The EMBO Journal vol.12 no.5 pp.1805-1810, 1993

The X-ray structure of an atypical homeodomain present in the rat liver transcription factor LFB1/HNF1 and implications for DNA binding

T.A.Ceska, M.Lamers, P.Monaci[1], A.Nicosia[1], R.Cortese[1] and D.Suck[2]

EMBL, Biological Structures and Biocomputing Programme, Meyerhofstrasse 1, 6900 Heidelberg, Germany
[1]Present address: Istituto di Ricerche di Biologia Molecolare, Via Pontina KM 30 600, 00040 Pomezia, Roma, Italy

[2]Corresponding author

Communicated by R.Cortese

The transcription factor LFB1/HNF1 from rat liver nuclei is a 628 amino acid protein that functions as a dimer binding to the inverted palindrome GTTAATN-ATTAAC consensus site. We have crystallized a 99 residue protein containing the homeodomain portion of LFB1, and solved its structure using X-ray diffraction data to 2.8 Å resolution. The topology and orientation of the helices is essentially the same as that found in the engrailed, MATα2 and *Antennapedia* homeodomains, even though the LFB1 homeodomain contains 21 more residues. The 21 residue insertion is found in an extension of helix 2 and consequent lengthening of the connecting loop between helix 2 and helix 3. Comparison with the engrailed homeodomain-DNA complex indicates that the mode of interaction with DNA is similar in both proteins, with a number of conserved contacts in the major groove. The extra 21 residues of the LFB1 homeodomain are not involved in DNA binding. Binding of the LFB1 dimer to a B-DNA palindromic consensus sequence requires either a conformational change of the DNA (presumably bending), or a rearrangement of the subunits relative to the DNA.

Key words: DNA binding/homeodomain/LFB1/HNF1/X-ray structure

Introduction

High resolution X-ray structures of protein-DNA complexes solved in recent years, have shown that a number of different structural motifs are being used by nature for the specific recognition of DNA sequence. Examples include the helix-turn-helix, β-sheet, zinc finger, leucine zipper and helix-loop-helix motifs. A common mechanism to increase the affinity and/or the specificity of the interaction is the binding of dimers to palindromic sites or the simultaneous binding of several identical or different motifs present in the same protein molecule. Both mechanisms are being used by LFB1 (also called HNF1), a protein which has been implicated in the liver-specific transcription of several genes (Frain *et al.*, 1989). LFB1/HNF1 from rat liver nuclei binds as a dimer to the palindromic consensus site GTTAATNATTAAC (Tomei *et al.*, 1992).

The DNA binding region of LFB1 comprising the N-terminal 281 residues of the protein consists of three structurally and functionally distinct domains: a 32 amino acid dimerization domain which resembles sequences in the myosin heavy chain (Chouard *et al.*, 1990; Nicosia *et al.*, 1990) and appears to be folded into two helical segments preceded by an extended chain segment (Pastore *et al.*, 1991, 1992), a region related to the POU-specific A-box (Rosenfeld, 1991), and a highly diverged homeodomain that is 21 residues larger than classical homeodomains. The latter two domains are necessary and sufficient for specific recognition, while the dimerization domain increases the DNA binding affinity (Tomei *et al.*, 1992). Based on sequence alignments the extra 21 residues were proposed to form a loop between helices two and three of the homeodomain structure (Finney, 1990).

In POU-domain-containing transcription factors, e.g. the Oct 1 transcription factor, both subdomains (i.e. the POU-type homeodomain and the POU-specific domain) are required for specific high affinity binding to DNA (Verrijzer *et al.*, 1992). This is true for LFB1 as well. In both cases the N-terminal subdomain appears to make contacts to the 5'-end of the recognition site and to influence the relative orientation of the homeodomains. It is interesting to note that the DNA affinity of the LFB1 protein dimer or the POU-domain proteins is not much higher than that of a classical 'single'-homeodomain protein.

In this paper we describe the three-dimensional structure of a 99 residue protein containing the atypical homeodomain of LFB1 as determined by X-ray crystallography at 2.8 Å resolution. By comparing this structure with the known structures of classical homeodomains we are addressing the question of the function of the additional 21 residue insertion present in LFB1 and in particular how it might affect the interaction with DNA. The overall similarity between LFB1 and classical homeodomain structures allows us to propose a model for its interaction with DNA showing a number of conserved features, but also indicating distinct differences.

Results

Crystallization and structure determination

The N-terminal DNA binding regions of LFB1 can be expressed and purified from *Escherichia coli* and we have successfully produced large quantities of purified proteins from the cloned genes. Crystallization trials of the LFB1 homeodomain with and without a consensus oligomer yielded crystals that were shown to contain only protein. Two different crystal forms were obtained and the best crystals grew in space group $P6_122$ with cell dimensions a = b = 81.0 and c = 85.1 Å. Two heavy atom derivative data sets yielding interpretable difference Patterson maps were collected and the derived phases were modified by a solvent flattening procedure. The resulting electron density map is shown in Figure 1 with the refined positions of some residues.

