Volume 219, number 1, 189–196 FEB 04899 July 1987

Studies with the *Escherichia coli* galactose operon regulatory region carrying a point mutation that simultaneously inactivates the two overlapping promoters

Interactions with RNA polymerase and the cyclic AMP receptor protein

Sreenivasan Ponnambalam, Annick Spassky* and Stephen Busby

Department of Biochemistry, University of Birmingham, PO Box 363, Birmingham B15 2TT, England and **Département de Biologie Moléculaire, Institut Pasteur, 25 Rue du Dr Roux, 75724 Paris Cedex 15, France*

Received 22 April 1987

We report in vitro studies of the interactions between purified *E. coli* RNA polymerase and DNA from the regulatory region of the *E. coli* galactose operon which carries a point mutation that simultaneously stops transcription initiation at the two normal start points, *S1* and *S2*. In the presence of this point mutation, transcription initiates at a third start point 14/15 bp downstream of *S1*, showing that inactivation of the two normally active promoters, *P1* and *P2*, unmasks a third weaker promoter, *P3*. Transcription initiation in the *gal* operon is normally regulated by the cyclic AMP receptor protein, CRP, that binds to the *gal* regulatory region and switches transcription from *P2* to *P1*. With the point mutation, CRP binding switches transcription from *P3* to *P1*, although the formation of transcriptionally competent complexes at *P1* is very slow. The results are discussed with respect to the mechanism of transcription activation by the CRP factor and the similarities between the regulatory regions of the galactose and lactose operons.

Galactose operon; Tandem promoter; Promoter mutation; RNA polymerase; cyclic AMP receptor protein; Protein interaction; (*E. coli*)

1. INTRODUCTION

The regulatory region of the *Escherichia coli* galactose operon is unusual as it contains two overlapping but distinct promoters, *P1* and *P2*, that are regulated by the cyclic AMP receptor protein, CRP [1]. The transcription start point for the *P2* promoter is at *S2*, 5 bp upstream of *S1*, the transcription start point of the *P1* promoter (fig.1). In vivo, *gal* operon transcription usually starts at *S2*. However in conditions where the intracellular level of cAMP rises, the cAMP-CRP complex binds to the *gal* promoter region and switches transcription from *P2* to *P1* [2].

In recent work, we and others have measured the effects of over 60 different point mutations in the *gal* operon promoter region [3–8]. Surprisingly, with one exception, none of the mutations cause drastic reductions in expression from the *gal* promoter region. This can be readily explained as the DNA sequences necessary for *P1* and *P2* function are distinct and, thus, mutations that knock out *P1* leave *P2* active and vice versa. The exception is a point mutation at −12 that falls at the intersection of the *P1* and *P2* −10 hexamer sequences (see fig.1): in vivo, this mutation causes a >95% reduction in expression from the *gal* operon regulatory region [6]. Here we report the properties, in vitro,

Correspondence address: S. Busby, Department of Biochemistry, University of Birmingham, PO Box 363, Birmingham B15 2TT, England

Published by Elsevier Science Publishers B.V. (Biomedical Division)

of the *gal* operon regulatory region DNA carrying this unique mutation that simultaneously knocks out both promoters. We show that the absence of *P1* and *P2* unmasks a third overlapping promoter to which RNA polymerase can bind and initiate transcription 14–15 bp downstream of *S1*. Further, we have investigated the interactions between CRP and RNA polymerase at the *gal* regulatory region carrying this mutation at −12.

2. MATERIALS AND METHODS

The *gal* operon regulatory region was cloned on a 144 bp DNA fragment between the *Eco*RI and *Hin*dIII sites of pBR322 as described [9]. The *gal* fragment carried either the wild type promoter sequence [4], point mutations at −12 or −19 [6] or the Δ420 deletion [10]. Transcription and footprint experiments were performed on *Pst*I-*Hin*dIII or *Pst*I-*Bst*EII fragments isolated from these constructions as before [9,11].

