The EMBO Journal vol.13 no.1 pp.180-189, 1994

Structure and function of the b/HLH/Z domain of USF

Adrian R.Ferré-D'Amaré[1], Philippe Pognonec[2,3], Robert G.Roeder[2] and Stephen K.Burley[1,4,5]

[1]Laboratories of Molecular Biophysics, [2]Laboratory of Biochemistry and Molecular Biology, [4]Howard Hughes Medical Institute, The Rockefeller University, 1230 York Avenue, New York, NY 10021, USA

[3]Present address: Centre de Biochimie du Centre National de la Recherche Scientifique, Parc Valrose, 06108 Nice cedex 2, France

[5]Corresponding author

Communicated by I.W.Mattaj

The basic/helix-loop-helix/leucine zipper (b/HLH/Z) transcription factor upstream stimulatory factor (USF) and its isolated DNA binding domain undergo a random coil to α-helix folding transition on recognizing their cognate DNA. The USF b/HLH cocrystal structure resembles the structure of the b/HLH/Z domain of the homologous protein Max and reveals (i) that the truncated, b/HLH DNA binding domain homodimerizes, forming a parallel, left-handed four-helix bundle, and (ii) that the basic region becomes α-helical on binding to the major groove of the DNA sequence CACGTG. Hydrodynamic measurements show that the b/HLH/Z DNA binding domain of USF exists as a bivalent homotetramer. This tetramer forms at the USF physiological intranuclear concentration, and depends on the integrity of the leucine zipper motif. The ability to bind simultaneously to two independent sites suggests a role in DNA looping for the b/HLH/Z and Myc-related families of eukaryotic transcription factors.

Key words: DNA-protein interaction/helix-loop-helix/induced fit/leucine zipper/tetramerization

Introduction

The b/HLH/Z family of eukaryotic transcription factors is characterized by a highly conserved DNA binding domain composed of a basic (b) region, followed successively by helix-loop-helix (HLH) and leucine zipper (Z) motifs. The contiguous presence of the HLH and the leucine zipper, two putative dimerization interfaces, distinguishes these proteins from both the basic/helix-loop-helix (bHLH) and the basic/leucine-zipper (b/Z) transcription factors. b/HLH/Z family members occur widely, ranging from the mammalian nuclear proteins Myc (Murre *et al.*, 1989), Mad (Ayer *et al.*, 1993), Max (Blackwood and Eisenman, 1991; Prendergast *et al.*, 1991), Mxi1 (Zervos *et al.*, 1993), USF (Gregor *et al.*, 1990), TFEB (Carr and Sharp, 1990), TFE3 (Beckman *et al.*, 1990) and AP-4 (Hu *et al.*, 1990), to the yeast protein CBF-1 (Cai and Davis, 1990), among many others. All of these proteins bind to a common CANNTG element known as the E-box [see Bexevanis and Vinson (1993) for a recent review and sequence alignment]. Structure determinations of the DNA binding domains of GCN4, a b/Z protein (Ellenberger *et al.*, 1992), and Max, a b/HLH/Z protein (Ferré-D'Amaré *et al.*, 1993), both complexed with their respective cognate DNAs, have demonstrated that the basic regions and the leucine repeats of the two proteins are structurally and functionally analogous. The HLH moiety of Max, however, has no counterpart in GCN4, and dimerizes to form a parallel, left-handed four-helix bundle, which is globular and contains a well-defined hydrophobic core.

Upstream stimulatory factor (USF) was first characterized as an activity from HeLa cell nuclei which binds to an upstream element of the adenovirus major late promoter (Carthew *et al.*, 1985; Miyamoto *et al.*, 1985; Sawadogo and Roeder, 1985), and stimulates transcription, possibly by direct interaction with the basal factor TFIID (Sawadogo and Roeder, 1985; Sawadogo, 1988; Meisterernst *et al.*, 1990; Workman *et al.*, 1990; Bungert *et al.*, 1992). Extensive purification of the HeLa nuclear activity yielded two polypeptides (Sawadogo *et al.*, 1988), the smaller of which has been cloned and sequenced revealing a protein of molecular weight 34 kDa with a highly conserved b/HLH/Z DNA binding domain near its C-terminus (Gregor *et al.*, 1990). The recombinant protein expressed from this cDNA clone (which shall be henceforth referred to as USF) homo-oligomerized efficiently, bound DNA containing a CACGTG E-box motif with nanomolar affinity and activated transcription in a manner indistinguishable from that of material purified from HeLa nuclear extracts (Pognonec and Roeder, 1991).

We have investigated the function of the b/HLH/Z DNA binding domain of USF, using highly purified recombinant proteins corresponding to full size USF, the isolated b/HLH/Z domain and a truncated b/HLH protein. We demonstrate that the HLH motif is primarily responsible for mediating the protein-protein interactions stabilizing the minimal, homodimeric b/HLH DNA binding domain. In addition, we show by circular dichroism (CD) spectroscopy and X-ray crystallography that the two basic regions of this homodimer undergo random coil to α-helix folding transitions on binding to the specific DNA recognition sequence CACGTG. Although we confirm that the stoichiometry of the USF-DNA interaction is 2:1 (protein monomer:duplex DNA), we document that the presence of the leucine zipper motif allows the transcription factor to function as a bivalent homotetramer that can simultaneously bind two spatially distinct DNA recognition sites.

Results and discussion

USF, b/HLH/Z and b/HLH undergo a DNA-induced folding transition

Full size USF, its intact DNA binding domain (b/HLH/Z) and a construct of the DNA binding domain missing the entire heptad repeat or leucine zipper element (b/HLH) were

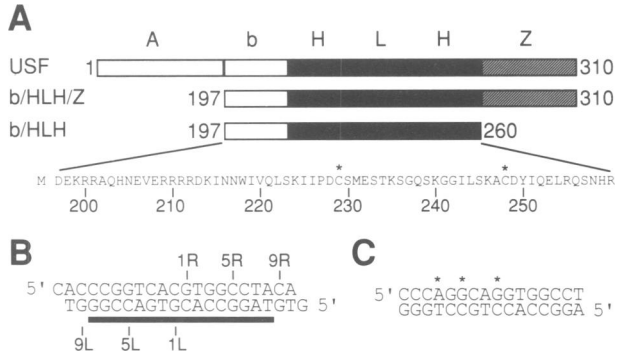

Fig. 1. Macromolecular reagents employed. (**A**) Cartoon representation of the three protein constructs used in this study. Regions of the proteins are indicated: A, activation; b, basic; HLH, helix−loop−helix; Z, leucine zipper. Numbering is according to the full size recombinant USF and the naming of the regions within the DNA-binding domain follows the alignment of Ferré-D'Amaré *et al.* (1993). The sequence for the b/HLH construct, including its initiation methionine, is shown in the inset. For this particular construct, the two cysteines marked with asterisks were mutated to serines. (**B**) Sequence of the DNA employed for crystallization, derived from the adenovirus major late promoter (Ziff and Evans, 1978). Numbering is to the left and right of the molecular pseudo-dyad. Solid bar indicates composition of the blunt 16mer duplex employed for EMSA, CD and PCS measurements, referred to throughout as sDNA. (**C**) Sequence of nsDNA. Asterisks indicate positions where this molecule differs from sDNA.

overexpressed and purified to homogeneity (Figure 1A). Recombinant USF binds specific DNA with an apparent dissociation constant of 1.3×10^{-9} M (Pognonec and Roeder, 1991). Electrophoretic mobility shift assays (EMSA) demonstrate that both USF and b/HLH/Z efficiently bind 16 bp oligonucleotides (Figure 2A), corresponding to the wild-type USF DNaseI footprint (Sawadogo and Roeder, 1985). In both cases, binding requires a protein to DNA duplex stoichiometry of 2:1, the apparent dissociation constants of b/HLH/Z and USF are comparable, and within experimental precision the purified, recombinant proteins have unit activity (Figure 2B, C and D).

USF binds specifically to a dyad-symmetric CACGTG E-box element (Prendergast and Ziff, 1991). The 2:1 stoichiometry suggests that the dyad axis of the E-box coincides with a dyad in the protein oligomer. Moreover, neither USF nor b/HLH/Z form an electrophoretically stable complex with a mutant DNA (non-specific DNA; nsDNA) that incorporates mutations in only one half of the symmetric element, implying a highly cooperative protein dimer−duplex DNA interaction (Figure 2A). In contrast, b/HLH does not form an electrophoretically stable complex with specific DNA under conditions used for EMSA studies of both USF and b/HLH/Z (data not shown). Nevertheless, CD spectroscopy demonstrates that b/HLH binds DNA in a sequence-specific manner indistinguishable from full-length, recombinant USF (see below).

