© 2000 Nature America Inc. • http://structbio.nature.com

Structural basis for the diversity of DNA recognition by bZIP transcription factors

Yoshifumi Fujii[1], Toshiyuki Shimizu[1], Takashi Toda[2], Mitsuhiro Yanagida[3] and Toshio Hakoshima[1]

[1]*Department of Molecular Biology, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0101, Japan.* [2]*Laboratory of Cell Regulation, Imperial Cancer Research Fund, P.O. Box 123, 44 Lincoln's Inn Fields, London, WC2A 3PX, UK.* [3]*Department of Biophysics, Faculty of Science, Kyoto University, Kitashirakawa-oiwake, Sakyo-ku, Kyoto 606-8502, Japan.*

The basic region leucine zipper (bZIP) proteins form one of the largest families of transcription factors in eukaryotic cells. Despite relatively high homology between the amino acid sequences of the bZIP motifs, these proteins recognize diverse DNA sequences. Here we report the 2.0 Å resolution crystal structure of the bZIP motif of one such transcription factor, PAP1, a fission yeast AP-1-like transcription factor that binds DNA containing the novel consensus sequence TTACGTAA. The structure reveals how the Pap1-specific residues of the bZIP basic region recognize the target sequence and shows that the side chain of the invariant Asn in the bZIP motif adopts an alternative conformation in Pap1. This conformation, which is stabilized by a Pap1-specific residue and its associated water molecule, recognizes a different base in the target sequence from that in other bZIP subfamilies.

Approximately three decades have passed since the question was first raised regarding what is the basis for specific DNA recognition by proteins. It is now clear that no general rule for DNA recognition will be found that can apply to all transcription factors. However, it seems possible that certain rules will emerge for select members of individual structural families[1–4]. The basic region leucine zipper (bZIP) proteins form one of the largest families of transcription factors in eukaryotic cells. These proteins dimerize through their leucine repeats to form a coiled coil with two flanking α-helices that constitute the basic region that contacts DNA bases. The dimers are capable of binding short palindromic or pseudo-palindromic target sequences. The AP-1 site (TGACTCA) and the CREB site (TGACGTCA) (the most conserved palindromic bases are in bold and underlined; Fig. 1) are typical consensus sequences recognized by members of the bZIP subfamily. Members of the other bZIP subfamilies, however, have preferences for divergent DNA sequences; one such bZIP transcription factor is Pap1 from fission yeast, which binds TTACGTAA and TTAGTAA sequences[5] and has important roles in multidrug resistance[6] and oxidative stress response[7]. Members of the yeast AP-1-like transcription factor (YAP)[8] and CAMP response element binding protein-2 (CREB-2)[9] subfamilies also prefer the AT-rich binding sites of Pap1. The DNA bases of both the CREB[10,11] and the AP-1 (ref. 12) sites are recognized through direct interactions with five conserved amino acid residues in the signature sequence of the bZIP basic regions, namely NXXAAXXCR (X represents variable amino acid residue type). Interestingly, in the signature sequence of Pap1 subfamily members, the second Ala is replaced with Gln and the Cys is replaced with Phe to give the sequence NXXAQXXFR. The characteristic DNA recognition by members of the Pap1 subfamily seems, therefore, to be an intrinsic property of their basic region. Based on the structure of a Pap1–DNA complex, we characterize the basis of stereochemical recognition that describes the binding of bZIP motifs to related but different DNA sequences.