'Run-off' transcription assays were performed exactly as previously described by us [9]. Footprint experiments were as described by Spassky et al. [11]. 1–2 nM end-labelled DNA fragments were incubated in the presence or absence of 100 nM CRP and 200 μM cAMP in 40 mM Tris, pH 8, 10 mM $MgCl_2$, 100 mM KCl, 0.1 mM dithiothreitol, 5% glycerol. After 15 min, 150 nM RNA polymerase was added and after a further 15 min the footprint was determined by adding 75 ng/ml DNase I or the copper-orthophenanthroline mix described by Sigman et al. [12]. Digestion was stopped after 20 s by phenol extraction and alcohol precipitation. Samples were analysed on 10% sequencing gels followed by autoradiography to visualise the footprint.

3. RESULTS

3.1. *Inactivation of* gal P1 and gal P2 *reveals a new transcription start point*

The sequence of the *gal* operon promoter region is shown in fig.1 together with the transcription start points and corresponding −10 hexamer sequences for the *P1* and *P2* promoters. The position of the mutation at −12 that disrupts both the *P1* and *P2* −10 hexamer sequences is also indicated. The 890 bp *Pst*I-*Hin*dIII fragment, illustrated in fig.1, was isolated carrying either the wild type *gal*

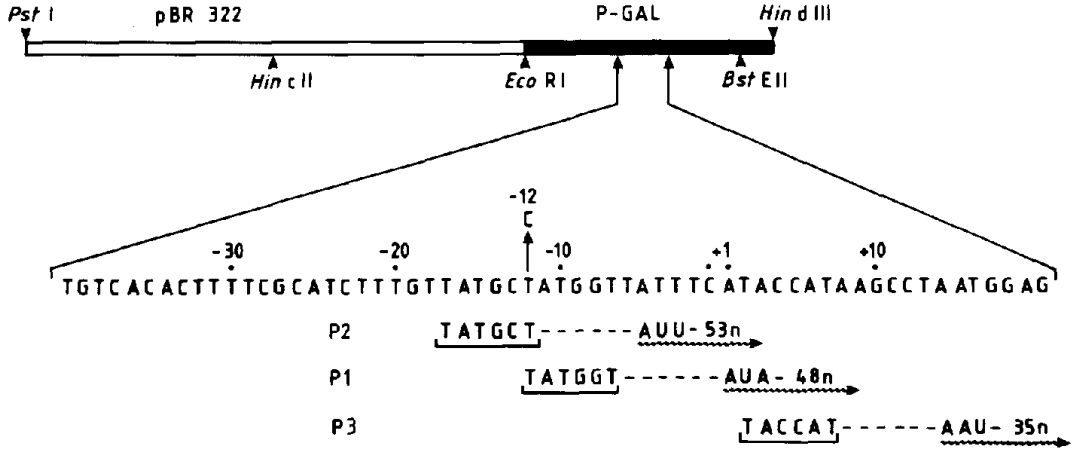

Fig.1. Diagram of the *Pst*I-*Hin*dIII fragments used in this study. The bar represents the DNA sequence consisting of 750 bp from the *Pst*I site to the *Eco*RI site of pBR322 (open) and a 144 bp *Eco*RI-*Hin*dIII fragment carrying the *gal* operon promoter region (shaded). The position of the *Bst*EII site just upstream of the *Hin*dIII site is marked. Below the bar is shown in detail an expanded segment of the *gal* promoter region with the nucleotide sequence of the upper strand of the DNA around the transcription start points. For *gal P1, P2,* and the *P3* promoter described here, the −10 hexamer sequences are indicated by brackets and the transcription start points are shown by the end of a wavy arrow. Above each arrow is marked the RNA sequence at the 5′-end of the transcript and the distance in bases from that start point to the *Hin*dIII end of the fragment. The *gal* sequence is numbered with respect to the *P1* transcription start point: the position of the transition at −12 that inactivates both *P1* and *P2* is indicated.

promoter sequence or the mutation at −12. After addition of purified RNA polymerase, run-off transcripts were made and analysed on polyacrylamide gels (fig.2). With the wild type *gal* promoter sequence, the major transcript starts at *S2* and runs 53 bases to the end of the fragment: a minor band 48 bases long is due to some initiation at *S1* (fig.2a). With the fragment carrying the mutation at −12, no transcripts initiate at *S1* or *S2* but a new band appears that is 35 bases long

