Structure of the DNA-binding domains from NFAT, Fos and Jun bound specifically to DNA

Lin Chen*, J. N. Mark Glover*†, Patrick G. Hogan‡, Anjana Rao‡ & Stephen C. Harrison*†

** Department of Molecular and Cellular Biology and † Howard Hughes Medical Institute, Harvard University, Cambridge, Massachusetts 02138, USA*
‡ Center for Blood Research, Harvard Medical School, Boston, Massachusetts 02115, USA

The nuclear factor of activated T cells (NFAT) and the AP-1 heterodimer, Fos–Jun, cooperatively bind a composite DNA site and synergistically activate the expression of many immune-response genes. A 2.7-Å-resolution crystal structure of the DNA-binding domains of NFAT, Fos and Jun, in a quaternary complex with a DNA fragment containing the distal antigen-receptor response element from the interleukin-2 gene promoter, shows an extended interface between NFAT and AP-1, facilitated by the bending of Fos and DNA. The tight association of the three proteins on DNA creates a continuous groove for the recognition of 15 base pairs.

Engagement of the T-cell antigen receptor (TCR) by presented antigen induces a cascade of transcriptional responses[1,2]. The synthesis of interleukin (IL)-2, an important early event, depends on a regulatory region that extends about 300 base pairs upstream of the transcriptional start site, and several signalling pathways, from the TCR and from co-stimulatory receptors, converge on this promoter[3,4]. The transcription factor known as nuclear factor of activated T cells (NFAT) binds to several sites within the regulatory region of the IL-2 gene and other inducible genes[5,6]. Its properties have been analysed in particular detail, because its activation is inhibited by the main immunosuppressive drugs in current clinical use, cyclosporin A and FK506 (ref. 7). There are four known mammalian NFAT genes[8–11], which encode proteins of about 700–1,100 amino-acid residues. There are two contiguous, well-conserved regions—the so-called NFAT homology region (NHR, 300 residues) and the DNA-binding region (about 270 residues)—sandwiched between more variable 'transactivation domains' at either end of the polypeptide chain[6]. The DNA-binding region bears a distant sequence similarity to the Rel homology region (RHR) of Rel-family transcription factors such as NF-κB[12,13].

In resting T cells, NFAT proteins are largely cytoplasmic. Stimulation of the TCR leads to increased intracellular Ca^{2+}, dephosphorylation of sites in the NHR by calcineurin (a calcium-activated phosphatase), and ultimately nuclear import and increased DNA binding of NFAT[14–18]. These steps are blocked by the calcineurin inhibitors CsA and FK506[7,14]. Full response at many NFAT sites requires concomitant activation of members of the AP-1 transcription-factor family[6,20] (the AP-1 partner is referred to as the 'nuclear component' of NFAT in much of the original literature on this factor)[1,2]. There is a discernible AP-1 (Fos–Jun) site immediately downstream of most NFAT sites in the promoter region of IL-2 and other cytokine genes, but it usually differs at more than one position from the consensus AP-1 sequence[6]. Indeed, relatively weak binding at their respective sites by NFAT and AP-1 individually, together with significant cooperative interaction between the two on composite sites, reflects the functional requirement for both transcription factors[21,22].

The DNA-binding domains of NFAT (the RHR) and AP-1 (the bZIP elements of the Fos and Jun partners) are necessary and sufficient for cooperative association on DNA[23]. Thus interactions between the adjacently bound proteins must be sufficiently favourable that cooperative binding (and the functional counterpart, synergistic activation) can occur at concentrations below the K_d of NFAT or AP-1 alone. To see directly the contact that two such unrelated structures might make, and to work out the mechanism for this part of the convergent, multipathway T-cell response, we have crystallized a quaternary complex containing the bZIP elements of Fos and Jun, the RHR of human NFAT1 (ref. 12), and a DNA fragment that has the sequence of the distal antigen-receptor response element (ARRE2) from the murine IL-2 promoter[23]. The structure reveals an extensive interface between NFAT and AP-1, facilitated by bending of the Fos helix and the DNA. The DNA-binding surfaces of NFAT and AP-1 together form a continuous groove for the recognition of 15 base pairs.

Overview of the structure

The NFAT DNA-binding region (residues 399–678) and the AP-1 bZIP elements form a tight complex with each other and with the DNA (Figs 1 and 2). The NFAT fragment is very similar in the folded structures of its domains to the RHR of NF-κB p50 (refs 24, 25; see also the published nuclear magnetic resonance (NMR) structure of the amino-terminal domain[26]), but the conformation of the linker is quite different, so that the cleft between the two domains closes up around the Fos–Jun heterodimer. Two structural adaptations ensure that AP-1 has an extended interaction with NFAT (Fig. 3a, b): the coiled coil inclines by about 15° toward the 5′ end of the site, and the DNA itself bends by about 20° at the centre of the AP-1 site. The Fos–Jun heterodimer is oriented specifically so that Jun binds to the half of the AP-1 site close to NFAT, as demonstrated in studies of the complex in solution[27].

NFAT

The two domains of the NFAT DNA-binding region identify it clearly as a Rel homologue, despite the lack of extensive sequence conservation (Fig. 1a). Both domains have immunoglobulin-like folds.

The N-terminal domain contains, in addition to the strands of a 'standard' s-type immunoglobulin module[28], two additional strands between C and C′. We designate these strands X and Y, to preserve the conventional immunoglobulin-strand nomenclature for the rest of the domain. The ABE sheet and the GFCC′Y sheet form the two faces of the β-sandwich, and the two-strand sheet XE′ and the AB loop seal the bottom. The E′F loop contains a short α-helix. In p50, this loop is the site of an approximately 68-residue, largely α-helical 'insert' into the minimal RHR. In NFAT, the loop is even shorter (21 residues) than in p65 or c-Rel, and forms an important part of the interface between NFAT and AP-1.

