Two distinct classes of FixJ binding sites defined by in vitro selection

Lionel Ferrières, Daniel Kahn*

Laboratoire de Biologie Moléculaire des Relations Plantes-Microorganismes, UMR 215 INRA-CNRS, Chemin de Borde-Rouge, P.O. Box 27, 31326 Castanet-Tolosan Cedex, France

Received 11 February 2002; accepted 19 February 2002

First published online 28 March 2002

Edited by Gianni Cesareni

Abstract The phosphorylated FixJ transcriptional activator is key to nitrogen fixation in *Sinorhizobium meliloti*, switching both the *nifA* and *fixK* genes on. Previously no consensus picture emerged concerning the nature of FixJ binding sites. Here we used in vitro DNA selection in order to systematically characterise FixJ binding sequences. This led to the definition of two classes of sites. Class I sites share the CTAAG-TAGTTTCCC sequence found in the *fixK* promoter, whereas class II sites are defined by a GTAMGTAG consensus octamer. Our results indicate that FixJ∼P binds DNA following two distinct binding modes. © 2002 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.

Key words: FixJ; Transcriptional regulator; Binding site; Two-component system; Nitrogen fixation

1. Introduction

The two-component FixLJ system of *Sinorhizobium meliloti* activates expression of nitrogen fixation genes in response to the low oxygen concentration prevailing in alfalfa root nodules. Under these conditions the FixL sensor protein autophosphorylates and transfers its phosphate to the FixJ response regulator. Once phosphorylated, FixJ activates transcription of the *nifA* and *fixK* genes, thereby switching the nitrogen fixation apparatus on [1,2].

FixJ comprises two domains with distinct functions. The C-terminal output domain displays a DNA binding HTH motif and is sufficient for activation of the *nifA* promoter both in vivo [3] and in vitro [4]. In the native FixJ protein the action of the output domain is inhibited by the N-terminal receiver domain. Under micro-aerobic conditions the FixJ receiver domain is phosphorylated by the FixL kinase, which both liberates the output domain and triggers receiver domain dimerisation. This results in a FixJ∼P dimer which is the active form of the protein with high affinity for target DNA [5].

Up to now only two genes were known to be directly regulated by FixJ. However their respective promoters exhibit strong dissimilarities concerning FixJ recognition. For instance a fully functional *fixK* promoter requires at least 63 nucleotides of upstream sequence [6] whereas 54 nucleotides only are required at the *nifA* promoter [7]. As the corresponding functional regions do not share any conspicuous sequence similarity, the nature of FixJ binding sites has remained elusive [2]. Using an in vitro selection (SELEX) approach on randomised oligonucleotides, here we identify two classes of sequences which are recognised by FixJ, allowing the definition of two consensus binding sites. Gel retardation and DNase I foot-printing studies indicate that phosphorylated FixJ (FixJ∼P) binds these two classes of sites following two distinct binding modes.

2. Materials and methods

2.1. Preparation of the glutathione S-transferase (GST)–FixJC protein

Plasmid pGMI2061 used for GST–FixJC expression was obtained as follows. The starting material was the pBluescript derivative pDK330, a plasmid expressing FixJC [3]. The *fixJC* coding sequence was amplified in order to engineer an *Nde*I site overlapping the initiation codon, and the 0.5 kb *Nde*I–*Pst*I fragment containing *fixJC* was cloned into pT7-7 and verified by sequencing. After subcloning the *Xba*I–*Pst*I fragment into pBluescript-KS+, the resulting plasmid was linearised with *Nde*I and treated with T4 DNA polymerase prior to *Eco*RI cleavage. The isolated 0.5 kb fragment was cloned between the *Sma*I and *Eco*RI sites of pGEX-2T (Amersham Pharmacia) to generate pGMI2061. This plasmid was transformed into *Escherichia coli* DH5 in order to overexpress the GST–FixJC fusion protein which was purified on glutathione-Sepharose beads according to the manufacturer's instructions (GST Gene Fusion System, Amersham Pharmacia).

