ments were done as described[24] using 34 pmol GTPase, 3.4 pmol of either Vav or Dbl, and, when appropriate, 3.5 pmol GST–Lck.

JNK assay and *in vivo* nucleotide labelling of GTPases. COS-7 cells were transfected using DEAE–dextran[12]. After transfection, cells were cultured for 48 h, serum-starved in phosphate-free DMEM medium for 18 h, labelled with [^{32}P]orthophosphate (100 μCi ml^{-1}) for 2 h, and disrupted in lysis buffer (50 mM Tris-HCl, pH 7.5, 20 mM $MgCl_2$, 150 mM NaCl, 0.5% Nonidet-P40, 1 mM sodium orthovanadate, 1 mM PMSF, 25 μg ml^{-1} leupeptin and 25 μg ml^{-1} aprotinin). Lysates were immunoprecipitated with an anti-AU5 monoclonal antibody (Babco) for 1 h and immunocomplexes recovered using Gamma-bind G-Sepharose beads (Pharmacia/LKB). Immunoprecipitates were washed three times in lysis buffer, twice in 50 mM Tris-HCl, pH 7.5, 20 mM $MgCl_2$, 500 mM NaCl, and finally resuspended in 1 M KH_2PO_4, 5 mM EDTA, pH 8.0. Bound nucleotides were released by heating, and fractionated using polyethyleneimine thin-layer chromatography plates[16]. JNK activity was determined using an *in vitro* immunocomplex assay[12,22].

Received 24 September; accepted 18 November 1996.

1. Katzav, S., Martin-Zanca, D. & Barbacid, M. *EMBO J.* **8,** 2283–2290 (1989).
2. Coppola, J., Bryant, S., Koda, T., Conway, D. & Barbacid, M. *Cell Growth Differ.* **2,** 95–105 (1991).
3. Bustelo, X. R., Ledbetter, J. F. & Barbacid, M. *Nature* **356,** 68–71 (1992).
4. Margolis, B. *et al. Nature* **356,** 71–74 (1992).
5. Bustelo, X. R. & Barbacid, M. *Science* **256,** 1196–1199 (1992).
6. Katzav, S., Sutherland, M., Packham, G., Yi, T. & Weiss, A. *J. Biol. Chem.* **269,** 32579–32585 (1994).
7. Wu, J., Motto, D. G., Koretzky, G. A. & Weiss, A. *Immunity* **4,** 593–602 (1996).
8. Zhang, R., Alt, F. W., Davidson, L., Orkin, S. H. & Swat, W. *Nature* **374,** 470–473 (1995).
9. Tarakhovski, A. *et al. Nature* **374,** 467–470 (1995).
10. Ridley, A. J. *Curr. Opin. Genet. Dev.* **5,** 24–30 (1995).
11. Boguski, M. S. & McCormick, F. *Nature* **366,** 643–654 (1993).
12. Crespo, P. *et al. Oncogene* **13,** 455–460 (1996).
13. Bolen, J. B. *Cell Growth Differ.* **2,** 409–414 (1991).
14. Bustelo, X. R., Suen, K. L., Leftheris, K., Meyers, C. A. & Barbacid, M. *Oncogene* **9,** 2405–2413 (1994).
15. Zheng, Y., Hart, M. J. & Cerione, R. A. *Methods Enzymol.* **256,** 77–84 (1995).
16. Laudanna, C., Campbell, J. J. & Butcher, E. C. *Science* **271,** 981–983 (1996).
17. Qiu, R. G., Chen, J., Kirn, D., McCormick, F. & Symons, M. *Nature* **374,** 457–459 (1995).
18. Qiu, R. G., Chen, J., McCormick, F. & Symons, M. *Proc. Natl Acad. Sci. USA* **92,** 11781–11785 (1995).
19. Wu, J., Katzav, S. & Weiss, A. *Mol. Cell. Biol.* **15,** 4337–4346 (1995).
20. Genot, E., Cleverley, S., Henning, S. & Cantrell, D. *EMBO J.* **15,** 3923–3933 (1996).
21. Spana, C., O'Rourke, E. C., Bolen, J. B. & Fargnoli, J. *Protein Expr. Purif.* **4,** 390–397 (1993).
22. Coso, O. A. *et al. Cell* **81,** 1137–1146 (1995).
23. Horii, Y., Beeler, J. F., Sakaguchi, K., Tachibana, M. & Miki, T. *EMBO J.* **13,** 4776–4786 (1994).
24. Shou, C., Farnsworth, C. L., Neel, B. G. & Feig, L. A. *Nature* **358,** 351–354 (1992).

ACKNOWLEDGEMENTS. We thank the Bristol-Myers Squibb Pharmaceutical Research Institute, and M. Barbacid in particular, for providing most of the reagents used in this study. We also thank J. Glaven and R. A. Cerione for purified Dbl, Y. Takai for smg21 GDS cDNA, and N. Reich and M. Dosil for their help and support.

CORRESPONDENCE and requests for materials should be addressed to X.R.B. (e-mail: xbustelo@path.som.sunysb.edu).