Interactions of the three proteins with sDNA and nsDNA were examined by CD spectroscopy. The spectra illustrated in Figure 3A demonstrate that USF contains 27% α-helix in the absence of DNA. Incubation of USF with sDNA (specific DNA), but not nsDNA, provokes a negative deflection of the signal at 208 and 222 nm which is consistent with an increase in α-helix content of 8%, or ~25 amino acids per monomer (Figure 3A). In addition, there is a

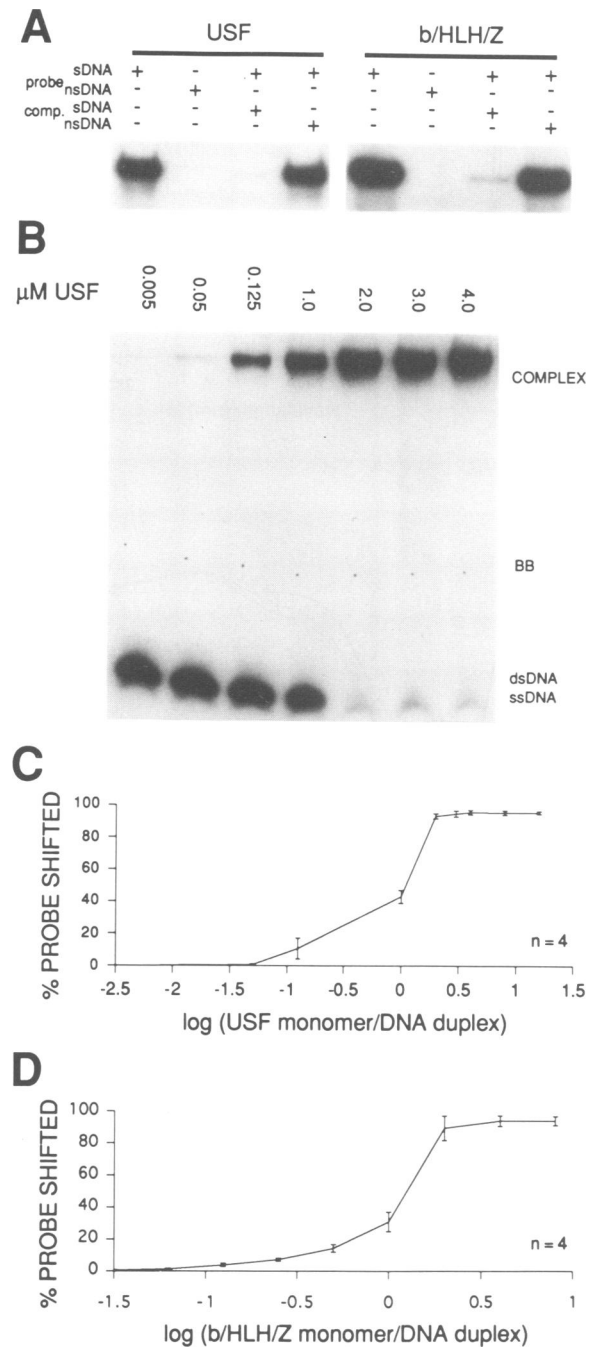

Fig. 2. Determination of protein−DNA stoichiometries. (**A**) Autoradiogram of a representative EMSA experiments with USF and b/HLH/Z. ^{32}P-labeled DNA duplex and protein monomer concentration was fixed at 1 μM in each binding reaction. Cold competitor was added in a 100-fold molar excess, where indicated. (**B**) Autoradiogram of a representative example of a titration with USF. Concentrations of USF monomer added to each binding reaction are indicated at the top of the figure. sDNA duplex concentration was 1 μM in each binding reaction. Migration positions of the protein−DNA complex, tracking dye (bromophenol blue, BB) double-stranded DNA (dsDNA) and single-stranded DNA (ssDNA) are indicated. (**C**) Average of quantitation of four independent experiments as shown in B. The ordinate is the percentage of the labeled DNA found in the shifted band in each lane; the abscissa is the Naperian logarithm of the molar ratio of protein monomer to duplex DNA. Bars represent standard errors of the mean. (**D**) Same as C but using b/HLH/Z instead of USF.

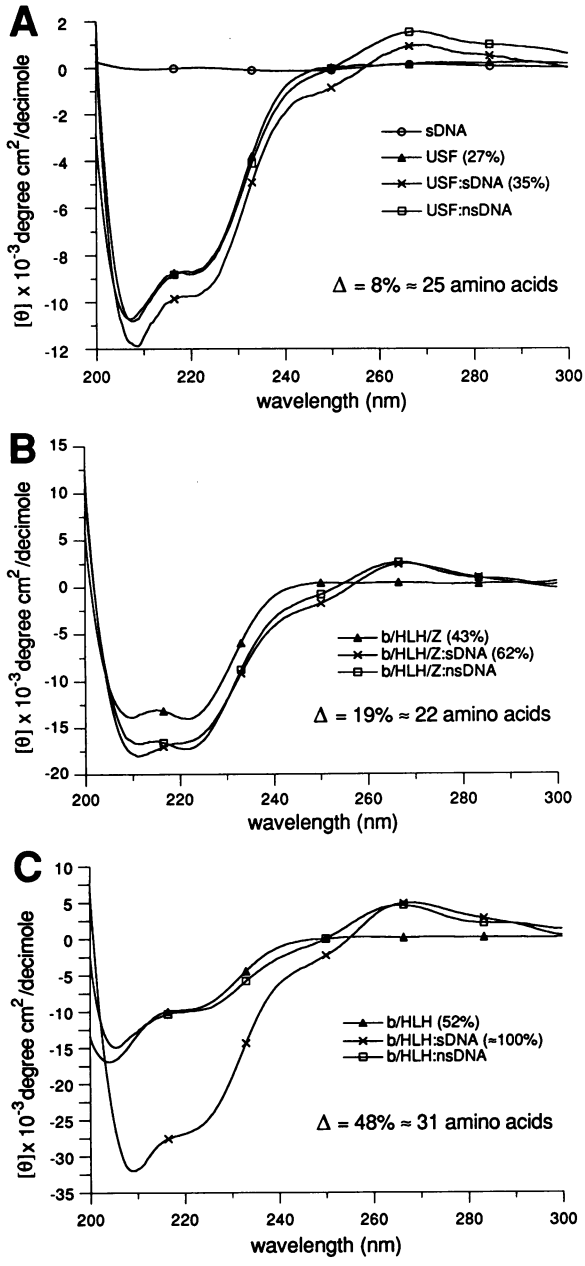

Fig. 3. Normalized CD spectra. (**A**) Spectra of free USF, free sDNA, a 2:1 molar mixture of USF and sDNA (USF:sDNA) and a 2:1 molar mixture of USF and nsDNA (USF:nsDNA). The concentration of USF was ~6 μM, on a monomer molecular weight basis, for each experiment. The estimated α-helical contents of USF and USF:sDNA are indicated in parentheses. (**B**) Spectra of free b/HLH/Z, a 2:1 molar mixture of b/HLH/Z and sDNA (b/HLH/Z:sDNA), and a 2:1 molar mixture of b/HLH/Z and nsDNA (b/HLH/Z:nsDNA). The concentration of b/HLH/Z was ~13 μM, on a monomer molecular weight basis, for each experiment. The estimated α-helical contents of b/HLH/Z and b/HLH/Z:sDNA are indicated in parentheses. (**C**) Spectra of free b/HLH, a 2:1 molar mixture of b/HLH and sDNA (b/HLH:sDNA), and a 2:1 molar mixture of b/HLH and nsDNA (b/HLH:nsDNA). The concentration of b/HLH was ~13 μM, on a monomer molecular weight basis, for each experiment. The estimated α-helical contents of b/HLH and b/HLH:sDNA are indicated in parentheses.

change in molar ellipticity in the vicinity of 264 nm, which may reflect a change in DNA structure on binding to protein (for reference, the CD spectrum of sDNA alone is illustrated in Figure 3A). b/HLH/Z undergoes a similar conformational change on DNA binding (Figure 3B). Addition of sDNA results in an increase in α-helix content from 43 to 62%, a difference of 19% or 22 amino acids. Surprisingly, given the specificity exhibited in the EMSA experiment (Figure 2A), this truncated form of USF appears to undergo a similar conformational change when mixed with nsDNA. In contrast, the minimal DNA binding unit (b/HLH) undergoes an increase in α-helix content of 48%, or 31 amino acids, upon mixing with sDNA, but exhibits no conformational change when incubated with nsDNA (Figure 3C). Furthermore, based on the dichroic signal at 264 nm, it appears to distort the DNA in a manner indistinguishable from b/HLH/Z. The illustrated b/HLH spectra were obtained at a complex concentration of 13 μM. No change in the spectra occurred on increasing the concentration of the complex 10-fold, implying no further increase in α-helix content (data not shown). The random coil to α-helix transition is fully saturated at a complex concentration of 13 μM. If specific DNA binding and the associated protein folding transition are one and the same phenomenon, our data imply that the affinity of b/HLH for sDNA is in the low micromolar range. Therefore, removal of the leucine repeat reduces the affinity for sDNA by ~1000-fold without altering the mode of DNA binding or its specificity.

