The EMBO Journal Vol.18 No.18 pp.5028–5041, 1999

Crystal structure of an IRF–DNA complex reveals novel DNA recognition and cooperative binding to a tandem repeat of core sequences

Yoshifumi Fujii, Toshiyuki Shimizu, Masahiro Kusumoto, Yoshimasa Kyogoku[1], Tadatsugu Taniguchi[2] and Toshio Hakoshima[3]

Department of Molecular Biology, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0101, [1]Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871 and [2]Department of Immunology, Graduate School of Medicine, University of Tokyo, 7-3-1 Hongo, Bunkyou-ku, Tokyo 113-0033, Japan

[3]Corresponding author
e-mail: hakosima@bs.aist-nara.ac.jp

There has been growing interest in the role of the IRF (interferon regulatory factor) family of transcription factors in the regulation of immune responses, cytokine signaling, and oncogenesis. These members are characterized by their well-conserved DNA binding domains at the N-terminal regions. Here we report the 2.2 Å resolution crystal structure of the DNA binding domain of one such family member, IRF-2, bound to DNA. The structure reveals its recognition sequence, AANNGAAA (here, recognized bases are underlined and in bold, and N indicates any base), and its cooperative binding to a tandem repeat of the GAAA core sequence induced by DNA structure distortions. These facts explain well the diverse binding properties of the IRF family members, which bind to both single and tandemly repeated sequences. Furthermore, we also identified the 'helix–hairpin–strand motif' at the C terminus of the recognition helix as a metal binding site that is commonly found in certain classes of DNA-interactive proteins. Our results provide new insights into the structure and function of this family of transcription factors.

Keywords: base recognition/cooperativity/IRF/transcription factor/X-ray

Introduction

The IRF (interferon regulatory factor) family of transcription factors has been extensively studied in the context of host defense and oncogenesis (Taniguchi *et al*., 1997). In fact, IRF-1 and IRF-2, the first members of this family to be identified, were originally discovered in their role as regulators of the interferon-α/β (IFN-α/β) genes (Miyamoto *et al*., 1988; Harada *et al.*, 1989). This discovery preceded the recent expansion of this family to include several other members: IRF-3, p48 (ISGF3-γ), ICSBP, IRF-4, IRF-5, IRF-6 and IRF-7 (Nguyen *et al.*, 1997). In addition, virally encoded IRF members (v-IRFs) have been reported in the human herpes virus 8 (Zimring *et al.*, 1998), which may regulate the IFN response and/or oncogenesis (Moore *et al.*, 1996; Gao *et al*., 1997). One typical feature of this family is that they all show extensive homology among their DNA binding domains in the N-terminal regions. Other interesting features include their DNA binding properties to both single and tandemly repeated sequences (Miyamoto *et al.*, 1988; Harada *et al.*, 1989; Uegaki *et al.*, 1993), and their interactions with other transcription factors such as TFIIB (Wang *et al.*, 1996).

IRF-2 was originally identified as a competitor of IRF-1, a transcriptional activator of the IFN system (Miyamoto *et al.*, 1988). In fact, these two closely related factors bind to the same DNA sequence at enhancer regions having similar affinities and regulate several IFN-inducible genes, as well as the IFN-α/β genes themselves (Maniatis *et al.*, 1992; Taniguchi *et al.*, 1995). Recent *in vitro* biochemical studies have shown that induction of human IFN-β gene expression requires the assembly of an enhanceosome bound to the enhancer DNA (Falvo *et al.*, 1995; Thanos and Maniatis, 1995). This macromolecular particle contains at least three transcription factors, IRF, NF-κB(p50/p65) and ATF-2–c-Jun, as well as the high-mobility-group protein HMG I(Y). Moreover, the formation of the enhanceosome–pre-initiation complex requires cooperative interactions between both the enhanceosome and the general transcription factors TFIID, TFIIA and TFIIB, and the cofactor USA (Kim and Maniatis, 1997), all of which are located around the TATA box at –29, ~25 bp away from the closest NF-κB site (Fujita *et al.*, 1985). In addition to playing essential roles in the IFN system, IRF-2 and IRF-1 are also involved in several regulatory processes, i.e. cell cycle regulation (Tanaka *et al.*, 1996), tumor suppression and oncogenic activities (Harada *et al.*, 1993; Tanaka *et al.*, 1994), and in several critical processes in the immune system (Taki *et al.*, 1997; Ogasawara *et al.*, 1998). Remarkably, IRF-2 also functions as a transcriptional activator when bound to the enhancer of the vascular cell adhesion molecule-1 (VCAM-1) gene (Jesse *et al.*, 1998) or to the enhancer of the cell-cycle-regulated histone H4 gene (Vaughan *et al.*, 1995). More recently, functional studies for other members have revealed that each factor carries functions distinct from the others (reviewed in Nguyen *et al.*, 1997). Notably, it has been demonstrated that direct triggering of the IFN-α and IFN-β systems by viral infection is mediated by a transcription factor complex containing IRF-3 and IRF-7 (Marie *et al.*, 1998; Sato *et al.*, 1998a,b; Wathelet *et al.*, 1998; Yoneyama *et al.*, 1998).

All the IRF members are believed to bind to similar DNA sequences. Using a polymerase chain reaction-assisted DNA binding site selection method, Tanaka *et al.* (1993) have determined that the interferon regulatory factor binding element (IRF-E), G(A)AAA$^{G}/_{C}{}^{T}/_{C}$GAA-

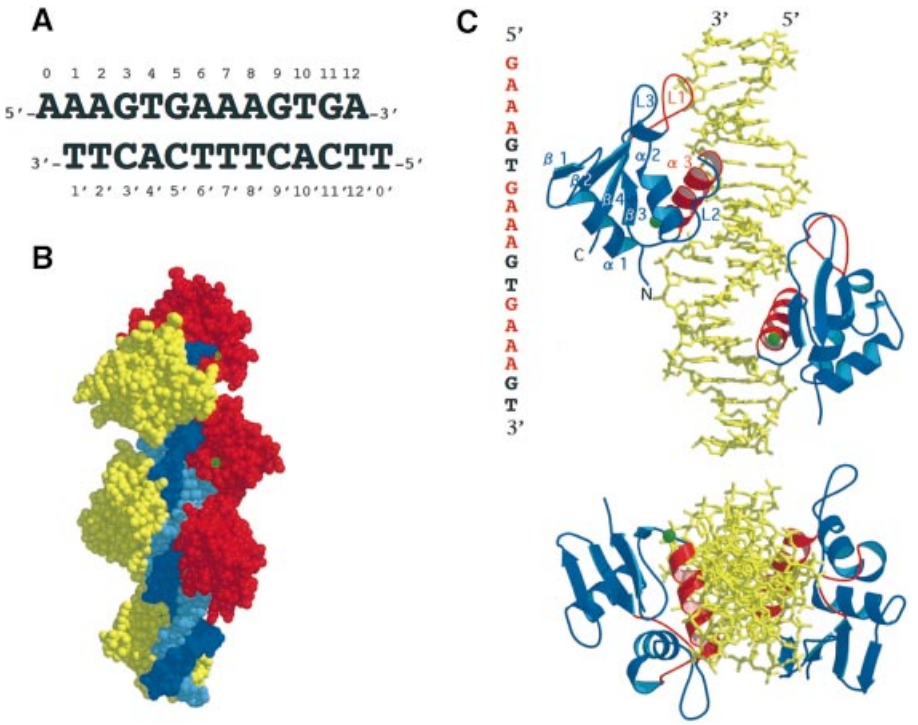

Fig. 1. Overall structure of the IRF-2–DNA complex. (**A**) The sequence of the DNA oligomer used in this study is shown with a numbering scheme. (**B**) Space-filled representation of the overall IRF-2–DNA complex structure containing six IRF-2 DNA binding domains and three DNA duplexes stacked atop each other to form a continuous double helix. The strands of DNA are colored in dark and light blue. The DNA binding domains are colored in red and light green. Potassium ions bound to the DNA binding domains are in green. (**C**) Two DNA binding domains tandemly bound to DNA are highlighted to show the DNA binding mode with both a side view (top) and an end view (bottom). The recognition helix α3 and loop L1 are colored in red, and the other parts are in blue. DNA and potassium ions are colored in light green and green, respectively. The secondary structure elements and the N and C termini are labeled. The side view is accompanied by the DNA sequence with the GAAA core sequences in red.

$A^G/_C{}^T/_C$, is the consensus sequence for IRF-1 and IRF-2, and have found it in the enhancers of several IFN-inducible genes in addition to the IFN-β enhancer. This sequence overlaps with the interferon-sensitive response element (ISRE), $^A/_G$NGAAANNGAAACT, which was deduced from IFN-α-stimulated genes (Darnell *et al.*, 1994). To date, many IRF binding sites have been identified and shown to have these consensus sequences, which are characterized by tandem repeats of GAAA with mostly two inserted bases. Interestingly, some of the binding sites have one or three insertion bases, whereas others have only one GAAA stretch. Moreover, another class of IRF binding sites repeats the GAAA sequences three or four times. Nevertheless, no rational interpretation has been provided for the binding properties of IRFs to these variable GAAA repeats.

Recently, the crystal structure of a DNA binding domain of IRF-1 bound to a PRD I oligomer was determined, and this, in turn, allowed the elucidation of the monomeric binding of IRF-1 to a single GAAA core sequence (Escalante *et al.*, 1998). However, the structural basis for the diverse binding properties of the IRF family members described above is largely uncharacterized. In addition, the analysis has been limited to a marginal resolution (3 Å). In this paper, we describe the highly resolved (2.2 Å) crystal structure of a DNA binding domain of IRF-2 complexed with DNA whose sequence is designed to form a tandem repeat of the core sequence (Figure 1A). The structure establishes the detailed framework of DNA recognition, especially those involving water-mediated interactions. We found that **AA**NN**GAAA** (here, recognized bases are underlined and in bold, and N indicates any base) is the sequence physically recognized by IRF-2, which explains the variety of the IRF binding sites. Our analysis also reveals DNA structural deformations, indicating a cooperative binding of IRFs to tandem repeats of the core sequence. The structure also facilitates detailed comparison of IRF-2 with other related transcription factors and serves as a high-resolution model that should be broadly applicable to other IRF proteins. Finally, this study provides new insights into the interactions of the IRF proteins with the other transcription factors.

Results

Overall structure

Two crystal forms, the orthorhombic and hexagonal forms, of a DNA binding domain of IRF-2 bound to DNA were determined by the multiple isomorphous replacement (MIR) and molecular replacement methods, respectively. The structure of the hexagonal form was refined to an *R* factor of 20.2% (R_{free} = 24.3%) for intensity data extending to 2.2 Å resolution, which enabled us to investigate the details of the molecular recognition with several water molecules and ions bound to protein or protein–DNA interfaces. In the asymmetric unit, the crystal structure contains six IRF-2 DNA binding domains and three DNA duplexes stacked together to form a continuous double helical DNA (Figure 1B). Moreover, the DNA duplexes are stabilized by stacking interactions with those of the symmetry-related complexes to form a finite continuous DNA helix in the crystal. The structures of these three

DNA duplexes are essentially the same and are superimposed with root mean square (r.m.s.) deviations ranging from 0.26 to 0.37 Å for all atoms. In addition to the GAAA core sequence within each oligomer, this arrangement of DNA duplexes produces an additional GAAA core sequence between two DNA duplexes to yield six repeats of a GTGAAA sequence, each of which is bound to the monomeric DNA binding domain. The crystal structure accordingly contains six crystallographically independent protein subunits bound to tandem repeats of this site which are formed only in the crystal by continuous base stacking and pushing out the unpaired bases of the 13mer used for crystallization (see Materials and methods). The fact indicates that the protein selects the correct binding site and by doing so induces this type of arrangement. The neighboring DNA binding domains bind to DNA on the opposite surfaces of the DNA helix with no direct intermolecular contact: the closest distance between the C_α-carbon atoms of the two domains is 14 Å. Similarly, there is no direct contact between any of the pairs of the DNA binding domains on the same DNA surface. The overall structure of the orthorhombic form, which was determined at 2.8 Å resolution, is basically the same as that of the hexagonal form.