Unusual Rel-like architecture in the DNA-binding domain of the transcription factor NFATc

Scot A. Wolfe*, Pei Zhou*, Volker Dötsch†, Lin Chen*, Angie You*, Steffan N. Ho‡, Gerald R. Crabtree‡, Gerhard Wagner† & Gregory L. Verdine*

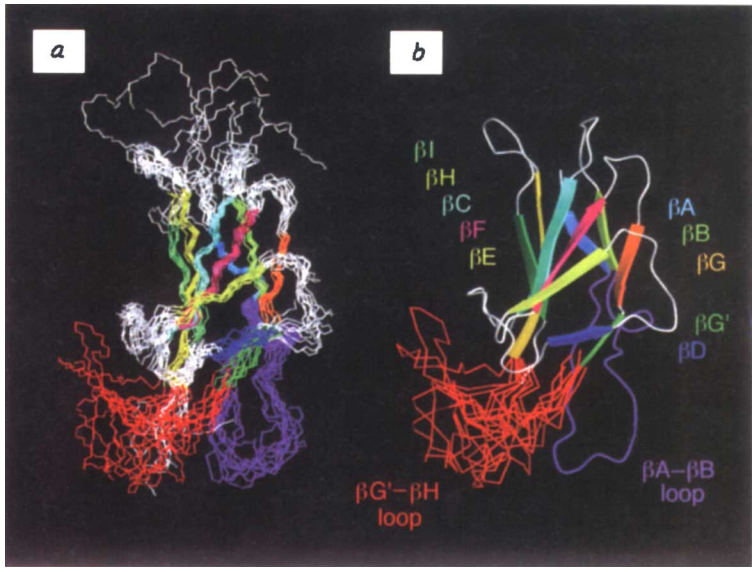
* Department of Chemistry and Chemical Biology, Harvard University, Cambridge, Massachusetts 02138, USA
† Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115, USA
‡ Departments of Developmental Biology and Pathology, Howard Hughes Medical Institute, Stanford University Medical School, Stanford, California 94305, USA

TRANSCRIPTION factors of the NFAT family regulate the production of effector proteins that coordinate the immune response[1]. The immunosuppressive drugs FK506 and cyclosporin A (CsA) act by blocking a Ca^{2+}-mediated signalling pathway leading to NFAT. Although FK506 and CsA have enabled human organs to be transplanted routinely, the toxic side-effects of these drugs limit their usage. This toxicity might be absent in antagonists that target NFAT directly. As a first step in the structure-based search for NFAT antagonists, we now report the identification and solution structure of a 20K domain of NFATc (NFATc-DBD) that is both necessary and sufficient to bind DNA and activate transcription cooperatively. Although the overall fold of the NFATc DNA-binding domain is related to that of NF-κB p50 (refs 2, 3), the two proteins use significantly different strategies for DNA recognition. On the basis of these results, we present a model for the cooperative complex formed between NFAT and the mitogenic transcription factor AP-1 on the interleukin-2 enhancer.

NFAT consists of two clearly separable functional domains. The N-terminal domain of NFATc, residues 1 to 415, controls its subcellular localization in response to Ca^{2+}-mobilization (C. R. Beals and G.R.C., manuscript submitted). Residues 416 to 591 of NFATc (NFATc-DBD), are necessary and sufficient for DNA binding and transcriptional activation on the distal antigen-receptor response element of the interleukin (IL)-2 enhancer (ARRE2)[4] (S.N.H. and G.R.C., unpublished results). *In vitro*, NFATc-DBD cooperates with AP-1 to form a tight, oriented complex on ARRE2 (ref. 5). Furthermore, NFATc-DBD alone binds specifically to ARRE2, albeit more weakly than full-length NFATc. NFATc-DBD thus represents the core functional domain of human NFATc, similar to that identified for murine NFATp (ref. 6).

The core structural motif of NFATc-DBD consists of a ten-stranded antiparallel β-barrel (Fig. 1*a*, *b*) assembled by the packing of three β-sheets. The two primary sheets (βIHCFE and βABG) form the core of the β-barrel, with the third sheet (βDG′) capping off one end. Devoid of α-helical structure, the remainder of NFATc-DBD is composed of loops, two of which are especially prominent (βA–βB and βG′–βH loops). As commonly observed for surface projections in NMR structures, the conformations of the two prominent loops in NFAT-DBD vary widely among the family of calculated structures shown in Fig. 1*a*. The origin of this behaviour, revealed through ^{15}N-relaxation measurements[7] (data not shown), was found to differ for the two loops. Whereas the βG′–βH loop has high conformational mobility throughout, the majority of the βA–βB loop is conformationally constrained, except at residues 28–35 near the tip (Fig. 1*c*), which appear to be only partially ordered.

NFAT family members share extensive sequence homology (~70%) over a ~300-amino-acid region comprising the entire DNA-binding domain, the N-terminal half of which corresponds to NFATc-DBD[8]. A potential sequence relationship between the DNA-binding domain of NFAT and members of the NF-κB/Rel protein family has been noted[9,10]; however, the extent of sequence similarity (<20%) falls below that considered to be statistically significant. Comparison of the solution structure of NFATc-DBD with the X-ray structure of NF-κB p50 DNA-binding domain (Fig. 2) reveals that the two share an essentially identical overall fold (Fig. 2*a*, *b*) and strand topology (Fig. 2*c*, *d*). Many of the residues that are invariant between NFAT family members and p50 (Fig. 1*c*, bold) form the hydrophobic core of the β-barrel; the packing interface between the βAB and βHI strands is nearly identical in the two classes of proteins. The sequence is also highly conserved throughout much of the βG′ and βH strands, because both faces of the small βG′H sheet are important: side chains on one face project inwards to establish the protein core, whereas those from the other face (Ile 123, Val 147) make stabilizing contacts to the lower end of the βA–βB loop (Val 38). Several other invariant residues, especially prolines 21, 46, 58, 110 and 156 in NFAT, are



