


FIG. 8. Expression of the transformed gene during development. Pooled populations of AX2 cells transformed with pB10D19G were analyzed. Total cellular RNA was extracted and used for Northern transfer analysis. The noncoding strand of the c-*myc* oligonucleotide was end-labeled with ^{32}P by using T4 polynucleotide kinase. This probe specifically detects mRNA from the transformed gene. (A) Times shown (h) indicate the interval after initiation of development. The corresponding morphological stages are as follows: 12 h, loose mounds; 14 h, tipped mounds; 16 h, first finger; 18 h, standing slug; 20 h, late preculminates; 24 h, culminates. (B) Disaggregated slug cells were separated by two rounds of Percoll density gradient centrifugation (33), and RNA was extracted from isolated prespore (psp) and prestalk (pst) cells.



FIG. 9. Immunostaining of transformants. Slug cells, disaggregated and fixed with methanol, were stained with the anti-c-*myc* MAb 9E10 and a rhodamine-conjugated second antibody before microscopy. Magnification, ×180.

Type 2 carbohydrate accounts for 12 kDa of the molecular mass of CsA (11) and for 7 kDa of the PsA apparent molecular mass as determined by SDS-PAGE (30 kDa in NC4 and 23 kDa in the *modB* mutant of NC4, HU2428; data not shown). Therefore, in PsA the *modB*-sensitive carbohydrate accounts for 7 kDa of the 12-kDa difference between the molecular mass predicted from the translated D19 gene and the apparent molecular mass determined by SDS-PAGE. The difference between the molecular mass predicted from the DNA sequence and the apparent molecular mass determined by SDS-PAGE may result from the unusually high content of hydroxyamino acids (21% threonine and 8% serine). A similar discrepancy, observed between the predicted size of VP9 (22 kDa), a minor structural component of a rotavirus capsid, and the apparent size of the nonglycosylated protein on SDS-PAGE (28 kDa), has been attributed to its unusually high content (20%) of serine residues (5). None of the Thr or Ser residues in the first 81 residues of PsA can be glycosylated, since glycosylated residues are not soluble in the sequencing reagents. Therefore, all of the carbohydrate is predicted to be on the Thr and Ser residues downstream of residue 81.

Despite the fact that PsA is a known plasma membrane protein, the sequence data does not predict a hydrophobic, transmembrane segment followed by a hydrophilic cytoplasmic "anchor" domain. Rather, the C terminus closely resembles the equivalent regions of a set of cell surface glycoproteins known to be attached to the membrane via a glycosyl-phosphatidylinositol link (23). At the extreme C terminus of these proteins, largely hydrophobic residues are interspersed with hydrophilic residues, often serine or threonine. In the mature polypeptide, the hydrophobic tails are removed by cleavage and the lipid is attached via an amide linkage to the new C-terminal amino acid. The PsA protein shows direct homology in the C-terminal region to two proteins involved in cell adhesion, *D. discoideum* CsA (31) and a form of chicken neural-cell adhesion molecule known to be linked to the membrane via a glycosyl-phosphatidylinositol link (Fig. 10) (14). It seems almost certain that both PsA and CsA are also attached to the cell membrane via a glycolipid. Direct analysis of CsA shows that it is attached in such a manner (J. Stadler, G. Bauer, and G. Gerisch, personal communication). The predicted glycosyl-phosphatidylinositol linkage would allow the rapid removal of PsA protein from the cell surface by the action of a specific phospholipase, releasing cleaved protein into the extracellular matrix. This potential rapid loss could be important during differentiation of *D. discoideum*, which involves rapid changes in the presentation of surface molecules (34).

There is homology between PsA and CsA upstream of the extreme C terminus in the proline-rich repeat region (31) (Fig. 10). In PsA from strain AX2, this sequence is 31 amino acids long and contains nine short, interspersed groups of alternating threonines and prolines. In the CsA protein, the comparable region is 32 amino acids long and contains three longer clusters of proline residues which alternate with threonine or serine. Such a primary sequence is capable of forming an extended chain in proline-proline conformation, which could be important in allowing the projection of these proteins into the extracellular space. It is possible therefore that the homology merely indicates a requirement for a similar stalk-like structure, but in combination with the extreme C-terminal homology to CsA and neural-cell adhesion molecules, this internal homology suggests that these three genes may belong to the same family. Direct evidence is, however, clearly required to prove a role of PsA in cell adhesion.