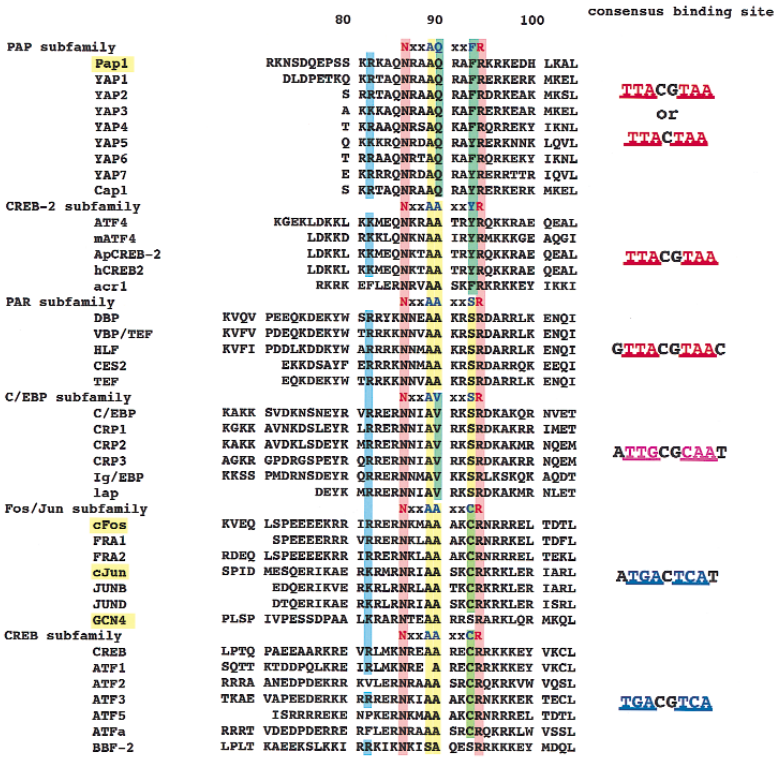

Fig. 1 Amino acid sequence comparison of the basic regions of members from several bZIP subfamilies. The sequences are aligned according to similarity of the basic regions and are terminated at the first Leu residue within the leucine zipper. For each subfamily, the signature sequence for DNA recognition, such as NXXAQXXFR for the Pap1 subfamily, is shown opposite the subfamily name. The proposed consensus DNA sequence for each subfamily is shown on the right. The two invariant amino acid residues in all subfamily members are highlighted in pink, and other colored residues are discussed in the text.

Overall structure

Crystals of the Pap1 bZIP motif (70 residues) in complex with a self-complementary 13-mer DNA (Fig. 2*a*,*b*) were obtained as described[26]. The structure was solved by multiple isomorphous replacement (MIR) and was refined at 2.0 Å resolution to an R-factor of 23.0% (R_{free} 25.3%). Like GCN4 and Fos-Jun, the Pap1 dimers grip the major groove of the DNA like a pair of chopsticks (Fig. 2*c*). The structures of two crystallographically independent Pap1–DNA complexes in the asymmetric unit are essentially the same. The structures of the two chains in each Pap1 dimer, however, are significantly different from each other; one chain is rather straight, the other bent. This bend is probably imposed on the chain by its different crystal packing environment, which was observed in the structures of Fos-Jun–DNA[12] and Fos-Jun–NFAT–DNA[13] complexes. A helical wheel diagram of the

