Volume 11 Number 6 1983 Nucleic Acids Research

Spacing mutations between the *Escherichia coli* p_{BAD} RNA polymerase binding site and the *araC* (I) induction site

William Hendrickson, Joanne Petithory and Robert Schleif

Biochemistry Department, Brandeis University, Waltham, MA 02254, USA

Received 18 November 1982; Revised and Accepted 11 February 1983

ABSTRACT

Mutations in the Escherichia coli promoter p_{BAD} have been constructed which alter the spacing of the adjacent RNA polymerase and araC inducer protein binding sites. While deletion of a single base-pair or small insertions do not detectably affect araC protein binding to DNA and they do not alter the conserved sequence of the RNA polymerase binding site, stimulation of p_{BAD} in vivo is greatly reduced. The experiments suggest that the distance or angle between the two proteins on the DNA is critical for promoter function.

INTRODUCTION

Transcription from the Escherichia coli L-arabinose operon promoter, p_{BAD}, requires araC protein in the presence of arabinose (1-3). AraC protein represses its own synthesis from p_C (4), and in the absence of arabinose, represses transcription from p_{BAD} (2,3,5). Both p_C and p_{BAD} are stimulated by the cAMP receptor protein (CRP)-cAMP complex (2-4,6). Between the divergently oriented promoters, the DNA binding sites for these effectors occupy a 147 base pair noncoding region in the order: RNA polymerase at p_C, which overlaps araC protein repressor (O); cAMP receptor protein; araC protein inducer (I); and RNA polymerase at p_{BAD} (7,8). Although the mechanism by which these proteins stimulate transcription is not known, models for positive activation encompass two extremes. The proteins may interact directly (9,10), or the effects may be indirect, mediated through changes in the DNA structure over significant distances (11,12). If protein-protein contacts or very short-range DNA alterations are necessary for activation of RNA polymerase, the distance and rotational relationship on the DNA between araC protein and polymerase should be critical. On the other hand, if araC and/or CRP·cAMP proteins exclusively affect DNA structure over appreciable distances, promoter activity should not be greatly impaired by small changes in the relative positions of the

binding sites. In order to test the alternatives, DNA was inserted or deleted between araC inducer and p_{BAD} polymerase binding sites. While the mutations do not detectably affect araC protein binding to DNA and they do not alter the conserved sequence of the RNA polymerase binding site, stimulation of p_{BAD} *in vivo* is greatly reduced or eliminated.

MATERIALS AND METHODS

Plasmids and strains

Plasmid pTD3 was constructed by Teresa Dunn (M.S. in preparation). It contains a 440 base-pair fragment with the arabinose operon control region flanked by HindIII and EcoRI linkers. This piece has been insertd into the EcoRI-HindIII large fragment of the promoter-cloning vehicle pKO1 (13). The resulting construct places the galK gene of pKO1 under control of p_{BAD} and retains ampicillin resistance (Fig. 1).

TMD29 is a galK derivative of C600 with a constitutive araC protein, $araC^c$ (T. Dunn, unpublished). SH313 is galK, araB, thr, his, thi, dcm, Str^r (S. Hahn, unpublished).

General methods

Deoxynucleotide triphosphates were obtained from PL Biochemicals, γ-^{32}P ATP was from ICN. All enzymes were from Bethesda Research Labs, New England Biolabs and PL Biochemicals.

All media, plates, and standard procedures including plasmid ligation, transformation and CsCl gradient plasmid isolation were as described in Schleif & Wensink (14).

Base insertions and deletions

Samples of BamHI cleaved pTD3 DNA (3 µg) were incubated with 0.5 unit DNA polymerase I large fragment in 30 µl of 50 mM Tris·HCl, pH 7.5, 10 mM MgC_2, 7 mM β-mercaptoethanol and 200 µM dNTP's for 30 minutes at room temperature (Fig. 2). After phenol extraction, the samples were digested with 300 Vogt units/ml SI nuclease in 50 µl of 30 mM Na Acetate, pH 4.5, 200 mM NaCl, 5 mM $ZnCl_2$ for 30 minutes at room temperature (15). After ligation and transformation into TMD29 ($AraC^c$, $GalK^-$) cells were plated on MacConkey galactose ampicillin plates and the GalK phenotype scored. Several $GalK^+$ and $GalK^-$ transformants from each of the three reactions were screened for loss of the BamHI restriction site and also tested for the presence of the 440 base-pair HindIII-EcoRI fragment. Several of the mutants which had undergone large deletions or insertions were discarded.