The IRF-2 DNA binding domain consists of four-stranded antiparallel β sheets (β1–β4) with three α helixes (α1–α3) and three long loops (L1–L3). This architecture resembles that of the winged helix–turn–helix (HTH) motif (Figure 1C). The structures of the six crystallographically independent domains are essentially the same, with r.m.s. deviations ranging from 0.2 to 0.4 Å for the C_α-carbon atoms, excluding residues of the mobile loop L3. The overall structures of the IRF-2 and IRF-1 DNA binding domains seem to share the same IRF fold (Figure 2). Superposition of these DNA binding domains using the secondary structure elements (Figure 2B), but not the corresponding residues in their sequences, gives r.m.s. deviations ranging from 1.04 to 1.14 Å for the C_α-carbon atoms, and ranging from 1.13 to 1.16 Å for the main chain atoms. Since the DNA binding domains of IRF-2 were highly homologous (76% sequence identity) with that of IRF-1, the entire sequences could be properly aligned. Nevertheless, superposition of the DNA binding domains using the corresponding residues in their sequences resulted in r.m.s. deviations of 4.06 Å for all C_α-carbon atoms and 3.71 Å for the C_α-carbon atoms of the secondary structural elements. These unexpectedly large deviations have been found to be due to the differences in assignments of the secondary structural elements. Compared with those of IRF-1, the residues forming helix α1 of IRF-2 are shifted by four residues toward the N terminus. In addition, the residues of strand β1 are shifted toward the N terminus by one residue and the residues of strand β4 are shifted toward the C terminus by two residues. These shifts cause differences in the residues of the hydrogen-bonded pairs of the β sheet, in addition to differences in the loop structures between the secondary structure elements. It is unlikely that these structural differences between the IRF-1 and IRF-2 DNA binding domains are induced by the crystal packing. In the present crystal, all crystallographic independent molecules, which have different crystal contacts, have essentially the same structure. In addition, the structures in the orthorhombic form are the same (r.m.s. deviations of ~0.50 Å) as those in the hexagonal form, which also have different crystal contacts. Moreover, recent NMR studies of a free form of the IRF-2 DNA binding domain in solution have verified our assignments of the secondary structure elements and the pattern of interstranded hydrogen bonds in the β sheet (Furui *et al.*, 1998) (Figure 2B) with reasonably small r.m.s. deviations of ~1.6 Å for C_α-carbon atoms of the secondary structural elements. This fact indicates that no significant structural changes involving the secondary structures occur on the DNA binding, although loops L1–L3 exhibit large conformational changes. These loops of the free form are highly disordered in solution. The differences in the DNA binding sites, as described below, are unlikely to induce these structural deviations between the IRF-2 and IRF-1 DNA binding domains because the regions displaying the large deviations are located at the opposite molecular surface of the DNA binding surface.

To clarify whether these unexpected differences between the IRF-2 and IRF-1 DNA binding domains reflect divergence among members of the IRF family or not, we need to determine the structure of another IRF member, in addition to further analysis of the IRF-1–DNA complex at higher resolution.

DNA sequence recognition at the major and minor grooves

All six IRF-2 domains interact with DNA in a similar manner. One of the most striking features of the DNA binding is the recognition of a 5′-flanking AA sequence which is located 2 bp upstream from the GAAA core sequence that is recognized by the recognition helix α3 (Figure 3A, left, and 3B). This AA sequence is part of another upstream core sequence which is also recognized by another DNA binding domain at the major groove. The recognition is due to the His-40 residue located at loop L1, which reaches into the minor groove of the upstream GAAA sequence. His-40 forms a hydrogen bond with a bridging water molecule (W1) between the AA base pair steps by hydrogen bonding to both the O2 atom of the paired thymine and the N3 atom of the adenine of the next step. These contacts have not been observed in the crystal structure of the IRF-1–DNA complex, although His-40 of IRF-1 is located at a position where it may be able to form water-mediated hydrogen bonds like that in the current structure. Since His-40 is completely conserved in the IRF family, we believe that a similar recognition would occur in all members of this family and propose **AA**NN**GAAA** as the consensus IRF recognition sequence (IRS) that is physically recognized by the IRF DNA binding domains.

Contacts with the major groove of the GAAA sequence are localized at the C-terminal half of the recognition helix and are mediated by four residues (Asn-80, Arg-82, Cys-83 and Asn-86). In the IRF-1–DNA complex (Figure 3A, right), Ser-87, which is variant in the IRF family, was used to recognize the GAAA core sequence, but was missing in the current complex. Several significant differences are found in the frameworks of hydrogen bonds and van der Waals contacts between these residues and the core sequence (Figure 3A). These differences are due primarily to an extensive network of water molecules located within the interface between the recognition helix

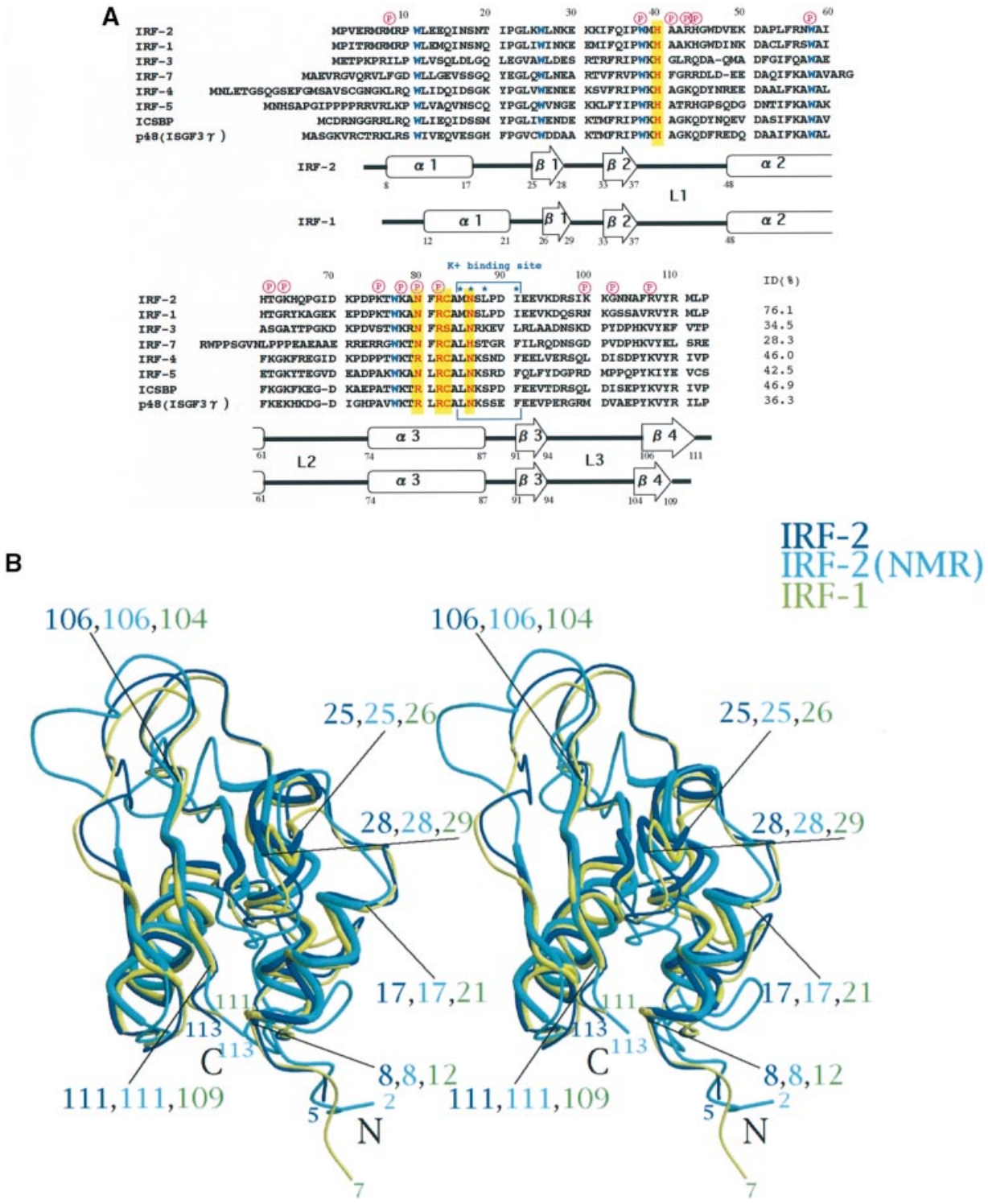

Fig. 2. Secondary structure elements and sequence alignment. (**A**) Sequence alignment of the DNA binding domains of IRF-2 and the related members of the IRF family from the mouse is shown with the secondary structure elements of IRF-2 and IRF-1. The sequential numbering of IRF-2 is shown at the top. Residues interacting with phosphate groups of the DNA backbones are marked by pink circles containing the letter P. The K^+ ion binding site forming the helix–hairpin–strand motif is marked by blue brackets with asterisks for the coordinated residues. The key residues for DNA sequence recognition are colored in red with highlights in yellow and the conserved tryptophans characteristic for the IRF family members are in blue. Pairwise identities with IRF-2 are shown at each end of the sequences. (**B**) A stereo view of superposition of the IRF-2 DNA binding domain bound to DNA in the crystal (blue), the free form of the IRF-2 DNA binding domain in solution determined by NMR (light blue) and the IRF-1 DNA binding domain bound to DNA in the crystal (light green). The residue numbers are indicated for the starting and end residues of the secondary structural elements, α1, β1 and β4, which exhibit significant deviations between IRF-1 and IRF-2. The N- and C-terminal residues are also indicated.