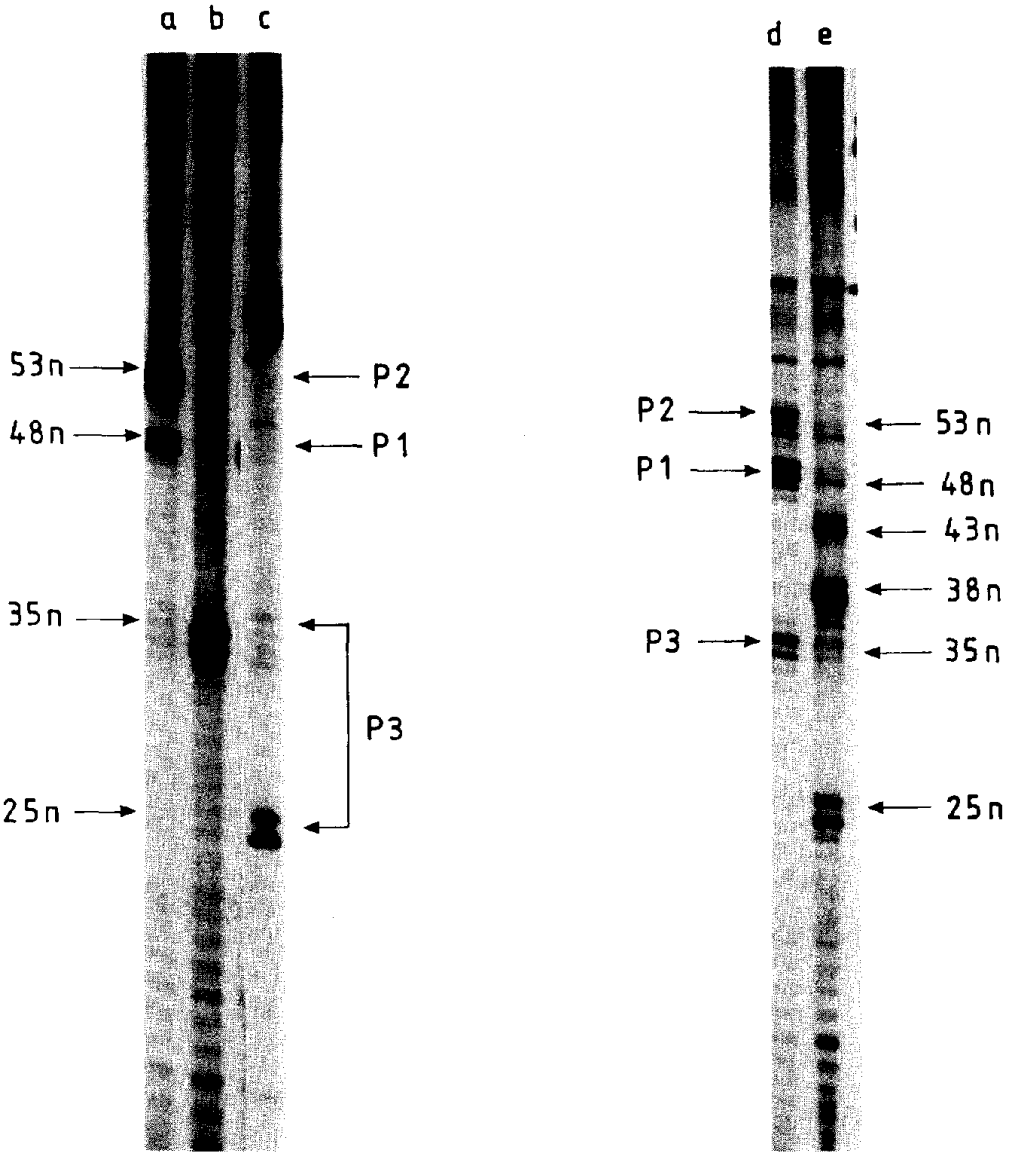

Fig.2. In vitro transcription experiments. The figure shows autoradiograms of gels run to analyse RNA made in run off transcription assays. The gels were calibrated and the length in bases of each band is indicated together with the promoter that is responsible for each band. The DNA fragments used were (a) *Pst*I-*Hin*dIII fragment carrying the wild type *gal* promoter sequence, (b) *Pst*I-*Hin*dIII fragment carrying the mutation at −12, (c) *Pst*I-*Bst*EII fragment carrying the mutation at −12, (d) *Pst*I-*Hin*dIII fragment carrying the wild type *gal* promoter sequence but a deletion from −29 to −92, (e) *Pst*I-*Bst*EII fragment carrying the deletion from −29 to −92.

