Mechanism of *Lac* repressor switch-off: orientation of the *Lac* repressor DNA-binding domain is reversed upon inducer binding

Dimitrii E. Kamashev[a,*], Natalia G. Esipova[a], Konstantin K. Ebralidse[b], A.D. Mirzabekov[a]

[a]*W.A. Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 32 Vavilov Street, 117984, Moscow B-334, Russian Republic*
[b]*Harvard University, Department of Biochemistry and Molecular Biology, 7 Divinity Avenue, Cambridge, MA 02138, USA*

Received 11 September 1995

Abstract ***Lac*** **repressor's DNA-binding domains contain helix-turn-helix motif which, though similar to those of phage λ *Cro* protein, are oriented differently with respect to DNA: in the specific complexes with *Lac* operator, N termini of the repressor's subunits are facing inwards. We demonstrate that, in the presence of an inducer, the repressor's N termini cross-link to the operator's outermost nucleotides. We suggest that the inducer fixes the repressor's DNA-binding domains in the Cro-type configuration and thus garbles its recognition surface. Since the Cro-type configuration is perfectly suitable for binding the DNA, this also explains how the switched-off repressor retains its non-specific DNA-binding.**

Key words: DNA recognition; *Lac* repressor; DNA–protein cross-linking; Helix-turn-helix motif

1. Introduction

E. coli's *lac* repressor regulates *lac* operon by recognizing and tightly binding the *lac* operator [1]. The affinity of the *lac* repressor for the *lac* operator is modulated by binding of inducer molecules, such as isopropyl-β-D-thiogalactopyranoside (IPTG). Binding of IPTG to the repressor decreases the repressor-operator binding constant from 10^{13} M^{-1} to 10^{10} M^{-1} without alteration of binding to non-specific DNA sequences (binding constant 10^{7} M^{-1}) [1,2]. So inducers can switch the repressor off by abolishing the recognition without influencing its non-specific interactions with DNA.

The N-terminal segment of the *lac* repressor (residues 1–51 or 1–59) forms a folded domain, 'headpiece', which can bind to DNA [3]. The three-dimensional structure of the *lac* repressor headpiece has been solved by NMR [4,5]. This structure contains the specific DNA-binding 'helix-turn-helix' motif, found in other DNA-binding proteins by X-ray studies [6] and predicted for the *lac* repressor [7]. The orientation of the *lac* repressor recognition helix relative to the operator was shown to be opposite [5,8] to that established for all other helix-turn-helix proteins for which the structure was determined by X-ray analysis, such as phage λ *cro* protein [9].

To study this peculiarity of *lac* repressor we have determined the residues close in *lac* repressor and DNA to each other in their complexes using DNA-protein zero-length cross-linking [10,11].

*Corresponding author. Fax: (7) (95) 135-1405.
E-mail: kamashev@imb.imb.ac.ru

2. Materials and methods

2.1. Reagents

The *lac* repressor was isolated from overproducing cells, kindly supplied by Prof. B. Mueller-Hill. *Lac* repressor headpiece (amino acids 1–51/59) was generated by limited tryptic digestion of the repressor [12] and purified by HPLC. Synthetic 'ideal' *lac* operator [13,14] (DNA International, USA), and poly[d(AT)] about 3000 bp long (Pharmacia, Sweden) were used.

2.2. DNA-protein cross-linking

Experimental procedures for covalent DNA-protein cross-linking were described in detail [10,11]. Depurination of DNA is a two-step reaction: (1) methylation with dimethyl sulfate in 0.3 M HEPES, pH 7.4, at 25°C for 30 min followed by removing dimethyl sulfate by DNA precipitation; (2) spontaneous elimination of methylated purines in 30 mM HEPES, pH 7.4, 100 mM NaCl, and 0.1 mM EDTA at 45°C for 16 h. Dimethyl sulfate concentration was 35 mM for 20-bp ideal *lac* operator, and 18 mM for poly[d(AT)].