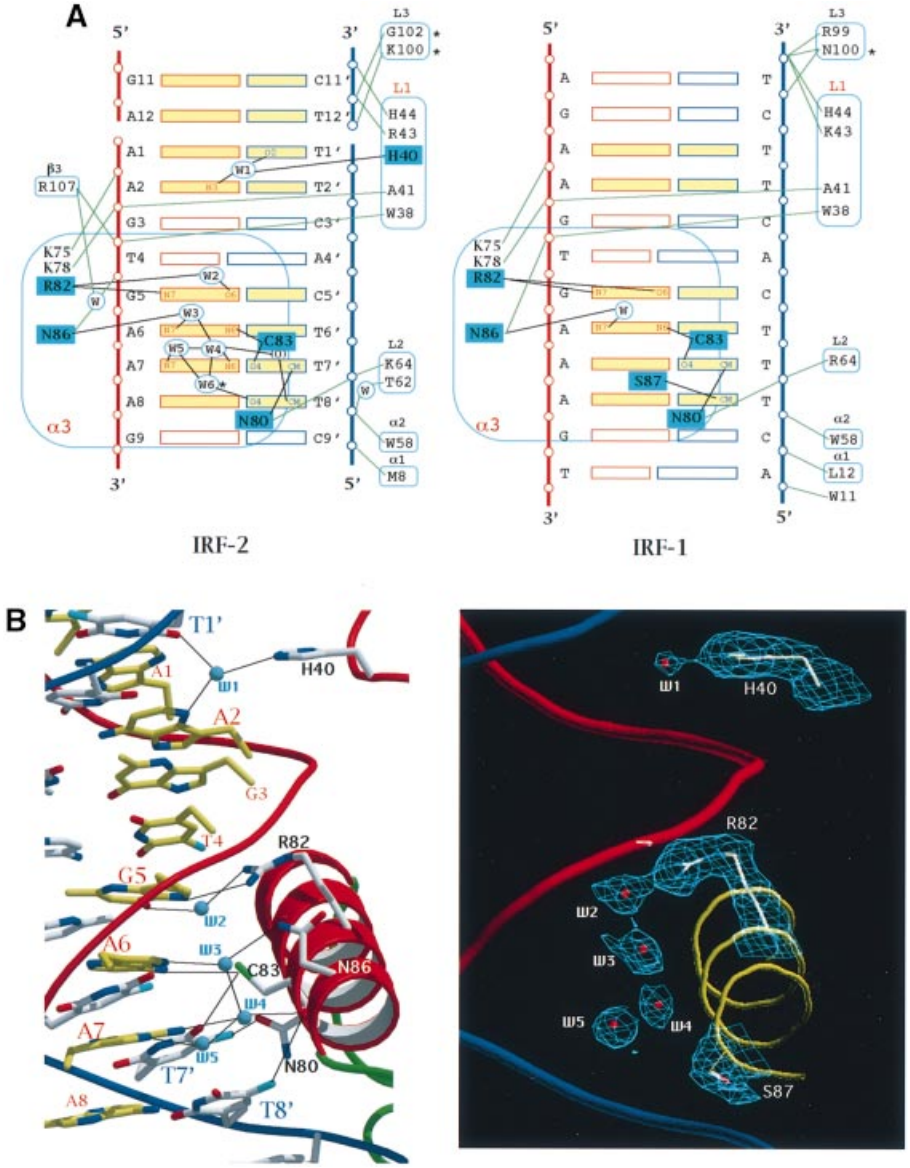

Fig. 3. Base recognition with water-mediated hydrogen bonds. (**A**) Schematic diagram of the protein–DNA contacts in the IRF-2–DNA (left) and IRF-1–DNA (right) complexes. Base pairs of the GAAA core sequence are colored in yellow. The labels of the amino acid residues contacting with bases are also highlighted with light blue. Hydrogen bonding and van der Waals contacts participating in the base recognition are represented by thin black lines. Hydrogen bonds and/or ion pairs to phosphate groups of the DNA backbones are represented by light green lines. The residues marked with asterisks do not participate in the DNA interactions in some of six molecules in the crystal. Similarly, W6 is missing at some of the complex interfaces. (**B**) Recognition of the 5′-flanking AA and the GAAA core sequences by loop L1 and recognition helix α3, respectively. A close-up view (left) of the structure at the protein–DNA interface. Hydrogen bonding and van der Waals contacts participating in the base recognition are represented by thin black lines. Water molecules are represented as light blue balls. For clarity, the water molecule W6 has been omitted. Methyl groups of thymines are colored in light blue with van der Waals contacts represented by thin black lines. Bases in light green are of the DNA strand (colored in red) running from top to bottom in a 5′ to 3′ direction. An omit map (right) contoured at the 1σ level around the side chains of His-40, Arg-82 and Ser-87, and the water molecules (W1–W5) at the protein–DNA interface.

and DNA. No significant changes in the position and orientation of the recognition helices (the r.m.s. deviation of 0.29 Å for C_α-carbon atoms) were observed. However, a few side chains in the recognition helix have different conformations from those of IRF-1, resulting in the differences in DNA recognition. The side chain of the completely conserved Arg-82 sticks into the major groove to form two hydrogen bonds with the first guanine of the core sequence (**G**AAA) in a manner similar to that with IRF-1, but with one of the hydrogen bonds mediated by a water molecule (W2). The recognition of the second adenine (G**A**AA) is mediated by a direct hydrogen bond with the side chain of conserved Cys-83 and water (W3)-mediated hydrogen bonds involving the side chain of conserved Asn-86. Specificity for the third adenine (GA**A**A) is conferred by a direct hydrogen bond between Cys-83 and the paired thymine, together with a van der Waals contact between the methyl group of thymine and the side chain of Asn-80. In addition, the adenine base is hydrated with two water molecules (W4 and W5); this hydration is part of the hydrogen bonding network linked to Asn-86. Among the water molecules, W3 is observed in the IRF-1–DNA complex, but the others are missing. The fourth adenine (GAA**A**) is recognized through a van

der Waals contact with the paired thymine as in the IRF-1–DNA complex, although the methyl group contacts with the main chain carbonyl of Cys-83, instead of the side chain of Ser-87 of IRF-1. The positions of these side chains (His-40, Arg-82, Ser-87) and the water molecules (W1–W5) were verified with an omit map (Figure 3B, right). It is noteworthy that the recognition of the fourth AT base pair is poor in comparison with the recognition of the other base pairs of the core sequence. We note that direct hydrogen bonds involving the side chains of Arg-82 and Cys-83 are common in the IRF-2–DNA and IRF-1–DNA complexes.

DNA structure deformations by IRF-2 binding

Several amino acid residues located at all three α helices and three loops plus strand β4 contact with the negatively charged DNA backbone via hydrogen bonds and ion pairs (Figure 3A). The contacts cover the DNA backbones of the 12 bp stretch and involve the main chain amide groups and the side chain indole rings of tryptophans, as well as the positively charged lysines and arginines. Interestingly, the conserved Trp-11 located at helix α1 is buried into the hydrophobic core, whereas in the IRF-1–DNA complex this residue interacted with the phosphate. Arg-82, which is a key residue for the base recognition as described above, forms a hydrogen bond with the phosphate group. Another key residue for the base recognition, Asn-86, also interacts with the 5′-phosphate group of the guanine nucleotide of the core sequence to form a water-mediated hydrogen bond together with Arg-107 of strand β4. The protein–phosphate interactions are localized in two sugar–phosphate backbones forming the major groove of the core sequence. These localized interactions result in a narrowing of the major groove with a local 20° bending of the DNA helix toward the protein, as observed in the IRF-1–DNA complex (Figure 4A). Superposition of the DNA oligomers in the two complexes gives a relatively small averaged r.m.s. deviation (0.94 Å) for the corresponding backbone atoms. It is of particular interest that all six DNA binding domains induce similar DNA distortions at every GAAA core sequence. Consequently, the helical axis of the continuously stacked DNA duplexes writhes, due to the tandem binding with each of the 6 bp, wherein every two neighboring DNA binding domains induce local bendings toward the nearly opposite sides. The bending toward the protein enables loop L1 to approach the minor groove of the 5′-flanking AA sequence.

To clarify the quantitative DNA distortions, the helical parameters of DNA were analyzed. Curiously, the mean axial rise per turn (3.32 Å) and the average helical twist (34.7°) resemble those of B-DNA. However, the structure is highly irregular, accompanied by both narrowing and widening of the grooves (Figure 4B). The helical twist per base pair varies from 23.3° to 43.5°. Superposition of the current DNA on B-DNA gives a large averaged r.m.s. deviation (2.18 Å) for the corresponding backbone atoms. The narrowing of the major groove of the TGAA sequence, which contains the 5′ part of the GAAA core sequence, results in a local widening of the minor groove at the GC base pair, which has a positive value of the base roll. In sharp contrast, the AAG sequence that contains the 3′ part of the core sequence has a narrow minor groove with negative base rolls (ranging from –3° to –5°) and large negative propeller twists (from –15° to 20°). We found that variations of several helical parameters are basically the same in both IRF-2–DNA and IRF-1–DNA complexes, as well as the variations of the groove widths (Figure 4B, left rows). These characteristics are reminiscent of the A-track sequences as a bending sequence (Wu and Crothers, 1984; Nelson *et al*., 1987).

To compare the DNA structures bound to IRFs with those of unbound DNA structures, the DNA structures containing GAAA sequence were searched in the Nucleic Acid Database (Berman *et al.*, 1998). We found one crystal structure (ID:BDJ081), that was determined at 1.85 Å resolution, which is sufficient for the structure comparison including hydration structures (Han *et al*., 1997). This DNA has a sequence, CAAAGAAAAG, that contains an A-tract repeat. The helical parameters are compared with those of the DNA bound to the IRF-2 DNA binding domain (Figure 4B, right row). The helical parameters of both DNAs show overall similarity as well as strong similarity of the groove widths. Like the IRF-2-bound DNA, the unbound DNA has large negative propeller twists, small base rolls and a narrow minor groove within the A-tract region. Excluding the terminal nucleotides, superposition of the IRF-2-bound DNA on the unbound DNA oligomer gives an r.m.s. deviation (1.82 Å) for the corresponding backbone atoms, which is smaller than that on B-DNA. Similar characteristics were also detected in the structure of another DNA oligomer, CGCG-AAAAAACG, which has been determined at 2.3 Å resolution (DiGabrieli and Steitz, 1993). In sharp contrast to the similarity, the IRF-2-bound DNA has exceptionally larger base rolls at the GA sequence, resulting in the widening of the minor groove at GC base pair, compared with the unbound DNA oligomer. From these facts, it is supposed that the DNA structure containing IRF recognition sequence in the unbound form has an A-tract-like conformation, while the protein induces the widening of the minor groove at every GA step with large positive base rolls by binding to DNA, resulting in a local 20° bending. Thus, the bend is largely the result of this variation in base roll.

Unfortunately, the water molecules (W1–W5) observed in the IRF-2–DNA complex cannot be assigned in these crystal structures of the unbound DNA oligomers having A-tracts, although the water molecule W1, which mediates the recognition of the 5′-flanking AA sequence, seems to correspond to a typical bridging water molecule of the primary spine of hydration at the minor groove of the A-tract (Drew and Dickerson, 1981). It is necessary to determine the unbound DNA structure containing the recognition sequence to clarify whether these water molecules exist in the unbound form.

Metal binding site

A metal ion binding site was found at the C terminus of the recognition helix α3 that is connected to a β hairpin structure followed by strand β3 (Figure 5A, left). We assumed that the ion was a potassium ion, since potassium is the only metal cation in the crystallization solution. The averaged B-factor of the potassium ions was 28.2 $Å^2$, which is comparable with the mean B-factor of this crystal. The coordination shell is formed by four main chain carbonyl groups of two residues (Met-85, Asn-86) of

