21. Pai, E. F. *et al. Nature* **341,** 209–214 (1989).
22. Amor, J. C., Harrison, D. H., Kahn, R. A. & Ringe, D. *Nature* **372,** 704–708 (1994).
23. Scheffzek, K., Klebe, C., Fritz-Wolf, K., Kabsch, W. & Wittinghofer, A. *Nature* **374,** 378–381 (1995).
24. Verotti, A. C. *et al. EMBO J.* **11,** 2855–2862 (1992).
25. Yu, H. *et al. Cell* **76,** 933–945 (1994).
26. Omichinski, J. G. *et al. Science* **261,** 438–446 (1993).
27. Brünger, A. T. *X-PLOR Version 3.0 Manual* (Yale Univ. Press, New Haven, 1992).
28. Carson, M. *J. appl. Crystallogr.* **24,** 958–961 (1991).
29. Nicholls, A., Sharp, K. A. & Honig, B. *Proteins Struct. Funct. Genet.* **11,** 281–296 (1991).

ACKNOWLEDGEMENTS. This research was supported by a grant from the National Institute of General Medical Sciences. S.L.S. is an Investigator at the Howard Hughes Medical Institute.

Structural basis for DNA bending by the architectural transcription factor LEF-1

John J. Love*†, Xiang Li*, David A. Case*, Klaus Giese‡, Rudolf Grosschedl‡ & Peter E. Wright*§

* Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037, USA
† Department of Chemistry and Biochemistry, University of California, San Diego, California 92037, USA
‡ Howard Hughes Medical Institute and Departments of Microbiology and Biochemistry, University of California, San Francisco, California 94143, USA

LYMPHOID enhancer-binding factor (LEF-1) and the closely related T-cell factor 1 (TCF-1) are sequence-specific and cell-type-specific DNA-binding proteins that play important regulatory roles in organogenesis and thymocyte differentiation[1–5]. LEF-1 participates in regulation of the enhancer associated with the T cell receptor (TCR)-α gene by inducing a sharp bend in the DNA and facilitating interactions between Ets-1, PEBP2-α, and ATF/CREB transcription factors bound at sites flanking the LEF-1 site[1,2,6,7]. It seems that LEF-1 plays an architectural role in the assembly and function of this regulatory nucleoprotein complex[7,8]. LEF-1 recognizes a specific nucleotide sequence through a high-mobility-group (HMG) domain[1,2]. Proteins containing HMG domains bind DNA in the minor groove, bend the double helix[6,9,10], and recognize four-way junctions and other irregular DNA structures[9,11]. Here we report the solution structure of a complex of the LEF-1 HMG domain and adjacent basic region with its cognate DNA. The structure reveals the HMG domain bound in the widened minor groove of a markedly distorted and bent double helix. The basic region binds across the narrowed major groove and contributes to DNA recognition.

The HMG domain of mouse LEF-1, corresponding to residues 296–380 of the full-length protein plus an initiator methionine residue, was complexed with a 15-base-pair oligonucleotide duplex containing the optimal binding site from the TCR-α gene enhancer (Fig. 1). Structures were calculated from multi-dimensional nuclear magnetic resonance (NMR) data (Table 1). LEF-1 forms an intimate complex with DNA, wrapping around and almost completely encompassing a highly distorted double helix (Fig. 2). The central framework of the protein is formed by the characteristic L-shaped arrangement of three helices and an extended region seen previously in structures of HMG-1 and HMG-D domains in the absence of DNA[12–14]. Helices 1 and 2 (residues 8–24 and 28–43, respectively) form one arm of the L, helix 3 (residues 46–66) and the extended N-terminal region the other. The angle between the axes of helices 2 and 3 ($60° \pm 2°$) is more acute than in the HMG-1 and HMG-D domains[12–14]. Helix 3 terminates at Pro 67, after which the polypeptide kinks

TABLE 1 NMR structure statistics

(a) NMR constraints	
Protein	
Distance constraints	1,393
Intra-residue	712
Sequential ($\|i-j\|=1$)	265
Medium-range ($\|i-j\|\leq 4$)	246
Long-range ($\|i-j\|\geq 5$)	170
Dihedral angle constraints	42 ϕ constraints
DNA	
Distance constraints	325
Intra-residue	164
Sequential	156
Inter-strand	5
Protein–DNA	
Distance constraints	332
(*b*) Structure statistics	
DIANA target function	4.7 ± 1.7 (32 structures)
Statistics for 12 SA structures	
NOE violations	
Number >0.2 Å	12.3 ± 3.8
Maximum violations (Å)	0.40 ± 0.06
Angle violations	
Number of violations >5°	2
Mean constraint violation energy	46.6 ± 9.9 kcal mol^{-1}
Mean AMBER energy	$-2{,}488 \pm 37$ kcal mol^{-1}
Mean deviation from ideal covalent geometry	
Bond length	0.005 Å
Bond angles	1.9°
(c) R.m.s. deviations from the mean structure	
Protein backbone heavy atoms	
Helices	0.52 Å
Residues Lys 4–Tyr 66	0.60 Å
Residues Met 0–Lys 85	1.76 Å
(superimposed on N, Cα, C′, O)	
DNA heavy atoms	
C4–G12, C19–G27	0.75 Å
(superimposed on all DNA heavy atoms)	
Protein + DNA	
Helices + C4–G12, C19–G27	0.75 Å
(superimposed on N, Cα, C′, O + all DNA heavy atoms)	