The repeat region in the C terminus could also be responsible for the polymorphism in size of the PsA protein. In different *D. discoideum* strains there is a difference in apparent molecular mass of ±2 kDa (12). The repeats in PsA are four amino acids long and there are four perfectly conserved copies. Unequal crossover would result in the loss of up to three copies of the repeat from one chromosome and the insertion of up to three extra copies into the other; at least one copy of the repeat would always be retained. In this way, exact insertion or deletion of up to 12 amino acids

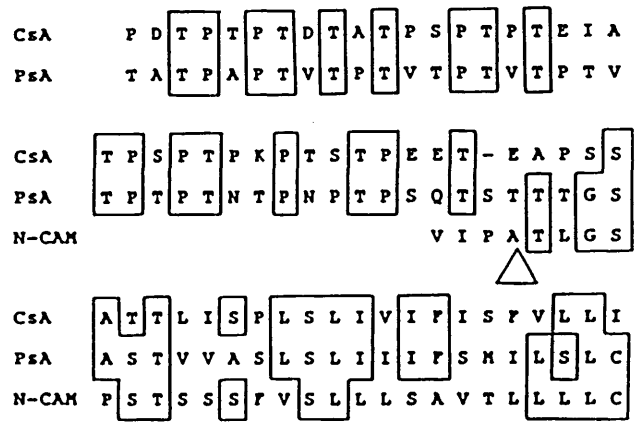


FIG. 10. Comparison of PsA amino acid sequence with CsA and neural-cell adhesion molecules. The C-terminal 63 amino acids of PsA are shown, aligned with the C-terminal 62 amino acids of CsA (31) and with the 25 amino acids unique to the form of chicken neural-cell adhesion molecule known to be attached to the membrane via phosphatidylinositol (14). The arrowhead indicates the boundary between this unique sequence and sequence from the previous exon. The boxed regions indicate the occurrence of identical amino acids.

might occur. The repeat contains threonines which are potential sites for O-linked glycosylation, and so these 12-amino-acid changes could account for the 2-kDa size polymorphisms.

The results of the transformation show that 990 base pairs upstream of the cap site of D19 are sufficient for correct temporal and spatial mRNA accumulation. It is not clear whether the signals for prespore-specific expression reside within the promoter or whether a generalized promoter element combined with posttranscriptional control leads to a specifically localized transcript. The upstream and 5′ noncoding sequences of one other prespore-specific gene, EB4, have been reported (2). The EB4 and PsA genes are expressed at similar times in development. The transcripts from both genes are rapidly lost because of selective destabilization when aggregates are disrupted, and exogenous cAMP acts to prevent this destabilization (24). One suggestion was that the extremely long, relatively GC-rich, 5′ noncoding region of EB4 might be involved in its posttranscriptional regulation. However, there is no obvious similarity to the 5′ noncoding region of PsA, which is much shorter and has a much lower GC content.

Since we have obtained correctly regulated expression of the transformed D19 gene and an appropriate localization of the marked protein, it will now be possible to use transformation both to localize the signals required for prespore-specific gene expression and to investigate the possible functions of the protein in the migratory slug.

ACKNOWLEDGMENTS

This research was supported by ARGS grant A08415732, Macquarie University Research grants to K.L.W. and E.S., and NH&MRC grant 860245. S.B.P. was supported by a Macquarie University Research Fellowship. The visit of A.A.G. to the laboratory of H.E.M. was supported by the Deutsche Forschungsgemeinschaft, SFB 168.

We thank S. McRobbie for his suggestion that the D19 gene might encode the PsA protein.

