Molecular Microbiology (1996) **21**(2), 257–266

Definition of a consensus DNA-binding site for the *Escherichia coli* pleiotropic regulatory protein, FruR

Didier Nègre, Christelle Bonod-Bidaud, Christophe Geourjon, Gilbert Deléage, Alain J. Cozzone and Jean-Claude Cortay*

Institut de Biologie et Chimie des Protéines, Centre National de la Recherche Scientifique, 7 Passage du Vercors, 69367 Lyon, France.

Summary

The FruR regulator of *Escherichia coli* controls the initiation of transcription of several operons encoding a variety of proteins involved in carbon and energy metabolism. The sequence determinants of the FruR-binding site were analysed by using 6× His-tagged FruR and a series of double-stranded randomized oligonucleotides. FruR consensus binding sites were selected and characterized by several consecutive rounds of the polymerase chain reaction-assisted binding-site selection method (BSS) using nitrocellulose-immobilized DNA-binding protein. FruR was demonstrated to require, for binding, an 8 bp left half-site motif and a 3 bp conserved right half-site with the following sequence: 5′-GNNGAATC/GNT-3′. In this sequence, the left half-site AATC/ consensus tetranucleotide is a typical motif of the DNA-binding site of the regulators of the GalR–LacI family. On the other hand, the high degree of degeneracy found in the right half-site of this palindrome-like structure indicated that FruR, which is a tetramer in solution, interacts asymmetrically with the two half-sites of its operator. However, potentially FruR-target sites showing a high degree of symmetry were detected in 13 genes/operons. Among these, we have focused our interest on the *pfkA* gene, encoding phosphofructokinase-1, which is negatively regulated by FruR.

Introduction

The fructose repressor, FruR, is a 334-amino-acid transcriptional regulator that controls the expression of several major pathways involved in carbon and energy metabolism in enteric bacteria (Chin *et al*., 1987; 1989; Geerse *et al*., 1986; 1989). Thus, FruR activates the initiation of transcription from the *pps*, *icd*, and *ace* operons, whereas it represses the expression of the *fru* and *pts* operons. The *fruR* gene has been cloned and sequenced (Jahreis *et al*., 1991), and the corresponding protein has been overproduced and purified to homogeneity (Cortay *et al*., 1994). In solution, this protein exists as a tetramer of four identical subunits, each with a predicted size of 37 kDa. On the basis of sequence similarity with other bacterial regulatory proteins, FruR has been classified in the GalR–LacI family (Leclerc *et al*., 1990; Weickert and Adhya, 1992). Genetic and biochemical studies have shown that each member of this family contains two functional domains, namely an N-terminal domain that exhibits a structural helix-turn-helix (HTH) motif responsible for the binding of the regulator on its operator, and a C-terminal portion which displays inducer-binding properties and subunit interaction (Ogata and Gilbert, 1978; Jahreis and Lengeler, 1993; Scarabel *et al*., 1995). The high degree of conservation of the HTH DNA-binding motif suggests that all the proteins of this family bind to their operator sequence in a similar manner. Detailed description of the structure of the LacI family member–DNA complexes has been proposed both from nuclear magnetic resonance (NMR) studies of the N-terminal domain (or headpiece) of the *lac* repressor complexed with an 11 bp *lac* half-operator (Chuprina *et al*., 1993), and from X-ray crystallography studies of the PurR repressor bound to its high-affinity *purF* operator (Schumacher *et al*., 1994).

Previous experiments carried out on 11 natural FruR-binding sites have determined the *fru* consensus operator sequence to be 5′-TGAAtC/GNTtNa-3′ (Ramseier *et al*., 1993; 1995). This consensus indicated that FruR essentially interacts asymmetrically with the two half-sites of its operator DNA.

In this work, we have used the general approach of ‘cyclic amplification and selection of targets’ (CASTing) (Nørby *et al*., 1992; Wright and Funk, 1993) to further characterize the binding specificity of FruR. By this method, we have identified 20 related palindrome-like DNA sequences that bind FruR with high affinity. Their analysis yielded a consensus sequence that allowed us to detect a FruR-binding site in the promoter regulatory region of the *pfkA* gene.

Results and Discussion

Binding-site selection for FruR protein

The FruR protein, tagged with 6× His residues at its C-terminal end, was used to determine its optimal DNA

Received 2 February, 1996; revised 12 April, 1996; accepted 18 April, 1996. *For correspondence. Tel. 72 72 26 87; Fax 72 72 26 01.