Interproton distance restraints were determined from nuclear Overhauser enhancement spectroscopy (NOESY) cross-peak volumes. Distances obtained from 3D ^{13}C-separated NOESY spectra[24] were assigned van der Waals lower bounds and upper bounds of 3.0, 4.0 and 5.0 Å. Additional upper bounds of 3.7 and 6.0 Å were used for NOEs from the ^{15}N-separated NOESY spectrum[24]. For protein–DNA constraints, upper bounds of 3.0, 3.5, 4.5 and 6.0 Å were used with van der Waals lower bounds. Interproton distances within DNA were calculated from a $\tau = 50$ ms NOESY spectrum at 750 MHz. The upper bound distance was increased 20% over the calculated r^{-6}-weighted distance for 2′ and 2″ protons, and by 15% for all other protons. Several constraints were obtained from a $\tau =$ 120 ms double half-filtered NOESY spectrum at 600 MHz, for which the upper bound was increased by 25% over the calculated distance. DNA lower bounds were set to 2.0 Å for calculated distances $r \leqslant 3.0$ Å; to 2.2 Å for $3.0 < r \leqslant 3.5$ Å; 2.5 Å for $3.5 < r \leqslant 4.0$ Å; and 3.0 Å for $4.0 < r \leqslant 6.0$ Å. Initial protein structures were calculated with the program DIANA[25] using the redundant dihedral angle constraint (REDAC) strategy[26]. Beginning with 40 randomized starting structures, three REDAC cycles were applied, followed by a single cycle using only experimental restraints and a final cycle of minimization. Initial structures were calculated using only spectroscopically unambiguous distance restraints and hydrogen-bond constraints in the known helical regions; additional NOE restraints were added during subsequent rounds of structure calculation. The 32 structures with lowest target functions were subjected to three cycles of molecular-dynamics simulated annealing (SA) using the AMBER 4.1 package and all-atom force field[27], modified to reduce charges to 20%. Additional sets of non-NMR restraints were used in the initial stages of SA: helical hydrogen bond restraints on protein residues 11–24, 31–43 and 48–66, Watson–Crick hydrogen bond constraints on DNA, and distance constraints between sugar and base heavy atoms to maintain B-form DNA. The first cycle involved docking of the DIANA-refined protein structures, placed about 50 Å away from the binding site, onto B-form DNA. The two subsequent SA cycles began from the end points of the preceding cycles. In each cycle, the temperature was increased from 0 to 600 K over 6 ps, then cooled to 0 K over 14 ps with a temperature relaxation parameter of 1.2 ps^{-1}. The force constant for violation of the distance restraints was 20 kcal mol^{-1} Å^{-2}. During the first SA cycle, the experimental protein–DNA restraints were slowly turned on while the distance constraints initially imposed to keep the DNA in B-form were slowly turned off for the central 9 bp. Of the 32 DIANA starting structures, 29 led to acceptable docked structures and were subjected to subsequent SA cycles. During the second cycle, the remaining B-DNA constraints (on the outer 6 bp) and the hydrogen bond constraints in the protein helices were also turned off. The final SA cycle included only the NMR-based restraints, the Watson–Crick hydrogen bond constraints, and the AMBER potential functions. Of the 29 acceptable structures, 12 with no constraint violations greater than 0.5 Å were minimized for an additional 200 steps with the same constraints as for the final SA cycle.

§ To whom correspondence should be addressed

sharply and extends away from the core of the protein, making contacts predominantly with DNA. Indeed, the carboxy-terminal region beyond Arg 72 contacts only DNA.

An extensive hydrophobic core is enclosed by helices 1 and 2 and the amino-terminal region of helix 3 (Fig. 2*b*). The aromatic side chains of Tyr 12, Trp 40, Tyr 51, and Tyr 52 are packed edge-to-face in the junction between the three helices; Ala 8, Leu 11, Leu 43 and Leu 54 complete this region of the hydrophobic core. The core is extended along the interface between helix 1 and 2 by Met 13, Met 16, Val 20 and Val 21 in helix 1, and by Leu 36, Ile 35 and Ile 32 in helix 2. There is a small hydrophobic cluster involving the side chains of Ile 2, Tyr 66 and Trp 69 at the end of the third helix.