(fig.2b). To identify the origin of this transcript the *Pst*I-*Hin*dIII fragment carrying the mutation at −12 was shortened by restriction with *Bst*EII; this cuts 10 bp before the *Hin*dIII site (fig.1). After restriction, the 35 base transcript is reduced to 25 bases (fig.2c), showing that this transcript is due to RNA polymerase initiating transcription at +14 or +15 and moving rightwards to the end of the frag-

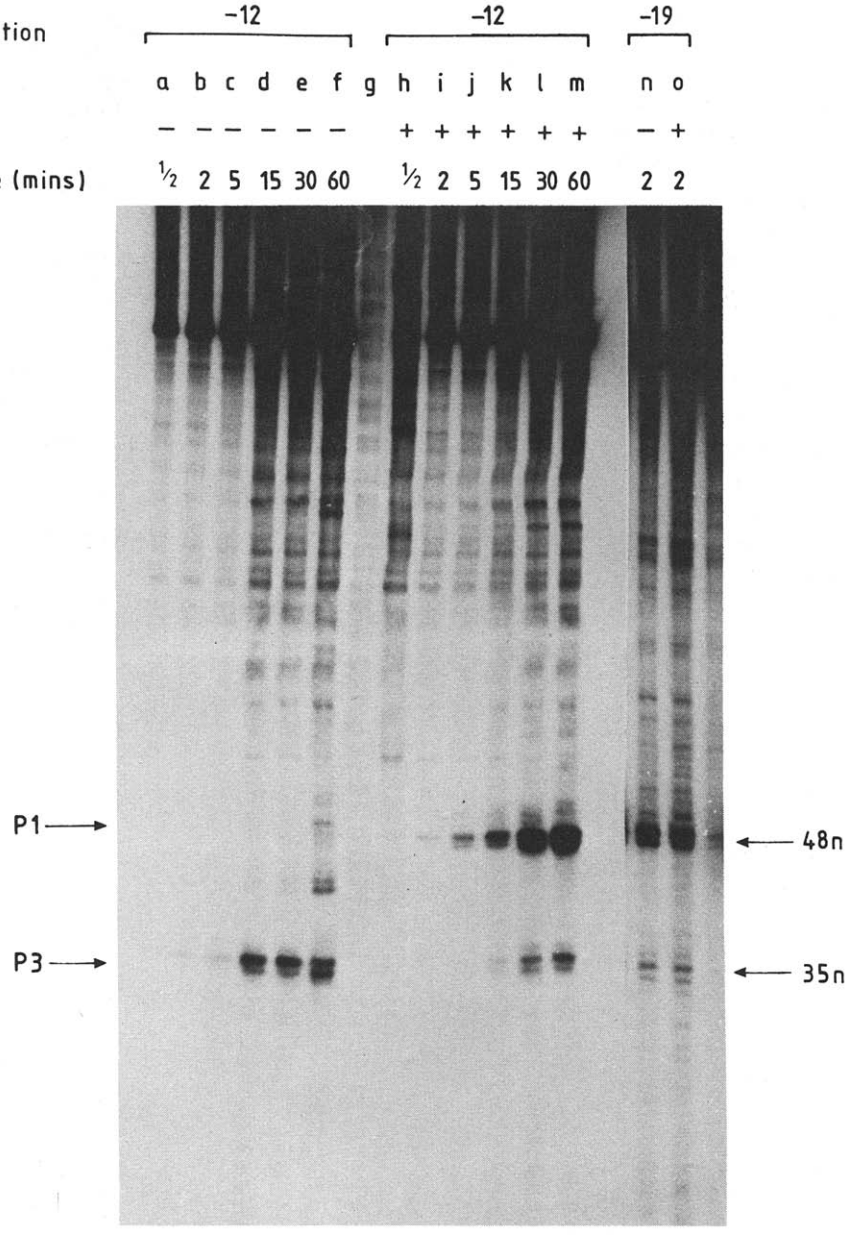

Fig.3. Kinetics of transcription initiation. RNA polymerase was incubated with the *Pst*I-*Hin*dIII fragment carrying the mutation at −12 in the absence (lanes a–f) or presence (lanes h–m) of cAMP-CRP. After the different times shown, nucleoside triphosphates and heparin were added and the transcripts were analysed by gel electrophoresis as shown in the figure. Lanes n and o show an experiment with DNA carrying a mutation at −19 that gives transcription from *P1* under all conditions.

ment. Fig.1 shows the position and orientation of the transcription start point and indicates a plausible −10 sequence, 5′-TACCAT-3′, from +2 to +7.