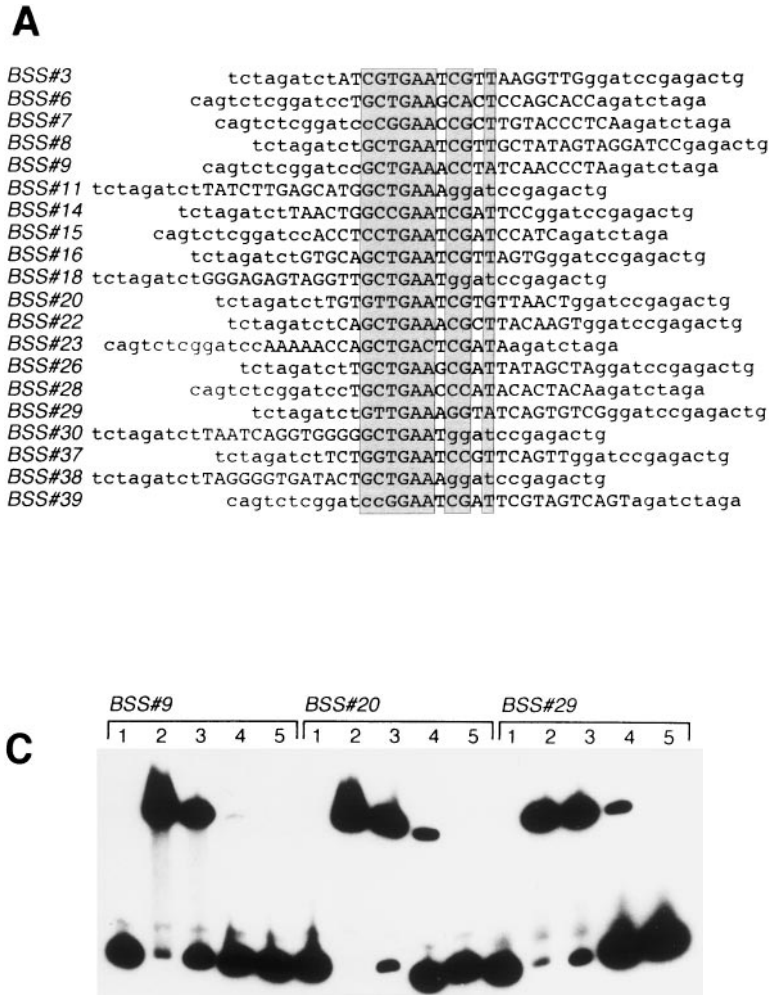

Fig. 1. Selection of FruR-binding sites from random-sequence oligonucleotides.
A. Alignment of FruR-binding sites selected in this study. The core sequence of the FruR-binding site is shaded. The residues present in the 20 bp randomized segment are shown in upper case, and the bases of the flanking sequences are shown in lower case.
B. Percentage of the most conserved base (y axis) at a particular position in the site (x axis). Bases present in more than 75% of the selected clones are shown in bold face. The numbering indicated below the consensus sequence refers to Schumacher *et al.* (1994). ◇: centre of symmetry.
C. Electrophoretic mobility shift assay (EMSA) for interaction of FruR with *BSS#9, 20* and *29*. Labelled DNAs for EMSA were amplified from recombinant pBluescript-II KS vectors by PCR with [α-^{32}P]-dATP and PT3 and PT7 primers. Lane 1, no protein; lanes 2–5, DNA with 20, 2, 0.2, and 0.02 nM FruR, respectively.

target sequence from a pool of oligonucleotides of random sequence. The oligonucleotides used for selection carried constant flanking sequences for polymerase chain reaction (PCR) amplification and cloning, whereas the central 20 nucleotides were randomized. The enrichment for FruR-binding sites was achieved by a combination of a filter-binding assay and PCR amplification of selected products.

The DNA recovered from the FruR–DNA complexes formed after five cycles of CASTing was digested by *Bam*HI and *Xba*I, and cloned between the corresponding sites in pBluescript-II KS vector. Several FruR-binding sites were sequenced and aligned (Fig. 1A). We checked whether each recovered CASTing double-stranded (ds)-oligonucleotide actually represented a FruR-binding site by testing its ability to bind to recombinant FruR in a band-shift assay. A series of corresponding results is presented in Fig. 1C. As the yield of complex formation obtained in these assays was about the same as that obtained with an equivalent amount of *ace* operator, which represents an authentic FruR-binding site (see Fig. 4A in Ramseier *et al*., 1993), it was concluded that the affinity for FruR of the selected sites was similar to that of the *ace* DNA.

Analysis of the selected fragments indicated that their sequence was very similar to that of the consensus FruR-binding site that we had determined previously from six natural *fru* operators (Ramseier *et al*., 1993). As shown in Fig. 1B, the alignment of the 20 selected sites revealed the FruR consensus sequence 5′-$N_1G_2C_3T_4G_5A_6A_7N_8C_9$/$G_{9'}N_{8'}T_{7'}$-3′, in which each position refers to the numbering used for the PurR operator site (Schumacher *et al*., 1994). In this consensus, each base was found to be present in at least 75% of the selected sequences. The DNA sequence thus recognized by FruR is an imperfect palindromic structure composed of an 8 bp left half-site and a 3 bp conserved right half-site. This particular organization suggests that, in general, FruR would interact asymmetrically with these two half-sites (Ramseier *et al*., 1993; Cortay *et al*., 1994). However, a more symmetrical arrangement of the nucleotide sequence within the operator was detected in some individual cases such as *BSS#14, BSS#37* and *BSS#39* (Fig. 1A). This observation favours the occurrence of interactions almost equally distributed between the two half-sites of the DNA-binding site.

In addition, it could be noted that most of the selected DNAs contained at least one AANC/tetranucleotide sequence in their left-half site, which is highly conserved in the DNA binding half-site of the bacterial regulators of the GalR–LacI family (Weickert and Chambliss, 1990; Schmid *et al*., 1991). In the FruR consensus sequence deduced from the analysis of the selected oligonucleotides, A is the preferred nucleotide at positions 6 and 7 with a percentage of occurrence of, respectively, 100% and 95% (Fig. 1B). Preference for A at the seventh position is obvious since only DNA *BSS#23* differs in this position (C versus A). To evaluate the importance of this difference, we prepared (by site-directed mutagenesis of the ds-oligonucleotide *BSS#23*), different mutant oligonucleotides that carried, respectively, $C_7 \rightarrow A$ and $C_7 \rightarrow G$ transversions, and a $C_7 \rightarrow T$ transition (Fig. 2). The DNA-binding properties of the FruR protein were then examined by a gel-shift experiment with these mutant oligonucleotides. Figure 2 shows that the mutant DNA corresponding to the $C_7 \rightarrow A$ transversion exhibited an increased affinity for FruR. In contrast, no significant difference was observed in the case of the $C_7 \rightarrow T$ transition relative to the wild-type *BSS#23* and, furthermore, the $C_7 \rightarrow G$ transversion appeared to totally impair the binding of FruR to the oligonucleotide. On the other hand, the T nucleotide at the symmetrical position 7′ in the degenerate right half-site of the FruR operator could be substituted by an A (*BSS#9* and *BSS#29*) or a G (*BSS#20*) without significantly affecting the interaction with FruR (Fig. 1C).

Position 6, in the *lac* and *pur* operators (Fig. 3), is known to interact with the second amino acid residue of the recognition helix of the protein. Thus, in the PurR–*purF* complex, Thr-16 establishes hydrogen bonds with the AT base pair at position 6, and in the LacI–*lacO* complex the corresponding Gln-18 residue, in co-operation with Ser-21, establishes specific contacts with GC base pair 6 of the operator (Lehming *et al*., 1987; Sartorius *et al*., 1989). Similarly, the Thr-15 residue of FruR (Fig. 3) would interact with the invariant AT base pair at position 6. A hydrophobic interaction involving the methyl group of thymine 7 and a hydrophobic pocket formed on the LacI protein by Tyr-17 and Ser-21 (the first and fifth residues of the recognition helix, respectively) (Caruthers, 1980; Chuprina *et al*., 1993) is known to be a crucial element in LacI repressor–operator recognition. In the PurR–*purF*

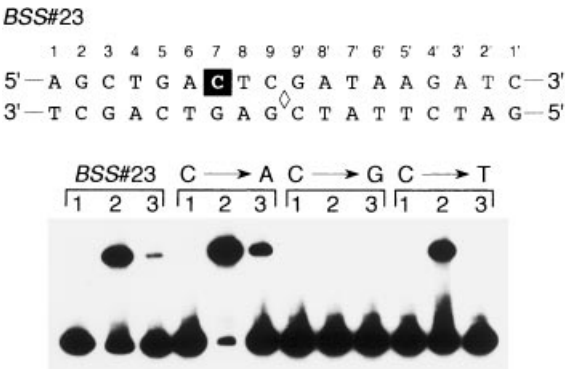

Fig. 2. EMSA for interaction of FruR with *BSS#23* variants. Incubation conditions were the same as in Fig. 1C. Lane 1, no protein; lanes 2 and 3, DNA with 2 and 0.2 nM FruR, respectively. The upper panel indicates the sequence of the FruR box detected in the *BSS#23* DNA. The C base, which was converted by site-directed mutagenesis into A, G or T, is shown in inverse printing. ◇: centre of symmetry.

