Isolation and characterization of a partial sequence encoding a protein with a putative Leucine .Zipper domain in *Dictyostelium discoideum.*

Anastasia S. Protopapas

A thesis

in

the Department

of

Biology

Presented in partial fulfillment of the requirements

for the degree of Master of Science at

Concordia University

Montreal, Quebec, Canada

January, 1995

ISBN 0-612-05124-2

Canada

ABSTRACT

The *zip*A gene of *Dictyostelium discoideum* encodes a protein with a leucine zipper motif (ZipA). A cDNA and a genomic clone corresponding to parts of *zip*A have been previously isolated. In the present studies, additional *zip*A cDNA clones were isolated. Genomic analysis by DNA blotting verified that the cDNA and genomic clones carry overlapping sequences of the *zip*A gene. Sequence analysis of these provided a composite 3197 bp long sequence, encoding a single open reading frame (ORF) of 1005 amino acids. The deduced amino acid sequence suggested structural features common to the bZip family of transcription factors. These include a leucine zipper motif with an adjacent basic region and a more distal acidic domain.

To identify the length of coding sequences in *zip*A we preformed RNA blot analysis. A single mRNA species of at least 8.0 kb in size was detected in Northern blots. In a complementary approach, specific anti-ZipA antibodies were generated. Immunoblot analysis revealed a nuclear protein with molecular weight of 60 kDa. The nuclear localization of ZipA is consistent with the idea that *zip*A encodes a transcription factor.

The size discrepancy between the ZipA polypeptide detected by Immunoblot analysis and the transcript detected by RNA blot analysis may suggest long untranslated regions in the transcript and/or postranslational modifications of the protein product.

ACKNOWLEDGEMENTS

I would like to acknowledge my supervisor Dr. A. Tsang for his guidance, help and advice on all aspects of the work. I am grateful to the members of my supervisory committee Drs. E. B. Newman and P.J. Gulick for generously offering their time and advice during the course of my studies. I would like to thank all my colleagues in the laboratory; special thanks go to Claire Bonfils, Michael Greenwood and Bruce Williams, who formed the team I initially joined, for the help and patient teaching both at the theoretical and practical level. I also thank Robert Wennington, Amalia Martinez, Abraham Shtevi, Jonathan Gisser, Guy Czaika, Pascale Gaudete and Zeina Saikali, for the help and sharing and for making the laboratory a truly enjoyable place to work. Also thanks are extended to all the colleagues -past and present- from neighbouring labs, for sharing their experience and knowledge, and for kindly offering their help.

CONTENTS

LIST OF FIGURES

ABBREVIATIONS

bp	base pair
BSA	bovine serum albumin
cAMP	cyclic AMP
cDNA	complementary DNA
Denhardt's	1X = 0.2% each of bovine serum albumin, Ficoll and polyvinylpyrrolidone
EDTA	ethylene diamine tetraacetic acid
kb	kilobases
mCi	milliCurie
min	minutes
mRNA	messenger ribonucleic acid
ORF	open reading frame
poly(A)+	polyadenlyated
SDS	sodium dodecyl sulphate
SDS-PAGE	SDS polyacrylamide gel electrophoresis
sec	seconds
µCi	microCurie
µg	microgram
µL	microliter

1. INTRODUCTION

In nature the cellular slime mold *Dictyostelium discoideum* is found in the forest litter and upper soil layers of the forest floor. The life cycle of *D. discoideum* occurs as follows. In their vegetative phase, *D. discoideum* cells grow as a population of solitary amoebae that live on bacteria and multiply by binary fission. When their food supply becomes low they enter a developmental program, starting with the aggregation of single cells into multicellular organisms that eventually develop into fruiting bodies consisting of only two cell types -spore and stalk cells (Loomis, 1982).

D. discoideum offers many advantages for studying both the molecular and cellular aspects of development in a relatively simple system. *Dictyostelium* is easily cultivated as single cells either axenically or in association with bacteria. It has a haploid genome of approximately 40,000 kb arranged on 7 chromosomes whose complexity is only 2% of that of mammalian cells. *D. discoideum* has a simple developmental cycle during which cells of a single type differentiate without growth into one of two alternate cell types. This distinct separation between the growth phase and the stages of morphogenesis and differentiation render *D. discoideum* an ideal system for the study of developmental processes. The entire developmental sequence can be studied under rigorously defined experimental conditions over about 24 hours.

study of developmental processes. The entire developmental sequence can be studied under rigorously defined experimental conditions over about 24 hours. Sufficient numbers of cells can be isolated to permit various biochemical assays. Development can be made more synchronous if cells are removed from nutrients by centrifugation and resuspended in an inorganic salt solution, then plated on filter paper or non-nutrient agar plates. To a certain extent differentiation in this organism is reversible because cells in the early stages of development, if disaggregated and placed in growth medium, can resume division and growth. Genetic analysis is possible although difficult. Parasexual genetics involve the rare fusion of two haploid strains to yield a diploid, followed by the rare chance production of haploid segregants. Parasexual genetics have been useful for mapping a number of mutations to complementation and linkage groups (Newell, 1982). There also exists a sexual cycle, where cells of opposite mating types fuse, eventually forming a large macrocyst. This cyst contains a diploid cell that germinates, yielding haploid daughters. The inability thus far to induce rapid germination, poses a limitation to genetic studies in this system (Firtel and Bonner, 1972).

A number of diffusible factors have been identified

that influence *D. discoideum* development, including cyclic 3': 5'-adenosine monophosphate (cAMP), Differentiation Inducing Factor (DIF), adenosine and ammonia. cAMP is the best studied and it plays multiple roles during development. It is the chemoattractant to which aggregating amoebae respond (Konijin *et al.*, 1967). cAMP has also been shown to function as a morphogen (Town *et al.*, 1976; Kay, 1982) and also to be involved in the control of expression of developmentally regulated genes (Williams *et al.*, 1980; Chung *et al.*, 1981; Kimmel and Carlisle, 1986). The unravelling of the molecular processes which mediate the effects of cAMP on *D. discoideum* development is an area of intense study. A novel cAMP binding protein (Tsang *et al.*, 1986) that may play a role in those processes has been identified. Characterization of this protein led to the identification of a number of other proteins with which it shares some similarity at the level of primary structure, but which have diverse biochemical properties (Tsang *et al.*, 1988). These proteins include Annexin VII, p34/31, p24, (Grant *et al.*, 1990; Bain *et al.*, 1992; Grenwood *et al.*, 1991; Greenwood *et al.*, 1993) and two proteins currently under characterization -a protein with predicted Ca^{++} binding activity and a protein with a leucine zipper motif.

This thesis deals with the isolation and characterization of the gene encoding the "leucine zipper"

protein and the immunological characterization of the protein. In this chapter the basic life cycle of *Dictyostelium discoideum* will be introduced. The regulation of cellular differentiation and pattern formation will be discussed in separate sections. A final section will deal with the developmental control of gene expression as this is mediated by cAMP.

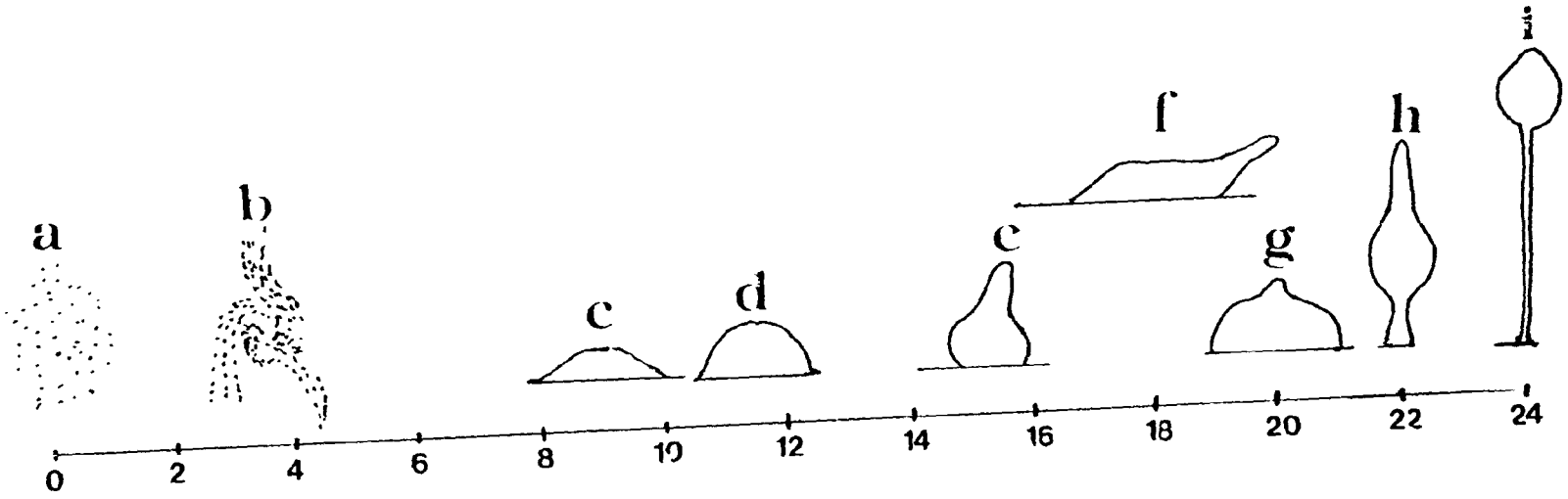
1.1 Development of the cellular slime mold *Dictyostelium discoideum.*

1.1.1 Aggregation

Vegetative amoebae (Fig. I, stage a) of *D. discoideum* begin to aggregate 6-7 hours after exhaustion of their food supply has occurred. Aggregation is initiated by a small number of cells in the starved population, that autonomously release periodic pulses of cAMP. Other cells respond to the cAMP by moving towards the signal and also by releasing their own cAMP, eventually forming pulsating streams of aggregating cells (Fig.I, stage b). Membrane bound cAMP receptors detect the signal. Each time a pulse of cAMP is detected by the cAMP membrane receptors, the cells undergo a surge of amoeboid movement in the direction where their cAMP receptors are first filled. Thus aggregation is a result of two processes: cAMP signal relay and chemotaxis.

Fig. I Diagram showing stages in *D. discoideum development*. Time is given in hours after the onset of starvation. a: vegetative stage; b: aggregating streams of cells; c: aggregate (~10^5 cells) in loose mount formation; d: aggregate (~10^5 cells) in tight mount formation; e: first finger; f: migrating slug; g: mexican hat; h: culmination; i: fruiting body.

Developmental stage



Time of development

Membrane bound cAMP receptors which detect the chemoattractant become temporarily insensitive to it after release of a cAMP pulse from their respective cells. Membrane bound or secreted cAMP phosphodiesterases that destroy the surrounding cAMP allow de-adaptation of the processes of signal relay and chemotaxis (Malchow *et al.*, 1972; Pannbacker and Bravard, 1972). cAMP receptors resensitized to the signal mediate the initiation of a new wave of cAMP, resulting in the pulsatile movement of aggregating cells. Separate classes of cAMP receptors are defined on the basis of the cAMP dissociation rate. The "slow" receptors are thought to be coupled to the chemotactic response (Janssens and van Haastert, 1987) while the "fast" receptors are thought to be associated with the activation of adenylate cyclase, and an increase in the levels of intracellular cAMP. After several minutes of continuous occupancy, phosphorylation of the receptors is induced leading to non-responsiveness to cAMP although the receptor maintains cAMP binding ability. Removal of cAMP by phosphodiesterase (PDE) (Kessin *et al.*, 1979) induces loss of receptor phosphorylation and resensitization of cells to cAMP. The chemotactic response is also regulated by the binding of cAMP to cell surface cAMP receptors. Chemotaxis describes the cell movement up the cAMP gradient; once the peak of cAMP concentration in a cAMP gradient is reached,

response of the cells to each wave lasts about 2 minutes. Activation of chemotaxis is also reflected in a transient 5-10 fold increase of cGMP, which peaks at ~10 sec, as a result of guanylate cyclase activation. The increase of cGMP concentration seems to have a control over the chemotactic process. Such a role was originally suggested from studies on an aggregation mutant called steamer F (Ross and Newell, 1981). These mutants exhibited unnaturally long periods of chemotactic movement during aggregation, coupled with prolonged intervals of elevated cGMP concentration, due to very low levels of intracellular cGMP phosphodiesterase (Van Haastert *et al.*, 1982).

Aggregating cells also adhere to each other by forming cell to cell contacts (Muller and Gerisch, 1978). By 10 hrs, an aggregate of up to 10^5 cells is formed (Fig.I, stage 3).

1.2 Morphogenesis and differentiation

At the end of aggregation, morphogenesis begins. A tip is formed at the top of the cell mound and this elongates to form a structure known as the first finger (Fig.I, stage 4). By this stage in development, the aggregate has become a discrete entity, surrounded by an extracellular matrix composed of protein and cellulose (Vardy *et al.*, 1986). Under the conditions of light, humidity and ionic strength prevailing in the forest floor, the first finger collapses to form a migrating slug (Fig.I, stage 5). The sensitivity of

this slug to light and temperature gradients directs it to the surface where it culminates to form a mature fruiting body (Fig. I, stage 8). Under the laboratory conditions however, the migrating slug stage can be brief. Thus the entire developmental process is accomplished within 24-28 hours.

The anterior one fifth of the slug is composed of cells which will under normal conditions form stalk cells. The posterior four fifths are predominantly composed of prespore cells, but approximately 10-15% of cells in this region display the morphological and biochemical characteristics of prestalk cells and are termed anterior-like cells (Sternfeld and David, 1981; Devine and Loomis, 1985). However, the commitment to anterior and posterior cell types is not irreversible in this stage. Culmination begins when cells at the first finger or slug form a compact structure that resembles a mexican hat (Fig.I, stage 6). A tube of cellulose is formed by the central apical cells, along which the cells of the tip migrate, transversing the hat (Fig.I, stage 7). In the process, they differentiate into stalk cells. This movement of cells through the middle of the structure pulls the remaining cells to the top of the stalk where they differentiate into spores.

An intriguing aspect of the multicellular stages of development is the role of the tip. The tip is thought to play the role of an embryonic organizer; removal of the tip at any time results in the cessation of migration and

structure pulls the remaining cells to the top of the stalk where they differentiate into spores.

An intriguing aspect of the multicellular stages of development is the role of the tip. The tip is thought to play the role of an embryonic organizer; removal of the tip at any time results in the cessation of migration and morphogenesis until a new tip is formed from the remaining cells (Raper, 1940; Sampson, 1976). If a slug is cut into thirds and then allowed to migrate, each third will form a normal, though small, fruiting body. The anterior third, however, if prevented to migrate, forms a fruiting body that is composed almost entirely of stalk cells. It has been suggested that after a slug is cut, the tip must undergo reorganization to direct the differentiation of smaller groups of cells. Evidence demonstrating the organizational function of the tip, results from experiments where severed tips were grafted onto the sides of host slugs (Raper 1940, Rubin and Robertson, 1975). Such a procedure results in the formation of several smaller slugs, each led by a grafted tip, a phenomenon known as tip dominance. The organizational properties of the tip are suggested to result from its ability to release periodic cAMP signals (Schaap and Wang, 1986). This oscillatory signalling propagated from the tip, directs both cell movement in the slug and tip dominance.

The two cell types that are found in the slug -prespore cells and prestalk cells- demonstrate morphogenetic and biochemical differences. Also, they are separable by centrifugation in percol gradients (Tsang and Bradbury, 1981; Ratner and Broth, 1983). Prespore cells can readily be identified by their characteristic morphology -the presence of prespore vesicles (PVS) that contain spore coat material (Muller and Hohl, 1973; Devine *et al.*, 1983). Also characteristic are immunological and mRNA markers particular to spore cell precursors: UDP-galactose polysacharide transferase, the enzyme necessary for production of the coat material is one of the proteins unique to this stage of development and this type of cell (Ratner and Broth, 1983; Morrissey *et al.*, 1984). Prestalk cells are characterized by relatively high levels of cell surface cAMP-phosphodiesterase (Tsang and Bradbury, 1981; Mee *et al.*, 1986). In addition prestalk cells contain an isozyme of acid phosphatase that is modified in a prestalk-specific manner (Oohata, 1983). Another gene specifically expressed in prestalk cells encodes extracellular matrix protein (Jermyn *et al.*, 1987; McRobbie *et al.*, 1988).

1.3 Pattern formation and morphogenesis.

Several experimenters agree that cell-fate determination in *Dictyostelium discoideum* is traced at that

point in the cell-cycle of vegetative amoebae, where cells are starved and challenged to differentiate (McDonald and Durston, 1984; Schaap and Wang, 1986; Gomer and Firtel, 1987; Maeda *et al.*, 1989). Cells at the early portion of the cell cycle or cells starved for longer periods of time (aged cells) will sort to the anterior region of the slug and differentiate into stalk cells. It is thought (Weijer et al., 1984) that the metabolic status of either "aged" cells or cells in the earlier stages of the cell cycle is the determining factor of their developmental fate: energy reserves in starved cells or cells in the early stages of the cell cycle are lower than those of cells at the later phases of the cell cycle or cells starved for shorter periods. Specifically in experiments using cells synchronized in the cell cycle it was shown that cells from the early stages of the cell cycle formed slugs with disproportionately larger prestalk regions. In separate experiments (Gomer and Firtel, 1987) it was shown that cells from the early cell cycle stages, when induced to develop as single cells, preferentially expressed a prestalk specific antigen. Moreover it has been shown (Krefft and Weijer, 1989) that cells in the early stages of the cell cycle preferentially express a prestalk specific cell-surface antigen and that these cells sort to the prestalk region of slugs.

Specific properties of the anterior and posterior regions in the slug are responsible for the differentiation of their respective cell populations and for pattern formation. It is thought that differences in the level of diffusible factors along the axis of the slug provide positional information to which cells respond by differentiation into either stalk or spore cells. As mentioned above, oscillatory cAMP signals produced in the tip, and transmitted through the slug via a cAMP relay mechanism, are responsible for cell movement in the slug. The anterior prestalk cells are more chemotactically responsive than the posterior prestalk cells, although the levels of cAMP binding sites are equivalent among these two cell populations. Prestalk cells also have higher levels of two membrane-associated enzymes, cAMP-phosphodiesterase and 5'-nucleotidase, which together convert cAMP to adenosine (Armant and Rutherford, 1979; Armant *et al.*, 1980; Tsang and Bradbury 1981; Mee *et al.*, 1986). It has been proposed that the combination of oscillatory cAMP signals produced by the tip, and the preferential breakdown of cAMP to adenosine by the prestalk cells produces a) a cAMP gradient that increases from the anterior to the posterior of the slug, and b) an adenosine gradient that decreases from the anterior to the posterior (Wang *et al.*, 1988).

At the molecular level, cAMP is involved in the

regulation of post-aggregative gene expression, particularly that of prespore markers, and its effects are mediated through a protein with the specificity of the cAMP cell surface receptor (Schaap and Van Driel, 1985; Oyama and Blumberg, 1986; Haribabu and Dottin, 1986;). In cells that have been rendered aggregation competent and subsequently shaken in suspension, prespore-specific gene expression is entirely dependent on exogenous cAMP (Schaap and Van Driel, 1985). Additionally, cells shaken in suspension are shown to express prespore markers much more rapidly in the presence of cAMP (Kay, 1979). When migrating slugs are disaggregated and the cells rapidly shaken, the rate of transcription of many genes is reduced, and their cognate mRNAs quickly disappear (Mangiarotti *et al.*, 1981; Chung *et al.*, 1981; Landfear *et al.*, 1982). Prevalent among these genes are prespore markers (Barklis and Lodish, 1983) and for many of them, the presence of cAMP acts to maintain and stabilize their cognate transcripts (Barklis and Lodish, 1983; Mehdy *et al.*, 1983; Oyama and Blumberg, 1986). Specifically, gradient purified prespore cells, after incubation in a simple suspension medium, lost prespore specific markers unless cAMP was added. If cAMP phosphodiesterase was also added, the stabilization of the prespore markers by cAMP was eliminated (Weijer and Durston, 1985).