c

```
                  βA                DNA binding loop                  βB
          10      •      20         •30       •40                    •
NFATc     MKDWQLPSHSGPYELRIEVQPKSH-HRAHYETEG-SRGAVKASAGG-----HPIVQLH
NFATp     PLEWPLSSQSGSYELRIEVQPKPH-HRAHYETEG-SRGAVKAPTGG-----HPVVQLH
NFAT3     PLDWPLPSQYEQLELRIEVQPRAH-HRAHYETEG-SRGAVKAAPGG-----HPVVKLL
NFAT4     PLDWPLPAHFGQCELKIEVQPKTH-HRAHYETEG-SRGAVKASTGG-----HPVVKLL
NF-κB p50 QPQMALPTADGPY-LQILEQPKQRGFRFRYVCEGPSHGGLPGASSEKNKKSYPQVKIC
                  50        60        70        80

          βC        70      βD        90      βE      100     βF
          •         •       •         •       •       •       •
GYLENEPLMLQLFIGTADDRLLRPHAFYQVHRITGKTVSTTSHEAILSNTKVLEIPLL
GYMENKPLGLQIFIGTADERILKPHAFYQVHRITGKTVTTTSYEKIVGNTKVLEIPLE
GYSE-KPLTLQMFIGTADERNLRPHAFYQVHRITGKMVATASYEAVVSGTKVLEMTLL
GYNE-KPINLQMFIGTADDRYLRPHAFYQVHRITGKTVATASQEIIIASTKVLEIPLL
NYVG--PAKVIVQLVTNG-KNIHLHAHSLVGK----------HCED----GICTVTAG
90        •       110       •       120

          βG      βG'     130       140     Insert region
110       •       •       •         •
PENSMRAVIDCAGILKLRNSDIELRKGETDIGRKNT----------
PKNNMRATIDCAGILKLRNADIELRKGETDIGRKNT----------
PENNMAANIDCAGILKLRNSDIELRKGETDIGRKNT----------
PENNMSASIDCAGILKLRNSDIELRKGETDIGRKNT----------
PKD-MVVGFANLGILHVTKKKVFETLEARMTEACIRGYNPGLLVHPDLAYLQAEGGGDR
130       150       160       180

                  βH              160     βI
                  •               •       •
                  RVRLVFRVHVPQPSGRTLS-L-QVASNPIECSQRS
                  RVRLVFRVHIPESSGRIVS-L-QTASNPIECSQRS
                  RVRLVFRVHVPQGGGKVVS-V-QAASVPIECSQRS
                  RVRLVFRVHIPQPSGKVLS-L-QIASIPVECSQRS
QLGDREKELIRQAALQQTKEMDLSVVRLMFTAFLPDSTGSFTRRLEPVVSDAIYDSKAP
190       210       240
```

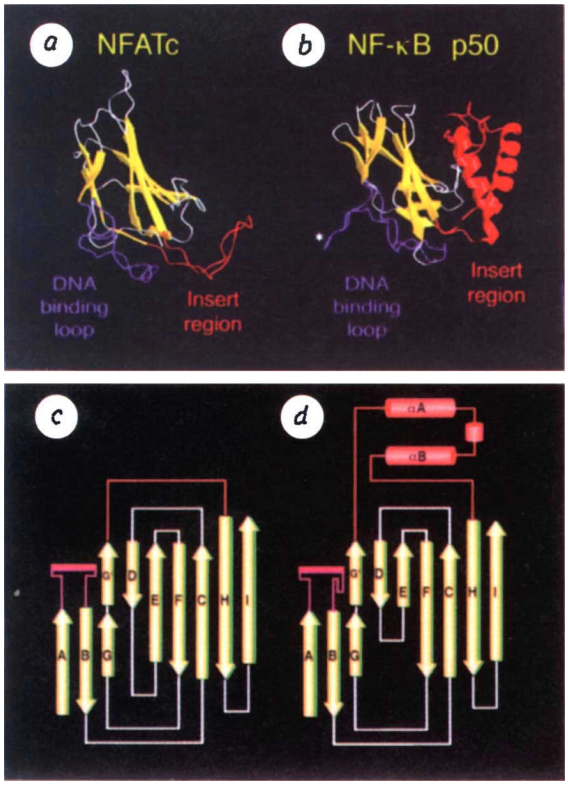


FIG. 1 *a*, *b*, Solution structure of NFATc-DBD. *a*, Backbone (N, Cα, C′) overlay of ten calculated structures of NFATc-DBD. The core domain consists of residues 13-172; residues 1–12 and 173–178 are highly disordered. The ten β-strands that comprise the β-barrel, as well as the two prominent loops, are colour-coded. *b*, MOLSCRIPT representation of the NFATc-DBD structure, viewed from the same perspective as in *a* and with the same colour coding. The highly mobile βG′–βH loop is shown in red as a family of Cα traces. Constituent strands that make up the three β-sheets of the structure are denoted accordingly in matching colours. Asterisk denotes a five-residue insert (residues 77–81) in the C-terminal end of the p50 DBD, which is without counterpart in the NFATc DBL (asterisk in Fig. 2*b*); this projection contains several basic residues that may contact phosphates outside the canonical DNA recognition site. *c*, Sequence and strand alignment of the four known NFAT isoforms[8–10,25,26] and human NF-κB p50 (refs 27,). β-strands are denoted above the sequence, colour-coded as in *a* and *b*. The numbering above the sequence corresponds to that of NFATc-DBD, which corresponds to that of human NFATc (ref. 10) minus 413. The italicized *MK* in the sequence of NFATc-DBD denotes amino acids added for purposes of overexpression, which are not present in NFATc itself.