2.3. Cross-linked amino acids identification

Synthetic trinucleotide pTpGpTp was fully depurinated by heating in 70% formic acid for 30 min at 70°C and freeze-dried twice. The *lac* repressor headpiece (8 mg) was incubated with 4 mg of completely depurinated trinucleotide in 50% acetonitrile, 10 mM pyridine–borane complex at 40°C for 2 h. The cross-links were additionally reduced with 10 mM sodium borohydride for 30 min at 25°C. The protein was precipitated with 20% trichloroacetic acid. The pellet was washed twice by ether, dried, redissolved in 100 mM NH_4CO_3, pH 8.0, 0.2% Triton X-100, 5 mM dithiothreitol and then digested with trypsin (1:25 w/w) at 37°C for 6 h and freeze-dried. The trypsin digest was redissolved and fractionated by reversed-phase HPLC (column Partisil ODS3) using the 0–80% acetonitrile gradient in Tris-trifluoroacetic acid, pH 7.0, with dual-wave length monitoring (210 and 260 nm) to discern peptides from nucleotide-peptides. The fractions were freeze-dried, dephosphorylated with bacterial alkaline phosphatase and rechromatographed under the same conditions. Fractions corresponding to the peaks that were eluted at higher acetonitrile concentrations were collected, the solvent was changed to 0.1% trifluoroacetic acid by HPLC, and nucleotide-peptides were sequenced by Edman degradation.

3. Results

DNA-protein cross-linking chemistry employed is mild random methylation-partial depurination of DNA, with ensuing covalent linking of either N-terminal α-amino, lysine ε-amino, or histidine imidazole groups in the vicinity of the aldehyde group of apurinic residues [10,11,15]. Micrococcal nuclease digestion of the covalent adducts leaves a short nucleotide tail which is tagged by ^{32}P. Trypsin digestion, gel-electrophoresis and autoradiography produce a nucleotide-peptide map of peptides labeled at the site of cross-linking.

Lac repressor with or without an inducer, isopropyl-β-D-thiogalactopyranoside (IPTG), was cross-linked to either poly-[d(AT)] or perfectly symmetric 20-base pair (bp) *lac* operator [13,14] as models of non-specific and specific DNA, respec-