We reasoned that the transcript starting at +14/+15 must be due to a third promoter (labelled *P3* in fig.1) that is masked in the presence of *P1* or *P2* but which becomes active when *P1* and *P2* are simultaneously inactivated. To test this hypothesis we repeated the run-off transcription experiments on a *Pst*I-*Hin*dIII fragment carrying the wild type *gal* promoter sequence with a deletion of the sequence from −29 to −92 that eliminates the −35 regions of both *P1* and *P2*. Fig.2d,e shows that, with the deletion, transcripts of 53, 48 and 35 bases are made which are reduced to 43, 38 and 25 bases when the fragment is shortened with *Bst*EII. The 53 and 48 base transcripts originate from *P2* and *P1* respectively, confirming our previous observation that both *gal* promoters are partially active in the absence of the normal −35 region sequences [10]. The 35 base transcript must start at +14/+15 confirming that the *P3* promoter becomes active when both *P1* and *P2* are disrupted.

To assess the strength of the *P3* promoter we determined the time for formation of transcriptionally-competent complexes after RNA polymerase was added to *gal* DNA carrying the mutation at −12. To do this, polymerase was mixed with DNA and, at different time points, a mix of the four nucleoside triphosphates was added together with heparin. Fig.3a–f shows that, under our conditions, incubation times of at least 15 min are required for the formation of complexes that initiate transcription at +14/+15.

3.2. *Interactions between CRP and RNA polymerase at the* gal *regulatory region carrying the mutation at −12*

Although the mutation at −12 destroys both the *P1* and *P2* −10 hexamer sequences, it does not affect the site for CRP binding which is located between −30 and −50 [13]. We and others have shown that, with the wild type *gal* promoter region, cAMP-CRP switches transcription from *P2* to *P1* [1,11]. Fig.3h–m shows that when cAMP-CRP is added to *Pst*I-*Hin*dIII fragments carrying the mutation at −12, transcription from the start point at +14/+15 is blocked and RNA

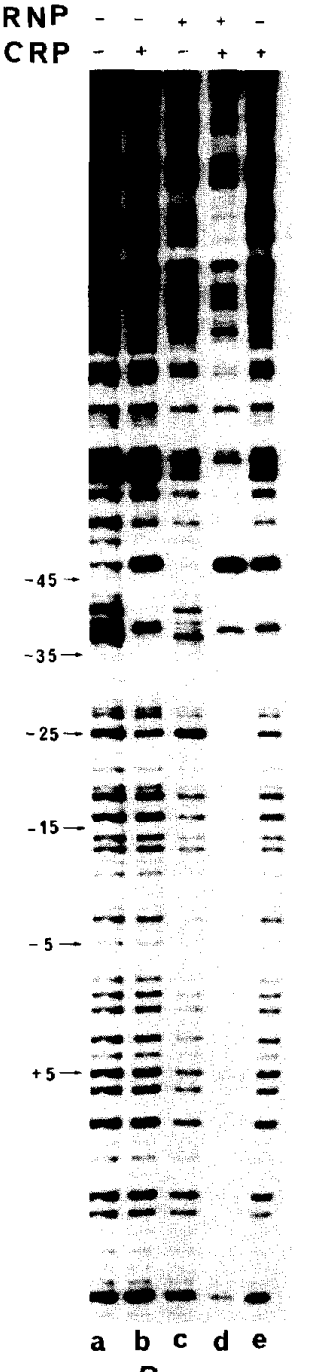

Fig.4. DNase I footprint analysis. Combinations of RNA polymerase and cAMP-CRP as shown were incubated with end-labelled *Pst*I-*Hin*dIII fragment carrying the mutation at −12. After incubation with DNase I the pattern of bands produced was analysed on a calibrated sequencing-type polyacrylamide gel and revealed by autoradiography as shown. The numbers refer to positions in the *gal* promoter sequence.