2.2. SELEX procedure

In vitro selection was performed on a pool of double-stranded DNA oligonucleotides randomised over 20 bp [8]. The initial oligonucleotide 5′-CATAGATAACTAGTAGCG-$(N)_{20}$-TTAGACTAGT-GACTCACA was made double-stranded using the Klenow fragment of DNA polymerase I with the OLB primer 5′-TGTGAGTCAC-TAGTCTAA, and purified on a 4% agarose gel. 100 ng of double-stranded DNA was incubated with 10 μM GST–FixJC for 15 min at 30°C in 100 μl of binding buffer (60 mM Tris-acetate, pH 8, 60 mM KCl, 30 mM K^+ acetate, 27 mM NH_4^+ acetate, 8.1 mM Mg^{2+} acetate, 1 mM dithiothreitol (DTT), 0.1 mM EDTA) containing 25 μg/ml polydI:dC, 100 μg/ml bovine serum albumin (BSA), 4% PEG6000 and 10% glycerol. Glutathione-Sepharose beads were added and the mixture was left for a further 10 min at room temperature to allow for adsorption of protein–DNA complexes. The beads were quickly rinsed with 1 ml of ice-cold binding buffer and bound DNA was eluted for 10 min at 100°C in 100 μl of binding buffer. The selected molecules were finally amplified by PCR using as primers OLB and OLA (5′-CATAGATAACTAGTAGCG) in a reaction mix containing 10 mM Tris–HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM each dNTP, 5% DMSO, 800 nM each primer and 0.02 U/μl Taq polymerase (Life Technologies). Amplification was carried out in an Eppendorf MasterCycler® Gradient instrument with the following cycle: 30 s at 94°C, 30 s at 52°C and 30 s at 72°C. The resulting DNA was subjected to further selection by iterating the same procedure until sufficient enrichment for FixJ binding sequences could be detected by gel retardation analysis. The resulting molecules were cloned in pGEM-T (Promega) and sequenced using the T7 universal primer.

*Corresponding author. Fax: (33)-561285061.
E-mail address: dkahn@toulouse.inra.fr (D. Kahn).

0014-5793/02/$22.00
PII: S0014-5793(02)02618-2

2.3. Gel retardation analysis

The FixJ protein was purified as described previously [5]. Phosphorylation was achieved by incubating 140 μM FixJ for 90 min at 28°C in a reaction mixture containing 50 mM Tris–HCl, pH 7.6, 0.1 mM EDTA, 0.1 mM DTT, 10 mM $MgCl_2$, 80 mM KCl, 25% glycerol and 20 mM acetyl-phosphate. The reaction was stopped by addition of 13 mM EDTA and kept on ice. The phosphorylation level ($62 \pm 5\%$) was verified by anion exchange chromatography as described earlier [5].

DNA fragments resulting from SELEX were radio-labelled by PCR using 5′-end-labelled T7 universal primer together with the SP6 universal primer. After 20 PCR cycles (30 s at 94°C, 30 s at 50°C and 30 s at 72°C), the resulting 215 bp PCR products were purified using the High Pure PCR Product Purification kit (Boehringer Mannheim). *fixK* promoter DNA was obtained as a 319 bp *Hin*cII–*Bam*HI fragment from pJMR300 [9] and end-labelled by filling in the *Bam*HI site using the Klenow fragment of DNA polymerase [10]. DNA was incubated with FixJ∼P for 15 min at 30°C in binding buffer containing 25 μg/ml polydI:dC, 100 μg/ml BSA, 4% PEG6000 and 10% glycerol. Reactions were loaded on a non-denaturing 10% polyacrylamide gel run in TBE buffer.

