Spatial arrangement of the three α helices in the solution conformation of *E. coli lac* repressor DNA-binding domain

Erik R.P. Zuiderweg, Martin Billeter†, Rolf Boelens, Ruud M. Scheek, Kurt Wüthrich† and Robert Kaptein*

Department of Physical Chemistry, University of Groningen, Nijenborgh 16, 9747 Groningen, The Netherlands and †*Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule, ETH-Hönggerberg, CH-8093 Zürich, Switzerland*

Received 15 June 1984; revised version received 12 July 1984

The relative orientations of the 3 helices in the DNA-binding domain ('headpiece') of *lac* repressor have been determined using distance constraints obtained from 2-dimensional ^{1}H nuclear Overhauser enhancement spectra. The relative orientations of its helices is similar to that of the central 3 helices in the DNA-binding domain of the λ repressor of the bacteriophage λ.

DNA-binding protein *2D NMR* *Helix topology* lac *repressor headpiece* *NOE-distance constraint*

1. INTRODUCTION

The *lac* repressor of *Escherichia coli* is a tetrameric protein of M_r 150000. It regulates the expression of genes coding for enzymes involved in the lactose metabolism by specific binding to the *lac* operator [1,2]. Each of the four identical monomeric subunits consists of two domains. The DNA-binding domain [3–5] or 'headpiece' obtained by limited proteolytic digestion of the intact repressor [6] retains its native 3-dimensional structure and sequence-specific DNA binding capacity [4–9]. Single crystal structures are not available either for intact repressor or for the headpiece. However, due to its small size the isolated headpiece (51 amino acid residues) is accessible for high-resolution ^{1}H NMR [9–11]. We describe a continuation of recent NMR work [12,13] on the elucidation of the solution conformation of headpiece. In those studies sequence-specific resonance assignments were obtained for a large proportion of the hydrogen atoms [12]. Furthermore, the sequence location of 3 helices extending approximately from the residues 6–13, 17–25 and 34–45 (referred to as helix I, II and III, respectively) was determined [13].

Here, we determine the relative spatial orientations of these helical segments using distance constraints between assigned protons as obtained from 2-dimensional nuclear Overhauser enhancement spectroscopy (NOESY) [14,15] experiments. Based on the assumption that the helices are regular, the topology of these elements could be obtained from a relatively small number of nuclear Overhauser enhancement (NOE) distance constraints.

Abbreviations: headpiece, N-terminal DNA-binding domain of *lac* repressor of *E. coli*, comprising residues 1–51

* To whom correspondence should be addressed

2. MATERIALS AND METHODS

Headpiece was obtained as an enzymatic digest of *lac* repressor from *E. coli* strain BMH 74-12 as described previously [9]. Solutions (~5 mM) of headpiece in a buffer containing 0.4 M KCl, 0.05 M potassium phosphate and 0.02% NaN_3 were prepared, with the pH at 6.5 or 6.9, depending on the experiment. NOESY spectra were recorded at 18°C on a Bruker HX 360 spectrometer and at 27°C on a Bruker WM 500 spectrometer.

Published by Elsevier Science Publishers B.V.
00145793/84/$3.00

The standard pulse sequence [14,15] was used, with phase cycling and data processing similar to previously described procedures [12,16]. The spectra were recorded with mixing times of 50, 80, 120 and 200 ms both in ^2H_2O and in H_2O.

The upper limits on intramolecular ^1H–^1H distances obtained from the NOESY spectra were used as constraints in two independent model building procedures. In one, Nicholson molecular models were used to position the three helices, taken as separate building blocks, in accordance with a number of key NOEs. Then the intervening loops were attached, which provided further constraints on the relative helix orientation. The molecular model obtained was further adjusted to be compatible with virtually all NOE constraints listed in table 1. The other procedure makes use of an interactive computer graphics program written for an Evans & Sutherland picture system connected to a PDP 11-34 computer. The program allows generation of different conformations of molecular structures with up to 1000 atoms by real-time variation of the torsion angles about single bonds. For the structural analysis of NOESY distance constraints, which may be between widely separated positions in the sequence, it is important that the complete structure was stored in memory. As a result, up to 8 independent torsion angles can be varied simultaneously and the complete conformations can be inspected on the screen. Violations of distance constraints imposed by the NOESY data are visualized by flashing arrows in the conformers thus obtained (detailed description of this program in preparation). The helices were again assumed to be regular and were preserved during the entire fitting procedure. Different relative spatial locations of these helices were generated by variation of the torsion angles ϕ and ψ in the intervening peptide segments minimizing the violations of the experimental distance constraints.

