DNA sequence determinants of LexA-induced DNA bending

Roland Lloubès, Claude Lazdunski, Michèle Granger-Schnarr[1] and Manfred Schnarr[1,*]
Centre de Biochimie et de Biologie Moléculaire, 31 Chemin Joseph-Aiguier, BP 71, 13402 Marseille and [1]Institut de Biologie Moléculaire et Cellulaire du CNRS, UPR 9005, 15 rue René Descartes, 67084 Strasbourg, France

ABSTRACT

The LexA repressor from *Escherichia coli* induces DNA bending upon interaction with the two overlapping operators which regulate the transcription of the colicin A encoding gene *caa*. Both *caa* operators harbor T-tracts adjacent to their recognition motifs. These tracts have been suggested to be especially favorable for the promotion of LexA-induced DNA bending. Here we show that this is indeed the case, since disruption of the TTTT-tract adjacent to operator O1 by the replacement of the two central thymine bases by AA, GA or CG markedly reduces LexA-induced DNA bending. Simple A·T-richness in this position is thus not sufficient to promote full LexA-induced bending, albeit a TAAT sequence is always more efficient to promote bending than those sequences containing one or two C/G base pairs.

INTRODUCTION

The LexA repressor from *Escherichia coli* is a sequence-specific DNA binding protein which negatively regulates the initiation of transcription of about 20 so-called SOS genes which are induced upon DNA damage (for recent reviews see 1–3). The regulatory strategy differs from one gene to the other in that LexA binds to either one, two, or three adjacent operators, and in that the LexA binding sites are located in variable positions within the different promoters. In those cases where a LexA binding site overlaps one of the two conserved hexameric promoter elements the protein should act as a competitive inhibitor of RNA polymerase binding as shown in the case of the *uvrA* gene (4). In a few cases, i.e. the *uvrD* gene and most of the genes coding for colicins, LexA binds downstream from the RNA polymerase binding site, suggesting that RNA polymerase and LexA might bind simultaneously to these promoters as shown earlier in the case of the *lac* repressor-promoter system (5,6). We have studied in some detail LexA binding to the promoter/operator region of the *caa* gene, coding for the bacterial toxin colicin A. In this case LexA binds to two overlapping operators situated between positions +2 and +31 with respect to the transcription start point.

Each operator interacts with a LexA dimer such that the two centrally located monomers will not adopt exactly the same binding geometry as the two external monomers (7,8). Electrophoretic mobility shift assays revealed the presence of two complexes for a low degree of DNA saturation. One species corresponds to a single occupied operator, the other to a complex where both operators are occupied. Both complexes show LexA-induced DNA bending, and the presence of a small intrinsic bent of the free DNA centered on the operator region suggested to us that a T_4- and an A_5-stretch adjacent respectively to the left and to the right side of the operator region might be involved in LexA-induced DNA bending (9).

Using single-operator mutations, which confine LexA to one or the other of the two operators, we show here that DNA flexure upon LexA binding to the left operator (O1) is more pronounced than upon binding to the right operator (O2). This difference is most likely due to the nature of the DNA sequence flanking the operator region as shown by symmetrizing the internal elements of the two operators. A mutational study of the TTTT-stretch situated 5' to the CTGT-recognition motif of operator O1 showed that this intrinsically bent sequence is indeed especially suitable to promote LexA-induced DNA-bending. Simple A·T-richness is not sufficient to promote full LexA-induced bending, albeit a sequence like TAAT is always more efficient to promote LexA-induced bending than sequences like TGAT and TCGT containing one or two G/C base pairs.

MATERIALS AND METHODS

Plasmid and mutant operator construction

Bending mutant operators within the T_4-stretch 5' to the CTGT-motif of operator O1 (see figure 1) were constructed by site-directed mutagenesis (10) in the context of a O1⁺O2⁻ operator in which binding to operator O2 is suppressed by a double-mutation of the external CTGT-motif of operator O2 (CTGT→CCAT). We used further a O1⁻O2⁺ operator (for which binding to O1 is abolished) to construct a new O2 operator identical to O1 called O1⁻O2⁺(sym) (see figure 1). The matrix for oligonucleotide-directed mutagenesis was a M13mp19 phage harboring the *caa* regulatory region (8).