Adenosine is amongst the breakdown products of cAMP, and is proposed to function as an antagonistic signal to cAMP, thus providing an intrinsic negative feedback loop in cAMP regulated processes. Adenosine appears to function as an antagonist to the cAMP dependent stimulation of prespore differentiation and to cAMP's role in tip dominance. During aggregation, the number of cAMP signalling centers is reduced in the presence of exogenous adenosine (Newell and Ross, 1982) while the size of aggregation territories increases. In late aggregation, adenosine acts as an inhibitor to the binding of cAMP to its cell-surface receptor thus limiting cAMP signal relay (Van Haastert, 1983). The previously mentioned phenomenon of tip dominance is also subjected to the antagonistic effects of adenosine over cAMP. The competitiveness of the transplanted tip in disrupting the development of the host and causing the formation of smaller slugs increases with its distance from the primary tip (Durston, 1976; Kopachik, 1983). Also, the size of the slug is a reflection of the balance between tip dominance and tip inhibition. Overly large slugs will not be sustained, but they will break into smaller slugs due to the formation of a secondary tip; reduction in the adenosine levels in the slug by treatment with adenosine deaminase reduces the size of the slug, while exposure to high levels of exogenous adenosine increases the size of the slug

(Schaap and Wang, 1986). In aggregation competent cells rendered dependent upon exogenous cAMP for induction of prespore markers, adenosine acts as a potent competitive inhibitor (Schaap and Wang, 1986). In intact slugs reduction of the adenosine concentration by enzymatic depletion, results in the appearance of prespore cells within the prestalk zone (Schaap and Wang, 1986). Adenosine was found to inhibit the cAMP-induced stabilization of prespore markers (Weijer and Durston, 1985) and furthermore to inhibit the expression of cAMP-induced prespore mRNAs (Spek *et al.*, 1988). The antagonistic role of adenosine to cAMP-mediated responses is thought to be the result of inhibition by the former of cAMP binding to the cell-surface receptor (Van Haastert, 1983; Newell and Ross, 1982; Theibert and Devreotes, 1984; Schaap and Wang, 1988). Adenosine has been found to promote expression of cAMP-induced genes preferentially found in prestalk cells. In this case, both cAMP and adenosine are required for the induction (Spek *et al.*, 1988).

Differentiation Inducing Factor (DIF) was detected as a diffusible factor which induces isolated cells, incubated at low densities, to form stalk cells (Town *et al.*, 1976). Cells of the strain V12M2 will differentiate into spore and stalk cells in submerged monolayer cultures, only if plated at high densities (Town *et al.*, 1976; Kay *et al.*, 1978).

Under the same conditions, sporogenous mutants of that strain will form spores. If plated at low densities, the sporogenous mutants will form spore cells in the presence of cAMP, but fail to differentiate into stalk cells (Kay, 1982). These experiments demonstrated that a spore differentiation pathway barrier exists, which in the mutants is eliminated or bypassed. Furthermore, these experiments showed that an additional factor, absent from cells plated at low densities, is necessary for stalk cell differentiation, and is supplied by plating cells at high densities. The DIF molecule has been purified and its structure has been determined to be 1-(3,5-dichloro-2,6-dihydroxy-4-methoxy-phenyl)-1-hexone (Kay *et al.*, 1983; Morris *et al.*, 1987). The accumulation of DIF is developmentally regulated; it reaches its peak level between the loose mount stage and the tipped aggregate stage of development (Brookman *et al.*, 1982). A mutant which produces lower levels of DIF (approximately 1-2% of the wild-type levels) does not form the normal percentage of stalk cells. The mutant becomes arrested in the loose mount stage and expresses only prespore markers. This deficiency is largely overcome by DIF addition (Kopachik *et al.*, 1983). DIF is also able to induce sporogenous mutants to differentiate into stalk cells (Gross *et al.*, 1983). In addition DIF suppresses the expression of prespore markers

induced by cAMP; thus DIF seems to function as a cAMP antagonist by acting as an inhibitor somewhere along the intracellular signal transduction pathway (Wang and Schaap 1989). DIF is also shown to induce transcription of prestalk specific genes whose products appear at the tipped aggregate stage when DIF levels are highest. The induction of these transcripts is quick and suggests that induction is mediated by a preexisting receptor or second-messenger pathway directly responsive to DIF.

There is good evidence that ammonia plays an important role in *D. discoideum* development. It can reach millimolar concentrations in the immediate environment of a developing aggregate, as a result of protein and nucleic acid catabolism, and has been shown to be one of the factors which prolongs slug migration (Schindler and Sussman, 1977). A slug that has been subjected to enzymatic removal of ammonia is promptly triggered to form a fruiting body, and prestalk cells are triggered to form mature vacuolated stalk cells (Wang *et al.*, 1989). It has also been proposed that ammonia acts during the initial establishment of the prespore/prestalk pattern to favour prespore cell differentiation (Gross *et al.*, 1983).

1.4 Signal Transduction Pathways in *D. discoideum*

1.4.1 The cAMP receptor

It has been demonstrated that the responses of adenylate cyclase and guanylate cyclase to cAMP induction differ in the rate of excitation, adaptation and de-adaptation, leading to the conclusion that the two responses are regulated via separate pathways. This conclusion is reinforced by the observation that there are two classes of cAMP receptors, which are defined by the rate of dissociation of cAMP (Van Haarstert and De Wit, 1984). The fast dissociating receptors are more abundant and thought to be associated with adenylate cyclase and the signal relay pathway, while the slowly dissociating receptors are less abundant and coupled to guanylate cyclase and the chemotaxis response (Van Haarstert *et al.*, 1986; Janssens and Van Haarstert, 1987). An additional property of the cAMP receptor that supports separate pathways underlying signal relay and chemotaxis is that the two kinetic classes of receptors behave differently during receptor down-regulation. Down-regulation is a process of receptor desensitization (much slower compared to receptor adaptation) that occurs during aggregation. Down-regulation involves loss of cAMP binding sites (Kesbeke and Van Haastert, 1985) at constant cAMP concentrations which overwhelm cAMP phosphodiesterase. In down-regulated cells,

the fast dissociating receptors are reduced by approximately 98%, while the slow dissociating receptors are reduced 10-fold. In SDS-PAGE analysis the cAMP receptor appears as a doublet with molecular weight of 40,000 and 43,000 Daltons (Klein *et al.*, 1985; Klein *et al.*, 1987). The kinetics and dose response of the fast dissociating receptor correlates with the adaptation of adenylate cyclase (Vaughan and Devreotes 1988), suggesting that cAMP induced receptor phosphorylation is an early event in the mechanism of adaptation of adenylate cyclase in the cAMP signal relay. The purified receptor has been used to raise specific antibodies recognising both forms, which allowed the isolation of a cDNA sequence encoding a cell-surface cAMP receptor (Klein *et al.*, 1988). The expression pattern of the receptor encoding gene is consistent with the role of the receptor during aggregation; the respective mRNA is absent from vegetative cells but begins to accumulate in early development, reaches its maximal level during the aggregation stage, and subsequently declines over the next few hours. Northern analysis using the isolated gene as a probe under reduced stringency conditions, reveals transcripts that are cross-reacting to the isolated receptor cDNA sequence. This suggests a family of receptors contained in *D. discoideum*. Analysis of the predicted amino-acid sequence of the isolated receptor-encoding cDNA

sequence suggests that the receptor has seven transmembrane domains and a serine rich carboxyl terminus. These structural motifs are common to G-protein-linked receptors. Ligand occupancy induces phosphorylation of serine residues at the carboxyl terminus (Vaughan and Devreotes, 1989) as is also seen in rhodopsin and β-adrenergic receptor. Supporting evidence for the G-protein association of the *Dictyostelium* cell-surface cAMP receptor comes from biochemical studies. In the presence of guanine nucleotides, cAMP binding by the receptor is inhibited. This observation is commonly diagnostic of G-protein linked receptors. Additionally, both adenylate and guanylate cyclase are stimulated by GTP in membrane preparations (Van Haarstert, 1984; Theibert and Devreotes, 1986; Van Haarstert *et al.*, 1986). Also, GTPase activity in membrane preparations is stimulated by cAMP (Snaar-Jagalska *et al.*, 1988). Evidence at the molecular level of the receptor's association with an heterotrimeric G-protein comes from the isolation of genes that are encoding several G-protein subunits. A sequence encoding a Gβ protein subunit and two genes encoding Gα protein subunits ($G\alpha_1$ and $G\alpha_2$) have been isolated. Their predicted amino-acid sequences are approximately 50% identical to those of their mammalian counterparts, but amino acid similarity increases to virtually 100% when the GTP binding and GTPase activity

domains are compared alone (Pupillo *et al.*,1988; Pupillo *et al.*, 1989). Genetic analysis sheds further light on transmembrane signal transduction via the cAMP receptor. This includes the isolation and characterization of two classes of mutants. A series of mutants designated *Frigid* fail to enter the developmental phase, even when induced with exogenous cAMP (Coukell *et al.*, 1983). In a class of such mutants designated *Frigid A*, guanine nucleotides do not influence cAMP binding. These mutants carry mutations or deletions in the $G\alpha_2$ protein subunit (Kumagai *et al.*, 1989). In a second class of aggregation deficient mutants, designated *synag*, adenylate cyclase and the cAMP signalling response are no longer sensitive to guanine nucleotides, but repeated stimulation with cAMP is still able to induce chemotaxis and gene expression. In this series of mutants there has been no defect identified, in the G-protein, adenylate cyclase or cAMP receptor. The defect in one strain, named *synag* 7, has been partially identified. In this mutant, stimulation of adenylate cyclase by GTP is prevented, but can be complemented by a soluble protein present in wild-type cells (Theibert and Devreotes, 1986). Thus, cAMP stimulated activation of adenylate cyclase requires additional factors apart from the receptor, the G-proteins and the enzymatic components.

1.4.2 Second Messenger Pathways

What are the molecular mechanisms involved in the intracellular propagation of the G-protein transduced signal? How do these intracellular processes link specific receptors to particular responses such as signal relay and chemotaxis? In the signal relay pathway, cAMP appears to play the role of a second messenger. cAMP binding to the fast dissociating class of receptors, transiently activates adenylate cyclase. This activation increases intracellular levels of cAMP and leads to cAMP secretion, mediated by secretory vesicles (Maeda and Gerisch, 1977) and cAMP signal relay. In mammalian and other eukaryotic cells, the effects of intracellular cAMP are mediated by the cAMP - dependent protein kinase or PKA (Krebs, 1989). The enzyme has a R_2C_2 subunit structure where R designates the regulatory subunit and C the catalytic subunit. In *D. discoideum* however, PKA is an RC dimer (Mutzel *et al.*, 1987). The R subunit is able to bind cAMP, upon which it releases the catalytic subunit, which acts to phosphorylate effector molecules. Expression of the regulatory subunit is very low in vegetative cells, but the levels increase when development is initiated and reaches peak levels just after aggregation (De Gunzberg *et al.*, 1986). These fluctuations in PKA levels are compatible with its role in development. Mutants that overexpress the

regulatory subunit of PKA (Simon *et al.*, 1989) do not aggregate and appear to be blocked within 3 hrs after entering development. These results suggest that the activity of the catalytic subunit is inhibited by the overwhelming levels of the PKA regulatory subunit present in the mutants.

The binding of extracellular cAMP to cells promotes the accumulation of other second messengers. The production of IP3, which in turn allows the mobilization of intracellular Ca^{++}, is suggested to promote cGMP synthesis (Europe-Finner and Newell, 1985). It has been suggested (Newell *et al.*, 1988; Berridge, 1989; Nishizuka, 1989) that a phosphatidylinositol pathway analogous to that observed in mammalian cells forms part of the developmental control system in *D. discoideum*, and specifically in the regulation of chemotaxis. On the other hand, activation of guanylate cyclase, cGMP accumulation, the polymerization of actin and the association of myosin with the cytoskeleton (Mato *et al.*, 1977; Wurster *et al.*, 1977; McRobbie and Newell, 1984; Berlot *et al.*, 1985) are cAMP dependent events that are associated with the chemotaxis pathway. Specifically, the addition of IP3 to permeabilized cells results in similar accumulation of cGMP and cytoskeletal actin polymerization (Europe-Finner and Newell, 1987). Treatment of permeabilized cells with Ca^{++} pulses also induces

cytoskeletal actin polymerization and cGMP accumulation (Small *et al.*, 1986). Guanylate cyclase cannot be stimulated by Ca^{++} *in vitro*, (Padh and Breuner, 1984) which suggests that there are a number of steps between IP3 production/Ca^{++} mobilization and guanylate cyclase activation. In addition, in the *stm F* streamer mutants, whose prolonged chemotactic response correlates with elevated cGMP levels (Ross and Newell, 1981; Van Haarstert *et al.*, 1982; Coukell *et al.*, 1983), the pattern of cytoskeletal myosin accumulation resembles that of cGMP accumulation. On the other hand, no alterations with respect to cytoskeletal actin levels are observed (Liu and Newell, 1988). It has been proposed (Newell *et al.*, 1988) that the signal transduction pathway that controls chemotaxis diverges passed the Ca^{++} mobilization step; events downstream of Ca^{++} mobilization, including actin polymerization, lead to pseudopodium formation, while guanylate cyclase activation, cGMP accumulation and cytoskeletal myosin accumulation lead to cell elongation. The two processes result in the chemotactic response.

The *D. discoideum* ras protein is encoded by Ddras, a gene with high degree of similarity with the human ras gene, and linked to the phosphatidylinositol pathway and chemotaxis. Site-directed mutagenesis of Ddras that results in amino acid substitutions found in the oncogenic form of

the human ras gene, yields mutants where the chemotactic response is hindered. At the molecular level, this mutation allows the cAMP-induced desensitization of guanylate cyclase. However the adenylate cyclase response is not altered in these mutants (Reymond *et al.*, 1986; Van Haarstert *et al.*, 1987; Europe-Finner *et al.*, 1988).

1.4.3 cAMP and gene expression

The developmental phase of *D. discoideum* consists of a series of morphogenetic and biochemical changes directed by a complex gene expression regulatory system. Developmentally controlled gene expression and the control mechanisms that underlie its pattern are fundamental aspects of *D. discoideum* development and are rigorously investigated.

A number of developmentally regulated genes underlie the accomplishment of chemotaxis and aggregation and consequently morphogenesis in *D. discoideum* cells. These gene activities are not found in vegetative cells but can be detected by 2-3 hrs after the onset of development and are maximally expressed during the peak of aggregation. Proteins such as the cAMP receptor, Gα2, contact sites A (cell adhesion molecules), and D2 serine esterase are products of cAMP pulse-induced genes, which are essential for aggregation (Gerisch, 1987; Kimmel 1987; Klein *et al.*, 1988). Under repeated stimulation with low levels of cAMP,

these genes are precociously induced, whereas continuous stimulation - which results in adaptation of the cell surface cAMP receptor- lowers or inhibits expression. It is therefore suggested that the oscillations of the receptor between the adapted and de-adapted states is an absolute requirement for the expression of these genes. Studies done on *Synag* and *Frigid A* mutants, suggest that the control of expression of these genes is coupled to the chemotactic pathway. *Synag* strains show a normal expression pattern of these genes, suggesting that activation of adenylate cyclase is not a component in their regulation. However, in the *Frigid A* strain, these genes are not expressed and can not be induced with repeated stimulation with cAMP pulses. Proteins such as the cAMP receptor, $G\alpha_2$ and serine esterase are present in pre-aggregative stages at basal levels since they are necessary for the generation of the signal transduction pathway that potentiates their developmentally regulated induction. In addition to the pulse-induced genes, there are two classes of pulse-repressed genes whose expression during the pre-aggregation stage is down-regulated by repeated stimulation with cAMP. One class of these genes is represented by K5, whose expression is maximal after ~2.5 hours after the onset of development, and is subsequently repressed during aggregation (Mann *et al.*, 1988; Mann and Firtel, 1987). The expression of this class

of genes in *Synag* and *Frigid A* mutants exhibits a pattern that is opposite to that of pulse-induced genes. In *Synag* strains, they are repressed in the presence of exogenous cAMP pulses, but are induced and continue to be expressed in the absence of cAMP. In *Frigid A* strains they are overexpressed and are found to be insensitive to cAMP pulses, suggesting that the regulation of these genes is linked to the chemotactic pathway. A second class of cAMP pulse-repressed genes includes M4-1, and appears to require intracellular cAMP for their regulation (Kimmel and Saxe, 1986; Kimmel, 1987). M4-1 is expressed during vegetative growth, is repressed during aggregation and is shown to be precauciously repressed by cAMP pulses. In contrast to K5, it is not repressed in *Synag* mutants, suggesting that in this case the regulation of expression may have a link to the signal relay pathway and may be mediated by cAMP fast-dissociating receptors and cAMP dependent protein kinase. The genes encoding extracellular PDE and the inhibitor of extracellular PDE (PDI) are both cAMP regulated (Coukell *et al.*, 1984; Kessin *et al.*, 1979). The PDI protein inhibits PDE activity by binding PDI and raising the Km of PDE for cAMP. *In vivo*, this complex is probably irreversible. While PDE expression is induced either by pulses of cAMP or by continuously elevated cAMP levels, PDI shows the opposite regulation. The regulation of PDE and PDI in opposite

directions is thought to provide the organism with an intrinsic negative feedback loop. PDE expression correlates with high cAMP levels, allowing effective removal of the ligand and adaptation of the processes of signal relay and chemotaxis. At lower levels of cAMP, PDE remains associated with the membrane, while at elevated cAMP concentrations it is secreted into the surroundings of the cells.

The prespore and prestalk genes whose expression is regulated by cAMP are referred to as the "late" regulated genes, as opposed to the class of "early" genes discussed above. Expression of both prespore and prestalk genes has been shown to be induced in cells stimulated with pharmacological agents such as caffeine and cAMP analogs, suggesting that control over their expression is mediated through the cell-surface cAMP receptor (Mehdy *et al.*, 1983; Schaap *et al.*, 1986; Kimmel,1987). Similar to the early pulse-induced genes,the regulation of these "late" genes, is not associated with increased intracellular cAMP levels. In permeabilized cells that are treated with DAG and 1,4,5-IP3 (Ginsburg and Kimmel, 1989) both prespore and prestalk genes are induced at high levels. This suggests that protein kinase C and/or Ca^{++} may be involved in their regulation. Contrasting the pattern of expression of pulse-induced genes, "late" genes are induced by continuous elevated cAMP levels, which result in the adaptation of the different

classes of cell-surface cAMP binding receptors. This finding, combined with the observation (Klein *et al.*, 1987) that the early class of receptors are present at very low levels in this stage in development, leads to the proposal that another class of receptors is expressed later in development and mediate control of cAMP over late gene expression. Presumably, either $G\alpha_2$ or a novel $G\alpha$ expressed later in development, mediate these effects.

1.5 Mechanisms of control of gene transcription by cAMP.

In bacteria, cAMP binds to catabolite repressor protein (CAP). The complex directly interacts with specific DNA sequences, leading to the transcription of several catabolite-sensitive operons (Peterkofsky,1976; Pastan and Adhya,1976).

In eukaryotes , several cAMP-induced proteins (growth hormone, somatostatin, prolactin, Fos) are known to be regulated at the level of transcription (Maurer, 1981; Hashimoto *et al.*, 1984; Montminy *et al.*, 1984; Kruijer *et al.*, 1984). So far cAMP-dependent protein kinase is the only recognized mediator of cAMP action in eukaryotic cells and it has in fact been suggested that all of the intracellular effects of cAMP are mediated via cAMP-dependent protein kinase (Kuo and Greengard, 1969). At the nuclear level, the mechanism of control of gene

transcription by cAMP and cAMP-dependent protein kinase remains to be elucidated.

Comparison of sequences near the promoters of genes regulated by cAMP reveals a highly conserved palindrome 5'-TGACGTCA-3' called CRE or cAMP-responsive element. (Montminy *et al.*, 1986). However the palindrome is not found near the promoters of several eukaryotic genes known to be responsive to cAMP (e.g. growth hormone, prolactin). Therefore, there may be multiple mechanisms by which cAMP regulates gene expression in eukaryotes. The cAMP responsive element has been proposed to bind a specific transcription factor whose level or binding affinity is regulated by cAMP (Montminy *et al.*, 1986). Mammalian CRE binding activity (CREB) has been cloned from human and rat and their structures determined from sequence analysis (Hoeffer *et al.*, 1988; Gonzalez *et al.*, 1989). The predicted CREB structure contains a bipartite region of primary sequence similarity with transcriptional factors of the bZIP family (C/EBP, Jun, Fos; Mitchell and Tjian, 1989). This region consists of a stretch of ~30 amino acids with a substantial net basic charge, immediately followed by a region containing four leucine residues positioned at intervals of several amino acids. The latter segment is termed the " leucine zipper" and is required for dimerization and for DNA binding. The basic region next to

the leucine zipper is necessary for DNA binding but not for dimerization (Landschultz *et al.*, 1988). The leucine zipper has been proposed to function as a dimerization interface between amphipathic α-helices containing the motif, thus enabling them to assume a coiled-coil conformation (O'Shea *et al.*, 1989). *In vitro* studies have shown that CREB dimerization, DNA binding, and transcription activity are promoted by phosphorylation of the protein (Yamamoto *et al.*, 1988; Gonzalez *et al.*, 1989; Hoeffler *et al.*, 1988). Thus, in mammalian cells, one mechanism by which cAMP induction of gene expression may occur is by direct phosphorylation (mediated by cAMP-dependent protein kinase A) of a *trans*-acting factor interacting with CRE (Yamamoto *et al.*, 1988). However, an effect of cAMP mediated by the regulatory subunits acting at the gene level must also be considered (Constantinou *et al.*, 1985). The regulatory subunits dissociate from the catalytic subunits and appear to be transported into the nucleus. The regulatory subunits of the type II isoform of cyclic AMP-dependent protein kinase have been shown to have topoisomerase activity (Constantinou *et al.*, 1985). These may regulate gene expression by altering chromatin structure near or within cAMP responsive genes.