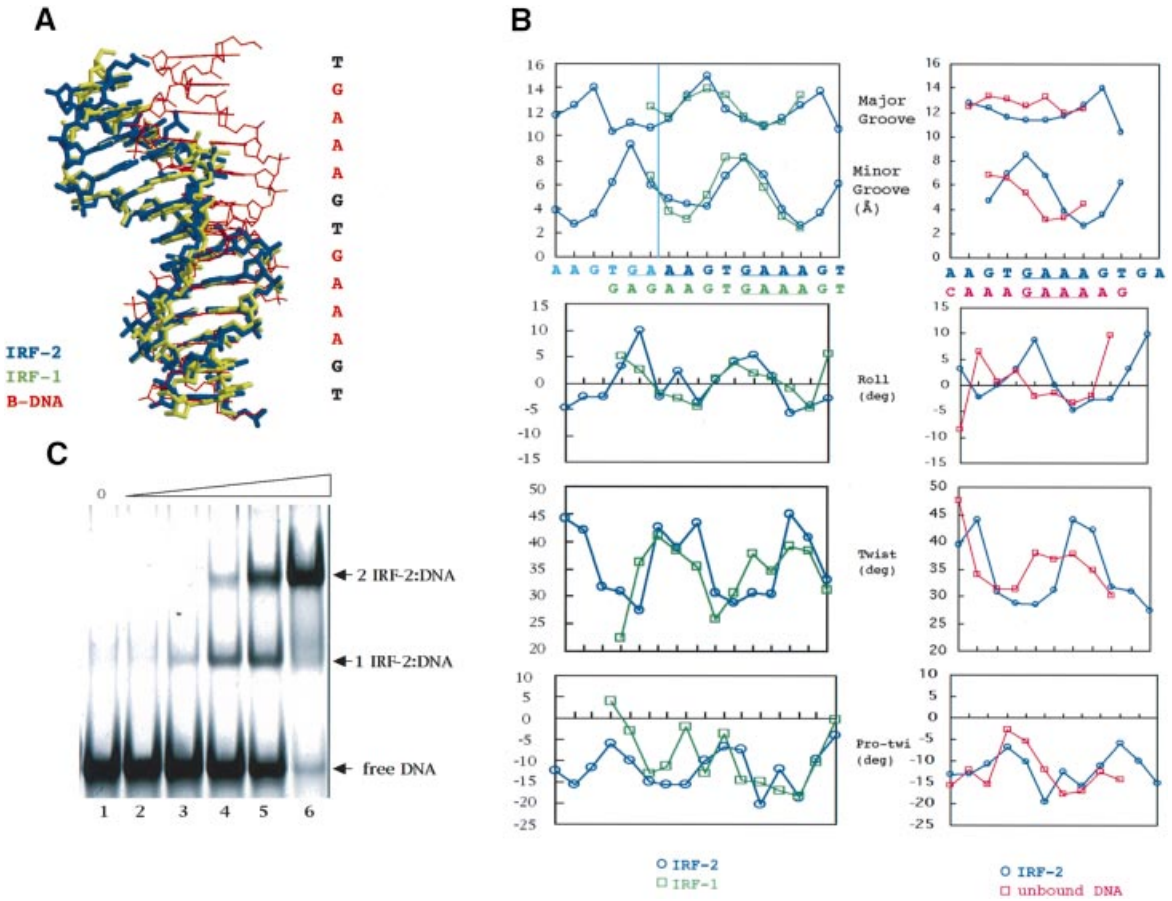

Fig. 4. DNA structural deformation. (**A**) The superposition of DNA oligomers in IRF-2–DNA (blue) and IRF-1–DNA (light green) complexes is compared with the standard B-DNA (red). The view is accompanied by the DNA sequence with the GAAA core sequence in red. (**B**) Plots of the variations in widths of the major and minor grooves, and helical parameters (roll, twist, propeller twist) of DNA in the IRF-2–DNA (blue circles) and IRF-1–DNA (light green rectangulars) are shown on the left, and those of IRF-2–DNA (blue circles) and unbound DNA (BDJ081, pink squares) are on the right. DNA sequences are also shown in the row. (**C**) Electrophoretic mobility shift assay of the IRF-2 DNA binding domain to the IRF recognition sequence (IRS). The DNA sequence 5′-AATGAC**AA**GT**GAAA**GT**GAAA**GTGTGCC-3′ contains two copies of IRS (**AA**NN**GAAA**). In this experiment, the IRF-2 DNA binding domain was prepared in a 2-fold serial dilution starting with a concentration of 3 μM (lanes 6–2). The DNA concentration is constant (1.2 μM).

the recognition helix and two from the β hairpin (Leu-88, Ile-91). On the current electron density map, two of the octahedral coordinations are invisible (Figure 5B). Surprisingly, a similar metal ion binding site was found by Clark *et al.* (1993) in the DNA binding domain of HNF-3γ, which possesses one of the closely related winged HTH domains with IRF. In this case, the metal ion was assigned as a magnesium ion. These ions may play the role of a C-terminal cap that neutralizes the helix dipole by the positive charge, while also contributing to electrostatic interaction with the phosphate groups of DNA. This 'helix–hairpin–strand (HhS) motif' for metal ion binding seemed to be related to the metal ion binding motif, i.e. the 'helix–hairpin–helix (HhH) motif', found in the DNA binding sites of endonuclease III and the related DNA repair enzymes (Thayer *et al.*, 1995) as well as human DNA polymerase β (Pelletier and Sawaya, 1996) (Figure 5A, right). The motif of DNA polymerase β has an affinity for biologically prevalent metal ions in the order $K^+ > Na^+ > Ca^{2+} > Mg^{2+}$, with the K^+ ion displaying the strongest binding. A K^+ ion bound to the site of DNA polymerase β has been shown to interact with a phosphate group of DNA (Pelletier *et al.*, 1996). The amino acid sequences of the HhH motifs are well conserved, but are somewhat different from those of the HhS motif (Figure 5C). Key residues in the formation of a hydrophobic core by the HhS motif of IRF-2 are Met, Leu and Ile. In HNF-3γ, this core is formed by Leu and a larger side chain of Phe. Remarkably, these residues are fairly conserved in the HhH motif, indicating that these two motifs are derivatives of a common 'helix–hairpin motif'.

Discussion

Cooperative binding

The present structure provides the first view of a tandem binding of IRF to a consensus repeated sequence. Several biochemical studies have suggested that IRF-1 and IRF-2 cooperatively bind to tandemly repeated sequences. A decade ago, the cooperative binding was analyzed using synthetic DNA oligomers containing AAGTGA repeats. By this method, a dimeric repeat was shown to provide relatively weak protection of footprinting, but multiply tandem repeats were found to provide complete protection (Fujita *et al.*, 1988; Miyamoto *et al.*, 1988; Harada *et al.*, 1989), since the dimeric repeat yields only a single IRS sequence. Similar results were also obtained for the DNA binding domain of IRF-2 by means of electrophoretic

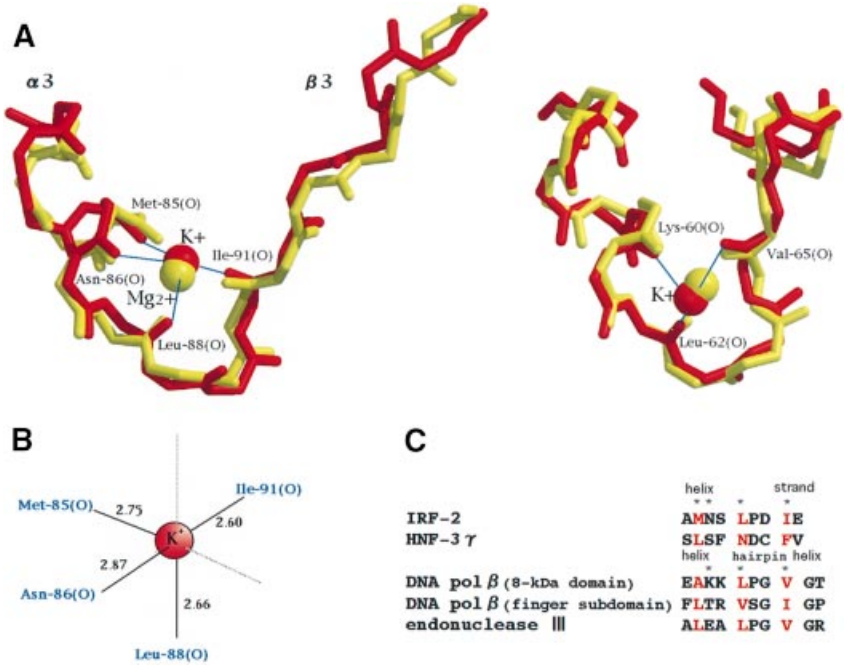

Fig. 5. Metal binding site of the recognition helix of IRF-2. (**A**) The HhS motif (left) of IRF-2 (red) for the binding of potassium ion is superimposed with that of HNF-3γ (light green), and the HhH motif (right) of DNA polymerase β (red) is superimposed with that of endonuclease III (light green). No side chains are shown for the purpose of clarity. The coordinations of the main chain carbonyl groups to the potassium ion of IRF-2 are indicated with blue lines with labels for the residues. (**B**) The coordination sphere of the potassium ion bound to the IRF-2 DNA binding domain with the coordination distances in angstroms. (**C**) Sequence alignment of the HhS and HhH motifs. Coordinated residues are marked with asterisks and conserved hydrophobic residues are colored in red.

mobility shift assay (EMSA) and footprinting (Uegaki *et al.*, 1993), implying that at least part of the cooperativity involves the DNA binding domains and/or their interactions with DNA. Our quantitative analyses of EMSA data also showed the cooperativity causing enhancement of the second binding: for the 27mer DNA shown in Figure 4C, the second binding (the obtained K_d value of 0.17 μM) was 6.1-fold stronger than the first binding (1.04 μM). Strikingly, the structure indicates that the cooperativity is not induced by contact between two adjacent DNA binding domains. Alternatively, the induction of the cooperativity by the DNA structural distortions is suggested by the fact that the DNA distortions at every binding site in the current complex are essentially the same as those in the IRF-1–DNA complex, where the DNA binding domain of IRF-1 bound monomerically to a single IRF binding site. The binding of a single DNA binding domain can introduce these DNA distortions, thus preparing a template for the cooperative interaction with the second DNA binding domain. This may be the first structure showing a cooperative DNA binding by DNA distortions without protein–protein contact. In the context of DNA looping on the transcription activation (Ptashne, 1988; Ptashne and Gann, 1997), the concept of cooperative binding mediated by DNA structural deformation has been introduced on finding a sharp DNA kink in the crystal structure of *Escherichia coli* CAP bound to DNA, where a DNA looping induced by CAP may enhance recruitment of RNA polymerase (Schultz *et al.*, 1991). Similar mechanisms have been proposed from the structures of TBP (Kim *et al*., 1993), LEF-1 (Love *et al*., 1995) and IHF (Rice *et al*., 1996) bound to DNA. Moreover, DNA bending-mediated gene repression has also been proposed (Schumacher *et al*., 1994; Lewis *et al*., 1996). The cooperativity or interference of these examples, however, are induced through long-range coupling such as DNA looping and are not induced by a direct effect of the DNA deformation. In fact, disruptions of the base pair stacking have no effect beyond about half a turn of the DNA double helix (Kim *et al*., 1993). Recently, cooperativity at shorter distances has been discussed in relation to a generic cooperativity resulting from structural distortions induced by the binding of a protein to DNA (Rudnick and Bruinsma, 1999). No structure, however, is reported to show these cooperative bindings mediated by DNA distortions. The cooperativity of the IRF-2 binding mediated by the DNA distortions seems to be consistent with the fact that the spacer between two repeated core sequences varies from 1 to 3 bp (see below), implying a lack of any specific protein–protein contact that may restrict the positions and orientations of two adjacent DNA binding domains. Model building studies indicate that spacers of either 1 or 3 bp produce no direct contact between two adjacent DNA binding domains. IRF achieves cooperativity in DNA binding, thereby enabling the recognition of naturally occurring IRF binding sites that are long enough to ensure binding strength and specificity.

IRF recognition sequence

The proposed IRS, **AA**NN**GAAA**, provides a rational interpretation of various sequences for the IRF family members, including IRF-E and ISRE consensus sequences. Some of the binding sequences that contain PRD I of the IFN-β gene have a single GAAA core sequence, but all these sequences are completely endowed with the 5′-flanking AA sequence that produces one complete IRS (Table I). Many IRF binding sites contain dimeric repeats of the core sequences with a spacer of two bases. Some

