Cloning of binding sequences for the *Escherichia coli* transcription activators, FNR and CRP: location of bases involved in discrimination between FNR and CRP

A.I.Bell, K.L.Gaston, J.A.Cole and S.J.W.Busby*

University of Birmingham, School of Biochemistry, PO Box 363, Birmingham B15 2TT, UK

ABSTRACT
      Expression from the E.coli melR promoter (pmelR) is normally totally
dependent on the transcription activator protein, CRP. We describe
experiments with a genetically engineered DNA fragment carrying pmelR in
which the wild type CRP binding site was replaced with synthetic
oligonucleotides containing either FNR or CRP binding sequences. When the
synthetic oligonucleotide contains the 22 bp consensus for FNR binding sites,
expression from pmelR is dependent on FNR but not CRP. Single changes at
either of two symmetrically-related positions create sites that are
recognised by both FNR and CRP. Changes at both positions result in a site
that is not recognised by FNR but which binds CRP tightly.

INTRODUCTION

      The Escherichia coli fnr gene encodes a transcription activator, FNR,

that is responsible for the regulated expression of many genes necessary for

adaptation to growth in anaerobic conditions (1). Spiro and Guest (2)

identified nucleotide sequences that were common to a number of FNR-

dependent promoters  and suggested that they corresponded to FNR binding

sites in these cases: comparison of these and further sequences allowed a

'consensus' sequence to be deduced (2-5). This consensus sequence contains

an inverted repeat, strongly suggesting that FNR binds as a dimer with each

subunit recognising one arm of the repeat. To date, however, there is little

direct evidence for this, as no system for studying FNR activity in vitro has

been found.

      Genetic evidence that the proposed consensus sequence constitutes an FNR

binding site has come from recent analysis of the effects of deletions and

point mutations at the FNR-dependent nirB promoter (3, 6). Here we describe

the synthesis of an oligonucleotide containing the consensus sequence for FNR

binding and its cloning upstream of a promoter that is not usually dependent

on FNR. We show that transcription initiation from this new promoter becomes

FNR-dependent.

It has long been known that the primary sequence of FNR is homologous to CRP, another transcription activator (7). Indeed Spiro and Guest demonstrated that 3 simple changes to the FNR sequence were sufficient for FNR to substitute for CRP at a number of promoters (8). Interestingly the proposed consensus sequence for FNR binding sites is similar to the known sequence for CRP binding sites (2,3). Here we show that CRP does not recognise the FNR- binding consensus sequence. However a one base pair transversion creates a site that is recognised by CRP and FNR. A second change related to the first by the twofold symmetry of the sequence makes a site that is recognised by CRP only.

MATERIALS AND METHODS

Standard methods for handling recombinant DNA were used throughout the work as described by Maniatis et al (9). The construction of the fragment carrying the melR promoter, together with the insertion of a BamHI linker at -25 has been described by Webster et al (10). BamHI-HindIII fragments carrying the melR promoter were cloned into the lacZ expression vector pAA182 (11). Transcription from the melR promoter was measured as β-galactosidase expression in the Δlac strains M182, M182 crp, M182 fnr and M182 crp fnr as previously described (3, 10): the crp allele is a total deletion of the crp gene whilst the fnr allele is a Tn10 insertion.

Synthetic oligonucleotides were synthesised by the University of Birmingham Macromolecular Analysis Service (Alta) as previously described (3). Complementary oligonucleotides carrying BamHI cohesive ends were hybridised and cloned into the BamHI site located upstream of -25 at the melR promoter. Nucleotide sequences were checked after transfer to mp10 and mp11 vectors using the Amersham cloning and sequencing kits.

Gel retardation assays to measure CRP binding to different sites were as described by Gaston et al (12) using purified CRP provided by Dr A Kolb of the Pasteur Institute, Paris. For these experiments, 110 bp EcoRI-HindIII fragments were purified from plasmids in which the FF, CC, CF and FC sites were cloned upstream of the nirB promoter (A Bell, unpublished constructions): the fragments were labelled using $\alpha$ $^{32}P$ dATP and Klenow enzyme.