The structure of the LFB1 homeodomain consists of three

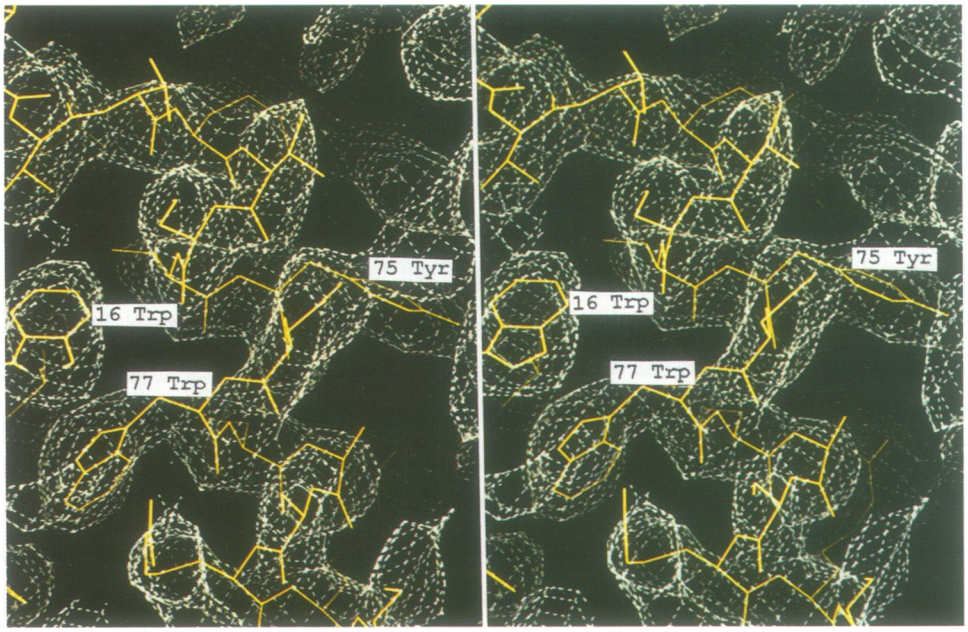

Fig. 1. Solvent flattened σ_A weighted (Read, 1986) MIRAS electron density map of LFB1 contoured at a 1σ level with the current model superimposed. Shown is part of the hydrophobic cluster in the interior of the protein as well as some residues on the third helix.

alpha helices packed tightly against each other (Figure 2) with a fold very similar to that of other known homeodomains. Residues found to be conserved in other homeodomains (Finney, 1990) are also conserved in the LFB1 homeodomain, for example residue LFB1-Leu24 and the WFXNXR motif in helix 3. Other conserved residues, including LFB1-Asn80 and LFB1-Arg82 make contact with DNA. The N and C-termini of the protein contain the residues that result from the cloning of the protein into *E.coli*, and are disordered. The residues between 13 and 89 have been built into the electron density and refined. Residues 1−12 are completely absent from the electron density map and residues 90−99 are ill-defined or absent. An interesting difference in the interior packing of the homeodomain includes LFB1-Trp16 which is a Phe or Tyr residue in most other homeodomains (Finney, 1990) and packs against the Leu40 sidechain. In LFB1 this Leu has been changed to Val69, a smaller residue, to compensate for the larger Trp16. The WF residues in the WFXNXR motif form part of the hydrophobic core sandwiched between residues LFB1-Leu24 and LFB1-Tyr28. These four residues are highly conserved in most homeodomains, with Phe most commonly found instead of Tyr28.

Although the crystals contain a large proportion of solvent, the orientation of the packed molecules precludes binding of duplex DNA to the homeodomain in the crystal (Figure 3). Also of note is the position of the long loop between helices 2 and 3, comprising residues 55−70. Residues 55−59 make a weak contact with the neighbouring molecule, but residues 60−69 are exposed to solvent, and we observe this region to be more poorly ordered than the rest of the molecule.