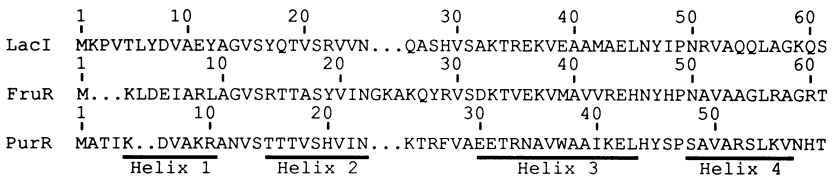

Fig. 3. Alignment of the N-terminal sequence of the FruR, LacI and PurR proteins and their respective consensus operator half-sites. The alignment was performed by using the ANTHEPROT package (Geourjon and Deléage, 1993).

	1	2	3	4	5	6	7	8	9
Lac	N	T	T	G	T	G	A	G	C ◊
FruR	N	G	C	T	G	A	A	T	C ◊
PurR	N	A	C	G	C	A	A	A	C ◊

complex, the Cγ methyl group of Ile-4, the first residue in helix 1, makes van der Waals contacts with the methyl group of Thy-7 (Schumacher *et al*., 1994). Similarly, the thymines at positions 6 and 7 in the natural FruR-binding sites display uracil interference sensitivity (Cortay *et al*., 1994). These contacts provide evidence for the conservation of an AT base pair at these two positions in the *fru* operator consensus. Further analysis showed that G is the preferred nucleotide at positions 2, 5, and 9′, with a percentage of occurrence of respectively 84%, 100%, and 79%. We had previously demonstrated that FruR binding interferes with the methylation of the N-7 position of G_2 and G_5 bases in the *fru* operator region of the *ace* operon (Cortay *et al*., 1994). These bases, which are more distant from the symmetry axis and seem to play a key role in the formation of the complex with FruR, would partly determine the specificity of the *fru* operator. Moreover, a putative role of the central dinucleotide C_9/$G_{9'}$ in the interaction with FruR can be proposed on the basis of the three-dimensional structure data obtained for the PurR–*purF* complex (Schumacher *et al*., 1994). In this complex, the side-chain of Leu-54 in the 'hinge-helix' of the protein (Leu-56 being the equivalent residue in FruR) is known to intercalate into the minor-groove central dinucleotide C_9/$G_{9'}$ of the *purF* operator (Fig. 3). Another important contact was described between the 'hinge-helix' residue Lys-55 in PurR, which corresponds to Arg-57 in FruR, and the TA base pair at position 8 in its operator (Schumacher *et al*., 1994).

Detection of a FruR-binding site in the 5′-end of the pfkA *gene*

A computer-based homology-search database on GenBank (Release 83), using the FASTA program (Pearson and Lipman, 1988) with the query sequence 5′-$T_1G_2C_3T_4G_5A_6A_7T_8C_9$/$G_{9'}A_{8'}T_{7'}T_{6'}C_{5'}A_{4'}A_{3'}$-3′, was performed to detect potential, unidentified FruR-binding sites almost identical to the consensus sequence described above. This 'improved consensus site' has helped us to find 13 new FruR DNA targets in genes/operons which could be under the control of FruR protein

Table 1. Computer search for FruR-binding sites in the *E. coli* genome.

Genomic region	Sequences (5′ to 3′)[a]	Accession no.[b]
Consensus sequence	1 TGCTGAATC/GATTCAA 16	
bent-5 DNA	39 TGCTGAATt/GATTCAg 54	X05960
betT, betI, betB, betA genes	879 TGCTGAATg/GATTCAg 864	X52905
Min 81.5 to 84.5	53 242 TGCTtAATC/GATTgAA 53 227	L10328
Conjugation proteins (*tra* genes)	3743 TGCTGAATC/GATaCgg 3728	U09868
Min 63 to 64	4055 TGCTGAATC/GAaTCtt 4040	U29581
manX and *manY* genes	147 atCTGAATC/GATTCgA 162	X17443
Min 76.0 to 81.5	199 462 aGCTGAATC/GATaaAA 199 447	U00039
Min 65 to 68	53 196 TGCTGAATC/GgcaCtt 53 211	U28377
artP, artI, artQ, artM and *artJ* genes	1042 gcaTcAATg/cATTCAA 1027	X86160
HK31IM gene	1066 gctTGAATC/GATTagA 1050	X82231
Ornithine and arginine decarboxylases antizyme genes	2048 aaagaAATC/GATTCAA 2063	L13078
DNA cytosine methylase (dcm) and patch-repair protein genes	572 ccgccAATC/GATTCAA 557	M32307
orf300 and *orf732* genes	914 aaCTGAATC/GATTtgt 924	U23723

a. Numbers refer to nucleotide position provided under the 'Accession no.' column.
b. Database releases: GenBank Version 93 and EMBL Version 45.

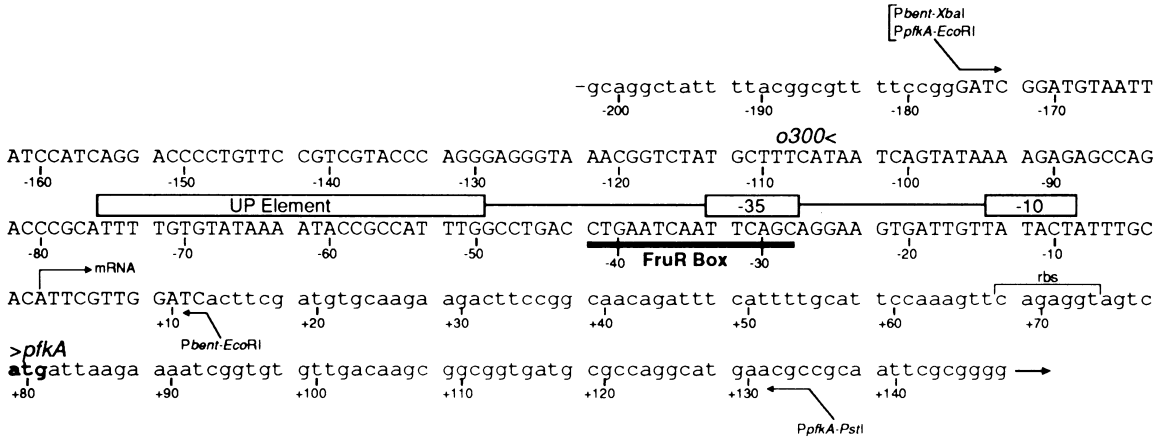

Fig. 4. Nucleotide sequence of the promoter region of the *pfkA* gene. The 3′-end of the *o300* gene (Plunkett *et al*., 1993) and the 5′-end of the *pfkA* gene are indicated (GenBank Accession No. L19201). *bent5* sequence (Mizuno, 1987) is in upper case. Possible transcription and translation control sites are indicated: −10 and −35 refer to the promoter region, and 'rbs' indicates the ribosome-binding site. The FruR-binding site is shown as FruR box.

