CCHX Zinc Finger Derivatives Retain the Ability to Bind Zn(II) and Mediate Protein-DNA Interactions*

Raina J. Y. Simpson‡, Edward D. Cram, Robert Czolij, Jacqueline M. Matthews, Merlin Crossley, and Joel P. Mackay§

From the School of Molecular and Microbial Biosciences, G08, University of Sydney, New South Wales 2006, Australia

Received for publication, October 31, 2002, and in revised form, May 2, 2003
Published, JBC Papers in Press, May 7, 2003, DOI 10.1074/jbc.M211146200

It is estimated that around 3% of all human gene products contain one or more zinc-binding domains, or zinc fingers (ZnFs).1 These domains are defined by the presence of one or more Zn(II) ions that stabilize the folded structure of the domain. More than 20 different classes of zinc fingers have been described, and these differ in the number of Zn(II) ions bound and in the spacing and identities of the ligands. The Zn(II) ion is normally coordinated by histidine and cysteine residues, although aspartate has been reported in LIM domains (1). The most common configuration is termed the classical (CCHH or C2H2) ZnF, which has the consensus sequence Cys-X2–4–Cys-X12–14-His. Over 14,000 of these domains are predicted in current sequence data bases.

The majority of classical ZnFs are found in transcription factors, and a single transcription factor may contain more than 30 ZnF domains. It is well established that many of these domains mediate specific protein-DNA interactions (2), although recent reports have also demonstrated that at least some classical ZnFs can act as protein recognition motifs (for a review, see Ref. 3). The three-dimensional structures of a number of ZnF-DNA complexes have been determined, and a great deal is understood about the roles of individual amino acids in determining DNA sequence specificity (4, 5). These ZnFs bind to DNA using a tandem array of more than one ZnF (often three), where each ZnF contacts three base pairs.

The identification of ZnF motifs in the ever-growing number of protein sequences is based primarily on the presence of conserved cysteine or histidine residues and the spacing between them. Further, DNA binding ability is sometimes inferred when three or more contiguous classical ZnFs are identified. This ability to infer function from sequence is becoming increasingly important as the amount of available sequence data increases. Interestingly, an examination of sequence data bases reveals that there are a number of proteins that contain sequences that correspond to one or more classical ZnFs, with the exception that the final zinc-ligating residue in the last ZnF is neither cysteine nor histidine (for examples, see Fig. 1). The question therefore arises as to whether these sequences are capable of folding and forming modules that are functional even in the absence of the typical final zinc-binding residue or whether the mutation leads to proteins that can no longer fold or (for example) bind to DNA.

In order to address this question, we have investigated the physical and functional properties of a panel of point mutants based on the transcriptional repressor basic Kruppel-like factor/Kruppel-like factor 3 (BKLF) (6). BKLF binds to DNA sequences containing CACCC motifs by means of three characteristic Kruppel-like ZnFs (Kruppel-like fingers are a subset of classical CCHH fingers with significant homology to those found in the archetypal protein Drosophila regulatory protein Krüppel). We chose to study the third or C-terminal zinc finger of BKLF (BKLF-F3, or BF3). We show that the third zinc finger is essential for high affinity DNA binding and that the final zinc-ligating histidine of BKLF-F3 can be substituted with a number of different residues without severely compromising the DNA binding ability of BKLF. Further, the mutant BF3

* This work was supported by a grant from the Australian Research Council. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ An Australian Postgraduate Awards.
§ An Australian Research Fellow. To whom correspondence should be addressed. Tel.: 61-2-9351-3906; Fax: 61-2-9351-4726; E-mail: j.mackay@mmb.usyd.edu.au.

1 The abbreviations used are: ZnF, zinc finger; EMSA, electrophoretic mobility shift assay; TCEP, tris(2-carboxyethyl)phosphine; HPLC, high pressure liquid chromatography; GST, glutathione S-transferase; BF3, BKLF-F3; NOESY, total correlation spectroscopy; NOE, nuclear Overhauser effect; NOESY, NOE spectroscopy; HSQC, heteronuclear single quantum correlation spectroscopy.