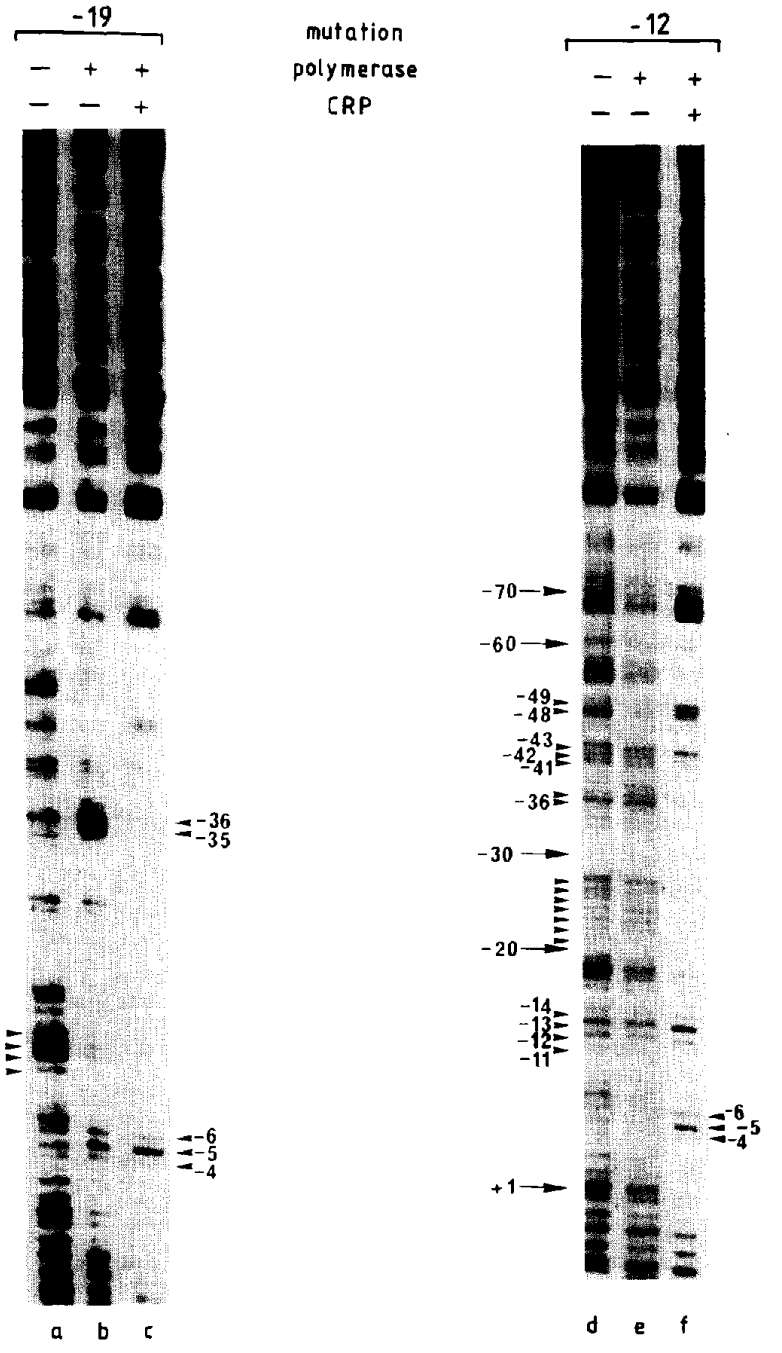

Fig.5. Footprint analysis using orthophenanthroline-copper ions as an artificial nuclease. End-labelled *Pst*I-*Hin*dIII fragments carrying a mutation at −19 (lanes a–c) or at −12 (lanes d–f) were incubated with combinations of RNA polymerase and cAMP-CRP as shown and subjected to footprint analysis. The pattern of fragments generated were analysed on calibrated polyacrylamide gels that were revealed by autoradiography.

polymerase switches to the *S1* start at +1. However, formation of complexes that initiate transcription at *S1* is slow, with a half-time of 20–25 min under these conditions. In parallel experiments with the wild type *gal P1* promoter, transcriptionally-competent complexes form in less than 2 min in the presence of cAMP-CRP ([11] and fig.3n–o).