FIG. 2 *a*, *b*, Comparison of the solution structure (*a*) of NFATc-DBD with the X-ray crystal structure (*b*) of the DNA-binding domain of NF-κB p50 (ref. 2). Yellow, β-strands; magenta, DNA-binding loop (DBL); red, insert region. Asterisk denotes a basic pentapeptide in the p50 DBL with no counterpart in NFATc-DBD (refer also to Fig. 1*c*); this segment in p50 is believed to make nonspecific contacts with the DNA backbone. *c*, *d*, Topological maps of NFATc-DBD (*c*) and the DNA-binding domain of NF-κB p50 (*d*). β-strands are shown as arrows, α-helices as cylinders. Colour scheme in *c*, *d* corresponds to that in *a*, *b*.

important in defining the junction between β-strands and their connecting loops. Many residues of the βA–βB loop are also identical in NFAT and p50 (see below).

The two structures do show significant differences, but these lie not in the strands that make up the β-barrel, but rather in the length and conformation of the segments connecting these strands. Most striking are the differences in the unusually long connector located between strands βG′ and βH, which in Rel proteins is known as the 'insert region'. In NF-κB p50, the 67 residues of the insert region form a compact helical bundle (Fig. 2*b*, red) which packs tightly against the immunoglobulin-like β-barrel[2,3], whereas the much smaller insert region of NFATc-DBD (21 residues) projects out into solution and is unstructured.

The sequence similarity between NFAT and Rel proteins is high throughout the βA–βB loop (residues 20–46 of NFATc-DBD; Fig. 1*c*). All but one of the sequence-specific contacts made