(Table 1). Of particular interest was the detection of a potential FruR-binding site in the *bent5* fragment whose position along the *Escherichia coli* genome corresponds to the nucleotide sequence near the 5′-end of the *pfkA* gene which encodes the glycolytic enzyme phosphofructokinase-1 (Fig. 4A) (Hellinga and Evans, 1985; Plunkett *et al*., 1993). The *bent5* fragment was originally isolated and characterized from a *Sau*3A digest of the chromosomal DNA of *E. coli* K-12 (Mizuno, 1987). The results of a circular permutation test have shown that this DNA presents an intrinsic curvature due to the periodic phasing of stretches of A_n and T_n ($n \geqslant 3$). Although previous genetic experiments provided evidence that FruR negatively regulates *pfkA* expression, nothing is known to date about the mode of action of this regulator (Chin *et al*., 1987; Saier and Chin, 1990), and it is not clear if the *bent5* DNA contains the promoter/operator region of the *pfkA* gene.

The binding of FruR to the 5′-flanking region of *pfkA* was measured by a gel-retardation assay. Figure 5A (lanes 1–5) reveals that in the presence of a 199 bp end-labelled *bent5* DNA fragment (4 pM) and a decreasing amount of FruR (see the *Experimental procedures*), a single and stable DNA–protein complex could be detected. The apparent dissociation constant (K_d) for FruR binding to its *bent5* operator was 3.5 nM at pH 7.9 and 25°C. The FruR–DNA complex was less abundant when the incubation was carried out in the presence of the specific inducer fructose-1-phosphate (lanes 6–8). These observations therefore indicated that the FruR repressor is capable of interacting in a specific manner with one operator sequence in the *bent5* DNA fragment.

DNase I footprinting experiments

In order to characterize the specific region of *bent5* DNA that interacts with FruR, the nucleotide sequence protected from DNase I digestion after binding of the protein was analysed. A 5′-end-labelled *bent5* DNA fragment was incubated with FruR at 25°C, then treated with DNase I. In the absence of FruR, DNase I cleavage produced a regular pattern of bands (Fig. 5B). In contrast, upon addition of the FruR protein, a protected region of 24 bp was detected, including the 12 bp palindromic structure (5′-TGAATT/GATTCA-3′) present in this particular region of DNA.

G-methylation-interference experiments

The *fru* operator region in the *bent5* DNA was characterized in more detail by determining the effect of the methylation of guanine residues on the binding of FruR (Ogata and Gilbert, 1978; Siebenlist and Gilbert, 1980; Siebenlist *et al*., 1980). As expected from our previous results (Cortay *et al*., 1994), methylation of guanines G_2 and G_5 in the top strand left half-site operator was shown to cause strong interference with the binding of FruR (Fig. 6). Interestingly, similar interference occurred for the symmetrically located guanine $G_{5'}$ in the bottom strand right half-site operator (Fig. 6). These three guanine residues were all present in the region of DNA protected by FruR against DNase I digestion. Furthermore, analysis of the footprinting data showed that FruR could protect both arms of this palindrome to almost the same extent. Therefore, as already reported in the case of the *O1 fruB* mutant operator (Cortay *et al*., 1994), the interaction between FruR and its operator seems to increase when the recognition DNA site exhibits a higher degree of symmetry. This could mean that a fully palindromic operator would be optimal for FruR binding. However, the methylation experiments described here raise the

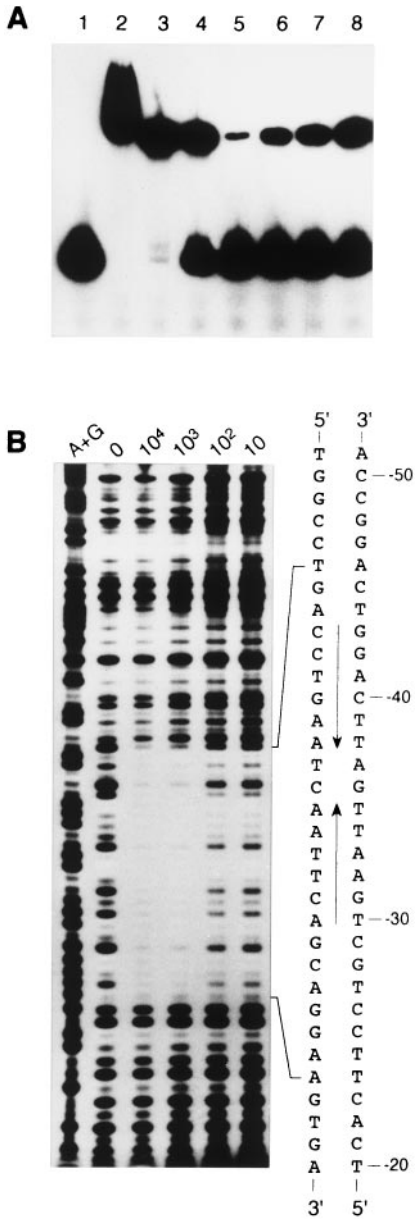

Fig. 5. Characterization of the FruR–*bent5* complex.
A. Relative strength of binding of FruR to the *bent5* fragment. The 199 bp [^{32}P]-*bent5* DNA (4 pM) was incubated at 25°C with a decreasing concentration of FruR protein, from 200 nM to 0.2 nM, and analysed by EMSA. Lane 1, no protein; lanes 2–5, DNA with 200, 20, 2 and 0.2 nM FruR, respectively; lanes 6–8, incubation was the same as in lane 4 but with 1, 0.1 and 0.01 mM fructose-1-phosphate, respectively.
B. FruR-mediated protection of the *bent5* region against digestion by DNAse I. In each assay, target DNA (10^5 c.p.m.) was incubated in the absence or in the presence of FruR (the number above each lane indicates the concentration of FruR protein in nM). Samples were separated on a denaturing 6% polyacrylamide gel. The vertical arrows beside the nucleotide sequence indicate a palindromic structure.

possibility that this may not actually be the case. In fact, FruR binds much more tightly to one side than the other on both symmetrical and asymmetrical operators.