DNA isolation and sequencing

A rapid plasmid preparation was modified from that described by Holmes & Quigley (16) as follows: 10 ml cultures were amplified by incubating 18 hours in 50 µg/ml chloramphenicol; samples were boiled for 45 seconds in lysis buffer containing 1% Triton X-100 and 500 µg/ml lysozyme and centrifuged 20 minutes at 20,000 x g. The resulting supernatants were further deproteinized by incubating 10 min. on ice in 0.6% diethylpyrocarbonate and 10 min. at 70°, followed by centrifugation at 10,000 x g for 5 minutes. For sequencing, RNA was removed by passing samples over 3 ml columns of Bio gel A50M agarose (Biorad) equilibrated in 10 mM Tris·HCl pH 8.0, 0.5 M NaCl, 1 mM EDTA. DNA was labeled at the HindIII restriction site (+44) with γ^{32}P-ATP using bacterial alkaline phosphatase and polynucleotide kinase, digested with EcoRI and the 440 base-pair fragment isolated as described by Maxam & Gilbert (17). Sequences of mutants were first determined using DNA isolated by this rapid preparation, then the sequence of the entire relevant region was confirmed using CsCl purified DNA (14).

DNase I protection and galactokinase assay

DNase I protection was as described (7). Approximately 20 ng of ara DNA fragment was incubated for 10 minutes at 37° with araC protein in 100 µl 20 mM Tris·HCl pH 7.5, 50 mM KCl, 2.5 mM $MgCl_2$, 0.1 mM DTT, 100 mM L-arabinose and 25 µg/ml BSA. DNase I (0.08 µg/ml) was added for 30 seconds and the reaction stopped by adjusting to 2.5 M NH_4-acetate, 3 mM EDTA, 10 µg/ml tRNA. AraC protein was purified as in (18,19).

Galactokinase assays were as in (13). Cultures were grown to a cell density of 2 to 4 x 10^8 in M10 minimal medium, assays performed on cells concentrated five-fold, and reactions terminated with 10 mM EDTA after 30 minutes. The values shown are averages of four samples; all replicate samples were within approximately ∓20% of each other.

RESULTS

Base insertions and deletions

The spacing between the p_{BAD} RNA polymerase and araC protein inducer DNA binding site, araI, was altered by taking advantage of a unique BamHI restriction site on plasmid pTD3. This is located at the boundary between the araI site and the RNA polymerase binding site (Fig. 1). The 3' ends of the BamHI-cut DNA were partially filled in with DNA polymerase I large fragment in the presence of two, three, or all four required

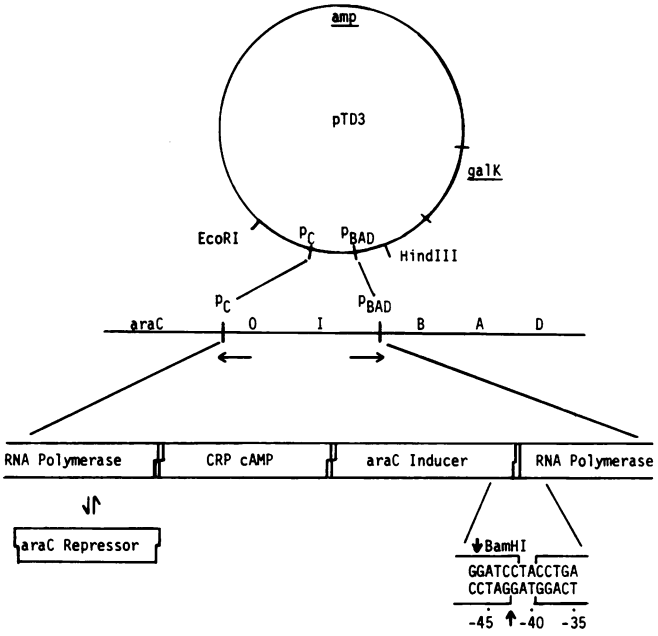

Figure 1. The arabinose operon. Top: plasmid with ara P_{BAD} fused to the galK gene. Arrows indicate the direction of transcription of the arabinose operon. DNA sites protected from nuclease digestion are enclosed in blocks. The BamHl site used for mutagenesis is shown at bottom.

deoxynucleotides (Fig. 2) and any remaining 5' protruding ends were removed with S1 nuclease. The resulting blunt ends were ligated and plasmid transformed into strain TMD 29 ($AraC^c$, $GalK^-$).