RESULTS

The consensus sequence for CRP binding sites is 5' aAaTGTGAtctagaTCACAtTt 3' (3,12,13) where bases present at 50% or more of sites are in capitals.

The aim of these experiments was to synthesise this sequence chemically, insert it upstream of a target promoter, and then determine the minimum number of changes required to alter the dependence of the promoter on activation by CRP to activation by FNR. For these experiments we chose the promoter of the melR gene which is normally absolutely dependent on CRP binding to a site located from  -52 to -31 bp upstream of the transcription start (10). This promoter, which is illustrated on the top line of Fig. 1, was isolated on a 388 bp fragment with a BamHI linker located upstream of -66 and a HindIII linker downstream of +266. To clone synthetic CRP and FNR sites we used the D25 derivative of this fragment (illustrated on line 2 of Fig. 1) where BAL31 exonuclease had been used to remove the CRP binding site and a BamHI linker had been placed just upstream of -25. Synthetic oligonucleotides could then be cloned into this BamHI site such that the nucleotide sequence from -52 to -31 would correspond to the CRP consensus sequence (line 3 of Fig. 1). However, because in previous work we found a perfect fit difficult to maintain in $crp^+$ cells (12), we synthesised a sequence that varied from the 22 bp consensus at positions 10 and 13, where bases are poorly conserved between different sites. Our attempts to clone the nucleotide sequence corresponding perfectly to the consensus were hindered by the frequent spontaneous appearance of short deletions in the CRP-binding sequence. Further, we deliberately reduced the two-fold symmetry, in the site by introducing G residues at positions 21 and 22. Despite these modifications, the resulting sequence (line 4 of Fig. 1) corresponds well to the CRP consensus binding sequence:  because the 22 bp sequence contains sites for the binding of 2 CRP subunits, this site was labelled CC.

The "consensus" sequence for FNR binding sites is 5' AAAtTTGATaTAtATCAAaTTT 3' (line 8 of Fig. 1: see refs 2, 3). The main differences between this and the CRP consensus are at positions 5 and 18:  other differences are at positions such as 10 and 13 where the base is poorly conserved between different sites. Hence we resynthesised the CC sequence with a T at position 5 and/or an A at position 18. The double change gave the sequence shown in line 7 of Fig. 1 that resembles the FNR consensus binding site and thus was labelled FF. The oligonucleotides carrying single changes gave a CRP half site either upstream or downstream of an FNR half site and are labelled CF and FC, respectively (lines 5 and 6 of Fig. 1).

BamHI-HindIII fragments carrying the CC, CF, FC and FF sequences in the melR D25 promoter were cloned into pAA182 together with the full length melR promoter (D66) and the melR promoter carrying the D25 deletion (Fig. 1). The

pmelRD66
```
        BamH1 -66        -52                           -31                              -20        -10        +1    +266  HindIII
[GGATCC]G----GTGAAAACCGTGCTCCCACTCGCAGTCATCCTCCCTCACTCCTGCCATAATTCTGATATT----C[AAGCTT]
```

pmelRD25
```
BamH1                                                            -20        -10        +1    +266  HindIII
[GGATCC]CCCCTCACTCCTGCCATAATTCTGATATT----C[AAGCTT]
```

Consensus CRP site          aAaTGTGAtctagaTCACAtTt

pmelRD25 CC          AAATGTGATGTACATCACATGG

pmelRD25 CF          AAATGTGATGTACATCAAATGG
                    GGATCAGGT          ATCCCCCCTCACTCCTGCCATAATTCTGATATT----C[AAGCTT]