2.4. DNase I footprint analysis

Complexes between FixJ∼P (20 μM) and radio-labelled DNA fragments were obtained as above. DNase I attack was initiated by adding 7 mM $MgCl_2$, 3 mM $CaCl_2$ and 50 ng/ml DNase I (Sigma). After 1 min incubation at 30°C, the reaction was stopped by addition of 25 mM EDTA and kept on ice. FixJ∼P–DNA complexes were separated on a non-denaturing 8% polyacrylamide gel. Both the complexes and free DNA were excised and eluted overnight at 37°C in 10 mM Tris–HCl, pH 8, containing 1 mM EDTA, 0.3 M NaCl, 0.2% SDS and 1 μg/ml tRNA. After ethanol precipitation, radioactivity was counted using Čerenkov emission and 25000 cpm was loaded on a denaturing 6% polyacrylamide gel. The G+A ladder was generated using the Maxam and Gilbert protocol [11].

3. Results and discussion

3.1. Two classes of FixJ binding sites defined by in vitro selection

In order to select for FixJ binding sites we used a fusion protein between GST and the C-terminal DNA binding domain of FixJ (FixJC). Preliminary experiments established that this fusion protein binds *fixK* promoter DNA and that the resulting complexes can be specifically adsorbed on glutathione-Sepharose beads (data not shown). Therefore this GST–FixJC fusion protein can be used in order to select for DNA sequences with affinity for FixJ. A complex mixture of double-stranded oligonucleotides randomised over 20 bp was incubated with GST–FixJC, protein–DNA complexes were adsorbed on glutathione-Sepharose beads, bound DNA was eluted, amplified and protein–DNA complexes were selected again. After six selection cycles the randomised DNA appeared highly enriched in FixJ binding fragments which were cloned and sequenced. Sixty-three resulting independent sequences were analysed for common motifs using the MEME algorithm [12]. From this analysis a rather degenerate consensus sequence GTASKTWS could be extracted. However it became evident that most sequences could be partitioned into two homogeneous classes.

In the first class (47 sequences), a conserved GTAGTTTC-CC sequence lies always directly adjacent to the OLB primer sequence (Fig. 1). This observation suggested that the flanking sequence may also be involved in FixJ recognition. Indeed DNase I foot-printing experiments in the presence of FixJ∼P showed that the last five 3′-nucleotides of OLB are strongly protected against DNase I attack (Fig. 2). We therefore propose that this FixJ recognition site should be extended to

A)