Discussion

Comparison with classical homeodomain structures

The protein topology is the same as has been determined for the three other known homeodomain structures; engrailed

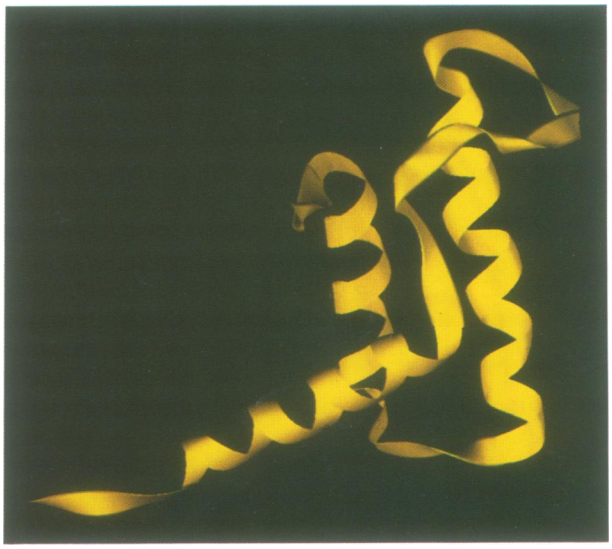

Fig. 2. A ribbon plot of the LFB1 homeodomain showing the overall topology of the protein, made with the RIBBONS program (Carson, 1987).

(Kissinger *et al.*, 1990), MATα2 (Wolberger *et al.*, 1991) and *Antennapedia* (Qian *et al.*, 1989; Otting *et al.*, 1990) homeodomains. The Cα backbones of LFB1, engrailed and MATα2 homeodomains are shown superimposed in Figure 4. The more closely related protein is the engrailed homeodomain, which not only has a very similar loop between helix 1 and helix 2, but also has the same TAAT core recognition sequence.

The LFB1 homeodomain is 21 residues longer than the homeodomains whose structures have been solved (Figure 5), and the structural effects of this major difference can now be addressed. We have found that the consequence of the insertion is a lengthening of the second helix by eight residues, and an extension of the loop between helix 2 and helix 3 by 13 residues. This long loop region has greater

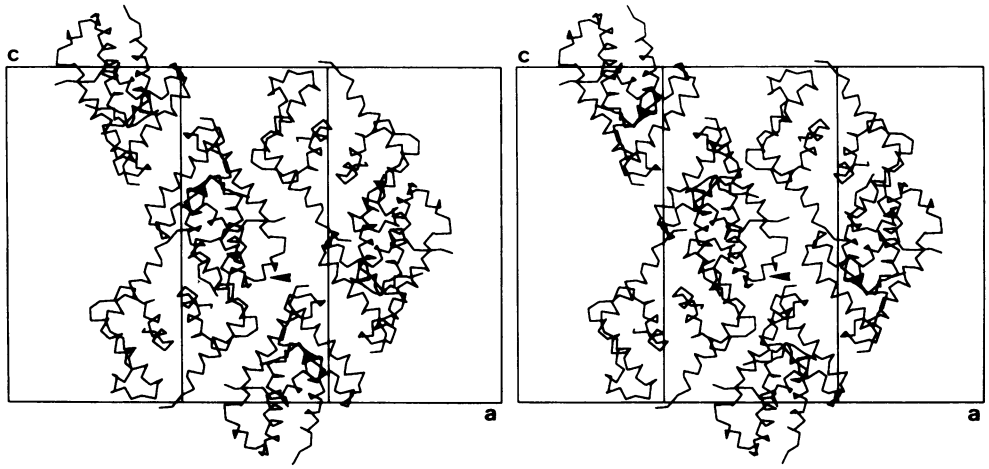

Fig. 3. A packing diagram illustrating the orientation of the homeodomain in the crystal. A view perpendicular to the c-axis is shown. The loop region between helices 2 and 3 is indicated, showing that the beginning of this loop makes weak contact to a neighbouring molecule, and the remaining part of the loop makes no crystal contacts. The conformation of the loop is not constrained by crystal contacts, and is not well defined in our maps.