Table I. IRF responsive elements of several genes

Gene	Sequence[a]	Position	Reference
IRS consensus	**AA** NN **GAAA**		this work
Single site			
IFN-β(PRD-I)	AG**AA** GT **GAAA** GT	(–78/–66)	Neish *et al*. (1995)
HLA-B7	AT**AA** GT **GAAA** CT	(–164/–175)	Neish *et al*. (1995)
ODC	GG**AA** CT **GAAA** CT	(2711/2722)	Manzella *et al*. (1994)
MHC class I	AG**AA** GT **GAAA** CT	(–142/–153)	Driggers *et al*. (1990)
H-2D^{d}	AG**AA** GT **GAAA** CT	(–140/–151)	Harada *et al*. (1989)
IgλB	GG**AA** GT **GAAA** CC	(377/387)	Eisenbeis *et al*. (1993)
Dimeric repeats with 2 bp spacer			
IRF-2	G***AA*** GC ***GAAA*** AT **GAAA** TT	(–262/–247)	Neish *et al*. (1995)
IFN-α1	***AA*** CA ***GAAA*** TG **GAAA** GT	(–88/–73)	Neish *et al*. (1995)
BGP/C-CAM-1	G***AA*** AG ***GAAA*** GA **GAAA** GT	(–228/–214)	Chen *et al*. (1996)
ISG54	A***AA*** GG ***GAAA*** GT **GAAA** CT	(–84/–100)	Tanaka *et al*. (1993)
gp91^{PHOX}	T***AA*** AA ***GAAA*** AG **GAAA** CC	(–96/–80)	Luo and Skalnik (1996)
AT2	A***AA*** GA ***GAAA*** GA **GAAA** AT	(–283/–267)	Horiuchi *et al*. (1995)
V-CAM-1	GGAGT ***GAAA*** TA **GAAA** GT	(–1/–17)	Neish *et al*. (1995)
EBNA1	TTTGC ***GAAA*** AC **GAAA** GT	(–21/–5)	Schaefer *et al*. (1997)
INDO	AACTA ***GAAA*** AT **GAAA** CC	(–115/–99)	Konan and Taylor (1996)
gp91^{PHOX}	TAGTG ***GAAA*** AT **GAAA** CC	(–208/–224)	Luo and Skalnik (1996)
IFN-A4/A11	GTAAA ***GAAA*** GT **GAAA** AG	(–103/–87)	Genin *et al*. (1995)
IFN-A6	TTAAA ***GAAA*** GT **GAAA** AG	(–103/–87)	Au *et al*. (1993)
iNOS	ATTAT ***GAAA*** GT **GAAA** TA	(–908/–924)	Neish *et al*. (1995)
MyD88	TCTCG ***GAAA*** GC **GAAA** GA	(924/940)	Harroch *et al*. (1995)
PKR	GCCGG ***GAAA*** AC **GAAA** CA	(–76/–160)	Tanaka and Samuel (1994)
ICE	ACT ***GAAA*** CT **GAAA** G	(–41/–28)	Casano *et al*. (1994)
CBP	TCAAG ***GAAA*** CA **GAAA** CT	(–124/–140)	Tanaka *et al*. (1993)
GBP	AATAT ***GAAA*** CT **GAAA** GT	(–113/–129)	Tanaka *et al*. (1993)
Dimeric repeats with 3 bp spacer			
IL-6	T***AA*** AA ***GAAA*** **A**AA **GAAA** GT	(–270/–253)	Sanceau *et al*. (1995)
IFN-β(PRD-III)	CATAG ***GAAA*** **A**CT **GAAA** GG	(–97/–79)	Neish *et al*. (1995)
INDO	GTAAG ***GAAA*** **A**CT **GAAA** CC	(–1109/–1126)	Konan and Taylor (1996)
H4	AGATT ***GAAA*** **A**CC **GAAA** GC	(–48/–65)	Ramsey-Ewing *et al*. (1994)
Dimeric repeats with 1 bp spacer			
2′-5′OAS	CTGAG ***GAAA*** C **GAAA** CC	(–103/–88)	Tanaka *et al*. (1993)
Mx	GCTCA ***GAAA*** C **GAAA** CT	(–116/–131)	Tanaka *et al*. (1993)
Multiple repeats			
ISG15	CTGGG **GAAA** GG ***GAAA*** CC **GAAA** CT	(–117/–95)	Tanaka *et al*. (1993)
IFI-56K	TAGG ***GAAA*** CC **GAAA** GGG ***GAAA*** GT **GAAA** CT	(–90/–118)	Wathelet *et al*. (1998)
P31	AA**AA** CT **GAAA** GGGAG**AA** GT **GAAA** GTG	(–93/–64)	Wathelet *et al*. (1998)

HLA-B7, major histocompatibility complex class I heavy chain; ODC, ornithine decarboxylase; MHC class I, major histocompatibility complex class I; H-2D^{d}, mouse H-2D^{d} gene; BGP/C-CAM, biliary glycoprotein/cell CAM-1; ISG, interferon-stimulated gene; gp21^{PHOX}, phagocyte NADPH oxidase 21 kDa subunit; AT2, angiotensin II type 2 receptor; V-CAM-1, vascular cell adhesion molecule-1; EBNA1, Epstein–Barr virus nuclear antigen-1; INDO, indoleamine 2,3-dioxygenase; iNOS, inducible NO synthetase; PKR, double-stranded RNA-activated protein kinase; ICE, interleukin-1β converting enzyme; CBP, complement binding protein; GBP, guanylate binding protein; H4, histone H4; 2′-5′OAS, 2′,5′-oligoadenylate synthetase.

[a]The bases of IRS are in bold. The underlining and italic type indicate bases recognized by individual IRF molecules. These lines are overlapped for sites having tandem repeats of the GAAA core sequence.

of them, e.g. the binding site of the IRF-2 gene, also possess the 5′-flanking AA sequence so as to produce two complete IRS sequences. However, a number of the dimeric repeats, such as the site of VCAM-1, lack the 5′-flanking AA sequence. This indicates that IRF binding to the 5′ core sequence of the dimeric repeats is weaker than binding to the 3′ core sequence. Depending on the concentration of IRF, it may be possible that a single IRF molecule binds to these dimeric repeats and exerts physiological functions *in vivo*.

Of particular interest is that a class of dimeric repeats, including the PRD III site of the IFN-β gene, is endowed with a spacer of three bases, with the first base of the spacers being completely conserved as adenine so as to enable the repeat to produce a complete IRS at the 3′ side. Spacers of one base, seen in a few dimeric repeats such as the 2′-5′ OAS gene, also produce one complete IRS at the 3′ side. Some other IRF binding sites contain multiple repeats. In most cases, these repeats possess a two-base spacer like that seen in the triplet repeats of the ISG15 gene, suggesting a cooperative binding of three IRF molecules. A closer inspection of the multiple repeats, however, reveals alternatively sized spacers between the repeats. One example, the IFI-56K gene, consists of two dimeric repeats with a spacer of three bases that have no conserved adenine at the first position. The P31 gene exhibits two IRS sequences, with a longer spacer of five bases. At present, it remains unclear whether IRF molecules bind cooperatively to these sites.

This AA sequence is essential for the recognition of a single repeat. For instance, IRFs do not bind to the NF-κB binding site (PRD II) of the IFN-β gene, which contains a GTGGGAAA sequence containing the GAAA core sequence, but no 5′-flanking AA sequence. Some

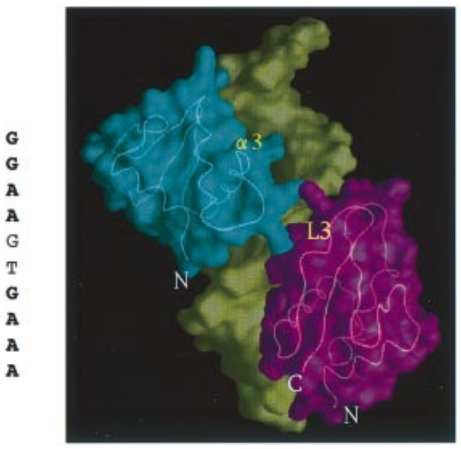

Fig. 6. Model for PU.1 and IRF-4 bound to DNA. A side view of surface representation showing the model for DNA binding domains of PU.1 (green) and IRF-4 (purple) bound to DNA (light green) containing the **GGAA**NN**GAAA** motif in the λB site of the enhancer of the immunoglobulin light chain gene.

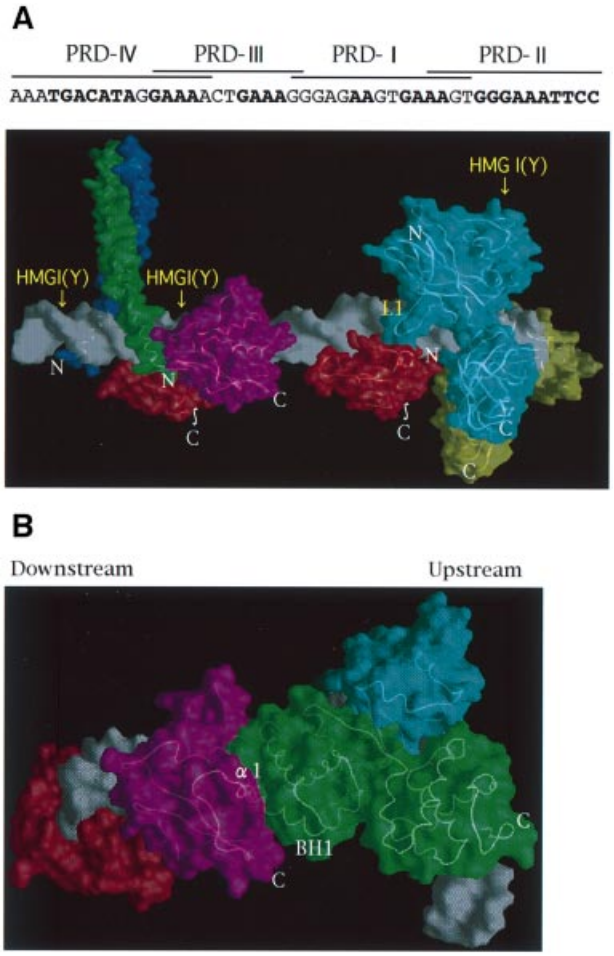

Fig. 7. Possible synergistic bindings at enhancer elements. (**A**) A side view of the surface representation of the IFN-β enhancer with the DNA binding domains of the transcription factors; IRFs (purple and wine red) bound to the PRD III and I sites, ATF-2–c-Jun (blue and green) bound to the PRD IV, and NF-κB, p65 (light green)–p50 (light blue) heterodimer bound to the PRD I sites. The DNA sequence and the binding sites are indicated at the top. Minor grooves of the HMG I(Y) binding sites are indicated with labels. The N and C termini are indicated to show the locations of their activation domains that link to the N termini of ATF-2–Jun and the C termini of IRF and p65. (**B**) A side view of the surface representation showing the IRF-2 DNA binding domains (purple and wine red) bound to the TATA-box region of the VCAM-1 gene, together with bound TBP (light blue) and the TFIIB core (green). An IRF-2 DNA binding domain (purple) bound to the upstream GAAA core sequence contacts with the N-terminal domain of the TFIIB core.

mutation studies also support the role of the 5′-flanking AA sequence (Eisenbeis *et al.*, 1995; Yee *et al.*, 1998). Moreover, EMSA experiments showed that mutations of the 5′-flanking AA sequence of a single IRS sequence reduce the binding of the IRF-2 DNA binding domain (data not shown). Similarly, reduced binding was observed by mutations of the conserved adenine of the 3 bp spacer of dimeric repeats. A full mutational analysis of these bases will be published elsewhere.

Strikingly, some of the single IRS sequences are overlapped with the Ets binding sequence, GGAA, producing a consensus sequence of **GGAA**NN**GAAA**, which is found in the immunoglobulin enhancers, $E_{\lambda 2\text{-}4}$, $E_{\kappa 3}$ and E_{μ}, as well as in the promoter of the macrophage scavenger receptor (Moreau-Gachelin, 1994). Biochemical studies have shown that PU.1 Ets and IRF-4/Pip/NF-EM5 bind cooperatively to the λB site of the immunoglobulin enhancer $E_{\lambda 2\text{-}4}$ (Eisenbeis *et al.*, 1995). This synergy is reconstituted in part with the DNA binding domains of the two transcription factors (Brass *et al.*, 1996; Yee *et al.*, 1998). The docking of the PU.1–DNA complex (Kodandapani *et al.*, 1996) onto this sequence indicates that there is no close contact between the two domains, but loop L3 could interact with the loop connected to the recognition helix of PU.1 (Figure 6). DNA distortions, however, might mediate the cooperativity since PU.1 also bends DNA toward the major groove. This model suggests that the N-terminal domain of PU.1 and the C-terminal domain of IRF might be located on the same DNA surface, and thus interact with each other to enhance the DNA binding of IRF-4. In contrast, the N-terminal tail of IRF-4, which inhibits the DNA binding, is located far from PU.1, suggesting no direct interaction between PU.1 and the inhibitory tail of IRF-4.