The DNA duplex binds to the concave surface of the LEF-1 domain and is bent severely towards the major groove (Fig. 2*c*). Cross-strand nuclear Overhauser enhancement (NOE) connectivities indicate retention of Watson–Crick base pairs. A wide and very shallow minor groove is formed between C5 · G26 and G9 · C22 that can accommodate helices 1 and 2 of LEF-1. The average minor groove width in this region is 11 Å with an average depth of ~1 Å; it is thus comparable in width to the minor groove of A-DNA but is even shallower. From A10 · T21 to C13 · G18 the minor groove becomes narrower (average width 5 Å) and deeper (average 4.6 Å). Bending and opening of the minor groove is accompanied by substantial narrowing (to <6 Å on average, cf. 11.7 Å for B-DNA) and deepening (to >7 Å on average) of the major groove from C5 · G26 to T8 · A23. The major groove subsequently becomes more like B-DNA in width although it remains rather deep. Bending occurs throughout the 9 base-pair (bp) recognition element, with an average total curvature summed over the 15-bp path of $117° \pm 10°$ in the 12 NMR structures. The curvature is not uniform but is most pronounced at the C5 · G26 : T6 · A25 step and the T7 · A24 : T8 · A23 step. The double helix is significantly underwound at the C5 · G26 : T6 · A25 and T6 · A25 : T7 · A24 steps, with average twist angles of $19.1° \pm 2.8°$ and $24.3° \pm 2.8°$, respectively, compared with 31.7° twist for regular B-DNA.

The LEF-1 HMG domain makes extensive and continuous contacts in the DNA minor groove from C4 · G27 to A11 · T20, encompassing the entire region implicated in binding by chemical footprinting[15,16] and mutagenesis[3,6]. The protein–DNA contacts are shown schematically in Fig. 3*b*. The most extensive interactions involve hydrophobic residues exposed on the concave surface formed by helices 1 and 2 and the extended N-terminal region. Further contacts are made by side chains from the C-terminal region where it crosses the minor groove (Fig. 3*a*). In particular, Tyr 75 is inserted into the narrowed region of the minor groove at the 5′ end of the TCR-α recognition element and appears to play a key role in DNA bending. The aromatic ring is in close contact with the ribose rings of G12 and T21 and the bases of G12, T20 and T21 (Fig. 3*d*). Interestingly, replacement of the G12 · C19 base pair by T · A has little effect on binding affinity but significantly diminishes the bending angle as measured by circular permutation assays (K.G. and R.G., unpublished observations). Polar side chains tend to contact the sugar-phosphate backbone rather than hydrogen bonding to bases. Exceptions are Asn 7 and Glu 28, which are located in the minor groove in close proximity to the 2-amino groups of G9 and G26/G27, respectively.

The side chain of Met 10 is partially inserted into the base stack between A23 and A24 (Fig. 3*c*), disrupting base stacking but not base pairing. Opening of the bases on the minor groove side leads to a large negative roll angle ($-52° \pm 6°$) and an increase in rise to 4.8 ± 0.3 Å at this step. Additional stabilizing contacts in this region are made by the side chain of Met 13, which is in close proximity to the A24 base, and the aromatic ring of Phe 9, which packs against the hydrophobic surface of the T8 ribose. These interactions probably play a major role in DNA bending through stabilization of the widened minor groove. An A · T to T · A transversion at position 24 results in

a

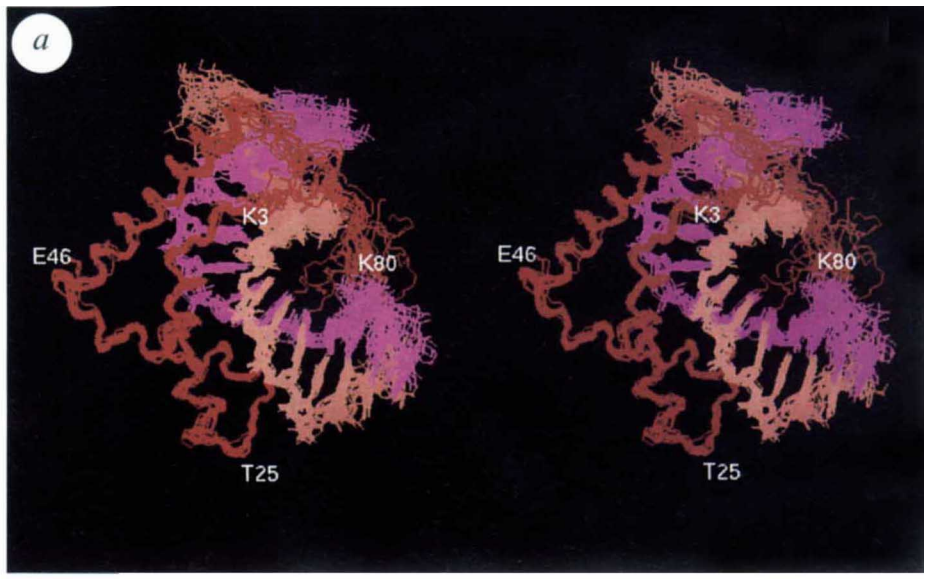
```
                1         10        20        30        40        50        60        70        80
LEF-1     MHIKKPLNAFMLYMKEMRANVVAESTLKESAAINQILGRRWHALSREEQAKYYELARKERQLHMQLYPGWSARDNYGKKKKRKREK
hSRY      RVKRPMNAFIVWSRDQRRKMALENPRMRNSEISKQLGYQWKMLTEAEKWPFFQEAQKLQAMHREKYPNYKYRPRRKAKMLP
STE-11    SVKRPLNSFMLYRRDRQAEI----PTSNHQSISRIIGQLWRNESAQVKKYYSDLSALERQKHMLENPEYKYTPKKRSTVRR
consensus ··K/RP·NAFML/IVφ····b·····T/P·····I/NS/AK··G··Wb·L····K/R··φ···A/S···b··H···Y/NP·YbY·····

HMG-D     KPKRPLSAYMLWLNSARESIKRENPGIKVTEVAKRGGELWRAMKD--KSEWEAKAAKAKDDYDRAVKEFEANGGSSAANGG
rHMG-1.B  APKRPPSAFFLFCSEYRPKIKGEHPGLSIGDVAKKLGEMWNNTAADDKQPYEKKAAKLKEKYEKDIAAYRAKGKPDAAKGV
hUBF.1    FPKKPLTPYFRFFMEKRAKYAKLHPEMSNLDLTKILSKKYKELPEKKKMKYIQDFQREKQEFERNLARFREDHPDQLIQNA
consensus ·PK/RP·SA/FY··φ····R·····P·····a··K··G··W··L/M····K··φ···A···K··Y·····φ·····
```