In the absence of DNA, CD spectra were obtained at increasing b/HLH and b/HLH/Z concentrations (data not shown). Even in the millimolar range there were no changes in the spectra at far ultraviolet wavelengths, implying that the largely α-helical oligomerization interface common to these two truncated forms of USF is fully folded at 13 μM. Thus, failure of b/HLH to gel-shift under these conditions is probably not due to weak binding to DNA, but instead results from association/dissociation rates that are fast relative to the time scale of the EMSA experiment giving a DNA smear instead of a sharp band (data not shown).

The simplest interpretation of these results is that an otherwise disordered segment of the DNA binding domain of USF is induced to fold by sDNA, becoming α-helical on recognizing the sequence CACGTG. Two other CD spectroscopic studies of related polypeptides have been reported. Both a b/HLH peptide derived from the MyoD (Anthony-Cahill *et al.*, 1992) and a truncated form of TFEB (Fisher *et al.*, 1993) have been shown to undergo similar DNA-induced, coil-to-helix folding transitions. Here we demonstrate that this property is present in the full size protein, is both qualitatively and quantitatively conserved in the isolated b/HLH/Z domain and the b/HLH construct, and is thus not an artifact of truncation.

The HLH region of USF is a parallel, left-handed four-helix bundle

The structure of the complex of b/HLH bound to sDNA was examined in detail by X-ray crystallography at 2.9 Å resolution (Figures 4, 5 and 6). The $(b/HLH)_2$–DNA complex folds into a parallel, left-handed four-helix bundle, which is topologically identical to the structure of another b/HLH/Z transcription factor Max (see Figure 4D) (Ferré-D'Amaré *et al.*, 1993). Comparison of the three-dimensional structures illustrated in Figure 4A and D suggests that the basic and loop regions of the b/HLH motif are remarkably plastic. Several differences between the Max

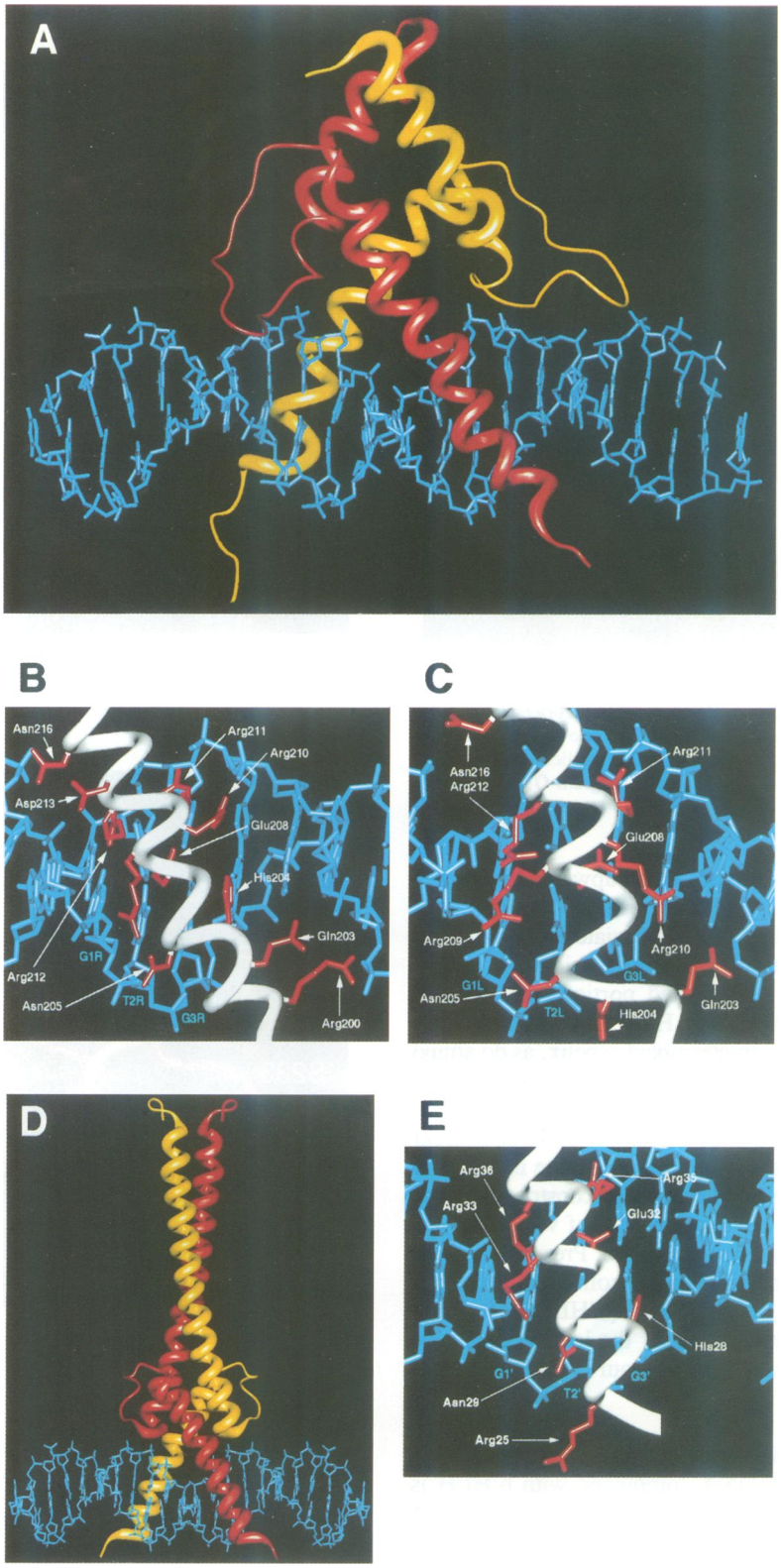

Fig. 4. Cocrystal structure of USF b/HLH. **(A)** Overall view of the complex. Helical regions are represented by thick tubes; non-regular secondary structure elements by thin tubes. Molecule 1 and molecule 2 are colored red and yellow, respectively. The N-termini are at the bottom of the figure. **(B)** View of the basic region of molecule 1 interacting with DNA. The path of the protein backbone is shown as a white tube; side chains which interact with DNA as red stick figures. **(C)** View of the basic region of molecule 2 interacting with DNA, as in B. Note that the first turn of the polypeptide backbone is not α-helical. **(D)** Overall view of the Max–DNA complex provided for comparison. The drawing convention is as in A. **(E)** View of the basic region of Max interacting with DNA provided for comparison. The drawing convention is as in B.

and USF structures represent distortions of the polypeptide chain that can be attributed to crystal packing interactions. The N-terminal 10 amino acids of molecule 2 of USF (yellow) have been pulled down from its cognate DNA by a lattice contact with a neighbouring DNA duplex. The top four residues of helix 1 of USF (H1) and the loop region

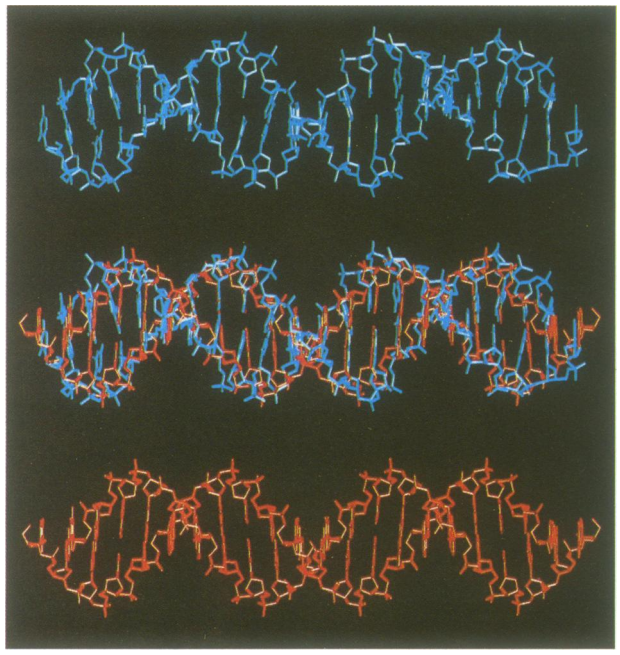

Fig. 5. Comparison of the DNA structure found in the b/HLH cocrystal (blue) to canonical B-form DNA (red). The orientation of the blue DNA is the same shown in Figure 4A. The two structures were overlaid using phosphorous coordinates with the 'explicit' and 'improve' options of the Lsq command in the program O (Jones *et al.*, 1991).

