Volume 15 Number 18 1987 Nucleic Acids Research

Structural and functional characterization of a *Dictyostelium* gene encoding a DIF inducible, prestalk-enriched mRNA sequence

A.Ceccarelli*, S.J.McRobbie, K.A.Jermyn, K.Duffy, A.Early and J.G.Williams

Imperial Cancer Research Fund, Clare Hall Laboratories, Blanche Lane, South Mimms, Potters Bar, Hertfordshire EN6 3LD, UK

Received July 24, 1987; Revised and Accepted August 20, 1987

ABSTRACT

The pDd56 mRNA sequence is highly enriched in prestalk over prespore cells and is inducible by DIF, the putative Dictyostelium stalk-specific morphogen. We show that the pDd56 gene is composed of forty one copies of a twenty four amino acid, cysteine rich repeat. This is highly homologus to a repeat which we have previously shown to compose the major fraction of the pDd63 mRNA, another DIF inducible, prestalk-enriched sequence. The predicted pDd56 protein contains a putative signal peptide but does not appear to contain a transmembrane segment. In combination these features suggest it to be an extrinsic protein and we confirm this elsewhere by showing that the pDd56 gene encodes a known, extracellular protein of the stalk. The pDd56 mRNA is dependent upon exogenous DIF for its accumulation. We show that this control is exerted at the transcriptional level and that a restriction fragment containing 1.7Kb of upstream sequence directs temporally-regulated expression of the gene.

INTRODUCTION

In the Dictyostelium culminant a stalk, composed of vacuolated, cellulose-ensheathed cells, supports a mass of spores. This differentiated structure derives from an initially uniform population of amoebae which aggregate in response to pulsatile emissions of cAMP from a signalling centre. The aggregate normally forms a migratory structure, the slug, which enters culmination only under environmental conditions appropriate for spore release. The anterior one-fifth of the slug is composed of cells which will ultimately differentiate to form stalk cells. The posterior four-fifths of the slug is predominantly composed of presumptive spore cells but contains a small proportion of cells, with the morphological and biochemical characteristics of prestalk cells, called anterior like cells (1, 2).

In addition to providing a convenient source of stalk and spore-cell precursors, the slug itself constitutes a differentiated structure which can be used to investigate pattern

© IRL Press Limited, Oxford, England.

formation. It is, however, difficult to study the signals inducing cellular differentiation in a multicellular structure and a major advance has been the development of an *in vitro* system, where isolated cells can be stimulated to differentiate down the stalk or spore-cell pathway (3). Using this system, a substance termed DIF has been identified (3, 4, 5, 6, 7). This is a small, hydrophobic molecule, (5) which will induce isolated amoebae to form stalk cells. A mutant strain HM44, which is defective in DIF production, synthesizes prespore markers but fails to accumulate a prestalk-specific marker unless DIF is added (8). These observations suggest that DIF acts as a stalk-specific morphogen.

In order to analyze the earliest events in cellular differentiation it is necessary to identify markers of the two cell types. This has been relatively straightforward for the spore cell pathway. There is a high degree of continuity of gene expression between prespore and spore cells and reliable morphological, biochemical and mRNA markers have been identified (reviewed by, 9, 10, 11). However, the great majority of prestalk-specific proteins are not present in mature stalk cells. A number of cDNA clones, derived from prestalk-enriched mRNA sequences, have been isolated (12, 13) but we have recently shown that these are not reliable markers of prestalk-cell differentiation (14). They are all expressed at a finite level in prespore cells and probably fall into the class of markers which are initially expressed in all cells and then selectively lost from prespore cells (15, 16).

We have isolated two cDNA clones, termed pDd56 and pDd63, derived from DIF-inducible mRNA sequences (14). These are much more highly enriched in prestalk over prespore cells than previously described "prestalk" mRNA sequences and they are possibly completely prestalk-specific. In the DIF-defective mutant, HM44, the pDd56 and pDd63 mRNA sequences are dependent upon DIF for their accumulation. The major rise in DIF concentration occurs after the tipped aggregate stage of development (17) and we find the pDd56 and pDd63 mRNA sequences both become detectable at this stage. Consistent with the relative kinetics of accumulation after DIF induction *in vitro*, the pDd63 gene is maximally expressed in the slug and the pDd56 gene is maximally expressed at the "mexican-hat" stage of culmination (18, 14).

