Structure of the LexA Repressor−DNA Complex Probed by Affinity Cleavage and Affinity Photo-Cross-Linking†

Pascal Dumoulin,‡ Richard H. Ebright,§ Ronald Knegtel,||,⊥ Robert Kaptein,|| Michèle Granger-Schnarr,‡ and Manfred Schnarr*,‡

Institut de Biologie Moléculaire et Cellulaire, UPR 9002 du CNRS, 15 rue Descartes, F-67084 Strasbourg Cedex, France, Department of Chemistry and Waksman Institute, Rutgers University, New Brunswick, New Jersey 08855, and Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584-CH Utrecht, The Netherlands

Received December 11, 1995⊗

ABSTRACT: The structure of the complex of full-length *Escherichia coli* LexA repressor with a consensus operator DNA fragment has been probed by affinity photo-cross-linking and affinity cleavage. These methods allow the determination of approximate intermolecular distances between a given protein residue and a base or sugar moiety within the operator. In a first step unique cysteine residues were introduced in positions 7, 28, 38, or 52 of the protein. In all four cases, the original amino acid was an arginine. The four amino acids in these positions were expected to be situated on the surface of LexA interacting with DNA, as infered from the structure of the LexA DNA binding domain [Fogh et al. (1994) *EMBO J. 13*, 3936−3944]. In a second step, these unique cysteine side chains of the purified proteins were chemically modified either with 4-azidophenacyl bromide or with *S*-(2-pyridylthio)cysteaminyl-EDTA. The first set of derivatives gives rise to UV-induced cross-linking which may be revealed by alkali/heat treatment; the second leads to direct DNA cleavage in the proximity of the derivatized amino acid. To reduce hydroxyl radical diffusion, the EDTA·iron cleavage reactions were done in the presence of high amounts of glycerol. The results indicate that amino acids 7 and 52 are near nucleotide pairs 8−12 of the operator and that amino acids 28 and 36 of LexA are near nucleotide pairs 5−8 of the operator. The results unambiguously define the orientation of the LexA DNA binding domain relative to the operator and provide support for the model of the LexA−operator complex proposed by Knegtel et al. [(1995) *Proteins 21*, 226−236]. Ethylation interference experiments further suggest that Arg-7 contacts the phosphate group between nucleotides 8 and 9 as predicted by the model.

The LexA repressor from *Escherichia coli* is a protein of 202 residues that regulates the transcription of about 20 genes known as SOS genes. These genes are involved in numerous cellular functions including DNA repair and replication, mutagenesis, and cell division [for reviews see Walker (1984), Little (1991), Schnarr et al. (1991), and Schnarr and Granger-Schnarr (1993)].

LexA is a two-domain protein that binds DNA via its amino-terminal domain (Little & Hill, 1985; Hurstel et al., 1986). The protein dimerizes by means of its carboxy-terminal domain with a rather small association constant of only 2×10^4 M^{-1} (Schnarr et al., 1985, 1988) and regulates transcription upon binding to its palindromic operator sequences such that, at least *in vitro*, two LexA monomers bind sequentially and cooperatively to the two operator half-sites (Kim & Little, 1992).

Sequence conservation among the different SOS operators suggests that an optimal operator should have a 2-fold symmetric $CTGT(AT)_4ACAG$ sequence (Wertman & Mount, 1985). This sequence binds LexA indeed more tightly than any naturally occurring SOS operator (Schnarr et al., 1991). The major recognition elements should be comprised in the two outer elements (the CTGT motifs) as judged from the distribution of operator-down mutations and methylation protection and interference data [compiled in Schnarr et al. (1991)].

The three-dimensional structure of the LexA DNA binding domain (DBD) has recently been solved by NMR spectroscopy (Fogh et al., 1994). The domain contains three α helices (residues 6−21, 28−35, and 41−52) and two antiparallel β strands (residues 56−58 and 66−68). Helices 2 and 3 form a variant helix−turn−helix (HTH) DNA binding motif with an unusual one-residue insert at residue 38. LexA shares this property with the c-Myb protooncogene product which contains also an additional amino acid in the turn region of the DNA binding repeat R3 (Ogata et al., 1992, 1994). Using amino acids 28−53 of LexA to probe the Swiss protein data bank with the BLAST program revealed also a significant primary structure homology of the LexA HTH motif and the mouse, human, and chicken c-Myb HTH motif of repeat 3 (P. Dumoulin and M. Schnarr, unpublished).

The LexA DNA binding domain is further a member of the "winged" HTH class of proteins (Holm et al., 1994) in which a β strand/loop/β strand element follows the HTH motif [for a review on winged HTH proteins, see Brennan (1993)]. To this class of $\alpha+\beta$ proteins belong both canonical HTH proteins like CAP (Schultz et al. 1991) and BirA

† This work was supported by an INSERM grant to M.G.-S., CNRS Grant 95N88/0034 to M.S., and NIH Grant GM41376 to R.H.E.

* Corresponding author. Tel: (33) 88 41 70 64. FAX: (33) 88 60 22 18. E-mail: schnarr@astorg.u-strasbg.fr.

‡ Institut de Biologie Moléculaire et Cellulaire.

§ Rutgers University.

|| Utrecht University.

⊥ Present address: Department of Pharmaceutical Chemistry, School of Pharmacy, UCSF, 513 Parnassus Ave., San Francisco, CA 94143-0446.

⊗ Abstract published in *Advance ACS Abstracts,* March 15, 1996.

0006-2960/96/0435-4279$12.00/0

(Wilson et al., 1993) and proteins with a longer turn between the two helices of the HTH motif like HNF-3γ (Clark et al., 1993), the GH5 human histone 5 (Ramakrishnan et al., 1993), and the HSF *Kluyveromyces lactis* heat shock transcription factor (Harrison et al., 1994).

The structure of the LexA−DNA complex has not yet been determined. Efforts to obtain the structure of the complex by X-ray analysis or by NMR spectroscopy failed so far due to poor crystallization and extensive line broadening.

Alternative approaches to probe the structure of a specific protein−DNA complex are affinity cleavage and affinity photo-cross-linking. In these approaches, a small DNA cleaving agent or a small photo-cross-linking agent is attached site-specifically to the protein, allowing identification of nucleotides positioned close to a particular amino acid (Chen & Sigman, 1987; Oakley & Dervan, 1990; Ebright et al., 1990; Pendergrast et al., 1992). In previous work, we have used this approach to show that helix 3, the putative recognition helix (Oertel-Buchheit et al., 1990) spanning amino acids 41−52, is in close contact with DNA with its NH_2 terminus most likely oriented toward the 2-fold axis of the operator (Dumoulin et al., 1993). This orientation is opposite to that generally observed for HTH proteins with the exception of *lac* repressor and homologues (Boelens et al., 1987; Lehming et al., 1990; Shin et al., 1991) and possibly *tet* repressor (Wissmann et al., 1991; Baumeister et al., 1992).

