An 80-bp cis-acting regulatory region controls cAMP and development regulation of a prestalk gene in Dictyostelium

Sumana Datta1 and Richard A. Firtel2

Department of Biology, Center for Molecular Genetics, M-034, University of California, San Diego, La Jolla, California 92039 USA

We have analyzed an 80-bp region containing the cis-acting sequences necessary for regulation of the Dictyostelium discoideum prestalk gene pst-cathepsin at the appropriate stage during multicellular development and in response to cAMP in single-cell culture. The region lies between approximately −280 and −200 bp upstream from the Cap site and contains two intertwined G/C-rich sequences, including a palindromic sequence and a direct repeat of the 3' portion of the palindrome. In a previous set of experiments, we showed that the direct repeat, or G-box, is important in the regulation of pst-cath expression. In this paper, we use a series of nested internal deletions to define other regions within the promoter required for cAMP and developmental expression, to further examine the importance of the two G-boxes, and to examine the functional relationship of the G-boxes with respect to the other regulatory sequences. Analysis of the expression of these constructs in transformed cells showed that both the 5' portion of the palindrome and the two G-boxes are required for promoter function and are capable of developmentally regulating pst-cath expression. An A/T-rich sequence located 5' to the G/C-rich sequence is also essential for maximal expression, whereas insertion of a linker 5' to this region suggests the presence of a negative element functional during multicellular development. The spacing between the G-box sequences proved to be important for the full induction of gene expression. Constructs containing the G-boxes at wild-type spacing or closer show wild-type or near wild-type levels of expression, whereas expansion of the region between the G-boxes leads to a substantial drop in the level of gene expression in response to cAMP. Insertion of an oligonucleotide containing one of the G-boxes and surrounding sequences can partially complement deletions in which this region has been removed. Analysis of the expression of the cassette constructs, in some cases, revealed significant differences in the quantitative level of expression under the two developmental conditions. This suggests that there are either qualitative or quantitative differences in the factors controlling the expression of this gene under these two conditions.

[Key Words: Dictyostelium; prestalk gene; gene expression; regulatory sequences; cAMP]

Received November 18, 1987; revised version accepted February 4, 1988.

Vegetative Dictyostelium discoideum cells grow as free-living single-celled amoebas on a food source. When starved, they initiate a developmental program that culminates in a multicellular fruiting body consisting primarily of two cell types, stalk cells and spores. Precursors to both cell types, prestalk and prespore cells, can be identified by cell-type-specific markers earlier in development. Multicellular differentiation is initiated when starvation induces a pulsatile relay system of the chemoattractant cAMP. In response to the cAMP pulses, amoebas stream together to form aggregates of ~10^5 cells. At ~10 hr after the onset of development, prestalk-specific mRNAs, markers for anterior cell differentiation, are induced by the rising levels of cAMP within the aggregates. By 13–14 hr, a tipped aggregate has formed, and prespore mRNAs are now expressed. By 15–16 hr, an additional set of mRNAs induced by the morphogen DIF are expressed in the prestalk cell population. The two precursor cell types become spatially organized within the organism, with prestalk cells comprising the anterior 20% of the slug and prespore cells comprising the posterior 80% of the migrating pseudoplasmodium. Approximately 24–26 hr after starvation, terminally differentiated stalk cells and spores can be found in a culminant or fruiting body [see Loomis 1975, 1982; Gomer et al. 1985; Williams et al. 1986, 1987; Datta et al. 1987; Mann et al. 1987].

A number of genes have been isolated that are differentially regulated through development and cell-type specific in their expression [Mehdy et al. 1983; Barklis and Lodish 1983; Chisholm et al. 1983; Dowds and Loomis 1984; Reymond et al. 1984; Pears et al. 1985; Datta et al. 1986; Gomer et al. 1986a; Williams et al. 1987]. One of the sequences isolated by Mehdy et al.

1Present address: Department of Biology, Kline Tower, Yale University, New Haven, Connecticut 06520 USA.
2Corresponding author.
(1983) encodes a cathepsin-like cysteine proteinase preferentially found in prestalk cells and has therefore been designated \textit{pst-cath} [Datta et al. 1986; Datta and Firtel 1987]. [The same gene has also been isolated by Williams and coworkers [Pears et al. 1985] and named cysteine proteinase 2.] \textit{pst-cath} protein and message are preferentially localized in the prestalk population, and nuclear run-on experiments indicate that regulation occurs at the level of transcription [A. Hjorth and R.A. Firtel, M. Levitan and R.A. Firtel, both unpubl.]. \textit{pst-cath} is inducible in single-cell culture by cAMP, and this induction is mediated via the cell-surface cAMP receptor [Mehdy et al. 1983; Mehdy and Firtel 1985; Pears et al. 1985; Gomer et al. 1986b].