Clone		SEQUENCE	
W41	actag	**tctaaGTAGTTGCCC**	TCGCTTGCCCcgc
W31	actag	**tctaaGTACTTGTCC**	GTAGTCGCCCcgc
W30b	actag	**tctaaGTAATGTCCG**	CAGCCCGCCcgct
W18	actag	**tctaaGTAGTTGTAC**	CGACCCGCCCcgc
W2a	actag	**tctaaGTAGAACCCC**	TTGGTCGCGCcgc
W30a	actag	**tctaaGTAGTTGTCC**	CGCTTGCTTCcgc
W38	actag	**tctaaGTATTCACCA**	CACGCCGCGCcgc
W1a	actag	**tctaaGTAGTTTCCC**	AGGTGCCGCCcgc
W37a	actag	**tctaaGTAGTTTCCC**	TGCCGCCCCCcgc
W39	actag	**tctaaGTAGTTCCCC**	GCATGACGCCcgc
W33	actag	**tctaaGTAGTTCCCC**	TCGGTCTGCCcgc
W1c	actag	**tctaaGTACTTTCCC**	TTGCCGCGGCcgc
W20c	actag	**tctaaGTAGTTTTCC**	GGATGTGGGCcgc
W11	actag	**tctaaGTATTTCCCC**	CAGGCCGGCCcgc
W22c	actag	**tctaaGTAGTTTCAC**	GTAGCCGGCCcgc
W26	actag	**tctaaGTAGTTTCAC**	CACGCCCACCcgc
W24a	actag	**tctaaGTACTTTCCC**	CCGACACCCCcgc
W3	actag	**tctaaGTAGTTTTCC**	TCGACGCGCCcgc
W24b	actag	**tctaaGTAGATGTAC**	CAACCGCGGCcgc
W20b	actag	**tctaaGTGGTTTCCA**	CAGGTGTGCCcgc
W21a	actag	**tctaaGTAGTTTCCC**	ACGCCCGAGCcgc
W14	actag	**tctaaGTAGTCTCCC**	GCACCCACGCcgc
W10	actag	**tctaaGTAGTTGCCA**	CGTCGCACCCcgc
W12	actag	**tctaaGTGTTTTCGC**	GTAGTGGTGCcgc
W8	actag	**tctaaGTAGTTTCCC**	AGCTTGCGCCcgc
W4a	actag	**tctaaGTAGTTACCA**	CAGCCGCCCCcgc
W21d	actag	**tctaaGTAGTTTCCA**	CCGGCGGCCCcgc
W28	actag	**tctaaGTACTTTCCG**	CACCCGGTCCcgc
W29	actag	**tctaaGTAGTATCCG**	CGCCCCGGGCcgc
W21b	actag	**tctaaGTACATGCCC**	CCTCCGTGCCcgc
W2c	actag	**tctaaGTGGTTGTGC**	GTAGTGCTGCcgc
W22b	actag	**tctaaGGGGAATCCC**	GTACTCGGCCcgc
J13b	actag	**tctaaGTAGTTTCCA**	CACGCTCGCCcgc
J13a	actag	**tctaaGTAGTTTCCG**	TACCGCGGCCcgc
A20a	actag	**tctaaGTAGTTCCCA**	CCATCGGCCcgct
J18	actag	**tctaaGTAGTTTCCC**	TGCATGCGGCcgc
J10	actag	**tctaaGTAGTTTCCC**	GGTGCCGTCCcgc
J4	actag	**tctaaGTAGTTTCCC**	GCCCTGGTGCcgc
J5	actag	**tctaaGTAGTTTCCC**	CGGGGTGCGCcgc
A18	actag	**tctaaGTACTTTCCC**	TCGGTGCGCCcgc
J61	actag	**tctaaGTACTTTCCG**	ACCCCCGTGCcgc
J91	actag	**tctaaGTAGTTTCCA**	CGGCGCGGCCcgc
J1	actag	**tctaaGTACTTTCCC**	CCACTCTGGCcgc
J8	actag	**tctaaGTAGTTGCCG**	CACGTCGACCcgc
J19	actag	**tctaaGTAGTTACCG**	CACCGGCAGCcgc
B3	actag	**tctaaGTAATGTCCG**	CAGCCCGCCcgct
A2	actag	**tctaaGTAGTTTCCC**	GGGCTGGTCCcgc
Consensus		**tctaaGTAGTTTCCC**	

B)

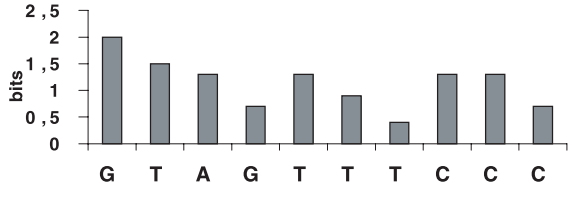

Fig. 1. Consensus for class I FixJ binding sites. A: Alignment of 47 selected FixJ binding sites. Sequences which were initially random are shown in capitals while constant flanking sequences are indicated in small case. B: Information content for each selected position in the consensus sequence.