Fig.4 shows an analysis, using DNase I footprinting [14], of the binding of CRP and RNA polymerase to *gal* promoter DNA carrying the mutation at −12. Lane a shows the pattern of bands formed when end-labelled *Pst*I-*Hin*dIII DNA fragments covering the *gal* promoter region were treated with DNase I. Lane b shows that cAMP-CRP alone binds normally giving a footprint between −30 and −50 identical to that reported [11,15,16]. Lane c shows that RNA polymerase alone gives a very weak footprint, unlike that found when RNA polymerase binds to *P2* [11,16,17], confirming that the complex between RNA polymerase and *P3* is unstable. Lane d shows the protection due to both RNA polymerase and cAMP-CRP: clearly a tight complex is formed that gives extensive protection from +17 to −65. Comparison of this footprint with published data for the CRP-induced binding of RNA polymerase to *gal P1* [11,16,17] shows that, in this case, RNA polymerase is indeed bound at *P1*. To confirm this, we compared the complexes formed between cAMP-CRP, RNA polymerase and *gal* DNA carrying either the mutation at −12 or a mutation at −19 that inactivates *P2* such that *P1* is the sole functional promoter [6,8]. In this experiment we used the nuclease activity of 1,10-phenanthroline-copper ions [12] as an alternative to the more bulky DNase I. The results in fig.5 show that, in the presence of cAMP-CRP, RNA polymerase creates an extensive footprint that is identical whether the *gal* promoter carries the mutation at −12 or at −19 (lanes c and f). In contrast, with polymerase alone, a strong footprint is seen with the mutation at −19 (lane b, due to polymerase at *P1*), whereas a weak footprint is found with the mutation at −12 (lane e, due to polymerase at *P3*).

4. DISCUSSION

Amongst the many point mutations that have been isolated throughout the *E. coli gal* operon promoter region, the mutation at −12 is unique as it simultaneously inactivates both *P1* and *P2* and almost totally eliminates expression in vivo from the *gal* regulatory region. However in the absence of *P1* and *P2* a third promoter, which we have labelled *P3*, is unmasked. This promoter initiates transcription slowly at +14/+15 and, like many weak promoters, forms complexes with RNA polymerase that give weak footprints. Most probably the low expression from this promoter in vivo is a reflection of its weakness. Interestingly the DNA sequence upstream of this start point carries plausible correctly spaced −10 and −35 hexamer sequences, 5′-TACCAT-3′ and 5′-TTGTTA-3′, that, according to the established consensus, should create a stronger promoter [18].

We and others have noted the striking similarities between the arrangement of signals in the *lac* and *gal* regulatory regions [11,19,20]: in both cases there are two overlapping promoters, *P1* and *P2*, one of which is stimulated and the other of which is blocked by cAMP-CRP. This work suggests a third common feature, as the *gal P3* promoter that initiates transcription at +14/+15, is strikingly similar to the P115 promoter in the *lac* operon that initiates transcription at +13 [21]. To explain the existence of this 'downstream' promoter, which was found after an analysis of transcripts made by the *lac* operon regulatory region carrying a mutation at +1, Peterson and Reznikoff [21] suggest three possible explanations: either it is an accident, an evolutionary remnant or an 'antenna' allowing loosely bound polymerase to associate near *P1*, the major functional promoter. In our view the existence of a similar promoter in the case of the *gal* operon allows us to eliminate the accident hypothesis but, to date, we have no evidence to favour one of the alternatives.