With the determination of the three-dimensional structure of the LexA DBD (Fogh et al., 1994) it was possible to design a series of LexA derivatives having the DNA cleaving agent EDTA·iron or the photo-cross-linking agent phenyl azide at positions likely to be situated on the surface of the protein that interacts with DNA. Here we report affinity cleavage and affinity cross-linking data for LexA derivatives having EDTA·iron or phenyl azide incorporated at positions 7, 28, 38, and 52.

MATERIALS AND METHODS

Site-Directed Mutagenesis and Protein Purification. A LexA derivative having a unique cysteine residue, LexA-C52, has been selected previously (Oertel-Buchheit et al., 1990). Three other LexA derivatives with unique cysteines (LexA-C7, LexA-C28, and LexA-C38) were constructed by site-directed mutagenesis according to the procedure described by Sayers et al. (1988a,b).

LexA-C38 and LexA-C52 were overexpressed using a *lacUV5* promoter (Schnarr et al., 1985) and LexA-C7 and LexA-C28 with the T7 expression system (Dubendorff & Studier, 1991; Roland et al., 1992). The mutations were introduced into pJWL228 (harboring the gene for the LexA wild-type repressor) (Roland et al., 1992) by ligation of a *Hpa*I−*Sna*BI fragment into the purified *Sna*BI (partial)/*Hpa*I-digested pJWL228 vector. The mutant proteins were purified essentially as described for wild-type LexA (Schnarr et al., 1985) except that NaCl was omitted from the dialysis buffer prior to loading the phosphocellulose column.

Modification of the Purified Proteins. Modification with 4-azidophenacyl bromide (Sigma) was done as described by Dumoulin et al. (1993) and modification with *S*-(2-pyridylthio)cysteaminyl-EDTA as described by Ebright et al. (1992).

The degree of modification of the unique cysteine in the different LexA mutant repressors was determined by measuring the amount of unmodified cysteine using 2-vinylpyridine (Friedman et al., 1970). In all cases the degree of modification with 4-azidophenacyl bromide or with *S*-(2-pyridylthio)-cysteaminyl-EDTA was greater than 90%.

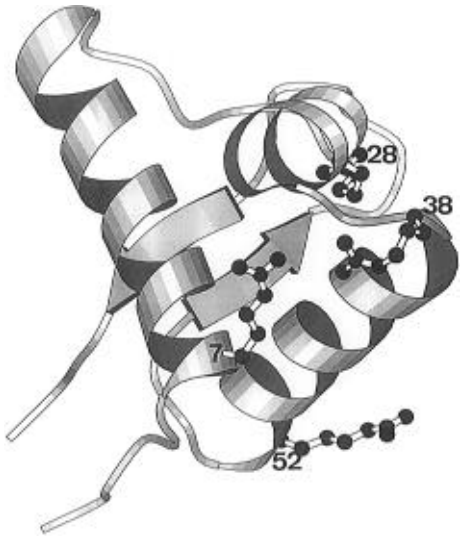

FIGURE 1: Structure of the LexA DNA binding domain (amino acids 1−72) showing the position and the side chains of the four arginine residues, 7, 28, 38, and 52, which have been individually mutated to cysteine and subsequently modified with a endonucleolytic or photo-cross-linking agent. The drawing was established using the program MolScript (Kraulis, 1991).

Phenyl Azide-Mediated Photo-Cross-Linking. LexA derivatives (about 3 μM) were incubated in the dark at 20 °C for 30 min in a binding buffer (20 mM MOPS, pH 7.3, and 100 mM KCl) with an operator DNA fragment (50-mer) ^{32}P-labeled at the 5′-end of either the top or the bottom strand. Fifty microliter aliquots were withdrawn and UV-irradiated with a 6W Vilber Lourmat VL-6MC UV lamp (312 nm) at a distance of 20 cm for 5 min. UV-irradiated samples were subjected to three phenol extractions (discarding aqueous phases), followed by ethanol precipitation and piperidine cleavage (1 M) for 30 min at 90 °C. G>A and G+A sequencing reactions were done according to Maxam and Gilbert (1980). High-resolution gel electrophoresis was carried out in 8 M urea−20% polyacrylamide denaturing gels.

EDTA·Iron-Mediated Affinity Cleavage. Reactions were done essentially as described by Ebright et al. (1992) using a different binding buffer (20 mM MOPS, pH 7.3, 100 mM KCl, 50 μg/mL BSA, and 40% glycerol). LexA derivatives ($\approx$3 μM) were incubated at 20 °C for 30 min with an operator DNA fragment ^{32}P-labeled at the 5′-end of either the top or the bottom strand. After addition of 1 mM ascorbic acid, the reaction proceeded for another 5 min at 4 °C. Further steps were done as described (Ebright et al., 1992).

RESULTS

Figure 1 presents the structure of the LexA DBD (Fogh et al., 1994) showing the four arginine residues which we have individually substituted by cysteine and subsequently subjected to chemical modification. These residues are located at the N-terminus of helix 1 (Arg-7), the N-terminus of helix 2 (Arg-28), the turn between helix 2 and helix 3 (Arg-38), and the C-terminus of helix 3 (Arg-52) as shown in Figure 2A.