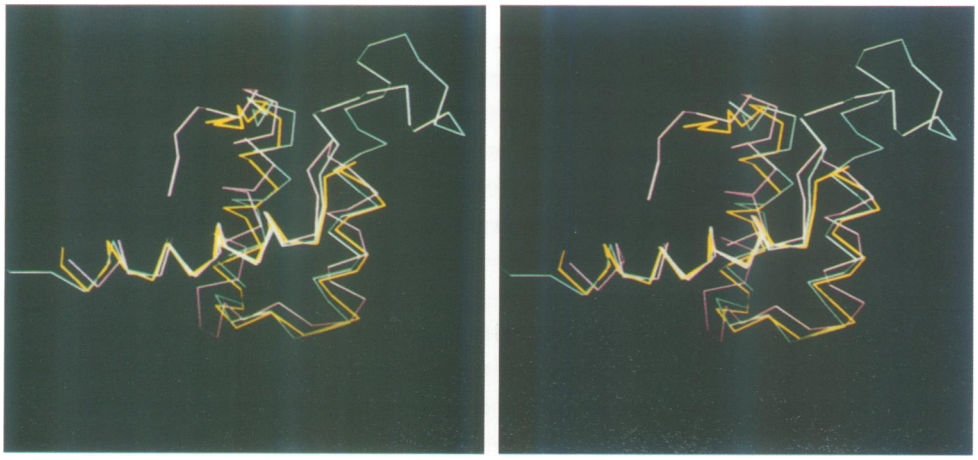

Fig. 4. A Cα trace of the LFB1 homeodomain (blue), overlayed on to the engrailed homeodomain (yellow) and the MATα2 homeodomain (red). The view is similar to that shown in Figure 2. The 21 residue extension is partly helical giving rise to an extension of helix 2 and partly in the extended loop between helix 2 and 3. The alignment of the structures was done by matching the helical residues in the LFB1 third helix with the corresponding residues of the MATα2 and engrailed homeodomains.

```
                5    10   15    20   25   30       35    40   45   50    55   60   65   70    75   80   85       90   95
                .    .    .     .    .    .        .     .    .    .     .    .    .    .     .    .    .        .    .
LF-B1     aridptkkGRRNRFKWG PASQQILFQAYER     QKNPS KEERETLVEECNRAECIQR GVSPSQAQGLGSNLVT EVRVYNWFANRRKEEA     FRHklamdtykln
Contact residues   ? ?  *                          *  *                                    ***  * ***
Insertion                                                  12345678 1234567890123

MATA2           TKPYRGHRFT KENVRILESWFAK NIENPYLD TKGLENLMKNT                    SLS RIQIKNWVSNRRRKEKT   I
ENGR          MDEKRPRTAFS SEQLARLKREFNE    NRYLT ERRRQQLSSEL                   GLN EAQIKIWFQNKRAKIKK   S
ANTP            RKRGRQTYT RYQTLELEKEFHF    NRYLT RRRRIEIAHAL                   CLT ERQIKIWFQNRRMKWKK   EN
                .    .    .     .    .    .        .     .    .    .     .    .    .    .     .    .    .        .    .
                5    10   15    20   25   30       35    40   45   50    55   60   65   70    75   80   85       90   95

                     < Helix 1  >          < Helix 2          >                < Helix 3      >
```

Fig. 5. A sequence alignment based on the structure of the LFB1 homeodomain and the MATα2, engrailed and *Antennapedia* homeodomains. Sequence identity between the LFB1 and engrailed homeodomains is ~25%. The breaks indicate the extent of the α-helical regions in each homeodomain. The 21 residue insertion is between the second and third helices, extending the second helix by eight residues and extending the turn by 13 residues. The residues in lower case characters at the beginning and end of the LFB1 sequence are present as a result of the cloning protocol, and are not present in the LFB1 gene. The residues in contact with DNA described in the text are highlighted. The symbol (*) denotes residues that are observed to contact the DNA, and the symbol (?) denotes a hypothetical minor groove contact based on the structural results from the engrailed and MATα2 homeodomain–DNA complexes.

flexibility, as determined by the quality of the electron density in the map and by the B-factors calculated in the refinement of the model (main chain B-factors of 30 Å^2 compared with 22 Å^2 for the helical regions). However, in view of the weak intermolecular contacts involving residues in this loop (see above), it appears unlikely that its conformation is determined to any great extent by crystal packing forces.