Implication of synergistic binding with other transcription factors

In the IFN-β enhancer, PRD IV is adjacent to PRD III, whereas no significant binding cooperativity has been observed between IRF-1 and ATF-2 (Thanos and Maniatis, 1995). To address this issue, a model of the IRF–PRD III complex was built alongside that of the Jun–Fos heterodimer (Glover and Harrison, 1995) bound to PRD IV (Figure 7A). Interestingly, the model reveals slight contacts of the bZIP basic region with the flexible loop L3 and the N-terminal region of helix α3 of IRFs, but no contact or steric clash between these DNA binding domains. This is in sharp contrast to the NFAT–Fos–Jun–DNA complex, in which the ZIP regions of both Fos and Jun are in several contacts with NFAT (Chen *et al.*, 1998a). Recently, it has been shown that HMG I(Y), an architectural protein, binds to the minor groove to bend the DNA toward the major groove (Huth *et al.*, 1997). Remarkably, one of the HMG I(Y) binding sites is overlapped with the 5′ repeat of PRD III, at a point where one of the IRF molecules contacts the minor groove of the 5′-flanking AA sequence of IRS. This dual binding to the minor groove explains

well the results that HMG I(Y) alone slightly inhibits binding of IRF-1 (Thanos and Maniatis, 1995). Furthermore, we point out that this HMG I(Y) binding site is tightly sandwiched between the ZIP region and the IRF DNA binding domain, suggesting that HMG I(Y) may induce binding synergy of IRF and ATF-2–c-Jun through contacts with both molecules and DNA distortions. We also tested the interferences between IRF and NF-κB at the PRD I and PRD II sites using the recently solved crystal structure of p50–p65 heterodimer–DNA (Chen *et al.*, 1998b). Interestingly, there is no steric clash between IRF-2 and p50–p65 heterodimer in our model, whereas the N-terminal residues of IRF stick into the interspace between DNA and the dimerization domains of the p65–p50 heterodimer and come into contact with the heterodimer. These contacts may result in strong interferences on the DNA binding of these two transcription factors. It is an interesting question where these transcription factors position the activation domains that interact with each other and/or with the other general transcription factors. Based on the present model, the locations of the C termini of IRFs, the N termini of ATF-2–c-Jun and the C terminus of p65 suggest that all regulatory domains of these transcription factors are positioned at the same surface of the DNA double helix. We note that all the minor grooves of the HMG I(Y) binding sites are faced toward the solvent region of the opposite DNA surface. Therefore, the DNA bendings by HMG I(Y), toward the major groove of the sites, might help in juxtaposing the regulatory domains closer together within the IFN-β enhanceosome. This seems to be consistent with the previous *in vitro* experiments (Falvo *et al.*, 1995).

The VCAM-1 gene has an IRF binding site located 10 bp downstream of the TATA box (Jesse *et al.*, 1998) and is activated by IRF-2 binding to this site, which is a dimeric repeat of the core sequence with a reverse 5′, 3′ orientation. Biochemical experiments have demonstrated physical interactions between the IRF-2 DNA binding domain and the TFIIB core (Wang *et al.*, 1996). The wealth of crystal structure determinations for TBP–TFIIB–DNA (Nikolov *et al.*, 1995) and TBP–TFIIA–DNA (Geiger *et al.*, 1996; Tan *et al.*, 1996) enables us to speculate on the protein contacts that might occur at the TATA-box region of this gene (Figure 7B). Our model indicates that one of the IRF-2 DNA binding domains, which binds to the upstream core sequence of the site, might contact the N-terminal domain of the TFIIB core, whereas this IRF-2 DNA binding domain would not come into any contact with the TFIIA subunits. The contact interfaces involve the N terminus and helix α1 of the IRF-2 DNA binding domain and helix BH1 and the BH2–BH3 loop of the TFIIB core. The present model reveals steric clashes involving these parts, suggesting that IRF-2 induces structural changes in the TFIIB–TBP–DNA complex. Interestingly, the C terminus of the IRF-2 DNA binding domain and the N terminus of the TFIIB core are closely positioned on the same surface. The C-terminal part of IRF-2 contains an acidic activation domain, which may interact with the putative Zn^{2+} binding region located in the additional N-terminal portion of TFIIB and modulate its essential role for assembly of pol II–TFIIF into the preinitiation complex (Hisatake *et al.*, 1993). To verify these hypotheses, further biochemical and structural studies will be required.

Conclusion

This study of the IRF-2 DNA binding domain bound to DNA provides the first view of a DNA-distortions-induced tandem binding of the DNA binding domains to a repeated consensus sequence and the first rational interpretation of various binding properties of the IRF family members. The high-resolution structure elucidates the water-mediated recognition of the base sequence at the major groove of the GAAA core sequence and the minor groove of the 5′-flanking AA sequence, and also suggests that the helix–hairpin motif of the metal binding site is a common structural module interacting with DNA. We propose IRS, **AA**NN**GAAA**, as the core sequence recognized by members of the IRF family. Our work also provides a clue for further biochemical and genetic studies of IRF-dependent transcriptional regulation of several genes. In particular, these data should aid directed, systematic analyses of the IFN-β enhanceosome assembly and the mechanisms by which RNA polymerase II enters the growing preinitiation complex to establish the transcriptional start state. Finally, possible contacts of the DNA binding domains of IRF-2 with TFIIB on the VCAM-1 genes provide a starting point for further analyses of how IRF-2 activates the gene expression.

Materials and methods

Protein and DNA preparation

The DNA binding domain of mouse IRF-2 was overexpressed in *E.coli* BL21(DE3) using a T7 expression system, and purified and crystallized as described previously (Kusumoto *et al.*, 1998). The protein, which consists of the N-terminal 113 residues (M_r 13 314), was purified by three column chromatographic steps, using SP-Sepharose, Mono-S (Pharmacia Biotech) and a gel filtration with Sephacryl S-100 (Pharmacia Biotech). The DNA oligomers used in the crystallization attempts were synthesized without the trityl group by standard phosphoramidite chemistry on an Applied Biosystem Model 394 synthesizer and purified by reverse-phase HPLC. The resulting single-stranded oligomers were quantified by UV spectrophotometry and mixed with an equimolar amount of a complementary strand for annealing in the presence of 100 mM KCl. The sequences of the DNA oligomers were based on the IRF-E consensus sequence and the binding sites of the IFN-β gene (Harada *et al.*, 1989; Tanaka *et al.*, 1993).

Crystallization and data collection

Crystals suitable for X-ray structural analysis were obtained at 4°C by the sitting-drop vapor diffusion method from solutions containing 50 mM Mes-K buffer pH 5.8, 50 mM KCl, 4% 2-methyl-2,4-pentanediol (2-MPD), and an equimolar mixture of the protein and the 13mer DNA shown in Figure 1A with a numbering scheme. The crystals belong to space group $P2_12_12_1$, with unit cell dimensions a = 90.7 Å, b = 101.0 Å, c = 171.6 Å. Heavy atom derivatives of the complex were obtained by using DNA oligomers with 5-iodouracil substituted for thymine and 5-iodocytosine substituted for cytosine at the positions indicated in Table I. In addition, a heavy atom derivative was prepared by soaking crystals in a solution of 0.3 mM $HgCl_2$. In the course of preparing the derivatives, we found that a heavy atom derivative obtained by using DNA oligomers with 5-iodouracil substituted at positions 10 and 2′ crystallized in hexagonal space group $P6_522$ with unit cell dimensions a = 132.3 Å, b = 132.3 Å, c = 296.5 Å. This crystal diffracted to 2.2 Å resolution. Following structural determination of the orthorhombic crystal form, the data set of this derivative was used at the final structure refinement. Diffraction data were collected using a Rigaku R-AXIS IV imaging plate detector mounted on a Rigaku FR-C generator, and also using the beamline 6B of the Photon Factory,

Table II. Data collection and refinement statistics

Diffraction data ($I > 1\sigma$)

Data set	Resolution (Å)	Reflections (measured/ unique)	Completeness (%) (overall/outer shell)	R_{sym} (%)	R_{deri} (%)	Number of sites	Phasing power (acentric/ centric	R_{Cullis}	Overall figure of merit
Orthorhombic form									
Native	2.8	143 615/35 660	92.4/81.2	9.3	–	–	–	–	
$HgCl_2$	3.3	40 906/24 384	77.5/54.0	8.1	31.1	7	0.85/0.56	0.89	
$I^5dC(11')$	3.0	68 898/24 384	74.7/54.9	10.0	13.9	7	0.74/0.65	0.81	
$I^5dC(3')$	3.0	97 980/24 641	75.6/59.6	9.2	12.9	7	0.66/0.48	0.87	
$I^5dU(7')$	3.1	58 702/22 125	73.6/50.2	11.7	16.3	7	0.59/0.45	0.92	
$I^5dU(10, 7')$	3.0	62 849/26 161	79.7/52.1	9.2	19.8	14	0.68/0.54	0.89	
$I^5dU(10)$	3.3	80 459/22 458	90.8/77.1	10.0	18.3	7	0.66/0.52	0.88	
									0.46
Hexagonal form									
$dU^5I(10, 2')$	2.2	2 494 554/69 536	88.4/76.1	6.2					

Refinement statistics

	Protein–DNA atoms	Solvent molecules (ions)	Resolution limit	R_{cryst}/R_{free} (%)	Mean B-factor	r.m.s. deviations			
						Bonds	Angles	Dihedrals	Impropers
Orthorhombic form									
	8141	0	10–2.8 Å	23.2/28.6	56.8 $Å^2$	0.012 Å	1.7°	24.0°	1.9°
Hexagonal form									
	6994	396 (6)	10–2.2 Å	20.2/24.3	26.1 $Å^2$	0.010 Å	1.6°	18.5°	1.6°

$R_{sym} = \Sigma \mid I - < I > \mid / \Sigma I$; $R_{deri} = \Sigma \parallel F_{PH} \mid - \mid F_P \parallel / \Sigma \mid F_P \mid$; phasing power = r.m.s. heavy atom structure factor/residual lack of closure; $R_{Cullis} = \Sigma \parallel F_{PH} - F_P \mid - \mid F_{H(calc)} \parallel / \Sigma \mid F_{PH} - F_P \mid$; R_{cryst} and $R_{free} = \Sigma \parallel F_o \mid - \mid F_c \parallel / \Sigma \mid F_o \mid$, where the free reflections (10% of the total used) were held aside for R_{free} throughout refinement.

Tsukuba, Japan. Using 25% 2-MPD as a cryoprotectant, the crystals were flash frozen in a nitrogen stream at –170°C. Data processing and reduction were carried out using the programs DENZO/SCALEPACK (Otwinowski and Minor, 1997) and PROCESS (Rigaku) (summarized in Table II).

Structure determination and refinement

Using the data of the orthorhombic form, the heavy atom parameters were refined, and the initial MIR phases as calculated with the program MLPHARE (Collaborative Computational Project Number 4, 1994) had a mean figure of merit of 0.46 to 3.5 Å. The phases were improved with solvent flattening and histogram matching with the program DM (Cowtan and Main, 1996). A model was built into the MIR electron density maps with the program O (Jones *et al*., 1991) and refined by simulated annealing with the program X-PLOR (Brunger, 1992). Refinement of the orthorhombic form resulted in a final crystallographic *R* factor of 23.2% and a free *R* factor of 28.6% for all data ($F > \sigma F$) between 10.0 and 2.8 Å. Using one DNA duplex and two IRF-2 DNA binding domains derived from the orthorhombic structure as a search model, molecular replacement solutions of the hexagonal form were found with the program AMoRe (Navaza, 1994) for three independent complexes in the asymmetric unit. The initial *R* factor was 35% at 2.5 Å resolution. The structure was refined in a manner similar to that used for the orthorhombic form. In the course of the analysis, the crystal structure of IRF-1 appeared (Escalante *et al.*, 1998). The appearance of a structure displaying a similar fold of IRF-2 seemed to provide a useful opportunity to review our structure, but significant disagreements also appeared among these structures. We were not able to reconcile the discrepancies in our structure refinement at 2.2 Å resolution (see Results). The final refinement converged to an *R* factor of 20.2% and an R_{free} of 24.3% for all data ($F > \sigma F$) between 10.0 and 2.2 Å. The average *B* factor was 26.1 $Å^2$. None of the protein residues were in disallowed regions of the Ramachandran plot, and 89.9% of residues were in the most favorable region as defined by the program PROCHECK (Laskowski *et al*., 1993). The final model was composed of six proteins of residues 5–113 (5484 atoms), three double-stranded DNA (1510 atoms), six potassium ions and 396 water molecules whose *B* factors were <35 $Å^2$. The N-terminal four residues of IRF-2 were disordered and therefore not included. The DNA duplexes were stacked together with both unpaired bases flipping out toward the solvent region. The flipped nucleotides were highly mobile. Only two of the six flipped nucleotides were included in the final model.