The carboxy-terminal domain has the h-type immunoglobulin

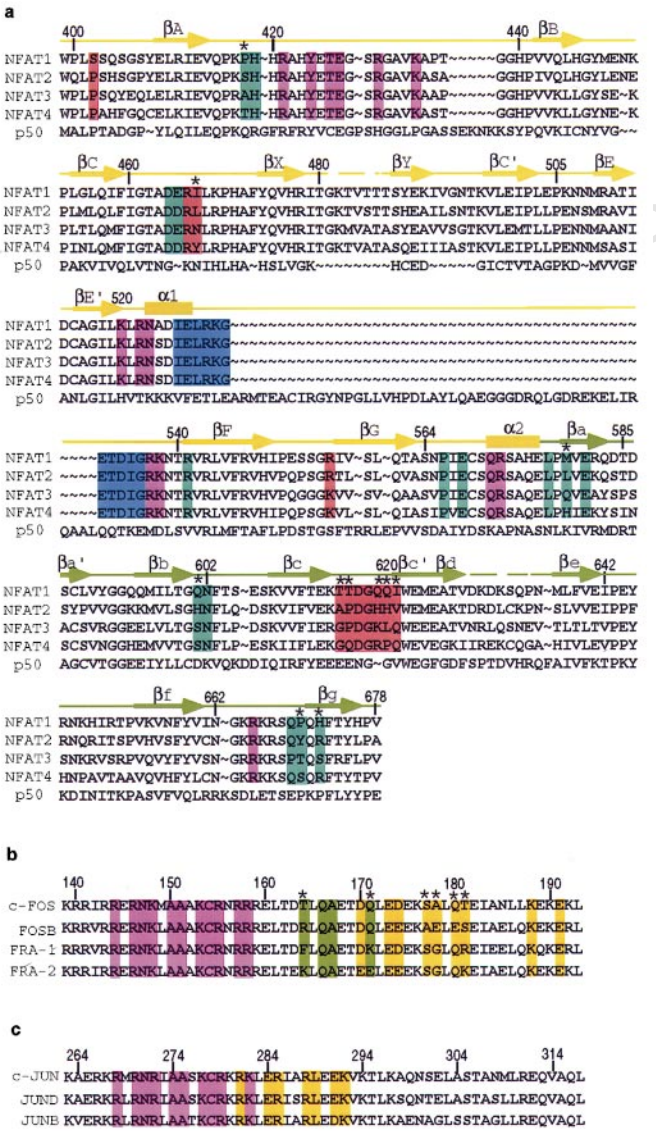


Figure 1 Amino-acid sequences of NFAT, NF-κB p50 and the bZIP regions of Fos and Jun. **a**, Amino-acid sequences of the DNA-binding regions of the four NFAT family members (human) and of NF-κB p50. The alignment with NF-κB (p50) is based on the structures. Secondary-structure assignments from our model of NFAT1 are shown as coloured bars (α-helices) and arrows (β-strands) above the aligned sequences. Broken lines denote disordered regions. Our lettering adheres essentially that of ref. 25, except that in the N-terminal domain we have used X and Y instead of H and I, to emphasize that these strands are insertions into the basic immunoglobulin design. The p50 lettering of ref. 24 simply assigned alphabetically sequential letters to the strands, thus departing from immunoglobulin convention after strand C. We suggest that the present convention be used for this entire superfamily of transcription factors. The colour code for the N-terminal domain is yellow and for the C-terminal domain green. The coloured blocks show residues that participate in contacts to DNA (magenta), Fos (red) or Jun (blue) or in interdomain contacts within NFAT (light green). Asterisks indicate variable contacting residues among family members. **b**, **c**, Amino-acid sequences of the bZIP regions of Fos (**b**) and Jun (**c**) family members (human). Coloured blocks show contacts with DNA (magenta) or with NFAT (yellow or green, according to the domain contacted). Asterisks indicate variable contacting residues among family members.

fold[28]. It folds back against the N-terminal domain (Fig. 2). The corresponding domain of Rel family members, which it closely resembles, would direct dimerization, but NFAT is clearly a monomer both in solution[11,27] and in this complex. The outer surface of the abed sheet is not significantly more polar in NFAT, but it is oriented in such a way that a partner with a Rel-like dimer contact could not bind DNA without major reconfiguration of the interdomain linker. This linker includes an α-helical segment that abuts the DNA backbone and contributes to base-pair recognition, and a linker rearrangement would therefore perturb both base-pair and backbone contacts. In the unbound state, however, it is likely that the linker is flexible. The inter-domain contacts do not appear sufficiently tight to maintain the relative orientation of the domains in the free protein. They may, however, contribute to the stability of the particular domain arrangement seen in the DNA-bound structure.

AP-1

The segments of Fos and Jun are each single α-helices, and the heterodimer forks across the DNA major groove, just as in the DNA complexes of other bZIP proteins[29–31]. The asymmetry of the coiled coil that was seen in our earlier crystal structure of a Fos–Jun–DNA complex[31] is evident here as well. The leucine-zipper part of Jun curls markedly around the straighter zipper element of Fos. The coiled coils in the two structures superpose well (Cα root-mean-square deviation 0.42 Å), showing that the asymmetry is a property of the Fos–Jun contacts and not a consequence of interactions with NFAT or other molecules in the crystal lattice. The Fos–Jun–DNA crystals studied previously contained two distinct complexes in the crystallographic asymmetric unit. These differed not only in their preferred orientation for the heterodimer on the DNA site but also in the way the two α-helices were bent in the fork between the basic regions and the coiled coil. In this quaternary complex with NFAT (Fig. 3), the entire Jun segment in the present structure can be superposed, together with the leucine zipper of Fos, on the corresponding part of one of the two complexes in the earlier structure (Fig. 3c), but the fork region of Fos flexes by about 15° around a pivot near Glu 160, so the axis of the coiled coil leans markedly away from a perpendicular to the axis of the DNA (Fig. 3a, b). Examination of the crystal packing shows that lattice contacts cannot be the cause of the distortion. Differences in the AP-1 site DNA conformation between Fos–Jun complexes with and without NFAT allow the local major-groove interactions of Fos and Jun to remain similar,

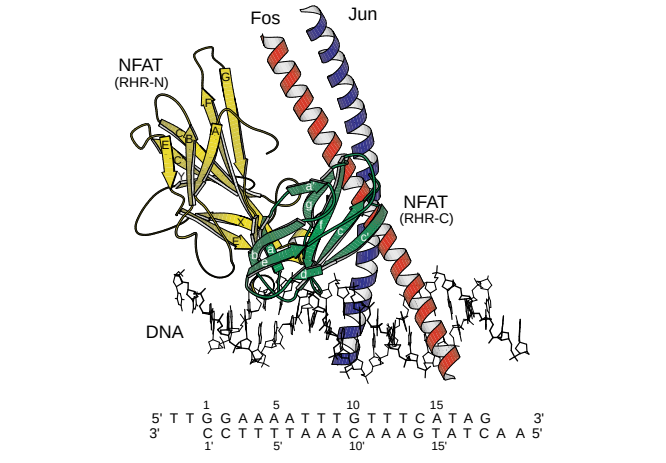


Figure 2 The NFAT–AP-1–DNA complex (drawn with MOLSCRIPT[48]). The colours here are used throughout the rest of the illustrations: the N-terminal domain (RHR-N) of the NFAT RHR is yellow; the C-terminal domain (RHR-C) is green; Fos is red; Jun is blue. Strands and helices are labelled; the lettering corresponds to that used for the secondary-structure indicators in Fig. 1. The ARRE2 nucleotide sequence in the crystals is shown below the figure, together with the numbering scheme used in the text.

even after such a significant adjustment in one of the partners (see below).