Figure 1. Restriction sites, sequence and operator mutations of the *caa* regulatory region. DNA restriction sites within the *caa* sequence are shown above the bold line, restriction sites within M13mp19 vector DNA (thin line) are shown below the line. The sequence of the HincII-HgiAI fragment is shown in detail, comprising the -35 and -10 elements, the transcription start point and the two LexA binding operators of the *caa* promoter. The different operator mutations used in this study are shown below the sequence.

Figure 2. Sequence of the tandem repeat of plasmid pHF constructed for the circular permutation analysis. Duplicate restriction sites (bold-faced) are labelled as follows: S=SacI; A=AvaII; N=NcoI; Pv=PvuI; M=MaeIII; B=BstNI; D=DraIII; P=PstI. The multiple cloning site is indicated by a grey line above the sequence. SmaI and XbaI were used respectively for the insertion of mutated operator DNA and for radiolabelling.

In order to facilitate a bending analysis by circular permutation a special vector (pHF) was constructed as follows: a pUC12 vector was linearized with PstI, dephosphorylated, and recut with SacI. Two purified DNA fragments were simultaneously inserted in this vector: a dephosphorylated SacI-PstI DNA fragment (161 bp: S(1) to P(1) in figure 2) from pColA9 (restriction sites located at position 216 and 377 in the sequence described in (11)), and a phosphorylated PstI DNA fragment (211 bp: P(1) to P(2) in figure 2) obtained from plasmid pGE. pGE was obtained by the

insertion of a synthetic DNA fragment (sequence comprised between base pairs 185 and 211 in figure 2) and a SacI-PstI DNA fragment of pColA9 (161 bp) into a BamHI-PstI linearized pTTQ19 vector (12). The pHF plasmid for bending analysis contains thus a tandem sequence with 7 restriction sites present in duplicate and a central multiple cloning site (see figure 2).

In order to insert different *caa* operator constructions into pHF, double-stranded DNAs from M13mp19 phages harboring respectively the *caa* operators $O1^+O2^-$, $O1^-O2^+$ and $O1^+O2^+$ were purified, 80 bp HincII-HgiAI fragments (see figure 1) were isolated, subjected to Klenow polymerase treatment and inserted into the SmaI site of pHF to give pHL, pHM and pHN plasmids respectively. These plamids were selected by DNA sequencing such that all three operator constructions were cloned in the same orientation.

Electrophoretic mobility shift assays

These assays were done as previously described using 5% polyacrylamide gels (8) with DNA fragments derived either from double-stranded phage DNA or from plasmids pHL, pHM, and pHN (see above). In the first case double-stranded M13mp19 DNAs harboring the *caa* operators $O1^+O2^-$, $O1^-O2^+$ and $O1^+O2^+$ were cut with EcoRI, 5' end-labelled with ^{32}P, and recut with PvuII giving 364 bp DNA fragments with centrally located operators (see figure 1) which were purified by polyacrylamide gel electrophoresis and electroelution.

For the circular permutation analysis plasmids pHL, pHM and pHN were linearized with XbaI, 5'end-labelled and recircularized with ligase. After ligation, dephosphorylation was performed to remove the 5' end-labels from the linear fragments. The 5' labelled circular plasmids were then recut with the following restriction enzymes: SacI, NcoI, PvuI, BstNI and PstI giving 291 bp DNA fragments.

For competition binding analysis two or three different operators were used in the same band shift assay. The different operator DNA fragments were: EcoRI-HincII (224bp) for $O1^+O2^+$; BstNI-HincII (177bp) for $O1^+O2^-$; BglI-HincII (358bp) for $O1^+O2^-$(TCGT); PvuI-HincII (333bp) for $O1^+O2^-$(TGAT); SacI-HincII (218bp) for $O1^+O2^-$(TAAT); HphI-HincII (201bp) for $O1^-O2^+$; EcoRI-HincII (224bp) for $O1^-O2^+$(sym) (see figure 1). All these fragments were 5'end-labelled at the HincII site.