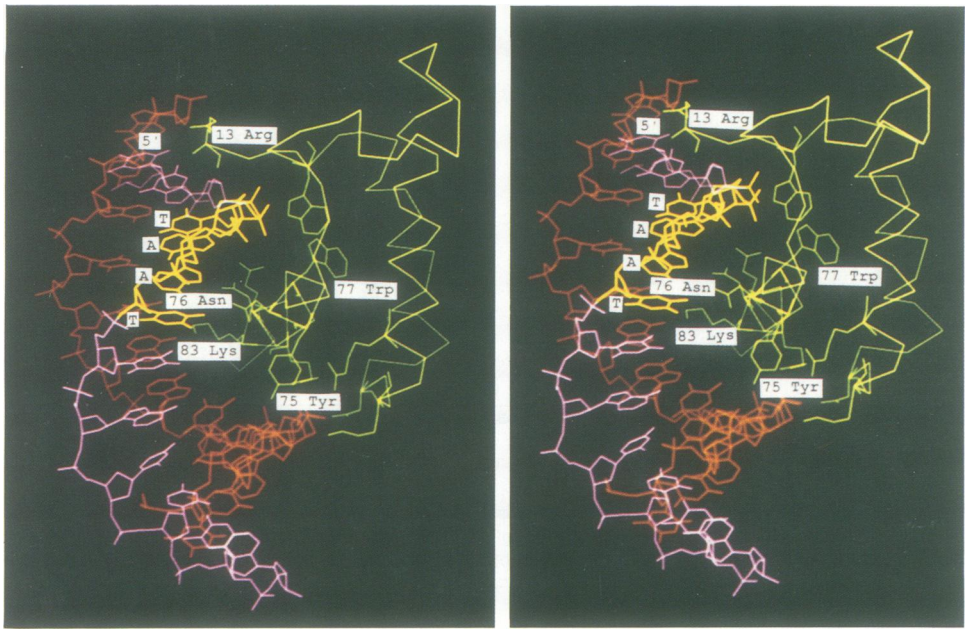

Fig. 6. A proposed model of the interaction of the LFB1 homeodomain with DNA. The recognition sequence for the engrailed homeodomain contains the same core **TAAT** sequence that is in the consensus sequence for LFB1. The **TAAT** sequence is highlighted in yellow. The 5′ start of the DNA strand is indicated at the top. The DNA segment as seen in the X-ray structure of the engrailed–DNA complex containing this **TAAT** sequence was docked to the LFB1 homeodomain, assuming the same positioning of the third (recognition) helix. The labelled residues were chosen for clarity. Asn80, a highly conserved residue is located one turn of a helix further down from Asn76. Considering the interactions between the third helix and the central TAAT core, residues LFB1-Arg82 and LFB1-Trp77 make contacts to the phosphate backbone. LFB1-Tyr75 makes an additional phosphate backbone contact. Several backbone contacts have been lost, such as those from LFB1-Ala86 and LFB1-Glu84 (which is in a position to make a potential van der Waal's contact to a base). LFB1-Asn76 makes contact to the 3′ thymine (T14) of the TAAT core and LFB1-Asn80 make contact to the second adenine of the TAAT core (A13). LFB1-Lys83 makes an additional contact to the adenine base complementary to T14. In addition, other residues with proposed contacts to the DNA are described in the text.

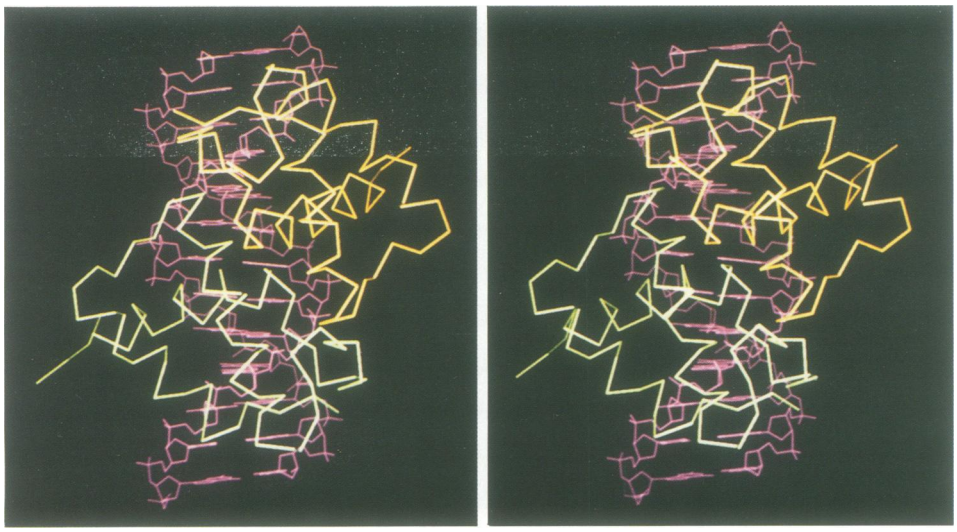

Fig. 7. Binding of dyad-related LFB1 homeodomains to the palindromic GT**TAAT**(N)**ATTAA**C consensus site (Frain *et al.*, 1989). The model was derived from the monomer binding shown in Figure 5 by applying a 2-fold rotation axis passing through the central base pair and assuming a regular B-type DNA conformation. Obvious steric clashes involving residues in the insertion between helices 2 and 3 and at the beginning of helix 3 can only be avoided if either the DNA conformation is changed, e.g. through bending, or the position of the monomers relative to the DNA is changed.

Postulated model of the interaction with DNA

We have chosen to use the engrailed homeodomain DNA cognate sequence which resembles the LFB1 consensus sequence for our modelling. We assume the same position of helix 3 relative to the DNA, and key residues that may interact with the DNA have been added to the Cα trace of the LFB1 homeodomain (Figure 6). Many proposed contacts are similar to those seen in the engrailed–DNA structure. As an example, LFB1-Arg39, LFB1-Arg82 and LFB1-Trp77 make contacts to the phosphate backbone similar to those seen in engrailed. There seem to be additional phosphate backbone contacts from residues LFB1-Lys36, LFB1-Tyr75 and LFB1-Trp16. Several backbone contacts have been lost, namely engr-Tyr25 (LFB1-Asn33), engr-