In *Dictyostelium*, cAMP induction of prestalk and other classes of genes is believed to be regulated by a signal-

transduction pathway that involves surface-receptor specific proteins and possibly activation of phospholipase C (Mann *et al.*, 1988). Because these are developmentally regulated events, one might expect that components of the signal transduction process such as *trans*-acting factors acting downstream of the effector molecule are regulated developmentally. It has been proposed that the cAMP-induced signal-transduction system could activate such factors either by post-translational modifications or by inducing the synthesis of the factor (Hjorth *et al.*, 1990). A nuclear activity has been identified that binds specifically to a GT-rich sequence (or G-box) that has been shown by deletion analysis to be required for the cAMP and developmentally induced expression of the prestalk gene pst-cathepsin (Hjorth *et al.*, 1990). The binding activity of G-box binding factor (GBF) is developmentally regulated and inducible by cAMP. Different models were proposed to explain increased binding-specificity of GBF to GBRE. GBF may bind alone or in association with an auxiliary factor to GBRE containing promoter regions. In the latter case, cAMP activation of the auxiliary factor may facilitate GBF binding to its respective sequence. This cAMP-dependent interaction results in the induction of cAMP responsive genes such as pst-cathepsin. Alternatively, cAMP activation of a transcription factor may result in the activation of

the gene encoding GBF. At higher levels of factor, stable complexes are formed with the G-box region of cAMP regulated genes such as pst-cathepsin, resulting in increased transcription of the gene. Cloning and expression of the genes encoding GBF and GBF related factors will shed light on these issues.

2. MATERIALS AND METHODS

2.1 Isolation and characterization of cDNAs encoding a putative Leucine Zipper *Dictyostelium discoideum* protein

Previous work in our laboratory identified a partial sequence that encodes a "leucine zipper" motif. To obtain additional sequence information, the partial sequence was used to screen a λ ZAPII cDNA library. The cDNA library was obtained from Dr. J. G. William of ICRF Clare Hall, United Kingdom, and was constructed with mRNA from vegetative *D. discoideum* cells. Also screened was a library of *D. discoideum* genomic DNA fragments (a gift from Dr. H. Ennis, Roche Institute of Molecular Biology, Nutley New Jersey).

2.1.1 Titering of the libraries

First, the concentration of plaque forming units in each of the phage libraries (the λ ZAP genomic library and the λ ZAPII cDNA library) was determined. Serial dilutions were prepared (10^{-1}-10^{-5}) of the respective phage suspensions and appropriate volumes (1-50 µl) were incubated with host bacteria (strain: XL1 Blue at OD_{600}=1) in a shaker/incubator at 37^{0}C for 30 min for infection to occur. At the end of incubation 3.5 ml of warm top agar (contains 0.7% Bacto-agar) was added to the test tubes containing the cells and phage. This was mixed by vortexing gently and poured onto LB agar

Petrie plates (1% NaCl, 1% tryptone, 0.5% yeast extract 15% agar). The plates were left at room temperature on a levelled surface, until the soft agar solidified. Plates were subsequently incubated at 37°C for 5-7 hours until plaques, approximately 0.5mm in diameter, appeared on the bacterial lawn.

2.1.2 Plating the libraries

Fifty milliliters of LB medium (1% NaCl, 1% tryptone, 0.5% yeast extract) containing 0.2% maltose and 10 mM $MgSO_4$ were inoculated with a small aliquot of an overnight culture of XL1 Blue cells. This was incubated with shaking at 37°C until an OD_{600} of 0.7 was reached. An aliquot of phage suspension containing approximately 5 x 10^6 pfu was mixed with an aliquot of XL1 Blue cell-suspension in a final volume of ~ 1 ml. The mixture was prepared in a sterile 10 ml tube and incubated at 37°C for 25 min with mild shaking, for infection to occur. At the end of incubation, 7.5 ml warm LB top agar was added to each tube, mixed gently and poured an LB agar plate. Plates were left at room temperature until the top agar solidified and then incubated at 37°C for 5-7 hours. At the end of incubation, the plates were wrapped with parafilm and placed in 4°C between 1 and 2 hours for the top agar to harden before proceeding with the plaque lifts.

2.1.3 Plaque lifts

Nitrocellulose filters (Schleischer & Schuell BA85) were laid on each plate for 2 and 4 min on original and duplicate filters, respectively (Maniatis et al., 1982). An 18-gauge needle was pushed through the filter and agar at 3 asymmetrical points, so that the filter-plate alignment could later be reproduced. The filters were submerged in 1.5 M NaCl, 0.5 M NaOH for 2 min, allowing phage lysis and phage DNA denaturation to take place. Subsequently, they were neutralized by immersion in 1.5 M NaCl, 0.5 M Tris-Cl (pH 8.0) for 5 min and finally rinsed in 0.2 M Tris-Cl (pH 7.5), 2 X SSC for 20 sec. Excess liquid was drained by blotting the filters on a clean 3MM paper. While still damp, membranes were placed in a UV Stratalinker 1000 (Stratagene) and irradiated for 1 min to fix the DNA onto the membranes. Finally, filters were wrapped in tin foil between 3MM papers and stored in the dark for up to a week before proceeding with hybridizations.

2.1.4 Prehybridizations

Nitrocellulose membranes were placed in prewarmed hybridization bottles with ~30 ml 2 X SSC. The bottles were rotated in a Hybaid oven (BIO/CAN SCIENTIFIC) and the whole system stabilized at 37°C for 30 min. Then, the solution was replaced with 5-10 ml of hybridization solution [5 X Denhardt's solution (2% BSA, 2% Ficoll, 2% PVP) 50% formamide, 5 X SSC, 0.5% SDS, 200 µg/ml salmon sperm DNA], and the

filters were prehybridized for 2-4 hours at 37°C.

2.1.5 Probe preparation

Probe was prepared by the random primers method (Bethesda Research Laboratories Random Primer Kit) using the *Bam*HI/*Eco*RI 1.8 kb pLZG1.8 fragment as template and [α-^{32}P] dCTP (ICN Biomedicals Canada Ltd., 3000 Ci/mmole) as labelled substrate. Labelled DNA were purified from unincorporated radioactivity by gel filtration through a G-50 Sephadex column. Probes with specific activity between 10^7-10^9 cpm/μg were considered to have satisfactory levels of incorporated radioactivity.

2.1.6 Hybridization

The purified probe was boiled for 10 min, chilled on ice and added to 5-10 ml of pre-hybridization solution. Hybridization was carried out at 37°C over 16-20 hours at maximum rotation in the Hybaid hybridization oven.

2.1.7 Washing conditions

Washing was done in a solution containing 2 X SSC, 0.1% SDS, for 15-20 min at room temperature, and 2 X 20 min at 68°C. Filters were wrapped in Saran Wrap, and exposed to X-ray film (Kodak) at -70°C with intensifying screens, between 72 and 96 hours.

2.1.8 Film processing

Films were developed in Kodak developer manually for 2 min, washed briefly and fixed in Kodak fixative for 2 min. Fixed films were thoroughly washed in water and air dried before examination.

2.1.9 Selecting positives

Positive signals found on duplicate filters were aligned on the plate. The agar containing the positive clones was cored using the back end of a Pasteur pipette. Agar plugs were transferred to a microfuge tube containing SM buffer (0.58 % NaCl, 0.2 % $MgSO_4$ $7H_20$, 2 % gelatin, 1 M Tris-Cl pH 7.5) and a drop of chloroform, and placed at room temperature for 4 hours for the phage to diffuse out of the agar. Subsequently, these mixtures were vortexed and stored at 4^0C.

2.1.10 Purification of potential positives

Phage from the positives were purified by screening with the labelled probe for a second time. Phage were diluted to approximately 100 plaques/plate and hybridization was carried out as explained above. Positive plagues were harvested, placed in 500 µl of SM buffer at room temperature for 4 hours and stored at 4^0C.

2.1.11 *In vivo* excision

In vivo excision was done on the independent positives

according to the procedure of Stratagene. In brief, a sample of the phage stock in the SM buffer was mixed with an equal volume of overnight culture of XL1 Blue cells, and incubated in a test tube in the presence of M13 helper phage in a shaker-incubator at 37°C for 15 min. The mixture was diluted with fresh medium. Infection and propagation were allowed to occur during 4-6 hours of further incubation at 37°C. During the latter incubation, the pBluescript vector and inserted sequences were excised from the λ ZAPII construct. Packaging of the resulting recombinant "phagemid" into M13 phage particles and secretion of the latter occurred during this time also. Bacteria were heat-killed by immersing the tubes in 70°C water for 20 min, then removed by spinning in a table-top centrifuge at 1000 g for 5 min. The supernatant, containing the Bluescript phagemid in filamentous phage particles, was stored at 4°C. A sample of this supernatant was mixed with an equal volume of overnight culture of XL1 Blue cells and incubated in a test-tube at 37°C for 15 min for infection to occur. Aliquots between 1-100 µl of a 10^{-1} and a 10^{-3} dilution of the mixture, were plated on LB agar plates and grown under ampicillin selection (100 µg/µl ampicillin) The infected cells gave rise to colonies, harbouring the products of the excision process, under the control of the ColE1 origin of replication. Single-stranded or double-stranded preparations of these plasmids were used for sequence analysis of the *D. discoideum* DNA inserts.

2.1.12 Plasmids

The plasmids used for sequence determination and subcloning in this study are plasmids pLZc1.0, pLZc1.7 and pLZc2.0, isolated from the cDNA library, and plasmid pLZG1.8, (the subclone of pLZG4.0) isolated from a genomic library. The relationship between the different inserts is shown in Figure II. In addition, the maps of the different clones are shown in Figure III, to indicate the various cloning sites.

2.1.13 Deletions and Sequencing

Nested sets of deletion mutants that lack progressively more sequence from each end of the 2.0 kb insert of the pLZc2.0 cDNA clone, were generated by *Exonuclease* III digestion (Maniatis *et al.*, 1982). The pLZc2.0 plasmid was linearized by restriction enzyme digestion at the polylinker sites *Sst*I or *Apa*I prior to *Exo*III treatment. DNA samples were taken at 30 sec intervals over a 5 min period. Overhungs were removed by S1 nuclease treatment. Recessed ends were repaired in the presence of dNTP's and the Klenow fragment of *E. coli* polymerase I, and the deleted molecules were recircularized with T4 DNA ligase.

Random derivatives of the pLZG1.8 genomic subclone were produced by digestion with limiting concentrations of pancreatic DNAase, with cleavage occurring randomly, once per plasmid molecule on average (Maniatis *et al.*, 1982). The linearized molecules were cleaved at polylinker restriction

Fig. II Restriction maps of clones used in characterization of a *D.discoideum* gene that encodes a leucine zipper motif. Restriction sites are, B: *Bam*HI, C: *Cla*I, E: *Eco*RI, G: *Bgl*II, H: *Hind*III, K: *Kpn*I, Q: *Sca*I, S: *Sst*I, V: *Eco*RV, X: *Xba*I.

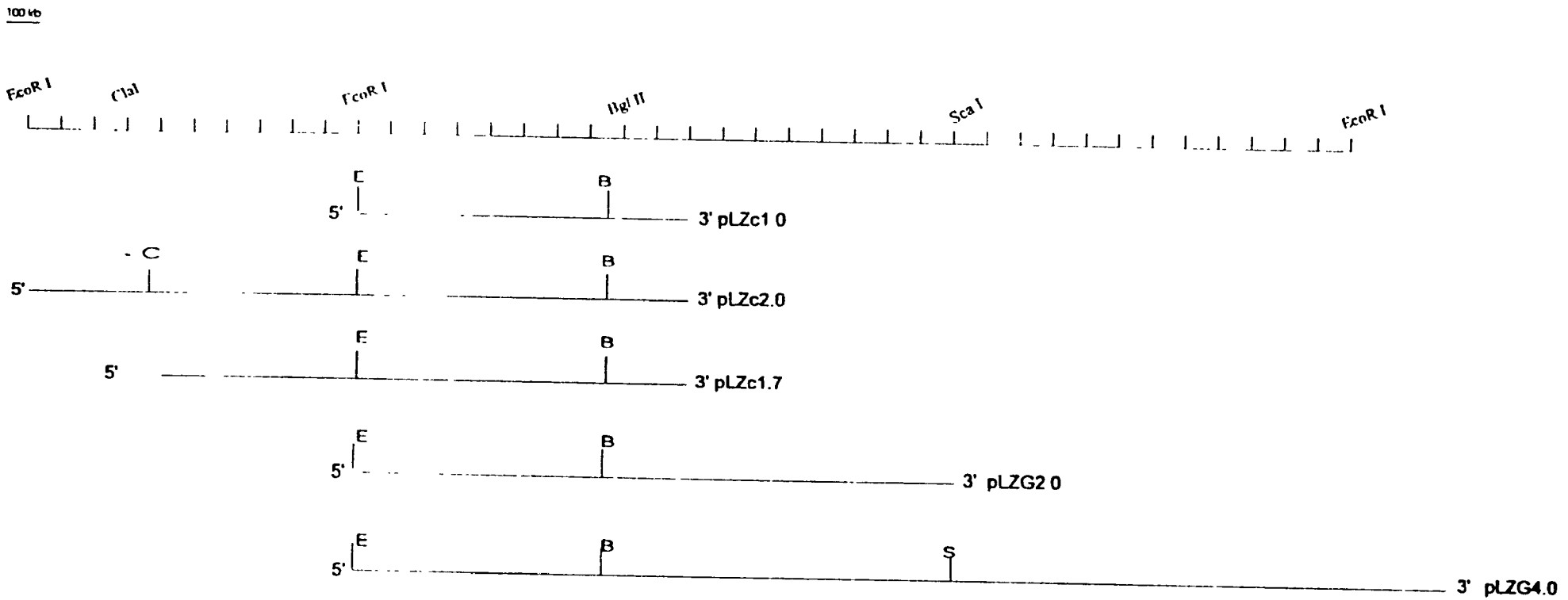
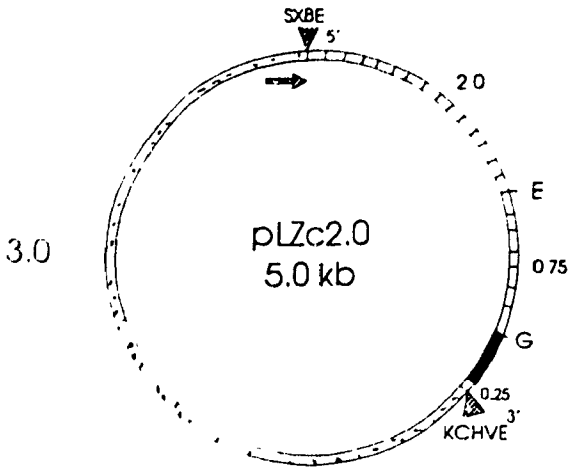
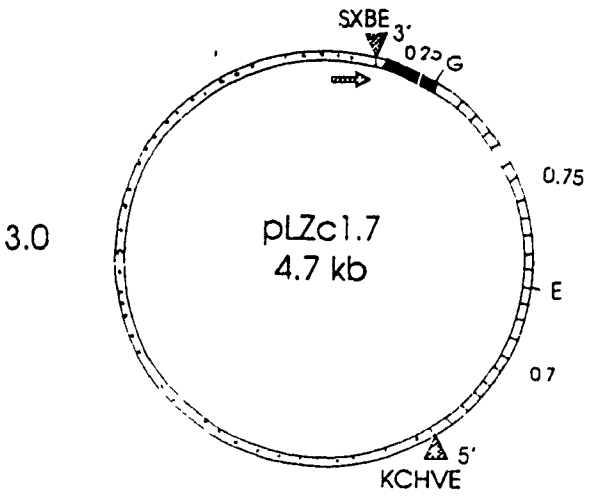
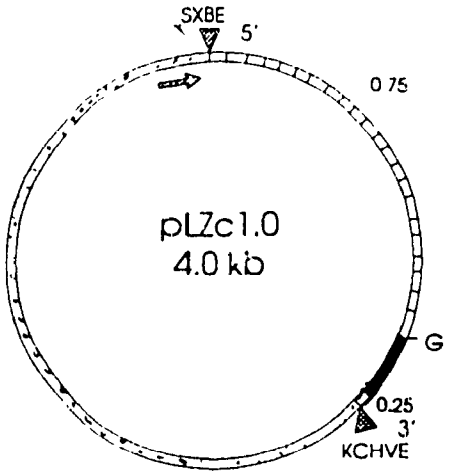
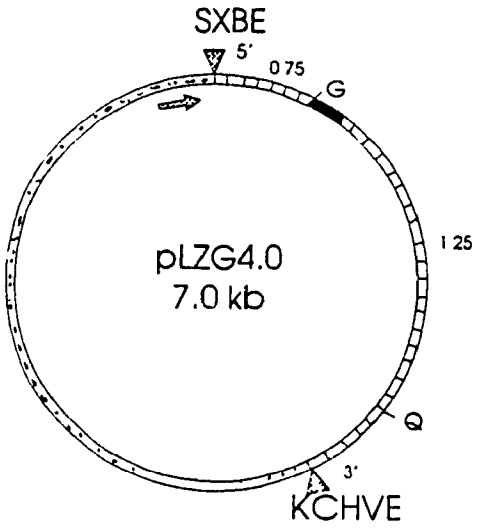

100 kb
EcoR I
Cla I
EcoR I
Hgl II
Sca I
EcoR I
E
B
5'
3' pLZc1 0
C
E
B
5'
3' pLZc2.0
E
B
5'
3' pLZc1.7
E
B
5'
3' pLZG2 0
E
B
S
5'
3' pLZG4.0

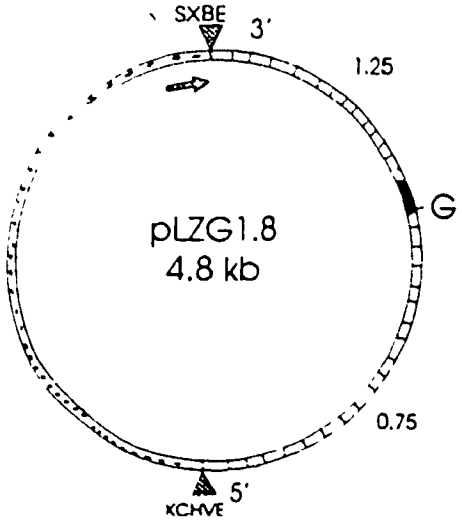
Fig. III Plasmid maps of clones used in analysis of a *D. discoideum* leucine zipper gene. Designated names include a G for genomic clone and a c for cDNA clones. Descriptions are given in text. Arrow indicates the universal primer binding site on pBluescript vector.

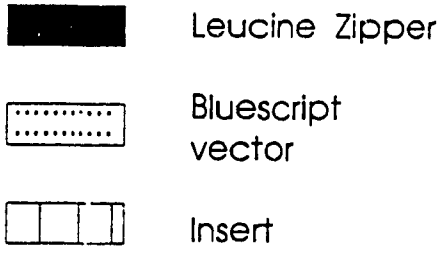

SXBE
5'
2.0
E
0.75
G
0.25
3'
KCHVE
3.0
pLZc2.0
5.0 kb


SXBE
3'
0.25
G
0.75
E
0.7
5'
KCHVE
3.0
pLZc1.7
4.7 kb


SXBE
5'
0.75
G
0.25
3'
KCHVE
pLZc1.0
4.0 kb


SXBE
5'
0.75
G
1.25
Q
3'
KCHVE
pLZG4.0
7.0 kb


SXBE
3'
1.25
G
0.75
5'
KCHVE
pLZG1.8
4.8 kb


Leucine Zipper
Bluescript vector
Insert

sites -*Sst*I or *Bam*HI- flanking the target DNA, yielding mutants of the original clone with a variety of sequences at either end of the insert missing. Repair of recessed ends took place in the presence of a buffer containing all four dNTP's with addition of both T4 DNA polymerase and the Klenow fragment of *E. coli* DNA polymerase. The deleted molecules were recircularized in the presence of T4 DNA ligase.

Sequencing of the deleted clones was done on single-stranded DNA propagated in and extracted from M13 phage, or on double-stranded DNA preparations. Sequencing was done by the dideoxy-termination method (Sanger *et al.*, 1977) using a T7 sequencing kit (Pharmacia) and following instructions therein. The sequencing ladder was labelled with α-^{35}S dATP (Amersham, 1000 Ci/mmole). DNA sequencing reactions were analyzed by electrophoresis through 6% polyacrylamide gels containing 8M urea. DNA sequence analysis was done using programs in PCGENE software.