pmelRD25 FC          AAATTTGATGTACATCACATGG

pmelRD25 FF          AAATTTGATGTACATCAAATGG

Consensus FNR site          AAATTTGATaTAtATCAAaTTT

recombinant plasmids were transferred to the Δlac strain, M182, which is crp⁺ fnr⁺ and to three derivatives that are c̣ṛp̣ fnr⁺, crp⁺ f̲n̲r̲ and c̲r̲p̲ f̲n̲r̲, and β galactosidase activities were measured to determine the activity of the melR promoter in each case. Fig. 2A shows that expression from the melR promoter carrying the D25 deletion is minimal in all four backgrounds. With the FF site, expression is activated in the crp⁺ fnr⁺ and c̲r̲p̲ fnr⁺ backgrounds but is minimal in the crp⁺ f̲n̲r̲ and c̲r̲p̲ f̲n̲r̲ strains, showing that this promoter is stimulated by FNR but not by CRP. In contrast, expression from the CC site is found in the crp⁺ fnr⁺ and crp⁺ f̲n̲r̲ backgrounds, but not in the c̲r̲p̲ fnr⁺ or c̲r̲p̲ f̲n̲r̲ backgrounds showing that this promoter is stimulated by CRP but not by FNR. With the FC and CF sites, expression is stimulated in the crp⁺ fnr⁺, crp⁺ f̲n̲r̲ and c̲r̲p̲ fnr⁺ backgrounds but not in the c̲r̲p̲ f̲n̲r̲ strain, showing that these sites are recognised by both CRP and FNR.

Figs 2B and 2C show experiments to compare the degree of transcription activation by FNR and CRP respectively at the FF, CF, FC and CC promoters. In Fig. 2B activation by FNR is measured in c̲r̲p̲ cells grown anaerobically. The order of activation is FF>CF=FC>CC. In Fig. 2C activation by CRP is measured in f̲n̲r̲ cells grown in minimal medium without glucose. The order of activation is CC>FC=CF>FF: in these conditions the melR promoter (D66) is also active.

We have independently assessed the binding of CRP to the CC, FC, CF and FF sites using gel binding assays. Increasing concentrations of CRP were incubated with labelled fragments carrying each of the sites, and the samples were loaded on gels which were run to resolve free and bound fragments. Fig. 3 shows an autoradiogram of the gel. With the CC site all the fragment was

Figure 1. Nucleotide sequences studied in this work. The top line (pmelR D66) shows the nucleotide sequence at the melR promoter. The sequence shown corresponds to the upper strand and is numbered with the transcription startpoint as +1. Parallel lines above and below the sequence identify the -10 hexamer sequence. This promoter was cloned on a BamHI-HindIII fragment with a BamHI linker at -66 and a HindIII linker at +266 as indicated. The second line (pmelRD25) illustrates a derivative of this fragment where the BamHI linker was moved to -25. This BamHI site allows synthetic oligonucleotides to be cloned upstream of -25. The sequences of four such derivates, pmelRD25CC, pmelRD25CF, pmelRD25FC and pmelRD25FF are shown on lines 4, 5, 6 and 7 of the figure respectively. These constructions create pmelR derivatives with differing sequences in the zone from -31 to -52. The 22 nucleotide sequence in each of the derivatives is shown in the extended brackets and is compared to the 22 base consensus sequences for CRP or FNR binding sites (lines 3 and 8 respectively). Positions 5 and 18 in this 22 base sequence are highlighted. The centre of symmetry in these 22 base sequences is marked by a dot flanked by two bars.
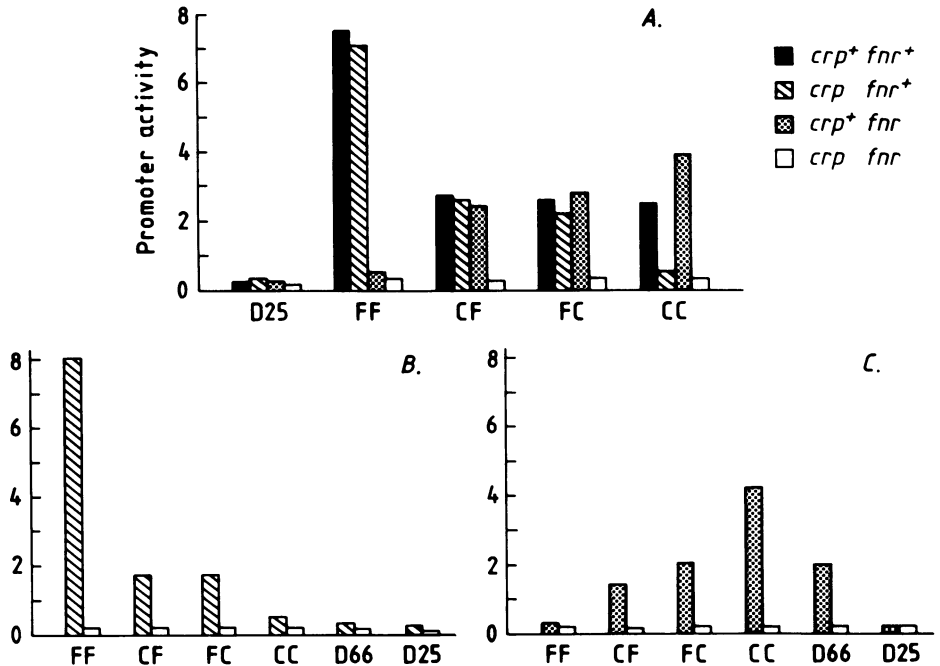
Figure 2. Promoter activity of pmelR derivatives. [A] The activities of the different promoters illustrated in Fig. 1 were measured in four derivatives of M182 that were $crp^+$ $fnr^+$, $crp$ $fnr^+$, $crp^+$ $fnr$ or $crp$ $fnr$: activities in the different strains are illustrated by the heights of bars shaded differently as shown in the inset. Activities were deduced from measurements of β-galactosidase activity in strains carrying pAA182 derivatives containing the different pmelR fragments as shown on the abcissa. Numbers on the ordinate refer to β galactosidase activities measured in standard units x $10^{-3}$ (3, 10). In this experiment cells were grown anaerobically in L broth containing 80 µg/ml ampicillin. The data represent the average of at least 3 independent determinations: the error is less than 10%. [B] The activity of a series of melR promoters is compared in $crp$ $fnr^+$ and $crp$ $fnr$ backgrounds. Cells were grown anaerobically in L broth containing 80 µg/ml ampicillin and 0.4% glucose. The data are presented as in section A. [C] The activity of a series of melR promoters is compared in $crp^+$ $fnr$ and $crp$ $fnr$ backgrounds. Cells were grown aerobically in standard minimal medium containing fructose as a carbon source (10, 12). The data are presented as in section A.