This paper is available on line at http://www.jbc.org

28011
domains still bind Zn(II) and form substantial secondary structure, although they are clearly not as well ordered as the wild-type domain. Remarkably, the His \( \rightarrow \) Asn mutant binds Zn(II) with an affinity that is essentially indistinguishable from that of the wild-type BF3.

These results demonstrate that three side chains can be sufficient to bind a Zn(II) ion. Further, our data show that such domains, even when partially folded, can act in concert with other ZnFs to bind DNA. These findings further our understanding of the basic ZnF scaffold and show that attempts to identify DNA-binding ZnFs from amino acid sequence data should not necessarily exclude apparently incomplete ZnF configurations.

**EXPERIMENTAL PROCEDURES**

Plasmid Construction—The region encoding zinc fingers 1–3 of BKLFL (BKLFL-F1–3 residues 254–344) was amplified from the murine BKLFL cDNA using primers CCGGATCCACATCCAGAAGGCAGGAGTAC and CGGAAATCAGTACTGCCGTTTGCAGGAGTT (A886), using the original murine BKLFL cDNA clone in the vector pMT2-BKLFL (9). The resulting fragment was digested with BamHI and EcoRI and inserted into pGEX-2T to generate in-frame fusions of BF3 with glutathione S-transferase, and the plasmids were transformed into *Escherichia coli* BL21 (DE3) cells. This plasmid also contains a thymol cleavage site between the GST and the insert. Mutant derivatives were similarly constructed using primers related to A886 but differing in the codon corresponding to BKLF H341, the final histidine.

Electrophoretic Mobility Shift Assays (EMSAs)—In order to confirm the sequences of all constructs.

**Analytical Ultracentrifugation**—Sedimentation equilibrium experiments were performed using an Optima™ XL-A analytical ultracentrifuge (Beckman Instruments) equipped with an An-60ti rotor. BF3E (at concentrations of 8.6, 20, and 33.6 \( \mu \)M) was centrifuged against a matched buffer at 25 °C for 2 h at 360,000 rpm, 4 °C, 20 min), was loaded onto glutathione-Sepharose beads. Unbound proteins were washed away from the beads with wash buffer (50 mM Tris, 100 mM NaCl, 1% Triton X-100, 1.4 mM phenylmethylsulfonyl fluoride, 1.4 mM \( \beta \)-mercaptoethanol, pH 8.0) and then eluted with thrombin buffer (50 mM Tris, 150 mM NaCl, 2.5 mM CaCl\(_2\), pH 8.0). Thrombin was added, and the mixture was incubated for either 2 or 24 h at 37 °C or overnight at 4 °C. The eluted peptides were lyophilized, redissolved in water, and further purified by reversed-phase HPLC (using a gradient of 5–95% acetonitrile in 0.1% trifluoroacetic acid). The purified peptides were lyophilized and stored at –20 °C. The identity of each peptide was confirmed using positive ion electrospray mass spectrometry.

**For UV CD Spectropolarimetry**—Each HPLC-purified peptide was dissolved in a solution containing TCEP (1 mM) and ZnSO\(_4\) (1 mM) to concentrations of either 30 \( \mu \)M (BF3, BF3D, BF3E, and BF3N) or 12 \( \mu \)M (BF3A). The pH of these solutions was 7.8. A far UV CD spectrum of each peptide was taken, and the pH was then adjusted to ∼5.5 with the addition of NaOH, and a second spectrum of each peptide was recorded. CD spectra were recorded at 25 °C on a Jasco J-720 spectropolarimeter equipped with a Neslab RTE-111 temperature controller. Spectra were recorded in a 1-mm path length cell with a resolution of 0.5 nm and bandwidth of 1 nm over the wavelength range of 190–250 nm. Each spectrum represented the average of three scans accumulated at a speed of 20 nm min\(^{-1}\) with a response time of 1 s.

**For the Zn(II) titration experiments, aliquots of a solution containing ZnCl\(_2\) (13–20 mM; pH 5.5) were added to solutions of the wild-type BF3 (13 \( \mu \)M) and the mutant BF3N (20 \( \mu \)M), each containing 0.5 mM TCEP, pH 5.5. CD spectra were taken at each point in the titration, allowing 5 min for equilibration after each Zn(II) addition. Spectra were recorded over the wavelength range of 195–200 nm with a resolution of 1 nm and an average of 50 scans. Spectra were base line-corrected by subtraction of a spectrum of TCEP/ZnSO\(_4\) buffer alone.