3. RESULTS AND DISCUSSION

Fig.1 shows part of a 360 MHz ^1H NOESY spectrum of a headpiece solution in ^2H_2O recorded with a mixing time of 80 ms. This region contains the NOE cross peaks between the tyrosine ring protons and aliphatic protons. The present structural interpretation of long-range NOEs, i.e., NOEs between protons in residues which are located 5 or more positions apart in the sequence, relied on qualitative interpretations of the NOESY cross peaks. First, the cross peaks in the spectra recorded with a mixing time of 200 ms were identified on the basis of the previously obtained resonance assignments [12]. Secondly, only those NOEs were retained which could also be observed in spectra recorded with a mixing time of 50 or 80 ms (fig.1). These were taken to manifest that the corresponding proton–proton distance is $\leqslant$4.0 Å [17]. With this procedure the entire NOESY spectra recorded in H_2O and $2H_2O$ provided the long-range distance constraints indicated in fig.3.

Table 1

Long-range NOEs in *lac* repressor headpiece which were used for the present study of the spatial arrangement of the 3 helical segments

Observed NOEs		Corresponding upper distance constraints (Å)[a]
4 Val $C^{\gamma}H_3$	45 Leu $C^{\delta}H_3$	C^{β}–C^{β} 9.6
4 Val $C^{\gamma}H_3$	47 Tyr $C^{\alpha}H$	C^{β}–$C^{\alpha}H$ 6.2
4 Val $C^{\gamma}H_3$	47 Tyr $C^{\delta}H$	C^{β}–C^{β} 8.9
5 Thr $C^{\alpha}H$	47 Tyr $C^{\epsilon}H$	$C^{\alpha}H$–C^{β} 8.7
6 Leu $C^{\delta}H_3$	17 Tyr $C^{\epsilon}H$	C^{β}–C^{β} 12.2
6 Leu $C^{\delta}H_3$	24 Val $C^{\gamma}H_3$	C^{β}–C^{β} 9.6
6 Leu $C^{\delta}H_3$	47 Tyr $C^{\epsilon}H$	C^{β}–C^{β} 12.2
7 Tyr $C^{\alpha}H$	17 Tyr $C^{\epsilon}H$	$C^{\alpha}H$–C^{β} 8.7
8 Asp $C^{\beta}H_2$	45 Leu $C^{\delta}H_3$	C^{β}–C^{β} 8.6
9 Val $C^{\gamma}H_3$	42 Met NH	C^{β}–NH 6.2
9 Val $C^{\gamma}H_3$	42 Met $C^{\alpha}H$	C^{β}–$C^{\alpha}H$ 6.2
9 Val $C^{\alpha}H$	45 Leu $C^{\delta}H_3$	$C^{\alpha}H$–C^{β} 7.5
9 Val $C^{\gamma}H_3$	47 Tyr $C^{\delta}H$	C^{β}–C^{β} 8.9
10 Ala $C^{\beta}H_3$	17 Tyr $C^{\alpha}H$	C^{β}–$C^{\alpha}H$ 5.1
10 Ala $C^{\beta}H_3$	17 Tyr $C^{\delta}H$	C^{β}–C^{β} 7.8
10 Ala NH	20 Val $C^{\gamma}H_3$	NH–C^{β} 6.2
10 Ala $C^{\alpha}H$	20 Val $C^{\gamma}H_3$	$C^{\alpha}H$–C^{β} 6.2
12 Tyr $C^{\delta}H$	41 Ala $C^{\beta}H_3$	C^{β}–C^{β} 7.8
12 Tyr $C^{\epsilon}H$	44 Glu $C^{\beta}H_2$	C^{β}–C^{β} 9.8
12 Tyr NH	45 Leu $C^{\delta}H_3$	NH–C^{β} 7.5
12 Tyr $C^{\beta}H_2$	45 Leu $C^{\delta}H_3$	C^{β}–C^{β} 8.6
13 Ala $C^{\beta}H_3$	38 Val $C^{\alpha}H$	C^{β}–$C^{\alpha}H$ 5.1
13 Ala NH	41 Ala $C^{\beta}H_3$	NH–C^{β} 5.1
13 Ala $C^{\alpha}H$	41 Ala $C^{\beta}H_3$	$C^{\alpha}H$–C^{β} 5.1
13 Ala $C^{\beta}H_3$	41 Ala NH	C^{β}–NH 5.1
13 Ala $C^{\beta}H_3$	41 Ala $C^{\beta}H_3$	C^{β}–C^{β} 6.2
20 Val $C^{\gamma}H_3$	38 Val NH	C^{β}–NH 6.2
24 Val $C^{\gamma}H_3$	47 Tyr $C^{\epsilon}H$	C^{β}–C^{β} 10.9