In their study, Peterson and Reznikoff [21] noted that cAMP-CRP diverted RNA polymerase from the alternative P115 promoter to *lac P1*. Similarly, in the *gal* case, cAMP-CRP blocks the *P3* promoter and stimulates, albeit slowly, open complex formation at *gal P1*. Presumably the lack of expression from the *gal* regulatory region carrying the mutation at −12 in vivo in crp^+ cya^+ cells is due to the fact that cAMP-CRP acts too slowly. However, from our studies in vitro we can draw

two clear conclusions concerning the action of CRP. Firstly, it is generally thought that CRP stimulates transcription by acting at an early stage during the recognition of promoter DNA by RNA polymerase probably by directly contacting incoming polymerase [20,22–24]. The −10 hexamer sequence cannot be essential for this step as CRP acts even when the hexamer is altered from 5′-TATGGT-3′ to 5′-CATGGT-3′. Secondly, the footprint data (figs 4 and 5) clearly show that CRP tightens polymerase binding to *gal* promoter DNA carrying the mutation at −12. We have previously studied CRP binding to the *gal* promoter region carrying mutations or deletions that weakened CRP binding: we showed that in the presence of RNA polymerase, CRP binding was tightened [11]. Here we have studied polymerase binding to the *gal* promoter region carrying a mutation that weakens RNA polymerase binding: we have shown that in the presence of CRP, polymerase binding is tightened. Hence there is a clear cooperativity between CRP and RNA polymerase binding to *gal* promoter DNA. We suggest that this is due to a direct contact between the two proteins that most probably is essential for CRP to stimulate transcription initiation.

ACKNOWLEDGEMENTS

We are grateful to Henri Buc for helpful discussions and interest throughout this work. The experiment shown in fig.3 was originally performed by Christine Perret and Michael Goranssön on an EMBO-sponsored practical course. This work was supported by SERC with grant no.GRD53494 and a research studentship to S.P.

REFERENCES

[1] Musso, R.E., Di Lauro, R., Adhya, S. and De Crombrugghe, B. (1977) Cell 12, 847–854.
[2] Aiba, H., Adhya, S. and De Crombrugghe, B. (1981) J. Biol. Chem. 256, 11905–11910.
[3] Busby, S., Aiba, H. and De Crombrugghe, B. (1982) J. Mol. Biol. 154, 211–227.
[4] Busby, S. and Dreyfus, M. (1983) Gene 21, 121–131.
[5] Busby, S., Truelle, N., Spassky, A., Dreyfus, M. and Buc, H. (1984) Gene 28, 201–209.
[6] Bingham, A., Ponnambalam, S., Chan, B. and Busby, S. (1986) Gene 41, 67–74.
[7] Kuhnke, G., Krause, A., Heibach, C., Gieske, U., Fritz, H.-J. and Ehring, R. (1986) EMBO J. 5, 167–173.
[8] Busby, S., Spassky, A. and Chan, B. (1987) Gene 53, 145–152.
[9] Ponnambalam, S. and Busby, S. (1987) FEBS Lett. 212, 21–27.
[10] Ponnambalam, S., Webster, C., Bingham, A. and Busby, S. (1986) J. Biol. Chem. 261, 16043–16048.
[11] Spassky, A., Busby, S. and Buc, H. (1983) EMBO J. 3, 43–50.
[12] Sigman, D., Spassky, A., Rimsky, S. and Buc, H. (1985) Biopolymers 24, 183–197.
[13] Kolb, A., Busby, S., Herbert, M., Kotlarz, D. and Buc, H. (1983) EMBO J. 2, 217–222.
[14] Galas, D.J. and Schmitz, A. (1978) Nucleic Acids Res. 5, 3157–3170.
[15] Taniguchi, T., O'Neill, M. and De Crombrugghe, B. (1979) Proc. Natl. Acad. Sci. USA 76, 5090–5094.
[16] Shanblatt, S. and Revzin, A. (1983) Proc. Natl. Acad. Sci. USA 80, 1594–1598.
[17] Taniguchi, T. and De Crombrugghe, B. (1983) Nucleic Acids Res. 11, 5165–5180.
[18] Hawley, D. and McClure, W. (1983) Nucleic Acids Res. 11, 2237–2256.
[19] Malan, T.P. and McClure, W. (1984) Cell 39, 173–180.
[20] Malan, T.P., Kolb, A., Buc, H. and McClure, W. (1984) J. Mol. Biol. 180, 881–909.
[21] Peterson, M. and Reznikoff, W. (1985) J. Mol. Biol. 185, 525–533.
[22] De Crombrugghe, B., Busby, S. and Buc, H. (1984) Science 224, 831–837.
[23] Meiklejohn, A. and Gralla, J. (1985) Cell 43, 769–776.
[24] Busby, S. and Buc, H. (1987) Microbiol. Sci. 4, in press.