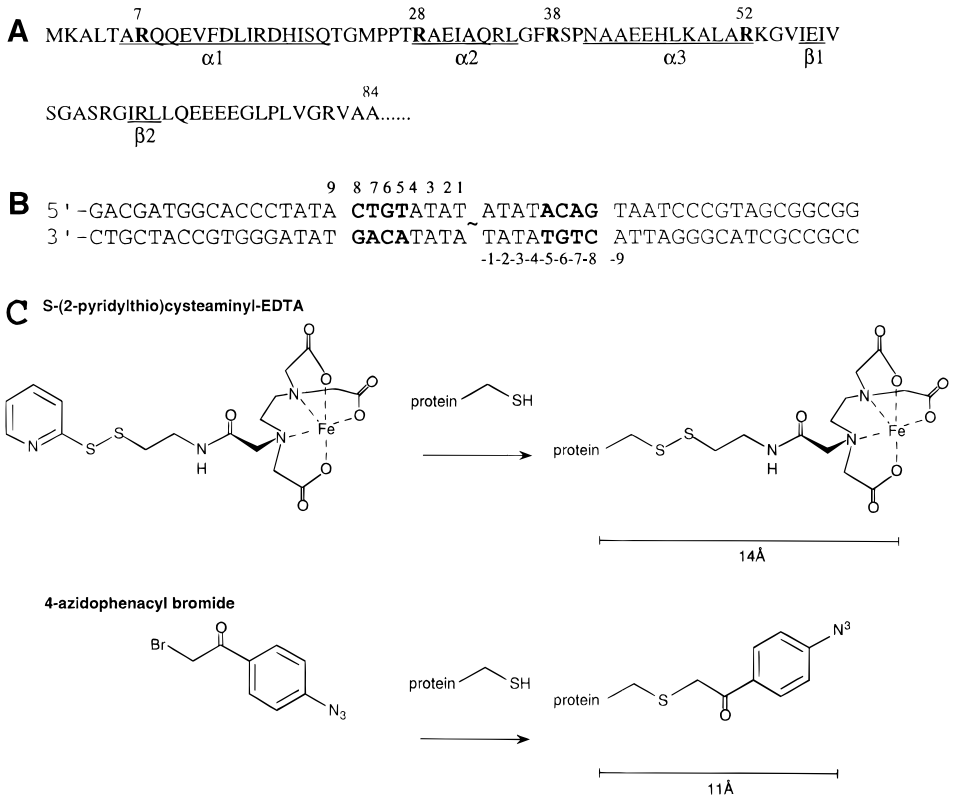

FIGURE 2: Panel A: Amino acid sequence of the LexA DNA binding domain. Panel B: Sequence of the DNA duplex (harboring a consensus SOS operator) used throughout this work. Panel C: Reaction of *S*-(2-pyridylthio)cysteaminyl-EDTA or 4-azidophenacyl bromide with a protein having a unique cysteine residue. If fully extended, the linker arm between the α carbon of the cysteine residue and the metal ion is 14 Å in the case of a [*S*-(2-pyridylthio)cysteaminyl-EDTA−metal−Cys] protein and 11 Å in the case of a [(4-azidophenacyl)-Cys] protein.

Assuming that helix 3 is the recognition helix (Oertel-Buchheit et al., 1990; Thliveris & Mount, 1992; Dumoulin et al., 1993), all four arginine residues shown in Figure 1 are likely to be on the contact surface between LexA and DNA. We have chosen to substitute arginines because arginine has a long and flexible side chain. We anticipated that the addition of a cleaving or cross-linking agent in these positions is less likely to perturb complex formation due to steric hindrance.

In Vivo Repressor Activity of Mutant LexA Derivatives Containing Unique Cysteines. The four mutant repressors LexA-C7, LexA-C28, LexA-C38, and LexA-C52 were tested for their *in vivo* repressor activity using the *E. coli* reporter strain JL806 (Little & Hill, 1985). This strain has a chromosomal *lexA* (Def) mutation and a *lacZ* gene (coding for β-galactosidase) under the control of the LexA-regulated *recA* promoter. JL806 allows thus to determine the repressor activity of plasmid-encoded LexA variants. The data in Table 1 show that all four LexA mutant derivatives retain substantial *in vivo* repressor activity.

In Vitro Affinity Cleavage and Photo-Cross-Linking Experiments. The four LexA derivatives were purified using standard *lacUV5* promoter driven expression for LexA-C38 and LexA-C52 (Schnarr et al., 1985). LexA-C7 and LexA-C28 were purified after expression under the control of a T7 promoter system (Dubendorff & Studier, 1991; Roland et al., 1992). All experiments were done with a synthetic 50 bp DNA duplex (see Figure 2B for the sequence and the numbering scheme) harboring a symmetric SOS consensus operator CTGT(AT)$_4$ACAG.

Table 1: *In Vivo* Repression Efficiency of LexA Mutant Repressors Containing Unique Cysteine Residues in Positions 7, 28, 38, or 52

	β-galactosidase units[a]	% repression[b]
LexA-C7	450	96
LexA-C28	1300	88
LexA-C38	300	97
LexA-C52	790	93
LexA wild type	280	97
no repression	11000	0

[a] β-Galactosidase units were determined in the reporter strain JL806 at 10^{-3} M IPTG. [b] The degree of repression was determined according to the equation $(1 - \text{units}^{+\text{repressor}}/\text{units}^{-\text{repressor}}) \times 100$.

EDTA·Iron-Mediated Affinity Cleavage. In EDTA·iron-mediated affinity cleaving (EIMAC),[1] one incorporates EDTA·iron at a single amino acid, *x*, not critical for protein−DNA complex formation but nevertheless close to DNA in the protein−DNA complex, one forms the derivatized protein−DNA complex, one initiates EDTA·iron-mediated affinity cleavage by addition of reducing agent, and one determines the nucleotide(s) at which EDTA·iron-mediated affinity cleavage occurs [reviewed in Dervan (1991)]. The results identify nucleotide(s) close to amino acid *x* in the protein−DNA complex. EDTA·iron-mediated affinity cleavage occurs through generation of diffusible hydroxyl radicals, followed by hydroxyl radical-mediated abstraction of the deoxyribose C1′ hydrogen atom or the C4′ hydrogen atom (Tullius et al., 1987; Dervan, 1991).

We prepared LexA−EDTA·iron conjugates by reaction of mutant LexA derivatives containing unique cysteines at

[1] Abbreviations: EIMAC, EDTA·iron-mediated affinity cleavage; PAMP, phenyl azide-mediated photo-cross-linking.

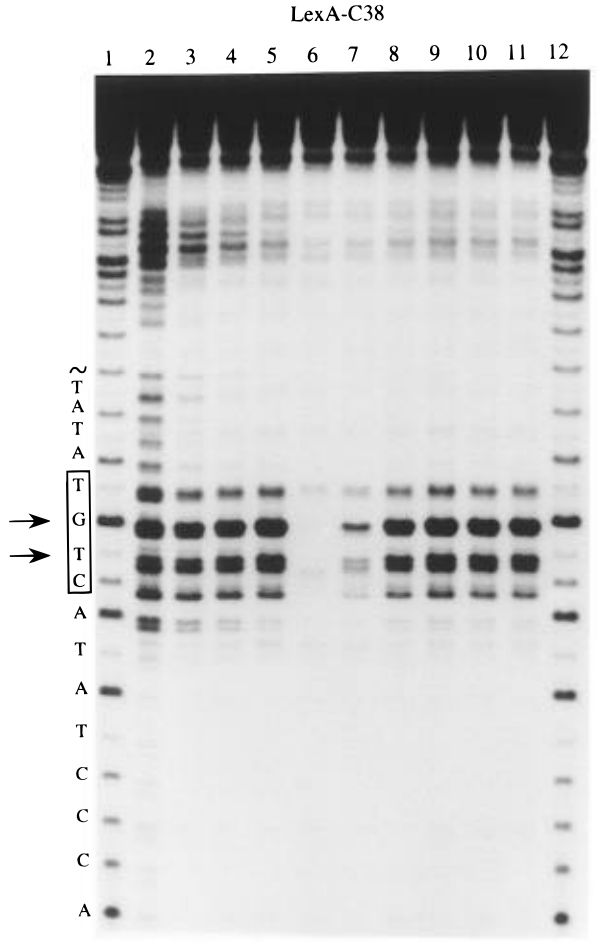

FIGURE 3: EDTA·iron-mediated affinity cleavage with the LexA-C38 derivative. Glycerol enhances the resolution of EDTA·iron-mediated affinity cleavage as shown here in the case of $LexA_{38}$−EDTA·iron: lane 2 (no glycerol), lane 3 (10% glycerol), lane 4 (25% glycerol), and lane 5 (40% glycerol). The two major cleavage products are indicated by arrows. Specific cleavage products disappear upon addition of unmodified LexA wild-type repressor: lane 11, no competing LexA repressor; lane 10, 10^{-10} M LexA; lane 9, 10^{-9} M LexA; lane 8, 10^{-8} M LexA; lane 7, 10^{-7} M LexA; and lane 6, 10^{-6} M LexA. Lanes 1 and 12 show G+A sequencing products according to Maxam and Gilbert (1980).