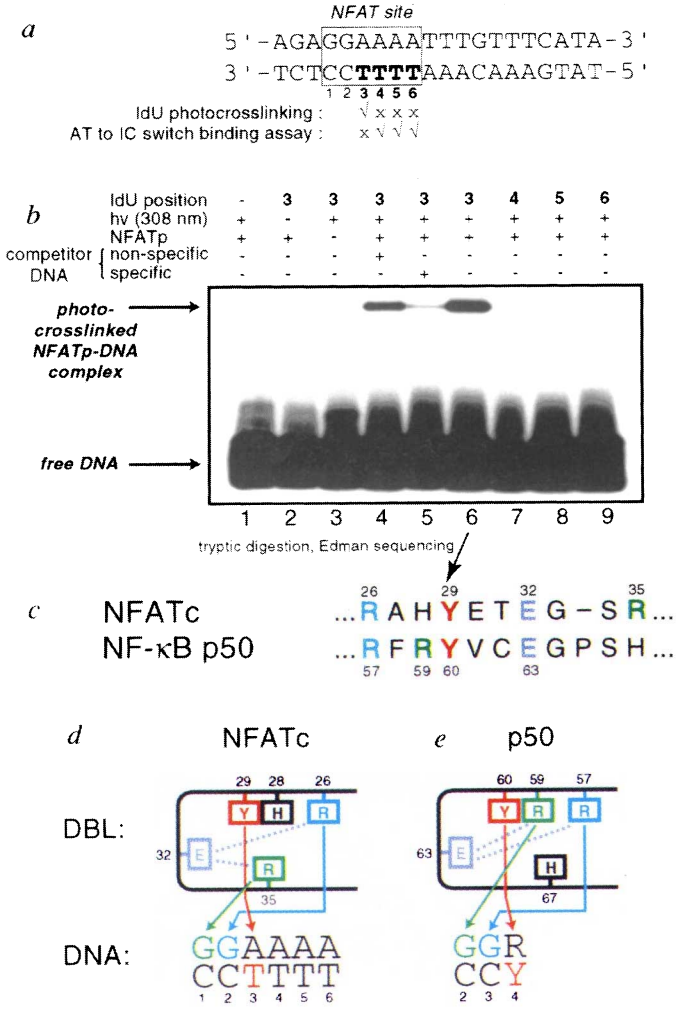


FIG. 3 *a*, Sequences of the ARRE2 oligonucleotides. Variants of the sequence shown containing individual dT residues substituted by 5-iododeoxyuridine (IdU) were used for *b*; variants having A·T base pairs switched to I·C were used in EMSA binding assays as a test for major-groove interactions. A tick indicates that photocrosslinking or binding (K_d < 20 nM) was significant; a cross indicates a lack of photocrosslinking or binding (K_d > 300 nM). *b*, Photocrosslinking of NFATp to ARRE2. ^{32}P end-labelled duplex oligonucleotides (top) containing individual IdU residues at the indicated positions within the NFAT site (see *a*) were irradiated with 308-nm UV light under various conditions. The products were separated by denaturing SDS–PAGE (below) and detected by phosphorimaging. A large-scale photocrosslinking reaction (lane 6) led to the identification of a covalent linkage between Tyr 29 and position 3, which is ordinarily occupied by a T (data not shown). *c*, Alignment of residues in the p50 DBL responsible for sequence-specific recognition of DNA (57–67) with the corresponding residues in NFATc (26–35). *d*, *e*, Diagram comparing the predicted major-groove contact from the DBL of NFATc-DBD (*d*) with those found in the X-ray structure of p50 (refs 2, 3) (*e*). Relevant portions of the respective binding sites are shown below the DBLs; arrows denote contact pairs. Bridging hydrogen bonds between a key conserved Glu (Glu 32 in NFATc-DBD; Glu 63 in p50) and Arg residues involved in DNA recognition are indicated by dotted lines. The red contact pair has been determined for both NFATp (*b*, *c*) and p50 (ref. 11) by photocrosslinking, and confirmed in the case of p50 by X-ray crystallography[2,3].

TABLE 1 Structural statistics for NFATc-DBD

Number of amino acids assigned (non-proline sequential assignments)	156 of 167
Total number of distance constraints	1,539
Number of effective distance constraints	1,087
Intramolecular	309
Sequential	321
Medium-range	71
Long-range	386
Number of dihedral angle constraints	331
Distance constraint violations >0.2 Å (per structure)	1.57 ± 1.31
Dihedral constraint violations >2.5° (per structure)	1.66 ± 1.21
X-PLOR potential energy (E_{total}, average per structure)	388 ± 15
R.m.s.d. to the mean for backbone heavy atoms, excluding poorly constrained regions (residues 13–26; 36–124; 144–172)	1.17 ± 0.22
R.m.s.d. to the mean for heavy atoms, excluding poorly constrained regions (residues 13–26; 36–124; 144–172)	1.60 ± 0.26
R.m.s.d. to the mean for backbone heavy atoms of all β-strands (residues 15–19; 47–51; 59–66; 81–83; 93–97; 102–108; 115–123; 146–155; 165–170)	0.70 ± 0.07
R.m.s.d. to the mean for heavy atoms of all β-strands	1.26 ± 0.11

All ϕ, ψ angles in the calculated structures lie within allowed regions of the Ramachandran plot.

by p50 involve residues on this DNA-binding loop (DBL)[2,3], and mutations within the corresponding segment of NFATp abrogate DNA binding[6]. To identify conclusively the segment of NFATp that contacts DNA in the major groove, we carried out photocrosslinking experiments using oligonucleotides substituted with 5-iododeoxyuridine (IdU) at various single sites within ARRE2 (Fig. 3*a*). Strong ultraviolet-dependent crosslinking occurred only when IdU was present at position T3 (Fig. 3*b*). Edman sequencing of the DNA-bound peptide derived from the tryptic digestion revealed that the peptide was crosslinked to DNA through the residue corresponding to Tyr 29 of NFATc-DBD (Fig. 3*c*). The analogous residue in p50, Tyr 60, undergoes highly efficient BrdU-mediated photocrosslinking to the fourth position of the NF-κB recognition site (Fig. 3*c*, *e*)[11]. In the crystal structure of p50, the aromatic ring of Tyr 60 is in direct van der Waals contact with the pyrimidine base to which it becomes photocrosslinked[2,3]. Taken together, these findings indicate that the aromatic ring of Tyr 29 in the NFATc DBL contacts T3 of its DNA-recognition site (Fig. 3*d*). Knowledge of this conserved contact pair thus permits alignment of the DBLs of NFATc and p50 on their respective DNA sites, and permits prediction of the major-groove DNA contacts used by NFATc (Fig. 3*d*, *e*).

Both the NFAT and p50 recognition sites contain a virtually invariant GG dinucleotide immediately adjacent to the cross-linked base pair (Fig. 3*d*, *e*). p50 recognizes these G residues through hydrogen-bonding interactions involving Arg 57 and Arg 59 (refs 2, 3). The former Arg is conserved in NFAT proteins (Arg 26 of NFATc-DBD), whereas the latter is replaced by a His (His 28) (Fig. 3*c*); Arg 26 is essential for DNA binding, whereas His 28 is not[6]. Nonetheless, the overall structure of the DBL in NFATc-DBD is similar to that of p50 (Fig. 2*a*, *b*), several key DNA contact residues in p50 are conserved in NFAT (Fig. 3*b*), and the critical interaction involving Tyr 29/60 is maintained in both. It is therefore likely that the DBLs of p50 and NFAT have similar DNA-bound conformations and make several conserved contacts, in addition to that involving Tyr 29. Thus, the side chain of His 28 in NFATc-DBD would be too short to contact G1, yet interference footprinting indicates that NFAT recognizes this guanine[12]. This issue was explained by modelling experiments, which suggested that the side chain of Arg 35 should ideally be positioned to contact G1. We thus propose that the DBLs of NFAT and p50 share a common scaffold for DNA recognition, and present the same side-chain functionality to the 5′ GGR sequence, but differ in the disposition of these residues along the scaffold.