In previous studies, we demonstrated that when the \textit{pst-cath} promoter fused to the \textit{Escherichia coli/B-glucuronidase} (\textit{pst-cath/B-gluc} construct pSD1) was introduced into \textit{Dictyostelium} cells, the construct showed appropriate temporal, spatial, cell-type, and cAMP regulated expression [Datta et al. 1986]. In that work, the \textit{pst-cath/B-gluc} fusion protein was shown to have the same spatial localization within the prestalk cell zone as the endogenous \textit{pst-cath} protein. We also showed that deletions to -792 and -313 nucleotides showed the same qualitative levels of expression, corrected for copy number of the vector, and that the deletions show the same qualitative pattern of expression with regard to cAMP induction in single cells and timing of expression relative to morphological stages of development [Mann et al. 1987; Datta 1987]. In subsequent experiments, we identified a G/C-rich palindromic sequence within the \textit{pst-cath} promoter that is followed by a repeat of the 3' half of the palindrome, denoted the G-box. A series of 5' and 3' deletions were combined to form internal deletions, duplications, and linker scanner mutations, which were used to delineate the function of this G-box sequence in the cAMP and developmental regulation of \textit{pst-cath}. The results from these 'cassette' constructs indicated that the downstream G-box is required to maintain the level of \textit{pst-cath} transcription. Deletion of both G-boxes resulted in a very low, barely detectable level of mRNA in developing cells, whereas deletion or interruption of the downstream G-box resulted in a drop of ~50-fold in the level of \textit{pst-cath/B-gluc} expression. Although the level of transcription was decreased in cassette constructs that affect this downstream G-box, the developmental and cAMP regulation of expression remained qualitatively unaffected [Datta and Firtel 1987]. Pears and Williams [1987] also examined a number of deletion mutations within the \textit{pst-cath} promoter region and observed similar results.

In this paper, we describe experiments using additional cassette constructs to further analyze the ~300-bp flanking region of the \textit{pst-cath} promoter necessary for proper \textit{pst-cath} expression. The results indicate that the ~300-bp 5'-flanking region has a complex structure containing a number of cis-acting regulatory sequences within an 80-bp region that, together, are necessary for proper cAMP, temporal, and quantitative expression of \textit{pst-cath} in multicellular aggregates and in single-cell cultures with added cAMP. In some cases, specific mutations within this region show dramatically different quantitative effects of expression under the two different assay/developmental conditions. We also show that small differences in the spacing between various regions drastically alter the quantitative level of expression. Insertion of short oligonucleotides containing a putative regulatory region into internal deletions can partially complement loss of function.

**Results**

**Palindrome and G-box analysis**

Visual examination of the \textit{pst-cath} promoter shows an island of G/C-rich sequence in the 80–90% A/T 5'-flanking region [see Fig. 1A]. The G/C-rich sequence consists of a palindrome [see arrows above sequence in Fig. 1A and a direct repeat of the 3' end of the palindrome (G-box) that is located 230 bases upstream of the mRNA Cap site. As a background for the experiments described in this paper, Figure 1A summarizes our previous results [Datta and Firtel 1987] in which we used a series of cassette constructs to examine the function of the downstream G-box in the expression of a \textit{pst-cath/B-gluc} fusion gene [see Fig. 2] (see Introduction). In those studies, and the ones described here, we used two different assays: multicellular development and a fast-shaking culture system with cAMP added at 5–6 hr after starvation. During development, \textit{pst-cath} mRNA can be detected first at 10–12 hr, which is the loose aggregate stage, and peaks at the finger stage (~15 hr). In single-cell fast-shaking culture conditions, the gene is rapidly induced in response to cAMP and is generally not detectably expressed [see below] in control cultures lacking exogenous cAMP. In general, these deletions showed a greater quantitative effect on expression using the fast-shake single-cell assay system.