2.1.14 Single-stranded DNA preparation

DNA from positive phage was isolated according to Maniatis *et al.*, 1982. Transformants (strain DH5αF') harbouring the deleted derivatives of the parent clone of interest were grown for 2-3 hours in liquid medium under ampicillin selection in a 37°C shaker-incubator. M13 phage was introduced in the growing bacterial cultures (OD_{600}=0.3). Phage infection and propagation were allowed to take place over 1-2 hours. Small

samples of the infected cultures were transferred into fresh medium and phage propagation was allowed to occur in a 37°C shaker-incubator for 18-20 hours. The phage particles were purified from bacterial cells by centrifugation at 9,000 g for 15 min. Phage particles were precipitated at 4°C in the presence of 5 % PEG/0.5 % NaCl and collected by centrifugation at 10,000 g for 20 min. Phage DNA was isolated by a series of phenol/chloroform extractions. One microgram of this DNA preparation was used in sequencing reactions.

2.1.15 Double-stranded DNA preparation

Transformants used for sequence analysis were grown as described in the section above. The cells were lysed by alkaline treatment, the cell extracts incubated in the presence of pancreatic RNAase, and subjected to two rounds of chloroform extractions. The DNA was precipitated with isopropanol and resuspended in dH_2O. Molecules in the supercoiled conformation were precipitated on ice in the presence of 7.5% PEG_{8000} according to the procedure of Applied Biosystems. Aliquots of 2-3 µg of these double-stranded DNA preparations were used in sequencing reactions.

2.2 Isolation of *D. discoideum* proteins

D. discoideum cells of the AX2 strain were grown in two membered cultures with *E. aerogenes* on SM agar plates [0.1% $MgSO_4$.7H2O [0.1% Na_2HPO_4, 0.22% KH_2PO_4, 1% glucose, 1%

bacteriological peptone (OXOID), 0.1% yeast extract, 1.5% agar] and were incubated at 22°C for 42 hours, until clearing on the plates was observed. Amoeba grown on bacteria were harvested in KK2 buffer (0.22% KH_2PO_4, 0.07% K_2HPO_4) and collected by centrifugation in a table top centrifuge at 4000 g for 3 min. Two to three washes followed to remove any contaminating bacteria. Amoebae were resuspended in KK2 at 10^7 cells/ml and left to develop for 16 hours at 22°C with moderate shaking.

2.2.1 Extraction of nuclear proteins from *D. discoideum*

Cells were washed in 50 ml ice-cold 0.2% NaCl and resuspended in 10 ml NIB [0.5 M $MgOAc(4H_2O)$, 10% sucrose, 2% Triton X 100] at 4°C at an approximate concentration of 10^8 cells/ml (Cocucci and Susman, 1970). A 500 µl sample (1 mg of protein) was removed, and treated as "total protein". The remaining NIB cell suspension was vortexed vigorously on ice for 1 min to lyse the cells, and centrifuged at 4°C for 5 min at 2000 g. The supernatant was centrifuged for 5 min at 4000 g (4°C) to remove any impurities and considered as the cytosolic fraction. The pellet, containing the nuclei and cell debris was resuspended in NIB, and centrifuged at 400 g to differentially pellet cell debris or unbroken cells. The supernatant was spun (4°C) at 2000 g for 5 min to collect the nuclei. Samples from different fractions were resuspended in SDS buffer, heated at 90°C for 5 min and stored at -20°C.

2.2.2 Protein determination

Determination of protein concentration in different fractions was estimated by the Esen assay (Esen 1978). In brief, different dilutions of protein were prepared in equivalent volumes and spotted onto a filter paper. The filter paper was stained in Coomassie blue stain (0.1% Coomassie blue R-250 in 25% isopropanol 10% acetic acid), rinsed in water and air dried. Protein was eluted from individual samples by incubating the respective area of the filter in 1% SDS with moderate shaking. Optical absorbance measurements at OD_{600} were taken and plotted onto a standard curve of OD_{600} vs. protein concentration values obtained from a series of BSA dilutions.

To visualize proteins separated by SDS polyacrylamide gel electrophoresis, the gels were stained with Coomassie blue R-250 (0.1% in 40% methanol 10% acetic acid). Destaining was carried out in 40% methanol 10% acetic acid.

2.3 Antibody preparation

2.3.1 Expressing the polypeptide encoded by the pLZG1.8 in pWR590-1 overexpression vector

The 3′ 1.2 kb portion of the pLZG1.8 insert was excised by *Bam*HI/*Bgl*II restriction enzyme digestion, and cloned in-frame in the pWR590-1 overexpression vector. Thus, cloned sequences are under the transcriptional control of the *lac*

promoter, and translated in fusion with approximately the first 590 amino acids of β-galactosidase (Guo *et al.* 1984). The construct was transformed into DH5αF' cells. Cell lysates from positive clones were prepared and analyzed by SDS-PAGE analysis for overexpression of a polypeptide (that is absent in untransformed cells) with molecular weight around 130 kDa. Protein extracts were prepared from pellets of overexpressing cells, and resolved in 14 X 20 cm (7%) SDS-polyacrylamide gels, 3mm in thickness. The polypeptide of interest was band-purified from the gel and electroeluted in SDS-running buffer at 60 A for 3-4 hours. The protein eluate was dialysed (3-4 hours at a time) against 3 changes of 10 mM Tris Cl pH 7.6). The yield and purity of the recovered protein were tested by SDS-PAGE analysis.

2.3.2 Immunizations

Prior to the primary injection, preimmune serum was sampled as control for further analysis. Protein samples injected into rabbits were Vortex mixed with an equal volume of incomplete Freund's adjuvant and injected as an emulsified solution. Total protein extract from bacterial cells expressing the fusion protein were used for the first injection. Subsequently, the rabbits were injected with approximately 150-200 ng of gel-purified protein at 4 week intervals, over a period of 3 months. Injections were given subcutaneously. Test-bleeds were taken 10 days after each

boost, and used in Western blot analysis.

2.3.3 Western blot analysis

Two micrograms of total, cytoplasmic, and nuclear protein extract of *D. discoideum* cells were resolved on a 10% SDS polyacrylamide gel, using a BIORAD mini protein electrophoresis apparatus. Gel preparation was done according to the instructions of Biorad for reagent and gel preparation. The protein samples were size-fractionated by electrophoresis at 3 watts per gel in SDS running buffer [0.025 M Tris base, 0.192 M glycine, 0.1% SDS] for 45 min. After electrophoresis the polypeptides were immobilized on Immobilon P membrane (MILLIPORE) by electroblotting in transfer buffer [25 mM Tris base, 192 mM glycine, 15% methanol] at 120 mA for 3 hours. The membrane was blocked in 1 X TBS, 0.5% Tween 20, 5% non-fat milk at room temperature for 1 hour. Different sections of the membrane were incubated with preimmune serum, immune serum or blocking buffer between 1 and 2 hours at room temperature.

The immune serum was preadsorbed to bacterial proteins from an acetone powder preparation, to remove cross-reacting materials. The unbound primary antibodies were removed by washing 4 times in 1 X TBS, 0.5% Tween 20, for 10 min at room temperature. All membrane sections were incubated with alkaline phosphatase conjugated goat anti-rabbit secondary antibodies at 1:1000 dilution in blocking buffer for - 1 hr at

room temperature. The secondary antibodies were incubated with methanol fixed slime-mould cells prior to use. This procedure reduces non-specific binding of the secondary antibody to *D. discoideum* proteins. Washing of the secondary antibodies was done as described above. The membranes are rinsed in alkaline phosphatase buffer [100 mM Tris.Cl (pH 9.5), 100 mM NaCl, 5 mM $MgCl_2$] and incubated in alkaline phosphatase substrate mixture [0.025% nitro blue tetrasolium (pNBT), 0.012% bromochloroindolyl phosphate (BCIP)] in alkaline phosphatase buffer. The reaction was stopped by incubating in alkaline phosphatase buffer containing 0.5 M EDTA.

2.4 Northern blotting

Eight micrograms of cellular RNA from vegetative *D.discoideum* cells, and from cells at 8 hours and 16 hours of development were denatured in the presence of 1.7 % formaldehyde, 50% formamide, 2.5% SDS and size-fractionated in a 1.5% agarose formaldehyde gel by electrophoresis in formaldehyde gel-running buffer [0.02 M MOPS, 1 mM EDTA (pH 7.0), 5% formaldehyde] (Maniatis *et al.*,1982). RNA was transferred to NYTRAN (Schleischer & Schuell) and fixed on the membrane by UV irradiation. Efficiency of RNA transfer was assessed by inspection of the membrane and gel under UV irradiation after completion of the transfer. The membrane was prehybridized at 37°C in [5 X Denhardt's, 50% formamide,

5 X SSC, 0.5% SDS, 200 µg/ml salmon sperm DNA] for 3-4 hours. Hybridization was carried out over 16-18 hours, in the same buffer at 37°C. The probe was the *Bam*HI/*Eco*RI 1.8 kb fragment of pLZG2.0, with estimated specific activity of 10^7 cpm/µg. The filter was washed for 15 min at room temperature in [2 X SSC, 0.1% SDS] and 2 X 20 min at 60°C in the same buffer, and exposed on X-ray film (Kodak) in -70°C with intensifyig screens.

2.5 Southern blotting

Genomic DNA (2 µg) of *D. discoideum* was digested with various restriction enzymes, separated by electrophoresis in a 1% agarose gel and (Maniatis et al., 1982) blotted on NYTRAN membranes (Schleischer & Schuell). Hybridizations were done in the same solution as described for the library screenings and Northern analysis, at 37°C for 16 hours. Washing was done as desribed in the section above.

3. RESULTS

The experimental results described in this section aim at the characterization of a novel *D. discoideum* gene. This project stems from a research program directed towards the analysis of the molecular processes underlying *D. discoideum* development, and focused on the regulatory role of cAMP in the developmental process. Specifically, the molecular mechanisms that mediate the effects of cAMP on developmentally regulated gene expression are investigated. In this context, a novel cAMP binding protein (CABP1) was isolated and monoclonal antibodies were raised against it. Western blot analysis revealed a number of cross-reacting polypeptides. Their cross-reactivity with the CABP1 monoclonal antibody was used as a means to screen expression libraries for cDNA clones encoding them. One of these clones harboured a 1.0 kb long insert that encodes a leucine zipper motif sequence. The clone was named pLZc1.0. Subsequently, the cDNA insert was used as a probe to screen a genomic library, for overlapping clones. This screening yielded a clone with a 4.0 kb genomic insert, that was designated pLZG4.0. This section describes the isolation and characterization of additional sequences of the pLZc1.0 clone. In addition, antibody was raised to characterize its gene product.

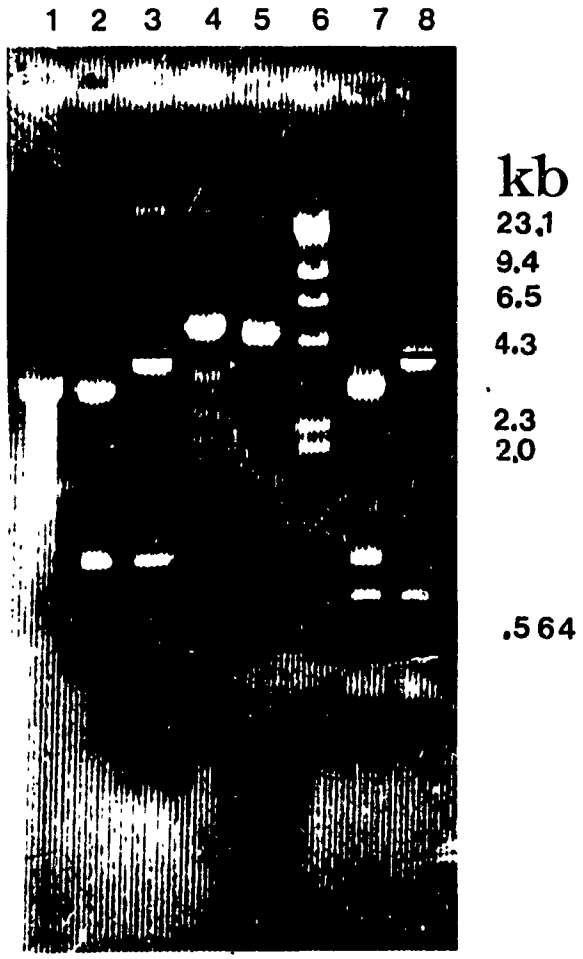
3.1 Library screening

Screening of the λ ZAP II cDNA library was conducted using the 1.8 kb subclone of the genomic pLZG4.0 to probe a total of 10^6 plaques on duplicate filters. The first round of screening revealed three putative positive clones which were purified from neighbouring plaques via a secondary screening. The second round of purification confirmed two of the purified clones as true positives. Positive plaques appeared at frequencies of approximately 5% in the second round of purification. cDNA sequences harboured in these positives were rescued from the λ ZAP II construct as pBluescript cDNA clones by *in vivo* excision.

3.2 Primary characterization of positives

To identify the approximate sizes of the different positives, they were subjected to restriction endonuclease analysis. DNA was prepared from individual colonies harbouring the phagemids, digested with *Eco*RI restriction endonuclease, and the digests were examined on a 1% agarose gel. DNA from the gel was blotted onto a nitrocellulose membrane and the blot probed with the pLZG1.8 kb insert. The resulting autoradiogram (data not shown) revealed strong hybridization to a 1.0 kb long *Eco*RI restriction fragment. The clones corresponding to these digests were interpreted as true positives. Analysis with *Pvu*II endonuclease showed the

Fig. IV Gel showing restriction endonuclease analysis of isolated positive clones pLZc1.7 and pLZc2.0: lanes 1-3, *Eco*RI digests of pBluescript vector, pLZc2.0 and pLZc1.7 respectively; lanes 4-5, *Bgl*II digests of pLZc2.0 and pLZc1.7 respectively; lane 6, *Eco*RI/*Hin*dII digest of ' DNA; lanes 7-8, *Eco*RI/*Bgl*II digests of pLZc2.0 and plZc1.7 respectively. Lane 2: The bottom band in lane 2 is in fact a doublet of two 1.0 kb *Eco*RI fragments. One of them is seen as the second band in lane 7; the other one contains a *Bgl*II site at 0.75 bp from its 5′ terminal end -and is represented by the 0.75 kb and 0.25 kb bands in lane 7. Lane 3: The top band in lane 3 is approximately 3.7 kb. Sequence analysis later revealed that a defective *Eco*RI site demarcates the 5′ terminal end of the pLZc1.7 insert. Thus *Eco*RI digestion yields a 1.0 kb insert fragment and 700 bp of 5′ terminal insert sequence fused to the 3.0 kb pBluescript vector.


1 2 3 4 5 6 7 8
kb
23.1
9.4
6.5
4.3
2.3
2.0
.564

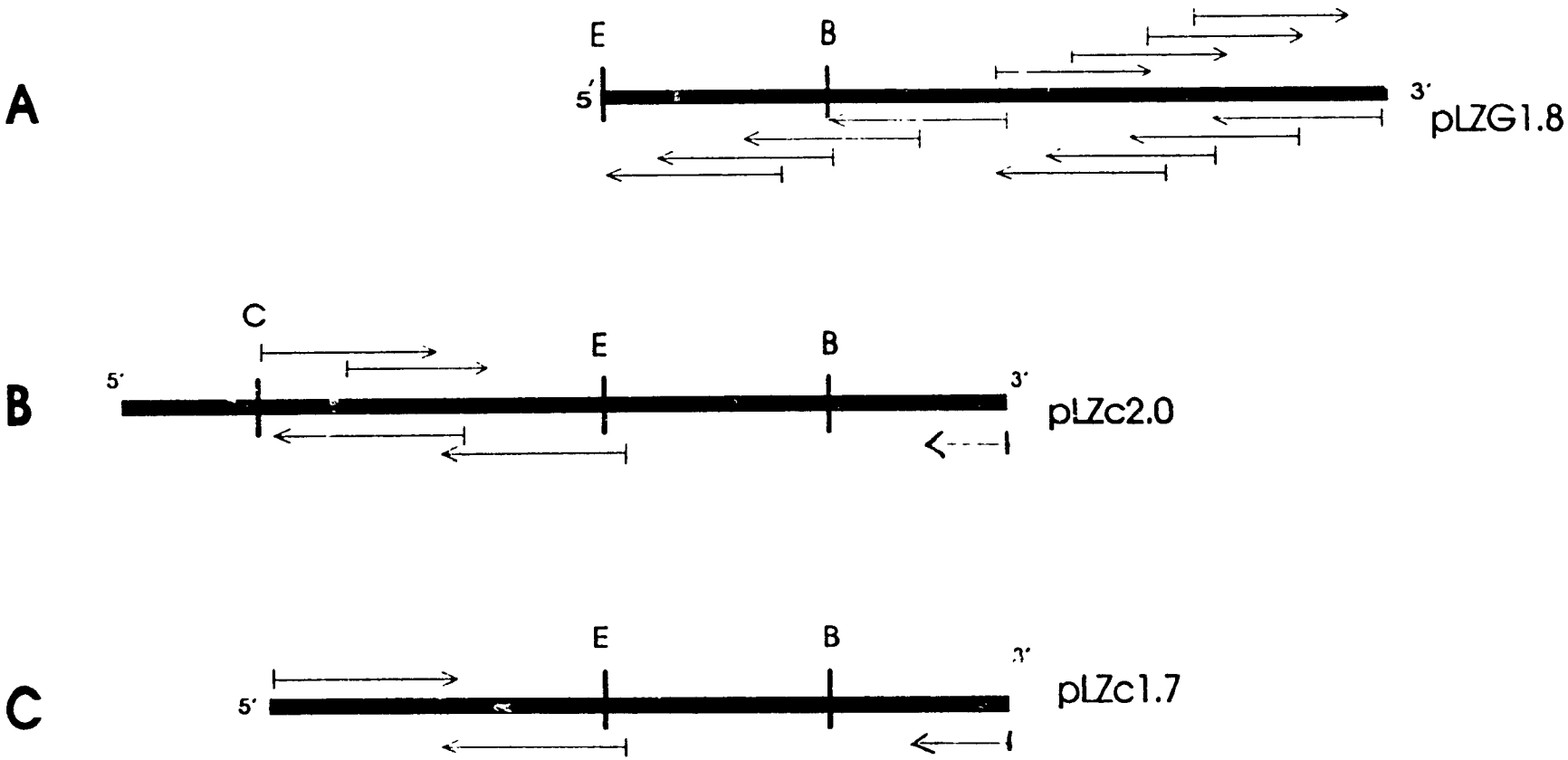
inserts from the two isolated cDNA clones to be approximately 2.0 kb and 1.7 kb long. The clones were named pLZc2.0 and pLZc1.7 respectively. Restriction analysis with *Bgl*II and *Eco*RI/*Bgl*II followed to further characterize the cloned cDNA sequences (Fig. IV). Both inserts were shown to have a 1.0 kb *Eco*RI fragment, harbouring a 0.25 kb *Eco*RI/*Bgl*II restriction subfragment (Fig. IV lanes 2-3 & 7-8). This part of the restriction digestion pattern was identical to that revealed when the pLZc1.0 insert is subjected to the same analysis (data not shown). This result, in conjunction with the positive hybridizations described above, showed inserts from both pLZc2.0 and pLZc1.7 share a 1.0 kb fragment of identical sequence between themselves and with pLZc1.0 (see Fig. II).

The restriction endonuclease analysis described above and similar analysis of the genomic clone revealed the following (see Fig.II): a) the *Eco*RI site at the 3′ end of all the cDNA clones was introduced by the adaptor used during cloning; b) the additional *Eco*RI site in all the clones was an integral part of the sequence; c) all the cDNA and genomic clones of this gene isolated so far shared in common approximately 1.0 kb downstream from the internal *Eco*RI site; d) pLZc2.0 and pLZc1.7 extended the 5′ sequence of the previously isolated clones by 1.0 kb and 0.7 kb, respectively.