bound at the lowest CRP concentration (1nM) whereas with the CF and FC sites concentrations above 10nM were required. In contrast, we could demonstrate no specific binding of CRP to fragments carrying the FF site. As controls we also measured binding of CRP to a synthetic tight-binding site (AA') that was described by Gaston et al (12) and the naturally-occurring weak binding site
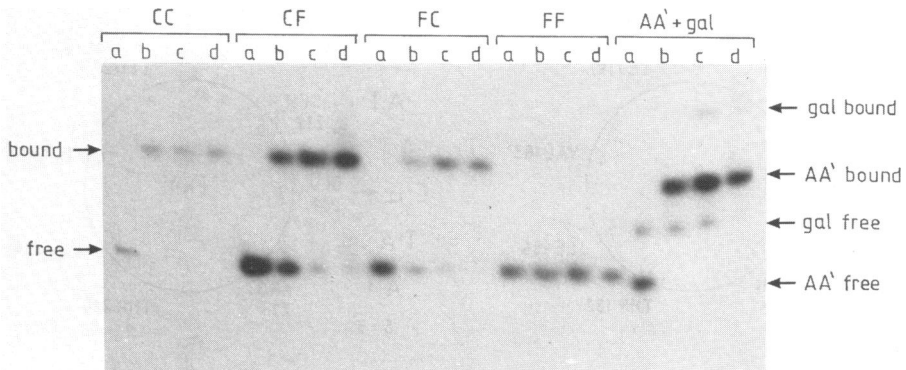
Figure 3. Gel binding assays to show CRP binding to different sites.
Fragments carrying the CC, CF, FC and FF sites illustrated in Fig. 1 were
incubated with different concentrations of purified CRP prior to gel
electrophoresis to separate free and bound fragments. The concentrations of
CRP used were: lanes a, 0; lanes b, 1 nM ; lanes c, 10 nM; lanes d, 100 nM.
As a control we used a mixture of EcoRI-HindIII fragments carrying the galP1
promoter either with the wild type CRP binding sequence or the tight-binding
AA' sequence (these control fragments are fully described in Figs. 2 and 5
respectively of Gaston et al (12)). The samples, gel and running buffer all
contained 0.2 mM cAMP. The figure shows an autoradiograph of the gel
annotated to show the position of migration of free and bound fragments.

at the galactose operon promoter (14). The results in Fig. 3 show that the
order of affinity of CRP for the different sites is CC=AA'>CF=FC>gal>FF.