DNA conformation in the complex

The entire DNA conforms to a B-type helix. The complementary two-base overhang at either end of the fragment establishes continuity of the double helix from one unit cell to the next. There is a bend of about 20° at the centre of the AP-1 site, towards the major groove and hence towards the coiled-coil side of the Fos–Jun heterodimer (Fig. 3a). Rather than narrowing the major groove, however, this bend appears to have the effect of displacing the sugar-phosphate backbone on the Jun side of the complex towards NFAT, so the minor groove in the vicinity of phosphate P7 is narrowed instead. This narrowing extends over the entire A:T-rich segment from base pairs 3 to 8. The reverse bend, required for continuity of DNA in the crystal, occurs in the based-paired overhang.

DNA recognition

Overall. The combined DNA-binding surfaces of NFAT, Fos and Jun contact the extended ARRE2 composite site continuously (Fig. 4). Bases in the core NFAT site (GGAAAA) and the AP-1 site (TGTTTCA) interact with NFAT and Fos–Jun, respectively, in the major groove, and the intervening spacer sequence (TT) interacts with an NFAT arginine residue in the minor groove (Fig. 4b, c).

NFAT recognition. The AB loop and the α-helical segment of the linker supply side chains that contact bases in the major groove. The AB loop is the homologue of the 'recognition loop' in NF-κB p50 (refs 24–26), and its conformation and contacts closely resemble those of its p50 counterpart (Fig. 4b, c). Arg 430 and Arg 421 participate in bidentate hydrogen-bond interactions with O6 and N7 of Gua 1 and Gua 2, respectively. The guanidinium groups of the arginines stack upon each other, following the stack of the corresponding guanine bases. In NFAT, the side-chain amide group of Gln 571 extends this stack and engages in bidentate hydrogen-bonding with Ade 3 (refs 32, 33). The GGA sequence thus recognized seems to be the essential element of an NFAT site. The remaining A:T base pairs in the GGAAAA sequence, which appears to be well conserved among many NFAT sites[6], are specified by van der Waals contacts with their thymine methyl groups. The ring of Tyr 424 is the partner for thymines 3′ and 4′; the aliphatic part of Arg 572 is the partner for thymines 5′ and 6′. Both residues are also anchored by hydrogen bonds to the DNA backbone.

The E′F loop bridges across the DNA backbone from the major to the minor groove. Arg 537, at the end of this loop, dips into the minor groove and donates a hydrogen bond to the carbonyl of Thy 8. The segment of DNA backbone between Thy 3′ and Ade 7′ is sandwiched by the linker α-helix in the major groove and the E′F loop in the minor groove and by the AB loop at one end and the fg loop at the other end. A strong ethylation interference footprint covers this region[34].

AP-1 recognition. The seven base-pair AP-1 site in the ARRE2 deviates from consensus at two positions, but many of the interactions between basic-region side chains and base-pair functional groups are the same as those found in the ternary Fos–Jun–DNA complexes[31] (Fig. 4b). The bend in the fork segment of Fos, the invariance of the zipper interactions between Fos and Jun, and the overall invariance of the Jun conformation might then have been expected to generate a displacement of the Jun basic region with respect to the DNA major groove. But underwinding and bending at the centre of the AP-1 site together restore the register of base pairs with residues in Jun (Fig. 4b), even though the orientation of the Jun basic region in the major groove is nearly perpendicular to the DNA axis. The conserved asparagine residue in the Jun basic region (Asn 271) accepts at least one hydrogen bond (from N4 of Cyt 10′) present in all the other GCN4 and Fos–Jun complexes, and Ala 274 also has a conserved contact with the methyl group of Thy 9.

Comparison with biochemical and mutational data. The base-pair and sugar-phosphate contacts shown in Fig. 4b readily account for most of the observed effects of chemical modifications in the DNA and mutational substitution in NFAT[12,26,34,35]. Methylation interference from Ade 3 to Ade 8′ can probably be ascribed to the minor-groove compression seen in the structure and to the contact between Arg 537 and Thy 8. Because the interference is seen also in the binary NFAT–DNA complex, the compression may be present even with NFAT alone. As explained above, minor groove narrowing

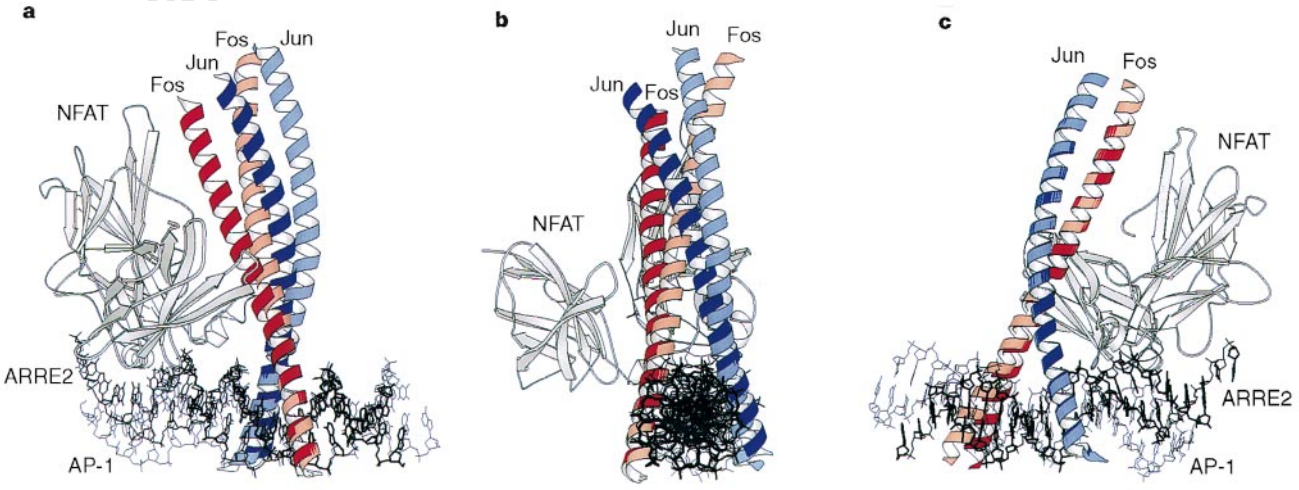


Figure 3 Superpositions of the NFAT–Fos–Jun–ARRE2 quaternary complex on the Fos–Jun–AP-1 site ternary complex (complex I of ref. 31). These comparisons (drawn with MOLSCRIPT) illustrate that bending of Fos and of DNA together allow the formation of the observed NFAT–Fos–Jun interface. **a**, The ternary (Fos–Jun–AP-1 site) and quaternary (NFAT–Fos–Jun–ARRE2) complexes have been superposed so that the basic region of Fos and the segment of DNA that it contacts are optimally aligned. Lighter colours and thin lines indicate the ternary complex[31]; darker colours and thick lines show the quaternary complex; NFAT is shown grey. The DNA bend can be seen by comparing the two superposed structures. **b**, End-on view of the same superposition. **c**, The two structures have been superposed so that the leucine zippers are optimally aligned. This superposition, when compared with the superposition used in **a** and **b**, shows that when shifting from the conformation in the ternary complex to that in the quaternary complex, the rest of the Fos–Jun–DNA complex can be imagined to pivot about the fork segment of Fos.