3.3 Sequence analysis

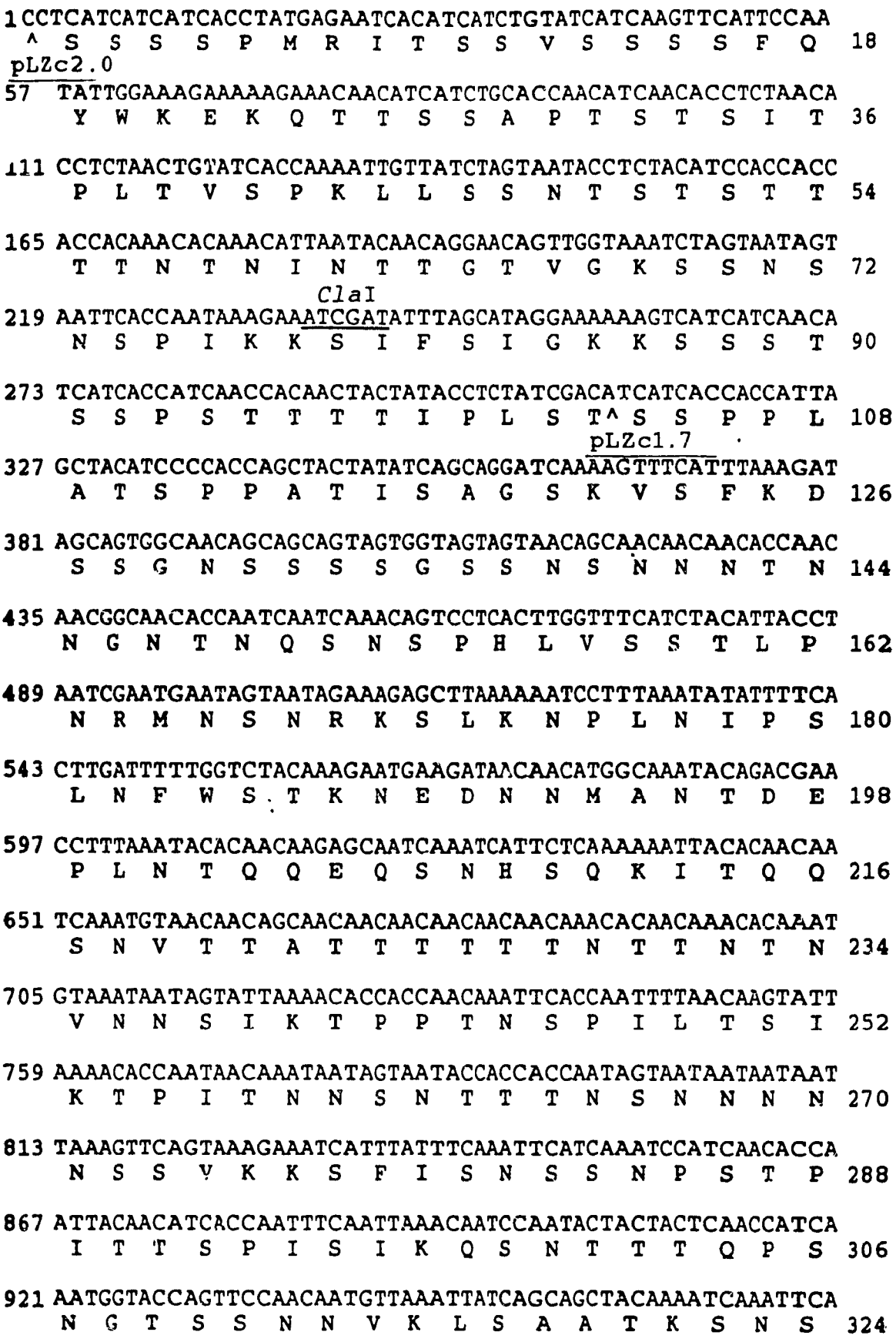
Sequence analysis was performed on both pLZc1.7 and

Fig. V Sequencing strategy for analysis of deletion clones from pLZG1.8 (A), from pLZc2.0 (B), or pLZc1.7 (C). Truncated end points of the target fragments are designated by short vertical lines. The length of each arrow indicates the length of sequence obtained from a particular clone. Arrowheads indicate that the target fragment extends to the (5′ or 3′) end of the parent clone.


A
E
B
5'
3'
pLZG1.8
B
C
E
B
5'
3'
pLZc2.0
C
E
B
5'
3'
pLZc1.7

pLZc2.0 cDNA clones as well as on the genomic insert of pLZG1.8. A bidirectional set of pLZG1.8 random deletion mutants were produced by subjecting the parent clone to pancreatic DNAse I digestion. Derivative clones falling in the desirable size range were identified by restriction endonuclease analysis. Selected mutants were ordered according to target fragment size on a scale where truncated end-points between successive deletions were 100-400 bp apart. Sequence analysis was preformed on these plasmids. Sequence data from successive deletions share 100-200 bp long overlaps. Target fragments in deletion clones are schematically represented in Fig. V (A). The extent of sequence information from each is also depicted in the figure. Sequence analysis verifies that the 5′ half of the genomic sequence in pLZG1.8 is for the most part identical to the cDNA insert of clone pLZc1.0. There are two discrepancies between the cDNA sequence recorded from pLZc1.0 and the corresponding region of the genomic sequence. First, the *Eco*RI restriction site that marks the 3′ terminus of the pLZc1.0 cDNA insert is not present in the genomic sequence, confirming that it is introduced during cloning by the addition of *Eco*RI adaptors. Second, a 15-A stretch that comprises the 3′ end of the cDNA clone gave the initial impression that it was the poly-A tail of the mRNA. However, the 3′ end of the cDNA does not contain stop codons. Therefore the poly-A stretch is probably an internal sequence. Analysis of the genomic sequence revealed

Fig.VI Nucleotide sequence and deduced amino acid sequence of a gene encoding a leucine zipper protein in *D. discoideum*. Nucleotides are numbered on the left and amino acids on the right. The leucine zipper motif is in brackets. For nucleotides 980-1980 the complementary sequence comes from the work of Simon Delagrave. Sequence information on nucleotides 2810-3119 comes from unidirectional analysis of the sense strand. The three GAAAAA repeats discussed in the text are double underlined. The primer used in sequence analysis is in bold characters. Restriction enzyme sites are designated on top of the sequence and underlined. Arrow-heads underneath the sequence mark the 5'-terminal ends of the designated clones. This is a composite sequence from analysis of pLZc1.7, pLZc2 0, pLZG1.8 and pLZG4.0.


1 CCTCATCATCATCACCTATGAGAATCACATCATCTGTATCATCAAGTTCATTCCAA
^ S S S S P M R I T S S V S S S S F Q 18
pLZc2.0
57 TATTGGAAAGAAAAAGAAACAACATCATCTGCACCAACATCAACACCTCTAACA
Y W K E K Q T T S S A P T S T S I T 36
111 CCTCTAACTGTATCACCAAAATTGTTATCTAGTAATACCTCTACATCCACCACC
P L T V S P K L L S S N T S T S T T 54
165 ACCACAAACACAAACATTAATACAACAGGAACAGTTGGTAAATCTAGTAATAGT
T T N T N I N T T G T V G K S S N S 72
ClaI
219 AATTCACCAATAAAGAAATCGATATTTAGCATAGGAAAAAAGTCATCATCAACA
N S P I K K S I F S I G K K S S S T 90
273 TCATCACCATCAACCACAACTACTATACCTCTATCGACATCATCACCACCATTA
S S P S T T T T I P L S T^S S P P L 108
pLZc1.7
327 GCTACATCCCCACCAGCTACTATATCAGCAGGATCAAAAGTTTCATTTAAAGAT
A T S P P A T I S A G S K V S F K D 126
381 AGCAGTGGCAACAGCAGCAGTAGTGGTAGTAGTAACAGCAACAACAACACCAAC
S S G N S S S S G S S N S N N N T N 144
435 AACGGCAACACCAATCAATCAAACAGTCCTCACTTGGTTTCATCTACATTACCT
N G N T N Q S N S P H L V S S T L P 162
489 AATCGAATGAATAGTAATAGAAAGAGCTTAAAAAATCCTTTAAATATATTTTCA
N R M N S N R K S L K N P L N I P S 180
543 CTTGATTTTTGGTCTACAAAGAATGAAGATAACAACATGGCAAATACAGACGAA
L N F W S T K N E D N N M A N T D E 198
597 CCTTTAAATACACAACAAGAGCAATCAAATCATTCTCAAAAAATTACACAACAA
P L N T Q Q E Q S N H S Q K I T Q Q 216
651 TCAAATGTAACAACAGCAACAACAACAACAACAACAAACACAACAAACACAAAT
S N V T T A T T T T T N T T N T N 234
705 GTAAATAATAGTATTAAAACACCACCAACAAATTCACCAATTTTAACAAGTATT
V N N S I K T P P T N S P I L T S I 252
759 AAAACACCAATAACAAATAATAGTAATACCACCACCAATAGTAATAATAATAAT
K T P I T N N S N T T T N S N N N N 270
813 TAAAGTTCAGTAAAGAAATCATTTATTTCAAATTCATCAAATCCATCAACACCA
N S S V K K S F I S N S S N P S T P 288
867 ATTACAACATCACCAATTTCAATTAAACAATCCAATACTACTACTCAACCATCA
I T T S P I S I K Q S N T T T Q P S 306
921 AATGGTACCAGTTCCAACAATGTTAAATTATCAGCAGCTACAAAATCAAATTCA
N G T S S N N V K L S A A T K S N S 324

EcoRI

975 ATGAATTCAGCAGTACCAAGAAAGAATAGTACAACTACCACTAATAATACATCT

M ▲N S A V P R K A S T T T T N N T S 342

pLZG4.0, pLZc1.0

1029 TTATCAACATCAACATCACCATCAATATCAGCCAATTCAATGCCACCACCATCA

L S T S T S P S I S A N S M P P P S 360

1083 TCAACAACTCCAATAATGAAAATCAAATTCAAGTTCTCATCTAAAAATTGTTTA

S T T P I M K I K F K F S S K N L L 378

1137 TCCATCACCAATACAACACCATCAACAACAACCACCAATATTAATAATAATAGT

S I T N T T P S T T T T N I N N N S 396

1191 AGTATTTCAGATGATACATCATCAGAATCAACACCACATAATACACCACCACTT

S I S D D T S S E S T P H N T P P L 414

1245 TCAGGTAGTGGTGGTACTATACCAAATGGGTTATCATCATCAATTGATCCTTTT

S G S G G T I P N G L S S S I D P D 432

1299 GATCCAGACGAATTAAATGTAACAAAGTTAAAGAGTAAAATTAATAAATCAGAT

N P D G L N V T K L K S K I N K S D 450

1353 TTTGAACGTTTAGTTAATTTAGAATTTTATACACGTGTTTTATCAAAGGATTTA

F G R L V N L E F Y T R V L S K N L 468

1407 GAAGATGTAACTTCAAAATTAGTGAATGAATCATCAATGAGAAAAGATTTAGAA

E A V T S K L V N E S S M R K N L E 486

1451 TCATCACATACTAAAGAATTGGAAGCAAGTTTAAGAATTTATGAAGCAAGATTA

S H T K E L E A R L R I Y E A R L K 504

1515 AAAAGATATTCAACAATGCCACCACCACCACAACCTGAACCAATTCCACAACAA

K R Y S T M P P P P P Q P E P I P Q Q 522

1569 CAACATCAATTAGCAATTGGTGATTTATCACCATTATCATCTTCACCACCAACT

Q H Q L A I G D L S P L S S S P P T 540

1623 ACTGGTATTTCAATTAATCATGATAGTGGTAATACAATAGTTTTAACAGCACCA

T G I S I N H D S G N T I V L T A P 558

BglII

1677 CCAAATGTTTCAACACCACCACCAGTTCCACCAAGATCTTATTTAAATTTAATT

P N V S T P P P V P P R S Y L N L I 576

1731 CAAGCAGAAAAGGATCAATTGGAATTGGAGGAGACTGAAATATCAAAAGCTAAA

Q A E K D Q L E L E E T E I S K A K 594

1785 GAGGAAAGAGAAGATCAAAGATTTAATTCATTGGTTGTAAGATTGGAACAATTG

E E R E D Q R F N S [L V V R L E Q L 612

1839 GAAAAGAAAGTTTATTCATTAGAGCAAAGAAATGGTCAATTGTTGGAAGAGATT

E K K V Y S L E Q R N G Q L L E E I 630

1893 CATCATTTACAAACTGAAAATCAGTTACTTAATCCAGATAGAAAGAAAACTTTT

H H L Q T E N Q L] L N P N R K K T L 648

1947 AAGAAATTAGAAAAAGAAAAAGAAAAAGAATTACAAAAGGAACAAGAAAGAATA
K K L E K E K E K E L Q K E Q E R I 666

2001 CTACGTGAAAAAGAGAAAGAATTGCAAAAAGAAAAAGAAAGGGAGAAACAAAGA
L R E K E K E L Q K E K E R E K Q R 684

2055 CAAAAGGACTTTTTAAAACAACAAAAAGAAAAAGAAAAAGAACAACAAAAACAT
Q K D F L K Q Q K E K E E K Q Q K H 702

2109 CAAAAACAATCTACTCATCAACTTTCACAACAAGATCATTTTGAAAATTCAAAT
Q K Q S T H Q L S Q Q D H F E N S N 720

2163 TCTGCCACGGCTCCAGATTTAGATACTTTTGATTCAGATTCAAATGATACAAAT
S A T A P D L D T F D S D S N D T N 738

2217 TCAGAGAATGATAATAGTGGTAATAATAATAATAATAATAATGAAAAGAAA
S N L K R K D Y P M S L V S D L S P 750

2271 TCAAATCTAAAGAGAAAGGATTACCCAATGTCATTAGTTAGTGATTTATCACCA
S N L K R K D Y P M S L V S D L S P 774

2325 AATAAAAAGAATGTAGTTTTTAATGATGGTAAATCAGCAGCACCACCAACAACA
N K K N V V F N D G K S A A P P T T 792

2379 GCCACCACCTCACTTAAATCTTCAAATTATTTATCATCAGTAACTTTAACAAAA
A T T S L K S S N Y L S S V T L R K 810

2433 GGTGCAGCAAAGAGAATGTCTGCTCATTATACTTCAACCAACAACAACAGGTCA
G A A K R M S A H Y T T S T T T G S 828

2487 CTAATTGATGATGATAGTAGTTTAGATGATAATCAATCAACTACCTCTGAAAAG
L I D D D S S L D D N Q S T T S E K 846

2541 GTTAAAGCTATTTTCCATGATATTTACGATGGTAATATTGAAGAAATGAAAAAT
V K A I F H D I Y D G N I E E M K N 864

2595 TCAAAATTAACTAGAATGAGAAGTAAATCACTATCAAAGGATGCTTCATCAGGC
S K L T R M R S K S L S K D A S S G 900

2649 AGTAATTTAGATGACGATGATGATGTTTATAATAGTCAGCCATCATCTGCTGCT
S N L D D D D D V Y N S Q P S S A A 918

2703 AAATTAAGTTTCACACAACAGCAACATAGTCAATTACCACCCACCCAACAACAA
K L S F T Q Q Q Q H S Q L P P T Q P 936

ScaI

2757 CCAAGTCCACCACTTCAATCAACTTCAACATCTAGTATAGCATCAGTACTTATT
P S P P L Q S T S T S S I A S V L I 954

2817 AGTAAAGAGAGTCTTTCAATTCCATTATGCACTTCAGTTGATACAAATGTTACA
S K E S L S I P L C T S V D T N V T 972

2871 ACTACAACTACAACCATTCCTACACCAGTGGTTCAACCACAACAGTTTCACCAA
T T T T T I P T P V V Q P Q Q F H Q 990

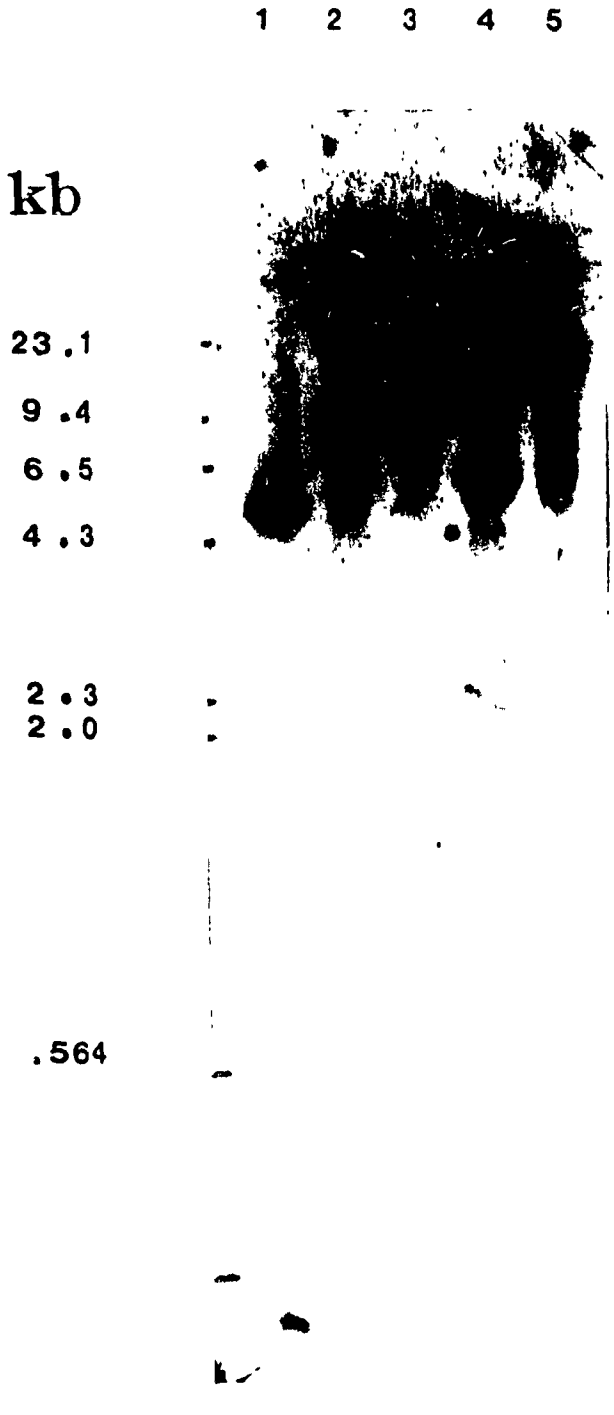
```
2925 ATGTACCAATTTCAAAACCAGATGCAATGGAAAAAATTAGACATTTAACAATTG
     M  Y  Q  F  Q  N  Q  M  Q  W  K  K  L  D  I  -  Q  L 1008

2979 GTAGAATGTCAAGTAAAAAGAAACCAGGTTATAAAGATGCTCAATTACAATTAA
     V  E  C  Q  V  K  R  N  Q  V  I  K  M  L  N  Y  N  -  1026

3033 GTGAAAAATTAAATCATAGAGCTTCGAAAAATGAATTAATTTTTAAAAATATTA
     V  K  N  -  I  I  E  L  R  K  M  N  -  F  L  K  I  L 1044

3087 TCAAAGCAGGTAAAGTATATTTATTTTTAAATTATTATTATTATAATAATAATC
     S  G  N  V  K  Y  I  Y  F  -  I  I  I  I  I  I  I  I 1062

3141 TATTTCAATTTTTTTTTTTTATTTTTATTTTTCAAATATTTGTTTGAAAGTTTT
     Y  F  N  F  F  F  L  F  L  F  F  K  Y  L  F  E  S  F 1080

3195 TTTT
     F
```

that the poly A stretch is not present. In its place there are three repeats of GAAAAA (see Fig. VI nucleotides 1956-1973). Our interpretation is that the oligo-dT primer used to initiate first strand synthesis during construction of the cDNA library was able to hybridize to the GAAAAA repeats and gave rise to a complementary second strand with a poly-A stretch at it's 3' end. Overall, sequencing data from the genomic clone, revealed a single uninterrupted ORF that runs through the entire length of the insert (Fig. VI nucleotides 982-2810).

A synthetic primer complementary to the pLZG1.8 3' terminus was designed and used to obtain information on downstream sequences from the parent clone pLZG4.0. The primer is comprised by nucleotides 2751-2770 in Fig.VI. The primer binding site is ~ 40 bp upstream the *Sca*I restriction site used in subcloning the 5' half of the genomic sequence. Sequence analysis identified an additional 400 bp of sequence downstream the *Sca*I restriction site. Translational analysis revealed an ORF of ~ 80 amino acids followed by multiple stop-codons encoded by the AT rich 3' region of the sequence. This region probably represents the 3' end of the gene. Clone pLZc2.0 was subjected to *Exo*III nuclease digestion to produce subclones containing progressive unidirectional deletions of the 2.0 kb insert. In addition, the internal *Cla*I site was utilized to generate subclones for sequencing of the parent clone. Successive deletions differing in size by 250-400 bp

spanning the 5' portion of pLZc2.0 are depicted in Fig. V (B). The pLZc1.7 was also subcloned into smaller fragments making use of the internal *Eco*RI site, and sequenced Fig. V (C). A synthetic primer which binds immediately downstream the internal *Eco*RI site was also used in sequence analysis of both pLZc2.0 and pLZc1.7. The primer binding site (5'GCAGTACCAAGAAAGAATA 3') corresponds to nucleotides 984-1004 (see Fig. VI). The sequences in the two cDNA clones are identical except that pLZc2.0 is longer than pLZc1.7 by ~300 bp. Sequence analysis of pLZc2.0 revealed a single uninterrupted ORF that runs through its entire length. Together, sequence analysis of pLZG1.8, pLZG4.0 and pLZc2.0/pLZc1.7 yielded a 3118 base pair long nucleotide sequence; a single continuous ORF of 1005 amino acids runs through almost the entire 3198 nucleotide sequence. This ORF is interrupted by multiple stop codons at the 3' terminal 230 bp of the sequence. The deduced amino acid sequence revealed structural features similar to the leucine zipper group of transcriptional activator proteins. Nucleotides 1815-1919 encode a leucine zipper motif. This is a stretch of 35 amino acids in an α-helical arrangement (amino acids 605-639) where every seventh amino acid is a leucine. Additional characteristic regions of the amino acid sequence include: amino acids 511-569 with a 23% content of prolines; amino acids 579-601 which harbour 11 acidic residues and only 3 basic residues; amino acids 644-661 which contain 50% of

Fig. VII Genomic analysis by Southern blotting. Genomic DNA (1 µg) is digested with *Cla*I (lane 1), *Hind*II (lane 2), *Hin*dII (lane 3), *Eco*RI (lane 4), or *Eco*RV (lane 5). The blot was probed with the genomic insert in pLZG1.8.