These two sequences are of considerable interest, both as the first authentic mRNA markers of prestalk cell differentiation and because they can be used to investigate the mode of action of DIF. We have recently shown (McRobbie et al., in preparation) that the pDd56 and pDd63 genes respectively encode the ST310 and ST430, stalk-specific proteins identified by Morrissey et al., (19). We have previously determined a part of the sequence of the pDd63 gene and shown it to be predominantly composed of a tandemly repeated 24 amino acid sequence (18). We have now determined the entire sequence of the pDd56 gene and we compare it with the pDd63 gene. We have also investigated the regulation of the pDd56 mRNA by assaying nuclear transcription in vitro and by analyzing a DNA transformant containing a 5' proximal segment of the gene.

MATERIALS AND METHODS

Growth and Development of Cells

D.discoideum, strain Ax-2 was grown axenically in suspension (20). HM44 was grown in association with Klebsiella aerogenes on SM nutrient agar (21). Cells were normally harvested and washed by centrifugation in 20mM KH_2PO_4/K_2HPO_4, pH6.0 (KK_2). Ax-2 cells were developed on Millipore filters (21). HM44 cells were induced according to the in vitro development conditions of Kopachik et al. (22) with the modifications of Williams et al. (18).

Bacterial Cloning and Sequence Analysis

A genomic library of Dictyostelium DNA fragments was prepared in the BamH1 site of pAT153 (23). The DNA was partially digested with Sau 3AI and fragments of 3 to 7Kb were isolated by centrifugation through a sucrose gradient. The library was screened with the entire insert from the pDd56 cDNA clone (14) and the pDd2052 clone was isolated. For sequence analysis, the insert from pDd56 was re-cloned from pUC8 (24) into the pEMBL18+ vector (25) and a uni-directional deletion series was generated. Single-stranded DNA preparations, prepared by super-infection with phage F1, were then subjected to sequence analysis by the di-deoxy procedure (26). The nucleotide sequence of selected regions of the genomic clone was determined, in part from a deletion series in a pEMBL vector and in part by the chemical degradation procedure (27). The start-site of transcription was determined with a synthetic oligo-nucleotide using the conditions of

hybridization and extension described by Devine et al. (28). The oligonucleotide was 22 nucleotides in length with a GC content of 36%, and was annealed to the RNA at 55°C.

Analysis of gene expression

Nuclear transcription in vitro and analysis of the synthesized RNA, was performed as described previously (18). The conditions for transformation of the pDd56-10 construct into the Ax2 strain of Dictyostelium were slightly modified (29) from those described by Nellen et al (30). After selection on plates in G418 (10 μg/ml), approximately 100 individual transformed clones were pooled and grown in suspension in the same concentration of drug. The cells were then subjected to development on agar and total cellular RNA was isolated and analyzed by Northern transfer as described previously (18).

RESULTS

The structure of the pDd56 gene

The pDd56 mRNA is 3.4 Kb in length and the pDd56 cDNA clone contains a 2.4Kb segment derived from its 3′ proximal region (14). The insert from the cDNA clone was used as a probe to isolate the cognate gene from a plasmid genomic library. The pDd2052 clone contains the entire pDd56 gene with several Kb of 5′ and 3′ flanking sequence (Figure 1). The nucleotide sequences of the insert in the cDNA clone, and of the region of the genomic clone proximal to the 5′ end of the mRNA, were determined. The start-site of transcription was established, by primer-extension using a synthetic oligonucleotide (data not shown). The first potential initiation codon lies 143 nucleotides downstream of the major cap site. The 5′ non-coding region (data not shown) displays the extremely high AT content typical of the non-coding regions of Dictyostelium genes (31). Restriction-mapping of the genomic clone, and comparison with the cDNA clone, indicated the possible existence of two introns and partial nucleotide sequence analysis of the genomic clone confirmed these to be present, both introns showing a typical very high AT content (31). The first intron is very near to the 5′ end of the coding region and the second intron is located 1.1Kb downstream (Figure 1 and Figure 2).