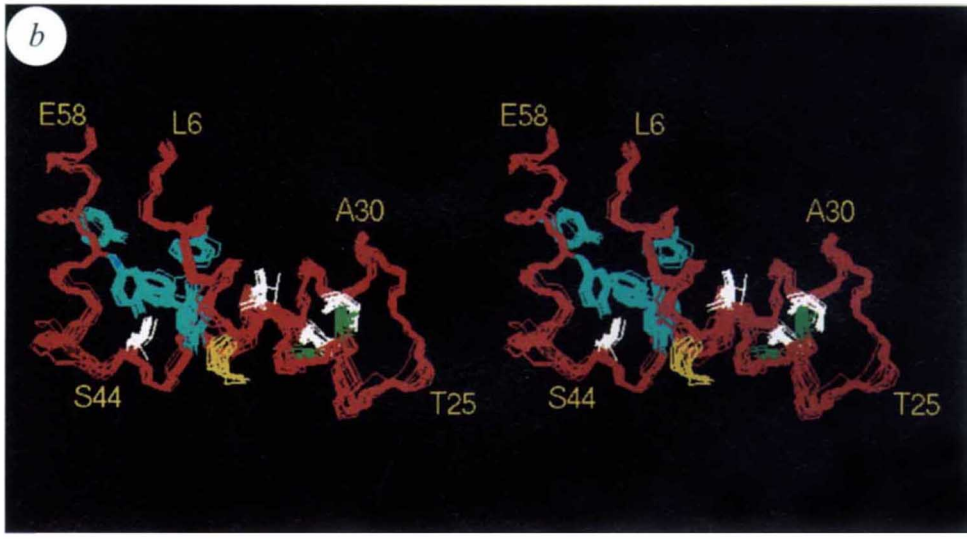
b

```
      1                  15
5′-CACCCTTTGAAGCTC-3′
3′-GTGGGAAACTTCGAG-5′
      30                 16
```

FIG. 1 *a*, Sequence alignments of representative members of the two subfamilies of HMG domain proteins (from ref. 8). Consensus sequences are shown for each subfamily. The exact numbering scheme of the murine LEF-1 HMG domain used here is indicated. The sequence differs from wild-type in that it retains the N-terminal initiator methionine, and Cys 24 has been mutated to serine to avoid potential problems during long-term NMR experiments. Neither modification has a discernible effect on the DNA binding affinity. *b*, Sequence and numbering scheme of the oligonucleotide duplex used for complex formation. The nine base pairs involved in LEF-1 binding are indicated in bold type.