To further define regulatory regions and spatial interactions between the two G-box sequences, we have used an additional 5' deletion and a nested series of deletions to examine the effects of the 5' portion of the palindrome, the A/T-rich upstream region, and the two G-boxes on expression of the \textit{pst-cath/B-gluc} gene fusion [see Methods; Datta and Firtel 1987]. The constructs were transformed into \textit{Dictyostelium} cells, and the expression of the \textit{pst-cath/B-gluc} fusion gene [B-gluc probe] and the endogenous \textit{pst-cath} (\textit{pst-cath} probe) gene were assayed by RNA blots as described previously [Datta and Firtel 1987]. Figure 1B [lower left] shows the expression of constructs pSD3 and pSD4 that carry sequences up to ~313 and ~281-bp, respectively, upstream of the \textit{pst-cath} mRNA Cap site. In fast-shaking culture with added cAMP, pSD3 and pSD4 are expressed at the same level and are not expressed when cAMP is omitted. [This level is very low or not detectable in all single-cell cultures lacking cAMP. In some experiments, a very low level of expression is observed and probably depends on the level of endogenous cAMP produced in any aggregates that
Figure 1. Schematic diagram of cassette constructs and regulation of starting control vectors. (A) Structure and summary of the regulation of some cassette constructs from Datta and Firtel (1987) are shown. Wild-type sequence is indicated above a drawing containing patterned boxes representing different portions of the sequence. The breakdown of the sequence into patterns derives from our interpretation of the results described in this paper and those presented previously (Datta and Firtel 1987). Distance to the mRNA Cap site is indicated above the sequence every 50 bp; dots above the sequence mark every 10 bp. Boldface letters indicate palindromic sequence; arrows above sequence indicate the repeated portion of the palindrome. Italicized letters indicate the downstream G-box; underlined letters indicate both 5' and downstream G-box sequences. The thick black bar in the drawing indicates the position of the HindIII linker. (+) Relative levels of gene fusion expression in fast-shaking culture with added cAMP [F + cAMP] or during development [Dev] in a nonlinear fashion (see subsequent results in figures 3-5). (-) No detectable expression (estimated −0.5% that of pSD3). Structure of gene fusion used in these and subsequent studies is shown in Fig. 2. (B) The structure and developmental and cAMP regulation of pSD3 and pSD4 constructs are shown. Cells transformed with the construct pSD4 were either plated for development or starved in fast-shaking culture and analyzed for pst-cath/β-gluc and endogenous pst-cath gene expression. For endogenous pst-cath mRNA, a probe from the 3' half of the gene was used, as described in Methods. RNA blots are shown. Fast-shaking cultures are indicated with (+) or without (−) added cAMP. (D) RNA from developing cells. (V) RNA from log-phase vegetative cells.
form under these culture conditions [Mehdy and Firtel 1983].] Examination of the regulation of these cells under normal developmental conditions, however, shows that pSD4 is expressed at a level ~3-fold higher than in fast-shaking cultures, whereas pSD3 is expressed at a level ~2.5-fold lower during development than it is in shaker cultures, or ~2-fold lower than pSD4 is expressed in fast-shaking culture. No expression is detectable in vegetative cells using these techniques [data for pSD3 not shown]. When pSD3 and pSD4 are compared during multicellular development, we see that pSD4 is expressed at a level ~6- to 6.5-fold higher. As can be seen in the lower right panel of Figure 1B, endogenous pst-cath mRNA is expressed at the same levels in pSD3 or pSD4 transformants in both fast-shaking +cAMP cultures and developing cells, is not expressed in vegetative cells, and is expressed at very low or undetectable levels in fast-shaking −cAMP cultures (pSD3 data not shown). These results demonstrate that the 32 bases deleted by pSD4 are not required for cAMP or developmental regulation of pst-cath expression but do affect the quantitative level of expression to some degree. This will be examined in more detail below. These results also indicate that placing pUC19 vector sequences closer to the promoter does not have a negative effect on the quantitative level of expression. All cassette constructs examined in this paper contain the region between ~313 and ~281 between the pst-cath promoter sequences and the E. coli vector sequences.

Figure 3 shows a drawing of the series of cassette constructs used to examine pst-cath regulatory sequences. The complete sequences of the constructs are given in the Methods [see Fig. 6]. Expression of the transformed construct and the endogenous pst-cath gene (an internal control) was examined at the tipped aggregate stage (~15 hr) and in fast-shaking cultures. In all cases, the endogenous pst-cath gene was expressed at the same level in all the transformed populations in either finger-stage aggregates or fast-shaking cultures + cAMP [data not shown]. Construct C17 is 'wild-type' except for the insertion of a HindIII linker at the point of the pSD4 deletion and is the starting point for the deletions examined. Succeeding constructs (C15, C16, and C18) deleted a 40-bp A/T-rich sequence and then increasingly larger portions of the G/C-rich sequences. Cells transformed with C17 showed cAMP induction of the gene fusion RNA in fast-shaking culture at levels ~2.5-fold lower than that in cells carrying the wild-type promoter pSD3, but overexpressed the fusion message approximately 5- to 6-fold with respect to pSD3 during multicellular development. When the 40-base A/T-rich sequences upstream from the palindromic region were deleted (construct C16), the level of expression during multicellular development dropped by a factor of 5-6, relative to C17, to approximately that of pSD3. Expression of C16 was barely detectable [down >50-fold] in fast-shaking culture. The next cassette in the series, C15, which deletes the 5’ portion of the palindrome, as well as the 40-bp A/T-rich sequence, gives the same level of expression as does C16 during multicellular development, but its expression in fast-shaking culture is now not detectable. This suggests that the 5’ half of the palindrome has a small quantitative effect on at least CAMP regulation in this context. Further deletion, which leaves only the downstream copy of the G-box in place (C18), results in regulated expression that is barely detectable during development (down 30- to 40-fold) and is not detectable from fast-shaking cultures −cAMP. Construct C14 has the linker in the same position as C15 relative to the G-boxes, but it does contain the 5’ portion of the palindrome upstream of the linker. C14 is expressed at slightly higher levels than C16 [see Fig. 2 drawing; RNA blot not shown]. None of the cassette constructs was detectably expressed in vegetative cells.

Cassette constructs C20, C22, C24, and C26 examine the function of the palindrome and downstream G-box in a promoter background containing the 40-bp A/T-rich region. C20 has a linker upstream from the G-boxes and palindromic sequence. It is expressed at about the same level as C17 in fast-shaking culture and at approximately wild-type levels during development [see Fig. 3 drawing; RNA blots not shown]. C22, which lacks only the 5’ portion of the palindrome, shows a level of expression 5-fold higher than pSD3 during development, similar to what is seen with C17, and a level 30-fold lower than pSD3 in fast-shaking culture. C24 lacks the entire palindrome and is expressed during development and in fast-shaking culture at levels similar to C15, which has two copies of the G-box but no A/T-rich sequence. C24 is expressed at ~10 times the level of C18.
Datta and Firtel
during development. Comparison of C24 and C18 sequence shows that C18 contains only the downstream G-box, whereas C24 has the G-box plus the 40-bp A/T-rich sequence. When the entire G/C-rich region has been deleted, however (see C26), there is no detectable expression. Thus, the A/T-rich sequence alone is not sufficient to direct expression of the gene fusion.