RESULTS

DNA curvature is more pronounced upon LexA-binding to operator O1 than to operator O2

In order to elucidate the DNA sequence requirements for LexA-induced DNA bending of the *caa* regulatory region, we have compared in a first step the relative degree of DNA deformation upon LexA-binding to operators O1 and O2. The two operators are generally occupied in a cooperative manner, but mutations within one or the other operator allow to study LexA binding to each operator independently (8).

A comparison of lanes 1 and 3 in figure 3A shows that a DNA fragment harboring the LexA binding sites roughly in the center of the fragment is more strongly retarded if LexA binds to operator O1 (mutant operator $O1^+O2^-$) than upon binding to operator O2 (mutant operator $O1^-O2^+$). This difference in migration upon O1 or O2 occupancy depends on the location of the LexA binding site(s) within the DNA fragment, since the two

A

B

A

B

Figure 4. Panel A shows that the nature of the nucleotide sequence upstream of the *caa* operator O1 influences the gel mobility of specific LexA-DNA complexes formed with *caa* operator O1. In each case the DNA was a 364 bp EcoRI-PvuII fragment harboring a $O1^+O2^-$ operator. The LexA concentrations were respectively 8 nM (for TTTT) and 16 nM (for TCGT, TGAT, and TAAT). **Panel B** shows a sequence comparison between the left half-site of the CAP binding site within the *lac* promoter and the left half-site of *caa* operator O1. The conserved nucleotides within each class of binding sites are underlined and the dinucleotide steps are numbered from 1 to 12 starting from the binding site dyad axis (♦). Below the sequence are shown the individual CAP-based bendability values (20) for the three mutated dinucleotide steps as well as the average of these values.

Figure 3. Panel A shows a comparison of the gel mobility of different LexA-DNA complexes formed with 364 bp EcoRI-PvuII DNA fragments harboring either the wild-type *caa* operator region ($O1^+O2^+$) or different single-operator mutations ($O1^+O2^-$, $O1^-O2^+$, $O1^-O2^+$(sym)). The LexA concentrations from lane 1 to 4 were respectively 10, 4, 25, and 25 nM leading to about fifty pourcent binding for the four DNA targets. As reported earlier (8) LexA binds to the $O1^+O2^-$ operator with about 2-fold higher affinity than to the $O1^-O2^+$ operator. Within experimental error the $O1^-O2^+$(sym) operator binds LexA under these conditions with the same affinity as the normal $O1^-O2^+$ operator. **Panel B** shows the gel mobility of two different LexA-DNA complexes (respectively with $O1^-O2^+$ and $O1^+O2^-$ operator DNA) as a function of the localisation of the LexA binding site on circularly permutated 291 bp DNA fragments obtained by the digestion of plasmids pHL ($O1^+O2^-$) and pHM ($O1^-O2^+$) with the following restriction enzymes: S=SacI; N=NcoI; Pv=PvuI; B=BstNI; P=PstI. LexA concentrations were 10 and 20 nM for respectively $O1^+O2^-$ and $O1^-O2^+$ operator DNA.

complexes comigrate if the binding sites are close to one end of the DNA fragment (see figure 2 in ref. 8).

In order to adress the question if the difference in LexA-induced curvature is due to DNA sequence elements inside or outside of the two operators we have mutated the internal sequence of operator O2 (i.e. the two underlined bases in the sequence of O2 shown below) such that it becomes identical to O1:

O1: 5'-TTTTACTGTATATAAACACATGTG
O2: 5'-AAAAACTGTATATA**TT**CACATGTG

A comparison of lanes 3 and 4 in figure 3A shows that LexA-induced curvature upon binding to operator O2 does not increase if the internal sequence of O2 between the two bold-faced recognition-motifs is identical to the internal sequence of O1.

We may conclude from this experiment that the observed mobility difference between the complexes formed at O1 and O2 most likely arises from the nature of the DNA sequences situated 5' to the conserved CTGT motif (the sequence 3' to the conserved

ACA motif is the same for both operators), i.e. the immediatly adjacent TTTTA and AAAAA sequences and possibly sequence elements even farther away from the center of the two operators.