1 2 3 4 5
kb
23.1
9.4
6.5
4.3
2.3
2.0
.564

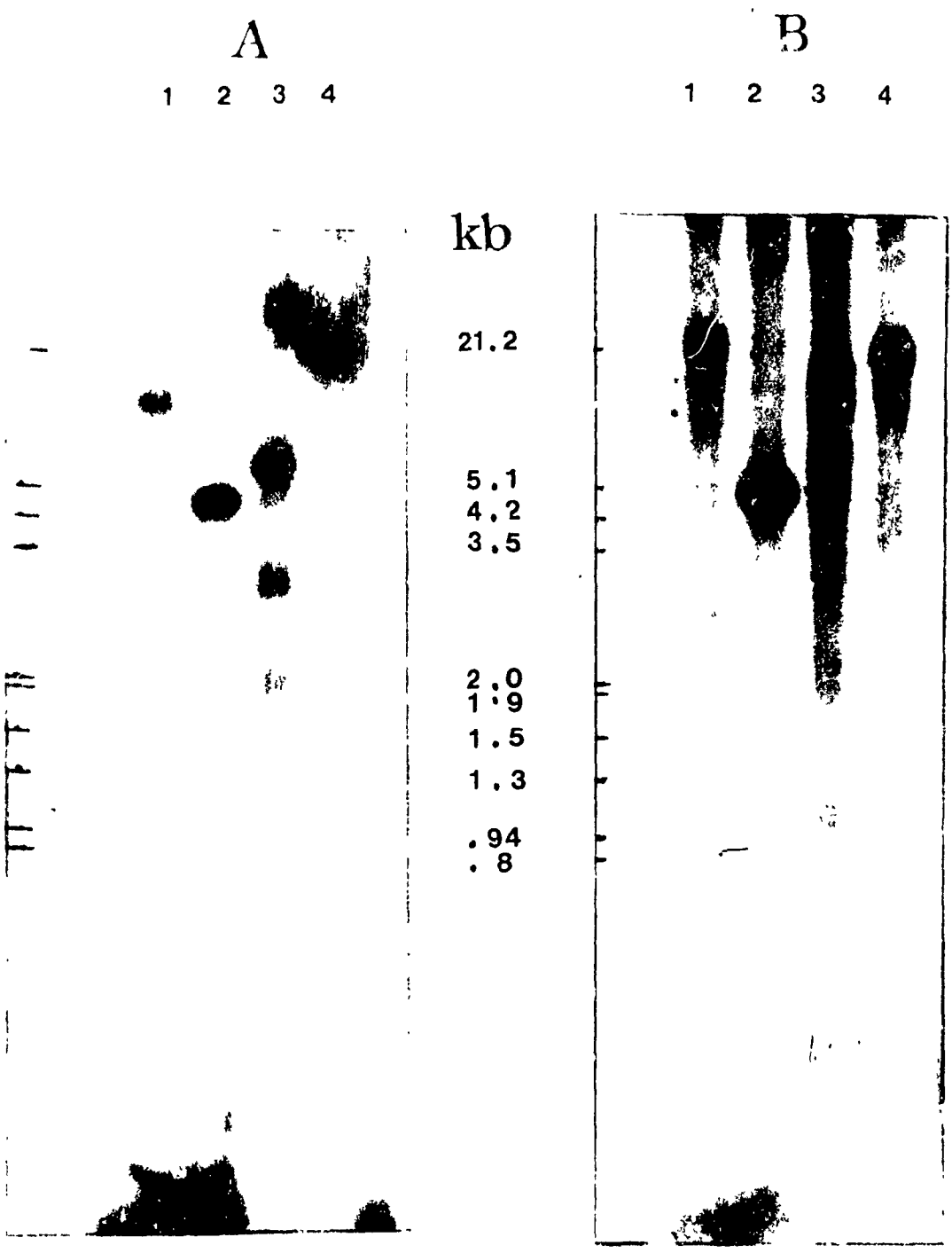
positively charged residues.

3.4 Blot analysis of genomic DNA.

As explained above, the recorded sequence on 1.0 kb of cDNA carried in pLZc1.0 was not in absolute agreement with the respective genomic sequence in pLZG1.8. The possibility exists that the genomic clone is the result of two unrelated fragments of DNA joined together during cloning. To rule out this possibility a Southern analysis was preformed. Genomic DNA that was cut with *Cla*I, *Hind*II, *Hind*III, *Eco*RI or *Eco*RV was size-fractionated in a 1.5% agarose gel and blotted onto a nitrocellulose membrane, which was subsequently probed with the 1.8 kb insert of the pLZG1.8 genomic clone (Fig. VII). The probe hybridized to a single species of *Cla*I, *Hind*II, *Hind*III, *Eco*RI or *Eco*RV genomic fragment, as shown in Fig. VII lanes 1-5. These fragments are ~5.0 kb, ~9.0 kb, ~18.0 kb, ~6.0 kb and ~22.0 kb respectively. The pattern from this Southern analysis was compared to results obtained when the same blot was probed with the pLZc1.0 kb cDNA insert (Michael Greenwood, Ph.D thesis, 1993) and was found to be exactly reproduced. These results demonstrated that the 5′ half of the insert in pLZG1.8 contains a contiguous fragment of genomic DNA.

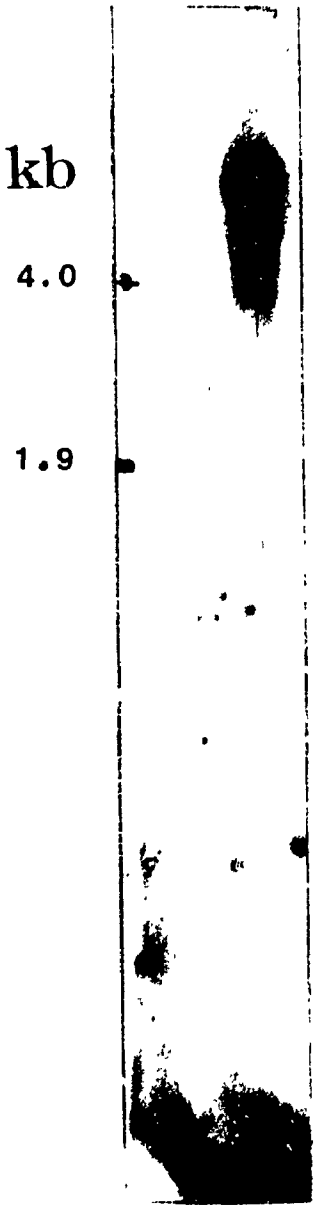
Similarly, the 5′ portions in clones pLZc2.0 and pLZc1.7 were tested with respect to their continuity with the 3′ portions of these clones. Genomic DNA was digested with

Fig. VIII Genomic analysis by Southern blotting using genomic and cDNA clones as probes. Sister digests of genomic DNA are from left to right: *Bgl*II (lane 1), *Cla*I (lane 2), *Eco*RI (lane 3), *Sca*I (lane 4). Five micrograms of genomic DNA were digested with each enzyme. Digestion reactions were split into two for analysis on sister blots with different probes. The probes used are: the genomic insert from pLZG1.8 (panel A), or the *Eco*RI/*Cla*I cDNA fragment from plZc2.0 (panel B).


A
1 2 3 4
B
1 2 3 4
kb
21.2
5.1
4.2
3.5
2.0
1.9
1.5
1.3
.94
.8

*Bgl*II, *Cla*I, *Eco*RI, or *Sca*I. The digests were split in two and each half was electrophoresed in a 1% agarose gel and blotted onto a nitrocellulose membrane. One of the blots was probed with the 1.8 kb genomic insert of pLZG1.8 -probe A (Fig. VIII A). The second blot was probed with the *Eco*RI/*Cla*I fragment of pLZc2.0 -probe B (Fig. VIII B). Figures II and III convey a schematic representation of the fragments used as probes in this analysis, and their relative positions in the sequence. Both cDNA and genomic probes hybridized with a single *Cla*I fragment and a single *Sca*I fragment ~5.0 kb and ~20.0 kb respectively (Fig. VIII A or B, lanes 2 & 4). Since there are no *Cla*I or *Sca*I restriction sites in either probe, this would be the expected result if the two probes are parts of a contiguous piece of DNA fragment. Two *Bgl*II genomic fragments of ~21.0 kb and ~15.0 kb were recognized by the genomic DNA fragment -probe A (Fig. VIII A, lane 1). As can be seen in Fig. II and Fig. III a *Bgl*II restriction site is contained in this fragment. Consequently, recognition of two fragments in a *Bgl*II DNA digest was the expected result. When the pLZc2.0 *Eco*RI/*Cla*I fragment was used as a probe, a single signal of ~21.0 kb was detected. This band coincides exactly in size with the larger of the *Bgl*II fragments detected by the genomic probe. Thus, the two probes which comprise distinct regions in tne sequence, are shown to be contiguous in a *Bgl*II fragment of ~21.0 kb. *Eco*RI digested DNA that was probed with the genomic fragment revealed a single hybridizing species of

Fig. IX RNA blot analysis: Total RNA from vegetative cells is probed with the pLZG1.8 genomic fragment. RNA markers are the 4.0 kb and 1.9 kb ribosomal RNAs of *D. discoideum*.


kb
4.0
1.9

~6.0 kb. When the cDNA sequence was used as a probe, a signal of ~13.0 kb was detected. Since the two probes constitute separate *Eco*RI fragments in the sequence, they are expected to hybridize to nonidentical *Eco*RI genomic fragments. These results suggest that the new sequences in the 5′ portion of pLZc2.0 are contiguous to the 3′ portion of this clone and to the 3′ portion of pLZG1.8.

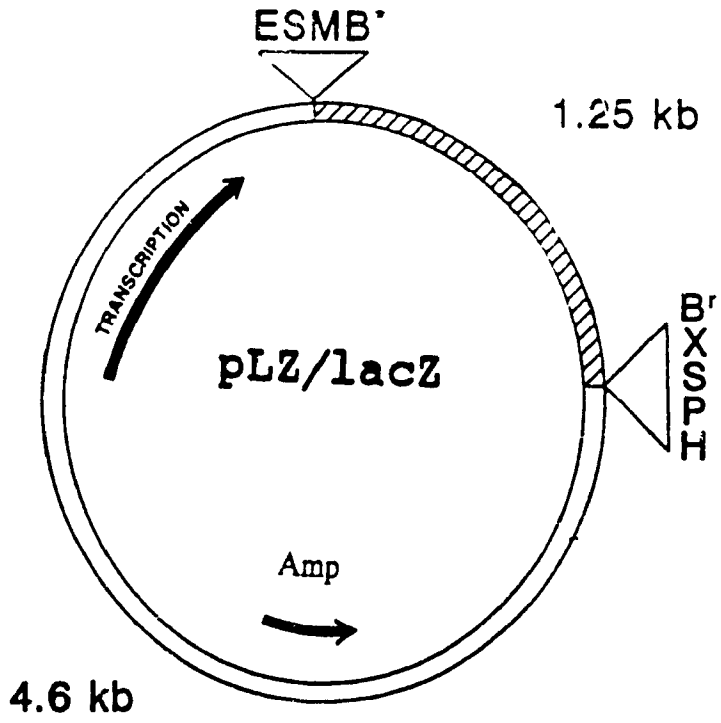
3.5 RNA blot analysis

To establish the length of the cDNA sequence that represents the gene under characterization, a Northern analysis was preformed. This information would assess the extent to which this gene is represented in the cloned sequences. Total cellular RNA from vegetative *D.discoideum* cells was size fractionated in a denaturing 1.5 % agarose-formaldehyde gel. RNA was transferred onto a nitrocellulose membrane and probed with the 1.8 kb insert of clone pLZc1.8 (Fig. IX). A single mRNA species, greater than 8.0 kb in size, hybridized to the DNA probe. The size markers used for the size determination were the 4.0 kb and 1.9 kb *D. discoideum* ribosomal RNAs. We have so far characterized approximately 3.0 kb of this gene which is less than half the transcribed sequences of the leucine-zipper gene.

3.6 Antibody production

An alternative way of predicting the extent of missing

Fig. X Overexpression plasmid pLZ/lacZ carries the *Bgl*II/*Bam*HI fragment of the genomic clone pLZG1.8 encoding the leucine zipper motif. The nucleotide sequence carried in the insert corresponds to nucleotides 1717-2810 in Fig. III.



B: *Bam*HI
E: *Eco*RI
H: *Hind*III
M: *Sma*I
P: *Pst*I
S: *Sac*I
X: *Xba*I

: Leucine zipper gene

: pWR590 vector

*: destroyed site
r: reconstructed site

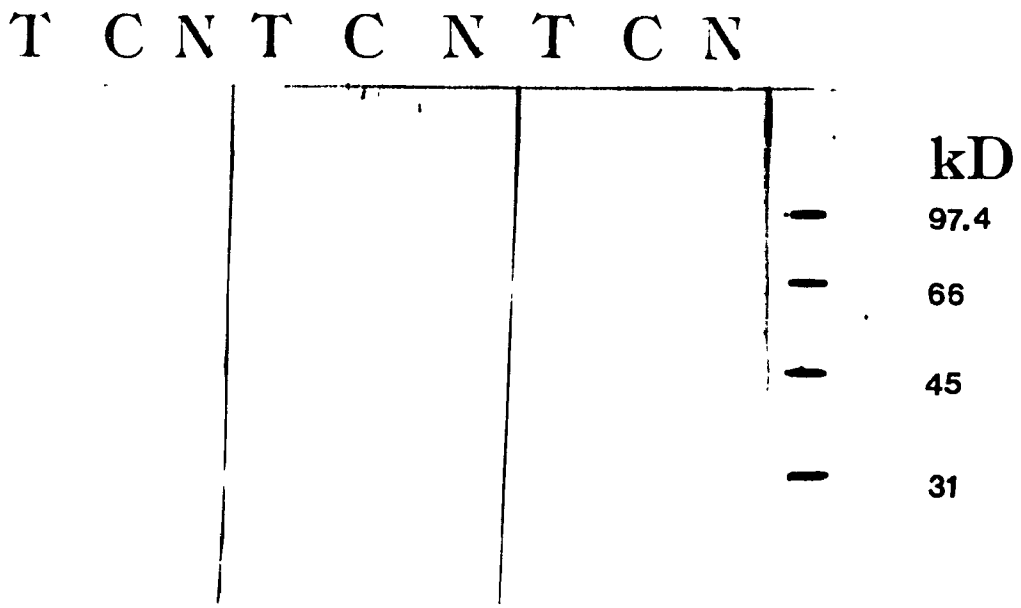
coding sequences was to identify the size of the protein product of the gene under characterization. Subcellular localization of the gene product was another objective. For these purposes, antibody was raised against the polypeptide product of a part of the genomic insert in clone pLZG1.8. The 1.2 kb 3'-terminal portion of the 1.8 kb pLZG1.8 insert was cloned in pWR590-1 overexpression vector in a sense orientation as a *Bgl*II/*Bam*HI fragment. (Fig. X). The cloned fragment, contains the sequence encoding the leucine zipper motif, but does not carry the sequence encoding a proline rich region that is thought to account for the cross-reactivity among the polypeptides that are recognized by the CABP1 monoclonal antibodies. The polypeptide product was overexpressed in fusion with the first 590 amino acids of β-galactosidase under the control of the Lac promoter in DHαF' transformants of the construct. SDS-PAGE analysis of these transformants revealed a dominant band with molecular weight of at least 120 kDa, which correlates with the predicted molecular weight of the fusion protein (Fig. XI lanes 1-4 & 8-10). The 120 kDa fusion polypeptide was absent from protein-extracts of DHαF' cells (Fig. XI lane 7), while the 65 kDa protein overexpressed in vector transformants (Fig. XI lane 8) was replaced in the fusion transformants by the 120 kDa protein. Polyclonal antisera specific to the fusion protein were prepared by repeated inoculation into rabbits. After 4 boosts with fusion peptide, sera were tested for antibody

Fig. XI SDS-PAGE analysis of transformants of the pLZ/lacZ fusion construct. SDS-polyacrylamide gel showing: Protein extracts of transformants of the fusion construct, lanes 1-4 & 8-10; untransformed DH5αF′ cells, lane 5; Standard Low Molecular Weight marker, lane 6; transformants of the vector pWR590-1, lane 7.


1
2
3
4
5
6
7
8
9
10
kD
200
116
97.4
66
45

reactivity and specificity in Western blot analysis (Fig. XII) Antibody reactivity was tested using whole cell extracts as well as cytoplasmic and nuclear protein extracts from vegetative *D. discoideum* cells. The non-specific materials in the antisera were removed by incubating the antiserum with protein extract (as an acetone powder preparation) from bacterial cells transformed with the expression vector alone. Additionally, the secondary antibodies were preadsorbed to methanol fixed *D. discoideum* cells to eliminate any possible non-specific interactions between the rabbit anti-mouse immunoglobulins and slime mold proteins. A single band, with an approximate molecular weight of 60 kDa was detected by the immune serum in the nuclear protein fraction. No signal was detected by the preimmune serum (Fig. XII lanes 1-3) or by the secondary antibodies (Fig. XII lanes 7-9). These results further support the specificity of the antibodies. Unexpectedly, no signal was observed in the total protein fraction. The data suggest a nuclear distribution for the 60 kDa protein that reacts to antibodies specific to the *BglII/ScaI* part of the putative leucine zipper protein.

Fig. XII Western blot analysis: Nylon membrane blot of total (T), cytoplasmic (C), and nuclear (N), fractions from *D. discoideum* vegetative cells. Proteins are challenged with preimmune serum (lanes 1-3), immune serum (lanes 4-6), or treated only with secondary antibody, goat anti-rabbit IGgs (lanes 7-9).


T C N T C N T C N
kD
97.4
66
45
31

4. DISCUSSION

In previous work monoclonal antibodies directed against *D.discoideum* cAMP-binding protein CABP1 were found to detect a number of cross-reacting polypeptides. The basis for immunological cross-reactivity among CABP1-related proteins was investigated by isolation and characterization of the genes encoding them. Anti-CABP1 antibodies were employed to isolate from a λ gt11 expression library the genes encoding cross-reacting proteins. Several cDNA clones were isolated and subjected to sequence analysis. One of the cDNAs contained a 1.0 kb long insert that encoded a leucine zipper motif and was designated pLZc1.0. The respective gene was termed *zip*A. The cDNA sequence was found to terminate shortly after the leucine zipper motif and contained no translation termination codon. The cDNA clone was used to isolate a genomic clone (pLZG4.0) with a 4.0 kb insert. The present study describes further characterization of the zipA gene and its protein product.

In an attempt to isolate full-length cDNA clones, we used a *zip*A genomic fragment as a hybridization probe in library screenings. Two cDNA clones were isolated. These contained 1.7 kb and 2.0 kb long inserts and were termed pLZc1.7 and pLZc2.0 respectively. Restriction digestion mapping, Southern blotting and sequence analysis suggested the 3′ portions of

pLZc1.7 and pLZc2.0 overlap with pLZc1.0. These data further suggested that clones pLZc1.7 and pLZc2.0 extended the 5′ side of pLZc1.0 by 0.7 kb and 1.0 kb respectively. Moreover, all cDNA clones were found to share identical 3′ ends. Sequence analysis of the genomic clone revealed that the cDNA inserts contained a structural anomaly. The inserts end with a polyA tract that is absent in the corresponding region of the genomic clone. In its place there are three repeats of GAAAAA. We believe that the oligo-dT primer used in first strand synthesis during construction of the cDNA library hybridized to these repeats in the message. Consequently cDNAs with polyA tails at their 3′ ends were created.

During construction of cDNA libraries unrelated sequences are occasionally joined together. Data obtained from these hybrid clones will lead to erroneous interpretation. Clones pLZc1.7 and pLZc2.0 are independent isolates. They have identical sequence in the 1.7 kb in which they overlap. These data suggest that both pLZc1.7 and pLZc2.0 are complementary sequences of the *zip*A mRNA. In addition, results of genomic analysis by DNA blotting indicate that the cDNA clones and the genomic clone carry overlapping sequences of the *zip*A gene.

The intergenic regions and introns of *Dictyostelium* are extremely rich in AT residues, usually over 90% (Sharp and Devine, 1989). Thus, open-reading frames are very easy to identify. The 3′ end of the sequence reported here has a high AT content. In this region there are no consensus sequences

demarcating the intron splice junction (.../GUAAGU) (Grant *et al.*, 1990). Therefore, the first stop codon is likely to be the 3′ end of the coding region.