Values of the association constant for zinc binding \((K_a)\) were determined by plotting the change in ellipticity at a single wavelength against the total Zn(II) concentration. Nonlinear least squares analysis was used to determine the Zn(II)-binding affinities, using the equations:

\[
\theta = \theta_0 + \frac{[PZn]}{K_a + [PZn]} \quad \text(Eq. 3)
\]

and

\[
K_a = \frac{[PZn][\theta_2 - \theta_1]}{[PZn][\theta_2 - \theta_1] - [PZn][\theta_2 - \theta_1]} \quad \text(Eq. 4)
\]

where \(\theta\) represents the observed ellipticity, \(\theta_0\) is the ellipticity of the apo-protein, \([PZn]\) is the concentration of the protein-Zn(II) complex, \([PZn]\_\text{tot}\) is the total protein concentration, \(\theta_2\_\text{tot}\) is the ellipticity of the holoprotein, \(K_a\) is the association constant for Zn(II) binding, and \([Zn]\_\text{tot}\) is the total Zn(II) concentration. The final values of \(K_a\) were obtained by averaging the values determined at 105, 195, 197, 199, and 200 nm.

**Nuclear Magnetic Resonance Spectroscopy**—Samples of BF3 were dissolved in a solution containing TCEP (1 mM) and ZnSO\(_4\) (1 mM) to concentrations of either 30 \( \mu \)M (BF3, BF3D, BF3E, and BF3N) or 12 \( \mu \)M (BF3A). The pH of these solutions was 7.8. A far UV CD spectrum of each peptide was taken, and the pH was then adjusted to ∼5.5 with the addition of NaOH, and a second spectrum of each peptide was recorded. CD spectra were recorded at 25 °C on a Jasco J-720 spectropolarimeter equipped with a Neslab RTE-111 temperature controller. Spectra were recorded in a 1-mm path length cell with a resolution of 0.5 nm and bandwidth of 1 nm over the wavelength range of 190–250 nm. Each spectrum represented the average of three scans accumulated at a speed of 20 nm min\(^{-1}\) with a response time of 1 s.

**For the Zn(II) titration experiments, aliquots of a solution containing ZnCl\(_2\) (13–20 mM; pH 5.5) were added to solutions of the wild-type BF3 (13 \( \mu \)M) and the mutant BF3N (20 \( \mu \)M), each containing 0.5 mM TCEP, pH 5.5. CD spectra were taken at each point in the titration, allowing 5 min for equilibration after each Zn(II) addition. Spectra were recorded over the wavelength range of 195–200 nm with a resolution of 1 nm and an average of 50 scans. Spectra were base line-corrected by subtraction of a spectrum of TCEP/ZnSO\(_4\) buffer alone.

Values of the association constant for zinc binding \((K_a)\) were determined by plotting the change in ellipticity at a single wavelength against the total Zn(II) concentration. Nonlinear least squares analysis was used to determine the Zn(II)-binding affinities, using the equations:

\[
\theta = \theta_0 + \frac{[PZn]}{K_a + [PZn]} \quad \text(Eq. 3)
\]

and

\[
K_a = \frac{[PZn][\theta_2 - \theta_1]}{[PZn][\theta_2 - \theta_1] - [PZn][\theta_2 - \theta_1]} \quad \text(Eq. 4)
\]

where \(\theta\) represents the observed ellipticity, \(\theta_0\) is the ellipticity of the apo-protein, \([PZn]\) is the concentration of the protein-Zn(II) complex, \([PZn]\_\text{tot}\) is the total protein concentration, \(\theta_2\_\text{tot}\) is the ellipticity of the holoprotein, \(K_a\) is the association constant for Zn(II) binding, and \([Zn]\_\text{tot}\) is the total Zn(II) concentration. The final values of \(K_a\) were obtained by averaging the values determined at 105, 195, 197, 199, and 200 nm.