accompanies DNA bending, which is in turn essential for the interaction between AP-1 and NFAT (see below). The narrow minor groove is probably stabilized in part by insertion of Arg 537 opposite Thy 8 and Arg 665 between Ade 7′ and Gua 10. Various NFAT–AP-1 composite sites exhibit a strong bias between positions 4 and 8 for A:T or T:A base pairs, and minor-groove narrowing occurs more readily in A:T-rich sequences. This apparent 'indirect' specification of A:T or T:A between base pairs 4 and 8 complements the 'direct' preference for T at 4′, 5′, and 6′, owing to van der Waals contacts from Tyr 424 and Arg 572.

Different NFAT family members have been reported to have distinct DNA-binding specificity and affinity[6]. The crystal structure shows, however, that residues involved in DNA recognition are highly conserved among family members (Fig. 1a).

The interface between NFAT and AP-1

The N-terminal domain of the NFAT RHR has a large contact surface with Fos and Jun, extending over much of the length of the fork and the coiled coil and merging with the protein–DNA interface (Fig. 5). The NFAT C-terminal domain also has a small contact with Fos. The interactions are primarily polar, creating a network of hydrogen bonds and salt links that connect Fos, Jun, NFAT and the DNA backbone (Fig. 5b). The contacts of the E′F loop with Jun and of the CX loop and N-terminal segment with Fos are each centred

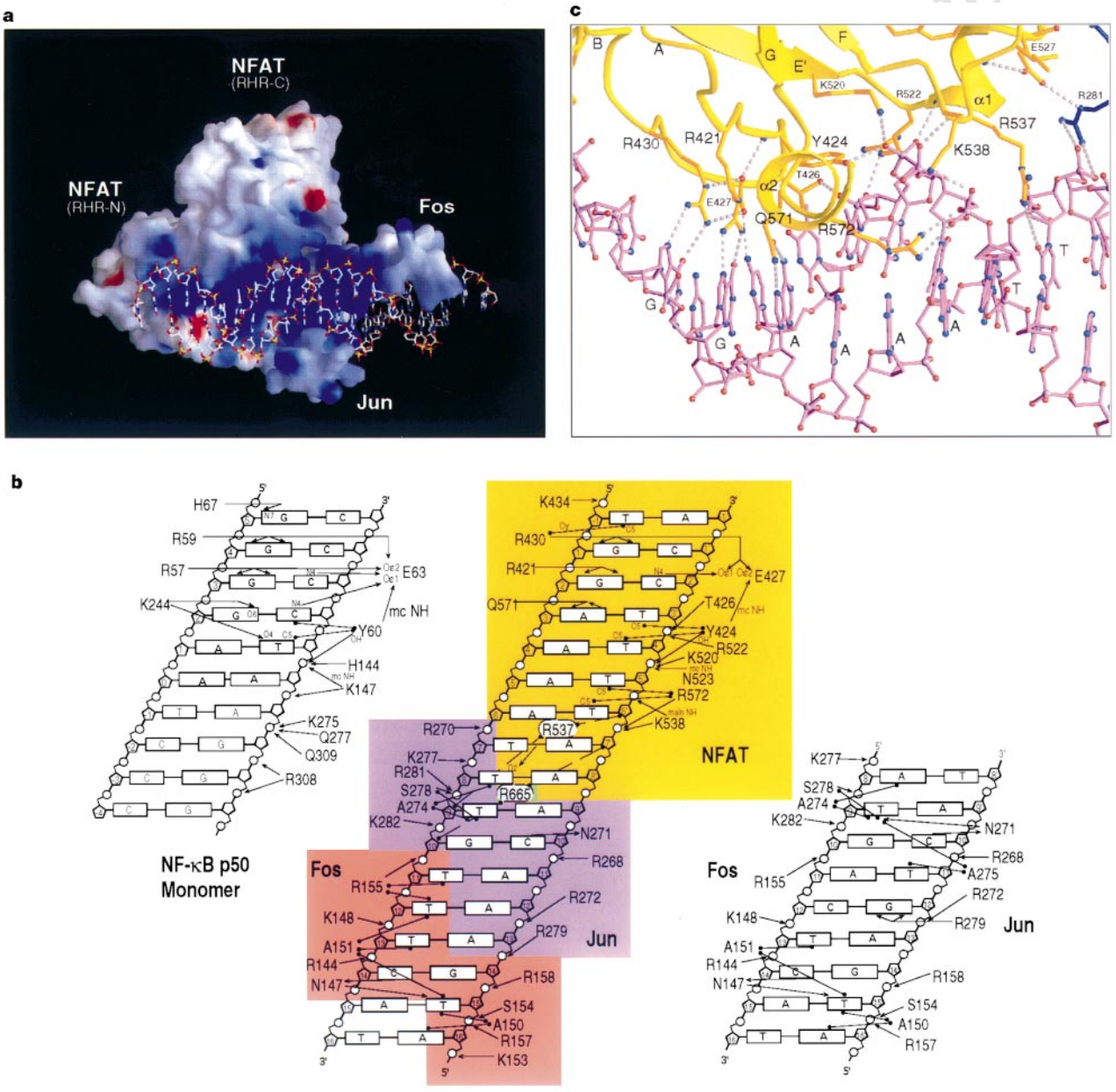
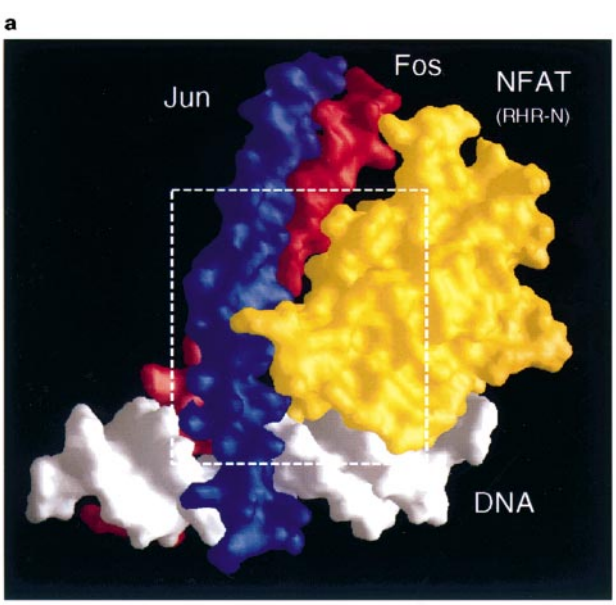