Our analysis defines the sequence-specific interactions between NFAT and the 5′-half site of its recognition element (5′-**GGA**AAA), but leave unanswered the basis for specific recognition of the 3′-half site (5′-GGA**AAA**). Template-directed interference footprinting[13,14] indicates that NFAT makes strong major-groove contacts only to G1, G2 and T3 (C. Min and G.L.V., manuscript in preparation). Switching of individual A·T base-pairs in the NFAT site to I·C, which drastically alters major-groove functionality but leaves the minor groove unchanged[15], disrupted binding at position 3 but had little or no effect at positions 4–6 (Fig. 3*a*). Thus, whereas NFAT clearly recognizes the 5′-half site through sequence-specific contacts in the major groove, the protein does not discriminate the 3′-half site through major-groove contacts. These data suggest that NFAT recognizes the 3′-half site either through direct or water-mediated contacts to DNA bases in the minor groove, or by 'indirect readout' of the unique structure of poly(A) DNA. Consistent with either possibility, N3-methylation of any adenine within the 3′-half site interferes with NFAT binding.

To gain insight into the overall architecture of the core NFATc–AP-1–DNA complex, we superimposed the NMR structure of NFATc-DBD on the X-ray structure of the p50–DNA complex, and used the Tyr 29–T3 photocrosslinking observation to position NFATc-DBD relative to AP-1 on ARRE2. A remarkable feature of the resulting model (Fig. 4) is the location of the

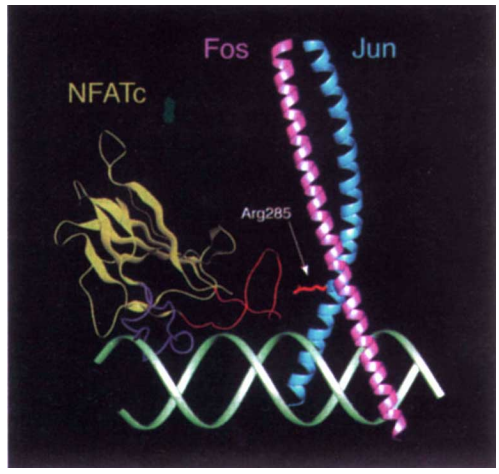


FIG. 4 Model of the core NFATc–AP-1–DNA structure. The solution structure of NFATc-DBD was positioned with respect to the DNA backbone (green) by least-squares superposition of the β-barrel (yellow) with that of a DNA-bound monomer of p50, and with respect to the ARRE2 sequence and AP-1 so as to align Tyr 29 and T3 (as defined by photocrosslinking). The X-ray coordinates of the AP-1 bZip domain[29] were used to place the heterodimer (pink, cFos; blue, cJun) over the non-consensus AP-1 site in ARRE2, using positional and orientational information derived from affinity cleaving experiments[30]. Arg 285 of c-Jun is indicated by an arrow (see text).

NFAT insert region, which ends up positioned directly over the minor groove of the 3′-half site. Although the structure of the insert region is not precisely defined from the NMR data, it is clear that residues of the insert region could readily move within reach of the minor-groove backbone and bases. Furthermore, the two stretches of basic residues at both ends of the insert region (residues 125–128 and 142–146) are reminiscent of minor-groove binding elements commonly used by proteins that recognize DNA by contacting both grooves[16–18]. Indeed, point mutations within the N-terminal basic stretch (Lys 125, Arg 127 or Asn 128, to Ala) disrupt binding to ARRE2 (L.-J. Sun, B. R. Peterson and G.L.V., manuscript in preparation). We conclude that the NFAT insert region is primarily responsible for sequence-specific recognition of the 3′-half site, and that the insert region of Rel proteins determines whether they bind DNA as monomers (NFAT family) or only as dimers (NF-κB/c-Rel family).

Whereas NFATc and NFATp exhibit similar specificity for ARRE2, NFAT3 and NFAT4 bind this site relatively weakly at physiological concentrations[8]. Although our discussion has focused on interactions common to all NFAT isoforms, we note that several residues of the DBL are not invariant, and these might help modulate DNA-binding specificity. In random site-selection experiments, NFATp shows a moderate preference for the two nucleotides immediately flanking the 5′-end of the site (positions −1 and −2; C. D. Vaughan, J. Triant and G.L.V., manuscript in preparation), raising the possibility that the characteristic specificity of each NFAT isoform is gained through differential recognition of positions outside the 5′-GGAAAA consensus sequence.

The model of the core NFATc–AP-1–DNA complex also provides insight into the structural basis for cooperativity between NFAT and AP-1. The locus of AP-1 interactions with NFATp maps to the junction of the leucine zipper and basic region[19]. Mutation of a single residue (Arg 285) in the c-Jun junction virtually abrogates cooperative interactions between NFATp and AP-1 (ref. 19). In our model, the c-Jun Arg 285 lies closest to the insert region of NFATc-DBD, suggesting that contacts between these loci mediate NFAT–AP-1 cooperativity; this holds true even if the insert structure is folded differently in the DNA-bound state from that in the solution structure of NFATc-DBD. In

agreement with this, point mutations in the NFATc insert region reduce cooperativity between NFAT and AP-1 but have negligible effect on NFATc–DNA interaction (Glu 132 and Thr 138 to Ala; L. J. Sun, B. R. Peterson and G.L.V., manuscript in preparation). Consistent with its critical role in DNA-binding and protein–protein interactions, the insert region is almost identical among the four known NFAT isoforms (Fig. 1*c*).