SSDI 0014-5793(95)01153-6

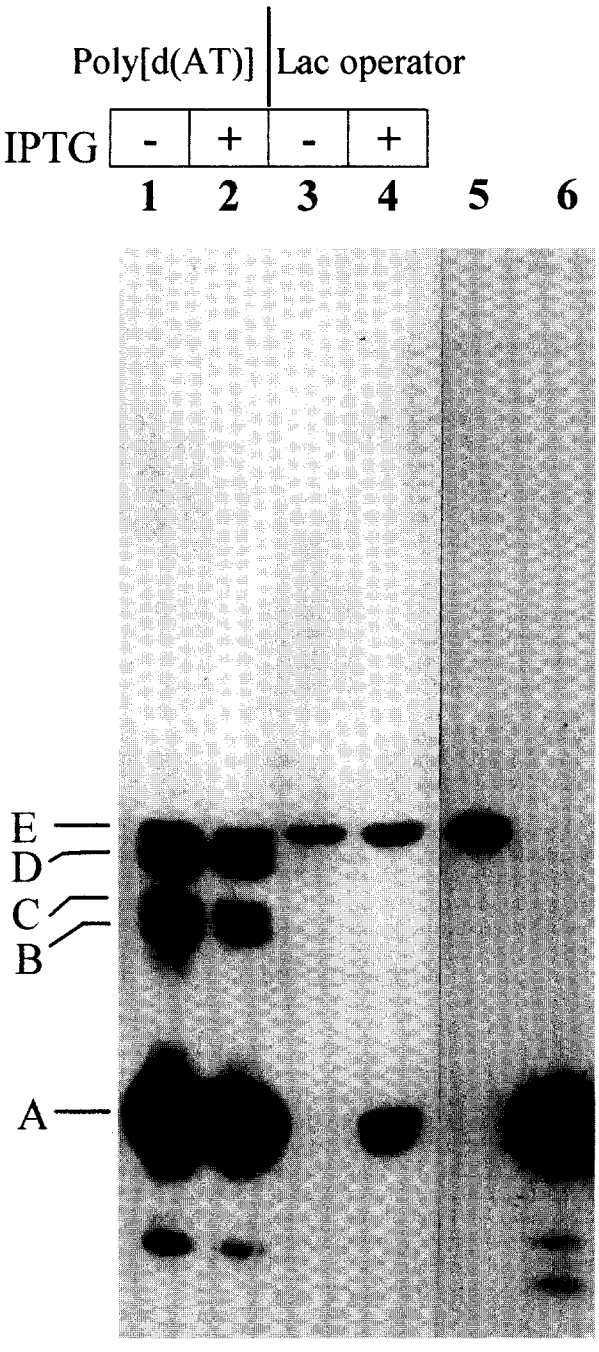

Fig. 1. Nucleotide-peptides separation. Lanes 1–4, *lac* repressor labeled at the site of cross-linking was digested with trypsin. Lanes 5 and 6, ^{32}P-labeled sequenced nucleotide-peptides 23–35 and 1–22, respectively.

tively. IPTG has no effect on non-specific repressor binding (lanes 1 and 2, Fig. 1), whereas the specific complex and operator-repressor-inducer complex differ from the above and from each other (lanes 3 and 4, Fig. 1).

Two major cross-links were identified directly by peptide sequencing. Bands A and E correspond to sequences XKPVTLYDVAEYAGVSYQTVSR (peptide 1–22 [16]) and VVNQASHVSAXTR (peptide 23–35), respectively. The actual sites of cross-linking (Xs) Met^1 of band A and Lys^{33} of band E were characteristically not detected because of a covalent modification by nucleotide residues(s).

Nucleotide-peptides B, C, and D, not obtainable in quantities sufficient for Edman degradation, were identified indirectly. Firstly, trypsin digestion of the shortest labeled cyanogen bromide-produced fragment of the cross-linked repressor generates nucleotide-peptides A, B, C, D, and E. This unambiguously assign this CNBr fragment to residues 1/2–42. Secondly, nucleotide-peptides A, B, C, and D, but not E, could be cleaved by *Staphylococcus aureus* endoprotease Glu-C (not shown). Now, the protein fragment 1/2–42 can only yield two triptic nucleotide-peptides, 1/2–22 and 23–35 (the cross-linked lysins are resistant to tryptic digestion). Only nucleotide-peptide 1/2–42 contains Glu and Asp residues [16]. Hence, nucleotide-peptides B, C, and D represent the cross-links via N-terminal α-amino group or Lys^2 ε-amino group.

Thus, the cross-links occur within the N-terminal domain of repressor, as expected from the known fact that the domain ('headpiece') is necessary and sufficient for the *lac* repressor DNA binding and recognition. In inducer-repressor-operator complex the *lac* repressor is linked to DNA almost exclusively through N-terminal amino acids (more than 90% of cases as judged by measurement of the relative radioactivity).

To identify the DNA site(s) in contact with the N termini, the repressor was cross-linked to a 5′-^{32}P-labeled perfectly symmetric 20-bp *lac* operator [13,14] in the presence of IPTG. After gel-electrophoresis (Fig. 2A) radioactive bands were cut out and digested with trypsin. The resultant DNA-linked peptides were separated by gel-electrophoresis (Fig. 2B) and eluted from the gel. DNA fragments labeled at 5′-terminus and bound to peptides (D1–D5) were subjected to A + G reaction of sequencing protocol [17]. The idea is that the sequencing ladder must