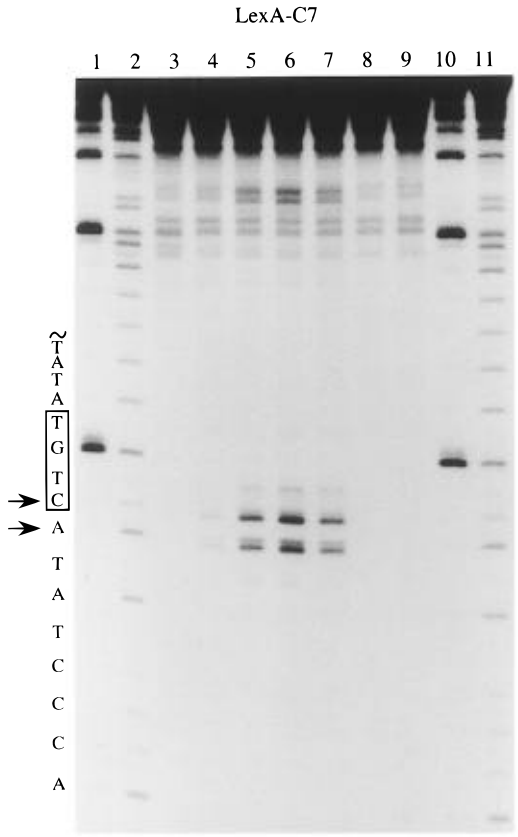

FIGURE 4: EDTA·iron-mediated affinity cleavage with the LexA-C7 derivative gives rise to two predominant cleavage products respectively within and 5′ to the CTGT recognition box (see arrows). These cleavage products disappear upon addition of unmodified LexA wild-type repressor: lane 7, no competing LexA repressor; lane 6, 10^{-9} M LexA; lane 5, 10^{-8} M LexA; lane 4, 10^{-7} M LexA; and lane 3, 10^{-6} M LexA. Lane 8 is a control corresponding to piperidine/heat treatment of DNA only; lane 9 is a control with unmodified LexA-C7 repressor. Lanes 2 and 11 show G+A sequencing products and lanes 1 and 10 G>A sequencing products.

position 7, 28, 38, or 52 with *S*-(2-pyridylthio)cysteaminyl-EDTA (Figure 2C; Ebright et al., 1992; Ermacora et al., 1992). We then analyzed DNA cleavage in the presence of a high concentration of glycerol (a known hydroxyl radical scavenger; Tullius et al., 1987). As documented in Figure 3 (lanes 2−5) this strategy appears to improve resolution. In the absence of glycerol (lane 2) cleavage is observed for at least 19 nucleotides. Upon addition of increasing amounts of glycerol the intensity of the principal cleavage products is essentially unchanged whereas the intensity of the other cleavage products is progressively diminished.

We find that three of the four LexA−EDTA·iron conjugates yield significant site-specific cleavage under our conditions: i.e., $LexA_7$−EDTA·iron, $LexA_{38}$−EDTA·iron, and $LexA_{52}$−EDTA·iron.

The data for $LexA_7$−EDTA·iron are presented in Figure 4. The main cleavage products are observed at nucleotides C8 and A9 (i.e., the cytosine of the CTGT motif and the preceding A residue). On top of the gel four additional minor bands are observed. However, these bands are also present in the control lanes (8 and 9) albeit for two of them with lower intensity.

In order to distinguish between possible nonspecific cleavage and signals due to the formation of the specific protein−DNA complex, the cleavage reaction was done in the presence of increasing amounts of LexA wild-type repressor. Only those bands which disappear upon addition of LexA will be considered as specific cleavage products. At 10^{-7} M wild-type LexA (lane 4 in Figure 4) the bands corresponding to the A9 and C8 cleavage products have essentially disappeared, and the intensities of two of the minor bands on top of the gel are diminished. Very similar results are obtained upon ^{32}P-labeling of the bottom strand.

Given the relative intensities of the cleavage products, the EDTA moiety of $LexA_7$−EDTA·iron should be close to the deoxyribose of A9 and C8. Two features suggest that the EDTA moiety approaches these two sugars from the minor groove. First, cleavage by $LexA_7$−EDTA·iron essentially gives rise to the 3′-phosphoglycolate species which migrates somewhat more rapidly than a DNA fragment harboring a 3′-phosphate terminus (Hertzberg & Dervan, 1984). The 3′-phosphoglycolate species is most likely generated by hydrogen abstraction from the 4′-deoxyribose position (Knapp-Pogozelski & Tullius, 1993). The 4′-deoxyribose hydrogen points into the minor groove. Second, the two minor cleavage products on the opposite strand are shifted to the 3′ side with respect to the cleavage positions in the top strand (see Figure 8 for a compilation of the cleavage and cross-

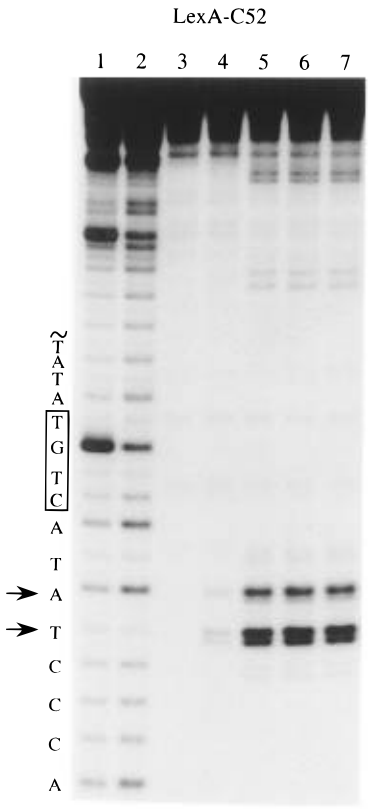

FIGURE 5: EDTA·iron-mediated affinity cleavage with the LexA-C52 derivative gives rise to two predominant cleavage products respectively 3 and 4 bases 5′ to the CTGT recognition box (see arrows). These cleavage products disappear upon addition of unmodified LexA wild-type repressor: lane 7, no competing LexA repressor; lane 6, 10^{-9} M LexA; lane 5, 10^{-8} M LexA; lane 4, 10^{-7} M LexA; and lane 3, 10^{-6} M LexA. Lane 1 shows G>A sequencing products and lane 2 G+A sequencing products.