The structural basis for the assembly of general transcription complexes on the eukaryotic promoter is becoming clearer[20,21]. Our results start to show how multiprotein complexes assemble to form the enhancesomes that activate transcription. □

Methods

Overexpression and purification of NFATc. The NFATc DBD used for gel-shift and NMR studies comprises residues 416–591 of human NFATc[10], fused to the N-terminal Met1-Lys2 dipeptide. This fragment was cloned into the overexpression plasmid pLM1 with the second codon encoding lysine to improve the level of overexpression. Overexpression was performed in the *E. coli* BL21 DE3 pLysS (Novagen). Transformed cells were grown in LB broth plus 100 mg l^{-1} ampicillin and 25 mg l^{-1} chloramphenicol to an A_{600} of 0.6–0.8, and then induced with 225 mg l^{-1} isopropylthio-β-D-galactoside (IPTG) at 30 °C for 5–6 h. NFATc was purified by cation-exchange chromatography over S-Sepharose (Pharmacia) followed by anion-exchange chromatography over Q-Sepharose (Pharmacia). ^{15}N- or $^{15}N/^{13}C$-labelled NFATc for NMR spectroscopy was prepared as described, except that 1 ml of cells grown in LBAC was used to inoculate 1 litre of M9 minimal medium containing 1g $^{15}NH_4Cl$ and 2g $^{12}C_6$- or $^{13}C_6$-glucose (Cambridge Isotope Labs). ^{14}N-reverse-labelled NFATc was produced according to ref. 22.

IdU photocrosslinking. IdU-containing oligonucleotides were synthesized by conventional solid-phase chemistry using an IdU phosphoramidite (Peninsula Labs). The IdU-containing strand was ^{32}P-end-labelled and annealed with its complement. Labelled duplex DNA was mixed with a recombinant fragment of murine NFATp containing the DNA-binding domain[9], in the presence or absence of a 100-fold molar excess of unlabelled specific or nonspecific competitor oligonucleotide. The mixture was allowed to equilibrate on ice for 1 h, then irradiated on a UV transilluminator (308 nm) for 1 h. The products were loaded directly onto an SDS–polyacrylamide gel (Fig. 3*a*). The DNA-linked amino acid was identified as described[11].

Binding of NFAT to ARRE2 oligonucleotides having A·T base-pairs switched to I·C. ARRE2 oligonucleotides containing individual A·T to I·C substitutions at positions 3 to 6 (Fig. 3*a*) were synthesized by conventional solid-phase chemistry. The K_d values of NFATc for these oligonucleotides were determined by non-denaturing electrophoretic gel mobility assays. Binding reactions (2 mM HEPES, pH 7.5, 90 mM KCl, 10% glycerol, 0.1% Nonidet-P40, 1 mM DTT, 5 μg ml^{-1} dI·dC, 0.1 mg ml^{-1} BSA) containing various concentrations of NFATc-RHR and 5′ ^{32}P-end-labelled probe (1 nM) were allowed to equilibrate at room temperature for 1 h before loading onto a 6% polyacrylamide gel (0.5 × TBE). Bands representing free and bound probe were quantified using a Fuji phosphorimager.

NMR spectroscopy. NMR experiments[23] on NFATc-DBD were done with sample concentrations of 1–2 mM in 25 mM deuterated Tris-acetate (Cambridge Isotopes), pH 6.5, 100 mM KCl, 5 mM deuterated DTT. Experiments were done at 27 °C on a Bruker DMX-500 instrument equipped with a triple-resonance probe and triple axis gradients, unless otherwise noted. Assignments for backbone ^1H, ^{13}C and ^{15}N chemical shifts were obtained using a combination of the following 3D experiments: HNCA, HN(CO)CA, HNCO, HN(CA)CO, HCACO, HN(CA)HA. These 3D experiments, in combination with amino-acid assignments determined from comparison of the ^1H-^{15}N HSQCs of the ^{14}N reverse-labelled samples[22] and from the Cα and Cβ chemical shifts determined from an HC(C)H-TOCSY (400 MHz Varian Unity Plus), allowed the sequential assignment of 156 of the 167 non-proline amino acids. $^3J_{HNHA}$ coupling constants were obtained from an HNHA experiment and used for backbone dihedral angle constraints. Stereospecific assignments for the methyl groups of 8 of 12 valines and 9 of 19 leucines were established based on the crosspeak pattern in a ^1H–^{13}C HSQC (without carbon decoupling) of a 10% ^{13}C-labelled sample. H^N exchange data were acquired at five time points (1, 2, 12, 36 and 72 h) using a ^1H–^{15}N HSQC experiment. NOE constraints for structural calculations were obtained from 4 NOESY experiments: 50 ms ^{15}N separated 3D NOESY-HSQC (750 MHz Varian Unity Plus), 50 ms 2D NOESY in D_2O (750 MHz Varian Unity Plus), 50 ms 2D NOESY in 90% H_2O/10% D_2O, 100 ms ^{13}C-separated 3D NOESY-HSQC in D_2O (400 MHz Varian Unity Plus).

Structural calculations. Structural calculations were performed based on 1,087 effective non-hydrogen-bond constraints using the program DIANA[24] with three REDAC cycles. Structures with target functions below 10 were further refined by simulated annealing using X-PLOR (version 3.1). 35 structures were produced with no NOE violations >0.4 Å and no dihedral angle violations >5.0 deg. This family of structures was used for the calculation of statistics, and the ten best structures were chosen for Fig. 1*a*.

Received 2 October; accepted 12 November 1996.