© 2000 Nature America Inc. • http://structbio.nature.com

© 2000 Nature America Inc. • http://structbio.nature.com

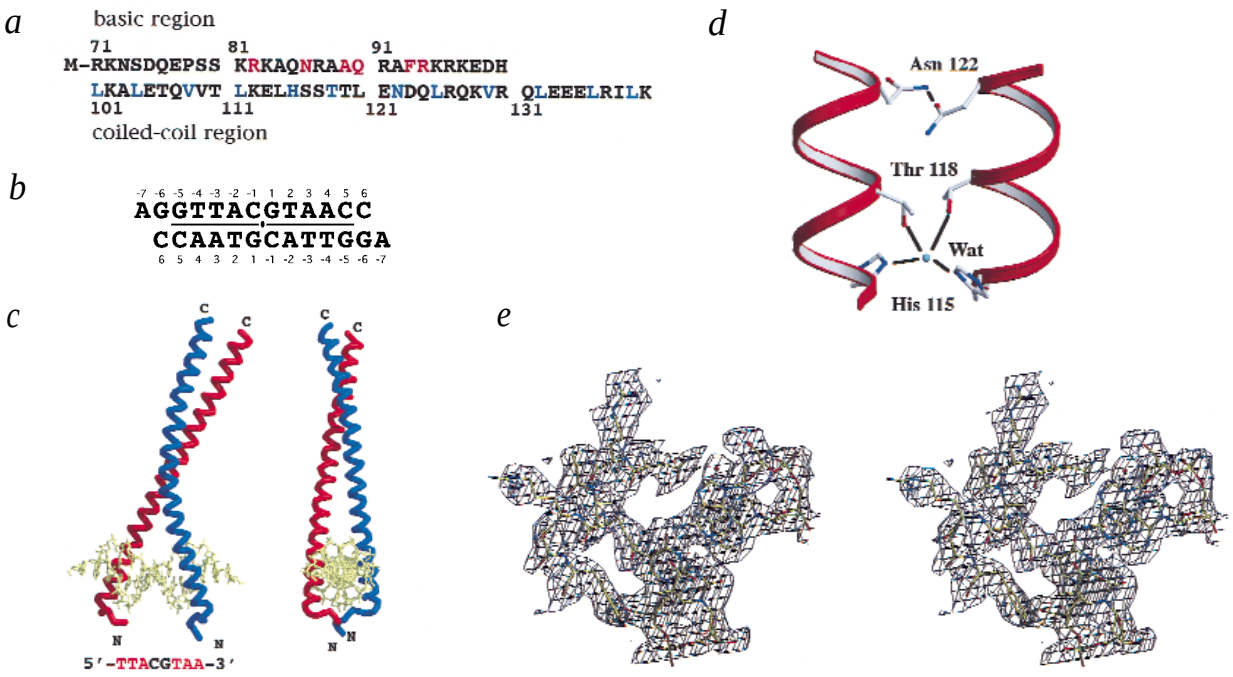

Fig. 2 Sequences of the Pap1 bZIP motif and its cognate DNA oligomer. ***a***, Sequence of the bZIP domain of Pap1, numbered according to the full length sequence. Residues that contact bases in the complex are in red. Residues at *a* and *d* positions that form the hydrophobic core in the coiled coil region are in blue. ***b***, Sequence and numbering of the self-complementary 13-base pair palindromic DNA in the crystal. The DNA two-fold axis is located at the C-1/G1 base pair step. Bases that contact the protein are underlined. ***c***, Overall structure of the Pap1 bZIP dimer bound to DNA shown from the side (left) and end on (right). ***d***, Internal interactions of polar residues at *a* and *d* positions in the leucine zipper region. ***e***, Stereo view of the solvent flattened MIR electron density map contoured at 1σ around Gln 90 superimposed on the refined model.

Pap1 coiled coil reveals a few unusual residues at the *a* (His 115, Asn 122) and *d* (Thr 118) positions of the heptad repeat (not shown). These hydrophilic residues are buried with low accessibility (<30%). A buried hydrogen bond between Asn 122 side chains from the coiled coil chains is present as observed for the Asn of GCN4 at the *a* position. Interestingly, a common bridging water molecule mediates interactions between His 115 and Thr 118 (Fig. 2*d*). These residues do not influence the bend because the torsion angles are in the normal range for an α-helix. The hinge point appears to be near the base of the coiled coil, where the two chains come into contact after emerging from the major groove, as is the case in the Fos-Jun–DNA structure. Unlike GCN4 and Fos-Jun, the N-terminal region (residues 75–79) of the Pap1 basic region unwinds its α-helical structure and wraps around the DNA. This unwinding is due to Pro 78, which interacts with the DNA backbone by packing its nonpolar side chain against the deoxyribose ring of the adenosine that is base paired with the second thymidine (TTAC). Similar van der Waals contacts involving prolines have been observed in the structures of the paired domain–DNA[14] and adenosine deaminase type 1 (ADAR1)–Z-DNA[15] complexes.

DNA recognition

The Pap1 dimer contacts the DNA backbone over a 12-base pair DNA stretch (six base pairs for each half site). These interactions resemble those in the GCN4–DNA complex[10,11]. Five bZIP residues interact with six DNA phosphate groups in each half site. In the structure of the Pap1–DNA complex, the DNA exhibits symmetrical deformation at its center, and is slightly bent toward the protein, as is observed in the GCN4–DNA complex. The mean axial rise per turn of DNA helix (3.31 Å) resembles that of B-DNA (3.4 Å), although the average helical twist of 32.5° suggests that some of its features are reminiscent of A-DNA (33°).

At the major groove of each half site (TTAC), six bZIP residues participate in the recognition of the DNA bases (Fig. 3*a*,*b*). In addition, the protein–DNA interface contains six water molecules that mediate intermolecular hydrogen bonding interactions. Two of the water molecules (W2 and W3) occupy the same position as in the high resolution (2.2 Å) structure of the GCN4–DNA complex[16], although several other protein–DNA interactions are different. The large hydrophobic side chain of Phe 93 in the signature sequence of Pap1 (NXXAQXXFR contacts two methyl groups, those of the first and second thymine (TTAC) of the major groove (Figs 3*c* (middle), 4*a*). The conserved Ala 89 in the signature sequence NXXAQXXFR of the Pap1 bZIP motif also contributes to the recognition of the first thymine, as observed in the GCN4–DNA and Fos-Jun–DNA complexes. Another residue characteristic of this motif is Gln 90 (NXXAQXXFR), which forms two hydrogen bonds that bridge the second thymine and the third adenine (TTAC) of the major groove. The specificity for this adenine is also due to nonpolar contacts between the Gln 90 side chain (Cβ and Cγ carbon atoms) and the thymine pair. Interestingly, Gln 90 is hydrogen bonded to a buried bridging water molecule (W1) between the first thymine (TTAC) and the adenine that forms a base pair with the second thymine (TTAC). This water molecule is also stabilized forming a hydrogen bond to the main chain carbonyl group of Asn 86.