The amino acid sequence of the pDd56 protein

The first ATG initiation codon defines an open reading frame of 3,138 nucleotides. The predicted sequence of the protein

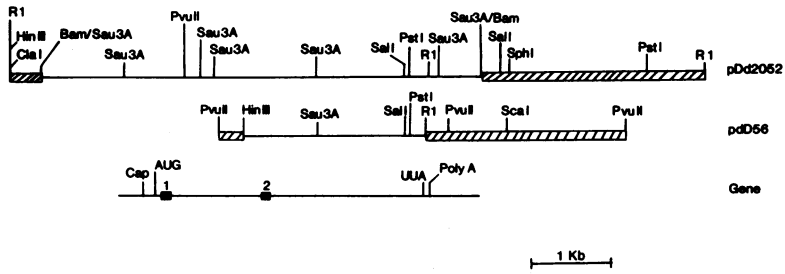

Fig. 1 The structure of the pDd56 gene. The pDd56 clone contains a cDNA insert of 2.4Kb in length inserted into the Pst1 site of pUC8 by homopolymeric dG.dC tracts. The pDd2052 genomic clone contains a partial Sau3AI restriction fragment inserted into the BamH I site of pAT153. The entire nucleotide sequence of the insert in the pDd56 cDNA clone was established as described in the methods section. The region of the pDd2052 genomic clone upstream of the 5' terminus of the cDNA clone and extending to the Sau3AI site upstream of the cap site, was also determined. The position of the major cap site was established by primer extension using a synthetic oligonucleotide complementary to a region between nucleotides 84 and 105 (relative to the ATG initiation codon).

High resolution restriction mapping, of the genomic and cDNA clone, suggested the existence of an intron (intron 2) downstream of the 5' proximal region for which sequence was initially established and this was confirmed by further sequence analysis. Because we have not determined the entire nucleotide sequence of the genomic clone, we cannot rule out the presence of other introns but, given the high resolution of the restriction-mapping we performed, these would have to be less than 25-50 nucleotides in length. The position of the 3' end of the gene was deduced from the cDNA clone and the 3' non-coding region is estimated to be 67 nucleotides in length. This region is, however, predominantly comprised of long tracts of A and T residues, hence we cannot rule out the possibility that the oligo dT priming used to generate the cDNA clone occurred upstream of the authentic poly(A) tail. Vector sequences in both constructs are indicated by cross-hatching and introns are represented by filled boxes.

between the initiation codon and intron 1 displays the characteristic features of a signal peptide, with fourteen hydrophobic or non-polar amino acids preceded by two charged residues Asn and Lys (Figure 2). Within approximately ten amino acids downstream of the position of intron 1, the sequence begins to display a repetitive pattern which extends through most of the length of the protein. This 24 amino acid, cysteine-rich repeat contains ten residues (indicated by an asterisk below the sequence in Figure 2) which are almost completely invariant. They are present in at least 36 out of the 41 copies of the repeat and one