Promoter activity of bent5 *DNA*

To evaluate the promoter activity of the *bent5* DNA, we constructed a gene fusion between this fragment and the *lacZ* gene in the promoter-detecting plasmid pNM481 (Minton, 1984). A 310 bp DNA fragment was amplified by PCR with P*pfkA-Eco*RI and P*pfkA-Pst*I primers (Fig. 4 and Table 2) and cloned into the corresponding sites of plasmid pNM481 so as to obtain translational phasing with the 8th codon of β-galactosidase. The differential rate of β-galactosidase synthesis was then measured in *E. coli* JM105 cells containing either the *bent5–lacZ* fusion or the control vector pNM481 alone. When the *bent5–lacZ* fusion was introduced into *E. coli* strain JM105, a 238-fold increase in the basal level of β-galactosidase activity was measured, indicating that the *bent5* fragment exhibited a relatively high promoter activity.

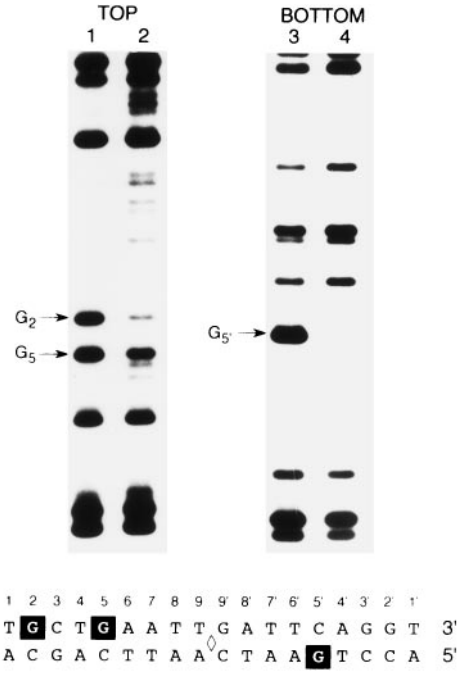

Fig. 6. Identification of FruR–DNA contacts by methylation-interference experiments. The 199 bp *bent5* fragment was ^{32}P-labelled at either the *Eco*RI site referred to as 'BOTTOM' or at the *Not*I site, referred to as 'TOP' (see the *Experimental procedures*). Each binding mixture contained the FruR protein (approximately 2 nM active protein) and end-labelled *bent5* DNA partially methylated by DMS. The free DNA (lanes 1 and 3) and bound DNA (lanes 2 and 4) were separated by EMSA, cleaved at methylated guanine residues with piperidine, and analysed on a 6% sequencing gel. The sequence designation is given at the bottom. The guanine residues interfering in FruR binding are shown in inverse printing. ◇: centre of symmetry.

Table 2. Designation and nucleotide sequence of primers.

Amplified DNA	Primer used	Oligodeoxynucleotide sequence (5′ to 3′)	Length of primer (nt)	Length of amplified DNA fragment (bp)
ds-BSS-1	BSS-1	CGGGCTGAGATCAGTCTAGATCT(N)$_{26}$GGATCCGAG ACTGAGCGTCGTCA[a]	65	65
	BSS-2	CGGGCTGAGATCAGTCTAGATCT	23	
	BSS-3	GACGACGCTCAGTCTCGGATCC	22	
bent-5	P*bent-Eco*RI	TATGAATTCGATCCAACGAATGTGCAAATAG[b]	31	204
	P*bent-Xba*I	TATTCTAGATCGGATGTAATTATCCATCAGG[c]	31	
bent–pfkA	P*bfkA-Eco*RI	TATGAATTCGATCGGATGTAATTATCCATCAGG[b]	33	327
	P*bfkA-Pst*I	TATCTGCAGCGTTCATGCCTGGCGCATCACCG[d]	32	
BSS#23	PT3	ATTAACCCTCACTAAAG	17	181
	PT7	AATACGACTCACTATAG	17	

a. N is A, C, G, or T (equimolar); *Xba*I site (TCTAGA) and *Bam*HI site (GGATCC) are underlined.
b. Three nucleotides (TAT) and an *Eco*RI site (GAATTC) were added at the 5′ end of the known DNA sequence.
c. Three nucleotides (TAT) and an *Xba*I site (TCTAGA) were added at the 5′ end of the known DNA sequence.
d. Three nucleotides (TAT) and a *Pst*I site (CTGCAG) were added at the 5′ end of the known DNA sequence.

The precise contacts between the RNA polymerase (RNAP) and the promoter of the *pfkA* gene were probed by the base-removal approach (Brunelle and Schleif, 1987). In these experiments, individual bases within the *bent5* DNA were removed by treatment with formic acid or hydrazine. This partially modified DNA was incubated with σ^{70}-saturated RNAP; free DNA and the complex were separated by electrophoresis on a non-denaturing gel at high ionic strength, then cleaved at the positions of missing bases, and separated on a sequencing gel to reveal the positions either irrelevant or, instead, crucial to binding. As shown in Fig. 7, the removal of any individual base along the promoter was insufficient to prevent RNAP binding. On the contrary, polymerase binding was enhanced when bases in the regions from −5 to −20, and from −50 to −75, on the bottom strand were removed. The first area could constitute the 'melting domain' of the transcriptional complex where the DNA structure is known to be determinant in mediating interactions with the RNAP (Werel *et al*., 1991). Hellinga and Evans (1985) had detected a putative mRNA start point of transcription in the 100 bp nucleotide sequence upstream from the *pfkA* ATG start codon (see Fig. 4), with a −10 region TATACT differing by only one base from the −10 promoter consensus sequence, TATAAT. In this regard, our results indicated that FruR binds to a single site, consisting of a 14 bp sequence, which would overlap the putative −35 region (TTCAGC) of the *pfkA* gene. As the operator and the promoter sequences are, in this case, essentially coincident, one may consider that the negative regulator FruR would inhibit transcription by steric hindrance, i.e. by preventing the access of RNAP to its binding site within the promoter of the *pfkA* gene. Of particular interest is the detection of the second region centred around the base at position −60 (Figs 4 and 7). This is the same area, previously characterized by DNase I footprinting, as the AT-rich α-subunit binding site on the *rrnB* P1 promoter (Ross *et al*., 1993). In this latter case, specific binding of the RNAP to this particular DNA, namely the UP module, is accompanied by a 30-fold increase in transcription from the promoter. We propose that base removal in the corresponding area of *pfkA* enhances structural flexibility of the DNA and thus constitutes the most critical factor in stimulating promoter binding by RNAP. In this regard, the UP-like element that we have characterized in the *pfkA* promoter region fully correlates with the major bend locus previously detected (Mizuno, 1987). Thus, the ability of this particular curved DNA upstream of the −35 region of the *pfkA* gene to affect the interaction of the *E. coli* RNAP with its promoter would facilitate the subsequent steps of the open-complex formation.