Surprisingly, the invariant residue, Asn 86 (NXXAQXXFR), confers distinct base specificity in comparison with the corresponding Asn of GCN4 and Fos-Jun. In the Pap1–DNA complex, Asn 86 forms one direct and one water (W4) mediated hydrogen

© 2000 Nature America Inc. • http://structbio.nature.com

© 2000 Nature America Inc. • http://structbio.nature.com

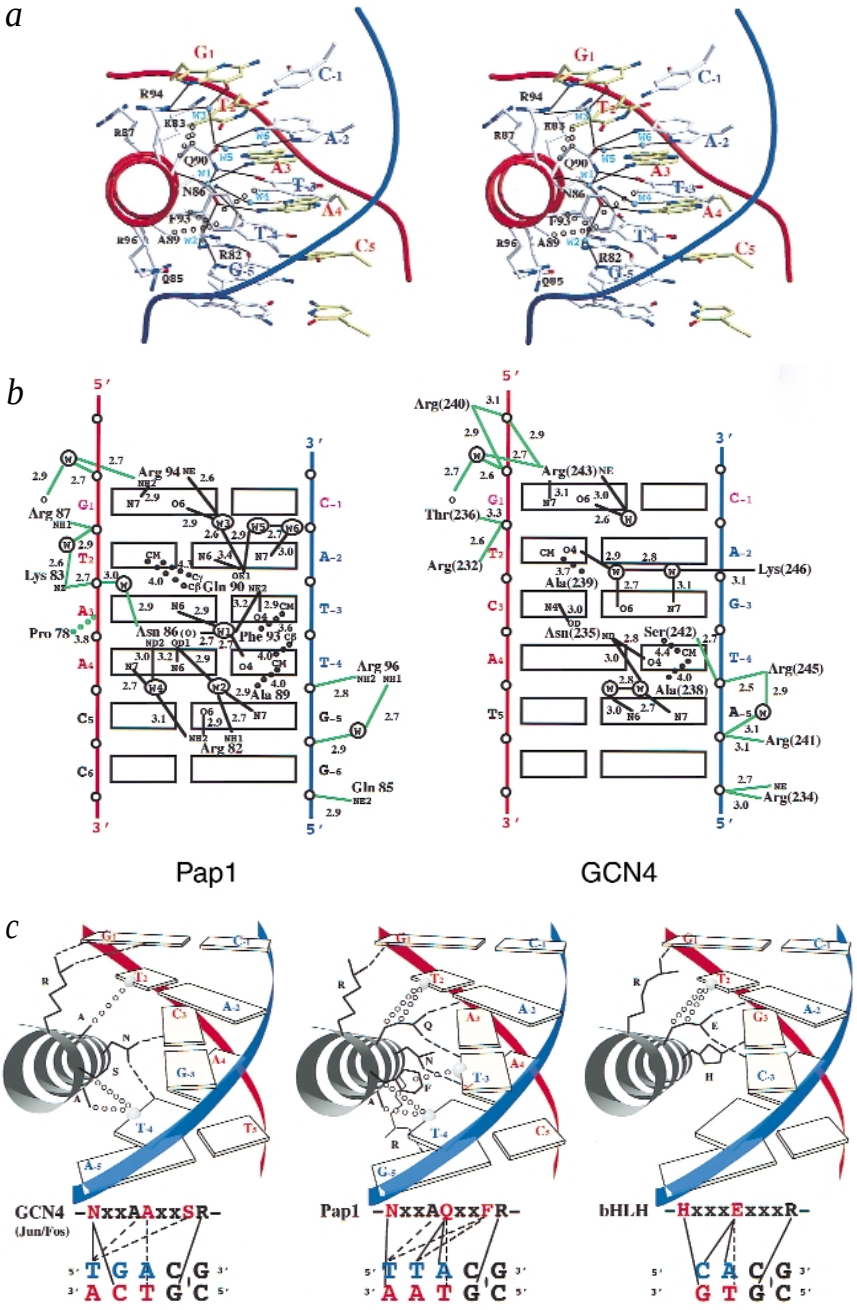

Fig. 3 DNA recognition of the Pap1 bZIP motif. ***a***, Stereo view of the basic region of the Pap1 bZIP bound to the major groove of its DNA half site. Hydrogen bonds and van der Waals contacts between protein and DNA are indicated by solid and dotted lines, respectively. ***b***, Diagrams showing protein–DNA interactions in the Pap1–DNA (left) and the GCN4–DNA (right) complexes. Each half site of the DNA is shown. Hydrogen bonds and van der Waals contacts between protein and DNA bases are indicated by solid and dotted black lines, respectively. Interactions with the DNA phosphate backbones are indicated with green lines. ***c***, Schematic representations of DNA recognition by the GCN4 bZIP (left), Pap1 bZIP (middle), and PHO4 bHLH (right) motifs. The conserved amino acid side chains of the motifs that make direct contacts with the DNA bases of the core sequences are shown. Hydrogen bonds are indicated with broken lines and van der Waals contacts with dotted lines. For clarity, water mediated hydrogen bonds are omitted.