1. Crabtree, G. R. & Clipstone, N. A. *Annu. Rev. Biochem.* **63,** 1045–1083 (1994).
2. Müller, C. W., Rey, F. A., Sodeoka, M., Verdine, G. L. & Harrison, S. C. *Nature* **373,** 301–317 (1995).
3. Ghosh, G., Van Duyne, G. D., Ghosh, S. & Sigler, P. B. *Nature* **373,** 303–310 (1995).
4. Durand, D. B. *et al. Mol. Cell. Biol.* **8,** 1715–1724 (1988).
5. Wolfe, S. A. thesis (Harvard Univ., 1996).
6. Jain, J., Burgeon, E., Badalian, T. M., Hogan, P. G. & Rao, A. *J. Biol. Chem.* **270,** 4138–4145 (1995).
7. Peng, J. W. & Wagner, G. *J. Mag. Res.* **98,** 308–332 (1992).
8. Ho, S. N. *et al. J. Biol. Chem.* **270,** 19898–19907 (1995).
9. McCaffrey, P. G. *et al. Science* **262,** 750–754 (1993).
10. Northrop, J. P. *et al. Nature* **369,** 497–502 (1994).
11. Liu, J., Sodeoka, M., Lane, W. S. & Verdine, G. L. *Proc. Natl Acad. Sci. USA* **91,** 908–912 (1994).
12. Chen, L. thesis (Harvard Univ., 1994).
13. Mascareñas, J. L., Hayashibara, K. C. & Verdine, G. L. *J. Am. Chem. Soc.* **115,** 373–374 (1993).
14. Min, C. H., Cushing, T. D. & Verdine, G. L. *J. Am. Chem. Soc.* **118,** 6116–6120 (1996).
15. Starr, D. B. & Hawley, D. K. *Cell* **67,** 1231–1240 (1991).
16. Pabo, C. O. & Sauer, R. T. *Annu. Rev. Biochem.* **61,** 1053–1095 (1992).
17. Blackwell, T. K., Bowerman, B., Priess, J. R. & Weintraub, H. *Science* **266,** 621–628 (1994).
18. Feng, J., Johnso, R. C. & Dickerson, R. E. *Science* **263,** 348–355 (1994).
19. Peterson, B. R., Sun, L. J. & Verdine, G. L. *Proc. Natl Acad. Sci. USA* **93,** 13671–13676 (1996).
20. Nikolov, D. B. *et al. Nature* **377,** 119–128 (1995).
21. Tan, S., Hunziker, Y., Sargent, D. F. & Richmond, T. J. *Nature* **381,** 127–134 (1996).
22. Shortle, D. *J. Mag. Res.* B **105,** 88–90 (1994).
23. *NMR of Proteins* (eds Clore, G. M. & Gronenborn, A. M.) (CRC Press, Ann Arbor, 1993).
24. Güntert, P., Braun, W. & Wüthrich, K. *J. Mol. Biol.* **217,** 517–530 (1991).
25. Hoey, T., Sun, Y.-L., Williamson, K. & Xu, X. *Immunity* **2,** 461–472 (1995).
26. Masuda, E. S. *et al. Mol. Cell. Biol.* **15,** 2697–2706 (1995).
27. Ghosh, S. *et al. Cell* **62,** 1019–1029 (1990).
28. Kieran, M. *et al. Cell* **62,** 1007–1018 (1990).
29. Glover, J. N. M. & Harrison, S. C. *Nature* **373,** 257–261 (1994).
30. Chen, L. *et al. Curr. Biol.* **5,** 882–889 (1995).

ACKNOWLEDGEMENTS. S.A.W. and P.Z. contributed equally to this work. We thank M. Oliger for assistance in processing relaxation data, S. Park for help with construction of overexpressing strains, and L. J. Sun, B. Peterson, C. Min, D. Erlanson and M. Chytil for discussions and for sharing unpublished data. This work was supported by grants from the NIH and NSF, Howard Hughes Medical Institute, and Hoffmann-LaRoche Institute of Chemistry and Medicine. NMR spectrometers were purchased with funding from the NSF and NIH. S.A.W. was partially supported by a predoctoral fellowship from the NSF; V.D. was supported by a postdoctoral fellowship from EMBO.

CORRESPONDENCE and requests for coordinates and materials should be addressed to G.L.V. (e-mail: verdine@chemistry.harvard.edu).

Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA

Alexey Bochkarev, Richard A. Pfuetzner, Aled M. Edwards & Lori Frappier

Institute for Molecular Biology and Biotechnology, Cancer Research Group, McMaster University, 1200 Main St. W., Hamilton, Ontario L8N 3Z5, Canada

The single-stranded-DNA-binding proteins (SSBs) are essential for DNA function in prokaryotic and eukaryotic cells, mitochondria, phages and viruses[1,2]. The structures of four SSBs have been solved[3–7], but the molecular details of the interaction of SSBs with DNA remain speculative. We report here the crystal structure at 2.4 Å resolution of the single-stranded-DNA-binding domain of human replication protein A (RPA) bound to DNA. Replication protein A is a heterotrimeric SSB that is highly conserved in eukaryotes. The largest subunit, RPA70, binds to single-stranded (ss)DNA[8,9] and mediates interactions with many cellular and viral proteins[10]. The DNA-binding domain, which lies in the middle of RPA70, comprises two structurally homologous subdomains oriented in tandem. The ssDNA lies in a channel that extends from one subdomain to the other. The structure of each