Figure 4 DNA recognition in the NFAT–AP-1–ARRE2 complex. **a**, View of the composite binding groove generated by the DNA-facing surfaces of NFAT, Fos and Jun. The colour scheme indicates electrostatic potential (blue, positive; red, negative) as calculated by the program GRASP[49]. **b**, Schematic diagram of detailed protein–DNA contacts, compared with similar contacts in the NF-κB p50–DNA complex (upper left)[24] and in the ternary Fos–Jun–AP-1 site complex (lower right)[31]. Lines with arrowheads, hydrogen bonds; lines with filled circles at either end, van der Waals contacts; solid lines, major-groove and sugar-phosphate interactions; dashed lines, minor-groove interactions. The coloured blocks indicate which protein is involved. **c**, Detailed view (using Ribbons[50]) of NFAT interactions in the major groove, showing the roles of Arg 421, Glu 427, Arg 430, Gln 571 and Arg 572 in specifying base pairs. Arg 430 is supported by salt links to Glu 427, which also accepts hydrogen bonds from the main-chain NH of Tyr 424 and from N4 of Cyt 2′. The view is essentially the same as that in Fig. 2.

on a small non-polar patch, surrounded by an extended cluster of hydrogen bonds. The hydrogen-bond networks communicate with each other, but there is not a fully continuous van der Waals complementarity of the AP-1 and NFAT surfaces between one patch and the next. That is, the extended interaction region contains small hydrated pockets between these major foci of contact (Fig. 5a).

The E′F loop contacts the fork segment of Jun close to the DNA surface. The loop extends away from a DNA phosphate (P4′) as a short α-helix (α1), and it returns toward the adjacent phosphate (P5′) as a strand (Fig. 5b). This loop appears to be a segment of structural variability in Rel-family proteins, and our structure suggests that this variability corresponds to different partners for



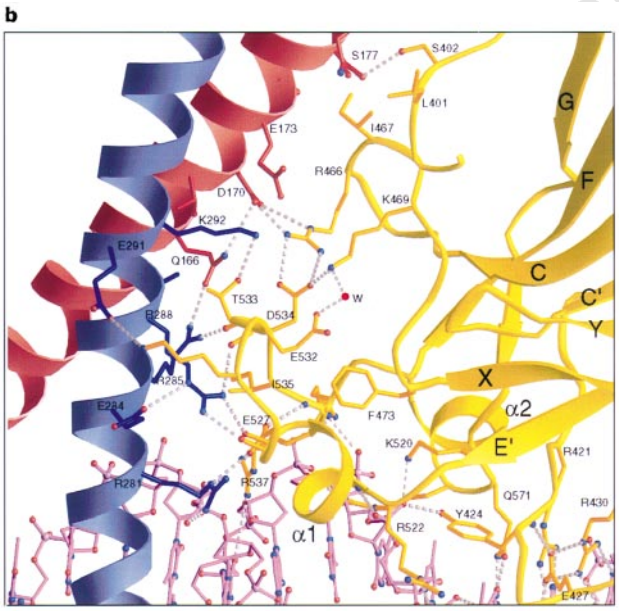


Figure 5 The interface between NFAT and AP-1. **a**, Surface representation (made with GRASP) showing the partly discontinuous character of the contacts. The box indicates the region shown in detail in **b**. The orientation is essentially the same as in Fig. 3c. **b**, Detailed view (using Ribbons[50]) of the contacts between the E′F and CX loops of NFAT (right) and Jun and Fos (left), respectively. Note the very extensive network of hydrogen bonds and salt bridges.

Table 1 Data collection, phase determination, and refinement statistics

Crystallographic data

Data set	Cell	*a* (Å)	*b* (Å)	*c* (Å)	β (°)	Resolution (Å)	Coverage (redundancies)	R_{sym}
Native I	$P2_1$	64.76	85.55	83.37	112.41	30–2.85	90.0% (1.9)	2.5%
Native II	$P2_1$	65.19	87.23	84.10	112.66	30–2.85	96.7% (2.5)	4.6%
Native III	$P2_1$	64.55	85.46	83.37	112.00	20–2.70	98.3% (3.1)	8.0%

Phase determination (Native I as native)

	Resolution (Å)	Cover (redun)	R_{sym}*	R_{deriv} (sites)†	R_{cullis} (ano)‡	Phasing power§	f.o.m. (res)‖
MIR							0.54 (3.1 Å)
Iodine 3	20–2.85	79% (2.7)	5.1%	9.2% (4)	0.79	1.27	
Iodine 5	20–2.95	89% (2.9)	5.1%	15.0% (4)	0.86	1.17	
Hg	20–3.00	92% (2.4)	8.6%	24.0% (4)	0.79	1.29	
MAD							0.32 (3.5 Å)
λ1 (0.9650 Å)	20–3.00	90% (4.5)	9.9%	(6)	(0.95)		
λ2 (0.9795 Å)	20–3.00	90% (4.5)	10.5%	(6)	0.93 (0.92)	0.66	
λ3 (0.9809 Å)	20–3.00	90% (4.5)	10.2%	(6)	0.92 (0.95)	0.60	

Refinement (Native I and II, 3,887 non-hydrogen atoms; native III, 3,975 non-hydrogen atoms)

	Resolution (Å)	Reflections (free set)	R_{cryst}/R_{free} (2σ)¶	R.m.s.d.# Bonds (Å)	R.m.s.d.# Angles (°)	Average *B*-factor (Å^2) Protein	Average *B*-factor (Å^2) DNA
Native I	10–2.85	16531 (1215)	23.6%/29.5%	0.009	1.51	57.7	46.2
Native II	10–2.85	17893 (1950)	24.0%/30.1%	0.011	1.62	60.8	48.4
Native III	10–2.7	20408 (1671)	24.6%/30.3%	0.010	1.62	54.3	39.4

Native I, II and III refer to three distinct data sets from slightly non-isomorphous crystals. Native III is the complex crystal in which NFAT is selenomethionine-substituted. Iodine 3 is a derivative in which T7, T8, T16 and T17′ have been replaced by 5-iodo-dU; in iodine 5, T (–1), T7, T16 and T17′ have been replaced by 5-iodo-dU. The refined structure contains 266 residues of NFAT, 53 of Fos, 52 of Jun, 20 DNA base pairs and 88 water molecules.