[a] All NOEs to sidechain protons are referred to C^{β} (see text)

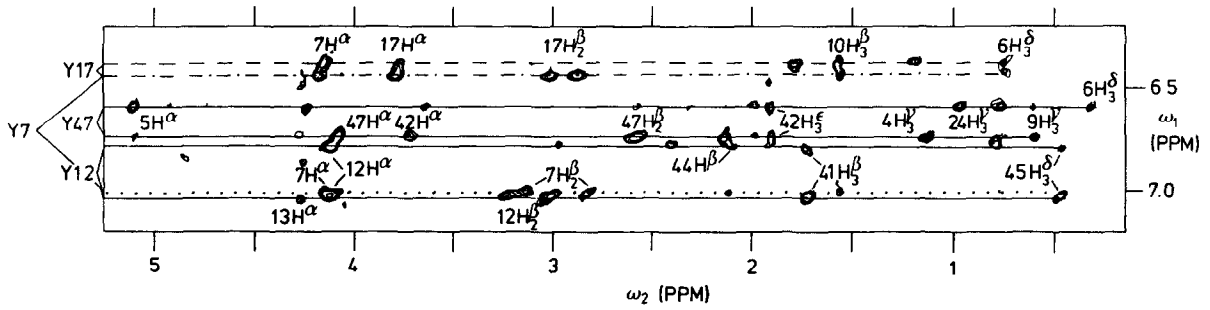

Fig.1. Contour plot of the region $\omega_1 = 6.2–7.2$ ppm, $\omega_2 = 0.1–5.2$ ppm of an absorption mode 360 MHz ^{1}H NOESY spectrum of 6 mM headpiece in ^2H_2O buffer, pH 6.9, 18°C, recorded with a mixing time of 80 ms. The time domain data consisted of 393 free induction decays of 1024 data points, with t_1 values from 0–63 ms. The spectrum was baseline corrected and symmetrized, the digital resolution is 6.1 Hz/point. This region contains the cross peaks from the ring proton resonances of the tyrosyl residues (indicated with horizontal lines and identified on the left) to aliphatic protons (identified by the sequence position (fig.2) and type of hydrogen atom).

```
1          10         20
MKPVTLYDVAEYAGVSYQTVSRVVN

30         40         50
QASHVSAKTREKVEAAMAELNYIPNR
```

Fig.2. Amino acid sequence [18] from residues 1–51 of *E. coli lac* repressor with IUB/IUPAC one-letter symbols.

The presently collected NOESY distance constraints are not sufficient to determine the conformation of the amino acid side chains. Therefore all experimental distance constraints were referred to backbone hydrogens or C^β atoms. The NOEs to amide and C^α protons were used without correc-