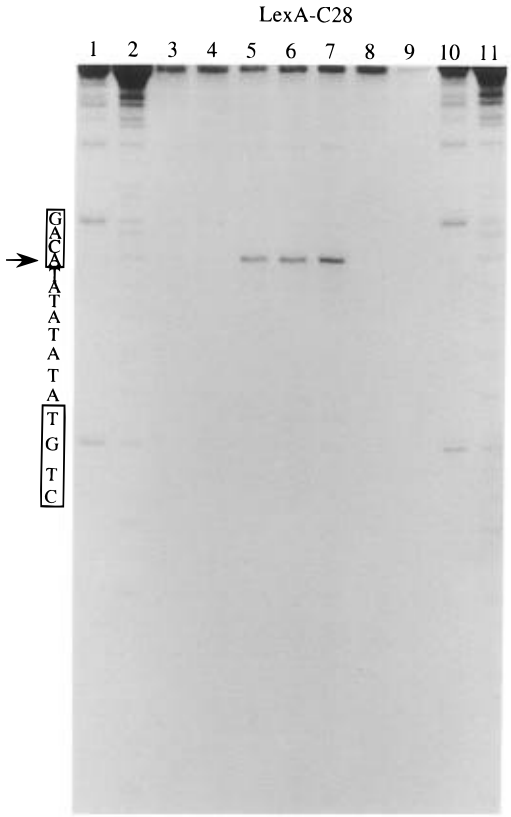

FIGURE 6: Phenyl azide-mediated photo-cross-linking with the LexA-C28 derivative reveals a single predominant photo-cross-linking product at adenine A5 within the LexA recognition motif (see arrow). This cross-linking product disappears upon addition of increasing amounts of unmodified LexA wild-type repressor: lane 7, no competing LexA repressor; lane 6, 10^{-9} M LexA, lane 5, 10^{-8} M LexA; lane 4, 10^{-7} M LexA; and lane 3, 10^{-6} M LexA. Lane 8 is a control with unmodified LexA-C28 repressor; lane 9 is a control corresponding to piperidine/heat treatment of DNA only. Lanes 2 and 11 show G+A sequencing products and lanes 1 and 10 G>A sequencing products.

linking results). This pattern is characteristic for an EDTA·iron located in the minor groove (Oakley & Dervan, 1990).

The data for $LexA_{38}$−EDTA·iron are presented in Figure 3. The two major cleavage products of this derivative are observed at nucleotides G6 and T7 within the $\mathbf{C_8T_7G_6T_5}$ motif (see lane 11 in Figure 3). Additional bands are observed for the two flanking nucleotides 8 and 5. The main cleavage 3′-terminal ends change within the motif: mainly or exclusively 3′-phosphate cleavage products are observed at T5 and G6, about equal amounts of 3′-phosphate and 3′-phosphoglycolte at T7, and mainly 3′-phosphoglycolate at C8 (the repartition of the cleavage products is most clearly seen in the presence of 10^{-7} M competing unmodified LexA repressor where the cleavage intensity is lower).

$LexA_{38}$−EDTA·iron exhibits a rather large cleavage pattern including four adjacent nucleotides even in the presence of 40% glycerol. This differs from the $LexA_7$−EDTA·iron (Figure 4) and $LexA_{52}$−EDTA·iron (Figure 5) conjugates which cleave only at two adjacent nucleotides. Most likely the derivatized side chain at position 38 has a higher degree of mobility than the derivatives at positions 7 and 52. This may be due to the fact that amino acid 38 is located in a turn (between helix 2 and helix 3), whereas amino acids 7 and 52 are respectively part of helix 1 and helix 3.

The data for $LexA_{52}$−EDTA·iron are presented in Figure 5. Cleavage occurs principally at nucleotides A11 and T12, i.e., respectively three and four base pairs upstream of the CTGT motif. Minor cleavage products are observed at nucleotides 14 and 15 on the opposite strand. At A11, cleavage gives rise to 3′-phosphate end groups, whereas at T12 both 3′-phosphate and 3′-phosphoglycolate are observed. Positions 11 and 12 of the operator are not involved in site-specific recognition (there is no detectable sequence conservation in this region). However, the nature of the bases in these positions modulates the gel mobility of a LexA−DNA complex, suggesting that the DNA segment around positions 11 and 12 is bent (Lloubés et al., 1993). Earlier affinity cross-linking experiments with the LexA-C52 mutant repressor gave rise to photo-cross-linking at positions 10 and 11 (Dumoulin et al., 1993), i.e., one base pair closer to the CTGT motif. This difference is consistent with the fact that the extended linker of the EDTA·iron adduct is about 3 Å longer than the linker of the photo-cross-linking adduct (see Figure 2C).

Phenyl Azide-Mediated Photo-Cross-Linking. In phenyl azide-mediated photo-cross-linking (PAMP), one incorporates a phenyl azide photoactivatable cross-linking agent at a single amino acid, *x*, not critical for protein−DNA complex formation but nevertheless close to DNA in the protein−DNA complex, one forms the derivatized protein−DNA complex, one initiates phenyl azide-mediated protein−DNA cross-linking by UV irradiation, and one determines the nucleotide(s) at which phenyl azide-mediated protein-DNA cross-linking occurs (Pendergrast et al., 1992; Chen & Ebright, 1993; Dumoulin et al., 1993). The results identify

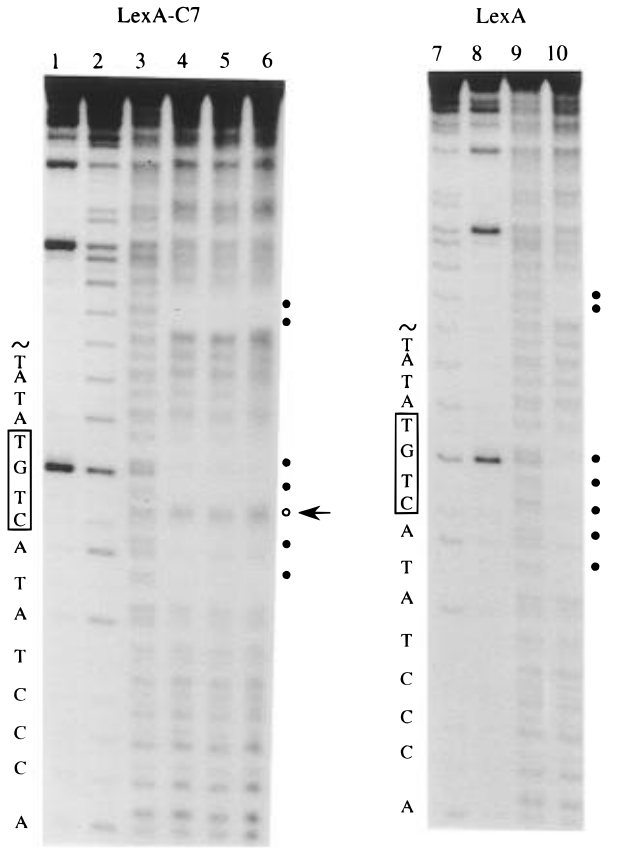

FIGURE 7: Ethylation interference experiments with the LexA-C7 mutant repressor and the LexA wild-type repressor. Lane 3 contains no protein; lane 4, 2×10^{-6} M LexA-C7; lane 5, 10^{-6} M LexA-C7; lane 6, 5×10^{-7} M LexA-C7; lane 9, no protein; lane 10, 10^{-7} M LexA wild-type repressor [for ethylation interference with the wild-type repressor, see also Hurstel et al. (1988)]. Lanes 2 and 7 show G+A sequencing products and lanes 1 and 8 G>A sequencing products.

nucleotide(s) close to amino acid *x* is the protein−DNA complex.