DISCUSSION

Comparison of the nucleotide sequence at several FNR-dependent promoters
by a number of laboratories has suggested a consensus sequence for FNR
binding sites (2 - 5): this 22 bp sequence greatly resembles the CRP binding
consensus, containing an 11 bp inverted repeat. Here we have shown that this
sequence is sufficient to confer FNR-dependence on expression from the melR
promoter. Comparison of the 22 bp FNR consensus sequence with that for CRP
shows that the only well-conserved nucleotides at which the sequences vary
are at positions 5 and 18, which are related by the two-fold axis of symmetry
(Fig. 1). Here we have shown that symmetric T:A to G:C and A:T to C:G
transversions at positions 5 and 18 respectively are sufficient to convert
the FNR site (FF) which is recognised by FNR but not CRP, to a site (CC) that
binds CRP tightly but not FNR.

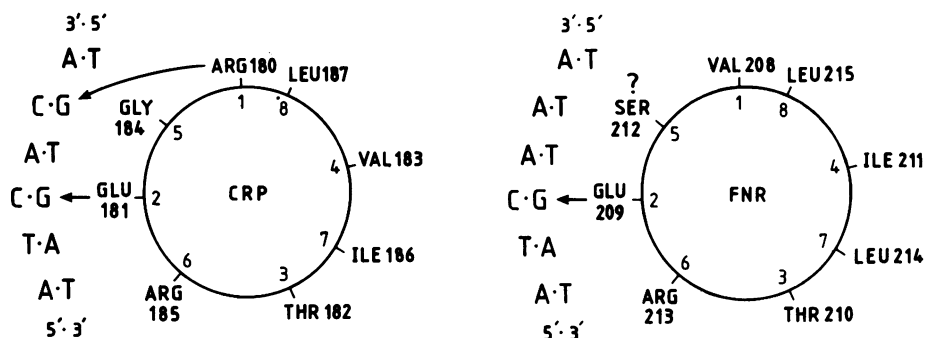The FC site, which contains a T:A at position 5 and a C:G at position

Figure 4. Comparison of CRP and FNR binding. The left hand diagram shows a helical wheel for the first eight residues of the DNA recognition helix F (amino acids 180-187) of CRP. The wheel is aligned with the core of the binding sequence (positions 4-9 of the consensus) with arrows to illustrate the principal contacts suggested by the Weber-Steitz-Ebright model (see text). The right hand diagram shows a helical wheel for the likely DNA binding region of FNR (amino acids 208-215) redrawn from Spiro et al (18). The FNR wheel is oriented with respect to the CRP wheel to emphasise sequence homologies. The FNR wheel is aligned with the consensus sequence from positions 4-9. The interaction of Glu 209 suggested by Spiro et al (18) is shown by an arrow. A question mark emphasises our ignorance of the precise interaction that occurs between the FNR helix and the T:A base pair at position 5.

18, binds both CRP and FNR. Similarly, both activators recognise the CF site with a G:C at position 5 and A:T at position 18, implying that both activators can recognise sites containing one poorly-binding half site. Clearly such sites could be used in vivo to ensure that a particular transcriptional unit is coupled to two activators, rather than one. In the crp background, expression dependent on the FC and CF sites is practically identical, strongly implying that FNR does indeed bind as a dimer with the final complex being symmetric around the centre of the site. Thus the organisation of bound FNR subunits appears similar to the organisation of CRP which also binds as a symmetric dimer (12). An interesting possibility arising from this is that FNR and CRP may form heterodimers which are inactive. One consequence of this may be that expression dependent on the FC and CF sites is not increased in the crp[+] fnr[+] host compared to the crp fnr[+] and crp[+] fnr backgrounds: if CRP and FNR acted independently their effects would be additive.