* $R_{sym} = \Sigma|I - \langle I\rangle|/\Sigma I$, where I is the observed intensity, $\langle I\rangle$ is the statistically weighted average intensity of multiple observations of symmetry-related reflections.

† $R_{deriv} = \Sigma||F_{PH}| - |F_P||/\Sigma|F_P|$, where $|F_P|$ is the protein structure factor amplitude and $|F_{PH}|$ is the heavy-atom derivative structure factor amplitude.

‡ $R_{cullis} = \Sigma||F_{PH}| - |F_P + F_H||/\Sigma|F_{PH} - F_P|$; $R_{cullis}(ano) = \Sigma|F_{PH^+} - F_{PH^-}|_{exp} - |F_{PH^+} - F_{PH^-}|_{calc}|/\Sigma|F_{PH^+} - F_{PH^-}|_{exp}$, where F_H is the heavy-atom structure factor amplitude.

§ Phasing power $= \langle|F_H|\rangle/\langle||F_{PH}| - |F_P + F_H||\rangle$, $\langle||F_{PH}| - |F_P + F_H||\rangle$ is also known as the residual lack of closure error.

‖ f.o.m. is figure of merit $= \langle\Sigma P(\alpha)e^{i\alpha}/\Sigma P(\alpha)\rangle$, where α is the phase and $P(\alpha)$ is the phase probability distribution.

¶ $R = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c are observed and calculated structure factor amplitudes, respectively. R_{free} is calculated for a randomly chosen 7.5% of reflections; R_{cryst} is calculated for the remaining 92.5% of reflections used for structure refinement.

R.m.s.d. is the root mean square deviation from ideal geometry.

cooperative interactions. The loop is supported by a small hydrophobic patch, centred on Phe 473. Several basic residues and two main-chain amides from the E′F loop contact the DNA backbone. The N terminus of α1 is capped by phosphate 4′; the C terminus is capped by Arg 288 of Jun. Thr 533 at the tip of the loop packs closely against the Jun helix, in a pocket created by the aliphatic segments of Arg 285, Arg 288 and Leu 289, and it also participates in the hydrogen-bond network. The E′F loop thus provides a continuous link at the protein interface between contacts to DNA backbone and contacts to Jun, and the presence of both DNA and Jun appear to stabilize its projecting conformation.

The CX loop, the N-terminal segment and the FG loop of NFAT contact Fos in successive 'layers' along the zipper helix. The tip of the CX loop (Arg 466 and Ile 467) interacts in the region between Asp 170 and Ser 177. The FG loop lies opposite the end of the Fos polypeptide we have used, and there is a salt link between Arg 556 and Glu 191. The cc′ loop of the NFAT C-terminal domain is adjacent to residues 164 to 171 of Fos. The actual non-covalent interactions appear to be limited, however, with only one clear hydrogen bond and a small patch of van der Waals contact.

There are four known members of the NFAT family and several members each of the Fos and Jun families. Residues involved in the NFAT–Jun contacts are invariant in both families (Fig. 1). By contrast, there is some variation of sequence on the two sides of the Fos–NFAT interface, particularly at the contact with the C-terminal domain of the NFAT RHR.

Alanine scanning mutagenesis has indicated that Arg 285 of Jun and Thr 533, Glu 527 and Ile 535 of NFAT participate in cooperative DNA binding by NFAT and AP-1 (refs 35, 36). Directed mutations in Arg 466 and Ile 467 also influence the interaction between NFAT and AP-1 (F. Macian and A. Rao, unpublished results). All these residues are in the observed interface (Figs 1 and 5b). However, contacts between NFAT and AP-1 are not strong enough to stabilize the complex when not bound to its DNA site, as shown by protein crosslinking experiments in the presence and absence of specific DNA[27].

Combinatorial regulation of transcription

Composite sites such as ARRE2 integrate signals from several pathways[6]. Weak independent binding of NFAT or AP-1 and strong cooperativity imply that both partners must be activated to generate a transcriptional response. A noteworthy characteristic of transcription-factor cooperativity on composite sites is the direct participation of the DNA-binding domains[23,37–39]. For example, mutations that affect interactions between AP-1 and the glucocorticoid receptor on the proliferin promoter map to outward-facing residues in the basic region of Fos[37]. Our structure illustrates that interactions between DNA-binding domains will constrain the design and evolution of the contacting surfaces. The interface between NFAT and AP-1 is strikingly extended, and some distortions, both of the AP-1 partner and of the DNA, appear to be required to achieve the observed match. The actual cooperativity (about a 10-fold increase in affinity of AP-1 for DNA[36]) is therefore the difference between these presumably unfavourable distortion components and favourable contributions from the various NFAT–AP-1 contacts.

Could NFAT function with partners other than AP-1? For example, could a transcription factor with a totally different sort of DNA-binding domain supply adequate cooperativity? The NFAT–AP-1 complementarity, although extensive, is partly discontinuous. There are discrete patches of interactions, and the polar side chains at the contacting surface of NFAT give it significant adaptability. It is therefore possible that a different kind of transcription factor could provide an interaction surface for a different subset of the loops on the 3′ facing side of NFAT. Contacts on the opposite, 5′-facing surface (the C′E and XY loops) may also occur in a fully occupied regulatory region such as the IL-2 promoter. There is evidence that active promoters and enhancers are essentially covered with bound factors and that their DNA may be looped into a compact, three-dimensionally ordered arrangement[39–41]. The quaternary complex we describe here is therefore likely to be a substructure in a still larger assembly, and some further components could associate with the combined surfaces of Fos, Jun and NFAT. □

Methods

Preparation of NFAT complex. Expression and purification of recombinant human NFAT1 (396–678, with an 18-residue N-terminal His-tag), Fos (c-Fos, 139–193), and Jun (c-Jun, 263–317) are as described previously[31,34]. Selenomethionine (SeMet)-substituted NFAT was prepared by growing NFAT-producing *Escherichia coli* in minimal medium in which methionine was replaced by SeMet. Oligonucleotides and their iodo derivatives were synthesized on a Milligen DNA synthesizer and purified by reverse-phase high-performance liquid chromatography. The complex was prepared by mixing 1.5:1.5:1.5:1 equivalents of NFAT, Fos, Jun and DNA at a concentration of 0.2–0.4 mM in a storage buffer containing 10 mM HEPES (pH 7.5), 1 mM DTT, 100 mM NaCl, 20% glycerol and 500 mM NH_4OAc.