Effects of spacing between promoter elements

The above results that the A/T and G-box regions affect

| Function of Palindrome and G-box sequences |
|--------------------------------------------|
| -213                                      |
| GATCCAAATAGATAATTATTAAATAATCACCAGTTTTTTTTCTATTTTTATGAA |
| -213                                      |
| AACACAGCGGGTGTTAGATGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT...
the level of \textit{pst-cath} expression. Because the two G-boxes are in close proximity to each other, we decided to examine the effect of spacing between these two elements. Cassette C13 contains a \textit{HindIII} linker between the two G-boxes, which changes the sequence between these two regions. This change results in 2 additional nucleotides between the two sequences and a change in the 3' nucleotide of the 5' G-box (T → C) (see Fig. 4). The internal 4 nucleotides of the linker are missing, in cassette C13S, whereas they are duplicated in C13F. In C13S, this results in additional nucleotide changes from the wild-type sequence. The constructs were co-transformed into KAx-3 cells and analyzed as above.

RNA blot hybridization of the expression of C13, C13S, and C13F in fast-shaking culture plus cAMP is shown in Figure 4. C13 is expressed at a level ~25-fold lower than the control pSD3. When 4 bases are removed from the \textit{HindIII} linker between the palindrome and the downstream G-box (C13S), the level of expression rises to near wild-type levels. When 4 extra bases are added (C13F), however, the level of expression drops below the level of C13 expression of >50-fold relative to pSD3. These results suggest that spacing between regulatory regions may be crucial for the regulation of \textit{pst-cath} expression. We have also examined the expression of the C13 spacing series during development. As seen in the lowest panel of Figure 4, C13 is expressed at a level approximately the same as pSD3, whereas C13F and C13S are expressed at levels only slightly lower (<1.5-fold).

\textit{Effect of G-box oligonucleotide on \textit{pst-cath} expression}

Deletion or interruption of the 3' G-box results in a major drop in the level of expression of \textit{pst-cath}. We have examined the ability of a 31-bp synthetic double-stranded oligonucleotide containing this G-box and surrounding sequence to complement cassette constructs C1 and C9 in which this region has been removed. Figure 5 shows the drawing of the constructs along with the DNA sequence. Construct C1 has a partial 3' G-box and is missing an A/T-rich sequence immediately downstream. Construct C9 is missing the region between the two G-boxes, the 3' G-box, as well as an A/T-rich sequence. Both of these constructs show extremely low levels of expression during multicellular development and no detectable expression in fast-shaking culture with cAMP (see Fig. 5). The 31-mer oligonucleotide containing the 3' G-box and some downstream A/T-rich sequence (see Fig. 5) was inserted into the \textit{HindIII} linker within C1 and C9 in the same orientation as in the wild-type gene, and the resultant vectors (C1G and C9G) were then transformed into \textit{Dictyostelium} cells and the expression of the constructs examined. C9G and C1G are expressed at a level 4- to 5-fold higher than the parents C9 and C1 in developing cells, although the level of expression is seven-fold lower than pSD3. A similar qualitative effect of the oligonucleotide on the expression of C9 and C1 is observed in fast-shaking culture with cAMP, but the quantitative level of expression is reduced severalfold. The level of endogenous \textit{pst-cath} expression is the same in the different transformants (data not shown).

\textbf{Discussion}

We have used a series of cassette and spacing constructs to examine the \textit{cis-acting} regions controlling \textit{pst-cath} expression. Analysis of \textit{Dictyostelium} cell lines transformed with the cassette constructs shows that \textit{pst-cath} has an 80-bp regulatory region containing at least three regions or sequence elements that are required to regulate the level and pattern of \textit{pst-cath} expression properly.

One of the important observations is that there are major quantitative differences in the relative level of expression of some cassette constructs under the two different developmental conditions, although the endogenous \textit{pst-cath} gene shows a consistent level of expression. For example, we see that construct pSD3 shows a level of expression reproducibly lower in fast-shaking cultures than in developing cells, whereas pSD4 shows the opposite effect. In the case of some cassette constructs, the relative differences in expression under these two physiological conditions can be great. This is highlighted by constructs C17 and C22, which are expressed, respectively, and 2.5- and ~30-fold lower than pSD3 in fast-shaking culture + cAMP but 5- to 6-fold higher than pSD3 in multicellular aggregates. In both of these, the linker interrupts the continuity of the sequence around the 40-bp A/T-rich region and, in the case of C22, deletes the 5' half of the palindrome of the upstream G-box. One possible explanation is that the linker in these constructs interrupts a regulatory sequence that acts as a negative element in developing cells. In contrast, in cAMP-induced cells in fast-shaking culture, the same region may either act as a positive element or the linker may interrupt interactions between proteins binding to surrounding sequences. Presumably differences in expression under these conditions arise from qualitative or quantitative differences in one or more \textit{trans-acting} regulatory factors or its (their) activation under different developmental conditions. Because the RNA expressed from each construct is identical, differences are not due to differences in mRNA stability. The presence of multiple factors or binding sites is also suggested by multiple bands in gel mobility shift experiments using this region as a probe (A. Hjorth, N. Khanna, and R.A. Firtel, unpubl.). It is possible that cell-cell interactions that are present during multicellular development, but not in the single-cell fast-shaking culture, play a role in inducing maximal expression of \textit{pst-cath} and may regulate the number, level, or activity of specific \textit{trans-acting} proteins differentially.