Compilations of sequences reveal that eukaryotic transcripts do not normally contain AUG codons upstream from the functional AUG triplet, nor have eukaryotic polycistronic mRNAs been reported (Kozak, 1983). Thus it has been proposed that eukaryotic translation initiates only at the most 5′ AUG codon (Mueller and Hinnebusch, 1986). In 95% of the cases AUG codons are not found in the 5′ untranslated regions of eukaryotic mRNAs; thus the first AUG of the open reading frame is easily identified to be the translation start site. In the reported sequence the most 5′ proximal AUG codon encodes for methionine at position 6 of the open reading frame. The DNA sequence obtained does not exhibit an obvious 5′ leader non-coding region; thus, the relative position of the functional translation initiation codon cannot be inferred directly from the sequence.

To determine the length of the *zip*A mRNA we examined the cellular RNA by Northern blotting. Our results revealed an mRNA species that was estimated to be at least 8.0 kb in length. Thus, we concluded that sequences determined so far make up < 50% of the mature mRNA. However, the extent to which the present sequence represents the coding capacity of the *zip*A message is not known. Many transcripts contain long untranslated regions. Examples are provided by the cDNAs of

a number of leucine zipper proteins that are classified as members of the bZIP family of transcriptional regulators e.g c-jun, c-fos, GCN4 (Vogt *et al.*, 1987). The respective cDNAs contain long (500 bp-1,000 bp) non-coding sequences flanking their 5′ and/or 3′ ends. The complete nucleotide sequence of the *zip*A message can become available from the analysis of additional cDNA clones. Alternatively, sequence analysis of the entire genomic sequence would help identify introns and coding sequences that are differentially represented in the mature mRNA. This information may help identify the coding capacity of the *zip*A gene.

An alternative approach to gain insight into the protein coding sequences in *zip*A was to characterize its protein product by immunological techniques. CABP1-related polypeptides contain proline-glutamine rich regions. This region in CABP1 has been characterized as the antigenic epitope of the monoclonal antibodies used in previous studies. To avoid generating antibodies to this common region, a region specific to the ZipA is employed as the antigen. The part of the *zip*A sequence used for antibody generation includes the leucine zipper motif and sequences downstream. The sequence was fused to the first 590 amino acids of β-gal. Sequences encoding a proline-rich region are eliminated in the course of cloning. Serum from rabbits immunized with the fusion protein was collected and used in Western blot analysis. Western blot analysis revealed a single band of ~60 kDa associated with the

nuclear protein fraction on blots of *D. discoideum* total, cytoplasmic and nuclear protein extracts. The observed molecular weight of 60 kDa is in contrast to the predicted molecular weight of at least 120 kDa by sequences characterized thus far. Post-translational modifications of the protein may explain the discrepancy between the observed and deduced molecular weights. Additionally, the observed molecular weight of 60 kDa suggests that a substantial portion of the sequence reported here is non-coding. The 60 kDa polypeptide was not detected in the total protein fraction. However, the concentration of nuclear proteins is not equivalent in the nuclear and total protein fractions. The nuclear protein fraction contains nuclear proteins at 10X their concentration in the total protein fraction. Thus, Western blot analysis may not detect the low level of the 60 kDa polypeptide in the total fraction.

Sequence analysis of the clones representing *zip*A revealed a number of structural characteristics resembling the bZip family of transcription activator proteins. These include: a) A periodic repeat of leucines at every seventh position through a stretch of 35 residues (amino acids 605-639), b) a region of basic amino acid residues (amino acids 644-661), c) a region rich in acidic amino acids (amino acids 579-501), and d) a stretch with a high proline content (residues 511-569).

Classical transcription regulators contain combinations

of similar or analogous features. Generally, transcription by RNA polymerase II requires the formation of a stable reinitiation complex over the proximal promoter region of each target gene. Regulation of this event is imposed by more distal control regions (or combinations of these) such as upstream promoter regions, enhancers or locus activation regions. Sub-elements in these control regions serve as binding sites for transcription factors. Such interactions provide fine-tuning of the control over the event of transcription initiation. Proteins that regulate gene expression generally recognize specific DNA sequences through the binding properties of a distinct domain. Several classes of such domains have been identified (Pabo, 1984). These include the zinc finger motif, the helix-loop-helix motif and the hypothetical motif of the leucine zipper. A shared characteristic among sequence-specific DNA binding proteins is their ability to form dimers, and employ bipartite DNA contact areas to bind selectively to their DNA targets. Dimerization mediates the assembly of a DNA contact surface and the formation of a three-dimensional "scaffold" that dictates appropriate positioning of the interacting protein surface into the grooves of double helical DNA. Dimerization domains may be involved in the formation of heterodimers and/or homodimers. Some heterodimers are competent of mediating positive transcriptional effects, but dominant negative effects have also been described in which a non-functional

partner suppresses the activity of a functional partner (Kouzarides and Ziff, 1988; Halazonetis *et al.*, 1988). Transcriptional regulators also contain one or more trans-activation domains that are required to mediate positive effects on the general transcriptional machinery (Hope and Struhl, 1986; Courey and Tjian, 1988; Mermod *et al.*, 1989).

One feature that was immediately obvious upon inspection of the predicted *zip*A protein sequence is that it contains a leucine zipper motif. This is a structure that contains leucines occurring at intervals of every seventh amino acid in a region predicted to be α-helical. The leucine zipper is a hypothetical structure facilitating dimer formation in a variety of proteins.

Examples of "Leucine Zipper" proteins include nuclear DNA-binding proteins that exert transcriptional control, membrane receptors, ligand gated ion channels or glucose transporter glycoproteins (Cormack *et al.*, 1989; Britton, 1990; Rodriguez and Park, 1993). Known leucine zipper domains have a minimum of 4 leucine repeats organized into α-helices, which are believed to form coiled-coil ropes during dimerization. The geometry of the coiled-coil conformation relies on a repeating heptapeptide sequence of the type $(a\text{-}b\text{-}c\text{-}d\text{-}e\text{-}f\text{-}g)_m$ where residues a and d generally have a hydrophobic character but no specific limitation is imposed in the remaining residues. A fundamental premise of the leucine zipper model is that a dimerization interface is formed

between polypeptide chains via hydrophobic interactions between two α-helices. (Parry *et al.*, 1977; Landschulz *et al.*, 1988; O'Shea *et al.*, 1989; Vinson *et al.*, 1989; Adel and Maniatis, 1989; Gentz *et al.*, 1989). In this model leucines occur along the same face of an amphipathic α-helix; the two helices are thought to attain a parallel association; in the dimer interface leucines are thought to interact with their side-chains and lend stability to the complex.

In the deduced amino acid sequence of *zip*A amino acids 603-639 comprise a segment of leucine heptad repeats characterized by the absence of prolines and glycines and by a high density of potential ion-pairs. It is then compatible with α-helical secondary structure. The distribution of leucines in amino acids 605-640 appears not to be simply a function of the abundance of this residue, since the 35 amino acid long motif contains only two more leucines. The occurrence of an Asn at position 603 conforms to the strong, highly position-specific preference for this residue at the N-terminal end of α-helices. The geometry of this amino acid facilitates H-bonding between the Oxygen at its side-chain and the backbone NH of residue N_0+2 or N_0+3 (where N_0 refers to the Asn). The role of the Asn residue is therefore one of determining the N-terminus of the helix, stabilizing the first helical turn, and providing definition to the α-helical structure (Richardson and Richardson, 1988). In amino acids 605-639 leucines are conserved at position d and position a is

occupied by a hydrophobic residue (phenylalanine, isoleucine, leucine or valine) with the exception of two occurrences of asparagine at that position. However, leucine zippers unlike prototypical coiled-coils may contain destabilizing residues such as an asparagine in the dimer interface (O'Shea *et al.*, 1991; Abel and Maniatis, 1989). Destabilization of the leucine zipper may facilitate reversible dimerization and/or control over the concentration of dimerized molecules that might be crucial to the pleiotropism of protein function (Kouzarides and Ziff, 1989; Turner and Tjian, 1989; Sassone-Corsi *et al.*, 1988). Positions e and g in adjacent heptads in coiled coils are commonly occupied by charged residues which is consistent with the formation of intrahelical and/or interhelical ion-pairs. Inspection of the ZipA leucine zipper motif reveals oppositely charged amino acids (e.g Glu and Arg or Glu and Lys) at positions e and g of adjacent heptades. However, the functional significance of this leucine zipper motif remains to be determined.

Several known transcription factors that bind to DNA in a dimerization dependent manner are known to form dimers via leucine zipper structures. Nuclear factors of this group (bZip family of transcriptional regulators) have DNA binding domains with some well conserved features. In most examples of such regulatory proteins the DNA binding domains contain clusters of basic amino acids, and are adjacent to the leucine zipper on its amino side e.g c-Fos, c-Jun, GCN4, CREB, c/EBP

(Vinson *et al.*, 1989).

It has been proposed (O'Shea *et al.*, 1989; Vanson *et al.*, 1989) that these basic regions generate the DNA contact surface -a bipartite surface analogous to the arms of a Y- at the bifurcation point of two "zippered" helices which correspond to the stem of the Y. It is furthermore thought that DNA binding proteins that utilize the leucine zipper motif dock in such a way as to allow the basic regions to track in opposite directions around dyad half-sites. The *zip*A deduced amino acid sequence exhibits the two hallmarks common to leucine zipper DNA binding proteins: an α-helix permissive region containing an heptad repeat of leucines, and a highly basic region located immediately adjacent to the leucine-repeat helix (residues 644-661). Compilation of sequence data has revealed that basic regions in bZip proteins exhibit a consensus, 16 residue sequence [BB-BN--AA-B-R-BB] -where B signifies a basic amino acid- that starts exactly 7 residues on the amino side of the first leucine of the zipper (Turner and Tjian, 1989). The basic region of ZipA appears unusual in more than one way. First, the basic region is on the carboxyl site of the leucine zipper. A cursory inspection of the ZipA basic region reveals clusters and pairs of basic amino acids which however do not conform to the distribution of positive charge in the consensus sequence sited above. Examples have been reported of bZip regulatory proteins with a basic region that deviates in amino acid composition and

overall charge from the consensus. These were found to exhibit dependence for DNA binding on additional sequences quite far removed from the bZip motif (Subramanian *et al.*, 1991). Thus, unusual features in the primary level of sequence in a bZip protein may suggest additional, particular constraints in the proper coupling of information in dimers of that species. However, it is not known whether the contiguous arrangement of the ZipA leucine zipper motif and basic region has a functional role. Consequently the significance of the basic region and its characteristics cannot be evaluated at this point.

It is believed that transcription regulatory segments of many gene-expression regulators contact components of the transcription apparatus via acidic domains. Protein-protein interactions are thought to be mediated by electrostatic interactions involving the negatively charged residues (Hope and Struhl, 1986; Ma and Ptashne, 1987). A 22 amino acid segment (residues 579-601) in the deduced amino acid sequence of zipA contains 11 acidic residues and only 3 basic residues. This is an additional feature of the *zip*A gene product that is reminiscent of a classical transcriptional regulatory domain.

The deduced zipA sequence exhibits a stretch of 58 amino acids (residues 511-569) with a high proline content (23%). Closer examination of the sequence reveals that a portion of the proline-rich stretch is interdigitated with glutamine residues. Glutamines have polar side-chains that can be both

hydrogen-bond donors and acceptors. This combination of features makes this part of ZipA a candidate for mediating protein-protein interactions via hydrogen-bonding. Known transcriptional activators contain proline-rich regions that have been shown to contribute to transcriptional activation (Mermod *et al.*, 1989). A feature that may be particular to the proline rich region is that it may provide a malleable structure that can participate in various interactions enhancing transcription, in concert with one or more components of the transcription machinery. However, the kind of interactions -if any- in which the proline-rich region of ZipA may participate in, is not known at this preliminary stage of analysis.

On the basis of the above structural clues and its nuclear localization, we think that the product of *zip*A may be participating in dimer formation via a leucine zipper structure and may have dimerization dependent DNA binding activity. We also hypothesize that an acidic region and a proline-rich region in the protein may mediate protein-protein interactions. The possibility of a dual potential for both protein-protein interactions and protein DNA interactions may suggest the potential for involvement in modulation of gene expression. Further studies should establish whether ZipA functions in DNA sequence recognition and/or in direct interactions with other transcriptional factors. Also interactions between ZipA protein and CABP1 and CABP1-related

polypeptides should be investigated. Such studies should reveal whether and how zipA may function in *D. discoideum* development.

LITERATURE CITED

Abel, T., and Maniatis, T. 1989. Action of leucine zippers. Nature **341:** 24-25

Armant, D.R., and Rutherford, C.I. 1977. 5′ AMP nucleotide is localized in the area of cell-cell contact of prespore and prestalk regions during culmination of *Dictyostelium discoideum*. Mech. Ageing Dev. **10:** 199-217.

Armant, D.R., Stetler, D.A., and Rutherford, C.I. 1980. Cell surface localization of 5′ AMP nucleotidase in prestalk cells of *Dictyostelium discoideum*. J. Cell. Sci. **45:** 119-129.

Bain G., Grant C E., and Tsang A. 1991. Isolation and characterization of cDNA clones encoding polypeptides related to a *Dictyostelium discoideum* cyclic AMP binding protein. Journ. of General Microbiol. **137:** 501-508.

Barklis, E., and Lodish, H.F. 1983. Regulation of *Dictyostelium discoideum* mRNA's specific for prespore or prestalk cells. Cell **32:** 1139-1148.

Berlot, C.H., Spudich, J.A., and Devreotes, P.N. 1985. Chemoattractant elicited increase in myosin phosphorylation in *Dictyostelium*. Cell **43:** 307-314

Berridge M.J. 1989. Inositol trisphosphate, calcium, lithium and cell signalling. JAMA **262:** 1834-1841

Bonner, J.T. 1980. Comparative biology of the cellular slime moulds p. 1-33. In Loomis, W.F. (ed.) The Development of *Dictyostelium discoideum*. Academic Press, New York.

Brookman J.J., Town C.D., Jermyn K.A., Kay, R.R. 1980. Developmental regulation of a stalk-cell differentiation-inducing factor in *Dictyostelium discoideum*. Dev. Biol. **91:** 191-196.

Chung, S., Landfear, S.M., Blumberg, D.D., Cohen, N.S., and Lodish, H.F. 1981. Synthesis and stability of developmentally regulated *Dictyostelium discoideum* mRNAs are affected by cell-cell contact and cAMP. Cell **24:** 785-797

Cocucci, S.M., and Sussman, M. 1970. RNA in cytoplasmic and nuclear fractions of cellular slime mold amebas. J. Cell Biol. **45:** 399-407.

Coukel, M. B., Lappano, S., and Cameron, A. M. 1983. Isolation and characterization of cAMP unresponsive (Frigid) aggregation-deficient mutants of *Dictyostelium discoideum*. Dev. Genet. **3:** 283-297.

Coukel, M.B., Cameron, A.M., Petrie, C.M., and Mee, J.D. 1984. Developmental regulation and properties of the cGMP-specific phosphodiesterase in *Dictyostelium discoideum*. Dev. Biol. **103:** 246.

Courey A. J., Tjian R. 1988. Analysis of Sp1 *in vivo* reveals multiple transcriptional domains, including a novel Glutamine-rich activation motif. Cell **55:** 887-898.

DeGunzberg, J., Franke, J., Kessin, R.H., and Veron, M. 1986. Detection and developmental regulation of the mRNA for the regulatory subunit of the cAMP-dependent protein kinase of *D. discoideum* by cell-free translation. EMBO J. **5:** 363-367.

Devine, K.M., Bergmann, J.E., and Loomis, W.F. 1983. Spore coat proteins are packaged in prespore vesicles. Dev. Biol. **107:** 364-372.

Devine, K.M., and Loomis, W.F. 1985. Molecular characterization of anterior-like cells in *Dictyostelium discoideum*. Dev. Biol. **107:** 364-372.

Devreotes, P. 1989. *Dictyostelium discoideum*: A model system for cell-cell interactions in development. Science **245:** 1054-1058.

Durston, A.J. 1976. Tip formation is regulated by an inhibitory gradient in *Dictyostelium discoideum*. Nature **263:** 126-129.

Europe-Finner, G.N., and Newell, P.C. 1985. Inositol, 1, 4, 5-triphosphate induces cyclic GMP formation in *Dictyostelium discoideum*. Biochem. Biophys. Res. Commun. **130:** 1115-1122.

Europe-Finner, G.N., and Newell, P.C. 1987. Cyclic cAMP stimulates accumulation of inositol triphosphate in *Dictyostelium discoideum*. J. Cell Sci. **87:** 513-518.

Europe-Finner, G.N., Gammon, B., Wood, C.A., and Newell, P.C. 1988. Mutant *ras* gene induces elevated levels of inositol tris- and hexakis phosphate in *Dictyostelium*. J. Cell Sci. **89:** 13-20.

Europe-Finner, G.N., and Newell, P.C. 1988. Inositol 1,4,5-trisphosphate and calcium stimulate actin polymerization in *Dictyostelium discoideum*. J. Cell Sci. **82:** 41-51.

Firtel, R.A., and Bonner, J. 1972. Characterization of the genome of the cellular slime mold *Dictyostelium discoideum*. J. Mol. Biol. **66:** 339-361.

Gentz, R., Ransher, F.J. III, Abate, C., Curran, T. 1989. Parallel association of *fos* and *jun* leucine zippers juxtaposes DNA binding domains. Science **243:** 1695-1699.

Gerisch, G. 1987. Cyclic AMP and other signals controlling cell development and differentiation in *Dictyostelium*. Ann. Rev. Biochem. **56:** 853-879.

Ginsburg, G., and Kimmel A.R. 1989. Inositol trisphosphate and diacylglycerol can differentially modulate gene expression in *Dictyostelium*. Proc. Natl. Acad. Sci. USA **86:** 9332-9336.

Gomer, R.H., and Firtel, R.A. 1987. Cell-autonomous determination of cell-type choice in *Dictyostelium* by cell cycle phase. Science **237:** 758-762.

Gonzalez, G.A., Yamamoto, K.K., Fischer, W.H., Karr,D., Menzel,P., Biggs III, W., Vale, W.W., and Montminy M.R. 1989. A cluster of phosphorylation sites on the cyclic AMP-regulated factor CREB predicted by its sequence. Nature **337:** 749-752.

Grant, C.E., Bain,G., and Tsang A. 1990. The nuclear basis for alternative splicing of the CABP1 transcripts in *D. discoideum*. Nucl. Acid. Res. **17:** 5457-5463.

Greenwood, M.T. 1991. Sequence and expression of annexin VII of *Dictyostelium discoideum*. Bioch. Biophys. Act. **1088:** 429-432.

Greenwood, M.T. 1992. Regulation of the gene encoding the p24 actin -bindfing protein in *Dictyostelium discoideum*. Biochem. Cell Biol. **70:** 1057-1059

Gross, J.D., Bradbury, J., Kay, R.R., Peacy, M.J. 1983. Intracellular pH and the control of cell differentiation in *Dictyostelium discoideum*. Nature **303:** 244-245.

Halazonetis, T.D., Georgopoulos, K., Greenberg, M.E., Leder, P. 1988. *c-jun* dimerizes with itself and with *c-fos*, forming complexes of different DNA binding affinities. Cell **55:** 917-924

Haribabu B., and Dottin, R.P. 1986. Pharmacological characterization of cyclic AMP receptors mediating gene regulation in *Dictyostelium discoideum*. Mol. Cell Biol. **6:** 2402-2408.

Hashimoto, S., Schmid, W., and Schutz, G. 1984. Transcriptional activation of the rat liver tyrosine aminotransferase gene by cAMP. Proc. Natl. Acad. Sci. USA **81:** 6637-6641.

Hjorth A.L., Khanna, N.C., and Firtel R.A. 1989. A *trans*-activation factor required for cAMP-induced gene expression in *Dictyostelium* is regulated developmentally and induced by cAMP. Genes and development **3:** 747-759.

Hoeffler, J.P., Meyer, T.E., Yun, Y., Jameson J.L., and Habener J.F. 1988. Cyclic AMP-responsive DNA-binding protein structure based on a cloned placental cDNA. Science **242:** 1430-1432.

Hope, I.A., and Struhl K. 1986. Functional dissection of a eukaryotic transcriptional activator protein, GCN4 of yeast. Cell **46:** 885-894.

Janssens, P.M.W., and VanHaastert, P.J.M. 1987. Molecular basis of transmembrane signal transduction in *Dictyostelium discoideum.* Microbiol. Rev. 396-418.

Jermyn, K.A., Berks, M.M., Kay, R.R., and Williams, J.G. 1987. Two distinct classes of prestalk-enriched mRNA sequences in *Dictyostelium discoideum.* Development **100:** 745-755.

Kay, R.R. 1979. Gene expression in *Dictyostelium discoideum*: mutually antagonistic roles of cyclic-AMP and ammonia. J. Embryol. Exp. Morph. **52:** 171-182.

Kay, C.A., Noce, T., and Tsang, A.S. 1987. Translocation of an unusual cAMP receptor to the nucleus during development of *Dictyostelium discoideum.* Proc. Natl. Acad. Sci. **84:** 2322-2326.

Kay, R.R., and Jermyn, K.A. 1983. A posible morphogen controlling differentiation in *Dictyostelium.* Nature **303:** 242-244.