Table I. Summary of X-ray data

Data set	Res.	R_{sym}[a]	% complete	R_{diff}[b]	F_H/E[c]	R_{Cullis}[d]	No. sites	Anomalous
Native	2.8 Å	5.4	63	–	–	–	–	–
(1) $K_3UO_2F_5$	3.0 Å	7.3	69	16.6	1.57	56.1	1	Yes
(2) K_2PtBr_4	3.8 Å	10.4	86	13.8	1.55	56.6	4	No
(3) $K_3UO_2F_5$	3.0 Å	8.3	68	16.8	1.39	59.6	1	Yes

[a] $R_{sym} = \left\{ \frac{\Sigma_h \Sigma_i |\bar{I}h - I_{h,i}|}{\Sigma_h \Sigma_i |\bar{I}h|} \right\}$

[b] R_{diff}(*fractional isomorphous difference*) $= \left\{ \frac{\Sigma |FP_H - FP|}{\Sigma FP} \right\}$

[c] F_H/E(*phasing power*) = *r.m.s.*(F_H)/*r.m.s.*(*lack of closure error*), where F_H is the calculated heavy atom structure factor

[d] $R_{Cullis} = \left\{ \frac{\Sigma ||F_{H(obs)}| - |F_{H(calc)}||}{\Sigma |F_{H(obs)}|} \right\}$,

where $|F_{H(obs)}|$ is the observed heavy atom structure factor amplitude and $|F_{H(calc)}|$ is the calculated heavy atom structure factor amplitude, for centric reflections.

Lys57 (LFB1-Ala86), engr-Thr6 (LFB1-Phe14) and engr-Lys55 (LFB1-Glu84 is in a position to make a potential van der Waal's contact to a base).

There are changes to the major groove contacts, the most notable being a change to LFB1-Ala79 from engr-Gln50. This residue at position 9 of helix 3 has been implicated as an important residue for DNA recognition (Laughon, 1991; Riddihough, 1992) and for determining differences in specificity of related homeodomains. According to our model, LFB1-Asn76 makes contact to the 3′ thymine (T14) of the TAAT core in place of engr-Ile47 and LFB1-Asn80 and engr-Asn51 make similar contacts to the second adenine of the TAAT core (A13). LFB1-Lys83 could make an additional contact to the adenine base complementary to T14.

The minor groove contacts found in the engrailed homeodomain are not directly seen in our proposed model. The N-terminus is too far away from the DNA to make any contact. However, we would suggest that at least LFB1-Arg13 and perhaps LFB1-Arg11 will make contacts with the DNA minor groove *in vivo*, based on the contacts seen in engrailed, MATα2 and *Antennapedia* homeodomains and the strong conservation of the RXR motif in most other homeodomain sequences. The proposed contacts made by the LFB1 homeodomain to a DNA model are summarized in Figure 5.

The second question we can address is how a homeodomain dimer might interact with DNA. A dimer of the DNA binding region of LFB1 binds (Frain *et al.*, 1989; Tomei *et al.*, 1992) to the inverted palindrome consensus sequence GT**TAAT**(N)**ATTA**AC. We have modelled this DNA sequence in a B-DNA conformation, and have docked two LFB1 homeodomains on to the appropriate DNA binding sites as before. Figure 7 shows the relative orientation of the molecules. Although any details of the interaction would be highly speculative, there are two striking features of the model. First, are the relative locations of the extended loops between helices 2 and 3. There has been a suggestion that these residues could be a mediator of dimerization (Finney, 1990), and we observe an interface between the two domains which consists of the residues in the loop between helix 1 and helix 2, a few residues at the end of the loop between helix 2 and helix 3 and a few residues at the beginning of the third helix. The end of the loop between helix 2 and helix 3 is solvent exposed in the crystals we have obtained (Figure 3) and the structure appears to be relatively flexible, and may change conformation upon dimer formation.

A second feature is the close distance between these regions. This suggests that some change to the position of the domains and/or a change of conformation of the DNA would be required for packing with no steric clashes. In agreement with this it has been shown that POU/homeodomain DNA binding proteins bend DNA (Verrijzer *et al.*, 1991). We would anticipate that changes in the dimer contact residues could alter the binding affinity of the LFB1 dimer to its consensus sequence.