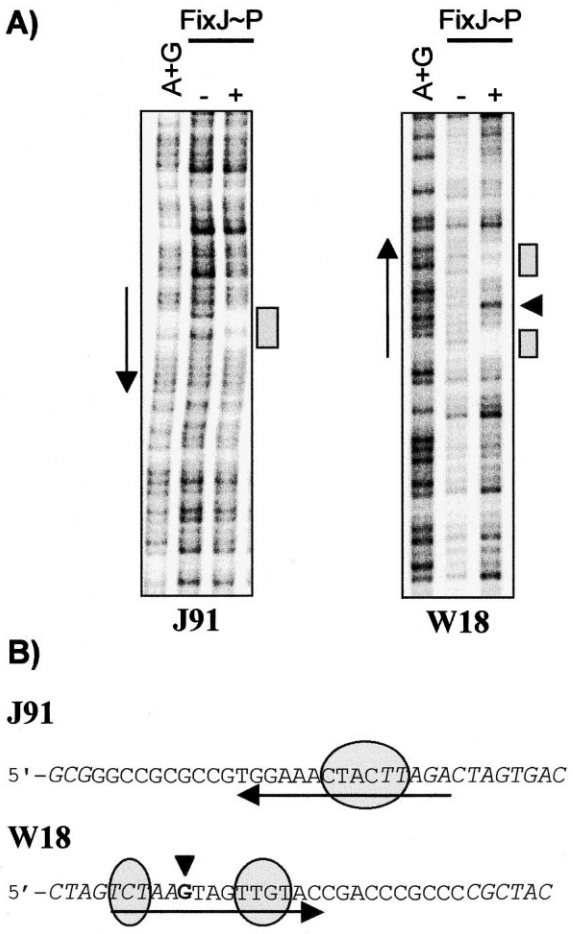

Fig. 2. FixJ∼P footprint on class I binding sites. A: End-labelled DNA from clones J91 and W18 was incubated without protein (−) or in the presence of 20 μM FixJ∼P (+). The A+G ladder was generated using the Maxam and Gilbert protocol. Fully protected regions and the DNase I hypersensitive site are indicated by shaded boxes and an arrowhead, respectively. The arrow corresponds to the consensus defined in Fig. 1. B: Location of FixJ∼P footprints. Same symbols as in A. Constant sequences flanking the selected variable sequence are italicised.

include the 3′-terminal OLB sequence (lower case): tctaaGTAGTTTCCC.

In the second class (16 sequences), a motif search with the MEME program [12] provided evidence for a conserved GTAMGTAG octamer fully contained within the randomised part of the original oligonucleotides (Fig. 3). Half of the selected sequences exhibit an exact TACGTA palindromic hexamer, including two sequences with an exact palindromic octamer (sequences A15 and J15, Fig. 3), suggesting a symmetrical binding mode for FixJ∼P.

3.2. Characterisation of FixJ∼P binding

DNase I foot-printing assays were performed on several selected fragments from both classes. On class I sequences FixJ∼P protected only the 15 nucleotide consensus site (Fig. 2). The characteristic alternating protection pattern between top and bottom strand suggests that the FixJ∼P dimer binds DNA on a single face of the DNA helix. The central hypersensitive site on the top strand indicates a widening of the minor groove on the opposite face as a result of DNA deformation in the complex.

On class II sequences FixJ∼P provided strongest protection on both strands over a 40 nucleotide stretch centred on the consensus octamer (Fig. 4). This extensive DNase I protection indicates minor groove interactions over three full DNA helical turns, which would require the binding of additional FixJ∼P dimers. Therefore we propose that the binding of one FixJ∼P dimer on a class II site triggers FixJ∼P oligomerisation on both sides of the consensus binding site.

Considering the differences between the two classes of FixJ binding sites, we further investigated whether these differences would affect FixJ∼P binding affinity. Gel-shift titration experiments showed that FixJ∼P binds class I sites with high cooperativity and a mid-titration point lying between 5 and 10 μM (Fig. 5). FixJ∼P binding to class II sequences showed a different pattern, with a slightly higher affinity and the appearance of at least one other complex type detected between 1 and 5 μM FixJ∼P. The latter observation indicates the binding of multiple FixJ∼P molecules on class II sites, consistent with FixJ∼P oligomerisation as suggested above from the extended footprint.