PAMP occurs through photogeneration of a phenylnitrene, followed by rearrangement of the phenylnitrene to an electrophilic species (probably a dehydroazepine) and reaction with a DNA nucleophile (Chen & Ebright, 1993). Cross-link formation requires direct contact between the photogenerated reactive species and its target.

We prepared LexA−azidophenacyl conjugates by reaction of mutant LexA derivatives containing unique cysteines at position 7, 28, or 38 with azidophenacyl bromide (Figure 2C; Pendergrast et al., 1992; Chen & Ebright, 1993; Dumoulin et al., 1993). We then analyzed protein−DNA photo-cross-linking by the resulting LexA−azidophenacyl conjugates.

We find that one of the three tested LexA−azidophenacyl conjugates yielded significant site-specific protein−DNA photo-cross-linking, i.e., $LexA_{28}$−azidophenacyl (Figure 6). Cross-linking occurs at the adenine in base pair 5. This base pair is part of the CTGT recognition box. The cross-link disappears upon addition of 10^{-7} M LexA wild-type repressor, similar to what we observe in the case of the EIMAC competition experiments.

Ethylation Interference. The wild-type LexA repressor shows ethylation interference at seven phosphate positions in each operator half-site (numbered from P1 to P7 in Figure 8) (Hurstel et al., 1988). The two principal cleavage products of $LexA_7$−EDTA·iron are situated on both sides of the

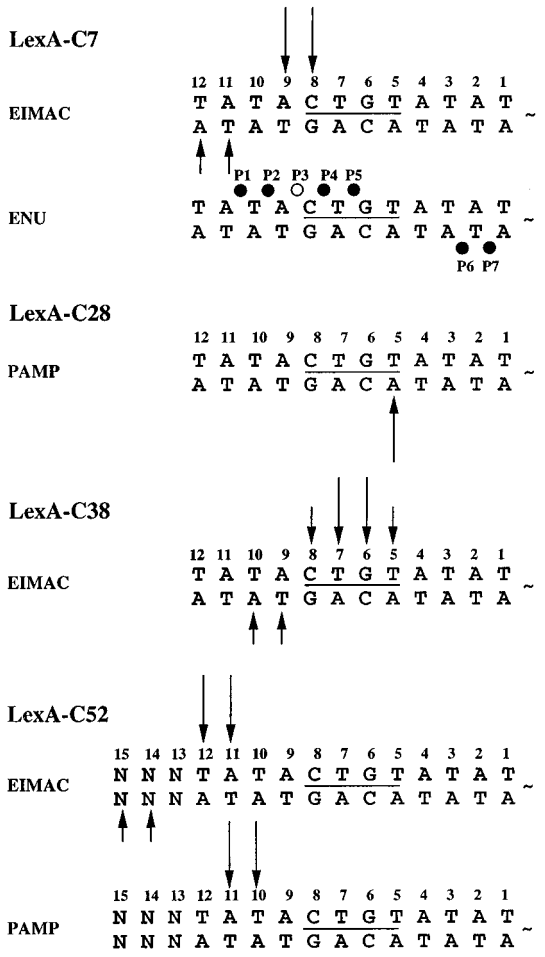

FIGURE 8: Compilation of EIMAC, PAMP, and ethylation interference data for LexA-C7, LexA-C28, LexA-C38, and LexA-C52.

Table 2: Comparison between Biochemical Distance Constraints and Corresponding Distances in the Model of Knegtel et al. (1995)

C_α	C1′	distance range (Å) (EIMAC)	distance (Å) (model)	strand
Arg-7	C8	7−16	12.26	top
	A9	7−16	9.00	top
	T11	7−16	12.22	bottom
	A12	7−16	9.77	bottom
Arg-38	T5	7−16	16.10	top
	G6	7−16	10.86	top
	T7	7−16	8.59	top
	C8	7−16	7.33	top
	T9	7−16	14.82	bottom
	A10	7−16	12.35	bottom
Arg-52	A11	7−16	14.73	top
	T12	7−16	17.32	top
	A14	7−16	21.20	bottom
	G15	7−16	24.48	bottom

C_α	base	distance range (Å) (PAMP)	distance (Å) (model)	strand
Arg-28	A5	9−12	10.29 (N7) 11.78 (N6)	bottom

interfering phosphate P3. It was thus tempting to speculate that the arginine side chain in position 7 of the wild-type repressor might be responsible interference at P3.

This is indeed the case, since DNA binding of the LexA-C7 mutant repressor (harboring a cysteine instead of arginine in position 7) is not suppressed upon P3 ethylation (Figure 7A), whereas all other interfering phosphates are conserved as compared to the wild-type repressor (Figure 7B). This result strongly suggests that Arg-7 of LexA contacts the