bonds with the adenine that is base paired with the first thymine. Moreover, it forms one water mediated hydrogen bond with the guanine flanking the TTAC sequence (GTTAC) , which also interacts with Arg 82 by direct and water mediated hydrogen bonds. In contrast, the Asn in the GCN4–DNA and Fos-Jun–DNA complexes confers base specificity for the first and second bases of the TGAC sequence. This difference in base recognition is induced by a flip of the Asn 86 side chain to accommodate the bridging water molecule W1, which is stabilized by the side chain of Gln 90 (Fig. 4*b*,*c*).

Another invariant residue, Arg 94 (NXXAQXXFR), makes two hydrogen bonds with the guanine that forms a base pair with the fourth cytosine in the TTAC sequence, in a manner similar to the Arg in GCN4 and Fos-Jun (Fig. 3*b*,*c*). One of the two hydrogen bonds is mediated by a water molecule (W3) that is also hydrogen bonded to the side chain of Gln 90. In the GCN4–DNA complex, Lys 246 makes a hydrogen bond to one of the water molecules that hydrogen bonds with the second guanine and the thymine that forms a base pair with the third adenine of the TGAC sequence. This interaction is missing in the Pap1–DNA complex. Instead, Glu 90 is involved in a hydrogen bond network containing two water molecules (W5 and W6), which are linked to the third adenine of the TTAC sequence).

Similarity with HLH motifs

Interestingly, the stereochemical mode of DNA recognition by the Pap1 bZIP motif has isostructural features comparable with those of the basic helix-loop-helix (bHLH) motif and its derivative, the bHLH/ZIP motif, which recognize a common DNA sequence (CANNTG; known as the E-box)[17] in which the central bases are mostly CG or GC. These motifs in MAX[18,19], upstream stimulatory factor (USF)[20], and PHO4 (ref. 21) recognize the E-box through direct interactions with three conserved residues in the core signature sequence (HXXXEXXXR) of the basic region. Like the bZIP motifs, the conserved Arg of the bHLH signature sequence contacts the central guanine that is base paired with the third cytosine of the E-box (CACGTG; Fig. 3*c* (right)). Moreover, the conserved Glu of bHLH proteins recognizes the first cytosine and the second adenine of the sequence CACGTG with its side chain methylene groups, making nonpolar contacts with the methyl group of the thymine that is base paired with the second adenine in a manner similar to that of the conserved Gln in Pap1. Despite this similarity, the different chemical features of the side chain terminal groups of Glu and Gln provide DNA sequence specificity. The bHLH Glu and the bZIP Gln thus provide an elegant mechanism of conferring alternative DNA binding specificities.