Structure inspection and model building

The helical parameters of DNA were analyzed with the program FREEHELIX (Dickerson, 1998). The models for the transcription factors bound to the IgλB, VCAM-1 and IFN-β enhancers were built by joining together the crystal structures of the protein–DNA complexes by superimposing the flanking 3 or 4 bp using a least squares fitting method. The ribbon representation of the protein was drawn using the program MOLSCRIPT (Kraulis, 1991), and the surface representation of the protein was drawn using the program GRASP (Nicholls *et al*., 1991).

Electrophoretic mobility shift assay

The 16–27 bp synthetic oligonucleotides, purchased from Nisshin Seihun, were used for EMSAs. Each binding solution (a final volume of 5 μl) contained 6 pmol of the DNA oligomer (1.2 μM final concentration), 1.5–15 pmol of IRF-2(113) (0.3–3 μM final concentration), 50 mM KCl, 1 mM NaN_3, 10 mM dithiothreitol and 20% (w/v) glycerol. Poly(dI-dC) at 40 μg/ml was used as a non-specific DNA competitor. The binding solutions were incubated for 1 h at 4°C and were electrophoresed at 4°C on a native 8% polyacrylamide gel at 40 mA for 1 h. The running buffer solution contained 40 mM Tris–HCl pH 7.0, 20 mM acetic acid, 1 mM EDTA and 50 mM KCl. Followed by quantification of the shift bands on the gel using an imaging analyzer (Atto AE6900MF), K_d values for the first and second bindings were calculated from the shift bands. The DNA sequence used in Figure 4C was 5′-AATGAC**AA**GT-**GAAA**GT**GAAA**GTGTGCC-3′.

Acknowledgements

The authors thank Drs Gourisankar Ghosh and Aneel K.Aggarwal for providing the coordinates of the p65–p50–DNA and IRF-1–DNA

complexes, respectively, and J.Kuriyan for critical reading. This work was supported by Grants in Aid for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan to T.H. (Nos 09308025, 10359003, 10179104, 10129221 and 09277102) and in part by the Tsukuba Advanced Research Alliance (TARA) project. T.H. is a TARA guest researcher for the Sakabe project. Coordinates of the refined IRF-2–DNA complex has been deposited in the Protein Data Bank (Brookhaven Laboratory), accession coded 2irf.

References

Au,W.C., Su,Y., Raj,N.B.K. and Pitha,P.M. (1993) Virus-mediated induction of interferon A gene requires cooperation between multiple binding factors in the interferon-α promoter region. *J. Biol. Chem.*, **268**, 24032–24040.

Berman,H.M., Zardecki,C. and Westbrook,J. (1998) The nucleic acid database: A resource for nucleic acid science. *Acta Crystallogr. D*, **54**, 1095–1104.

Brass,A.L., Kehrli,E., Eisenbeis,C.F., Dtroub,U. and Singh,H. (1996) Pip, a lymphoid-restricted IRF, contains a regulatory domain that is important for autoinhibition and ternary complex formation with the Ets factor PU.1. *Genes Dev.*, **10**, 2335–2347.

Brunger,A. (1992) *X-PLOR (Version 3.1) Manual*. Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney Avenue, New Haven, CT.

Casano,F.J., Rolando,A.M., Mudgett,J.S. and Molineaux,S.M. (1994) The structure and complete nucleotide sequence of the murine gene encoding interleukin-1β converting enzyme (ICE). *Genomics*, **20**, 474–481.

Chen,C.J., Lin,T.T. and Shively,J.E. (1996) Role of interferon regulatory factor-1 in the induction of biliary glycoprotein (cell CAM-1) by interferon-γ. *J. Biol. Chem.*, **271**, 28181–28188.

Chen,L., Glover,J.N.M., Hogan,P.G., Rao,A. and Harrison,S.C. (1998a) Structure of the DNA-binding domains from NFAT, Fos and Jun bound specifically to DNA. *Nature*, **392**, 42–48.

Chen,F.E., Huang,D.-B., Chen,Y.-Q. and Ghosh,G. (1998b) Crystal structure of p50/p65 heterodimer of transcription factor NF-κB bound to DNA. *Nature*, **391**, 410–413.

Clark,K.L., Halay,E.D., Lai,E. and Burley,S.K. (1993) Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature*, **364**, 412–420.

Collaborative Computational Project Number 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D*, **50**, 760–763.

Cowtan,K. and Main,P. (1996) Phase combination and cross validation in iterated density-modification calculation. *Acta Crystallogr. D*, **52**, 43–48.

Darnell,J.E.,Jr, Kerr,I.M. and Stark,G.R. (1994) Jac-STAT pathways and transcriptional activation in response to IFNs and other extracellular signaling proteins. *Science*, **264**, 1415–1421.

Dickerson,R.E. (1998) DNA bending: the prevalence of kinkiness and the virtues of normality. *Nucleic Acids Res.*, **26**, 1906–1926.

DiGabriele,A.D. and Steitz,T.A. (1993) A DNA dodecamer containing an adenine tract crystallizes in a unique lattice and exhibits a new bend. *J. Mol. Biol.*, **231**, 1024–1039.

Drew,H.D. and Dickerson,R.E. (1981) Structure of a β-DNA dodecamer: III. Geometry of hydration. *J. Mol. Biol.*, **151**, 535–556.

Driggers,P.H., Ennist,D.L., Gleason,S.L., Mak,W.-H., Marks,M.S., Levi,B.-Z., Flanagan,J.R., Appella,E. and Ozato,K. (1990) An interferon-γ-regulated protein that binds the interferon-inducible enhancer element of major histocompatibility complex class I genes. *Proc. Natl Acad. Sci. USA*, **87**, 3743–3747.

Eisenbeis,C.F., Singh,H. and Storb,U. (1993) PU.1 is a component of a multi-protein complex which binds an essential site in the murine immunoglobulin 2-4 enhancer. *Mol. Cell. Biol.*, **13**, 6452–6461.

Eisenbeis,C.F., Singh,H. and Storb,U. (1995) Pip, a novel IRF family member, is a lymphoid-specific, PU.1-dependent transcription activator. *Genes Dev.*, **9**, 1377–1387.

Escalante,C.R., Yie,J., Thanos,D. and Aggarwal,A.K. (1998) Structure of IRF-1 with bound DNA reveals determinants of interferon regulation. *Nature*, **391**, 103–106.

Falvo,J.V., Thanos,D. and Maniatis,T. (1995) Reversal of intrinsic DNA bends in the IFNβ gene enhancer by transcription factors and the architectural protein HMG I (Y). *Cell*, **83**, 1101–1111.

Fujita,T., Ohno,S., Yasumitsu,H. and Taniguchi,T. (1985) Delimitation and properties of DNA sequences required for the regulated expression of human interferon-β gene. *Cell*, **41**, 489–496.

Fujita,T., Sakakibara,J., Sudo,Y., Miyamoto,M., Kimura,Y. and Taniguchi,T. (1988) Evidence for a nuclear factor (s), IRF-1, mediating induction and silencing properties to human IFN-β gene regulatory elements. *EMBO J.*, **7**, 3397–3405.

Furui,J., Uegaki,K., Yamazaki,T., Shirakawa,M., Swindells,M.B., Harada,H., Taniguchi,T. and Kyogoku,Y. (1998) Solution structure of the IRF-2 DNA-binding domain: a novel subgroup of the winged helix-turn-helix family. *Structure*, **6**, 491–500.

Gao,S.J., Boshoff,C., Jayachandra,S., Weiss,R.A., Chang,Y. and Moore,P.S. (1997) KSHV ORF K9 (vIRF) is an oncogene which inhibits the interferon signaling pathway. *Oncogene*, **15**, 1979–1985.

Geiger,J.H., Hahn,S., Lee,S. and Sigler,P.B. (1996) Crystal structure of the yeast TFIIA/TBP/DNA complex. *Science*, **272**, 830–840.

Genin,P., Braganca,J., Darracq,N., Doly,J. and Civas,A. (1995) A novel PRD I and TG binding activity involved in virus-induced transcription of IFN-A genes. *Nucleic Acids Res.*, **23**, 5055–5063.

Glover,J.N.M. and Harrison,S.C. (1995) Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. *Nature*, **373**, 257–261.

Han,G.W., Kopka,M.L., Cascio,D., Grzeskowiak,K. and Dickerson,R.E. (1997) Structure of a DNA analog of the primer for HIV-1 RT second strand synthesis. *J. Mol. Biol.*, **269**, 811–826.

Harada,H., Fujita,T., Miyamoto,M., Kimura,Y., Maruyama,M., Furia,A., Miyata,T. and Taniguchi,T. (1989) Structurally similar but functionally distinct factors, IRF-1 and IRF-2, bind to the same regulatory elements of IFN and IFN-inducible genes. *Cell*, **58**, 729–739.

Harada,H., Kitagawa,M., Tanaka,N., Yamamoto,H., Harada,K., Ishihara,M. and Taniguchi,T. (1993) Anti-oncogenic and oncogenic potentials of interferon regulatory factors-1 and -2. *Science*, **259**, 971–974.

Harroch,S., Gothelf,Y., Revel,M. and Chebath,J. (1995) 5′ upstream sequences of MyD88, an IL-6 primary response gene in M1 cells: detection of functional IRF-1 and Stat factors binding sites. *Nucleic Acids Res.*, **23**, 3539–3546.

Hisatake,K., Roeder,R. and Horikoshi,M. (1993) Functional dissection of TFIIB domains required for TFIIB-TFIID-promoter complex formation and basal transcription activity. *Nature*, **363**, 744–747.

Horiuchi,M., Koike,G., Yamada,T., Mukoyama,M., Nakajima,M. and Dzau,V.J. (1995) The growth-dependent expression of angiotensin II type 2 receptor is regulated by transcription factors interferon regulatory factor-1 and -2. *J. Biol. Chem.*, **270**, 20225–20230.

Huth,J.R., Bewley,C.A., Nissen,M.S., Evans,J.N.S., Reeves,R., Gronenborn,A.M. and Clore,G.M. (1997) The solution structure of an HMG-I (Y)-DNA complex defines a new architectural minor groove binding motif. *Nature Struct. Biol.*, **4**, 657–665.

Jesse,T.L., LaChance,R., Iademarco,M.F. and Dean,D.C. (1998) Interferon regulatory factor-2 is a transcriptional activator in muscle where it regulates expression of vascular cell adhesion molecule-1. *J. Cell Biol.*, **140**, 1265–1276.

Jones,T.A., Zou,J.-Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110–119.

Kim,J.L., Nikolov,D.B. and Burley,S.K. (1993) Co-crystal structure of TBP recognizing the minor groove of a TATA element. *Nature*, **365**, 520–527.

Kim,T.K. and Maniatis,T. (1997) The mechanism of transcriptional synergy of an in vitro assembled interferon-β enhanceosome. *Mol. Cell*, **1**, 119–129.

Kodandapani,R., Pio,F., Ni,C.-Z., Piccialli,G., Klemsz,M., McKercher,S., Maki,R.A. and Ely,K.R. (1996) A new pattern for helix-turn-helix recognition revealed by the PU.1 ETS-domain-DNA complex. *Nature*, **380**, 456–460.