of molecule 1 (red) have been stretched upwards by another crystal packing interaction. Finally, the C-terminal four residues of both molecules of USF deviate from α-helical geometry because of lattice contacts with a pair of symmetry-related complexes. The undistorted portions of the USF complex show that the basic and H1 regions (amino acids 199 through 225) form an uninterrupted α-helix, as do amino acids 243 through 256 of H2. The hydrophobic core of the four-helix bundle closely resembles that of Max, and the last conserved hydrophobic position of H2, Leu254 (equivalent to Met74 in Max), packs against its dimer mate in standard coiled-coil fashion. The b/HLH expression construct encoded another six amino acids beyond Leu254, of which the last four appear as random coil in the structure. Presumably these residues do not adopt the left-handed coiled coil of right-handed α-helices seen in the Max $(b/HLH/Z)_2$–DNA structure because the first leucine of the leucine zipper or heptad repeat region is not present in our b/HLH construct of USF.

Binding of USF does not produce a net bend in DNA

The conformation of the DNA complexed with b/HLH is compared with that of canonical B-form DNA in Figure 5 [root-mean-square (r.m.s.) deviation between phosphorous atoms = 2.4 Å]. Although binding produced a subtle S-shaped distortion of the DNA, there is no systematic variation of the stereochemical parameters. The mean rise per base pair is 3.32 Å and the average helical twist is 32.9°, implying 10.9 bp/turn. The DNA stacks 5′ to 3′ in the crystal, producing a pseudo-continuous, B-form double helix which is stabilized by Watson–Crick base pairing of the overhanging C and G (Figure 1B). In contrast to results from phasing analyses employing the homologous proteins Myc and Max, which suggested bend angles ranging from 50 to

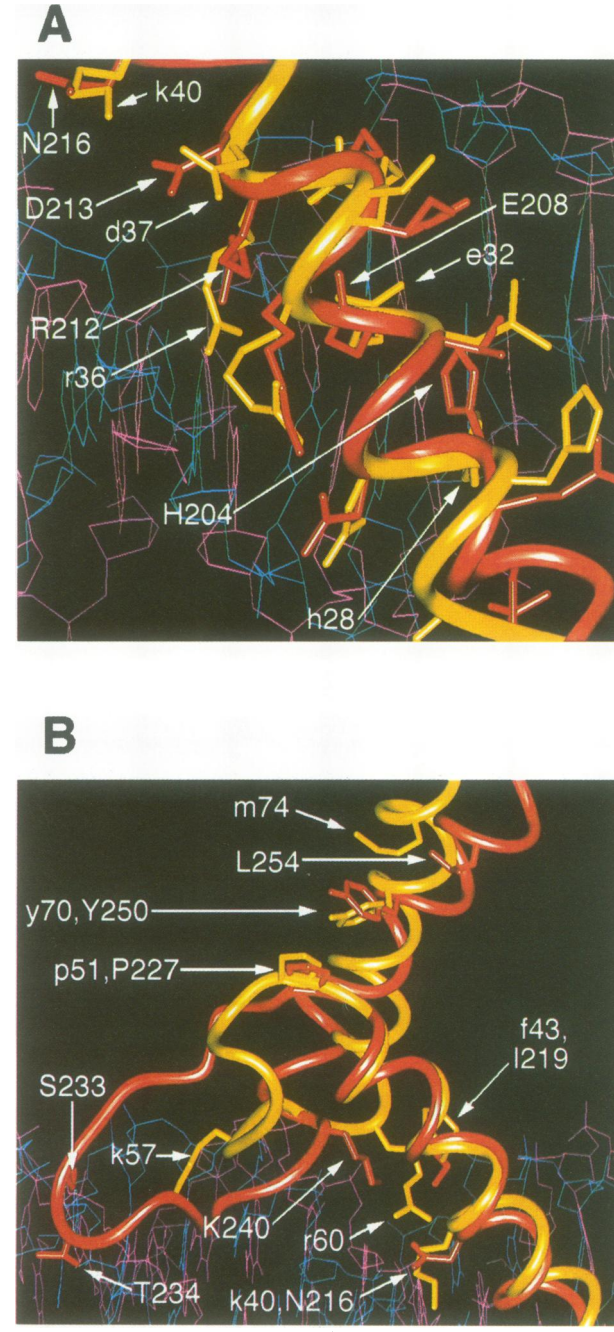

Fig. 6. Three-dimensional comparison of the USF b/HLH structure (protein in red, DNA in blue; present work) and the Max b/HLH/Z structure (protein in yellow, DNA in magenta; Ferré-D'Amaré *et al.*, 1993). Amino acid residues are identified in single letter code, employing upper and lower case letters for USF and for Max, respectively. At the current resolution limit of 2.9 Å, the precision of the atomic coordinates implied by a Luzzati plot (Luzzati, 1952) is between 0.3 and 0.4 Å (data not shown). (**A**) Alignment of the basic regions of molecule 1 of USF and Max. The superposition was achieved by fitting the α-carbon coordinates of USF residues 204 through 212 with the corresponding atoms of Max. The r.m.s. deviation of the nine pairs of atoms is 0.59 Å. (**B**) Alignment of the H1, loop and H2 regions of molecule 2 of USF with the corresponding regions of Max. The superposition was achieved by fitting the α-carbon coordinates of USF residues 216 through 225 with the corresponding atoms of Max. The r.m.s. deviation of the 10 pairs of atoms is 0.63 Å.

80° (Fisher *et al.*, 1992; Wechsler and Dang, 1992), there is virtually no net bend in the double-helical axis. A similar linear DNA conformation was observed in the Max–DNA cocrystal structure (Ferré-D'Amaré *et al.*, 1993).

The basic region becomes α-helical as it recognizes specific DNA

In the orthorhombic crystal form examined here, the two basic regions of the b/HLH dimer lie in different crystalline environments and the basic region of molecule 2 (yellow) appears to be distorted by lattice contacts (Figure 4A, B, C, D and E). Comparing interactions made by the two polypeptide chains with DNA, stronger, presumably specific, interactions can be distinguished from weaker, or secondary, interactions. Two critical side chain–base contacts are made by both basic regions. Arg212, which confers specificity for CACGTG (class B) versus CAGCTG (class A) E-boxes (Dang *et al.*, 1992), contacts N7 of the guanine adjacent to the palindrome's dyad axis through its η1 nitrogen. This guanine was mutated in one half-site in the nsDNA, a substitution which abolishes binding. Glu208, which is absolutely conserved among b/HLH and b/HLH/Z family members and is critical for DNA binding (Fisher and Goding, 1992), makes contact with the N4 of the outer C of the palindrome. The outer C:G bp in both half-sites was shown to be indispensable for native USF binding (Lennard and Egly, 1987). Both these side chain–base interactions as well as backbone contacts involving Asn205, Arg209 and Arg211 are seen in equivalent positions in the cocrystal structure of the complex of Max with CACGTG (Ferré-D'Amaré *et al.*, 1993) (see Figures 4E and 6A). In addition, Nε2 of His204 of molecule 1 is located 3.8 Å away from O6 of the outer G (G3). In the Max structure, the equivalent His28 is located a similar distance away from N7 of the same G.

Comparison of the undistorted basic region of molecule 1 of USF with the corresponding region of Max (Figure 6A) demonstrates a high degree of structural similarity, with most equivalent residues lying in approximately equal conformations. Nonetheless, the orientations of the helices relative to the major groove of DNA appear to differ considerably. This may result, for instance, in the difference in contacts made by the histidine residues, mentioned above. Because of the current modest resolution limits and because the Max DNA coordinates are based on the rotationally averaged electron density, detailed comparison of DNA conformation between the two structures must await further crystallographic investigation.