Specificity of binding

With respect to the specificity of binding to DNA, it is now clear that the isolated homeodomain of LFB1 is not sufficient for specific recognition of the LFB1 consensus sequence. Recent experiments with recombinant proteins containing only one or a combination of two subdomains of the tripartite LFB1 DNA binding domain have shown that the A-box POU-related portion of LFB1 is essential for specificity (Tomei *et al.*, 1992). The LFB1 homeodomain alone binds to a consensus oligonucleotide with a dissociation constant of $\sim 2 \times 10^{-9}$ M and other DNA sequences with only a 10-fold loss of affinity. In the natural cognate sequences the GT dinucleotide immediately 5′ to the **TAAT** recognition sequence is strictly conserved (Frain *et al.*, 1989) (i.e. **GTTAAT**(N)ATTAAC). In our model of the LFB1 homeodomain–DNA complex there are no residues contacting these bases. The A-box related domain precedes the homeodomain in the sequence, and the N-terminus of the homeodomain is at the 5′ end of the DNA sequence. We would suggest that this is where the A-box related domain interacts with DNA, and we would anticipate specific contacts between residues in this domain and the **GT** bases in the DNA.

Materials and methods

Protein purification

LFB1 homeodomain was overexpressed in *E.coli* using a T7 expression system. Cells were grown at 37°C to an absorbance of 0.7 at 600 nm induced with 0.44 mM isopropyl-β-D-thiogalactopyranoside for 3 h. Following

sonication, nucleic acid was removed by 1% protamine sulfate precipitation. The cleared supernatant was subjected to a 55% ammonium sulfate fractionation. The pellet containing the LFB1 homeodomain was dissolved in a high-salt buffer (20 mM Tris−HCl pH 9.0, 1.5 M $(NH_4)_2SO_4$, 2 mM DTT, 1 mM EDTA, 1 mM PMSF) and fractionated on a Phenyl Sepharose column with a linear 1.5−0.0 M $(NH_4)_2SO_4$ gradient. The homeodomain eluted at 1.1 M salt was further purified on a FPLC Superdex G-75 gel filtration column using a high salt buffer (20 mM Tris−HCl pH 9.0, 500 mM NaCl, 2 mM DTT). Peak fractions were concentrated to 40 mg/ml for crystallization.

Crystallization and data collection

The crystals were grown at room temperature in a hanging drop with a well solution containing 100 mM acetate pH 5.0 and 42% ammonium sulfate. The drops initially contained 2 μl of protein solution and 4 μl of well solution. Pyramidal crystals grew out of precipitate after 2 weeks to a typical size of 0.3×0.3×0.2 mm. The space group and cell dimensions were determined to be $P6_122$ with a = b = 81.0 and c = 85.1 Å. The Matthews' parameter assuming one molecule in the asymmetric unit was determined to be 3.6 $Å^3$/dalton corresponding to a solvent content of ~65%.

The native data and first two derivative data sets (Table I) were collected on a MAR image plate scanner (Marresearch), and the third derivative data set ($K_3UO_2F_5$) was collected on a FAST area detector (Enraf-Nonius), with the crystal oriented to maximize the anomalous signal (data processed with MADNES). The first two derivative data sets were processed with MOSFLM. Native data from two crystals were processed with the XDS package (Kabsch, 1988). A total of 43 872 reflections were collected to 2.8 Å which gave 4344 unique reflections. The derivatives were obtained by transferring the crystals to a buffer containing 100 mM acetate pH 5.0 and 50% ammonium sulfate with 0.5 mM of the heavy atom compound.

Structure determination

The Patterson maps were interpreted with the aid of the VECSUM program in the CCP4 crystallographic package (CCP4, 1979), and one heavy atom site for the uranyl derivative and four heavy atom sites for the platinum derivative were identified (see Table I). In addition, an anomalous Patterson difference map of the uranyl derivative showed a peak on a Harker section in a position identical to the peak in the Patterson difference map, and this peak was the highest peak in the section. The heavy atom positions determined were used in a phased refinement procedure (MLPHARE; Otwinowski *et al.*, 1991) to obtain phases. The MIR map was solvent flattened (with the envelope determined by the main-chain 'BONES' atoms traced automatically into the MIR map with the bones option of O; Jones *et al.*, 1991; M.Noble, personal communication). At a later stage when a reasonable model was available better phases were obtained by phase combination of the model phases with the experimentally determined phases. Difference Fourier maps using these improved phases were used to check the positions of the heavy atoms and to search for weak heavy atom sites. The interpretation of one weak platinum site was changed, new phases determined and a new experimental map calculated. This improved the map slightly. This new experimental map was solvent flattened as before and is shown here (Figure 1).