Deletion of the A/T-rich region and 5' portion of the palindrome has significant effects on the expression in fast-shaking culture but not in developing cells. In contrast, subsequent deletion of either G-box has a major effect on expression in either assay. Although the A/T-rich region appears to have a positive effect on expression, a construct containing only the A/T-rich region is...
Effect of Spacing on *pst-cath* Regulation

Figure 4. Effect of spacing on *pst-cath* regulation. A diagram of spacing construct sequence is shown. Thick black bars indicate the *HindIII* linker, length of the bar indicates relative length of the linker due to deletion or insertion of nucleotides. Below is shown the sequence of the C13 series constructs. Italicized letters indicate the downstream G-box, underlined letters indicate both G-boxes, and boldface letters indicate the palindrome. Pluses and minuses indicate relative levels of expression in fast-shaking culture, as described in the legend to Fig. 1.

not detectably expressed during development (C26). Construct C5 (see Fig. 1A; Datta and Firtel 1987) is expressed at a low level, suggesting that the 5' portion of the palindrome has a positive effect on the level of expression when combined with the A/T-rich region. These results must be interpreted with care because the full importance of sequence context, the effect of sequences flanking the putative regulatory regions, and the effect of spacing between regulatory regions is not known for these two elements.

One of the obvious features of the regulatory region is the two G-boxes. We have used a number of approaches to examine their function and relative importance. Comparison of the expression constructs C8 and C24, lacking the upstream G-box to C1, C2, or C4, which lack or have an interrupted downstream G-box, suggests that the upstream G-box, which has two nucleotide changes relative to the downstream G-box, may not be as active in enhancing gene expression as the downstream G-box. We have also shown in the analysis of the C13 series (C13, C13S, C13F; see Fig. 4) that spacing between the G-boxes is important for expression in fast-shaking cul-

![Figure 4](image-url)
ture. In the case of C13S, a deletion that brings the two G-boxes into closer proximity than the wild-type spacing, expression in fast-shaking culture rises to wild-type levels. A previously examined cassette construct, C8, has a deletion of the palindromic G-box, bringing the downstream G-box closer to the upstream regulatory regions (see Fig. 1A). This results in another spacing-type construct that is expressed at wild-type levels in fast-shaking culture [Datta and Firtel 1987]. Taken together, these results suggest that either the spacing between upstream regulatory elements and the downstream G-box, or between upstream regulatory elements and downstream promoter elements such as the TATA box or oligo(dT) run [present in Dictyostelium polymerase II genes between the TATA box and Cap site [Kimmel and Firtel 1983]] are important in the regulation of the level of \( \text{pst-cath} \) expression. We do not believe that the nucleotide sequence changes caused by the linker insertion cause the change in expression because C14, with a low level of expression, appears to have a sequence more similar to the parental wild-type sequence than the highly expressed C13S construct. This suggests that the wild-type spacing may not be optimum and that there is an extended binding site or multiple binding sites for trans-acting factors within this region and that spacing is important for their interaction.

\[ \text{Effect of G-box Oligonucleotide on Pst-cath Expression} \]

\[ \begin{array}{c}
\text{G-box Oligonucleotide} \\
pSD3 \\
C1 \\
C1G \\
C2 \\
C9G \\
pSD3 \\
C1 \\
C1G \\
C2 \\
C9G \\
\end{array} \]

\[ \begin{array}{c}
\text{FacAMP} \\
\text{Dec} \\
+++ \\
+++ \\
++ \\
++ \\
+++ \\
++ \\
++ \\
++ \\
++ \\
\end{array} \]

\[ \text{Figure 5. (See following page for legend.)} \]
When we insert the downstream G-box and surrounding sequences back into constructs in which part or all of this region has been deleted, we observe a four- to fivefold enhancement of expression with two constructs, but the level of expression is still below wild-type levels. In both cases, we are inserting the oligonucleotide that contains the downstream G-box into a construct that contains the upstream G-box. In addition, the oligonucleotide has additional sequences used in the cloning (HindIII linker) and subsequent orientation (EcoRV site) of the fragment. Thus, the spacing, as well as the nucleotide sequence, between the G-boxes is altered. It is quite possible that one of the reasons why reinsercion of the oligonucleotide did not restore full, wild-type level of expression could be due to sequence context of the surrounding nucleotides and spacing interactions of the oligonucleotide in relation to adjacent sequences.