Crystallization and data collection. Hanging drops of 1–2 μl NFAT complex stock solution were equilibrated with 700 μl reservoir solution that had the same chemical composition as the storage buffer, but a lower NH_4OAc concentration (300 mM). Trials with 5 NFAT RHR variants, Fos and Jun zippers of 4 different lengths, and about 60 DNA duplexes ultimately yielded crystals that diffracted to a spacing of 2.7 Å. The NFAT1, Fos and Jun sequences are derived from the human proteins (Fig. 1); the ARRE2 sequence (Fig. 2) is murine but differs at only two positions from the human ARRE2 (ref. 6). At room temperature, crystals grew to maximum size ($0.4 \times 0.2 \times 0.03\ mm^3$) as thin plates within 24 h. Crystals were stable in the harvest/cryoprotectant buffer: 10 mM HEPES (pH 7.5), 100 mM NaCl, 10% PEG3K, 15% glycerol, 15% ethyleneglycol. There is one complex in the asymmetric unit (68% solvent). Isomorphous heavy-atom derivatives were obtained with iodinated DNA (Table 1) and by soaking crystals in methyl mercury nitrate. All native and derivative crystals were flash frozen in liquid nitrogen for storage and data collection under cryogenic conditions (100 K).

Data were collected with a Mar Research image plate scanner at NSLS (beamline X-25) and at SSRL (beamline 7-1). MAD data were collected using the inverse beam mode. Data were reduced using the programs DENZO and SCALEPACK[42], then processed using the CCP4 suite[43].

Phase determination and structure refinement. We used a combination of multiwavelength anomalous dispersion (MAD)[44] from selenomethionyl residues (SeMet) incorporated into the NFAT and multiple isomorphous replacement (MIR) from iodinated DNA bases and a conventionally prepared methylmercury derivative. Native I was isomorphous to all of the derivatives, and this data set was used for initial phase determination and model building. Mercury positions (four sites) were found by inspecting the anomalous difference Patterson map. Subsequently, iodine and selenium positions were determined by the difference Fourier method. Refinement of heavy-atom parameters in MIR and MAD and phase calculation were performed using MLPHARE[43]. MAD and MIR phases were combined using SIGMAA[43]. The phase-combined map was further improved by solvent flattening using DM[43].

The overall figure of merit was 0.543 at 20-3.1 Å resolution for the initial map. About 85% of the total model was built into this map. DNA was built by fitting a standard B DNA to the density with appropriate adjustments. Fos and Jun were built by matching the previously determined model with the map as rigid bodies. Whereas Jun could be fit as one piece, the basic region and leucine zipper of Fos had to be fit as two separate pieces. NFAT was built directly in O by the baton-build utility[45]. Iterative rebuilding and refinement (positional and simulated annealing) were performed using the program O and XPLOR[46]. A partially refined model in the native I lattice was transformed to other lattices (native II and native III) by one round of rigid body refinement in XPLOR. Maps calculated in the three different lattices were averaged by the program MAVE[47], and the model was further adjusted using the triple-averaged map. Phase and refinement statistics are given in Table 1. Final models have very good geometry as examined by PROCHECK[43] (all residues have ϕ/ψ angles in the 'allowed regions' of a Ramachandran plot, with 86% of them in the 'most

favoured region'). Simulated-annealing omit maps have been computed to check several regions of the model, including the E′F loop, the CX loop and the AB loop.

Received 17 September; accepted 23 December 1997.