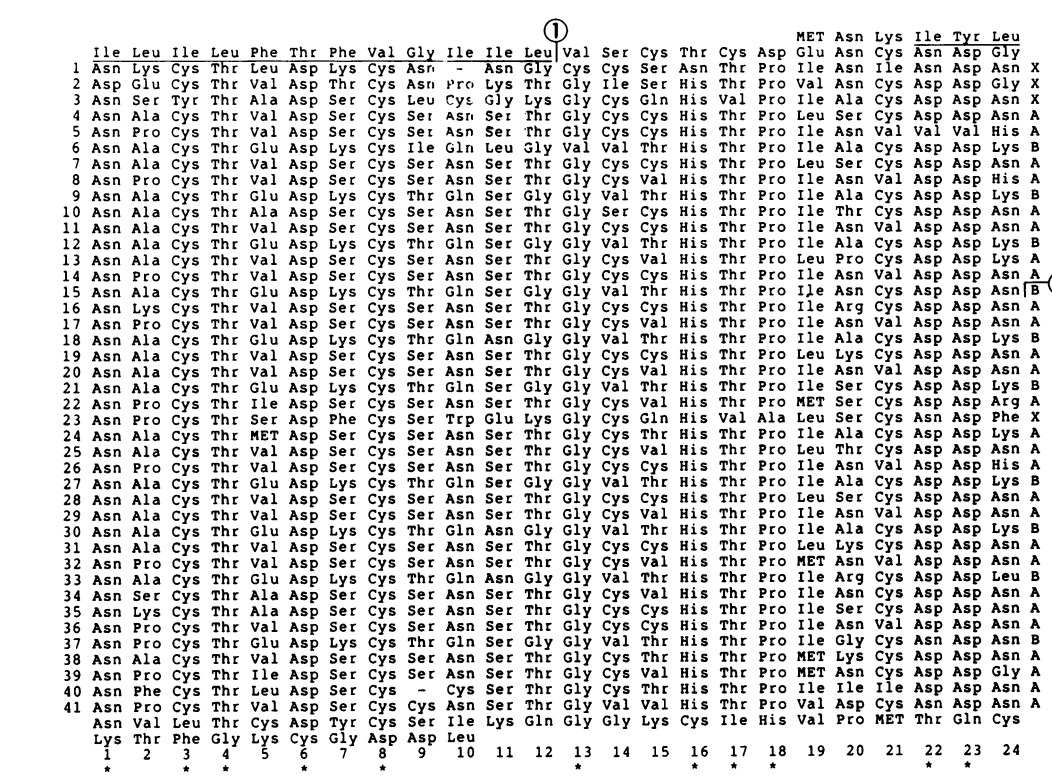

Fig. 2 The complete amino acid sequence of the pDd56 protein. The amino acid sequence was predicted from the nucleotide sequence of the pDd56 cDNA and pDd2052 genomic clones. In order to obtain an optimal alignment of the repeats, two deletions have been posited, labelled (-) in repeats 1 and 40. The underlined portion at the N-terminus is a putative signal peptide. The positions at which intron 1 (1) and intron 2 (2) interrupt the coding region are indicated. The residues marked with an asterisk are those present in >90% of the copies of the repeat.

of the ten amino acids (Cys at position 8) is present in every copy. Based upon the other, more variable, residues the repeats fall into two types, which we term A and B. The consensus sequence of an A and a B repeat is shown in Figure 3A. A comparison with the consensus repeat sequence of the related gene, pDd63, (18) indicates that the pDd63 repeat has much more homology with the type A than the type B repeat (Figure 3A). The position of these repeats within the pDd56 protein is shown on the right in Figure 2 and is represented linearly in Figure 3B. They are

A

* Asn	Pro Ala Lys	* Cys	Thr	Val Ala	* Asp	† Ser	* Cys	† Ser	Asn	Ser	Thr	* Gly	† Cys	Cys Val	* His	* Thr	* Pro	Leu Ile	† Asn Ser	Cys Val	* Asp	* Asp	* Asn Lys	pDd 56 "A"
Asn	Ala	* Cys	Thr	Glu	Asp	† Lys	* Cys	† Thr	Gln	Ser	Gly	* Gly	† Val	Thr	* His	* Thr	* Pro	Ile	† Ala	Cys	* Asp	* Asp	Lys	pDd 56 "B"
* Asn	Pro Lys Ala	* Cys	Thr	Ile Val	* Asp	Ser Ala	* Cys	Ser Asp	Lys	Ser	Thr	* Gly	Cys Val	Thr Cys	His Asn	* Thr	* Pro	Ile Val	Asp Ser	Cys Val	* Asp	* Asp	Asn Lys	pDd 63

B

X X X A A B.A A B.A A B.A A B.A A B.A A B.A X A A A B.A A B.A A B.A A A.B A A A A

Fig. 3 (A) Consensus sequences for the A and B repeats of pDd56 and comparison with the pDd63 repeat. The residues marked with an asterisk are "invariant" residues, defined as being present in >90% of copies of the repeat (Figure 2). The residues marked with a cross (†) are diagnostic of an A or a B repeat. (B) Linear representation of the organization of repeats in the pDd56 gene. The repeats labelled X, show no strong bias towards either the type A or type B repeat. The internal X repeat is unusually diverged throughout. The arrows over the sequence indicate the start points of two regions of the gene which are highly homologous to each other at the nucleotide level. Optimal alignment of these two regions, which involves omitting two copies of repeat in the N terminal half of the molecule, shows 87% homology between them (data not shown).

organized in the pattern AAB throughout most of the protein. This pattern breaks down only at the N and C termini and at one point near the middle of the sequence where there are two A repeats flanking a repeat which we have termed X, since it displays no strong homology to either the A or B form of the repeat and contains three substitutions of the ten "invariant" residues. This repeat lies near to the approximate centre of two tandemly duplicated regions (indicated by arrows in Figure 3B) between which there is a very high overall degree (87%) of nucleotide conservation (data not shown).

The pDd56 protein has the potential for very heavy N-glycosylation, with 26 copies of the modification signal Asn-Ser-Thr. Every A repeat, apart from an atypical A repeat near the C terminus, contains this signal at residues 10-12 (Figure 2).