From the MIRAS map it was clear that there were three helices, a short connecting loop between helices 1 and 2 and a long loop between helices 2 and 3. Three perfect alpha helices with the LFB1 sequence were created using the model building option in FRODO (Jones, 1978). The boundaries of the helices were estimated, and the helices were made longer than the length of the observed electron density. There were several residues in the map that were clear and these were used as starting points. In the first helix, the electron density for the side-chain residues of Phe25 and Tyr28 provided the orientation of the first helix, and densities for residues Tyr75, Trp77 and Phe78 in the third helix defined the orientation of the third helix. The second helix was positioned based on the length of the observed short loop. The loop residues were built approximately into the observed electron density. Residues between 18 and 86 were included in the initial model. After a couple of cycles of XPLOR (Brünger *et al.*, 1987) refinement it was clear that the second helix was positioned incorrectly, and was shifted toward the C-terminus by one residue. The loop residues were then rebuilt using the BONES created with the auxiliary program of O. The side-chains of the molecule between residues 16 and 86 were examined carefully and side-chains with a poor fit were changed by using the O rotamer database, followed by some manual intervention. The refinement was continued by alternating cycles of model building (with O, checking the symmetry related contacts with FRODO), and automatic refinement (with XPLOR). Finally residues 13−15 and residues 87−89 were added to the model.

The final R-factor for the current model including residues 13−89 was 21.2% for data between 6 and 2.8 Å (2487 reflections) with bond deviations of 0.013 and angle deviations of 3.36 degrees. Examination of the Ramachandran plot showed three residues with ϕ/ψ angles outside the accepted region for non-glycine residues: Phe14, which has the ring in clear density, and Phe87 and Arg88 which are in regions of weak density.

Acknowledgements

We would like to thank Paul Tucker for data collection advice, Martin Noble for computing help in the unix environment, and C.Wolberger and C.Pabo for providing coordinates for the MATα2 and engrailed homeodomain−DNA complexes.

References

Brünger,A.T., Kuriyan,J. and Karplus,M. (1987) *Science*, **235**, 458−460.
Carson,M. (1987) *J. Mol. Graphics*, **5**, 103−106.
CCP4 package (1979) Daresbury Laboratory, UK.
Chouard,T., Blumenfeld,M., Bach,I., Vandekerckhove,J., Cereghini,S. and Yaniv,M. (1990) *Nucleic Acids Res.*, **18**, 5853−5863.
Finney,M. (1990) *Cell*, **60**, 5−6.
Frain,M., Swart,G., Monaci,P., Nicosia,A., Stämpfli,S., Frank,R. and Cortese,R. (1989) *Cell*, **59**, 145−157.
Jones,T.A. (1978) *J. Appl. Crystallogr.*, **11**, 268−272.
Jones,T.A., Zou,J.-Y., Cowan,S.W. and Kjeldgaard,M. (1991) *Acta Crystallogr.*, **A47**, 110−119.
Kabsch,W. (1988) *J. Appl. Crystallogr.*, **21**, 916−924.
Kissinger,C.R., Lui,B., Martin-Blanco,E., Kornberg,T.B. and Pabo,C.O. (1990) *Cell*, **63**, 579−590.
Laughon,A. (1991) *Biochemistry*, **30**, 11357−11367.
Nicosia,A., Monaci,P., Tomei,L., De Francesco,R., Nuzzo,M., Stunnenberg,H. and Cortese,R. (1990) *Cell*, **61**, 1225−1236.
Otting,G., Qian,Y.Q., Billeter,M., Müller,M., Affolter,M., Gehring,W.J. and Wüthrich,K. (1990) *EMBO J.*, **9**, 3085−3092.
Otwinowski,Z. (1991) In Wolf,W., Evans,P.R. and Leslie,A.G.W. (eds), *Isomorphous Replacement and Anomalous Scattering*. Daresbury Laboratory, UK, pp. 80−85.
Pastore,A., De Francesco,R., Barbato,G., Morelli,M.A.C., Motta,A. and Cortese,R. (1991) *Biochemistry*, **30**, 148−153.
Pastore,A., De Francesco,R., Morelli,M.A.C., Nalis,D. and Cortese,R. (1992) *Protein Engng*, **5**, 749−757.
Qian,Y.Q., Billeter,M., Otting,G., Müller,M., Gehring,W.J. and Wüthrich,K. (1989) *Cell*, **59**, 573−580.
Read,R.J. (1986) *Acta crystallogr.*, **A42**, 140−149.
Riddihough,G. (1992) *Nature*, **357**, 643−644.
Rosenfeld,M.G. (1991) *Genes Dev.*, **5**, 897−907.
Tomei,L., Cortese,R. and De Francesco,R. (1992) *EMBO J.*, **11**, 4119−4129.
Verrijzer,C.P., van Oosterhout,J.A.W.M., van Weperen,W.W. and van der Vliet,C.P. (1991) *EMBO J.*, **10**, 3007−3014.
Verrijzer,C.P., Alkema,M.J., van Weperen,W.W., Van Leeuwen,H.C. Strating,M.J.J. and van der Vliet,C.P. (1992) *EMBO J.*, **11**, 4993−5003.
Wolberger,C., Vershon,A.K., Liu,B., Johnson,A.D. and Pabo,C.O. (1991) *Cell*, **67**, 517−528.

Received on January 19, 1993; revised on March 3, 1993