Our results clearly indicate that interactions involving bases at positions 5 and 18 in the 22 bp binding sequence are crucial in determining

the specificity of the binding of FNR and CRP to their cognate sites. The
genetic and crystallographic data currently available for CRP suggest that
the G:C base pair at position 5 (and position 18) interacts with Arg-180
(position 1 of the recognition helix) whilst the G:C at position 7 (and 16)
interacts with Glu-181 (position 2) (15-17). This is illustrated
schematically on the left hand side of Fig. 4: note that as the complex is
symmetric we need only consider one subunit. On the basis of the homologies
between CRP and FNR, Spiro et al (18) recently proposed a similar scheme for
the DNA recognition helix of FNR which is shown on the right hand side of
Fig. 4. Spiro et al (18) suggested that Glu-209, which is located at
position 2 of the helix, interacts with the G:C base pair at position 7 of
the FNR binding sequence, and that this interaction is a common feature of
recognition by CRP and FNR. By analogy with CRP, Val-208 must therefore be
implicated in interactions at the T:A base pair at position 5 that is
responsible for specificity (compare the left and right hand diagrams in Fig.
4). However, as it is unlikely that an interaction with valine could provide
sufficient specificity, we propose that the neighbouring Ser-212 might also
be involved. It is possible that the mode of binding of the DNA recognition
helix in FNR is quite different from that of CRP. Clearly a study of the
effects of mutations at key positions is required to resolve how a single
base change at this position determines the specificity of protein binding.


## ACKNOWLEDGEMENTS

*To whom correspondence should be addressed


## REFERENCES
1.    Lambden, P. and Guest, J. (1976) J. Gen. Microbiol. 97, 145-160.
2.    Spiro, S. and Guest, J. (1987)  J. Gen. Microbiol. 133, 3279-3288.
3.    Jayaraman, P.-S., Cole, J. and Busby, S. (1989)  Nucl. Acids Res. 17,
      135-145.
4.    Nees, D., Stein, P. and Ludwig, R. (1988)  Nucl. Acids Res. 16,
      8839-8853.
5.    Eiglmeier, K., Honore, N., Iuchi, S., Lin, E. and Cole, S. (1989)  Mol.
      Microbiol. in press.
6.    Jayaraman, P.-S., Gaston, K., Cole, J. and Busby, S. (1988)  Mol.
      Microbiol. 3 , 147-151.
7.    Shaw, D., Rice, D. and Guest, J. (1983)  J. Mol. Biol. 166, 241-247.
8.    Spiro, S. and Guest, J. (1987)  Mol. Microbiol. 1, 53-58.

9.  Maniatis, T., Fritsch, E. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 1982.
10. Webster, C., Gaston, K. and Busby, S. (1988) Gene 68, 297-305.
11. Jayaraman, P.-S., Peakman, T., Busby, S., Quincey, R. and Cole, J. (1987) J. Mol. Biol. 196, 781-788.
12. Gaston, K., Chan, B., Kolb, A., Fox, J. and Busby, S. (1988) Biochem. J. 253, 809-818.
13. Berg, O. and von Hippel, P. (1988) J. Mol. Biol. 200, 709-723.
14. Kolb, A., Busby, S., Herbert, M., Kotlarz, D. and Buc, H. (1983) EMBO J. 2, 217-222.
15. Weber, I. and Steitz, T. (1984) Proc. Natl. Acad. Sci. USA. 81, 3973-3977.
16. Ebright, R., Cossart, P., Gicquel-Sanzey, B. and Beckwith, J. (1984) Proc. Natl. Acad. Sci. USA 81, 7274-7278.
17. Gent, M., Gronenborn, A., Davies, W. and Clore, G. (1987) Biochem. J. 242, 645-653.
18. Spiro, S., Todd, M., Artymiuk, P. and Guest, J. (1988) Biochem. Soc. Trans. 16, 755-756.