Plasmids from reactions one and two (Fig. 2) would be expected to contain a 4 base-pair and a two base-pair insertion, respectively, while reaction three should reconstitute the wild type promoter. Although $GalK^+$ and $GalK^-$ transformants were obtained from all three samples, all plasmids containing the 440 base pair insert and having lost the BamHl site were $GalK^-$ (P_{BAD}^-). Isolates from each of the mutant classes were selected for sequence analysis. The expected GATC and AT insertions from classes one and two were found, and a $GalK^-$ mutant from class three contained a deletion of the deoxyadenosine at position -45 presumably from overdigestion by the S1 nuclease. All three mutations occur just within the DNA region protected from nuclease digestion by araC protein plus arabinose.

Promotor activity

The mutants were further characterized by quantitative measurements of

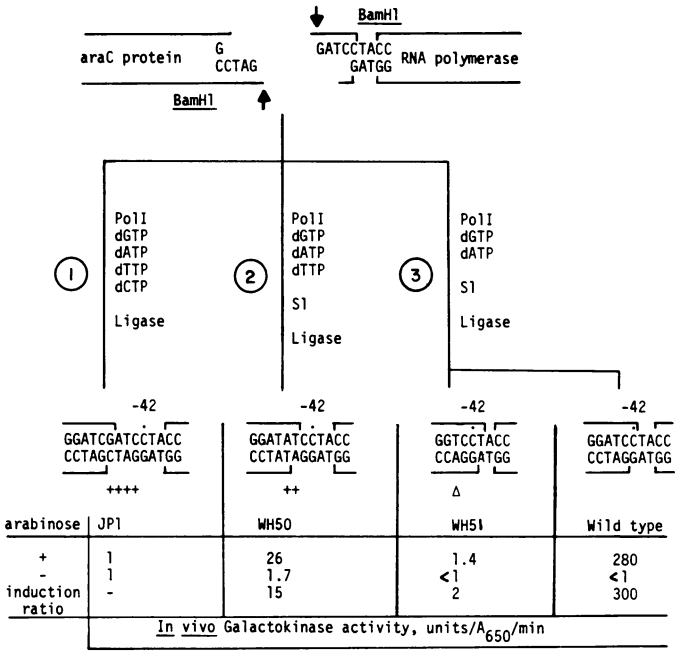

Figure 2. Construction, structure and properties of deletion and insertion mutations. Sequences protected from nuclease digestion are indicated by brackets as in Figure 1. Number above DNA sequences indicates distance from transcriptional start site. Inserted or deleted base-pairs are identified by + or Δ below the sequences. Both wild type and the single base-pair deletion were obtained from procedure 3. Below is shown basal activity and inducibility of the wild type and mutant *ara* promoters.

galactokinase activity *in vivo* (Fig. 2). Plasmid DNA was transformed into GalK$^-$, AraC$^+$ strain SH313, and enzyme levels determined for arabinose induced and uninduced cultures. The four base-pair insertion mutant, JP1, and the deletion, WH51, were not inducible and had enzyme levels similar to that of the uninduced wild type. The two base-pair insertion mutant, WH50, was different. Its uninduced level was only slightly higher than the uninduced wild type, but its induced level was 10% the level of induced wild type. On galactose indicator plates this mutant appeared p_{BAD}^-(Gal$^-$) when transformed into an *araC*^{c} strain, but it was slightly p_{BAD}+ in the presence of AraC$^+$.

DNase I protection

To determine if the mutations impair *araC* protein binding to the *araI* site, DNase I protection studies were performed. In parallel experiments, mutant and wild type 440 base-pair *ara* fragments were incubated with