Conformational differences between the two b/HLH basic regions of USF underscore the relative plasticity of this portion of the motif. The Max-like basic region of molecule 1 participates in a dense network of side chain–DNA contacts involving Gln203, His204, Arg210, Asp213 and Asn216, in addition to those mentioned above. It is possible that the observed Asp213–phosphate contact (2.9 Å) is an artifact of the low pH (4.75) at which the crystals were grown (Materials and methods). However, given the highly negative electrostatic potential of the neighborhood due to the phosphate groups, the pKa of the carboxyl group may be increased by several units and the hydrogen-mediated phosphate–carboxylic acid interaction might very well occur at physiological pH (Sawyer and James, 1982). The distorted basic region of molecule 2 makes substantially fewer contacts with DNA. Arg210 and Asn216 make the same contacts as in molecule 1. Because of polypeptide backbone distortion, Asn205 contacts a phosphate one nucleotide removed from that contacted by both the undistorted molecule 1 and the corresponding residue in the Max–DNA complex structure. Finally, the side chains of Arg200 of molecule 1 and Gln203 of the unwound basic region of molecule 2 make contacts with bases outside the central, palindromic CACGTG element. Although we cannot exclude the formal possibility that conformational differences between the basic regions of molecules 1 and 2 are biologically relevant, the similarity of molecule 1 to Max and easy attribution of structural differences to crystal packing effects suggest otherwise.

As in the Max–DNA complex structure, both the loops and the four-helix bundle of USF interact with DNA. The USF loop region consists of 12 amino acids, four residues longer than in Max. As a result of this added length, the loop of molecule 2 traverses the adjacent minor groove and makes two phosphate contacts (Ser233 and Thr234; Figure 6B) and a contact with a sugar oxygen within the minor groove (Gln238 Nε to C4R O4′; 2.8 Å). In contrast, the eight-residue Max loop makes a sole lysine–phosphate contact (Lys57; Figure 6B). The loop of molecule 1 (red) adopts a different conformation than either the Max loop or the loop of molecule 2 of USF (Figure 4A and 4D). In this case, the loop is involved in a lattice contact. The backbone amide of the conserved basic residue at the beginning of H2, Lys240, packs against the phosphate backbone, as does the equivalent residue of Max (Arg60).

Three-dimensional alignment of the HLH regions of USF and Max (Figure 6B) shows that the crossing angle between H1 and H2 is substantially different in the two structures. As mentioned above, the hydrophobic cores of the four-helix bundles are similar and a number of interactions which presumably affix H1 to H2 are conserved. For instance, Ile219 from H1 packs against Lys240 at the beginning of H2 in the USF structure as does Pro227 at the end of H1 against Tyr250 on the outer face of H2, interactions analogous to the Phe43–Arg60 and Pro51–Tyr70 contacts seen in Max. The differing orientations of H2 can be attributed partly to the fact that in the USF sequence, Lys240 is followed by two successive glycines, providing conformational flexibility at the beginning of H2. In addition, the deletion of the leucine zipper from the USF sequence could also have contributed to the differing packing angle of the four-helix bundle. A comparison of the USF b/HLH construct structure with the structure of a b/HLH protein which lacks a heptad repeat in the natural sequence could be illuminating in this regard.

Three segments of the USF b/HLH polypeptide contact its cognate DNA: the basic region, the loop and the first residue of H2. Of these three only the basic region is α-helical. The increased helical content of USF, b/HLH/Z and b/HLH observed upon association with specific DNA by CD must result from the basic regions of these proteins undergoing a random coil to α-helix folding transition upon binding to CACGTG. Therefore, molecular recognition depends on DNA-induced folding of the basic region. Comparison of the two basic regions in the b/HLH structure suggests that 10 or 12 protein–DNA contacts per dimer are required for recognition of the 6 bp element. It is remarkable that only four (or six, including His204) of these occur between amino acid side chains and nucleic acid bases. The

Table I. Summary of PCS results

Sample	D_T ($10^{-13}m^2s^{-1}$)	R_H (nm)	Estimated mol. wt (kDa)	Oligomerization state	Calculated mol. wt (kDa)
b/HLH/Z	737.5 (21.3)	3.35 (0.12)	55.8 (4.16)	$(b/HLH/Z)_4$	54.0
b/HLH	1055.7 (29.3)	2.18 (0.07)	19.6 (1.59)	$(b/HLH)_2$	15.2
b/HLH/Z + sDNA	614.9 (11.0)	3.74 (0.05)	72.8 (1.54)	$(b/HLH/Z)_4-(DNA)_2$	73.4
b/HLH + sDNA	980.0 (15.4)	2.40 (0.04)	24.8 (1.04)	$(b/HLH)_2-DNA$	24.9

Averages (and standard deviations) of translational diffusion coefficients (D_T), hydrodynamic radii of gyration (R_H) and estimated molecular weights for four monodisperse protein and protein–DNA complexes as determined by PCS. Deduced oligomerization states, and corresponding molecular weights are also given. The covalent molecular weights (determined by laser desorption mass spectrometry) of b/HLH/Z and b/HLH are 13.5 and 7.6 kDa, respectively, of the DNA duplex 9.7 kDa. Small discrepancies between observed and calculated molecular weights could arise from partial unfolding of the protein in the absence of DNA, as indicated by our CD results.

remainder involve the DNA backbone. In addition, water-mediated contacts, which would not be visualized at the current resolution limit, cannot be excluded. The density of contacts made with the DNA backbone suggests that subtle geometric and torsional properties of the B-form double helix may play roles in sequence specific recognition. The same may be true of contacts with the adjacent base pairs made by amino acids in the loop and Lys240. The dichroic signal in the CD spectra at 264 nm discussed above provides evidence independent of X-ray crystallography that the DNA binding domain of USF provokes a distortion in DNA. Thus, induced fit of both the protein and the DNA appears to be required for specific, high affinity binding.

b/HLH/Z is a bivalent homotetramer, b/HLH a monovalent homodimer

Our CD results imply that the mode of interaction of the three tested constructs with DNA is very similar. However, CD is a technique that is sensitive fundamentally to local effects: the spectra in the far ultraviolet are a measure of the average local conformation, secondary structure, of the peptide backbone. To study the effect of the various motifs present in USF on the oligomerization state of the proteins and protein–DNA complexes, we turned to photon correlation spectroscopy (PCS) or dynamic light scattering. This approach measures laser light scattered by macromolecules in solution that are undergoing Brownian motion. The fluctuations in intensity of the scattered light can be analyzed to give the macromolecular diffusion coefficient, and estimates of the hydrodynamic radius of gyration and the apparent molecular weight (reviewed in Schmitz, 1990). The method is also sensitive to sample polydispersity because the intensity of light scattering is proportional to the square of solute molecular weight. Hence, high order aggregation of macromolecules in solution can be readily detected by this technology.

Under a variety of solvent conditions, purified, recombinant USF exists in a highly aggregated and polydisperse state with an average oligomer mass exceeding 1000 kDa (data not shown). Removal of the activation domain to give b/HLH/Z and b/HLH eliminates high order aggregation, and both smaller proteins are monodisperse in solution. Presumably aggregation does not occur in the HeLa nucleus and is an artifact of folding errors within the bacterially expressed activation domain, or a result of the absence of additional proteins which associate with this domain *in vivo*. The molecular weights measured by PCS for free and DNA-bound b/HLH and b/HLH/Z (at complex concentrations where the full folding transition is observed by CD spectroscopy) are given in Table I. The macromolecular