Induction of the pDd56 mRNA by DIF is regulated at the transcriptional level

In the mutant strain HM44, which is defective in DIF production, accumulation of the pDd56 mRNA is totally dependent upon exogenous DIF. In the standard DIF induction assay HM44 cells are incubated for 10 hr in the presence of cAMP only and then for a further 12 hr in the presence of cAMP and DIF. Morphologically recognizable stalk cells appear about 6 hr after the addition of DIF. The pDd56 mRNA displays sigmoidal kinetics of accumulation (14). A low level of mRNA accumulates in the first few hours of induction but the major rise in concentration does not occur until four hours after the addition of DIF. In contrast, the pDd63 mRNA displays linear kinetics of induction; the major rise in concentration occurring within two hours of the addition of DIF (14). We have previously shown that induction of the pDd63 mRNA is regulated at the transcriptional level. There is a detectable rise in the rate of nuclear "run-on" transcription within 15 minutes of the addition of DIF, with maximal transcription occurring 90 minutes after induction (18). We have performed a similar analysis for pDd56 and find it also to be transcriptionally induced (Figure 4). The kinetics for the induction of pDd56 gene transcription fit very well with the pattern of mRNA accumulation, the maximal rate of transcription occuring four hours after the addition of DIF.

Analysis of a Dictyostelium transformant containing a pDd56 fusion gene

We have initiated identification of the DNA sequence elements responsible for cell-type specific gene expression, by analyzing the expression of a fusion gene containing 5′ proximal sequences of the pDd56 gene. The pDd56-10 construct was created in the G418 resistance vector pB10TP10 (29), a derivative of the B10 vector of Nellen et al., (30). The pDd56-10 construct contains 1.7 Kb of upstream sequence, and 500 nucleotides of the coding region of pDd56, fused to the 3′ proximal portion of the pDd10 gene. This is a constitutively transcribed Dictyostelium gene of unknown function which, in this construct, provides termination and polyadenylation signals for the pDd56 gene (Figure 5; and 29).

A transformed population of cells containing the pDd56-10 construct was subjected to development on filters and total cellular RNA was extracted. The endogenous gene is maximally

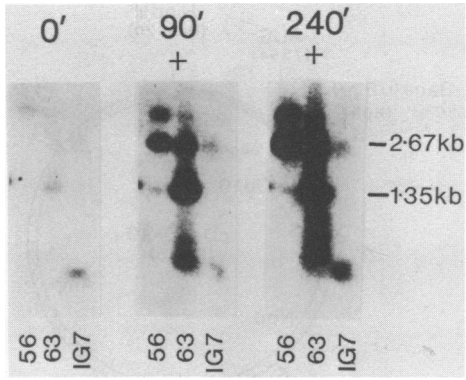

Fig. 4 The induction of pDd56 gene transcription by DIF. A DIF induction was performed using HM44 cells and at the times indicated nuclear "run-on" transcription was performed. The labelled transcripts were hybridised to Southern blots of digested plasmid DNA. The digests enabled internal insert fragments to be separated from vector DNA. pDd2052, a genomic clone of 56, was digested with PvuII and SalI to give an internal fragment of 2.67kb. The pDd63 cDNA clone was digested with PvuII and ClaI, yielding an internal fragment of 1.35Kb. IG7, a cDNA clone expressed constitutively during development was included as a control for the hybridization.

expressed at the mexican hat stage of culmination, there being a somewhat lower level of mRNA in the slug and a vanishingly low level in vegetative cells (14). This was confirmed for these RNAs by probing the filters with the pDd56 cDNA clone (Figure 6, Panel A). We do not detect a signal from the fusion mRNA, presumably because there is only a very small portion of the pDd56 coding region in the fusion gene and the probe derives from downstream of this region. The failure to detect the fusion RNA could also be in part due to a low level of the fusion RNA relative to the authentic pDd56 transcript. There are several precedents for fusion transcripts containing the pDd10 sequence being relatively unstable (D. Driscoll, C.J. Pears and J.G. Williams, unpublished data). As an alternative method of detection we used a synthetic oligonucleotide containing vector-derived sequences located between the pDd56 and pDd10 sequences. Using this probe, we detect transcripts from the fusion gene which are of the expected size (1.2Kb) and which are precisely co-regulated with transcripts from the endogenous gene (Figure 6, Panel B).