C.J. Pears and J.G. Williams [pers. comm.] have done similar experiments. They have inserted a 68-bp-long synthetic oligonucleotide containing the regulatory region from cysteine proteinase 1 (CP-1), another cAMP-inducible gene [Pears et al. 1985], into a G-box deletion of pst-cath. This sequence has the structure of a palindrome with substantial nucleotide sequence divergence and has a G/T-rich [or complementary C/A-rich] sequence. Insertion of the 68-mer into a pst-cath construct carrying a deletion of the two G-boxes restores the full level of cAMP-inducible expression. Insertion of one copy of a 28-mer containing the 3′ G-box of CP-1 results in an increase in the level of expression from 15% of wild-type to 30%, whereas insertion of two copies of the 28-mer results in ~150% activity. These results are consistent with ours and suggest that the G-rich boxes in the two genes may serve similar functions.

We have identified three regions in the pst-cath 5′-flanking region with different regulatory effects on pst-cath expression and have examined the effect of spacing between the two G-box sequences. The results described here will aid us in identifying trans-acting regulatory proteins involved in these processes, as well as, analysis of the basis for differences in expression in shaking cultures and multicellular development. This difference, including the identification of other physiological regulatory factors, will be important in defining these factors further and in determining how cellular differentiation is regulated in Dictyostelium.

**Methods**

**Growth and development of D. discoideum**

Strain KAx-3 (see Poole and Firtel 1984) was grown axenically in suspension and was used for all experiments [Firtel and Bonner 1972]. Cells for development were plated on Whatman 50 filters, which were floated on buffer, as described previously [Datta et al. 1986; Sussman 1966; Datta and Firtel 1987]. Morphology was examined, and cells were harvested at the finger stage (~15 hr), the peak of pst-cath mRNA expression. The same cultures of transformed cell populations were used to examine expression during multicellular development and in fast-shaking cultures.

**Construction of cassette vectors and spacing constructs**

The pSD3 construct, as well as the deletion and reconstruction processes used to generate the cassette constructs described in this paper, have been described previously (Datta and Firtel 1987). Additional internal deletions were made by combining B′ deletions with the appropriate 3′ deletions by means of a common HindIII site, as shown in Figure 6.

The constructs to examine spacing are derived from the cassette vector C13. C13 was linearized at the unique HindIII site, either filled in with the large fragment of DNA polymerase I to add 4 bp (C13F) or digested with nuclease S1 to delete 4 bp (C13S), and recircularized. Clones were screened for the destruction of the HindIII site. The addition or removal of nucleotides was verified by sequencing.

Figure 6 shows the sequence of all constructs used in this paper. Sequences of C13S and C13F and those containing an oligonucleotide are given in Figures 4 and 5 within Results.

**Transformation**

DNA-mediated transformation of *D. discoideum* cells was carried out as described previously [Nellen et al. 1984; Nellen and Firtel 1985]. Cassette constructs were cotransformed with the vector B10SX, which confers G418 resistance. In all single clones examined, there was cotransformation of the *pst-cath/B-gluc* construct with B10SX, and the copy number of the *pst-cath/B-gluc* was similar. In other experiments with other vectors, one sees a lower frequency of cotransformation (S. Datta and R.A. Firtel, unpubl.). Although we did not see differ-

---

**Figure 5.** Effect of G-box oligonucleotide on *pst-cath* expression. Two 31-bp complementary oligonucleotides were synthesized using an Applied Biosystems 380B DNA Synthesizer and were inserted into the HindIII site of C1 and C9, which had been blunt-ended. To confirm the orientation of this sequence, an EcoRV site was inserted downstream from the G-box. This restriction site was chosen because it was unique within the clone, and the hexonucleotide sequence defining the EcoRV site is relatively A/T-rich like the majority of 5′-flanking sequences in *Dictyostelium* genes [Kimmel and Firtel 1983]. The C/G ClaI overhangs were placed at the ends because oligonucleotide is also being used to examine the expression in a number of *Dictyostelium* genes and represented the most useful linker sequence for a number of different studies. It is possible that the additional nucleotides for the EcoRV restriction endonuclease site might reduce the level of expression, either because of the sequence or possibly because of changing spacing requirements, although they lie downstream from the G-box. The sequence and drawing of the construct is shown. In the drawing of C1G and C9G, the oligonucleotide (shown by the open box and arrow to denote orientation) is not drawn to scale. In the sequence, the terminal GGGTT is not italicized because it is derived from nucleotides at position −201, rather than −213, but because these nucleotides do reconstruct the 3′ end of the downstream G-box, they are underlined. Expression in fast-shaking cultures and development was examined as described in Methods. *Note:* The level of expression of C9′+ and C1+ are too low to quantitate. Therefore, we cannot determine the relative increase in level of expression in C9G+ and C1G+.
Prestalk gene regulatory sequences