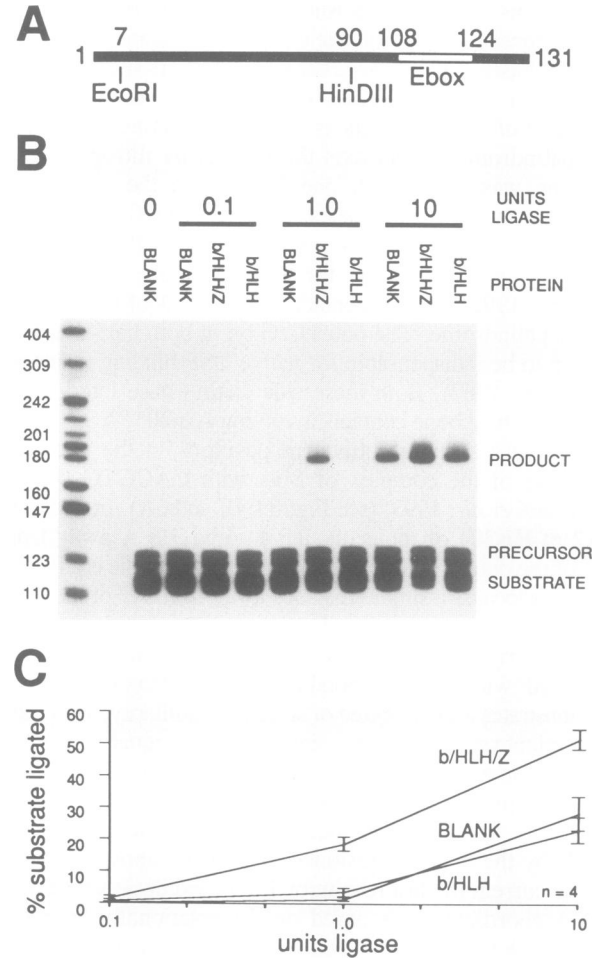

Fig. 7. Bimolecular ligation experiment. (**A**) Cartoon portrayal of the DNA employed. (**B**) Representative example of the experiment. Size marker lengths (in bases) are indicated on the left. The migration positions of ligation product, PCR product precursor and *Eco*RI-cleaved PCR product substrate are indicated on the right. All ligation reactions contained 1 μM DNA, 50 μg/ml BSA and 2 μM (calculated using the monomer molecular mass) b/HLH/Z or b/HLH where indicated. The ligation product (excised from polyacrylamide gels) was characterized by cleavage with *Eco*RI and *Hin*dIII enzymes, and demonstrated to consist exclusively of dimers of the substrate ligated via the *Eco*RI site (data not shown). The ligation product migrates at an anomalously fast rate on 8 M urea gels, presumably because its true palindromic nature favors hairpin formation. (**C**) Average of four independent experiments as shown in B. Error bars represent standard errors of the mean.

mass of b/HLH bound to sDNA predictably corresponds to a protein dimer complexed with one DNA duplex. However, the presence of the leucine zipper alters the behavior of

b/HLH/Z dramatically. Under various solvent conditions, the complex of b/HLH/Z with sDNA occurs as a complex of four polypeptide chains and two DNA duplexes (i.e. b/HLH/Z is a bivalent homotetramer capable of binding simultaneously to two sDNA duplexes). This striking oligomerization state is the only one detected by this highly sensitive technique under the experimental conditions. In the absence of sDNA, b/HLH and b/HLH/Z exhibit masses consistent with those of protein dimer and tetramer, respectively.

Although the results of our PCS studies of DNA binding by b/HLH and b/HLH/Z are secure, they were obtained at concentrations of the protein–DNA complexes in excess of the intranuclear concentration of USF in HeLa cells. We therefore designed an experiment to test whether b/HLH/Z could function as a bivalent homotetramer at physiological intranuclear concentration. USF is highly abundant in eukaryotic nuclei. In the HeLa nucleus the USF concentration can be estimated as 0.5 μM, assuming homogeneous protein distribution throughout a spherical nucleus of radius 2.5 μm with 20 000 molecules/cell, as shown by Sawadogo *et al.* (1988). This estimate represents a lower bound because the distribution of USF within the nucleus may not be homogeneous. To provide independent conformation of simultaneous binding to two spatially separate DNA sites by the tetrameric leucine zipper-containing b/HLH/Z protein at a concentration of 0.5 μM, we studied the effects of exogenously added proteins on the rate of bimolecular ligation of a long duplex DNA with a USF binding site at one end and at the other a cohesive end generated by *Eco*RI cleavage (Figure 7A). After purification, the ligation substrate was incubated either with a blank buffer containing a high concentration of bovine serum albumin (BSA), or with the same buffer containing either b/HLH or b/HLH/Z at a protein–DNA complex concentration of 0.5 μM. DNA ligase was then added, the reaction mixtures incubated for a fixed time and the ligation product and the substrate resolved by denaturing PAGE. Results of ligation reactions performed with three different concentrations of DNA ligase are shown in Figure 7B. Average quantities derived from four independent, replicate experiments are plotted in Figure 7C. Presence of the tetrameric b/HLH/Z protein in the ligation reaction greatly enhances the rate of bimolecular ligation of the substrate over the rate observed in the control experiment where only BSA is present. In contrast, and as expected, the rate of bimolecular ligation in the presence of the dimeric b/HLH protein does not differ, within experimental precision, from that observed with BSA alone.

The combination of the results of the EMSA titrations, CD, PCS and bimolecular ligation experiments demonstrate that b/HLH/Z, the USF DNA binding domain, exists as a bivalent homotetramer at micromolar concentrations, and its enhancement of the rate of bimolecular ligation is demonstrable at a protein–DNA complex concentration of 0.5 μM. Therefore, the dissociation constant for tetramerization both in the presence and absence of DNA containing the USF binding site must be micromolar or less, suggesting that USF can function in HeLa nuclei as a bivalent homotetramer.

Tetramerization of b/HLH/Z proteins

The b/HLH/Z proteins Myc (Dang *et al.*, 1989), AP-4 (Hu *et al.*, 1990) and TFEB (Fisher *et al.*, 1991) have been shown to exist as tetramers in the absence of DNA. We show that the b/HLH/Z DNA binding domain of USF forms stable tetramers when its leucine repeat is intact. The experimentally observed 2:1 protein monomer to DNA stoichiometry, the palindromic nature of the recognized E-box and the specific recognition of DNA by a dimer of the construct missing the heptad repeat, which we have demonstrated both spectroscopically and crystallographically, imply that the tetramer either dissociates into dimers upon binding DNA or binds two independent sites simultaneously. Sharp and coworkers have presented evidence that unpurified, *in vitro* translated TFEB dissociates into dimers when it is mixed with DNA (Fisher *et al.*, 1991). The presence of large quantities of endogenous b/HLH/Z proteins in reticulocyte lysate (Prendergast *et al.*, 1991), which may form a variety of DNA binding hetero-oligomers with the translated protein, complicates the interpretation of those results. In this paper we demonstrate that the tetrameric b/HLH/Z domain of USF, purified to homogeneity, simultaneously binds to two independent pieces of DNA and exhibits a solution mass consistent with a bivalent homotetramer.

Because of the high similarity between the Max and USF sequences within the DNA binding and oligomerization moieties, both proteins are likely to exhibit the same oligomerization states on and off the DNA. The b/HLH/Z domain of Max complexed with its cognate DNA crystallized in a high symmetry space group as a dimer (Ferre-D'Amare *et al.*, 1993). This finding suggests that the propensity for Max tetramer formation is comparable with or weaker than the crystal packing forces responsible for stabilizing the Max–DNA cocrystalline lattice.

The recent discovery that USF can activate transcription by associating with the initiator element in addition to the upstream E-box and that ablation of the initiator element in the context of an otherwise intact promoter decreased transactivation (Du *et al.*, 1993), as well as the observation based on kinetic measurements that USF can efficiently transfer from spatially separated binding sites by passing through a doubly bound intermediate (Sawadogo, 1988) can both be reinterpreted in the light of our results. DNA looping is of critical importance in transcriptional control (Ptashne, 1988) and a bivalent homotetramer provides a compact molecular tethering device that can effectively accomplish DNA looping. The variety and abundance of b/HLH/Z proteins, and the fact that the leucine zippers of some of them associate only with the zippers of a defined subset of family members (Beckmann and Kadesch, 1991), increase the combinatorial range and specificity that obtains from the selective dimerization seen in HLH proteins.

In conclusion, we have demonstrated that the fundamental DNA binding unit for a b/HLH/Z protein USF is, as with the better characterized HLH proteins, a dimer, stabilized by the HLH motif in both cases. Our cocrystal structure demonstrates that sequence-specific DNA binding by this protein proceeds through an induced fit mechanism, in which >40 amino acids per dimer change conformation. Comparison of this structure with the homologous structure of Max shows that only a dozen or so protein–DNA contacts per dimer suffice for recognition. We also demonstrate that the b/HLH/Z domain exists as a bivalent homotetramer in solution at physiological intranuclear concentration, and that the leucine zipper element is required for the interaction of the dimeric b/HLH elements. This work constitutes the first

rigorous demonstration that a b/HLH/Z protein can simultaneously bind two distinct sites and opens the way for further investigations of the biological role of b/HLH/Z-mediated DNA looping, and interactions, in various oligomerization states, of these proteins with other classes of transcription factors.