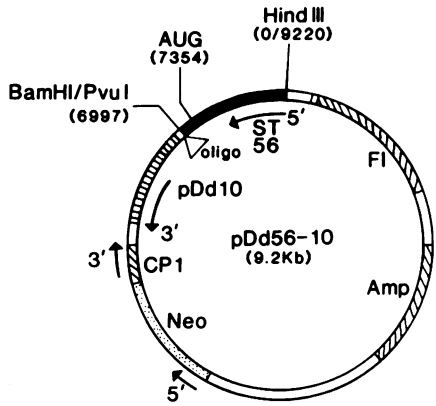

Fig. 5 The structure of the pDd56-10 fusion gene. The pB10TP1 vector contains a Neomycin-Actin fusion gene inserted into a pEMBL vector with a modified multilinker. Termination and polyadenylation signals for the G418 resistance gene derive from the Cysteine Proteinase 1 (CP1) gene of Dictyostelium. The pDd10 sequences derive from a gene encoding a constitutively transcribed mRNA of unknown function and is convergently transcribed with respect to the CP1 gene. Transcription termination and 3' processing signals must, therefore, lie within this fragment and 5' proximal fragments fused to the pDd10 gene will generate a fusion mRNA. The derivative of pB10TP1 used in this construction, pB10TP10, contains a synthetic "marker" oligonucleotide of 33 residues in the length inserted into the unique Xba1 site. A HindIII-PvuII fragment from pDd2052 (Figure 1), of 2.2kb in length, was inserted into pB10TP10. The vector was prepared for cloning by cleavage with BamH1, end-filling using the Klenow fragment of DNA polymerase 1, and then cleavage with HindIII. The ATG initiation codon of the pDd56 gene lies 357 nucleotides upstream of the PvuII site and the fusion is in frame with the pDd10 coding sequence. The major start-site of transcription of the pDd56 gene is 502 nucleotides upstream of the PvuII site and the poly(A) addition site of the pDd10 gene lies 690 nucleotides downstream of the BamH1 site. The resultant fusion mRNA is therefore expected to be 1,190 nucleotides in length.

DISCUSSION

We show that the pDd56 and pDd63 genes clearly belong to the same gene family. In both genes, the putative N-terminal leader sequences are bounded on the terminal side by introns (18) and Figure 2). There is no obvious homology between the sequence of the two N-terminal peptides but the repeat in pDd56 is highly homologous to the pDd63 repeat, with nine of the ten "invariant" residues in the pDd56 repeat also being invariant in pDd63; i.e. present in >90% of copies (Figure 3A). The sole exception is His 16, which is either His or Asn in the pDd63 repeats.

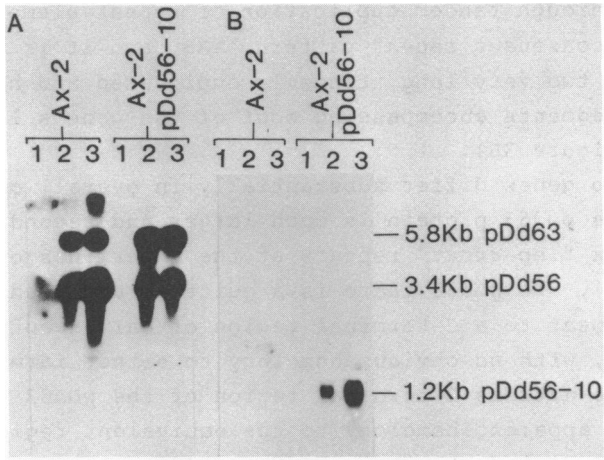

Fig. 6 Analysis of the accumulation of the pDd56-10 fusion mRNA by Northern Transfer. A pooled population of transformants, derived from approximately 100 individual G418 resistant colonies, was subjected to development on agar. Total cellular RNA was extracted from vegetative cells (Lanes 1), from slugs (Lanes 2) and culminants at the mexican hat stage (Lanes 3). Aliquots (5 μg) of the RNA were electrophoresed through a denaturing gel, transferred to nitrocellulose and detected by hybridization with the insert from the pDd56 cDNA clone (Panel A) or with a synthetic oligonucleotide (Panel B). The oligonucleotide derives from a region of the vector lying between the ST56 and pDd10 sequences (Fig. 5). It is 33 nucleotides in length and has a GC content of 33%. Hybridization was performed in 6xSSC at 32°C and the filter was washed in 6xSSC at 37°C. The lanes marked Ax2 contain RNA from untransformed, control cells and the lanes marked Ax2, pDd56-10 from transformants expressing the fusion mRNA. Using the pDd56 cDNA probe at this stringency of washing (3xSSC, 50% formamide 37°C) there is cross hybridisation to the pDd63 mRNA.

Interestingly, intron 2 in both genes lies in an identical position relative to the frame of the repeats (Figure 2 and (18), although in the pDd63 gene it is located much nearer the N-terminus. This suggests that the intron was present in a gene ancestral to both pDd56 and pDd63. Hence we have used the position of this intron to set the frame of the repeat (i.e. Asn1 of the repeat is positioned to the immediate C-terminal side of the intron).