Kay, R.R. 1982. cAMP and spore differentiation in *Dictyostelium discoideum.* Proc. Natl. Acad. Sci. USA 79: 3228-3231.

Kesbeke, F., and VanHaastert, P.J.M. 1985. Selective down regulation of cell surface cAMP-binding sites and cAMP-induced responses in *Dictyostelium discoideum.* Biochem. Biophys. Acta. **847:** 33-39.

Kessin, R.H., Orlow, S.J., Shapiro, R.I., and Franke, J. 1979. Binding of inhibitor alters kinetic and physical properties of extracellular cyclic AMP phosphodiesterase from *Dictyostelium discoideum.* Proc. Natl. Acad. Sci. **76:** 5450-5454.

Kimmel, A.R., and Saxe, C.L. 1986. cAMPcontrolled gene regulation in *Dictyostelium* mutants exhibiting abnormal cAMP metabolism. Dev. Biol. **117:** 541-552.

Kimmel, A.R., and Carlisle, B. 1986. A gene expressed in undifferentiated vegetative **Dictyostelium** is represed by developmental pulses of cAMP and reinduced during differentiation. Proc. Natl. Acad. Sci. USA. **83:** 2506-2510.

Kimmel, A.R. 1987. Different molecular mechanisms for cAMP regulation of gene expression during *Dictyostelium disccideum* development. Dev. Biol. **122:** 163-171.

Klein, C., Lubs-haukeness, J., and Simmons, S. 1985. cAMP induces

a rapid and reversible modification of the chemotactic receptor in *Dictyostelium discodeum*. 1987. J. Cell. Biol. **100:** 715-720.

Klein, P., Vaughan, R., Borleis, J., and Devreotes, P.N. 1987. The surface cyclic AMP receptor in *Dictyostelium*: levels of ligand induced phosphorylation, solubilization, identification of primary transcript, and developmental regulation of expression. J. Biol. Chem. **262:** 358-364.

Klein, P.S., Sun, T.J., Saxe, C.L.III, Kimmel, A.R., Johnson, R.L., and Devreotes, R.N. 1988. A chemoattractant receptor controls development in *Dictyostelium discoideum*. Science **241:** 1467-1472.

Konijinm, T.M., Van de Meene, J.G.C., Bonner, J.T., and Berkley, D.S. 1987. The acrasine activity of adenosine 3′:5′ cyclic monophosphaqte. Proc. Natl. Acad. USA **58:** 1152-1154.

Kopachik, W. 1982. Size regulation in *Dictyostelium*. J. Embryol. Exp. Morphol. **68:** 23-35.

Kopachik, W., Oohata, A., Dhokia, B., Brookman, J. J., and Kay, R.R. 1983. *Dictyostelium* mutants lacking DIF, a putative morphogen. Cell **33:** 397-403.

Kozak, M., and Shatkin A.J. 1978. Characterization of translational initiaton regions from eukaryotic messenger RNA. In Methods in Enzymology, **60.** (New York, Academic press).

Kozak. M. 1988. How do eukaryotic ribosomes select initiation regions in messenger RNA? Cell **15:** 1109-1123.

Kouzarides T., and Ziff E. 1989. Leucine zippers of *fos*, *jun* and GCN4 dictate dimerization specificity and thereby control DNA binding. Nature **340:** 568-570.

Krebs, E.G. 1989. Role of the cyclic AMP-dependent protein kinase in signal transduction. JAMA **262:** 1815-1818.

Krefft, and Weijer, C.J. 1989. Expression of a cell surface antigen in *Dictyostelium discoideum* slugs. J. Embryol. Exp. Morph. **88:** 15-24.

Kruijer,W., Cooper, J.A., Huster, T., and Verma, I.M. 1984. Platelet-derived growth factor induces rapid but transient expression of the *c-fos* gene and protein. Nature **312:** 711-716.

Kumagai, A., Pupillo, M., Gundersen, R., Miake-Lye, R., Devreotes, P.N., and Firtel, R.A. 1989. Regulation and function of G protein subunits in *Dictyostelium*. Cell **57:** 265-275.

Kuo, J.F., and Greengard, P. 1969. Cyclic nucleotide-dependent

protein kinases. IV. Widespread occurence of adenosine 5'-monophosphate-dependent protein kinase in various tissues and phyla of the animal kingdom. Proc. Natl. Acad. Sci. USA **64:** 1349-1355.

Landfear, S.M., Lefebvre, P., Chung, S., and Lodish, H.F. 1982. Transcriptional control of gene expression during development of *Dictyostelium discoideum*. Mol. Cell. Biol. **2:** 1417-1426.

Landschultz, W.H., Tohnson, P.F., and McKnight, S.C. 1988. The leucine zipper protein: A hypothetical structure common to a new class of DNA bining proteins. Science **240:** 1759-1764.

Liu, G., and Newell, P.C. 1988. Evidence that cyclic GMP regulates myosin interaction with the cytoskeleton during chemotaxis of *Dictyostelium*. J. Cell. Sci. **90:** 123-129.

Loomis, W.F. 1982. (ed.) The development of *Dictyostelium discoideum*. Academic press, New York.

Ma, J., and Ptashne M. 1987. Deletion analysis of GAL4 defines two transcriptional activating segments. Cell **48:** 847-853.

Maeda, Y., and Gerisch, G. 1977. Vesicle formation in *Dictyostelium discoideum* cells during oscillations of cAMP synthesis and release. Exp. Cell. Res. **110:** 119-126.

Malchow D., and Gerisch, G. 1974. Short-term binding and hydrolysis of cyclic 3':5'-adenosine monophosphate by aggregating *Dictyostelium discoideum* cells. Proc. Natl. Acad. Sci. USA **71:** 2423-2427.

Mangiarrotti, G., Chung, S., Zuker, C., and Lodish, H.F. 1981. Selection and analysis of cloned developmentally-regulated *Dictyostelium discoideum* mRNAs. Nature **301:** 616-618.

Mangiarotti, G., Ceccarelli A.M., and Lodish H.F. 1983. Cyclic AMP stabilizes a class of developmentally regulated *Dictyostelium discoideum* mRNAs Nature **301:** 616-618.

Maniatis, T., Fritsch, E.F., and Sambrook, J. 1982. Molecular cloning. A laboratory manual. Cold Spring Harbour Laboratory, Cold Spring Harbour, NY.

Mann, S.K.O., and Firtel, R.A. 1987. Cyclic AMP regulation of early gene expression in *Dictyostelium discoideum*: mediation via the cell surface cyclic AMP receptor. Mol. Cell. Biol. **7:** 458-469.

Mann, S.K.O., Pinko, C., and Firtel, R.A. 1988. cAMP regulation of early gene expression in signal transduction mutants of *Dictyostelium*. Dev. Biol. **130:** 294-303.

Mato, J.M., Krens, F., VanHaastert, P.J.M., Konijn, T.M. 1977. cAMP dependent cGMP accumulation in *Dictyostelium discoideum*. Proc. Natl. Acad. Sci. **74:** 2348-2351.

Maurer, R.A. 1981. Transcriptional regulation of the prolactin gene by ergocryptine and cyclic AMP. Nature 294: 94-97.

McDonald, S.A., and Durston, A.J. 1984. The cell cycle and sorting behavior in *Dictyostelium discoideum*. J. Cell. Sci. **66:** 195-204.

McRobbie, S.J., and Newell, P.C. 1984. Chemoattractant-mediated changes in cytoskeletal actin of cellular slime-molds. J. Cell. Sci. **68:** 139-151.

McRobbie, S.J., Jermyn, K.A., Duffy, K., Blight, K., and Williams, J.G. 1988. Two DIF-inducible, prestalk-specific mRNAs of *Dictyostelium* encode extracellular matrix proteins of the slug. Development **104:** 275-284.

Mee, J.D., Tortolo, D.M., and Coukell, M.B. 1986. Chemotaxis associated properties of separated prestalk and prespore cells of *Dictyostelium discoideum*. Biochem. Cell. Biol. **125:** 59-66.

Mehdy, M.C., Ratner, D., and Firtel, R.A. 1983. Induction and modulation of cell-type specific gene expression in *Dictyostelium discoideum*. Cell **32:** 763-771.

Mermod, N., O'Neill, E.A., Kelly, T.J., Tjian, R. 1989. The proline-rich transcriptionl activator of CTF/NF-1 is distinct from the replication and DNA binding domain. Cell **58:** 741-753.

Mitchell P.J., and Tjian, R. 1989. Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science **245:** 371-379.

Montminy M.R., Sevarino, K.A., Wagner J.A., Mandel, G., and Goodman K.H. 1986. Identification of a cyclic-AMP responsive element within the rat somatostatin gene. Proc. Natl. Acad. Sci. USA **83:** 6682-6684

Morris, H.R., Taylor, G.W., Masento, M.S., Jermyn, K., and Kay, R.R. 1987. Chemical structure of the morphogen differentiation inducing factor from *Dictyostelium discoideum*. Nature **328:** 811-814.

Morrisey, J.H., Devine K.M., and Loomis W.F. 1984. The timing of cell-type specific differentiation in *Dictyostelium discoideum*. Dev. Biol. **103:** 414-424

Mueller P.P., and Hinnebusch A.G. 1986. Multiple upstream AUG

codons mediate transcriptional control of GCN4. Cell **45:** 201-207.

Muller, M.M., and Hohl, H.R. 1973. Pattern formation in *Dictyostelium discoideum:* Temporal and spatial distribution of prespore vacuoles. Diff. **1:** 267-276.

Mutzel, R., Lacombe, M.-L., Simon, M.-N., de Gunzberg, J., and Veron, M. 1987. Cloning and cDNA sequence of the regulatory subunit of cAMP-dependent protein kinase from *Dictyostelium discoideum*. Proc. Natl. Acad. Sci. USA **84:** 6-10.

Newell, P.C. 1982. Genetics p. 35-70. In Loomis, W.F. (ed.) The Development of *Dictyostelium discoideum*. Academic Press, New York.

Newell, P.C., and Ross, F.M. 1982. Inhibition by adenosine of aggregation centre initiation and cyclic AMP binding in *Dictyostelium*. J. Gen. Microbiol. **128:** 2715-2724.

Oohata, A. 1983. A prestalk cell specific acid phosphatase in *Dictyostelium discoideum*. J. Embryol. Exp. Morph. **74:** 311-319.

Oyama, M., and Blumberg, D.D. 1986. Interaction of cAMP with the cell surface receptor induces cell type specific mRNAs in the cellular slime mould *Dictyostelium discoideum*. Proc. Natl. Acad. Sci. **83:** 4819-4823.

O'Shea, E.K.,Klem, J.D., Kim, P.S., Alber, T. 1991. X-ray structure of the GCN4 leucine zipper, a two stranded, parallel coiled-coil. Science **254:** 539-544.

Pabo, C.O. 1984. Protein-DNA recognition. Ann. Rev. Biochem. **53:** 293-321.

Padh, H., and Brenner, M. 1984. Studies on the guanylate cyclase of the social amoebae *Dictyostelium discoideum*. Arch. Biochem. Biophys. **229:** 73-80.

Pastan, I., and Adhya, S. 1976. Cyclic adenosine monophosphate in *Escherichia coli*: Bacteriol. Rev. **40:** 527-551.

Peterkofsky, A. 1976 Cyclic nucleotiodes in bacteria. Adv. Cyclic Nucleotide Res. **7:** 1-48.

Pupillo, M., Kumagai, A., Pitt, G.S., Firtel, R.A., and Devreotes, P.N. 1989. Multiple α subunits of guanine nucleotide-binding proteins in *Dictyostelium*. Proc. Natl. Acad. USA. **86:** 4892-4896.

Pupillo, M., Klein, P., Vaughan, R., Pitt, G., Lilly, P., Sun, T., Devreotes, P., Kumagai, A., and Firtel, R. 1988. cAMP receptor and G-protein interactions control development in *Dictyostelium*.

Cold Spr. Hbr. Quan. Biol. **53:** 657-663.

Raper, K.B. 1940. Pseudoplasmodium formation and organization in *Dictyostelium discoideum*. J. Elisha Mitchell Sci. Soc. **56:** 241-282.

Ratner, D., and Broth W. 1983. Comparison of differentiating *Dictyostelium discoideum* cell types separated by an improved method of density gradient centrifugation. Exp. Cell Res. **143:** 1-13.

Reymond, C.D., Gomer, R.H., Nellen, W., Theibert, A., Devreotes, P., and Firtel, R.A. 1986. Phenotypic changes induced by a mutated *ras* gene during development of *Dictyostelium* transformants. Nature **336:** 83-86.

Richardson, J.S., and Richardson C. 1988 Amino acid preferences for specific locations at the ends of α helices. 1988. Science **240:** 1648-1652

Rodriguez, G.A., and Park M. 1993. Dimerization mediated through a leucine zipper activates the oncogenic potential of the *met* receptor tyrosine kinase. Mol. Cel. Biol. **13:** 6711-6722.

Ross, F.M., and Newell, P.C. 1981. Streamers: chemotactic mutants in *Dictyostelium* with altered cyclic GMP metabolism. J. Gen. Microbiol. **127:** 339-350.

Rubin, J., and Robertson, A. 1975. The tip of the *Dictyostelium discoideum* pseudoplasmodium as an organiser. J. Embryol. Exp. Morph. **33:** 227-241.

Sampson, J. 1976. Cell patterning in migrating slugs of *Dictyostelium*. J. Embryol. Exp. Morph. **36:** 663-668.

Sanger, F., Nicklen, G., and Coulsen, A.R. 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74:** 5463-5467.

Sassone -Corsi, P., Ransone, L.Z., Lamph W.W., Verma, M.I. 1988. Direct interaction between *fos* and *jun* nuclear oncoprotein: role of the "leucine zipper" domain. Nature **336:** 692-695.

Schaap, P., and Van Driel, R. 1985. Induction of post-aggregative differentiation in *Dictyostelium discoideum* by cAMP. Exp. Cell Res. **159:** 388-398.

Schaap, P., and Wang, M. 1986. Interactions between adenosine and oscillatory cAMP signaling regulate size and pattern in *Dictyostelium*. Cell **45:** 137-144.

Schaap, P., Van Lookeren Campagne, M.M., Van Driel, R., Spek, W., Van Haastert, P.J.M., and Pinas, J. 1986. Postaggregative

differentiation induction by cyclic AMP in *Dictyostelium*: intracellular transduction pathway and requirement for additional stimuli. Dev. Biol. **118**: 52-63.

Sharp P.M., and Devine K.M. 1989. Codon usage and gene expression in *D. discoideum*. Nucl. Acid. Res. **17**: 5029-5039.

Schindler, J., and Sussman, M. 1977. Ammonia determines the choice of morphogenetic pathways in *Dictyostelium discoideum*. J. Mol. Biol. **116**: 161-170.

Simon, M.N., Driscoll, D., Mutzel, R., Part, D., Williams, J., and Veron, M. 1989. Overproduction of the regulatory subunit of the cAMP-dependent protein kinase blocks the differentiation of *Dictyostelium discoideum*. EMBO J. **8**: 2039-2043.

Small, N.V., Europe-Finner, G.N., and Newell, P.C. 1986. Calcium induces cyclic GMP formation in *Dictyostelium*. FEBS Lett. **203**: 11-14.

Snaar-Jagalska, B.E., Jokabs, K.H., and Van Haastert, P.J.M. 1988. G-proteins in the signal-transduction pathways of *Dictyostelium discoideum*. Dev. Genet. **9**: 215-226.

Spek, W., Van Drunen, K., van Eijk, R., and Schaap, P. 1988. Opposite effects of adenosine on two types of cAMP-induced gene expression in *Dictyostelium* indicate the involvement of at least two different intracellular pathways for the transduction of cAMP signals. FEBS Lett. **228**: 231-234.

Sternfeld, J., and David, C.N. 1981. Cell sorting during pattern formation in *Dictyostelium discoideum*. Differ. **10**: 1-5.

Subramanian V., Davis, D.L., Seal, S.N., and Burch J.B.E. 1991. Chicken vitellogenin gene-binding protein, a leucine zipper transcription factor that binds to an important control element in the chicken vitellogenin II promoter, is related to rat DBP. Mol. and Cell. Biol. **11**: 4863-4875.

Theibert, A., and Devreotes, P.N. 1984. Adenosine and its derivatives inhibit the cAMP signalling response in *Dictyostelium discoideum*. Dev. Biol. **106**: 166-173.

Theibert, A., and Devreotes, P.N. 1986. Surface receptor-mediated activation of adenylate cyclase in *Dictyostelium*: regulation by guanyl nucleotides, mutant characterization and *in vitro* mutant reconstitution. J. Biol. Chem. **259**: 12318-123231.

Thireos, G., Driscoll Penn, M., and Greer, H. 1984. 5′ untranslated sequences are required for the translational control of a yeast regulatory gene. Proc Natl. Acad. Sci. USA **81**: 5096-5100.

Town, C.D., Gross, J.D., and Kay, R.R. 1976. Cell differentiation without morphogenesis in *Dictyostelium discoideum*. Nature **262**: 717-719.

Tsang A.S., and Bradbury, J.M. 1981. Separation and properties of prestalk and prespore cells of *Dictyostelium discoideum*. Exp. Cell Res. **132**: 433-441.

Tsang, A.S., Mahbubani, H., and Williams, J.G. 1982. Cell-type-specific actin mRNA populations in *Dictyostelium discoideum*. Cell **31**: 375-382.

Tsang, A.S., and Tasaka, M. 1986. Identification of multiple cyclic AMP binding proteins in developing *Dictyostelium discoideum*. J. Biol. Chem. **261**: 10753-10759.

Tsang, A.S., Kay, C.A., and Tasaka, M. 1987. Expression of an altered cAMP binding protein by rapid developing strains of *Dictyostelium discoideum*. Dev. Biol. **120**: 294-298.

Turner, R., and Tjian, R. 1989. Leucine repeats and an adjacent DNA binding domain mediate the formation of functional *c fos-c jun* heterodimers. Science **243**: 1689-1694.

VanHaastert P.J.M., Van Lookeren K., Campagne M.M., and Ross F.M. 1982. Altered cGMP phosphodiesterase activity in chemotactic mutants of *Dictyostelium discoideum*. FEBS Lett. **147**: 149-152

Van Haastert, P.J.M. 1983. Binding cAMP and adenosine derivatives to *Dictyostelium discoideum* cells: Relationship between binding, chemotactic receptor and antagonistic derivatives. J. Biol. Chem. **258**: 9643-9648.

Van Haastert, P.J.M., and De Wit, R.J.W. 1984. Demonstration of receptor heterogeneity and affinity modulation by nonequilibrioum binding experiments. J. Biol. Chem. **259**: 13321-13328.

Van Haarstert, P.J.M., De Wit, R.J.W., Janssens, P.M.W., Kesbeke, F., and DeGoede, J. 1986. G-protein-mediated interconversions of cell-surface cAMP receptors and their involvement in excitation of guanylate cyclase in *Dictyostelium discoideum*. Proc. Natl. Acad. Sci. USA **261**: 6904-6911.

Van Haarstert, P.J.M., Kesbeke, F., Reymond, C.D., Firtel, R.A., Luderus, E., and VanDriel, R. 1987. Abberrant transmembrane signal transduction in *Dictyostelium* cells expresing a mutated *ras* gene. Proc. Natl. Acad. Sci. USA **84**: 4905-4909.

Vardy, P.H., Fisher, L.R., Smith, E., and Williams, K.L. 1986. Traction proteins in the extracellular matrix of *Dictyostelium discoideum* slugs. Nature **320**: 526-529.

Vaughan, R.A., and Devreotes, P.N. 1989. Ligand-induced phosphorylation of the cAMP receptor from *Dictyostelium discoideum*. J. Biol. Chem. **263:** 14538-14543.

Wang, M., Aerts, R.J., Spek, W., and Schaap, P. 1988. Cell cycle in *Dictyostelium discoideum* is correlated with the expression of cyclic-AMP signal production, detection and degradation. Dev. Biol. **128:** 72-77.

Wang, M., and Schaap P. 1989. Ammonia depletion and DIF trigger stalk cell differentiation in intact *Dictyostelium discoideum* slugs. Development **105:** 569-574

Weijer, C.J., Duschl, G., and David, C.N. 1984. Dependence of cell-type proportioning and sorting on cell cycle phase in *Dictyostelium discoideum*. J. Cell. Sci. **70:** 133-145.

Weijer, C.J., and Durston, A.J. 1985. Influence of cAMP and hydrolysis products on cell type regulation in *Dictyostelium discoideum*. J. Embryol. Exp. Morph. **86:** 19-37.

Wurster, B., Schubigerm, K., Wick, U., and Gerisch G. 1977. cGMP in *Dictyostelium discoideum*: Oscillations and pulses in response to folic acid and cAMP signals. FEBS LETT. **76:** 141-144.

Yamamoto, K.K., Gonzalez, G.A., Biggs III, W.H., and Montminy M.R. 1988. Phosphorylation-induced binding and transcriptional efficiency of nuclear factor CREB. Nature **334:** 494-498.