Konan,K.V. and Taylor,M.W. (1996) Importance of the two interferon-stimulated response element (ISRE) sequences in the regulation of the human indoleamine 2,3-deoxygenase gene. *J. Biol. Chem.*, **271**, 19140–19145.

Kraulis,P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.

Kusumoto,M., Fujii,Y., Tshukuda,Y., Ohira,T., Kyogoku,Y., Taniguchi,T. and Hakoshima,T. (1998) Crystallographic characterization of the DNA-binding domain of interferon regulatory factor-2 complexed with DNA. *J. Struct. Biol.*, **121**, 363–366.

Laskowski,R.A., MacArthur,M.W., Moss,D.S. and Thornton,J.M. (1993) PROCHECK: a program to check the stereochemical quality of protein structure. *J. Appl. Crystallogr.*, **26**, 283–291.

Lewis,M., Chang,G., Horton,N.C., Kercher,M.A., Pace,H.C., Schumacher, M.A., Brennan,R.G. and Lu,P. (1996) Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science*, **271**, 1247–1254.

Love,J.J., Li,X., Case,D.A., Giese,K., Grosschedl,R. and Wright,P.E. (1995) Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature*, **376**, 791–795.

Luo,W. and Skalnik,D.G. (1996) Interferon regulatory factor-2 directs transcription from the gp91^{phox} promotor. *J. Biol. Chem.*, **271**, 23445–23451.

Maniatis,T., Whittemore,L.A., Du,W., Fan,C.M., Keller,A., Palombella,V. and Thanos,D. (1992) Positive and negative control of human interferon-β gene expression. In McKnight,S.L. and Yamamoto,K. (eds), *Transcription Regulation, Part 2*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Manzella,L., Giuffrida,M.A., Pilastoro,M.R., Girlando,G., Picardi,G., Malaguarnera,L. and Messina,A. (1994) Possible role of the transcription factor interferon regulatory factor 1 (IRF-1) in the regulation of ornithine decarboxylase (ODC) gene expression during IFN-γ macrophage activation. *FEBS Lett.*, **348**, 177–180.

Marie,I., Durbin,J.E. and Levy,D.E. (1998) Differential viral induction of distinct interferon-α genes by positive feedback through interferon reguratory factor-7. *EMBO J.*, **17**, 6660–6669.

Miyamoto,M., Fujita,T., Kimura,Y., Maruyama,M., Harada,H., Sudo,Y., Miyata,T. and Taniguchi,T. (1988) Regulated expression of a gene encoding a nuclear factor, IRF-1, that specifically binds to IFN-β gene regulatory elements. *Cell*, **54**, 903–913.

Moore,P.S., Boshoff,C., Weiss,R.A. and Chang,Y. (1996) Molecular mimicry of human cytokine and cytokine response pathway genes by KSHV. *Science*, **274**, 1739–1744.

Moreau-Gachelin,F. (1994) Spi-1/PU.1: an oncogene of the Ets family. *Biochim. Biophys. Acta*, **1198**, 149–163.

Navaza,J. (1994) AMoRe: An automated package for molecular replacement. *Acta Crystallogr. A*, **50**, 157–163.

Neish,A.S., Read,M.A., Thanos,D., Pine,R., Maniatis,T. and Collins,T. (1995) Endothelial interferon regulatory factor 1 cooperates with NF-κB as a transcriptional activator of vascular cell adhesion molecule 1. *Mol. Cell. Biol.*, **15**, 2558–2569.

Nelson,H.C.M., Finch,J.T., Luisi,B.F. and Klug,A. (1987) The structure of an oligo (dA)-oligo (dT) tract and its biological implications. *Nature*, **330**, 221–226.

Nguyen,H., Hiscott,J. and Pitha,P.M. (1997) The growing family of interferon regulatory factors. *Cytokine Growth Factor Rev.*, **8**, 293–312.

Nicholls,A., Sharp,K.A. and Honing,B. (1991) Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins*, **11**, 281–296.

Nikolov,D.B., Chen,H., Halay,E.D., Usheva,A.A., Hisatake,K., Lee,D.K., Roeder,R.G. and Burley,S.K. (1995) Crystal structure of a TFIIB-TBP-TATA element ternary complex. *Nature*, **377**, 119–128.

Ogasawara,K., Hida,S., Azimi,N., Tagaya,Y., Sato,T., Yokochi-Fukuda,T., Waldmann,T.A., Taniguchi,T. and Taki,S. (1998) Requirement for IRF-1 in the microenvironment supporting development of natural killer cells. *Nature*, **391**, 700–703.

Otwinowski,Z. and Monor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.

Pelletier,H. and Sawaya,M.R. (1996) Characterization of the metal ion binding helix-hairpin-helix motifs in human DNA polymerase β by X-ray structural analysis. *Biochemistry*, **35**, 12778–12787.

Pelletier,H., Sawaya,M.R., Wolfle,W., Wilson,S.H. and Kraut,J. (1996) Crystal structure of human DNA polymerase β complexed with DNA: implications for catalytic mechanism, processivity and fidelity. *Biochemistry*, **35**, 12742–12761.

Ptashne,M. (1988) How eukaryotic transcriptional activators work. *Nature*, **335**, 683–689.

Ptashne,M. and Gann,A. (1997) Transcription activation by recruitment. *Nature*, **386**, 569–577.

Ramsey-Ewing,A., Wijen,A.J.V., Stein,G.S. and Stein,J.L. (1994) Delineation of human histone H4 cell cycle element in vivo: The master switch for H4 gene transcription. *Proc. Natl Acad. Sci. USA*, **91**, 4475–4479.

Rice P.A., Yang,S., Mizuuchi,K. and Nash,H.A. (1996) Crystal structure of an IHF-DNA complex: a protein-induced DNA U-turn. *Cell*, **87**, 1295–1306.

Rudnick,J. and Bruinsma,R. (1999) DNA-protein cooperativity binding through variable-range elastic coupling. *Biophys. J.*, **76**, 1725–1733.

Sanceau,J., Kaisho,T., Hirano,T. and Wietzerbin,J. (1995) Triggering of the human interleukin-6 gene by interferon-γ and tumor necrosis factor-α in monocytic cells involves cooperation between interferon regulatory factor-1, NF-κB and Sp1 transcription factors. *J. Biol. Chem.*, **270**, 27920–27931.

Sato,M., Tanaka,N., Hata,N., Oda,E. and Taniguchi,T. (1998a) Involvement of the IRF family transcription factor IRF-3 in virus-induced activation of the IFN-β gene. *FEBS Lett.*, **425**, 112–116.

Sato,M., Hata,N., Asagiri,M., Nakaya,T., Taniguchi,T. and Tanaka,N. (1998b) Positive feedback regulation of type I IFN genes by the IFN-inducible transcription factor IRF-7. *FEBS Lett.*, **441**, 106–110.

Schaefer,B.C., Paulson,E., Strominger,J.L. and Speck,S.H. (1997) Constitutive activation of Epstein-Barr virus (EBV) nuclear antigen 1 gene transcription by IRF-1 and IRF-2 during restricted EBV latency. *Mol. Cell. Biol.*, **17**, 873–886.

Schultz,S.C., Shields,G.C. and Steitz,T.A. (1991) Crystal structure of a CAP-DNA complex: the DNA is bent by 90°. *Science*, **253**, 1001–1007.

Schumacher,M.A., Choi,K.Y., Zalkin,H. and Brennan,R.G. (1994) Crystal structure of LacI member, PurR, bound to DNA: minor groove binding by α-helices. *Science*, **266**, 763–770.

Taki,S. *et al*. (1997) Multistage regulation of Th-1 type immune responses by the transcription factor IRF-1. *Immunity*, **6**, 673–679.

Tan,S., Hunziker,Y., Sargent,D.F. and Richmond,T.J. (1996) Crystal structure of a yeast TFIIA/TBP/DNA complex. *Nature*, **381**, 127–134.

Tanaka,H. and Samuel,C.E. (1994) Mechanism of interferon action: structure of the mouse PKR gene encoding the interferon-inducible RNA-dependent protein kinase. *Proc. Natl Acad. Sci. USA*, **91**, 7995–7999.

Tanaka,N., Kawakami,T. and Taniguchi,T. (1993) Recognition DNA sequences of interferon regulatory factor 1 (IRF-1) and IRF-2, regulators of cell growth and the interferon system. *Mol. Cell. Biol.*, **13**, 4531–4538.

Tanaka,N. *et al*. (1994) Cellular commitment to oncogene-induced transformation or apoptosis is dependent on the transcription factor IRF-1. *Cell*, **77**, 829–839.

Tanaka,N. *et al*. (1996) Cooperation of the tumour suppressors IRF-1 and -p53 in response to DNA damage. *Nature*, **382**, 816–818.

Taniguchi,T., Harada,H. and Lamphier,M.S. (1995) Regulation of the interferon system and cell growth by the IRF transcription factors. *J. Cancer Res. Clin. Oncol.*, **121**, 516–520.

Taniguchi,T., Lamphier,M.S. and Tanaka,N. (1997) IRF-1: the transcription factor linking the interferon response and oncogenesis. *Biochim. Biophys. Acta*, **1333**, M9–M17.

Thanos,D. and Maniatis,T. (1995) Virus induction of human IFN-β gene expression requires the assembly of an enhanceosome. *Cell*, **83**, 1091–1100.

Thayer,M.M., Aheren,H., Xing,D., Cunningham,R.P. and Tainer,J.A. (1995) Novel DNA binding motifs in the DNA repair enzyme endonuclease III crystal structure. *EMBO J.*, **14**, 4108–4120.

Uegaki,K., Shirakawa,M., Fujita,T., Taniguchi,T. and Kyogoku,Y. (1993) Characterization of the DNA binding domain of the mouse IRF-2 protein. *Protein Eng.*, **6**, 195–200.

Vaughan,P.S., Aziz,F., van Wijnen,A.J., Wu,S., Harada,H., Taniguchi,T., Soprano,K.J., Stein,J.L. and Stein,G.S. (1995) Activation of a cell-cycle-regulated histone gene by the oncogenic transcription factor IRF-2. *Nature*, **377**, 362–365.

Wang,I.-M., Blanco,J.C.G., Tsai,S.Y., Tsai,M.-J. and Ozato,K. (1996) Interferon regulatory factors and TFIIB cooperatively regulate interferon-responsive promoter activity in vivo and in vitro. *Mol. Cell. Biol.*, **16**, 6313–6324.

Wathelet,M.G., Lin,C.H., Parekh,B.S., Ronco,L.V., Howley,P.M. and Maniatis,T. (1998) Virus infection induces the assembly of coordinately activated transcription factors on the IFN-β enhancer in vivo. *Mol. Cell*, **1**, 507–518.

Wu,H.-M. and Crothers,D.M. (1984) The locus of sequence-directed and protein-induced DNA bending. *Nature*, **308**, 509–513.

Yee,A.A., Yin,P., Siderovski,D.P., Mak,T.W., Litchfield,D.W. and Arrowsmith,C.H. (1998) Cooperative interaction between the DNA-binding domains of PU.1 and IRF-4. *J. Mol. Biol.*, **279**, 1075–1083.

Yoneyama,M., Suhara,W., Fukuhara,Y., Fukuda,M., Nishida,E. and Fujita,T. (1998) Direct triggering of the type I interferon system by virus infection: activation of a transcription factor complex containing IRF-3 and CBP/p300. *EMBO J.*, **17**, 1087–1095.

Zimring,J.C., Goodbourn,S. and Offermann,M.K. (1998) Human herpesvirus 8 encodes an interferon regulatory factor (IRF) homolog that represses IRF-1-mediated transcription. *J. Virol.*, **72**, 701–707.

Received January 11, 1999; revised June 1, 1999; accepted August 2, 1999