Experimental procedures

FruR purification

The FruR protein fused with 6× His residues at its C-terminal end was expressed in BL21(DE3)[pJCD2] cells after induction for 3 h by IPTG, and purified as already described (Cortay *et al*., 1994).

Binding-site selection method

The method for binding-site selection was adapted from Nørby *et al*. (1992). One hundred picomoles of the 65-mer *BSS-1* oligonucleotide were mixed with an equimolecular amount of *BSS-3* primer (Table 2) in 10 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 50 mM NaCl, and 1 mM DTE, heated for 5 min at 65°C, then cooled slowly to room temperature. A mixture of dNTPs (50 μM each) and 10 units of Klenow enzyme were added, and the reaction was incubated at 16°C for 1.5 h. The incubation mixture was then applied to a 4.6 × 100 mm Gen-Pak Fax anion-exchange column (Waters) equilibrated with 25 mM Tris-HCl, pH 8.0, 1 mM EDTA. After

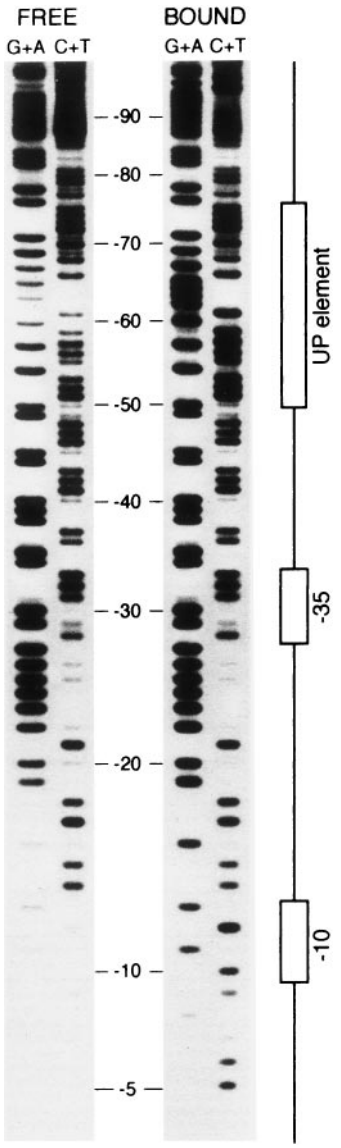

Fig. 7. Identification of RNAP–DNA contacts by base-removal footprinting. The binding mixture contained RNAP (approximately 8.2 nM active protein) and the 199 bp *bent5* fragment ^{32}P-labelled at the *Eco*RI site (see the legend to Fig. 6). The DNA was then treated with either formic acid to remove G and A or hydrazine to remove C and T, before electrophoresis and autoradiography. 'FREE', gel lanes with the DNA that had dissociated or was free of complexes; 'BOUND', gel lanes with the modified DNA isolated from the RNAP–DNA complexes.

extensive washing, the ds-*BSS-1* oligonucleotide was eluted with the same buffer containing 0.5 M NaCl and concentrated in a Microcon-3 filtration device (Amicon). Ten picomoles of the ds-random oligonucleotide mixture were resuspended in 200 μl of binding buffer (25 mM HEPES–KOH, pH 7.9, 40 mM KCl, 3 mM $MgCl_2$, 1 mM DTT), and exposed to the FruR protein (200 ng) immobilized on nitrocellulose BA 83 filters (Schleicher and Schuell) for 2 h at 4°C. After four successive washes with 200 μl of binding buffer, the bound ds-oligonucleotides were dissociated from the DNA-binding protein by washing with 200 μl of 0.5 M KCl. Ten microlitres of these filter-eluted oligonucleotides were mixed with PCR buffer without KCl (10 mM Tris-HCl, pH 8.3, 1.5 mM $MgCl_2$) containing 10 μM each of *BBS-2* and *BSS-3* primers (Table 2) and 2.5 units of *Taq* DNA polymerase in a final volume of 100 μl. Twenty cycles of amplification were performed, each of them including 95°C for 50 s, 64°C for 50 s, and 72°C for 1 min. Five microlitres of amplified ds-oligonucleotides were used without further purification for the next round of the filter-binding procedure.

After three rounds of filter-binding and amplification, the recovered DNA was amplified and labelled in a 20 μl PCR reaction medium containing 2 μM each of *BSS-2* and *BSS-3* primers, 10 μCi of [α-^{32}P]-dCTP (3000 Ci mmol^{-1}), 20 μM unlabelled dCTP, 50 μM each dATP, dTTP, dGTP, 1.5 mM $MgCl_2$, 10 mM Tris-HCl, pH 8.3, and 1 unit of *Taq* DNA polymerase. Amplification was performed successively at 94°C for 1 min, at 62°C for 1 min, at 72°C for 1 min, and a total of 15 cycles was performed. A small amount of the amplified product was checked for binding with the FruR protein in a gel-shift assay (see below).

After five rounds of iterative selection and amplification, ds-oligonucleotides were digested with *Xba*I and *Bam*HI, and cloned between the corresponding sites in the pBluescript-II KS vector (GenBank Accession No. X52327, Stratagene). Recombinant plasmids were sequenced by the chain termination method using the Sequenase Version 2.0 kit (Amersham Int. plc).

Gel-shift assays

Gel-retardation assays were performed essentially as already described (Garner and Revzin, 1981). A typical assay mixture contained in 25 μl: 12 mM HEPES–NaOH, pH 7.9, 4 mM Tris-HCl, pH 7.9, 95 mM KCl, 1 mM EDTA, 1 mM DTT, 9% (v/v) glycerol, 0.2% (v/v) Nonidet P-40, 0.5 μg of poly(dI-dC).poly(dI-dC) as bulk carrier DNA, 2.5 μg of BSA, radioactive DNA probe (4 pM), and FruR protein. After 10 min of incubation at 25°C, 10 μl of this mixture was loaded onto a 4% (w/v) polyacrylamide gel at high ionic strength and electrophoresed for 1 h at 30 V cm^{-1}. Radioactive compounds were detected by autoradiography after overnight exposure to Kodak Biomax MR film at −70°C.

Preparation of the bent5 *fragment*

bent5 DNA was obtained by PCR amplification from *E. coli* K-12 chromosomal DNA by using two primers (Table 2) containing, respectively, an *Eco*RI site and a *Xba*I site at their 5′-end, and was cloned into plasmid pBluescript-II KS digested by the same enzymes. The *Eco*RI–*Not*I digest containing *bent5* was end-labelled with the Klenow fragment of DNA polymerase I, either at the *Eco*RI site by [α-^{32}P]-dATP (bottom strand) or at the *Not*I site by [α-^{32}P]-dGTP (top strand). These labelled DNAs were further purified by using a MicroSpin S-400 column (Pharmacia Biotech) before incubation and gel-retardation analysis.