3.3. Comparison with natural FixJ targets

Consensus I appears quite similar to the high affinity FixJ binding site found in the *fixK* and *fixK′* promoters [6,10] with only one and two mismatches, respectively (Fig. 6A), corresponding to the three least conserved positions in class I sites (Fig. 1). Moreover a DNase I hypersensitive site is found at a matching position in *fixK* promoter DNA at position G_{-62} and in class I sites at position G_6. We conclude that the high affinity FixJ binding site in the *fixK* promoter belongs to class

A)

Clone	SEQUENCE		
A15	aaGGGCTGTTGC	**GTACGTAC**	cgctactagt
W27a	gtagcgGCCCGC	**GTAAGTGG**	AAACCCttag
W34	gtagcgGCGCGG	**GTAAGTAG**	TAGCCCttag
J15	gtagcgGGCCGG	**CTACGTAG**	AAGTGCttag
W4b	gtagcgGCCCCG	**GTACGCAG**	AAACCCttag
W22a	tagtctaaCCCC	**GTACGTAG**	TGCCCCTCcg
A3	tagtagcgGCCC	**GTACGTAG**	ACACCGCCtt
W9	gtctaaGGGATT	**GTGCGTAG**	TCGTGCcgct
A11	tagtagcgGCCC	**GTACGTAG**	AGCGCACAtt
J11	tagtagcgGCCC	**GTACGCAG**	ACACCGCAtt
D4	tagtctaaCCCC	**GTACGTAG**	TGCCCCCCcg
D5	tagtagcgGCAC	**GTAAGTAG**	AGCCCGGCtt
T29	tagtagcgGCCC	**GTAAGTAT**	GGTCCCCGtt
J3b	tagtagcgGCCC	**GTACGTGG**	ACCCCACAtt
W37b	tagtagcgGCCC	**GTACGTAG**	CGATACCGtt
W44	tagtctaaCCCC	**GTACGTAG**	TTGCCTCCcg
Consensus		**GTACGTAG**	
		A	

B)

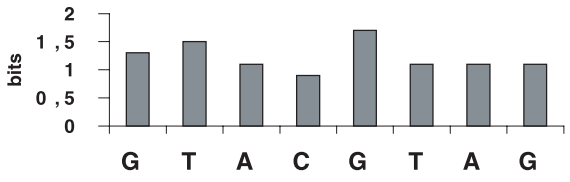

Fig. 3. Consensus for class II FixJ binding sites. A: Alignment of 16 selected FixJ binding sites. Sequences which were initially random are shown in capitals while constant flanking sequences are indicated in small case. B: Information content for each selected position in the consensus sequence.