Materials and methods

Reagent production and purification

DNA was synthesized, purified, quantitated and annealed as described (Ferré-D'Amaré *et al.*, 1993). All protein constructs were overexpressed in *Escherichia coli* [Bl21(DE3)pLysS] using the T7 RNA polymerase system (Studier *et al.*, 1990). The USF construct is the one described by Kaulen *et al.* (1991); this produces a protein of molecular weight 33.5 kDa, which migrates on SDS gels with an apparent molecular weight of 43 kDa and has been referred to in the literature as the '43 kDa form of USF'. b/HLH/Z was derived from this sequence essentially by the methodology described by Pognonec and Roeder (1991). b/HLH was derived in the same way from the double cysteine to serine mutant described by Pognonec *et al.* (1992). Cells transformed with the appropriate expression vector, grown to an optical density of 1.0 at 596 nm and induced by the addition of IPTG to 0.5 mM for 5–8 h at 30°C, were collected by low speed centrifugation and lysed as in Pognonec *et al.* (1991). USF was fractionated with ammonium sulfate and then purified to apparent homogeneity by a combination of anion and cation exchange chromatography. All other constructs were first fractionated by chromatography over heparin–Sepharose (Pharmacia) and then further purified by cation exchange. The purified, recombinant proteins were characterized by laser desorption mass spectrometry and chemical N-terminal sequencing. These analytical techniques revealed that except for b/HLH, which entirely retains its initiation methionine, the proteins are heterogeneous at the N-terminus, with variable proportions possessing the methionine. Otherwise, all proteins are unmodified by proteolysis. Care was taken to remove traces of NP-40 from the proteins by performing successive washes with 0.5% *n*-octyl-β-glucopyranoside containing buffers and detergent-free buffers while the proteins were bound to the ion exchange columns. Purified proteins were quantitated by UV spectrophotometry in the native state, using calculated extinction coefficients (based on chromophore contributions from denatured proteins, Edelhoch, 1967), since denaturation in 6 M guanidine hydrochloride produced absorbance shifts which were insignificant relative to the precision of all other experiments (data not shown).

EMSAs

Single-stranded DNA was phosphorylated with [γ-^{32}P]ATP, phenol–chloroform extracted, ethanol precipitated, purified over Sephadex G-25 and quantitated by UV spectrophotometry. Binding reactions (10 μl each) contained 100 mM KCl, 10 mM Tris–HCl pH 8.0, 10% (v/v) glycerol, 5 mM DTT, 1 mM $MgCl_2$, 1 μM annealed labeled DNA and USF. Binding reactions for b/HLH/Z and b/HLH also contained 25 μg/ml BSA. Binding reactions for USF were incubated on ice for 1 h, whereas binding reactions for the other constructs were incubated at room temperature for 30 min. After incubation, 5 μl of 20% (v/v) glycerol were added to each reaction and the mixture was loaded onto a pre-run gel. USF was resolved on 4% acrylamide 39:1 (w/w) acrylamide–*bis*acrylamide gels cast and run in 380 mM glycine, 50 mM Tris base, 2.7 mM EDTA, at 10 V/cm for 1.5 h at room temperature. b/HLH/Z and b/HLH binding reactions were resolved on 10% acrylamide 39:1 (w/w) acrylamide–*bis*acrylamide gels cast in the same buffer and run at 10 V/cm for 2 h at room temperature. Gels were dried and visualized by autoradiography on imaging plates. Bands were quantitated using the volume integration function on ImageQuant v.3.15 (Molecular Dynamics).

CD spectroscopy

CD spectra were obtained with an AVIV model 62DS spectropolarimeter at room temperature, using 1, 0.5, 0.1 or 0.01 mm path-length cylindrical quartz cells (Hellma) and 1.5 nm bandwidth, 1 s averaging time per point, measurements every 0.5 nm. Ten spectra were taken for each sample, averaged, blanked and smoothed using software provided by the manufacturer. Blanks were taken before and after each data collection and subtracted from each other to monitor instrument drift. All spectra were collected with the proteins, DNA or DNA–protein complexes in the same buffer used for the binding reactions for EMSA (without BSA), described above. The proteins and DNA–protein complexes were taken into the buffer used for CD by microdialysis and the actual buffer against which samples were equilibrated was used as a blank, both for quantitation of the sample by UV spectrometry and for CD. Spectra were analyzed using the program CONTIN (Provencher and Glöckner, 1981).

Cocrystallization and structure determination

$(b/HLH)_2$–DNA complex prepared by mixing two molar equivalents of b/HLH monomer with annealed, double-stranded DNA of the sequence depicted in Figure 1B, was concentrated by microfiltration to 1.4 mM in a buffer containing 100 mM KCl and 10 mM HEPES pH 7.5. Cocrystals were grown at 4°C by vapor diffusion of hanging drops prepared by mixing equal volumes of complex and a reservoir solution consisting of 15% PEG 400, 15% glycerol, 100 mM KCl, 2.8 mM $MgCl_2$, 1.4 mM Cd(II) acetate and 100 mM sodium acetate pH 4.75. Orthorhombic cocrystals (space group $P2_12_12_1$; a = 136.6 Å, b = 54.7 Å, c = 44.4 Å; one $(b/HLH)_2$–DNA complex per asymmetric unit) grew over the course of several days to typical sizes of 0.3 × 0.3 × 0.2 mm^3. Oscillation photographs were collected at −20°C, from one crystal using CuKα X-radiation with a Rigaku RAXIS-IIc area detector, from three crystals with 0.91 Å X-radiation and imaging plates at beam-line F1 of the Cornell High Energy Synchrotron Source (CHESS) (Finkelstein, 1992), and from three crystals with 0.95 Å X-radiation and imaging plates at the wiggler beam-line X25 of the National Synchrotron Light Source (NSLS; Brookhaven National Laboratory). Diffraction data were reduced to structure factor amplitudes and scaled using the programs DENZO and SCALEPACK (Z.Otwinowski, personal communication). Fresh crystals diffract isotropically to minimum Bragg spacings of 1/2.6 Å^{-1}, but severe radiation sensitivity limited useful data to 3.5 Å in our laboratory and 2.9 Å at CHESS and NSLS. The merged data set is 76.4% complete between 15 and 2.9 Å (32 049 measurements of 6038 unique reflections; R_{merge} = 7.6%, for all data $>1\sigma$). Between 3.0 and 2.9 Å, the data are 49.3% complete with $<I/\sigma(I)>$ = 3.2. The structure was solved by molecular replacement using a dimer of amino acids 22 through 80 and the central 10 bp (i.e. a total of 20 bp) of the refined Max–DNA structure (Ferré-D'Amaré *et al.*, 1993) as a search model. A rotation search between 10.0 and 4.5 Å, followed by Patterson correlation refinement with XPLOR (Brünger, 1992) yielded a solution 4.8σ above mean peak height with θ_1, θ_2, θ_3 equal to 272.1°, 72.5° and 177.9°, respectively. This solution was employed in a translation search with XPLOR, resulting in a family of symmetry-related solutions 3.2σ above mean peak level. The model was modified at this stage by incorporating the correct DNA sequence and the overhanging nucleotides (for which electron density was present in 2fo-fc maps). All side chains were stripped and the polyalanine model without the loop region with the full 21 bp DNA was subjected to positional refinement to yield an R-factor of 36.7% from 10.0 to 3.1 Å. Side chains, loop and N-terminal residues were gradually added to the model by inspecting omit and annealed-omit maps. Manual rebuilding (Jones *et al.*, 1991) and positional refinement brought the R-factor to 28.9%. At this stage, the DNA sequence was inverted, resulting in a decrease of the R-factor by 1.5%; this model had an R-factor of 27.1% between 6.0 and 3.1 Å, using an overall isotropic temperature factor of 17.0 Å^2. Further manual rebuilding, positional refinement, phase extension to 2.9 Å and refinement of tightly restrained individual isotropic temperature factors (r.m.s. deviation of temperature factors between bonded atoms = 2.71 Å^2) resulted in the current model which contains 1923 non-hydrogen atoms (the complete amino acid sequence from residues 197 to 260 plus the initiation methionine for both polypeptides and the 21 bp duplex DNA; no solvent molecules were added) and has an R-factor of 23.6%, between 6.0 and 2.9 Å, for reflections with $F>2\sigma(F)$. The model has deviations of 0.019 Å and 3.05° from ideal bond lengths and bond angles, respectively. DNA conformation was analyzed using the program NEWHEL92 (R.E.Dickerson, personal communication). Atomic coordinates will be deposited in the Brookhaven Protein Data Bank.

PCS

PCS was performed with a Biotage model dp-801 molecular size detector. Apparent translational diffusion coefficients, molecular masses, hydrodynamic radii of gyration and degree of sample polydispersity were calculated from the autocorrelation function using the manufacturer's software. Samples were taken into 100 mM KCl, 10 mM HEPES pH 7.5 and 10 mM (USF and b/HLH/Z) or no (b/HLH) DTT by microdialysis and filtered through 200 Å filters to remove any dust particles. Sample concentrations ranging from 4.3 to 62 μM were chosen empirically to ensure that the dynamic light scattering signal exceeded the instrumental detection threshold. Twenty or more independent measurements were made from each sample and the reported values are calculated arithmetic means.