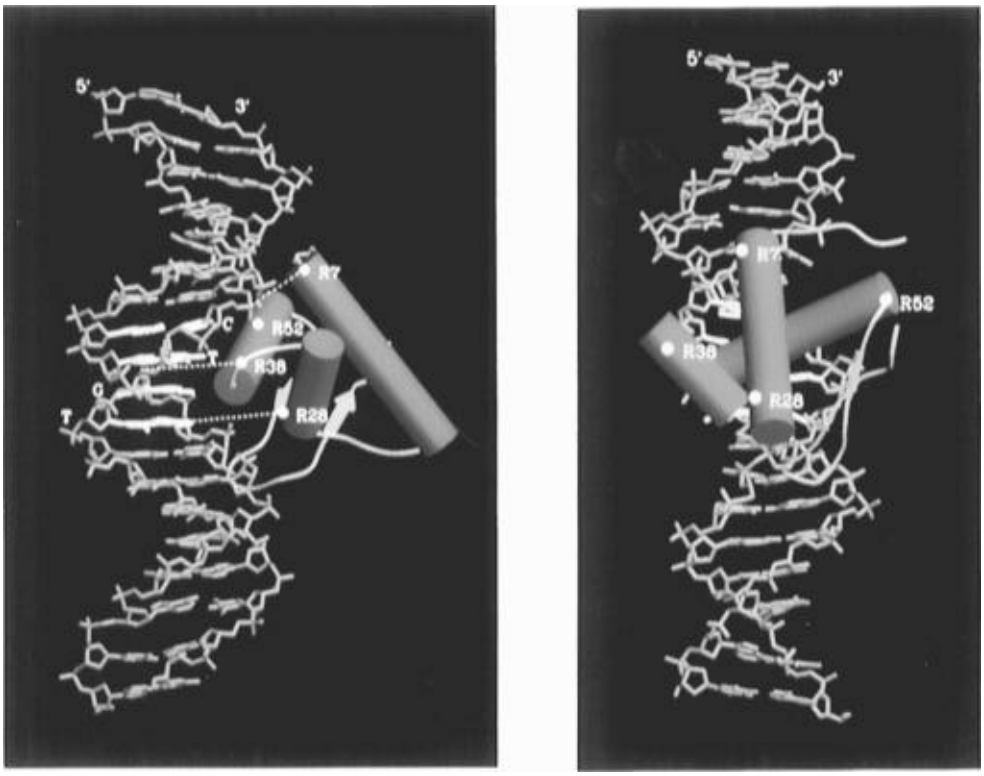

FIGURE 9: Two perpendicular views of the complex between the LexA DNA binding domain and a half-operator [according to Knegtel et al. (1995)]. The four arginine residues which have been mutated to cysteine are labeled. The principal biochemical constraints from the present work are shown as stippled lines.

phosphate group P3 located 5′ to the CTGT recognition motif. This contact has been predicted in the model of Knegtel et al. (1995).

DISCUSSION

The LexA−DNA complex has been so far recalcitrant to structural investigation by both NMR and X-ray crystallography. Here we use affinity cleavage and affinity photocross-linking to probe the structure of the complex and to compare these biochemical data with a model of the LexA−DNA complex derived from Monte Carlo docking and energy minimization (Knegtel et al., 1995).

The results presented here unambiguously define the orientation of the LexA DNA binding domain relative to the operator and provide support for the principal features of the model of the LexA−operator complex proposed by Knegtel et al. (1995). The data obtained with the $LexA_{52}$−EDTA·iron derivative suggest, however, that the DNA may be more strongly bent than anticipated by the model.

In the past EIMAC and PAMP analyses have been carried out for two protein−DNA complexes for which crystallographic structures are available, i.e., the CAP−DNA complex and the cro−DNA complex (Ebright et al., 1992, 1993; Pendergrast et al., 1992; Chen & Ebright, 1993). Comparison of the EIMAC and PAMP results to the crystallographic structures indicates that (i) efficient DNA cleavage using the EDTA·iron derivative used in this work requires that C_α of the derivatized amino acid be 7−16 Å (mean of 12 examples = 13 Å) from the C1′ atom of the target nucleotide and (ii) efficient protein−DNA cross-linking using the phenyl azide derivative used in this work requires that C_α of the derivatized amino acid be 9−12 Å (mean of 5 examples = 11 Å) from a base O or N atom of the target nucleotide (Ebright et al., 1992; Chen & Ebright, 1993).

We infer that in the case of the LexA−DNA complex the distances between the C_α atoms of residues 7, 38, and 52 and the C1′ atoms of the target nucleotides should be in the range of 7−16 Å as based on EIMAC experiments and that the distance between the C_α atom of residue 28 and its target nucleotide should be in the range of 9−12 Å as based on PAMP experiments.

Table 2 shows a detailed comparison of these biochemical distance constraints with the corresponding distances of the model developed by Knegtel et al. (1995). Figure 9 shows two perpendicular views of this model. The links between the derivatized side chains and their principal target nucleotides (this work) are shown as stippled lines. In agreement with earlier experimental work (Oertel-Buchheit et al., 1990; Thliveris & Mount, 1992; Dumoulin et al., 1993) the model predicts that DNA recognition occurs mostly via helix 3 and that the HTH motif is positioned such that the amino-terminal end of helix 2 is oriented toward the dyad axis of the operator. Additional nonspecific protein−DNA contacts are predicted to be formed via the amino-terminal extremities of helix 1 and 2 and the loop between the two antiparallel β strands.

In the model, all distances between the Arg-7 C_α atom and the C1′ sugar atoms of the target nucleotides are within the range of 9.0−12.3 Å. These distances are in excellent agreement with the affinity cleavage experiments. In the model, Arg-7 approaches the phosphate between C8 and A9 from the minor groove, as was suggested by the EDTA-induced cleavage experiments, and its side chain forms a hydrogen bond to this phosphate in agreement with the ethylation interference pattern of the LexA-C7 mutant repressor.

In the model, the C_α atom of Arg-28 is at 10.3 Å from the N7 atom of target nucleotide A5, i.e., close to the average cross-linking distance reported by Chen and Ebright (1993).

The model can thus account for the observed cross-linking at the A5 base.

The $LexA_{38}$−EDTA·iron derivative displays a fairly large region of cleavage centered around G6 and T7, which is likely to be due to its placement in a flexible turn of the LexA DBD. In the model, all distances between the Arg-38 C_α atom and the C1′ atoms of the target nucleotides are within 7.3−16.1 Å, in good agreement with the results of the EDTA cleavage experiments.

The $LexA_{52}$−EDTA·iron derivative cleaves the DNA at positions 11, 12, 14, and 15, i.e., outside the CTGT recognition motif (base pairs 5−8). Lloubés et al. (1993) have shown that his region is involved in LexA-induced DNA bending. Although the MONTY program used for docking allows for some DNA bending, the extent of bending that could be introduced in the Monte Carlo docking simulations is limited (Knegtel et al., 1995). Table 2 shows that in particular A14 (21.2 Å) and G15 (24.5 Å) are too far apart from the Arg-52 C_α atom, suggesting that the DNA in this region may be more strongly bent toward the protein than predicted by the model.

Despite this bending problem, the overall agreement of the biochemical data reported here and the Knegtel et al. (1995) model is satisfying. The biochemical data confirm the orientation of the LexA DBD on the DNA and a particular nonspecific contact between Arg-7 and the interfering phosphate P3. This general agreement adds to the credibility of the specific protein−DNA contacts that were predicted by the model, involving in particular Ser-39, Asn-41, Ala-42, Glu-44, and Glu-45.

ACKNOWLEDGMENT

We are grateful to Dr. Y. W. Ebright for the synthesis of *S*-(2-pyridylthio)cysteaminyl-EDTA, to Dr. J. Reinbolt for determining the degree of modification of the LexA derivatives, and to Dr. J. W. Little for some of the strains and plasmids used in this work.