If these external sequence elements are indeed a major determinant of LexA-induced curvature of the *caa* regulatory region, then the mobility difference observed in figure 3A between complexes formed at O1 and O2 might in principle be simply due to the fact that the bending loci of the two operators would be potentially displaced by as much as 35 base pairs. The operator O2 bending locus would be closer to the EcoRI end of the DNA fragment than the O1 bending locus (see figure 1) giving potentially rise to a position-dependent increased gel mobility of the LexA-O2 complex as compared to the O1 complex.

In order to adress this possibility the $O1^+O2^-$ and $O1^-O2^+$ mutant operators have been subcloned into a plasmid (pHF) suitable for a convenient application of the 'circular permutation assay', which consists in creating a family of DNA fragments of identical length but variable position of the protein binding site (13). Figure 3B shows the expected bell-shaped migration behaviour for LexA-binding to both operator families, i.e. those DNA fragments for which the LexA binding site is close to the center of the fragment are more strongly retarded upon complex formation than those fragments harboring the binding site close to one extremity. Figure 3B shows further that, regardless of the position of the binding site, LexA binding to operator O1 always induces stronger gel retardation than binding to operator O2. We may conclude from these experiments that the difference in LexA-induced curvature upon binding to O1 and O2 is an intrinsic property of the two operators, most likely linked to sequence elements situated 5' to the conserved CTGT recognition motifs. The sequence elements immediatly preceeding these motifs are TTTTA for O1 and AAAAA for O2. Both sequences should promote intrinsic curvature of the DNA (for reviews on intrinsic and protein-induced DNA curvature see 14−18) albeit not necessarily with the same efficiency.

A T_4-stretch adjacent to the LexA binding site is particularly favorable for protein-induced bending

In the following we have subjected the TTTT-stretch preceeding operator O1 to a mutational study, first, because this sequence seems to support LexA-induced curvature more efficiently than the A_5-stretch adjacent to O2, albeit an influence of sequence elements even farther apart from the dyad axis may not be excluded. Second, the T_4-stretch is localized just in front of the transcriptional start point at the following adenine base (11) and LexA-induced DNA bending in this region might contribute to transcriptional repression, since RNA polymerase has been reported to induce DNA bending and/or increased flexibility 3 base pairs upstream of the initiation point of RNA synthesis in the case of the A1 promoter of bacteriophage T7 (19), i.e. in a position which would be situated in the case of the *caa* promoter within the TTTT motif.

Three different $O1^+O2^-$ mutant operators have been constructed which contain respectively **TAAT**, **TGAT** and **TCGT**- motifs (instead of the wild-type TTTT-motif) in front of operator O1. The choice of the mutated dinucleotides in the center of the T_4-stretch was based on the dinucleotide ranking deduced from CAP-induced DNA bending (see table 2 in reference 20). According to this ranking an AA-step is as favorable for DNA bendability towards the minor groove as a TT-step, whereas GA- and CG-steps are increasingly unfavorable for this type of DNA bending.

Figure 4A shows however that all three mutant motifs are less efficient in supporting LexA-induced DNA bending than the original TTTT-motif. The order of bendability is such that:

$$TTTT > TAAT > TGAT \cong TCGT,$$

i.e. simple A·T-richness of this sequence is not sufficient to support full LexA-induced bending. However an A·T-rich sequence in this position is still more favorable than a sequence containing one or two G/C-base pairs. This latter finding is in agreement with the DNA sequence determinants of bending induced by the CAP protein, whereas the preference of LexA for an intrinsic bent sequence over simple A·T-richness seems to be a particular feature of the LexA complex.

DNA binding affinity

We asked further if mutations within the T_4-stretch have a measurable influence on the DNA binding affinity of LexA. Since the differences in binding affinity are expected to be small, because these mutations are located outside the recognition sequence, we have used a competition gel retardation assay. DNA fragments containing mutant binding sites were mixed with a DNA fragment of different size containing the wild-type *caa* regulatory region and incubated with variable amounts of the LexA repressor. The different species were resolved by gel electrophoresis and the relative amounts of free and bound DNA were determined by densitometry of the autoradiographs. The relative amount of LexA, with respect to the wild-type sequence, necessary to bind half of the DNA was 0.9 for TAAT, 1.5 for TGAT and 2.2 for TCGT. Those DNA fragments which support DNA bending poorly (TGAT and TCGT) have thus an about two-fold smaller binding affinity for the LexA repressor. However as in the case of the CAP protein (17, 20) the correlation between binding strength and bending is not perfect in that TGAT and TCGT show essentially the same degree of bending but different binding strength, whereas TTTT and TAAT show a different degree of bending but similar binding affinity.