Bimolecular ligation experiment

A DNA template with the sequence 5′ cgttcggaat tcccagctct ccaagataat cccagactgc tctatggaga gacccaagtc tggccagagt aaagatggga ttctatccaa agcttggtat tatgttttta ggcc**acgtg**a ccggatccgc g 3′ was synthesized and purified with a

denaturing 8 M urea gel after phosphorylation and amplified by PCR using two 23 nucleotide primers complementary to either end of the corresponding strands (the E-box CACGTG is highlighted with bold letters). PCR was carried out under standard conditions, and ^{32}P was incorporated to a final specific activity of 3.5×10^{15} Bq/mol by including [α-^{32}P]dATP in the PCR reaction. The product was then cleaved using *Eco*RI restriction endonuclease (New England Biolabs) and, after phenol–chloroform extraction, all short sequences (PCR failure products and the short restriction fragments) were removed by gel purification under denaturing conditions. The purified DNA was quantitated by Cerenkov counting and annealed by linear cooling as with the EMSA probes. EMSA studies documented efficient binding to b/HLH/Z through one specific site per DNA under conditions described above (data not shown). Ligation reactions (10 μl each) all contained 1 μM annealed template, 50 mM Tris–HCl pH 7.8, 10 mM $MgCl_2$, 20 mM DTT, 1 mM ATP, 50 μg/ml BSA and 0, 0.1, 1 or 10 units of DNA ligase (New England Biolabs) as defined by the manufacturer. Ligation reactions were started by adding the ligase (diluted in the same buffer) to the buffered DNA–protein mixture, incubated at 16°C for 30 min and stopped by adding phenol–chloroform and vortexing. The DNA was extracted, ethanol precipitated and examined by electrophoresis on 10% acrylamide 8 M urea denaturing gels. Phosphorylated pBR322 DNA *Msp*I digest was used for molecular size markers. Visualization and quantitation were as described for the EMSA experiments.

Acknowledgements

We thank S.Jacques for help with protein production, S.Cohen and B.Chait for mass spectroscopy, D.Cowburn for advice on CD measurements and data analysis, A.Gazes and the staff of the Rockefeller University Computing Services for technical assistance, T.Irving for assistance in using beam-line F1 at CHESS, and R.Sweet for help with beam-line X25 at NSLS. We are grateful to D.Baltimore, K.L.Clark, J.L.Kim, M.M.Konarska, X.-P.Kong, J.Kuriyan, P.Model, D.B.Nikolov, Z.Otwinowski and G.A.Petsko for valuable suggestions. A.F.-D. is a David Rockefeller Fellow. P.P. was supported by an European Molecular Biology Organization post-doctoral long-term fellowship. This work was supported by the N.I.H. (R.G.R.), and the general support from the Pew Trusts to the Rockefeller University.

References

Anthony-Cahill,S.J., Benfield,P.A., Robert,F., Wasserman,Z.R., Brenner,S.L., Stafford,W.F.I., Altenbach,C., Hubbel,W.L. and DeGrado,W.F. (1992) *Science*, **255**, 979–983.

Ayer,D.E., Kretzner,L. and Eisenman,R.N. (1993) *Cell*, **72**, 211–222.

Beckmann,H. and Kadesch,T. (1991) *Genes Dev.*, **5**, 1057–1066.

Beckmann,H., Su,L.-K. and Kadesch,T. (1990) *Genes Dev.*, **4**, 167–179.

Bexevanis,A. and Vinson,C.R. (1993) *Curr. Opin. Genet. Dev.*, **3**, 278–285.

Blackwood,E.M. and Eisenman,R.N. (1991) *Science*, **251**, 1211–1217.

Brünger,A.T. (1992) *X-PLOR*. Version 3.1 manual. Yale University, New Haven, CT.

Bungert,J., Düring,F. and Seifart,K.H. (1992) *J. Mol. Biol.*, **223**, 885–898.

Cai,M. and Davis,R.W. (1990) *Cell*, **61**, 437–446.

Carr,C.S. and Sharp,P.A. (1990) *Mol. Cell. Biol.*, **10**, 4384–4388.

Carthew,R.W., Chodosh,L.A. and Sharp,P.A. (1985) *Cell*, **43**, 439–448.

Dang,C.V., McGuire,M., Buckmire,M. and Lee,W.M.F. (1989) *Nature*, **337**, 664–666.

Dang,C.V., Dolde,C., Gillison,M.L. and Kato,G.J. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 599–602.

Du,H., Roy,A.L. and Roeder,R.G. (1993) *EMBO J.*, **12**, 501–511.

Edelhoch,H. (1967) *Biochemistry*, **6**, 1948–1954.

Ellenberger,T.E., Brandl,C.J., Struhl,K. and Harrison,S.C. (1992) *Cell*, **71**, 1223–1237.

Ferré-D'Amaré,A.R., Prendergast,G.C., Ziff,E.B. and Burley,S.K. (1993) *Nature*, **363**, 38–45.

Finkelstein,K.D. (1992) *Rev. Sci. Instrum.*, **63**, 305–308.

Fisher,F. and Goding,C.R. (1992) *EMBO J.*, **11**, 4103–4109.

Fisher,D.E., Carr,C.S., Parent,L.A. and Sharp,P.A. (1991) *Genes Dev.*, **5**, 2342–2352.

Fisher,D.E., Parent,L.A. and Sharp,P.A. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 11779–11783.

Fisher,D.E., Parent,L.A. and Sharp,P.A. (1993) *Cell*, **72**, 467–476.

Gregor,P.D., Sawadogo,M. and Roeder,R.G. (1990) *Genes Dev.*, **4**, 1730–1740.

Hu,Y.-F., Lüscher,B., Admon,A., Mermod,N. and Tjian,R. (1990) *Genes Dev.*, **4**, 1741–1752.

Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) *Acta Crystallogr.*, **A47**, 110–119.

Kaulen,H., Pognonec,P., Gregor,P.D. and Roeder,R.G. (1991) *Mol. Cell. Biol.*, **11**, 412–424.

Lennard,A.C. and Egly,J.M. (1987) *EMBO J.*, **6**, 3027–3034.

Luzzati,P.V. (1952) *Acta. Crystallogr.*, **5**, 802–810.

Meisterernst,M., Horikoshi,M. and Roeder,R.G. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 9153–9157.

Miyamoto,N.G., Moncollin,V., Egly,J.M. and Chambon,P. (1985) *EMBO J.*, **4**, 3563–3570.

Murre,C., McCaw,P.S. and Baltimore,D. (1989) *Cell*, **56**, 777–783.

Pognonec,P. and Roeder,R.G. (1991) *Mol. Cell. Biol.*, **11**, 5125–5136.

Pognonec,P., Kato,H., Sumimoto,H., Kretzschmar,M. and Roeder,R.G. (1991) *Nucleic Acids Res.*, **19**, 6650–6650.

Pognonec,P., Kato,H. and Roeder,R.G. (1992) *J. Biol. Chem.*, **267**, 24563–24567.

Prendergast,G.C., Lawe,D. and Ziff,E.B. (1991) *Cell*, **65**, 395–407

Prendergast,G.C. and Ziff,E.B. (1991) *Science*, **251**, 186–189.

Provencher,S.W. and Glöckner,J. (1981) *Biochemistry*, **20**, 33–37.

Ptashne,M. (1988) *Nature*, **335**, 683–689.

Sawadogo,M. (1988) *J. Biol. Chem.*, **263**, 11994–12001.

Sawadogo,M. and Roeder,R.G. (1985) *Cell*, **43**, 165–175.

Sawadogo,M., Van Dyke,M.W., Gregor,P.D. and Roeder,R.G. (1988) *J. Biol. Chem.*, **263**, 11985–11993.

Sawyer,L. and James,M.N.G. (1982) *Nature*, **295**, 79–80.

Schmitz,K.S. (1990) *An Introduction to Dynamic Light Scattering by Macromolecules*. Academic Press, Boston, MA.

Wechsler,D.S. and Dang,C.V. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 7635–7639.

Workman,J.L., Roeder,R.G. and Kingston,R.E. (1990) *EMBO J.*, **9**, 1299–1308.

Zervos,A.S., Gyuris,J. and Brent,R. (1993) *Cell*, **73**, 223–232.

Ziff,E.B. and Evans,R.M. (1978) *Cell*, **15**, 1463–1475.

Received on July 26, 1993; revised on October 7, 1993