DNase I footprinting

The footprinting method with bovine pancreas deoxyribonuclease I was used (Galas and Schmitz, 1978). After

10 min of preincubation at 25°C of the FruR protein and the end-labelled *bent5* DNA probe in binding buffer, the mixture was treated for 30 s at the same temperature with 1.6 units of DNase I (Stratagene) in the presence of 2 mM $CaCl_2$ and 6 mM $MgCl_2$. The reaction was stopped by the addition of 10 mM EDTA. Nucleic acids were precipitated with ethanol, dissolved in formamide dye mix, and heated at 80°C for 2 min prior to electrophoresis through a 7 M urea/6% (w/v) acrylamide sequencing gel.

Interference experiments

Pre-modification of the end-labelled *bent5* DNA fragment was achieved by using two different approaches: (i) base methylation with dimethyl sulphate (DMS) as described by Siebenlist and Gilbert (1980), and (ii) base removal with formic acid as a depurinating reagent and hydrazine as a depyrimidinating reagent (Brunelle and Schleif, 1987). Methylation-interference experiments were performed by incubating modified DNA (approximately 10^5 c.p.m.) with FruR (2 nM) while the base-removal assay was carried out in the presence of 8.2 nM *E. coli* RNAP holoenzyme (Epicentre Technologies Corp.). After preparative electrophoresis in the gel-retardation conditions described above, bands corresponding to free and complexed DNA were visualized by autoradiography on the wet gel after overnight exposure to Kodak Biomax MR film at 4°C. Labelled DNA was cut out of the gel, eluted for subsequent piperidine cleavage, and analysed in a sequencing gel.

Oligonucleotide-directed site-specific mutagenesis of BSS#23

The mutagenic oligodeoxynucleotides were 26 bases in length, with the targeted mismatched base in the centre. Mutagenesis was achieved by the procedure of Kammann *et al*. (1989). In the first PCR round, the sequence between the primer PT7 and the mutagenic primer was amplified. In a second round, this amplified fragment was used as a megaprimer in combination with primer PT3 (Table 2). The mutant sequences thus obtained were checked by dideoxy sequencing.

Assay of β-galactosidase

Beta-galactosidase activity in *E. coli* strain JM105 (*endA1 thi rpsL sbcB15 hsdR4* Δ(*lac–proAB*), [F′ *traD36 proAB lacI*^{q}*Z*Δ*M15*] (Yanisch-Perron *et al*., 1985), carrying the bent5–lacZ fusion was measured by the absorbance change at 420 nm associated with ONPG hydrolysis. This assay was similar to that previously described (Miller, 1972), as modified by Platko *et al*. (1990) and Ernsting *et al*. (1993). The activity of β-galactosidase was measured in arbitrary units defined previously (Miller, 1972): $1000 \times \Delta A_{420}$ $min^{-1} ml^{-1}$.

Computer-aided calculation of consensus pattern

To calculate a consensus pattern in a fully automatic manner, the binding-site selection program (BSSP) (requests to geourjon@ibcp.fr) was written. This program allowed the alignment of the different clones and the calculation of a quality index by using the Needleman and Wunsch algorithm (1970). The alignment quality index was used to select the nucleotides that shared the highest similarity score with the whole set of aligned clones. The comparison was performed by using a unity matrix and a high gap penalty value so as to minimize the frequency of gaps.

Acknowledgements

We thank L. Panaye for reading the manuscript, and C. Van Herrewege and A. Bosch for help with the preparation of photographs. This work was supported by grants from the CNRS (UPR 412) and the University of Lyon.

References

Brunelle, A., and Schleif, R.F. (1987) 'Missing contact' probing of DNA–protein interactions. *Proc Natl Acad Sci USA* **84:** 6673–6676.

Caruthers, M.H. (1980) Deciphering the protein–DNA recognition code. *Acc Chem Res* **13:** 155–160.

Chin, A.M., Feucht, B.U., and Saier, Jr, M.H. (1987) Evidence for regulation of gluconeogenesis by the fructose phosphotransferase system in *Salmonella typhimurium*. *J Bacteriol* **169:** 897–899.

Chin, A.M., Feldheim, D.A., and Saier, Jr, M.H. (1989) Altered transcriptional patterns affecting several metabolic pathways in strains of *Salmonella typhimurium* which overexpress the fructose regulon. *J Bacteriol* **171:** 2424–2434.

Chuprina, V.P., Rullmann, J.A.C., Lamerichs, R.M.J.N., van Boom, J.H., Boelens, R., and Kaptein, R. (1993) Structure of the complex of *lacI* repressor headpiece and an 11 base-pair half-operator determined by nuclear magnetic resonance spectroscopy and restrained molecular dynamics. *J Mol Biol* **234:** 446–462.

Cortay, J.C., Nègre, D., Scarabel, M., Ramseier, T.M., Vartak, N.B., Reizer, J., and Cozzone, A.J. (1994) *In vitro* asymmetric binding of the pleiotropic regulatory protein, FruR, to the *ace* operator controlling glyoxylate shunt enzyme synthesis. *J Biol Chem* **269:** 14885–14891.

Ernsting, B.R., Blumenthal, R.M., and Matthews, R.G. (1993) Regulation of the *gltBDF* operon of *Escherichia coli*: how is a leucine-insensitive operon regulated by the leucine-responsive regulatory protein? *J Bacteriol* **175:** 7160–7169.

Galas, D.J., and Schmitz, A. (1978) DNase footprinting: a simple method for the detection of protein–DNA binding specificity. *Nucl Acids Res* **5:** 3157–3170.

Garner, M.M., and Revzin, A. (1981) A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the *Escherichia coli* lactose operon regulatory system. *Nucl Acids Res* **9:** 3047–3060.

Geerse, R.H., Ruig, C.R., Schuitema, A.R.J., and Postma, P.W. (1986) Relationship between pseudo-HPr and the PEP:fructose phosphotransferase system in *Salmonella typhimurium* and *Escherichia coli*. *Mol Gen Genet* **203:** 435–444.

Geerse, R.H., Izzo, F., and Postma, P.W. (1989) The PEP:fructose phosphotransferase system in *Salmonella typhimurium*: FPr combines enzyme III^{Fru} and pseudo-HPr activities. *Mol Gen Genet* **216:** 517–525.