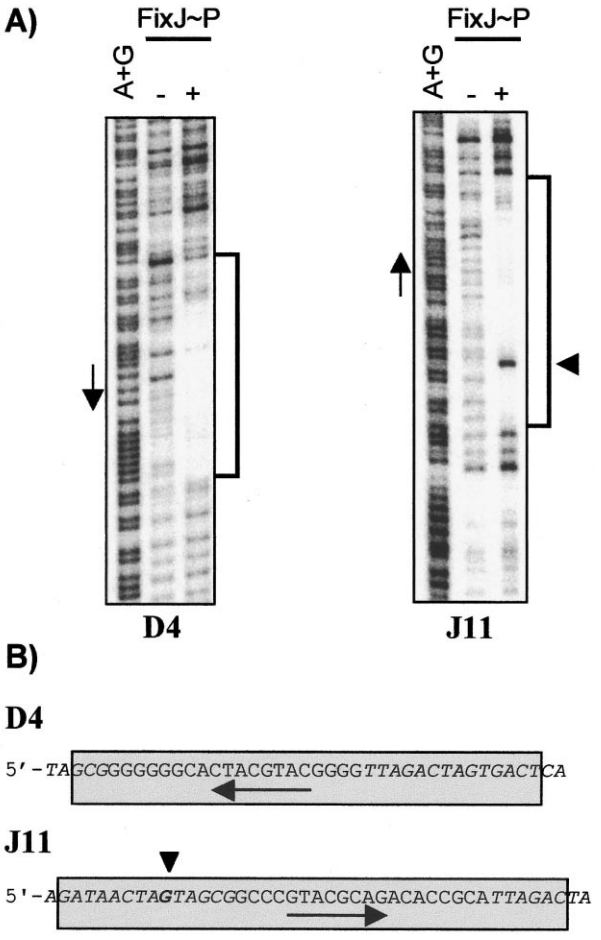

Fig. 4. FixJ∼P footprint on class II binding sites. A: End-labelled DNA from clones D4 and J11 was incubated without protein (−) or in the presence of 20 μM FixJ∼P (+). The A+G ladder was generated using the Maxam and Gilbert protocol. The protected region and the DNase I hypersensitive site are indicated by a bracket and an arrowhead, respectively. The arrow corresponds to the consensus defined in Fig. 3. B: Location of FixJ∼P footprints. The shaded box corresponds to the protected region. Constant sequences flanking the selected variable sequence are italicised.

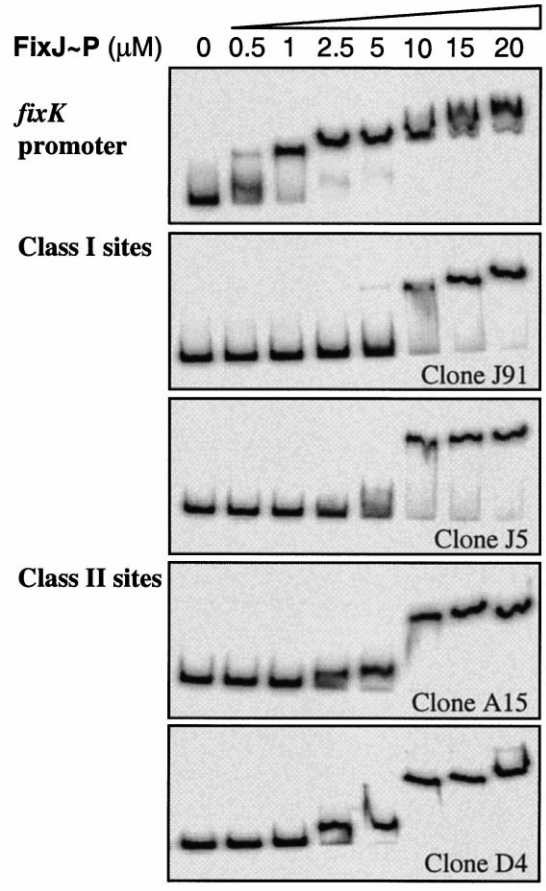

Fig. 5. FixJ∼P binds class I and class II sites with comparable affinities. Band shift titration experiments were performed on end-labelled DNA containing the *fixK* promoter (top), the J91 and J5 class I sites (middle), or the A15 and D4 class II sites (bottom). FixJ∼P molarities are expressed in terms of FixJ monomers.

I. This site plays a major role in *fixK* activation as mutations in the [−59,−67] region exhibit a strong down phenotype [6]. The function of this site is to recruit FixJ∼P to the low affinity downstream site which in turn is the essential functional site for promoter activation [10]. The 10-fold higher affinity of *fixK* promoter DNA, when compared to in vitro selected class I sites, can be accounted for by the existence of this second site which is known to bind FixJ∼P synergistically [10]. Note however that this downstream site does not share detectable sequence similarity to either class I or class II sites, which can be related to its low binding affinity. It is intriguing that the functionally essential, promoter proximal FixJ binding site is so atypical and exhibits low affinity for FixJ∼P. It is tempting to speculate that such a low affinity may be functionally important for *fixK* transcription.