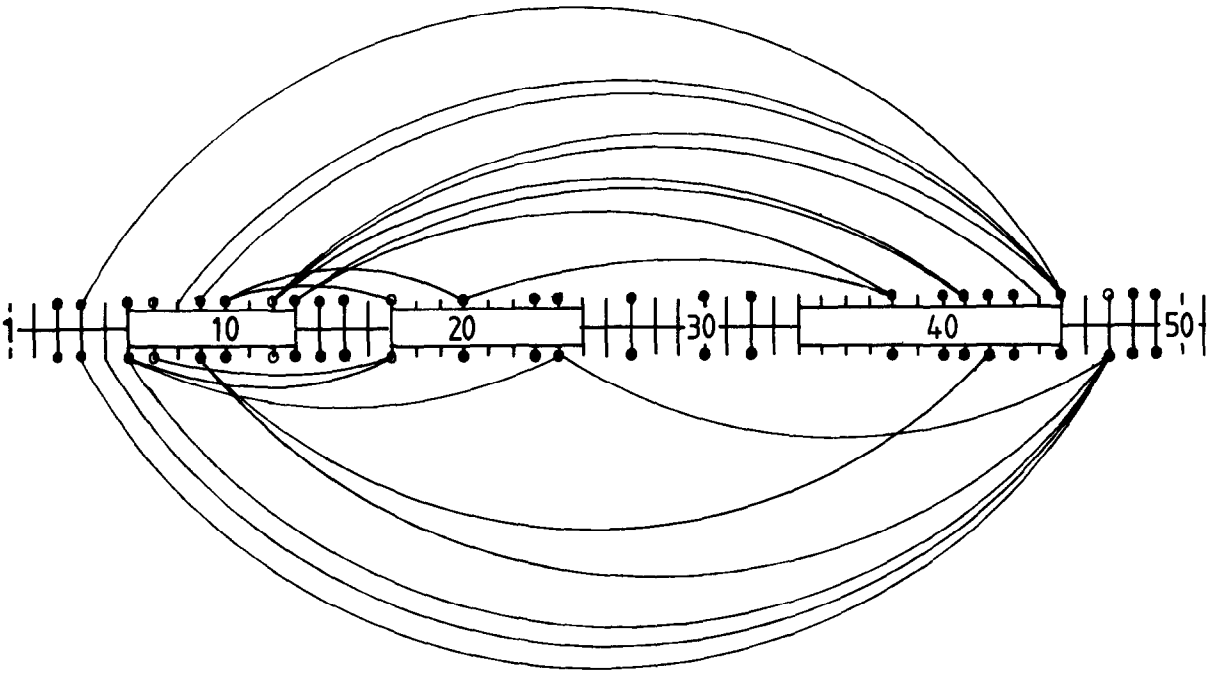

Fig.3. Schematic presentation of the primary and secondary structure of headpiece with indication of the long range NOEs. Helices are indicated with boxes. Non-polar side chains are marked with filled circles and tyrosyl residues with open circles. The division of the lines indicating long-range NOEs above and below the sequence was chosen for presentation purposes only.

tion. For the NOEs referred to the C^{β} atoms, the maximum sterically allowed distance from the observed hydrogen atoms to C^{β} was added to the NOE upper distance constraint of 4.0 Å. The resulting input for the determination of the topology of the three helices is listed in table 1.

In addition to the NOE constraints of table 1, the covalent linkage consisting of only 3 residues strongly restricts the possible relative spatial arrangements of the first two helices. (Previously, the first helix was proposed to extend from positions 6 to 15 [13] rather than from 6 to 13. This apparent discrepancy arises because the previously used pattern recognition of secondary structure cannot usually distinguish between residues at the ends of a helix and those in tight turns adjacent to the helix.) The linkage by an octapeptide segment, however, hardly restricts the allowed relative orientations of the second and third helices.

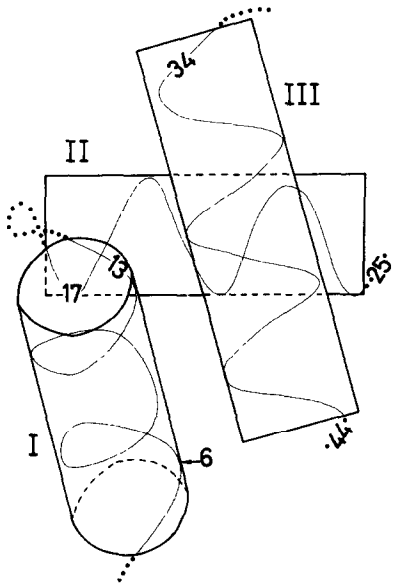

Fig.4. Computer drawing illustrating the spatial arrangement of the 3 helices in *lac* repressor headpiece. The molecule is oriented so that the helices II and III are parallel to the projection plane and helix I runs from back to front at an angle of ~50° relative to that plane. The polypeptide backbone in the helices is represented by a smooth line through the C^{α} positions, where those of residues 6, 13, 17, 25, 34 and 44 are identified. Outside the helices the backbone is indicated by dotted lines, where most of the loop from residues 26–33 and of the chain termini is omitted.

The helix topology shown in fig.4 was obtained by the procedures outlined above. Its compatibility with the experimental data can readily be checked by inspection of table 1 and fig.3. There are 4 NOEs which imply that the two chain termini are near each other in space. Helix I has close contacts with all other structure elements, i.e., there are 7 NOEs to helix II, 13 NOEs to helix III and 2 NOEs to the C-terminal segment. In contrast there is only one single NOE between helix II and helix III.

Consideration of only a small number of long-range distance constraints is sufficient to determine the topology of helices in a unique way. The argument runs as follows. First, helix I and helix II form a V-shaped unit which follows from NOEs Ala 10–Val 20 and Leu 6–Val 24. Secondly, helix I and helix III run approximately antiparallel, which is inferred from NOEs Tyr 12–Ala 41, Asp 8–Leu 45 and Val 9–Leu 45, and from the contacts Val 4–Tyr 47, Thr 5–Tyr 47 and Leu 6–Tyr 47 which show that the C-terminus of headpiece folds back onto its N-terminus. The localization of helix III as sketched in fig.4 appears to be the only conformation in which Val 20 in helix II can be close to Val 38 in helix III.