METHODS. The LEF-1 HMG domain was expressed in *Escherichia coli* and labelled uniformly with ^{13}C and/or ^{15}N. The purified protein was shown to have the correct sequence by ion-spray mass spectrometry and was fully functional in DNA binding. DNA was synthesized by the phosphoramidite method and purified by reverse-phase HPLC on a C-18 column (0.1 M triethylamine–acetic acid buffer, pH 6.5, eluted with an acetonitrile/water gradient) before removal of the dimethoxytrityl protecting groups. After deprotection, the duplex was formed by mixing equimolar amounts of each oligonucleotide and annealed by slow cooling from 75 °C over 12 hours. HiTrap Q anion exchange was used for final purification. The complex was formed by titration of protein (in 10 mM KCl) into the DNA duplex (in 200 mM KCl, pH 5.5), using changes in the DNA imino proton spectrum to indicate 1:1 stoichiometry, and was desalted and concentrated using a Centricon to 1–2 mM in 10 mM KCl, pH 6.7, for NMR. Spectra were acquired at 303 K using Bruker DMX750, AMX600 and AMX500 spectrometers and processed with Felix 2.3 (Biosym Technologies). Protein resonance assignments (to be described elsewhere) were obtained by using extensive 2D and 3D NMR experiments (^{15}N-separated total correlation spectroscopy (TOCSY), ^{15}N-separated NOESY, constant time–HNCA, HNCACB, CBCA(CO)NH, HBHA(CO)NH, C(CO)NH–TOCSY, H(CCO)NH–TOCSY, HCCH–COSY, HCCH–TOCSY, ^{13}C heteronuclear multiple bond correlation (HMBC) and aromatic ^{13}C-^1H correlation spectroscopy). Distance restraints were derived from ^{13}C-separated NOESY spectra (mixing times $\tau = 60$ and 120 ms), a ^{15}N-separated NOESY spectrum[27] ($\tau = 120$ ms) and 750 MHz ^1H homonuclear NOESY spectra ($\tau = 50$ and 120 ms). Backbone ϕ dihedral angle constrains were obtained from $^3J_{HN,H\alpha}$ coupling constants measured in an HNCA-J experiment[28]. DNA resonances were assigned from 2D[^{15}N, ^{13}C]-double half-filtered NOESY and [^{15}N, ^{13}C]-half-filtered TOCSY spectra[29]. DNA distance restraints were derived from double half-filtered NOESY and ^{13}C-decoupled homonuclear ^1H–^1H NOESY spectra at 750 MHz ($\tau = 50$ and 120 ms). NOEs between the protein and the DNA were identified specifically in a 3D $^{13}C(\omega_1, \omega_2)$-edited, $^{12}C(\omega_3)$-selected NOESY experiment ($\tau = 120$ ms) (to be reported elsewhere).





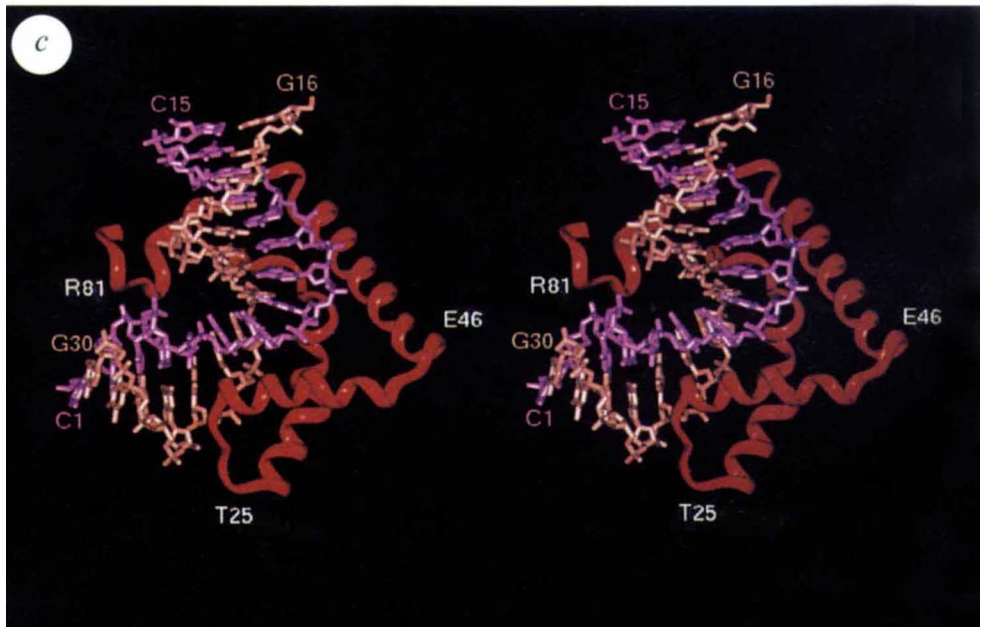
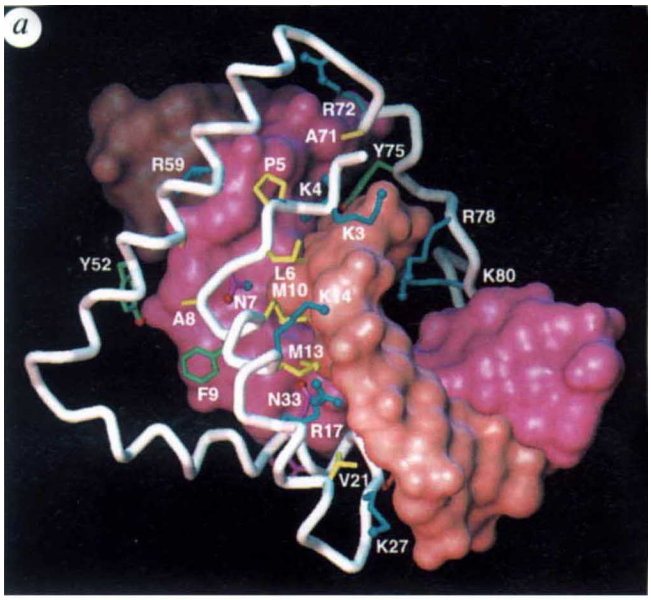
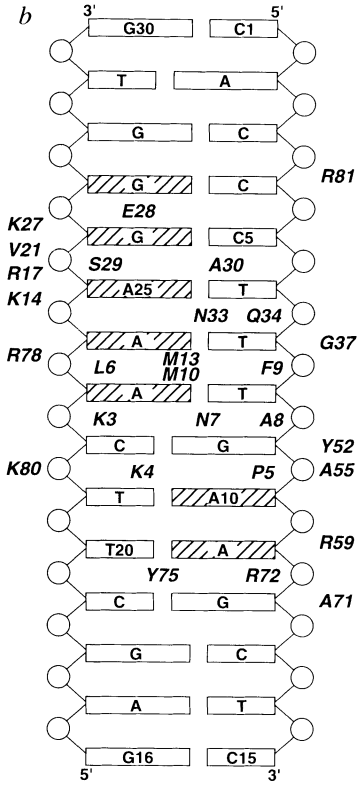