LITERATURE CITED

1. Barklis, E., and H. F. Lodish. 1983. Regulation of *Dictyostelium discoideum* mRNAs specific for prespore or prestalk cells. Cell 32:1139–1148.
2. Barklis, E., B. Pontius, K. Barfield, and H. F. Lodish. 1985. Structure of the promoter of the *Dictyostelium discoideum* prespore EB4 gene. Mol. Cell. Biol. 5:1465–1472.
3. Bartles, J. R., W. A. Frazier, and S. D. Rosen. 1982. Slime mold lectins. Int. Rev. Cytol. 75:61–99.
4. Bause, E. 1983. Structural requirements of N-glycosylation of proteins. Biochem. J. 209:331–336.
5. Both, G. R. 1987. Replication of the Reoviridae: information derived from gene cloning and expression. *In* P. Ahlquist, E. Domingo, and J. Holland (ed.), RNA Genetics 1987. CRC Press, Inc., Boca Raton, Fla.
6. Briggs, M. S., and L. M. Gierasch. 1986. Molecular mechanism of protein secretion: the role of the signal sequence. Adv. Protein Chem. 38:109–180.
7. Chisholm, R. L., E. Barklis, and H. F. Lodish. 1984. Mechanism of sequential induction of cell-type specific mRNAs in *Dictyostelium discoideum*. Nature (London) 310:67–69.
8. Early, A. E., and J. G. Williams. 1987. Two vectors which facilitate gene manipulation and a simplified transformation procedure for *Dictyostelium discoideum*. Gene 59:99–106.
9. Erdos, G. W., K. B. Raper, and L. K. Vogen. 1973. Mating types and macrocyst formation in *Dictyostelium discoideum*. Proc. Natl. Acad. Sci. USA 70:1828–1830.
10. Evan, G. I., G. K. Lewis, G. Ramsay, and J. M. Bishop. 1985. Isolation of monoclonal antibodies specific for human c-*myc* proto-oncogene product. Mol. Cell. Biol. 5:3610–3616.
11. Gerisch, G., U. Weinhart, G. Bertholdt, M. Claviez, and J. Stadler. 1985. Incomplete contact site A glycoprotein in HL220, a *modB* mutant of *Dictyostelium discoideum*. J. Cell Sci. 73:49–68.
12. Grant, W. N., D. L. Welker, and K. L. Williams. 1985. A polymorphic, prespore-specific cell surface glycoprotein is present in the extracellular matrix of *Dictyostelium discoideum*. Mol. Cell. Biol. 5:2559–2566.
13. Grant, W. N., and K. L. Williams. 1983. Monoclonal antibody characterisation of slime sheath: the extracellular matrix of *Dictyostelium discoideum*. EMBO J. 2:935–940.
14. Hemperley, J. J., G. M. Edelman, and B. A. Cunningham. 1986. cDNA clones of the neural cell adhesion molecule (N-CAM) lacking a membrane-spanning region consistent with evidence for membrane attachment via a phosphatidylinositol intermediate. Proc. Natl. Acad. Sci. USA 83:9822–9826.
15. Hohmann, H. P., S. Bozzaro, R. Merkl, E. Wallraff, M. Yoshida, U. Weinhart, and G. Gerisch. 1987. Posttranslational glycosylation of the contact site A protein of *Dictyostelium discoideum* is important for stability but not for its function in cell adhesion. EMBO J. 6:3663–3671.
16. Hunkapillar, M. W., R. M. Hewick, W. J. Dreyer, and L. E. Hood. 1983. High-sensitivity sequencing with a gas-phase sequenator. Methods Enzymol. 91:399–413.
17. Jermyn, K. A., M. Berks, R. R. Kay, and J. G. Williams. 1981. Two distinct classes of prestalk-enriched mRNA sequences in *Dictyostelium discoideum*. Development 100:745–755.
18. Krefft, M., L. Voet, J. H. Gregg, H. Mairhofer, and K. L. Williams. 1984. Evidence that positional information is used to establish the prestalk-prespore pattern in *Dictyostelium discoideum* aggregates. EMBO J. 3:201–206.
19. Krefft, M., L. Voet, H. Mairhofer, and K. L. Williams. 1983. Analysis of proportion regulation in slugs of *Dictyostelium discoideum* using a monoclonal antibody and a FACS-IV. Exp. Cell Res. 147:235–239.
20. Kyhse-Anderson, J. 1984. Electroblotting of multiple gels: a simple apparatus without buffer tank for rapid transfer of protein from polyacrylamide to nitrocellulose. J. Biochem. Biophys. Methods 10:203–209.
21. Kyte, J., and R. F. Doolittle. 1982. A simple method for displaying the hydropathic character of a protein. J. Mol. Biol. 157:105–132.
22. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) 227:680–685.
23. Low, M. G., M. A. J. Ferguson, A. H. Futerman, and I. Silman. 1986. Covalently attached phosphatidylinositol as a hydrophobic anchor membrane proteins. Trends Biochem. Sci. 11:212–215.
24. Mangiarotti, G., A. Ceccarelli, and H. F. Lodish. 1983. Cyclic AMP stabilises a class of developmentally regulated *Dictyostelium discoideum* mRNAs. Nature (London) 301:616–618.
25. Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
26. Maxam, A., and W. Gilbert. 1980. Sequencing end-labelled DNA with base-specific chemical cleavages. Methods Enzymol. 65:499–560.
27. Mehdy, M. C., D. Ratner, and R. A. Firtel. 1983. Induction of cell-type-specific expression in *Dictyostelium*. Cell 32:763–771.
28. Meyer, H. E., E. Hoffmann-Posorske, C. C. Kuhn, and L. M. G. Heilmeyer. 1988. Microcharacterisation of phosphoserine-containing proteins. Localisation of the autophosphorylation of skeletal muscle phosphorylase kinase, p. 185–212. *In* H. Tschesche (ed.), Modern methods in protein chemistry, vol. 3. Walter de Gruyter, Verlag, Holland.
29. Munro, S., and H. R. B. Pelham. 1986. An Hsp70-like protein in the ER: identity with the 78kd glucose-regulated protein and immunoglobulin heavy chain binding protein. Cell 46:291–300.
30. Nellen, W., C. Silan, and R. A. Firtel. 1984. DNA-mediated transformation in *Dictyostelium discoideum*: regulated expression of an actin gene fusion. Mol. Cell. Biol. 4:2890–2898.
31. Noegel, A., G. Gerisch, J. Stadler, and M. Westphal. 1986.