**Cassette Constructs**

| Construct | Sequence |
|-----------|----------|
| WILD TYPE | ATGAAAACACGGGCTGGTTTAAGTGAAGGGTTGTTTTATATAGGGCGAGCAGGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| G1 | GATCCAATAGAATTTATTAAAATAATCACCAATTGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| | ATGAAAACACGGGCTGGTTTAAGTGAAGGGTTGTTTTATATAGGGCGAGCAGGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| G2 | GATCCAATAGAATTTATTAAAATAATCACCAATTGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| | ATGAAAACACGGGCTGGTTTAAGTGAAGGGTTGTTTTATATAGGGCGAGCAGGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| G3 | GATCCAATAGAATTTATTAAAATAATCACCAATTGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| | ATGAAAACACGGGCTGGTTTAAGTGAAGGGTTGTTTTATATAGGGCGAGCAGGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| G4 | GATCCAATAGAATTTATTAAAATAATCACCAATTGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| | ATGAAAACACGGGCTGGTTTAAGTGAAGGGTTGTTTTATATAGGGCGAGCAGGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| G5 | GATCCAATAGAATTTATTAAAATAATCACCAATTGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| | ATGAAAACACGGGCTGGTTTAAGTGAAGGGTTGTTTTATATAGGGCGAGCAGGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| G6 | GATCCAATAGAATTTATTAAAATAATCACCAATTGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| | ATGAAAACACGGGCTGGTTTAAGTGAAGGGTTGTTTTATATAGGGCGAGCAGGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| G7 | GATCCAATAGAATTTATTAAAATAATCACCAATTGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| | ATGAAAACACGGGCTGGTTTAAGTGAAGGGTTGTTTTATATAGGGCGAGCAGGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| G8 | GATCCAATAGAATTTATTAAAATAATCACCAATTGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| | ATGAAAACACGGGCTGGTTTAAGTGAAGGGTTGTTTTATATAGGGCGAGCAGGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| G9 | GATCCAATAGAATTTATTAAAATAATCACCAATTGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| | ATGAAAACACGGGCTGGTTTAAGTGAAGGGTTGTTTTATATAGGGCGAGCAGGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| G10 | GATCCAATAGAATTTATTAAAATAATCACCAATTGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| | ATGAAAACACGGGCTGGTTTAAGTGAAGGGTTGTTTTATATAGGGCGAGCAGGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| G11 | GATCCAATAGAATTTATTAAAATAATCACCAATTGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| | ATGAAAACACGGGCTGGTTTAAGTGAAGGGTTGTTTTATATAGGGCGAGCAGGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| G12 | GATCCAATAGAATTTATTAAAATAATCACCAATTGTTTTTTTTATTCTTTAAATTTTTTCTATTT |
| | ATGAAAACACGGGCTGGTTTAAGTGAAGGGTTGTTTTATATAGGGCGAGCAGGTTTTTTTTATTCTTTAAATTTTTTCTATTT |

The 5' deletion pSD4, cassette, and spacing constructs were cotransformed into *Dictyostelium* cells with B10SX (encoding resistance to G418), as described previously [Datta and Firtel 1987]. Populations of stable transformants were examined. Copy numbers of *pst-cath/β-gluc* cassettes were determined from DNA blots of DNA isolated from transformed cells, digested with appropriate restriction enzymes to release the *pst-cath/β-gluc* fusion, size fractionated, blotted, and probed with β-gluc sequences, as described previously [Datta and Firtel 1987]. All transformed cell lines carried the 5' deletion or the cassette constructs at a copy number of -10–15 copies per cell [data not shown].

RNA analysis

RNA was extracted from cells, size fractionated on a 1.4% agarose denaturing gel, blotted onto GeneScreen [New England Nuclear] or Paul nylon membranes [ICN], and probed with nick-translated DNA fragments, as described previously [Mehdy et al. 1983; Mehdy and Firtel 1985]. Some of the data were quantitated by densitometric scanning of the autoradiographs using an LKB Ultrascan XL laser densitometer. In some cases, ranges are given because of differences between experiments and/or between tracings.
Datta and Firtel

Fast-shaking culture analysis

Fast-shaking cultures were prepared according to Mehdy et al. (1983). cAMP was initially added at 6 hr to 300 μM after starvation, and then every 2 hr to 300 μM for +cAMP samples. Samples labeled –cAMP had no endogenous cAMP added. Both cultures were made 1 mM EDTA to reduce cell–cell association. Cells were harvested 12 hr after initiation of starvation.

Acknowledgments

We would like to thank Kevin Ahern and James Posakony for helpful discussions and Sandra Mann for critically reading the manuscript. S.D. was a recipient of a National Science Foundation predoctoral fellowship and has been supported by a U.S. Public Health Service predoctoral training grant. This work has been supported by a U.S. Public Health Service grant (GM24279) to R.A.F. from the National Institutes of Health.

References

Barklis, E. and H.F. Lodish. 1983. Regulation of Dictyostelium discoideum mRNAs specific for prespore and prestalk cells. Cell 2032: 1139–1148.

Chisholm, R.L., E. Barklis, and H.F. Lodish. 1984. Mechanism of sequential induction of cell-type specific mRNAs in Dictyostelium differentiation. Nature 310: 67–69.

Datta, S. 1987. "The regulation of prestalk-specific gene expression in Dictyostelium discoideum". Ph.D. thesis, University of California, San Diego.

Datta, S. and R.A. Firtel. 1987. Identification of the sequences controlling cyclic AMP regulation and cell-type specific expression of a prestalk-specific gene in Dictyostelium discoideum. Mol. Cell. Biol. 7: 149–159.

Datta, S., S.K.O. Mann, A. Hjorth, C.M. Silan, and R.A. Firtel. 1987. cAMP-regulated gene expression during Dictyostelium development is mediated by the cell-surface cAMP receptor. In Genetic regulation of development, 45th symposium for the society of developmental biology (ed. W.F. Loomis), pp. 373–383. Alan R. Liss, New York.