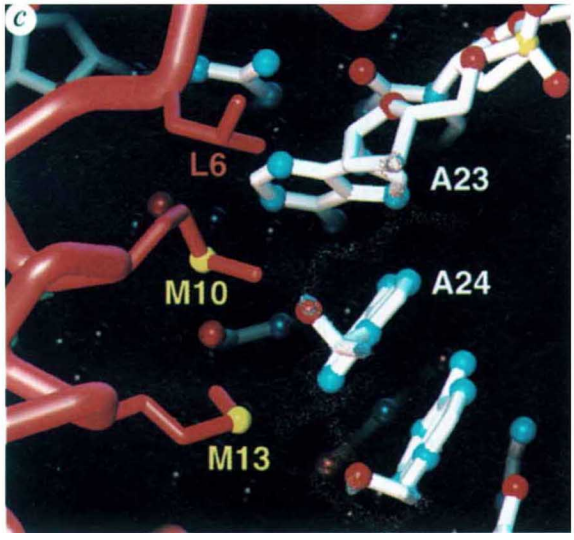
FIG. 2 *a*, Stereo view showing the best-fit superimposition of 12 structures of the LEF-1 HMG domain complexed with DNA. The structures were superimposed on all backbone heavy atoms of residues 6–66 and all heavy atoms of base pairs 4–12. The backbone N, Cα and C′ atoms of the protein are shown in red. All non-hydrogen atoms of the DNA are shown, coloured bright pink for nucleotides C1–C15, and salmon pink for G16–G30. The C1 · G30 base pair is at the lower right and the C15 · G16 base pair is at the top centre. *b*, Stereo view showing the packing interactions in the hydrophobic core of the LEF-1 HMG domain. The structures were superimposed on all backbone heavy atoms of residues 6–66 of the protein. The polypeptide backbone is shown in red, the aromatic side chains of Phe 9, Tyr 12, Trp 40, Tyr 51 and Tyr 52 in cyan, the side chains of Val 20 and Val 21 in green, Ile 32, Ile 35, Leu 36 and Leu 43 in white, and Met 16 in yellow. *c*, Stereo view of the complex from the opposite side of the DNA from that shown in *a*. The HMG domain is represented by the red tube drawn through the backbone N, Cα and C′ atoms. The pronounced bend in the DNA (117° ± 10°) is evident in this view. The DNA curvature and helical parameters were calculated with the program CURVES[30].

a severe loss of binding affinity[3]. Moreover, transversion of A25 to thymine substantially decreases the bending angle but has little effect on binding affinity (K.G. and R.G., unpublished observations). The side chain of Ile 68 in the sex-determining factor SRY, corresponding to Met 10 of LEF-1, also penetrates partly into the base stack and is required to induce specific DNA bending but is not required for binding to an intrinsically bent four-way junction[17].

In addition to the minor groove interations, residues from the highly basic C-terminal region bind across the narrowed major groove, making contacts primarily with the sugar-phosphate backbone (Figs 2*c* and 3*a*). Mutation of Lys 79 and Lys 80 or