Geourjon, C., and Deléage, G. (1993) Interactive and graphic coupling between multiple alignments, secondary structures prediction and motif/pattern scanning into proteins sequences. *Comput Applic Biosci* **9:** 87–91.

Hellinga, H.W., and Evans, P.R. (1985) Nucleotide sequence and high-level expression of the major *Escherichia coli* phosphofructokinase. *Eur J Biochem* **149:** 363–373.

Jahreis, K., and Lengeler, J.W. (1993) Molecular analysis of two ScrR repressors and of a ScrR–FruR hybrid repressor for sucrose and D-fructose specific regulons from enteric bacteria. *Mol Microbiol* **9:** 195–209.

Jahreis, K., Postma, P.W., and Lengeler, J.W. (1991) Nucleotide sequence of the *ilvH–fruR* gene region of *Escherichia coli* K12 and *Salmonella typhimurium* LT2. *Mol Gen Genet* **226:** 332–336.

Kammann, M., Laufs, J., Schell, J., and Gronenborn, B. (1989) Rapid insertional mutagenesis of DNA by polymerase chain reaction (PCR). *Nucl Acids Res* **17:** 5404–5404.

Leclerc, G., Noel, G., and Drapeau, G.R. (1990) Molecular cloning, nucleotide sequence, and expression of *shl*, a new gene in the 2-minute region of the genetic map of *Escherichia coli. J Bacteriol* **172:** 4696–4700.

Lehming, N., Sartorius, J., Niemöller, M., Genenger, G., von Wilcken-Bergmann, B., and Müller-Hill, B. (1987) The interaction of the recognition helix of *lac* repressor with *lac*. *EMBO J* **6:** 3145–3153.

Miller, J.H. (1972) *Experiments in Molecular Genetics.* Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Minton, N.P. (1984) Improved plasmid vectors for the isolation of translational *lac* gene fusions. *Gene* **31:** 269–273.

Mizuno, T. (1987) Random cloning of bent DNA segments from *Escherichia coli* chromosome and primary characterization of their structures. *Nucl Acids Res* **15:** 6827–6840.

Needleman, S.B., and Wunsch, C.D. (1970) A general method applicable to the search of similarities in the amino acid sequences of two proteins. *J Mol Biol* **48:** 443–453.

Nørby, P.L., Pallisgaard, N., Pedersen, F.S., and Jørgensen, P. (1992) Determination of recognition-sequences for DNA-binding proteins by a polymerase chain reaction assisted binding site selection method (BSS) using nitrocellulose immobilized DNA binding protein. *Nucl Acids Res* **20:** 6317–6321.

Ogata, R.T., and Gilbert, W. (1978) An amino-terminal fragment of the *lac* repressor binds specifically to *lac* operator. *Proc Natl Acad Sci USA* **75:** 5851–5854.

Pearson, W.R., and Lipman, D.J. (1988) Improved tools for biological sequence comparison. *Proc Natl Acad Sci USA* **85:** 2444–2448.

Platko, J.V., Willins, D.A., and Calvo, J.M. (1990) The *ilvIH* operon of *Escherichia coli* is positively regulated. *J Bacteriol* **172:** 4563–4570.

Plunkett, III, G., Burland, V., Daniels, D.L., and Blattner, F.R. (1993) Analysis of the *Escherichia coli* genome. III. DNA sequence of the region from 87.2 to 89.2 minutes. *Nucl Acids Res* **21:** 3391–3398.

Ramseier, T.M., Nègre, D., Cortay, J.C., Scarabel, M., Cozzone, A.J., and Saier, Jr. M.H. (1993) *In vitro* binding of the pleiotropic transcriptional regulatory protein, FruR, to the *fru, pps, ace, pts* and *icd* operons of *Escherichia coli* and *Salmonella typhimurium*. *J Mol Biol* **234:** 28–44.

Ramseier, T.M., Bledig, S., Michotey, V., Feghali, R., and Saier, Jr, M.H. (1995) The global regulatory protein FruR modulates the direction of carbon flow in *Escherichia coli*. *Mol Microbiol* **16:** 1157–1169.

Ross, W., Gosink, K.K., Salomon, J., Igarashi, K., Zou, C., Ishihama, A., Severinov, K., and Gourse, R.L. (1993) A third recognition element in bacterial promoters: DNA binding by the α subunit of RNA polymerase. *Science* **262:** 1407–1413.

Saier, Jr, M.H., and Chin, M. (1990) Energetics of the bacterial phosphotranferase system in sugar transport and the regulation of carbon metabolism. In *Bacterial Energetics.* Krulwich, T.A. (ed.). New York: Academic Press, pp. 273–299.

Sartorius, J., Lehming, N., Kisters, B., von Wilcken-Bergmann, B., and Müller-Hill, B. (1989) *lac* repressor mutants with double or triple exchanges in the recognition helix bind specifically to *lac* operator variants with multiple exchanges. *EMBO J* **8:** 1265–1270.

Scarabel, M., Penin, F., Bonod-Bidaud, C., Nègre, D., Cozzone, A.J., and Cortay, J.C. (1995) Overpurification, purification and structural characterization of the functional N-terminal DNA-binding domain of the *fru* repressor from *Escherichia coli* K-12. *Gene* **153:** 9–15.

Schmid, K., Ebner, R., Jahreis, K., Lengeler, J.W., and Titgemeyer, F. (1991) A sugar-specific porin, ScrY, is involved in sucrose uptake in enteric bacteria. *Mol Microbiol* **5:** 941–950.

Schumacher, M.A., Choi, K.Y., Zalkin, H., and Brennan, R.G. (1994) Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by α helices. *Science* **266:** 763–770.

Siebenlist, U., and Gilbert, W. (1980) Contacts between *Escherichia coli* RNA polymerase and an early promoter of phage T7. *Proc Natl Acad Sci USA* **77:** 122–126.

Siebenlist, U., Simpson, R.B., and Gilbert, W. (1980) *E. coli* RNA polymerase interacts homologously with two different promoters. *Cell* **20:** 269–281.

Weickert, M.J., and Adhya, S. (1992) A family of bacterial regulators homologous to Gal and Lac repressors. *J Biol Chem* **267:** 15869–15874.

Weickert, M.J., and Chambliss, G.H. (1990) Site-directed mutagenesis of a catabolite repression operator sequence in *Bacillus subtilis*. *Proc Natl Acad Sci USA* **87:** 6238–6242.

Werel, W., Schickor, P., and Heumann, H. (1991) Flexibility of the DNA enhances promoter affinity of *Echerichia coli* RNA polymerase. *EMBO J* **10:** 2589–2594.

Wright, W.E., and Funk, W.D. (1993) CASTing for multicomponent DNA-binding complexes. *Trends Biochem S*ci **18:** 77–80.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33:** 103–119.