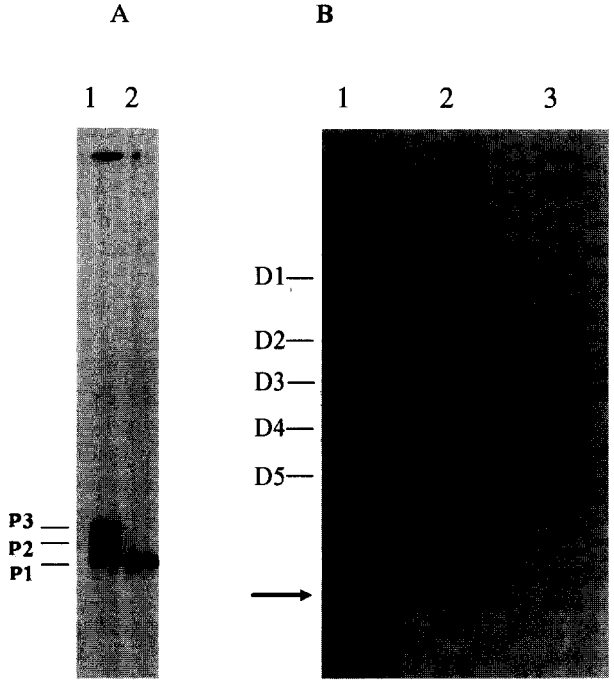

Fig. 2. Purification of the DNA-linked peptides. Panel A. Discontinuous gel-electrophoresis in the presence of urea. Lane 1, *lac* repressor cross-linked to 5′-labeled ideal *lac* operator. Lane 2, *lac* repressor cross-linked to the *lac* operator, digested with micrococcal nuclease, and ^{32}P-labeled at the site of cross-linking. Panel B. Bands P1–P3 indicated on the panel A were digested with trypsin, loaded onto lanes 1–3, respectively, and gel-electrophoresed at the presence of urea. The arrow indicates the position of the *lac* operator.

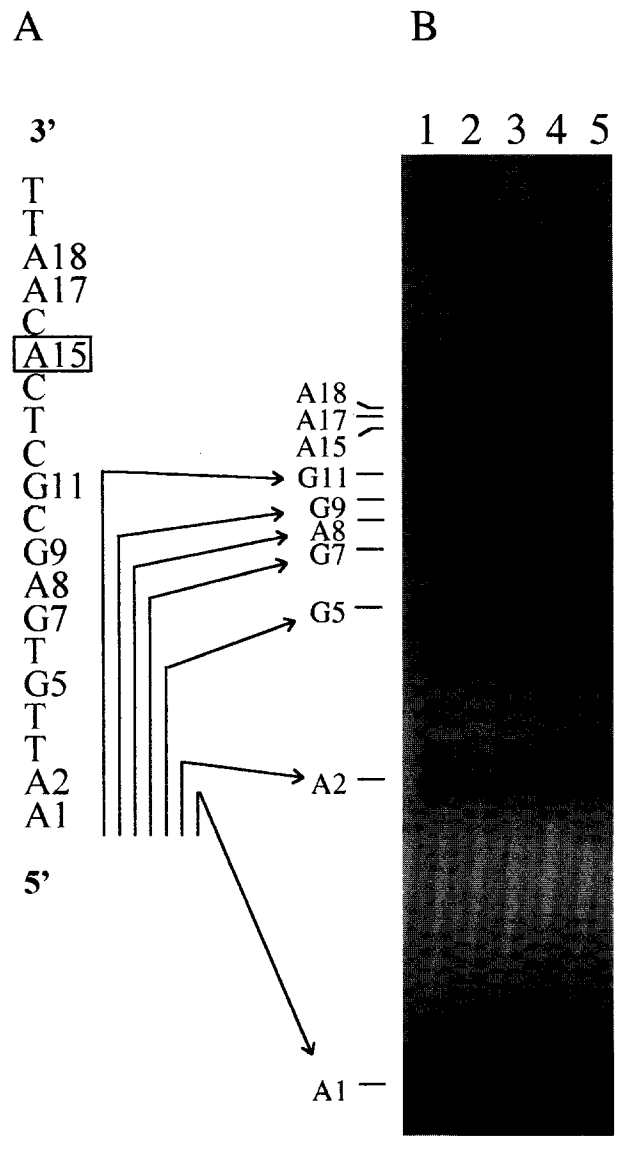

Fig. 3. Panel A. Ideal 20-bp *lac* operator cross-linked to *lac* repressor through A15, vertical lines indicate the DNA fragments revealed after the 'A + G' reaction. Panel B. Purified DNA-linked peptides (D1–D5 on the Fig. 2) were fragmentated with 2% diphenylamine and 70% formic acid ('A + G' reaction [17]) and gel-electrophoresed. Fragmentated DNA-linked peptides D1–D5, were loaded onto lanes 1–5, respectively.

be interrupted at and above the position of cross-linked purine because of the retardation by a covalently attached peptide moiety. The ladder is interrupted (Fig. 3B) at A15 (lanes 1, 2 and 3), and at A18 (lane 4). Thus, within the inducer-*lac* repressor-*lac* operator complex, the respressor's N-termini cross-link to purines A15 and A18.

4. Discussion

Within the specific repressor-operator complex only Lys^{33} is cross-linkable to DNA, while the *lac* repressor N-terminus is not. This is in agreement with NMR data [5,18,19].