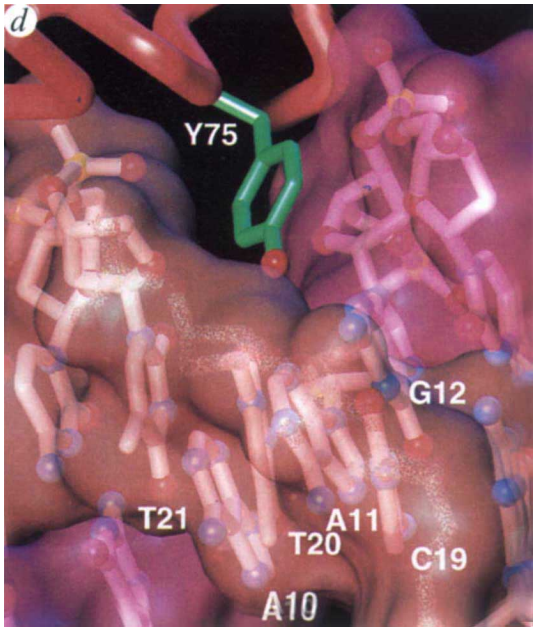


FIG. 3 *a*, View of the protein–DNA interactions in the minor groove. The polypeptide backbone is shown as a white tube through the Cα atoms. Many of the side chains that contact DNA are shown. The DNA is represented by a surface, coloured bright pink for the C1–C15 strand and salmon pink for the G16–G30 strand. *b*, Schematic representation of the protein–DNA contacts. The bases are represented as boxes. The hatched boxes represent bases implicated in DNA binding by methylation interference and diethylpyrocarbonate (DEPC) carbethoxylation interference footprinting[15,16]. *c*, Close-up view of the site of methionine insertion into the base stack. The protein backbone and side chains are represented as red tubes. The yellow balls represent the sulphur atoms of Met 10 and Met 13. The DNA is shown as a ball-and-stick model, and the bases of A23 and A24 are labelled. The sugar-phosphate backbone is ghosted to provide a clearer view of the interactions. *d*, Close-up view of the side chain of Tyr 75 (green) inserted into the minor groove. The surfaces of each DNA strand are shown (bright pink, C1–C15; salmon pink, G16–G30), together with the outline of the bases and the sugar-phosphate backbone. The backbone of one strand is ghosted for clarity.

truncation at Lys 77, Lys 78 or Lys 80 significantly impairs DNA binding (refs. 15, 18, 19, and results not shown), thus delineating the minimal domain for high-affinity binding as His 1–Arg 81. A polypeptide containing the LEF-1 HMG domain plus the basic residues bends DNA in circular permutation assays by ~130° (ref. 6), but only by 50° when truncated at Lys 77 (ref. 19). We conclude that the basic region C-terminal to the HMG domain contributes to both DNA recognition and bending. Sequence variation in this region may, in part, account for the difference in the previously estimated angle of DNA bending by LEF-1 (130°) and the mammalian testis-determining factor SRY (83°)[6,9].

The NMR structure forms a basis for understanding mutagenesis experiments on LEF-1 and provides insights into the likely function of several conserved amino acids. Mutation of residues 3, 4, 6, 10, 13, 37, 40, 51, 79 and 80 impairs or abrogates DNA binding (ref. 15 and unpublished observations). The majority of these mutations involve DNA contact residues; Trp 40 and Tyr 51 form the hydrophobic core at the junction of the three helices and are probably essential for stabilization of the folded structure. Several DNA contact residues are highly conserved (see ref. 8 and Fig. 1 for sequence alignments). Basic side chains are favoured at positions 3, 4, 17, 34 (Gln in LEF-1) and 59, which interact predominantly with the DNA backbone in the LEF-1 complex. An aromatic residue is strongly conserved at position 9, presumably to retain optimal ribose packing, and a bulky hydrophobic side chain is favoured at the site of insertion into the base stack (position 10). There is a marked preference for a small polar side chain at position 7, where there are base contacts in the LEF-1 complex. Finally, small side chains are conserved at positions 8, 37 and 55, with a strong preference for Ala, Gly and Ala, respectively. All three sites are in close proximity to the sugar-phosphate backbone and substitution by residues with bulky side chains would be expected to interfere with DNA binding. Indeed, replacement of Gly 37 by Ala in LEF-1 and by Arg in SRY impairs (LEF-1) or abrogates (SRY) DNA binding (ref. 20, and unpublished work).

HMG-domain proteins can be classified into two subfamilies[8]. Members of one subfamily, which includes the abundant HMG-1 and HMG-2, have multiple HMG domains, bind DNA with little or no specificity and are found in all cell types. Members of the other subfamily, which includes the lymphoid enhancer-binding factors LEF-1 and TCF-1 (refs 1–3), contain a single HMG domain, bind DNA sequence-specifically and are expressed in few cell types. Comparison of the structure of the LEF-1 HMG domain bound to DNA with structures of the unbound HMG-1 and HMG-D domains[12–14] leads to a highly plausible structural model to explain the different DNA-binding activities of the two subfamilies. In LEF-1, helix 3 is significantly shorter than in the HMG-1 and HMG-D domains, being limited in length by Pro 67, which induces a pronounced kink in the polypeptide chain and allows the C-terminal region to make extensive contacts with DNA, thereby greatly extending the binding surface. Contacts include insertion of Tyr 75 deep into the minor groove and interactions of the basic amino acids adjacent to the HMG domain with the sugar-phosphate backbone. Pro 67 is invariant in HMG-domain proteins that bind DNA with sequence-specificity but is absent from the subfamily that exhibits little or no specificity in binding DNA (see ref. 8 for sequences). Further, many members of the HMG-1 subfamily lack the highly basic region immediately adjacent to the HMG domain. However, these proteins retain the characteristic L-shaped fold and many of the DNA contact residues in the HMG core, which explains their ability to bind non-specifically but with high affinity to four-way junctions and other intrinsically bent DNA structures. Thus it appears that additional protein–DNA interactions, facilitated by the Pro 67-induced kink in the HMG domain and by the adjacent basic region, are important for high-affinity DNA binding and induction of DNA bending.