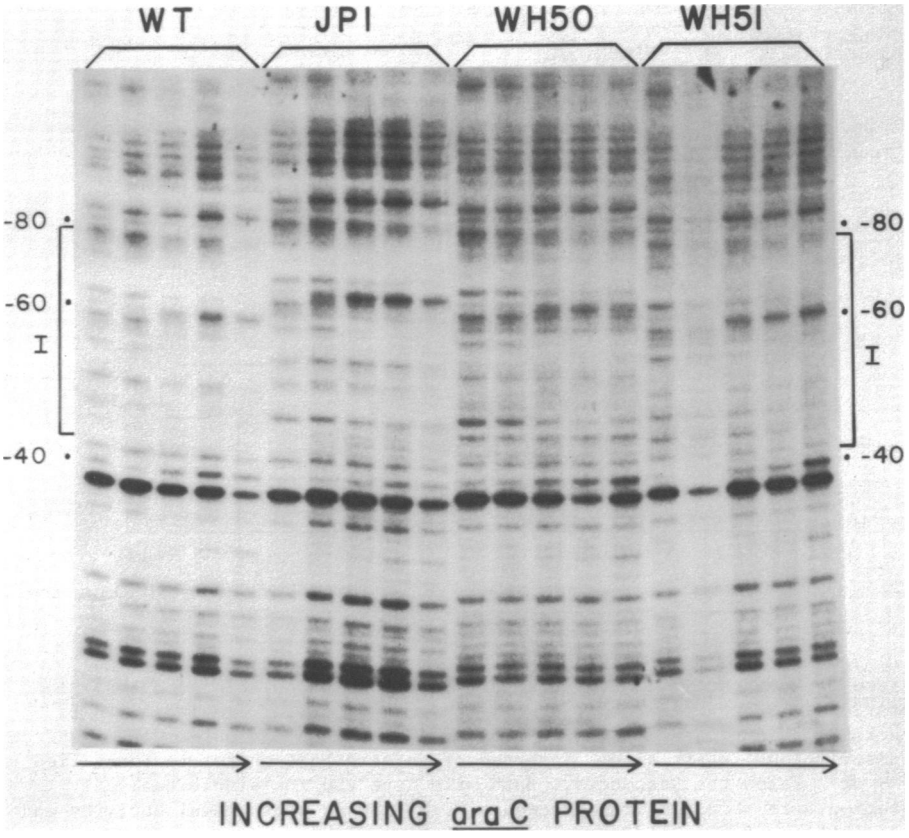

Figure 3. Nuclease protection of the arabinose control region by araC protein. Brackets indicate DNA region protected by araC protein bound to the I site. The amount of protein added to each sample increases as indicated by the arrows: 0; 1 ng; 3 ng; 6 ng; and 18 ng.

increasing levels of araC protein in the presence of arabinose (Fig. 3). The level of araC protein required for protection from DNase I was the same in all cases. Also, the pattern of protected regions and enhanced cleavage of DNA was the same among the samples. These results suggest that the mutations do not directly affect the binding of araC protein to the araI site.

DISCUSSION

These experiments show that 2 base-pair or 4 base-pair insertions or a 1 base-pair deletion at the junction of the RNA polymerase and araC protein DNA binding sites strongly reduce ara P_{BAD} promoter activity. The deletion and larger insertion completely eliminate the ability of araC protein to

induce the promoter, and the 2 base-pair insertion reduces induction to 10% of the wild type level. All the mutations are located a few base-pairs within the DNA region protected by araC protein.

The mutations could reduce promoter activity by: 1) altering the distance or angle between RNA polymerase and araC protein on the DNA; 2) reducing the interaction of araC protein with its DNA binding site; 3) changing DNA contact sites required for RNA polymerase binding. Several facts suggest the first alternative for the insertion mutations, with somewhat less evidence supporting this possibility for the deletion mutation. None of the mutations measurably affect the binding of araC protein to the DNA in vitro. Both the level of protein required for binding and the pattern of protection from nuclease digestion remain unchanged, however it is not yet known whether these experiments, under the conditions employed, could detect some physiologically important alterations in binding. While all the mutations lie just within the DNA region protected by araC protein (7,8), the insertion mutations duplicate DNA sequence so that the binding sites remain largely intact (Fig. 2). In particular, the four base-pair insertion would leave the sequence unchanged upstream of the RNA polymerase binding site to position -46. From the araI site, the sequence would be unchanged downstream to position -43. It seems unlikely that RNA polymerase would absolutely require DNA contacts upstream of position -46 since araC protein strongly protects both DNA strands in this region.