The consensus repeat in pDd63 shows a much closer fit to the type A repeat of pDd56 than to the B repeat (Figure 3A), hence the ancestral gene presumably contained repeats of the A type. There is particularly clear evidence that evolution of the pDd56 gene

proceeded through tandem duplication of repeat elements. It contains a consensus repeat pattern, AAB, and it is possible to distinguish two very long, tandemly duplicated and highly conserved segments encompassing most of the gene's length (see legend to Figure 3B).

The two genes differ substantially in overall organization, firstly, the pDd63 protein is much larger and secondly there are at least six "imperfect" repeats at the C-terminus of the pDd63 protein (18). In pDd56 there is a quite abrupt transition from the last repeat to a C-terminal region of only twenty to thirty amino acids, with no obvious homology to either form of the repeat. The extreme C-terminal region of the pDd63 gene does not display any apparent homology to the equivalent region in pDd56.

The predicted structure of the pDd56 polypeptide is consistent with its being an extracellular protein. The N terminal leader indicates that this a non-cytosolic protein. The absence of an anchor segment and the presence of multiple potential glycosylation sites suggest that it might be extracellularly located. A high cysteine content is also often a feature of extracellular protein domains (32, 33, 34, 35, 36) and a highly re-iterated repeat is a characteristic feature of a number of extracellular proteins of lower eukaryotes (37, 38, 39, 40). We confirm this elsewhere (McRobbie et al., in preparation) by showing that the pDd56 gene encodes the ST310 polypeptide (19) which appears to be an extracellular stalk protein (J.S. Wallace and P.C. Newell, personal communication)

The accumulation of developmentally regulated mRNA sequences in Dictyostelium may be regulated at the transcriptional (41, 42, 43) or the post-transcriptional level (44, 45, 46). We have previously shown, for the pDd63 gene (18), that DIF induces an elevated rate of transcription in isolated nuclei and we obtain similar results here for the pDd56 gene. This suggests that the control of gene expression on the prestalk-stalk pathway may be primarily mediated at the transcriptional level. It will be of interest to identify the cis-acting control sequences which confer DIF-inducibility. As a first step towards this, we have shown that a fragment containing 1.7 Kb of sequence upstream of the start site of transcription directs correctly regulated gene transcription. By analyzing in detail the DIF-responsive elements

and the trans-acting factors which recognize them we hope to elucidate the mechanism whereby DIF induces prestalk-cell differentiation.

*Present address: The Hygiene Institute, Cattedra di Biologie Generale, Via Santena 5 bis, Torino 10126, Italy

REFERENCES

1. Sternfeld, J. and David, C.N. (1982) Dev. Biol. **93**, 111-118.
2. Devine, K.M. and Loomis, W.F. (1985) Dev. Biol. **107**, 364-372.
3. Town, C.D., Gross, J.D. and Kay, R.R. (1976) Nature **262**, 717-719.
4. Town, C.D. and Stanford, E. (1979) Proc. Natl. Acad. Sci. U.S.A. **76**, 308-312.
5. Gross, J.D., Town, C.D., Brookman, J.J., Jermyn, K.A., Peacey, M.J. and Kay, R.R. (1981) Phil. Trans. R. Soc. Lond. B. **295**, 497-508.
6. Kay, R.R., Garrod, D. and Tilly, R. (1978) Nature, **271**, 58-60.
7. Kay, R.R. and Jermyn, K.A. (1983) Nature **303**, 242-244.
8. Kopachik, W., Oohata, A. Dhokia, B., Brookman, J.J. and Kay, R.R. (1983) Cell **33**, 397-403.
9. Morrissey, J.H. (1982) In Loomis, W.F. (ed.), The development of Dictyostelium discoideum. Academic Press, New York, pp411-449.
10. Loomis, W.F. (1982) Dev. Biol. **93**, 279-284.
11. Williams, J.G., Pears, C.J., Jermyn, K.A., Driscoll, D.M., Mahbubani, H. and Kay, R.R. (1986) In Booth, I. and Higgins, C. (eds.), Symposium of the Society for General Microbiology. Regulation of Gene Expression. Cambridge University Press, pp277-298.
12. Barklis, E. and Lodish, H.F. (1983) Cell **32**, 1139-1148.
13. Mehdy, M.C., Ratner, D. and Firtel, R.A. (1983) Cell **32**, 763-771.
14. Jermyn, K.A., Berks, M., Kay, R.R. and Williams, J.G. (1987) Development, in press.
15. Tasaka, M., Noce, T. and Takeuchi, I. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 5340-5344.
16. Krefft, M., Voet, L., Gregg, J.H. and Williams, K.L. (1985) J. Embryol. Exp. Morph. **88**, 15-24.
17. Brookman, J.J., Town, C.D., Jermyn, K.A. and Kay, R.R. (1982) Dev. Biol. **91**, 191-196.
18. Williams, J.G., Ceccarelli, A., McRobbie, S., Mahbubani, H., Kay, R.R., Early, A., Berks, M. and Jermyn, K.A. (1987) Cell, **49**, 185-192.
19. Morrissey, J.H., Devine, K.M. and Loomis, W.F. (1984) Dev. Biol. **103**, 414-424.
20. Watts, D.J. and Ashworth, J.M. (1970) Biochem. J. **119**, 171-174.
21. Sussman, M. (1966) Methods in Cell Physiol. **2**, 397-410.
22. Kopachik, W., Dhokia, B. and Kay, R.R. (1985) Differentiation **28**, 209-216.
23. Twigg, A.J. and Sherrat, D. (1980) Nature **283**, 216-218.
24. Vieira, J. and Messing, J. (1982) Gene **19**, 259-268.
25. Dente, L. Cesareni, G. and Cortese, R. (1983) Nucl. Acids Res. **11**, 1645-1655.

26. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A., **74**, 5463-5467.
27. Maxam, A.M. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 560-564.
28. Devine, J.M., Tsang, A.S. and Williams, J.G. (1982) Cell **28**, 793-800.
29. Early, A. and Williams, J.G. (1987) Gene, in press.
30. Nellen, W., Silan, C. and Firtel, R.A. (1984) Mol. Cell Biol. **4**, 2890-2898.
31. Kimmel, A. and Firtel, R.A. (1983) Nucl. Acids Res. **11**, 541-551.
32. Petersen, T.E., Thogesen, H.C., Skortesgaard, K., Vive-Pederson, K., Sottrup-Jensen, L. and Magnusson, S. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 137-141.
33. Wharton, K.A., Johansen, K.M., Xu, T. and Artavanis- Tsakonas, A. (1985) Cell **43**, 567-581.
34. Greenwald, I. (1985) Cell, **43**, 583-590.
35. Ullrich, A., Coussens, L., Hayflick, J.S., Dull, T.J., Gray, A., Tam, A.W., Lee, J., Yardon, Y., Libermann, T.A., Schlessinger, J., Downward, J., Mayes, E.L.V., Whittle, N., Waterfield, M.D. and Seeburg, P.H. (1984) Nature 309, 418-425.
36. Ullrich, A., Bell, J.R., Chen, E.Y., Herrera, R., Petruzelli, L.M., Dull, T.J., Gray, A., Coussens, L., Liao, Y.C., Tsubokawa, M., Mason, A., Seeberg, P.H., Grunfeld, C., Rosen, D.M. and Ramachandran, J. (1985) Nature **313**, 756-761.
37. Ozaki, L.S., Svec, R. Nussenzweig, R.S. Nussenzweig, V. and Godson, G.N. (1983) Cell **34**, 815-822.
38. Prat, A., Katinka, M. Caron, F. and Meyer, E. (1986) J. Mol. Biol. **189**, 47-60.
39. Dowds, B.C.A. and Loomis, W.F. (1986) Biochem. Biophys. Res. Commun. **135**, 336-339.
40. Kochan, J., Perkins, M. and Ravetch, J.V. (1986) Cell **44**, 689-696.
41. Williams, J.G., Lloyd, M.M. and Devine, J.M. (1979) Cell, **17**, 903-913.
42. Rowekamp, W. and Firtel, R.A. (1980) Dev. Biol. **79**, 409-418.
43. Landfear, S.M., Lefebvre, P., Chung, S. and Lodish, H.F. (1982) Mol. Cell Biol., **2**, 1417-1426.
44. Chung, S., Landfear, S.M., Blumberg, D., Cohen, N.S. and Lodish, H.F. (1981) Cell **24**, 785-797.
45. Mangiarotti, G., Lefebvre, P. and Lodish, H.F. (1982) Dev. Biol. **89**, 82-91.
46. Mangiarotti, G., Ceccarelli, A. and Lodish, H.F. (1983) Nature **301**, 616-618.