DISCUSSION

Protein-induced DNA bending has been observed for numerous sequence specific DNA binding proteins including transcriptional activators and repressors (for reviews see 15, 17). Only in very few cases the DNA sequence determinants for protein-induced DNA bending have been adressed, the most thoroughly studied example being the CAP protein (17, 20). In this case A·T-richness in a sequence element flanking the recognition element is necessary and sufficient to support DNA bending. An earlier study on LexA-induced bending of the colicin A regulatory region suggested that in this case a simple A·T-rich sequence flanking the recognitized base pairs might not be sufficient to support full bending activity, but that a segment of intrinsically bent DNA might be required. Here we show that this hypothesis was essentially correct, since the disruption of a T_4-tract flanking the *caa* operator O1 increases the gel mobility of the LexA-DNA complex.

Unexpectedly the A_5-tract flanking operator O2 seems to be less efficient than the T4-tract adjacent to operator O1 in supporting LexA-induced DNA-bending. This difference is only observed upon inactivation of one of the operators, since the wild-type *caa* operator does not give rise to two different single-operator complex species (see figure 3A, lane 2 and reference 9). This finding is likely to be due to dissociation and reassociation events during gel electrophoresis, which may lead to the appearence of a single intermediate complex or even a complete disappearence of any intermediate complex as shown in the case of the Tet repressor-operator system (21). In the LexA case the intermediate complex with the wild-type *caa* operator $(O1^+O2^+)$migrates close to the position of the O1 complex (see figure 3A) in agreement with earlier findings that O1 binds LexA more tightly than O2 (8).

The finding that the A_5-tract adjacent to O2 seems to support bending less efficiently than the T_4-tract adjacent to O1 is unexpected in view of the observation that the gel mobility of intrinsically bent sequences like $(A_6xxxxA_6xxxx)_n$ is in general similar to that of sequences like $(A_6xxxxT_6xxxx)_n$ suggesting that the direction and magnitude of bending is largely independent of the orientation of the $(dA)\cdot(dT)$ tract (22). However, if the gel mobility of these sequences is examined in the presence of a ligand like Mg^{2+}, the gel mobility of these sequences is not anymore identical (23) and one may suppose that a protein ligand like LexA might also disrupt this apparent symmetry.

In the case of the CAP protein, Gartenberg and Crothers (20) have shown that maximal bending occurs when a site situated at about one helical turn from the binding site dyad is A·T rich, whereas minimal bending occurs when this site is G·C rich. The reverse pattern, though less pronounced, is observed at a site centered roughly 16 base pairs from the dyad axis. Alignment of the dyad axes of the CAP and the LexA binding sites (figure 4B) shows that the three dinucleotide steps of the T_4-tract adjacent to the *caa* operator O1 (i.e. steps 12, 11 and 10 with respect to the dyad axis) coincide with the three dinucleotide steps of the CAP binding site which are most sensitive to the incorporation of G/C base pairs in terms of reduction of DNA bending (17). Figure 4B summarizes further the bendability values for the three dinucleotide steps as deduced from CAP-induced bending (20). The CAP-based average values show that TTTT and TAAT should support bending with equal efficiency. This is clearly not the case for LexA-induced bending (as shown in figure 4A) suggesting that the DNA segment with the property to form an intrinsic bent is more favorable to support LexA-

induced bending. LexA is likely to enhance the natural tendency of T-tract bending towards the minor groove as suggested earlier (9) in order to form additional, most likely non-specific electrostatic contacts with nucleotides outside the specific recognition sequence.