The central TAMGTA hexamer of class II consensus shares similarity with the functionally important [−51,−46] region of the *nifA* promoter (Fig. 6B). The proposed imperfect match is congruent with the weak affinity of FixJ∼P for *nifA* promoter DNA (our unpublished observations). We note that the matching nucleotides include the key positions at −51, −48 and −46 defined earlier by genetic means, whereas mismatches at −52 and −45 correspond to unessential nucleotides in the *nifA* promoter [7]. We conclude that the alignment shown in Fig. 6B matches indeed the functional FixJ binding

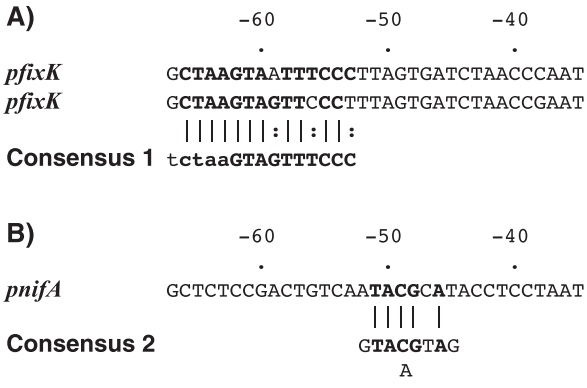

Fig. 6. Alignment of consensus sequences with natural FixJ targets. A: Alignment of class I consensus sequence with both copies of the *fixK* promoter found in *S. meliloti* [6]. Sequence numbering relates to the *fixK* transcription start [13]. B: Alignment of class II consensus sequence with the functional region of the *S. meliloti nifA* promoter [7].

site in the *nifA* promoter. Thus the two different classes of FixJ binding sites match two different types of *fixJ* regulated promoters, consistent with the existence of two distinct binding modes for FixJ∼P.

Acknowledgements: We wish to thank Joëlle Fourment for technical assistance. L.F. was the recipient of a 'Bourse de Docteur Ingénieur' from CNRS. This work was supported by Grants from the European Union (BIO4 CT 97-2143), the Ministère de l'Education Nationale, de la Recherche et de la Technologie (Programme de Recherche Fondamentale en Microbiologie et Maladies Infectieuses et Parasitaires) and INRA (Programme prioritaire 'Microbiologie').

References

[1] David, M., Daveran, M.L., Batut, J., Dedieu, A., Domergue, O., Ghai, J., Hertig, C., Boistard, P. and Kahn, D. (1988) Cell 54, 671–683.
[2] Fischer, H.M. (1994) Microbiol. Rev. 58, 352–386.
[3] Kahn, D. and Ditta, G. (1991) Mol. Microbiol. 5, 987–997.
[4] Da Re, S., Bertagnoli, S., Fourment, J., Reyrat, J.M. and Kahn, D. (1994) Nucleic Acids Res. 22, 1555–1561.
[5] Da Re, S., Schumacher, J., Rousseau, P., Fourment, J., Ebel, C. and Kahn, D. (1999) Mol. Microbiol. 34, 504–511.
[6] Waelkens, F., Foglia, A., Morel, J.B., Fourment, J., Batut, J. and Boistard, P. (1992) Mol. Microbiol. 6, 1447–1456.
[7] Agron, P.G., Ditta, G.S. and Helinski, D.R. (1992) J. Bacteriol. 174, 4120–4129.
[8] Tuerk, C. and Gold, L. (1990) Science 249, 505–510.
[9] Reyrat, J.M., David, M., Blonski, C., Boistard, P. and Batut, J. (1993) J. Bacteriol. 175, 6867–6872.
[10] Galinier, A., Garnerone, A.M., Reyrat, J.M., Kahn, D., Batut, J. and Boistard, P. (1994) J. Biol. Chem. 269, 23784–23789.
[11] Maxam, A.M. and Gilbert, W. (1980) Methods Enzymol. 65, 499–560.
[12] Bailey, T.L. and Elkan, C. (1994) Proc. Int. Conf. Intell. Syst. Mol. Biol. 2, 28–36.
[13] Batut, J., Daveran-Mingot, M.L., David, M., Jacobs, J., Garnerone, A.M. and Kahn, D. (1989) EMBO J. 8, 1279–1286