DNA specificities of the bZIP subfamilies

The essential roles of the conserved Gln and Phe of the Pap1 sig-

© 2000 Nature America Inc. • http://structbio.nature.com

© 2000 Nature America Inc. • http://structbio.nature.com

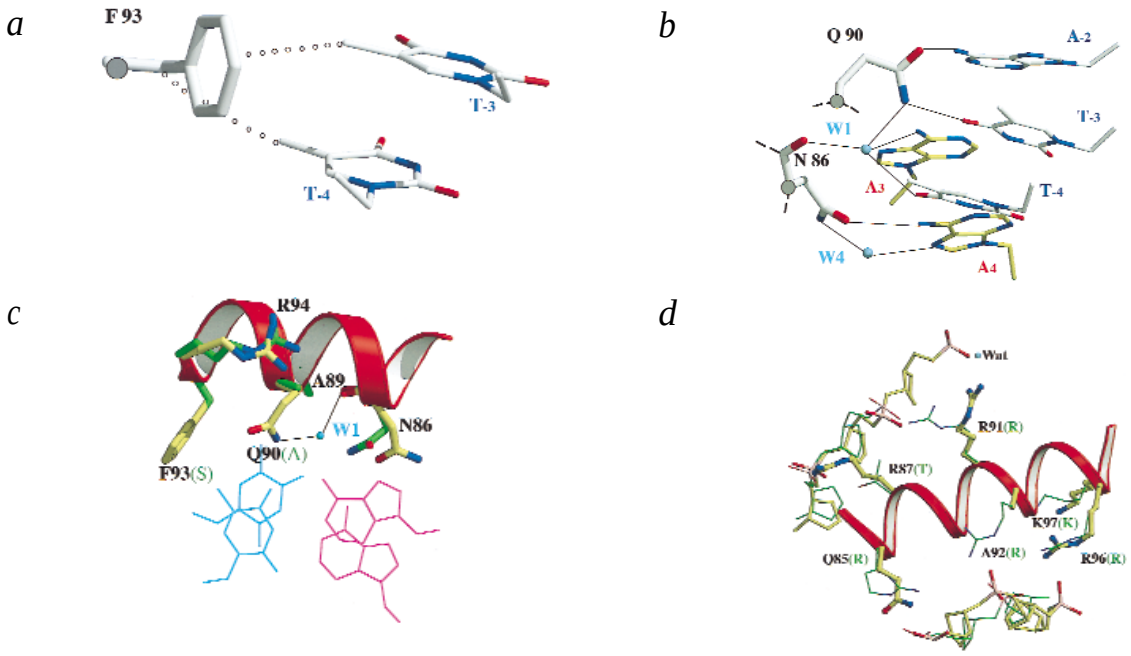

Fig. 4 Details of PAP1–DNA interactions. ***a***, Close up view of the hydrophobic interactions (dotted lines) between Phe 93 and the two thymine methyl groups of the TT base pair step. ***b***, Comparison of the side chain conformational changes in the Pap1 and GCN4 bZIP basic regions. The Pap1 side chains (yellow) are superimposed on the GCN4 side chains (green). Hydrogen bonds are indicated by solid lines, with the buried water molecule bridging Gln 90 and the main chain of Asn 86. ***c***, Close up view of the Pap1 bZIP basic region, with the buried water molecule bridging Gln 90 and the main chain of Asn 86. The water molecule also forms hydrogen bonds with the DNA bases. ***d***, Comparison of the side chains of conserved residues of Pap1 (yellow) with the corresponding residues of GCN4 (green).

© 2000 Nature America Inc. • http://structbio.nature.com

nature sequence are supported by mutagenesis experiments that have shown increased binding of GCN4 to TTAGTAA when the corresponding bZIP residues (Ala and Cys/Ser) are substituted with Gln and Phe, respectively[8]. In addition, a V285Q mutant in CCAAT/enhancer binding protein (C/EBP) (Gln 90 in Pap1) binds to TTACGTAA exclusively[22]. In contrast to CREB1 (identical to CREB), which binds the AT-rich bZIP-binding sequence poorly, members of the CREB2 subfamily are characterized by the signature sequence NXXAAXXYR, in which the Tyr residue, at the position corresponding to Phe 93 of Pap1, may make nonpolar contacts with the TT bases similar to those in the PAP1–DNA complex. In fact, some members of the Pap1 subfamily have Tyr instead of Phe at this position. The C/EBP subfamily members possess a characteristic Val residue in the bZIP signature (NXXAVXXSR) and bind preferably to the sequence TTGCGCAA[23]. This Val residue has been shown to be an important residue for the specificity of C/EBP[24]. The structure of the Pap1–DNA complex suggests that the contact between Val at this position and the methyl group of the thymine forming a base pair with the third adenine in the Pap1 binding sequence (TTAC) is unfavorably short. PAR and C/EBP subfamilies have a conserved Ser residue in the signature sequence NXXAAXXSR. The binding mode of this Ser would be similar to that in Fos–Jun and GCN4. Both Fos and Jun have a Cys at the corresponding position, but this Cys residue was replaced with a Ser in the Fos and Jun peptides used for the X-ray analysis to prevent formation of an artificial disulfide bond. GCN4 has a Ser residue at this position. The Ser residue in GCN4 and Fos–Jun makes a contact with the methyl group of the first thymine in the sequence TGA, which is common to the recognition sequence of PAR (TTA) and C/EBP (TTG).

It is interesting to consider the roles of the conserved residues in members of the Pap1 subfamily (Gln 85, Arg 87, Arg/Lys 91, Ala 92, Arg 96 and Lys 97; Fig. 4*d*). Three of them (Gln 85, Arg 87 and Arg 96) make contacts with phosphate backbones and Arg 91 forms a water mediated hydrogen bond with phosphate. But the others (Ala 92 and Lys 97) are not involved in any interactions with the DNA. The residues in GCN4 corresponding to Gln 85, Arg 87, Arg 91 and Arg 96 are Arg, Thr, Arg, and Arg, respectively, all of which make contacts with the phosphate backbone. In Pap1, Lys 97 is pushed away by Phe 93 and cannot contact the phosphate backbone, while the corresponding residue in Fos–Jun can because the residue in Fos–Jun corresponding to Phe 93 is not bulky. The Pap1-specific residue Ala 92 may play a structural role by eliminating a short contact with phosphate where the width of the major groove is narrow.