Complete sequence and transcript regulation of a cell adhesion protein from aggregating *Dictyostelium* cells. EMBO J. 5:1473–1476.

32. Oyama, M., and D. D. Blumberg. 1986. Changes during differentiation in requirements for cAMP for expression of cell-type-specific mRNAs in the cellular slime mould, *Dictyostelium discoideum*. Dev. Biol. 117:550–556.
33. Ratner, D., and W. Borth. 1983. Comparison of differentiating *Dictyostelium discoideum* cell types separated by an improved method of density gradient centrifugation. Exp. Cell Res. 143:1–73.
34. Sadeghi, H., K. L. Williams, and C. Klein. 1987. Characterization of 117 antigen: a developmentally regulated cell surface glycoprotein of *Dictyostelium discoideum*. J. Biol. Chem. 262: 16294–16299.
35. Sandring, R., and A. F. Williams. 1978. Glycoproteins and antigens of membranes prepared from rat thymocytes after lysis by shearing or with the detergent Tween-40. Biochim. Biophys. Acta 508:85–96.
36. Schaap, P., M. M. Van Looheren Campagne, R. Van Driel, W. Spek, P. J. M. Van Haastert, and J. Pinas. 1986. Postaggregative differentiation induction by cAMP in *Dictyostelium*: intracellular transduction pathway and requirement for additional stimuli. Dev. Biol. 118:52–63.
37. Schaap, P., and M. Wang. 1986. Interactions between adenosine and oscillatory cAMP signaling regulate size and pattern in *Dictyostelium*. Cell 45:137–144.
38. Shively, J. E. 1986. Reverse phase HPLC, p. 41–87. *In* J. E. Shively (ed.), Methods of protein microcharacterisation: a practical handbook. Humana Press, Clifton, N.J.
39. Smith, E., K. Roberts, G. W. Butcher, and G. Galfre. 1984. Monoclonal antibody screening: two methods using antigens immobilized on nitrocellulose. Anal. Biochem. 138:119–124.
40. Sussman, M. 1966. Biochemical and genetic methods in the study of cellular slime mold development. Methods Cell Physiol. 2:397–410.
41. Towbin, H., T. Staehelin, and J. Gordon. 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA 76:4350–4354.
42. Tsang, A., and J. M. Bradbury. 1981. Separation and properties of prestalk and prespore cells of *Dictyostelium discoideum*. Exp. Cell Res. 132:433–441.
43. Watts, D. J., and J. M. Ashworth. 1970. Growth of myxamoebae of the cellular slime mould *Dictyostelium discoideum*. Biochem. J. 119:171–174.
44. West, C. M., and W. F. Loomis. 1985. Absence of a carbohydrate modification does not affect the level or subcellular localisation of three membrane glycoproteins in *modB* mutants of *Dictyostelium discoideum*. J. Biol. Chem. 260:13803–13809.