The structure of the LEF-1 HMG domain complexed with DNA provides new insights into the molecular basis for minor groove binding and protein-induced DNA bending. The LEF-1 HMG domain, the TATA-binding protein[21,22], and the purine repressor[23] induce large angle bends in DNA by a common mechanism, involving opening of the minor groove and partial insertion of one or more hydrophobic side chains into the base stack from the minor groove side. In each case the DNA is bent towards the major groove and away from the bound protein. DNA bending induced by proteins with single or multiple HMG domains appears to play an architectural role in the assembly of higher-order nucleoprotein complexes by juxtaposing non-adjacent factor-binding sites and/or positioning activation domains for interaction with other regulatory proteins[8].

Note added in proof: After submission of this paper, the structure of a DNA complex of human SRY has been reported[31]. The protein–DNA interactions observed for SRY are generally similar to those of LEF-1, but differ in several details, mainly at the C terminus. □

Received 31 May; accepted 14 July 1995.

1. Travis, A., Amsterdam, A., Belanger, C. & Grosschedl, R. *Genes Dev.* **5,** 880–894 (1991).
2. Waterman, M. L., Fischer, W. H. & Jones, K. A. *Genes Dev.* **5,** 656–669 (1991).
3. Oosterwegel, M. *et al. J. exp. Med.* **173,** 1133–1142 (1991).
4. Van Genderen, C. *et al. Genes Dev.* **8,** 2691–2703 (1994).
5. Verbeek, S. *et al. Nature* **374,** 70–74 (1995).
6. Giese, K., Cox, J. & Grosschedl, R. *Cell* **69,** 185–195 (1992).
7. Giese, K., Kingsley, C., Kirshner, J. R. & Grosschedl, R. *Genes Dev.* **9,** 995–1008 (1995).
8. Grosschedl, R., Giese, K. & Pagel, J. *Trends Genet.* **10,** 94–100 (1994).
9. Ferrari, S. *et al. EMBO J.* **11,** 4497–4506 (1992).
10. Paull, T. T., Haykinson, M. J. & Johnson, R. C. *Genes Dev.* **7,** 1521–1534 (1993).
11. Pil, P. M. & Lippard, S. J. *Science* **256,** 234–237 (1992).
12. Weir, H. M. *et al. EMBO J.* **12,** 1311–1319 (1993).
13. Read, C. M., Cary, P. D., Crane-Robinson, C., Driscoll, P. C. & Norman, D. G. *Nucleic Acids Res.* **21,** 3427–3436 (1993).
14. Jones, D. N. M. *et al. Structure* **2,** 609–627 (1994).
15. Giese, K., Amsterdam, A. & Grosschedl, R. *Genes Dev.* **5,** 2567–2578 (1991).
16. van de Wetering, M. & Clevers, H. *EMBO J.* **11,** 3039–3044 (1992).
17. Peters, R. *et al. Biochemistry* **34,** 4569–4576 (1995).
18. Carlsson, P., Waterman, M. L. & Jones, K. A. *Genes Dev.* **7,** 2418–2430 (1993).
19. Read, C. M., Cary, P. D., Preston, N. S., Lnenicek-Allen, M. & Crane-Robinson, C. *EMBO J.* **13,** 5639–5646 (1994).
20. Harley, V. R., Lovell-Badge, R. & Goodfellow, P. N. *Nucleic Acids Res.* **22,** 1500–1501 (1994).
21. Kim, Y., Geiger, J. H., Hahn, S. & Sigler, P. B. *Nature* **365,** 512–520 (1993).
22. Kim, J. L., Nikolov, D. B. & Burley, S. K. *Nature* **365,** 520–528 (1993).
23. Schumacher, M. A., Choi, K. Y., Zalkin, H. & Brennan, R. G. *Science* **266,** 763–770 (1994).
24. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A. & Bax, A. *J. Am. chem. Soc.* **111,** 1515–1517 (1989).
25. Güntert, P., Braun, W. & Wüthrich, K. *J. molec. Biol.* **217,** 517–530 (1991).
26. Güntert, P. & Wüthrich, K. *J. Biomol. NMR* **1,** 447–456 (1991).
27. Weiner, S. J., Kollman, P. A., Nguyen, D. T. & Case, D. A. *J. comput. Chem.* **7,** 230–252 (1986).
28. Seip, S., Balbach, J. & Kessler, H. *J. magn. Reson.* B**104,** 172–179 (1994).
29. Otting, G. & Wüthrich, K. *Q. Rev. Biophys.* **23,** 39–96 (1990).
30. Lavery, R. & Sklenár, V. *J. biomolec. Struct. Dyn.* **6,** 63–91 (1988).
31. Werner, M. H., Huth, J. R., Gronenborn, A. M. & Clore, G. M. *Cell* **81,** 705–714 (1995).

ACKNOWLEDGEMENTS. J.J.L. and X.L. contributed equally to this work. We thank J. Dyson, I. Radhakrishnan, W. Chazin, E. Komives and D. Hornby for valuable discussions and assistance; T. Macke for help with DNA structure calculations; M. Pique for help with computer graphics, and L. Tennant and M. Reymond for technical assistance. This work was supported by grants from the National Institutes of Health (P.E.W. and D.A.C.), and by the Howard Hughes Medical Institute (R.G.). X.L. received a postdoctoral fellowship from the Cancer Research Institute.