Binding flexibility of the Pap1 bZIP motif

Although Pap1 binds other bZIP binding sequences such as the C/EBP site with a six-fold reduction in affinity compared to its cognate DNA, it retains two-thirds of its affinity when binding to the CREB site. This dual binding specificity could be explained by a possible rearrangement of the side chains of Pap1 residues at the protein–DNA interface to allow interaction with the CREB site bases in a way similar to that observed in the GCN4–DNA complex. Binding flexibility has been observed in other bZIP proteins that lack the characteristic Gln of the Pap1 subfamily, such as cAMP response element binding protein 2 (CRE-BP2 or CREB2) and activating transcription factor 2 (ATF2)[25], which binds to TTACGTAA and the DNA binding site of CREB with a similar affinity.

Conclusions

The Pap1–DNA structure reveals a mechanism for recognizing diverse DNA sequences by the bZIP motifs. The differences between DNA base recognition by Pap1, GCN4 and Fos–Jun are

© 2000 Nature America Inc. • http://structbio.nature.com

Table 1 Data collection and structure refinement statistics

Diffraction data	Native	IdU4[1]	IdU5[1]	IdC7[2]	IdU9[1]	IdC12[2]
Resolution (Å)	2.0	2.5	2.5	2.5	2.5	2.5
R_{sym} (%)[3]	5.5	8.4	10.6	9.0	8.4	10.0
Phasing power[4]		2.25	2.16	1.44	1.86	1.43
Overall FOM	0.64					

Refinement statistics			
Protein / DNA atoms	3,471	R.m.s. deviations	
Solvent molecules	831	Bonds (Å)	0.006
Resolution (Å)	20–2.0	Angles (°)	1.00
R_{cryst} / R_{free} (%)[5]	23.0 / 25.3	Dihedrals (°)	18.9
Mean B-factor (Å^2)	42.8	Impropers (°)	1.28

[1]IdU4, IdU5 and IdU9 are derivatives in which T-4, T-3 and T2, respectively, have been replaced by 5′-iodo-dU.
[2]IdU7 and IdU12 are derivatives in which C-1 and C5, respectively, have been replaced by 5′-iodo-dC.
[3]$R_{sym} = \Sigma|I - <I>| / \Sigma I$; $R_{deri} = \Sigma||F_{PH}| - |F_P|| / \Sigma|F_P|$.
[4]Phasing power = root mean square heavy atom structure factor / residual lack of closure; $R_{cullis} = \Sigma||F_{PH}| - |F_P| - |F_{calc}|| / \Sigma|F_{PH}| - |F_P|$.
[5]R_{cryst} and $R_{free} = \Sigma||F_o| - |F_c|| / \Sigma|F_o|$, where the free reflections (10% of the total used) were held aside for R_{free} throughout refinement.

primarily due to two characteristic residues in Pap1, Phe 92 and Gln 90, but also to the conserved Asn 86 that exhibits a conformational change induced by Gln 90 and its linked water molecule (W1). The structure also provides valuable information for understanding the common stereochemical features in the recognition of DNA by two related DNA binding motifs — the bZIP and bHLH motifs.

Methods

Crystallization and data collection. Crystals of the Pap1 bZIP motif (residues 71–140) in complex with the DNA oligomer were obtained as described[26]. The crystals belong to space group R3, with a = b = 240.91 Å, and c = 43.87 Å. Heavy atom derivative crystals were prepared in the same manner using DNA in which 5-iodo U or 5-iodo C had been substituted for T or C in the positions indicated in Fig. 1*b*. Intensity data were collected from flash frozen crystals with a Rigaku R-AXIS IV imaging plate detector mounted on a Rigaku FR-C generator. Data processing and reduction were carried out using DENZO/SCALEPACK[27]. The 138,216 reflections were merged to give 56,393 unique reflections up to 2.0 Å with a completeness of 87.9% (73.1% for the outermost shell) and an R_{merge} of 5.5% (Table 1).

Phase determination and structure refinement. Initial MIR phases calculated with MLPHARE[28] gave a mean figure of merit of 0.64 to 3.0 Å. The phases were improved with solvent flattening and histogram matching with DM[29] (Fig. 2*e*). The MIR map revealed that, in the asymmetric unit, the crystal contains two Pap1–DNA complexes bound to DNA. A model was built into the MIR electron density map with O[30] and refined by simulated annealing with CNS. Refinement resulted in a final crystallographic R-factor of 23.0% and R_{free} of 25.3% for data between 20.0 and 2.0 Å. The N-terminal four residues of each Pap1 bZIP chain are not included in the current model due to their disorder. Similarly, the last C-terminal residue of both chains is also missing. In the course of the structural refinement, we found residual electron densities representing part of the coiled coil architecture in the large crystal channel. The densities were interpreted as a disordered Pap1 dimer that does not bind DNA. However, the electron densities were so diffuse that only parts of the main chains were traced. These disordered chains have higher B-factors; the averaged value was 75.7 Å^2. Using the GCN4 structure (PDB code 2DGC) determined at higher resolution, we tried to solve the structure by molecular replacement (MR) but were unsuccessful. After refinement, we found that Pap1 is bent toward the axis of the DNA helix, while GCN4 is almost perpendicular to it. This difference, in addition to high symmetry, made it difficult to solve the structure by molecular replacement.