DNA backbone atoms situated at the 3'-edge of the TTTT-stretch are expected to be in contact with LexA even in the absence of intrinsic DNA bending, since in the case of the *recA* operator, those phosphates for which ethylation suppresses complex formation (24) extend until a position corresponding to the phosphate between the third and the fourth thymine of the TTTT-tract (i.e. the position corresponding to dinucleotide step 10 in figure 4B). Furthermore the two nucleotides belonging to this step may be crosslinked with a photoreactive agent attached site-specifically to a LexA mutant repressor harboring a cysteine in position 52 (25). It is thus plausible that even minor changes of the DNA path due to intrinsic and/or induced DNA bending might bring DNA segments situated further apart from the dyad axis (most probably: steps 11 and 12) in close contact with the protein.

Once the TTTT-tract (and thus intrinsic bending) has been disrupted, LexA seems to behave similar to the CAP protein in that G/C base pairs in these positions unfavour protein-induced bending with respect to A/T base pairs. In the case of the CAP protein this observation has been linked to the bendability of these sequence elements towards the minor groove allowing the DNA to wrap around the protein (20).

ACKNOWLEDGEMENTS

REFERENCES

1. Walker, G.C. (1987) In Neidhardt, F.C. (ed.), *Escherichia coli* and *Salmonella Thypimurium*, American Society of Microbiology, Washington, Vol. 2, pp. 1346−1357
2. Little, J.W. (1991) *Biochimie* **73**, 411−422.
3. Schnarr, M., Oertel-Buchheit, P., Kazmaier, M. and Granger-Schnarr M. (1991) *Biochimie* **73**, 423−431.
4. Bertrand-Burggraf, E., Hurstel, S., Daune, M. and Schnarr, M. (1987) *J. Mol. Biol.* **193**, 293−302.
5. Schmitz, A. and Galas, D.J. (1979) *Nucl. Acids Res.* **6**, 111−137.
6. Straney, S.B. and Crothers, D.M. (1987) *Cell* **51**, 699−707.
7. Granger-Schnarr, M., Lloubès, R., de Murcia, G. and Schnarr, M. (1988) *Anal. Biochem.* **174**, 235−238.
8. Lloubès, R., Granger-Schnarr, M., Lazdunski, C. and Schnarr, M. (1991) *J. Biol. Chem.* **266**, 2303−2312.
9. Lloubès, R., Granger-Schnarr, M., Lazdunski, C. and Schnarr, M. (1988) *J. Mol. Biol.* **204**, 1049−1054.
10. Sayers, J.R., Schmidt,W. and Eckstein, F. (1988) *Nucl. Acids Res.* **16**, 791−802.
11. Lloubès, R., Baty,D. and Lazdunski, C. (1986) *Nucl. Acids Res.* **14**, 2621−2636.
12. Stark, M. (1987) *Gene* **51**, 255−267.
13. Wu, H.M. and Crothers, D.M. (1984) *Nature* **308**, 509−513.
14. Diekmann, S. (1987) In Eckstein, F. and Lilley, D.M.J. (ed.) *Nucleic Acids and Molecular Biology*, Vol. 1, pp. 138−156.
15. Travers, A.A. (1989) *Annu. Rev. Biochem.* **58**, 427−452.
16. Hagerman, P.J. (1990) *Annu. Rev. Biochem.* **59**, 755−781.
17. Crothers, D.M., Gartenberg, M.R. and Shrader, T.E. (1991) *Meth. Enzymol.* **208**, 118−146.
18. Lane, D., Prentki, P. and Chandler, M. (1992) *Microbiol. Rev.* **56**, 509−528.
19. Heumann, H., Ricchetti, M. and Werel, W. (1988) *EMBO J.* **7**, 4379−4381.
20. Gartenberg, M.R. and Crothers, D.M. (1988) *Nature* **333**, 824−829.
21. Kleinschmidt, C., Tovar, K. and Hillen, W. (1991) *Nucl. Acids Res.* **19**, 1021−1028.
22. Koo, H.-S., Wu, H.-M. and Crothers, D.M. (1986) *Nature* **320**, 501−506.
23. Diekmann, S. (1987) *Nucl. Acids. Res.* **15**, 247−265.
24. Hurstel, S., Granger-Schnarr, M. and Schnarr, M. (1988) *EMBO J.* **7**, 269−275.
25. Dumoulin, P., Oertel-Buchheit, P., Granger-Schnarr, M. and Schnarr, M. (1993) *Proc. Natl. Acad. Sci. USA* **90**, 2030−2034.