REFERENCES

Baumeister, R., Helbl, V., & Hillen, W. (1992) *J. Mol. Biol. 226*, 1257−1270.
Boelens, R., Scheek, R. M., van Boom, J. H., & Kaptein, R. (1987) *J. Mol. Biol. 193*, 213−216.
Brennan, R. G. (1993) *Cell 74*, 773−776.
Chen, C.-H. B., & Sigman, D. S. (1987) *Science 237*, 1197−1201.
Chen, Y., & Ebright, R. H. (1993) *J. Mol. Biol. 230*, 453−460.
Cheo, D. L., Bayles, K. W., & Yasbin, R. E. (1993) *J. Bacteriol. 175*, 5907−5915.
Clark, K. L., Halay, E. D., Lai, E., & Burley, S. K. (1993) *Nature 364*, 412−420.
Dervan, P. (1991) *Methods Enzymol. 208*, 497−515.
Dubendorff, J. W., & Studier, F. W. (1991) *J. Mol. Biol. 219*, 45−59.
Dumoulin, P., Oertel-Buchheit, P., Granger-Schnarr, M., & Schnarr, M. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 2030−2034.
Ebright, R. H., Ebright, Y. W., Pendergrast, P. S., & Gunasekera, A. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 2882−2886.
Ebright, Y. W., Chen, Y., Pendergrast, S., & Ebright, R. H. (1992) *Biochemistry 31*, 10664−10670.
Ebright, Y. W., Chen, Y., Ludescher, R. D., & Ebright, R. H. (1993) *Bioconjugate Chem. 4*, 219−225.
Ermacora, M., Delfino, J., Cuenoud, B., Schepartz, A., & Fox, R. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6383−6387.
Fogh, R. H., Ottleben, G., Rüterjans, H., Schnarr, M., Boelens, R., & Kaptein, R. (1994) *EMBO J. 13*, 3936−3944.
Friedman, M., Krull, L. H., & Cavins, J. F. (1970) *J. Biol. Chem. 245*, 3868−3871.
Garriga, X., Calero, S., & Barbé, J. (1992) *Mol. Gen. Genet. 236*, 125−134.
Harrison, C. J., Bohm, A. A., & Nelson, H. C. M. (1994) *Science 263*, 224−227.
Hertzberg, R. P. & Dervan, P. B. (1984) *Biochemistry 23*, 3934−3945.
Holm, L., Sander, C., Rüterjans, H., Schnarr, M., Fogh, R., Boelens, R., & Kaptein, R. (1994) *Prot. Eng. 7*, 1449−1453.
Hurstel, S., Granger-Schnarr, M., Daune, M., & Schnarr, M. (1986) *EMBO J. 5*, 793−798.
Hurstel, S., Granger-Schnarr, M., & Schnarr, M. (1988) *EMBO J. 7*, 269−275.
Kim, B., & Little, J. W. (1992) *Science 255*, 203−206.
Knapp-Pogozelski, W., & Tullius, T. D. (1993) *J. Biomol. Struct. Dyn. 10*, a153.
Knegtel, R. M. A., Fogh, R. H., Ottleben, G., Rüterjans, H., Dumoulin, P., Schnarr, M., Boelens, R., & Kaptein, R. (1995) *Proteins 21*, 226−236.
Kraulis, J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
Lehming, N., Sartorius, J., Kisters-Woike, B., von Wilcken-Bergmann, B., & Müller-Hill, B. (1990) *EMBO J. 9*, 615−621.
Little, J. W. (1991) *Biochimie 73*, 411−422.
Little, J. W., & Hill, S. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 2301−2305.
Lloubés, R., Lazdunski, C., Granger-Schnarr, M., & Schnarr, M. (1993) *Nucleic Acids Res. 21*, 2363−2367.
Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499−560.
Oakley, M. G., & Dervan, P. B. (1990) *Science 248*, 847−850.
Oertel-Buchheit, P., Lamerichs, R. M. J. N., Schnarr, M., & Granger-Schnarr (1990) *Mol. Gen. Genet. 223*, 40−48.
Ogata, K., Hojo, H., Aimoto, S., Nakai, T., Nakamura, H., Sarai, A., Ishii, S., & Nishimura, Y. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 6428−6432.
Ogata, K., Morikawa, S., Nakamura, H., Sekikawa, A., Inoue, T., Kanai, H., Sarai, A., Ishii, S., & Nishimura, Y. (1994) *Cell 79*, 639−648.
Pendergrast, P. S., Chen, Y., Ebright, Y. W., & Ebright, R. H. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10287−10291.
Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L., & Sweet, R. M. (1993) *Nature 362*, 219−223.
Raymond-Denise, A., & Guillen, N. (1991) *J. Bacteriol. 173*, 7084−7091.
Roland, K. L., Smith, M. H., Rupley, J. A., & Little, J. W. (1992) *J. Mol. Biol. 228*, 395−408.
Sayers, J. R., Schmidt, W., & Eckstein, F. (1988a) *Nucleic Acids Res. 16*, 791−802.
Sayers, J. R., Schmidt, W., Wendler, A., & Eckstein, F. (1988b) *Nucleic Acids Res. 16*, 803−814.
Schnarr, M., & Granger-Schnarr, M. (1993) In *Nucleic Acids and Molecular Biology* (Eckstein, F., & Lilley, D. M. J., Eds.) Vol. 7, pp 170−189, Springer-Verlag, Berlin.
Schnarr, M., Pouyet, J., Granger-Schnarr, M., & Daune, M. (1985) *Biochemistry 24*, 2812−2818.
Schnarr, M., Granger-Schnarr, M., Hurstel, S., & Pouyet, J. (1988) *FEBS Lett. 234*, 56−60.
Schnarr, M., Oertel-Buchheit, P., Kazmaier, M., & Granger-Schnarr, M. (1991) *Biochimie 73*, 423−431.
Schultz, S. C., Shields, G. C., & Steitz, T. A. (1991) *Science 25*, 1001−1007.
Shin, J. A., Ebright, R. H., & Dervan, P. B. (1991) *Nucleic Acids Res. 19*, 5233−5236.
Thliveris, A. T., & Mount, D. W. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 4500−4504.
Thliveris, A. T., Little, J. W., & Mount, D. W. (1991) *Biochimie 73*, 449−455.
Tullius, T. D., Dombroski, B. A., Churchill, M. E. A., & Kam, L. (1987) *Methods Enzymol. 155*, 537−558.
Walker, G. C. (1984) *Microbiol. Rev. 48*, 60−93.
Wertman, K. F., & Mount, D. W. (1985) *J. Bacteriol. 163*, 376−384.
Wilson, K. P., Shewchuk, L. M., Brennan, R. G., Otsuka, A. J., & Matthews, B. M. (1993) *Proc. Natl. Acad. Sci. U.S.A. 89*, 9257−9261.
Wissmann, A., Baumeister, R., Müller, G., Hecht, B., Helbl, V., Pfleiderer, K., & Hillen, W. (1991) *EMBO J. 10*, 4145−4152.

BI9529162