Coordinates. Coordinates will be deposited in the Protein Data Bank (accession code1GD2).

Acknowledgments
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.). We thank Y. Kyogoku for support at the early stage of the project.

Correspondence should be addressed to T.H. *email: hakosima@bs.aist-nara.ac.jp*

Received 26 April, 2000; accepted 29 August, 2000.

1. Pabo, C.O. *et al. Science* **247**,1210–1213 (1990).
2. Wolberger, C., Vershon, A.K., Liu, B., Johnson, A.D. & Pabo, C.O. *Cell* **67**, 517–528 (1991).
3. Pomerantz, J.L. & Sharp, P.A. *Biochemistry* **33**, 10851–10858 (1994).
4. Choo, Y. & Klug, A. *Curr. Opin. Struct. Biol.* **7**, 117–125 (1997).
5. Toda, T. *et al. Mol. Cell. Biol.* **12**, 5474–5484 (1992).
6. Shimanuki, M., Saka, Y., Yanagida, M. & Toda, T. *J. Cell Sci.* **108**, 569–579 (1995).
7. Toone, W.M. *et al. Genes Dev.* **12**, 1453–1463 (1998).
8. Fernandes, L., Rodrigues-Pousada, C. & Struhl, K. *Mol. Cell. Biol.* **17**, 6982–6993 (1997).
9. Benbrook, D.M. & Jones, N.C. *Nucleic Acids Res.*. **22**, 1463–1469 (1994).
10. Ellenberger, T.E., Brandl, C.J., Struhl, K. & Harrison, S.C. *Cell* **71**, 1223–1237 (1992).
11. Koning, P. & Richmond, T.J. *J. Mol. Biol.* **233**, 139–154 (1993).
12. Glover, J.N.M. & Harrison, S.C. *Nature* **373**, 257–261 (1995).
13. Chen, L., Glover, J.N.M., Hogan, P.G., Rao, A. & Harrison, S.C. *Nature* **392**, 42–48 (1998).
14. Xu, W., Rould, M.A., Jun, S., Desplan, C. & Pabo, C.O. *Cell* **80**, 639–650 (1995).
15. Schwartz, T., Rould, M.A., Lowenhaupt, K., Herbert, A. & Rich, A. *Science* **284**, 1841–1845 (1999).
16. Keller, W., Konig, P. & Richmond, T.J. *J. Mol. Biol.* **254**, 657–667 (1995).
17. Baxevanis, A.D. & Vinson, C.R. *Curr. Opin. Gen. Dev.* **3**, 278–285 (1993).
18. Ferre-D'Amare, A.D., Prendergast, G.C., Ziff, E.B. & Burley, S.K. *Nature* **363**, 38–44 (1993).
19. Brownlie, P., *et al. Structure* **5**, 509–520 (1997).
20. Ferre-D'Amare, A.D., Pognonec, P., Roeder, R.G. & Burley, S.K. *EMBO J.* **13**, 180–189 (1994).
21. Shimizu, T. *et al. EMBO. J.* **16**, 4689–4697 (1997).
22. Koldin, B., Suckow, M., Seydel, A., Wilcken-Bergmann, B. & Muller-Hill, B. *Nucleic Acids Res.* **23**, 4162–4169 (1995)
23. Vinson, C.R., Sigler, P.B. & McKnight, S.L. *Science* **246**, 911–916 (1989).
24. Johnson, P.F. (1993). *Mol. Cell. Biol.* **13**, 6919–6930 (1993).
25. Benbrook, D.M. & Jones, N.C. *Nucleic Acids Res.* **22**, 1463–1469 (1994).
26. Fujii, Y. *et al. Acta Crystallogr. D* **54**, 1014–1016 (1998).
27. Otwinowski, Z. & Monor, W. *Methods Enzymol.* **276**, 307–326 (1997).
28. Collaborative Computational Project, Number 4. *Acta Crystallogr. D* **50**, 760–763 (1994).
29. Cowtan, K. & Main, P. *Acta Crystallogr. D* **52**, 43–48 (1996).
30. Jones, T.A., Zou. J.-Y., Cowan, S.W. & Kjeldgaad, M. *Acta Crystallogr. A* **47**, 110–119 (1991).

© 2000 Nature America Inc. • http://structbio.nature.com