In all non-specific complexes studied (with poly[d(AT)] and with 49-bp DNA without operator sequence (not shown) either in the presence and in the absence of an inducer) the *lac* repressor N-terminus is the major product of cross-linking. Similarly, mainly N-termini were cross-linked at the presence of inducer to the *lac* operator and the attachement was located at the operator boundaries. NMR studies showed that the *lac* repressor N-terminus resides close to the *lac* operator axis of symmetry, but does not make the direct contact with the operator [5,18–20] (Fig. 4A). Three α-helixes of the *lac* repressor headpiece are rigidly oriented relative to each other owing the contacts between nonpolar residues [4,21]. Therefore the drawing of the repressor N-terminus close to the operator boundaries can be explained by the following way: when inducer is bound to repressor, its DNA-binding domains have orientation relatively to operator (and relatively to each other) like λ *cro* protein [9] (Fig. 4B). Repressor's headpieces have the same orientation when interact with non-specific DNA. This reorientation can take place through the mutual arrangement of repressor's subunits or rotation of the DNA-binding domains relatively the protein core.

Thus, *lac* repressor realizes both orientations of the recognition helices relatively to the operator, pointed out in [20] as orientations of *lac* and *cro* types. Within the specific repressor–operator complex this helix has the 'specific' *lac* orientation [5,8,19,20,22,23]. This orientation arises upon the operator's influence. Binding of the inducer to the repressor fixes the reversed, *cro*-type, orientation. Now operator can not change this reversed orientation to 'specific', *lac*-type, orientation.

It has been anticipated that the solution of the puzzle of the

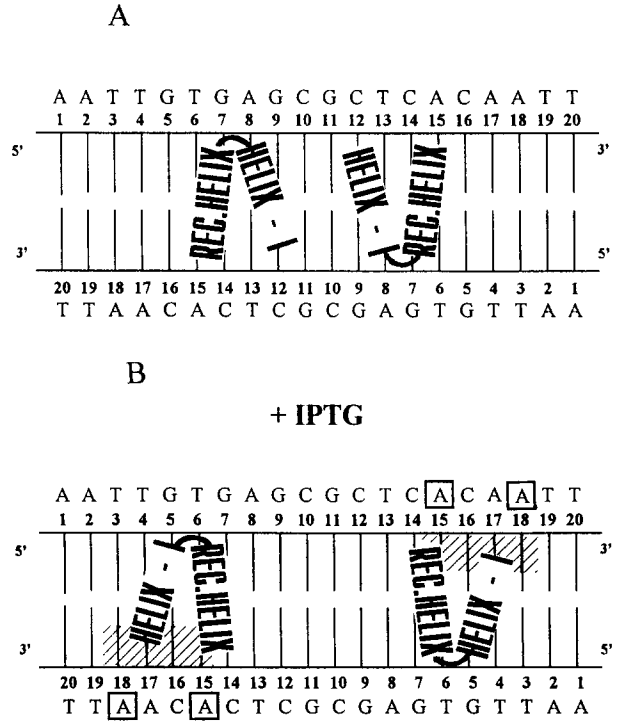

Fig. 4. Model illustrating how *lac* repressor determines the orientation of the two helix-turn-helix motifs relative to each other and an ideal *lac* operator. The recognition helix (residues 17–25) is depicted as REC. HELIX, the first helix (residues 6–13) as HELIX I. Panel A. Specific complex established by NMR and genetic studies (based on the figures from [24,25]). Panel B. Inducer-repressor-operator complex. The orientation of the motifs is reversed as compared with the specific complex. *Lac* repressor's N-termini contact operator DNA in the regions shown by hatching. Adenines A15 and A18 cross-linked to N-terminus are boxed.

repressor switch-off should be sought in the accuracy of the register of the two headpieces along DNA. Now, this prediction could be stated in simple molecular terms: the *cro*-type orientation freezes the headpieces in the configuration that is warped with respect to DNA recognition while still allowing the non-specific DNA-binding.

Acknowledgements: We are thankful to Dr. I. Teledsinskaya for peptide sequencing, to Dr. S. Grachev for the valuable advice, and to Prof. B. Müller-Hill for the gift of the *lac* repressor overproducing cells. This work was supported by Grant 94-04-12344 from Russian Fund for Basic Research.