To satisfy the NOEs between Tyr 17 in helix II and Leu 6 and Tyr 7 in helix I, the C terminus of helix I is shifted over the N terminus of helix II as indicated in the figure. Although it is not apparent from fig.4, in this orientation the long side chains of these residues can come within NOE distance.

It can be seen from fig.3 that the majority of the identified NOEs occur between protons of hydrophobic amino acid residues. As many of these residues make multiple contacts, this indicates that they are clustered in the interior of the protein. Tyr 45 is in contact with at least 5 non-polar residues and therefore is part of the hydrophobic cluster. This concurs with genetic and biochemical data showing that Tyr 47 can neither be mutated nor removed without disrupting the 3-dimensional structure of the protein [3,10].

Finally, it is noted that the relative orientation of the helices of *lac* repressor headpiece is very similar to that of the central 3 helices of the DNA-binding domain of the λ repressor of the bacteriophage λ, a protein that, as *lac* repressor, recognizes a specific DNA sequence [19].

ACKNOWLEDGEMENTS

The experimental work was carried out at the Dutch National NMR Facilities at Groningen and Nijmegen and was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO). We acknowledge financial support by the Schweizerischer Nationalfonds (project 3.284.82), a special grant of the ETH Zürich and use of the facilities at the Zentrum für Interaktives Rechnen (ZIR) of the ETH.

REFERENCES

[1] Jacob, F. and Monod, J. (1961) J. Mol. Biol. 3, 318–356.
[2] Caruthers, M.H. (1980) Acc. Chem. Res. 13, 155–160.
[3] Miller, J.H., Coulondre, C., Hofer, M., Schmeissner, U., Sommer, H., Smitz, A. and Lu, P. (1982) Proc. Natl. Acad. Sci. USA 79, 218–222.
[4] Ogata, R.T. and Gilbert, W. (1979) J. Mol. Biol. 132, 709–728.
[5] Schlottmann, M. and Beyreuther, K. (1979) Eur. J. Biochem. 95, 39–49.
[6] Geisler, N. and Weber, K. (1977) Biochemistry 16, 938–943.
[7] Wade-Jardetzky, N.G., Bray, R.P., Conover, W.E., Jardetzky, O., Geisler, N. and Weber, K. (1979) J. Mol. Biol. 128, 259–264.
[8] Nick, H., Arndt, K., Boschelli, F., Jarema, M.A., Lillis, M., Sadler, J., Caruthers, M. and Lu, P. (1982) Proc. Natl. Acad. Sci. USA 79, 218–222.
[9] Scheek, R.M., Zuiderweg, E.R.P., Klappe, K.J.M., van Boom, J.H., Kaptein, R., Rüterjans, H. and Beyreuther, K. (1983) Biochemistry 22, 228–235.
[10] Ribeiro, A.A., Wemmer, D., Bray, R.P., Wade-Jardetzky, N.G. and Jardetzky, O. (1981) Biochemistry 20, 818–823.
[11] Arndt, K.T., Boschelli, F., Lu, P. and Miller, J.H. (1981) Biochemistry 20, 6109–6118.
[12] Zuiderweg, E.R.P., Kaptein, R. and Wüthrich, K. (1983) Eur. J. Biochem. 137, 279–292.
[13] Zuiderweg, E.R.P., Kaptein, R. and Wüthrich, K. (1983) Proc. Natl. Acad. Sci. USA 80, 5837–5841.
[14] Jeener, J., Meier, B.H., Bachmann, P. and Ernst, R.R. (1979) J. Chem. Phys. 71, 4546–4553.
[15] Anil Kumar, Wagner, G., Ernst, R.R. and Wüthrich, K. (1981) J. Am. Chem. Soc. 103, 3654–3658.
[16] States, D.J., Haberkorn, R.A. and Reuben, D.J. (1982) J. Magn. Reson. 48, 286–292.
[17] Braun, W., Wider, G., Lee, K.H. and Wüthrich, K. (1983) J. Mol. Biol. 169, 921–948.
[18] Beyreuther, K. (1978) in: The Operon (Miller, J.H. and Reznikoff, W. eds) pp. 123–154, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[19] Pabo, C.O. and Lewis, M. (1982) Nature 298, 443–447.