Dowds, B.C.A. and W.F. Loomis. 1984. Cloning and expression of a cDNA that comprises part of the gene coding for a spore coat protein of Dictyostelium discoideum. Mol. Cell. Biol. 4: 2273–2278.

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

Gomer, R.H., S. Datta, and R.A. Firtel. 1986a. Cellular and subcellular distribution of a cAMP-regulated prestalk protein and prespore protein in Dictyostelium discoideum. J. Cell Biol. 103: 1999–2015.

Gomer, R.H., D. Armstrong, B.H. Leichtling, and R.A. Firtel. 1986b. cAMP induction of prespore and prestalk gene expression in Dictyostelium is mediated by the cell-surface cAMP receptor. Proc. Natl. Acad. Sci. 83: 8624–8628.

Gomer, R., S. Datta, M. Mehdy, T. Crowley, A. Sivertsen, W. Nellen, C. Reymond, S. Mann, and R.A. Firtel. 1985. Regulation of cell-type-specific gene expression in Dictyostelium. Cold Spring Harbor Symp. Quant. Biol. 50: 801–812.

Jefferson, R.A., S.M. Burgess, and D. Hirsh. 1986. β-Glucuronidase from Escherichia coli as a gene-fusion marker. Proc. Natl. Acad. Sci. 83: 8447–8451.

Kimmel, A.R. and R.A. Firtel. 1983. Sequence organization in Dictyostelium: Unique structure at the 5'-ends of protein coding genes. Nucleic Acids Res. 11: 541–552.

Loomis, W.F. 1975. Dictyostelium discoideum: A developmental system. Academic Press, New York.

———. 1982. The development of Dictyostelium discoideum. Academic Press, New York.

Maniatis, T., E. Fritsch, and J. Sambrook. 1982. Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Mann, S.K.O., S. Datta, P. Howard, A. Hjorth, C. Reymond, C.M. Silan, and R.A. Firtel 1987. Cyclic AMP regulation of gene expression during Dictyostelium development. In UCLA Symposium on molecular biology of development (ed. R.A. Firtel and E.H. Davidson), pp. 303–328. Alan R. Liss, New York.

Mehdy, M.C. and R.A. Firtel. 1985. A secreted factor and cAMP jointly regulate cell-type-specific gene expression in Dictyostelium. Mol. Cell. Biol. 5: 705–713.

Mehdy, M.C., D. Ratner, and R.A. Firtel. 1983. Induction and modulation of cell-type-specific gene expression in Dictyostelium. Cell 32: 761–771.

Nellen, W. and R.A. Firtel. 1985. High-copy-number transformants and co-transformation in Dictyostelium. Gene 39: 155–163.

Nellen, W., C. Silan, and R.A. Firtel. 1984. DNA-mediated transformation in Dictyostelium discoideum: Regulated expression of an actin gene fusion. Mol. Cell. Biol. 4: 2890–2898.

Pears, C.J. and J.G. Williams. 1987. Identification of a DNA sequence element required for efficient expression of a developmentally regulated and cAMP-inducible gene of Dictyostelium discoideum. EMBO J. 6: 195–200.

Pears, C.J., H. Mahbubani, and J.G. Williams. 1985. Characterization of two highly diverged but developmentally co-regulated cysteine proteinase genes in Dictyostelium discoideum. Nucleic Acids Res. 13: 8853–8866.

Poole, S.J. and R.A. Firtel 1984. Genomic instability and mobile genetic elements in regions surrounding two discoidin I genes of Dictyostelium discoideum. Mol. Cell. Biol. 4: 671–680.

Reymond, C.P., R.H. Gomer, M.C. Mehdy, and R.A. Firtel. 1984. Developmental regulation of a Dictyostelium gene encoding a protein homologous to mammalian ras protein. Cell 39: 141–148.

Sussman, M. 1966. Biochemical and genetic methods in the study of cellular slime mold development. In Methods in cell physiology (ed. D. Prescott), vol. 2, pp. 397–410. Academic Press, New York.

Williams, J.G., C.J. Pears, K.A. Jermyn, D.M. Driscoll, H. Mahbubani, and R.R. Kay 1986. The control of gene expression during cellular differentiation of Dictyostelium discoideum. In Regulation of gene expression (ed. I. Boothe and C. Higgins), pp. 277–298. Cambridge University Press, Cambridge, England.

Williams, J.G., A. Ceccarelli, S. McRobbie, H. Mahbubani, R.R. Kay, A. Early, M. Berks, and K.A. Jermyn. 1987. Direct induction of Dictyostelium prestalk gene expression by DIF provides evidence that DIF is a morphogen. Cell 49: 185–192.

304 GENES & DEVELOPMENT
An 80-bp cis-acting regulatory region controls cAMP and development regulation of a prestalk gene in Dictyostelium.

S Datta and R A Firtel

*Genes Dev.* 1988, 2:
Access the most recent version at doi:10.1101/gad.2.3.294

This article cites 20 articles, 10 of which can be accessed free at: http://genesdev.cshlp.org/content/2/3/294.full.html#ref-list-1

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.