References

[1] Miller, J.H. and Reznikoff, W. (Eds.), The Operon, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1979.
[2] Barkley, M.D., Riggs, A.D., Jobe, A. and Bourgrois, S. (1975) Biochemistry 23, 2933–2939.
[3] Ogata, R.T. and Gilbert, W. (1978) Proc. Natl. Acad. Sci. USA 75, 5851–5854.
[4] Kaptein, R., Zuiderweg, R.M., Scheek, R., Boelens, R. and van Gunsteren, W.F. (1985) J. Mol. Biol. 182, 179–182.
[5] Boelens, R., Scheek, R.M., van Boom, J.H. and Kaptein, R. (1987) J. Mol. Biol. 193, 213–216.
[6] Pabo, C.O. and Sauer, R.T. (1984) Annu. Rev. Biochem. 53, 293–321.
[7] Matthews, D.H., Ohlendorf, W.F., Anderson and Takeda, Y. (1982) Proc. Natl. Acad. Sci. USA 79, 1428–1432.
[8] Lehming, N., Sartorius, J., Oehler, S., von Wilken-Bergmann, B. and Müller-Hill, B. (1988) Proc. Natl. Acad. Sci. USA 85, 7947–7951.
[9] Brennan, R.G., Roderick, S.L., Takeda, Y. and Matthews, B.W. (1990) Proc. Natl. Acad. Sci. USA 87, 8165–8169.
[10] Ebralidse, K.K., Grachev, S.A. and Mirzabekov, A.D. (1988) Nature 331, 365–367.
[11] Mirzabekov, A.D., Bavykin, S.G., Belyavsky, A.V., Karpov, V.L., Preobrazhenskaya, O.V., Shick, V.V. and Ebralidse, K.K. (1989) Methods Enzymol. 170, 386–407.
[12] Geisler, N. and Weber, K. (1977) Biochemistry 16, 938–943.
[13] Sadler, J.R., Sasmor, H. and Betz, J.L. (1983) Proc. Natl. Acad. Sci. USA 80, 6785–6789.
[14] Simons, A., Tills, D., von Wilcken-Bergmann, B. and Müller-Hill, B. (1984) Proc. Natl. Acad. Sci. USA 81, 1624–1628.
[15] Nacheva, G.A., Guschin, D.Y., Preobrazhenskaya, O.V., Karpov, V.L., Ebralidse, K.K. and Mirzabekov, A.D. (1989) Cell 58, 27–36.
[16] Farabaugh, P.J. (1978) Nature 274, 765–769.
[17] Burton, K. and Petersen, G.B. (1960) Biochem. J. 75, 17–27.
[18] de Vlieg, J., Berendsen, H.J.C. and van Gunsteren, W.F. (1989) Proteins 6, 104–127.
[19] Kaptein, R., Lamerichs, R.M.J.N., Boelens, R. and Rullmann, J.A.C. (1990) Biochem. Pharm. 40, 89–96.
[20] Lamerichs, R.M.J.N., Boelens, R., van der Marel, G.A., van Boom, J.H., Kaptein, R., Buck, F., Fera, B. and Rüterjans, H. (1989) Biochemistry 28, 2985–2991.
[21] Boelens, R., Scheek, R.M., Lamerichs, R.M.J.N., de Vlieg, J., van Boom, J.H. and Kaptein, R. in DNA-ligand Interactions, Plenum, New York, pp. 191–215 (1987).
[22] Sartorius, J., Lehming, N., Kirsters-Woike, B., Wilken-Bergmann, B.V. and Müller-Hill, B. (1991) J. Mol. Biol. 218, 313–321.
[23] Chuprina, V.P., Rullmann, J.A., Lamerichs, R.M., van Boom, J.H., Boelens, R. and Kaptein, R. (1993) J. Mol. Biol. 234, 446–462.
[24] Kisters-Woike, B., Lehming, N., Sartorius, J., von Wilken-Bergmann, B. and Müller-Hill, B. (1991) Eur. J. Biochem. 198, 411–419.
[25] Kolkhof, P., Teichmann, D., Kisters-Woike, B., von Wilken-Bergmann, B. and Müller-Hill, B. (1992) EMBO J. 11, 3031–3038.