1. Weiss, A. & Littman, D. R. Signal transduction by lymphocyte antigen receptors. *Cell* **76,** 263–274 (1994).
2. Crabtree, G. R. & Clipstone, N. A. Signal transmission between the plasma membrane and nucleus of T lymphocytes. *Annu. Rev. Biochem.* **63,** 1045–1083 (1994).
3. Jain, J., Loh, C. & Rao, A. Transcriptional regulation of the IL-2 gene. *Curr. Opin. Immunol.* **7,** 333–342 (1995).
4. Serfling, E., Avots, A. & Neumann, M. The architecture of the interleukin-2 promoter: a reflection of T lymphocyte activation. *Biochim. Biophys. Acta* **1263,** 181–200 (1995).
5. Shaw, J. P. *et al*. Identification of a putative regulator of early T cell activation genes. *Science* **241,** 202–205 (1988).
6. Rao, A., Luo, C. & Hogan, P. G. Transcription factors of the NFAT family: regulation and function. *Annu. Rev. Immunol.* **15,** 707–747 (1997).
7. Schreiber, S. L. & Crabtree, G. R. The mechanism of action of cyclosporin A and FK506. *Immunol. Today* **13,** 136–142 (1992).
8. McCaffrey, P. G. *et al*. Isolation of the cyclosporin-sensitive T cell transcription factor NFATp. *Science* **262,** 750–754 (1993).
9. Northrop, J. P. *et al*. NF-AT components define a family of transcription factors targeted in T-cell activation. *Nature* **369,** 497–502 (1994).
10. Masuda, E. S. *et al*. NFATx, a novel member of the nuclear factor of activated T cells family that is expressed predominantly in the thymus. *Mol. Cell Biol.* **15,** 2697–2706 (1995).
11. Hoey, T., Sun, Y. L., Williamson, K. & Xu, X. Isolation of two new members of the NF-AT gene family and functional characterization of the NF-AT proteins. *Immunity* **2,** 461–472 (1995).
12. Jain, J., Burgeon, E., Badalian, T. M., Hogan, P. G. & Rao, A. A similar DNA-binding motif in NFAT family proteins and the Rel homology region. *J. Biol. Chem.* **270,** 4138–4145 (1995).
13. Chytil, M. & Verdine, G. L. The Rel family of eukaryotic transcription factors. *Curr. Opin. Struct. Biol.* **6,** 91–100 (1996).
14. Shaw, K. T. *et al*. Immunosuppressive drugs prevent a rapid dephosphorylation of transcription factor NFAT1 in stimulated immune cells. *Proc. Natl Acad. Sci. USA* **92,** 11205–11209 (1995).
15. Timmerman, L. A., Clipstone, N. A., Ho, S. N., Northrop, J. P. & Crabtree, G. R. Rapid shuttling of NF-AT in discrimination of Ca2+ signals and immunosuppression. *Nature* **383,** 837–840 (1996).
16. Luo, C. *et al*. Interaction of calcineurin with a domain of the transcription factor NFAT1 that controls nuclear import. *Proc. Natl Acad. Sci. USA* **93,** 8907–8912 (1996).
17. Shibasaki, F., Price, E. R., Milan, D. & McKeon, F. Role of kinases and the phosphatase calcineurin in the nuclear shuttling of transcription factor NF-AT4. *Nature* **382,** 370–373 (1996).
18. Beals, C. R., Clipstone, N. A., Ho, S. N. & Crabtree, G. R. Nuclear localization of NF-ATc by a calcineurin-dependent, cyclosporin-sensitive intramolecular interaction. *Genes Dev.* **11,** 824–834 (1997).
19. Liu, J. *et al*. Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell* **66,** 807–815 (1991).
20. Jain, J., McCaffrey, P. G., Valge-Archer, V. E. & Rao, A. Nuclear factor of activated T cells contains Fos and Jun. *Nature* **356,** 801–804 (1992).
21. Jain, J., Miner, Z. & Rao, A. Analysis of the preexisting and nuclear forms of nuclear factor of activated T cells. *J. Immunol.* **151,** 837–848 (1993).
22. Cockerill, P. N. *et al*. Human granulocyte-macrophage colony-stimulating factor enhancer function is associated with cooperative interactions between AP-1 and NFATp/c. *Mol. Cell. Biol.* **15,** 2071–2079 (1995).
23. Jain, J. *et al*. The T-cell transcription factor NFATp is a substrate for calcineurin and interacts with Fos and Jun. *Nature* **365,** 352–355 (1993).
24. Muller, C. W., Rey, F. A., Sodeoka, M., Verdine, G. L. & Harrison, S. C. Structure of the NF-κB p50 homodimer bound to DNA. *Nature* **373,** 311–317 (1995).
25. Ghosh, G., van Duyne, G., Ghosh, S. & Sigler, P. B. Structure of NF-κB p50 homodimer bound to a κB site. *Nature* **373,** 303–310 (1995).
26. Wolfe, S. A. *et al*. Unusual Rel-like architecture in the DNA-binding domain of the transcription factor NFATc. *Nature* **385,** 172–176 (1997).
27. Chen, L. *et al*. Only one of the two DNA-bound orientations of AP-1 found in solution cooperates with NFATp. *Curr. Biol.* **5,** 882–889 (1995).
28. Bork, P., Holm, L. & Sander, C. The immunoglobulin fold. Structural classification, sequence pattens and common core. *J. Mol. Biol.* **242,** 309–320 (1994).
29. Ellenberger, T. E., Brandl, C. J., Struhl, K. & Harrison, S. C. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex. *Cell* **71,** 1223–1237 (1992).
30. Konig, P. & Richmond, T. J. The X-ray structure of the GCN4-bZIP bound to ATF/CREB site DNA shows the complex depends on DNA flexibility. *J. Mol. Biol.* **233,** 139–154 (1993).
31. Glover, J. N. & Harrison, S. C. Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. *Nature* **373,** 257–261 (1995).
32. Seeman, N. C., Rosenberg, J. M. & Rich, A. Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl Acad. Sci. USA* **73,** 804–808 (1976).
33. Pabo, C. O. & Sauer, R. T. Transcription factors: structural families and principles of DNA recognition. *Annu. Rev. Biochem.* **61,** 1053–1095 (1992).
34. Chen, L. *Functional and Structural Studies of DNA Binding Proteins.* Thesis, Harvard Univ., (1994).
35. Sun, L. J., Peterson, B. R. & Verdine, G. L. Dual role of the nuclear factor of activated T cells insert region in DNA recognition and cooperative contacts to activator protein 1. *Proc. Natl Acad. Sci. USA* **94,** 4919–4924 (1997).
36. Peterson, B. R., Sun, L. J. & Verdine, G. L. A critical arginine residue mediates cooperativity in the contact interface between transcription factors NFAT and AP-1. *Proc. Natl Acad. Sci. USA* **93,** 13671–13676 (1996).
37. Yamamoto, K. R., Pearce, D., Thomas, J. & Miner, J. N. in *Transcriptional Regulation* (eds McKnight, S. L. & Yamamoto, K. R.) 1169–1192 (Cold Spring Harbor Laboratory Press, NY, 1992).
38. Johnson, A. in *Transcriptional Regulation* (eds McKnight, S. L. & Yamamoto, K. R.) 975–1006 (Cold Spring Harbor Laboratory Press, NY, 1992).
39. Tjian, R. & Maniatis, T. Transcriptional activation: a complex puzzle with few easy pieces. *Cell* **77,** 5–8 (1994).
40. Thanos, D. & Maniatis, T. Virus induction of human IFN beta gene expression requires the assembly of an enhanceosome. *Cell* **83,** 1091–1100 (1995).
41. Garrity, P. A., Chen, D., Rothenberg, E. V. & Wold, B. J. Interleukin-2 transcription is regulated in vivo at the level of coordinated binding of both constitutive and regulated factors. *Mol. Cell. Biol.* **14,** 2159–2169 (1994).
42. Otwinowski, Z. in *Proceedings of the CCP4 study weekend* (eds Sawyer, L., Isaacs, N. & Burley, S.) 56–62 (SERC Daresbury Laboratory, Warrington, 1993).
43. Collaborative Computational Project Number 4. The CCP suite: Programs for protein crystallography. *Acta Crystallogr. D* **50,** 760–776 (1994).
44. Hendrickson, W. A. Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science* **254,** 51–58 (1991).
45. Jones, T. A., Bergdoll, M. & Kjeldgaard, M. in *Crystallographic Computing and Modeling Methods in Molecular Design* (eds Bugg, C. & Ealick, S.) (Springer, New York, 1989).
46. Brunger, A. T. *X-PLOR Version 3.0: A System for Crystallography and NMR* (Yale Univ. Press, New Haven, CT, 1992).
47. Kleywegt, G. & Jones, T. A. in *From First Map to Final Model* (eds Bailey, S., Hubbard, R. & Walker, D.) (Daresbury Laboratory, Warrington, 1994).
48. Kraulis, P. J. MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.* **24,** 946–950 (1991).
49. Nicholls, A., Sharp, K. A. & Honig, B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins Struct. Funct. Genet.* **11,** 281–296 (1991).
50. Carson, M. Ribbons 2.0. *J. Appl. Crystallogr.* **24,** 958–961 (1991).

Acknowledgements. We thank T. Hsieh, B. Willard and N. Sinitskaya for help with protein and DNA preparation; K. Leong, T. Stehle, J. Wang, R. Chopra and M. Jacobs for help with data collection and computation; J. Jain, G. Verdine, S. Wolfe, D. Erlanson, P. Zhou, L. Sun, B. Peterson and C. Vaughan for discussions; and L. Berman, Z. Yin and M. Soltis for synchrotron technical assistance. L.C. was supported by the Cancer Research Fund of the Damon Runyon-Walter Winchell Foundation Fellowship, and by The Medical Foundation. S.C.H. is an investigator of the Howard Hughes Medical Institute.

Correspondence and requests for materials should be addressed to S.C.H. or L.C. (e-mail: schadmin@crystal.harvard.edu). Coordinates have been deposited at the Brookhaven Protein Data Bank, accession code 1a02, and are also available by e-mail from lchen@xta1200.harvard.edu.