Sequence comparisons between p_{BAD} and the similarly induced arabinose transport operon promoters, p_E and p_{FG}, suggest that the specific bases altered by the insertion and deletion mutations are not essential for RNA polymerase or araC protein binding (Detailed sequence comparisons are made elsewhere; Stoner, Hendrickson and Schleif, manuscript submitted). Both of these promoters contain araC protein and CRP•cAMP binding sites similar to those of p_{BAD}, and the three promoters contain regions of strong sequence homology. While p_E is most closely related to p_{BAD}, with 26 of 31 identical base-pairs from -30 to -44 of the RNA polymerase and -57 to -71 of the araC protein binding regions, the sequences from -45 to -56 are strikingly divergent (2/12 homology). The -45 adenosine deleted in mutant WH51 is not found in p_E or p_{FG}, and no common bases occur among all three promoters within this region.

The evidence presented here is not consistent with a model for positive activation of promoters in which structural alterations in DNA are

transmitted to the RNA polymerase binding site over a significant distance from the effector binding site. Consequently, protein-protein contacts and/or very short range DNA alterations are involved in the induction of the arabinose promoter by araC protein. Recent studies have shown that the DNA between the -10 and -35 RNA polymerase binding regions also appears to act as a spacer (20,21). It is the length of this DNA rather than its sequence which is important for proper protein recognition. Studies of another positively regulated promoter, λPRM, suggest protein interactions between RNA polymerase and λ repressor are required for induction (9,10). For ara P_{BAD} more direct measurements will be required to ultimately prove whether or not specific contacts occur between the proteins involved.

ACKNOWLEDGEMENT

This work was supported by United States Public Health Service research grant GM-18277. This paper is publication no. 1431 from the Department of Biochemistry, Brandeis University, Waltham, Massachusetts 02254.

REFERENCES

1. Sheppard, D. and Englesberg, E. (1967) J. Mol. Biol. 25, 443-454.
2. Greenblatt, J. and Schleif, R. (1971) Nature New Biol. 233, 166-170.
3. Wilcox, G., Neuris, P., Bass, R. and Englesberg, E. (1974) J. Biol. Chem. 249, 2946-2952.
4. Casadaban, M. (1976) J. Mol. Biol. 104, 557-566.
5. Englesberg, E., Squires, C. and Meronk, F. (1969) Proc. Natl. Acad. Sci. U.S.A. 62,1100-1107.
6. Hirsh, J. and Schleif, R. (1977) Cell 11, 545-550.
7. Ogden, S., Haggerty, D., Stoner, C., Kolodrubetz, D. and Schleif, R. (1980) Proc. Natl. Acad. Sci. U.S.A. 77, 3346-3350.
8. Lee, N., Gielow, W. and Wallace, R. (1981) Proc. Natl. Acad. Sci. U.S.A. 78, 752-756.
9. Guarente, L., Nye, J., Hochschild, A. and Ptashne, M (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 2236-2239.
10. Johnson, A., Poteete, A., Lauer, G., Sauer, R. Ackert, G. and Ptashne, M. (1981) Nature 294, 1-7.
11. Dickson, R., Abelson, J., Barnes, W. and Reznikoff, W. (1979) Science 187, 27-35.
12. Ebright, R. and Wong, . (1981 Proc. Natl., Acad. Sci. U.S.A 78, 4011-4015.
13. McKenney, K., Shimatake, H., Court, D., Schmeissner, U., Brady, C. and Rosenberg, M. (1981) Gene Amplification and Analysis, Vol.II: Analysis of Nucleic Acids by Enzymatic Methods, J. Chirikjian and T. Papas, Ed. pp383-415, Elsevier-North Holland.
14. Schleif, R. and Wensink, P. (1981) Practical Methods in Molecular Biology, Springer-Verlag, New York.
15. Vogt, V. (1973) Eur. J. Biochem. 33, 192-200.
16. Holmes, D. and Quigley, M. (1981) Anal. Biochem. 114, 193-197.
17. Maxam, A. and Gilbert, W. (1980) Methods in Enzymology, Vol. 65, L. Grossman and K. Moldave, eds., pp. 499-560, Adademic Press, New York.
18. Steffen, D. and Schleif, R. (1977) Mol. Gen. Genet. 157, 333-339.
19. Schleif, R. and Favreau, A. (1982) Biochem. 21, 778-782.
20. Mandecki, W. and Reznikoff, W. (1982) Nuc. Acids. Res. 10, 903-912.
21. Stefano, J. and Gralla, J. (1982) Proc. Natl. Acad. Sci. U.S.A. 79, 1069-1073.