**Analytical Ultracentrifugation**—Sedimentation equilibrium experiments were performed using an Optima™ XL-A analytical ultracentrifuge (Beckman Instruments) equipped with an An-60ti rotor. BF3E (at concentrations of 8.6, 20, and 33.6 \( \mu \)M) was centrifuged against a matched buffer at 25 °C at 30,000 and 42,000 rpm, using double sector cells. Data were collected as \(A_{\text{soc}}\) versus radius scans in 0.001-cm increments, and 10 scans were averaged for each data set. Base line correction was achieved by the subtraction of a spectrum of TCEP/ZnSO\(_4\) buffer alone.

Values of the association constant for zinc binding \((K_a)\) were determined by plotting the change in ellipticity at a single wavelength against the total Zn(II) concentration. Nonlinear least squares analysis was used to determine the Zn(II)-binding affinities, using the equations:

\[
\theta = \theta_0 + \frac{[PZn]}{K_a + [PZn]} \quad \text(Eq. 3)
\]

and

\[
K_a = \frac{[PZn][\theta_2 - \theta_1]}{[PZn][\theta_2 - \theta_1] - [PZn][\theta_2 - \theta_1]} \quad \text(Eq. 4)
\]

where \(\theta\) represents the observed ellipticity, \(\theta_0\) is the ellipticity of the apo-protein, \([PZn]\) is the concentration of the protein-Zn(II) complex, \([PZn]\_\text{tot}\) is the total protein concentration, \(\theta_2\_\text{tot}\) is the ellipticity of the holoprotein, \(K_a\) is the association constant for Zn(II) binding, and \([Zn]\_\text{tot}\) is the total Zn(II) concentration. The final values of \(K_a\) were obtained by averaging the values determined at 105, 195, 197, 199, and 200 nm.
preparation by dissolving either 3.5 mg of $^{15}$N-labeled BF3 or 11 mg of BF3 in H$_2$O/D$_2$O (95:5) containing 1.5 molar equivalents of both TCEP and ZnSO$_4$. The pH was adjusted to 5.5 using 0.1 and 0.01 m NaOH; this gave sample concentrations of 1 mM for $^{15}$N-labeled BF3 and 3 mM for BF3.

For the comparison of wild-type BF3 with BF3E, both peptides were dissolved in a solution containing TCEP (1 mM) and ZnSO$_4$ (1 mM). The pH was adjusted to 5.5, and each solution was supplemented with D$_2$O (5%, v/v) and d$_6$-(trimethylsilyl)propionic acid (1 µl). The final sample concentrations were 100 µM (BF3) and 85 µM (BF3E). The pH of the BF3E sample was subsequently dropped to pH ~2 to obtain the spectrum of the unfolded protein. NMR spectra were recorded on a Bruker DRX600 spectrometer equipped with a triple resonance (HCN) probe and three-axis pulsed-field gradients. NMR spectra used for the comparison of BF3 and BF3E were recorded at 298 K, whereas those used for structure determination were acquired at 280 K. The solvent signal was suppressed using pulsed field gradients. One-dimensional $^1$H spectra consisted of 128 scans collected at 8,192 complex data points over a spectral width of 7,200 Hz. The following homonuclear two-dimensional spectra were used to assign the $^{15}$N, $^1$H HSQC spectra: TOCSY (11) ($\tau_m = 70$ ms), DQFCOSY (12), and NOESY (13) ($\tau_m = 200$ ms). Three-dimensional HNHA (14) and three-dimensional TOCSY-HSQC (15) ($\tau_m = 70$ ms) experiments were used to assign the $^{15}$N,$^1$H HSQC spectra of the $^{15}$N-labeled BF3, whereas the HNHA spectrum was used to derive $^3$J$_{HNH}$ coupling constants. Spectra were processed as described previously (16). The $^1$H frequency scale was referenced to $d_6$-(trimethylsilyl)propionic acid at 0.00 ppm.

**Determination of the Structure of BF3—**Resonance assignment was carried out using the standard homonuclear sequential assignment method (17). Cross-peaks in the two-dimensional NOESY spectra were integrated in XEASY and converted to upper distance limits using the programs MOLMOL (30), PROCHECK (31), and WhatIf (32). Distance limits involving H methylene groups were corrected to 0.1 Å and those involving isopropyl groups were corrected to 0.5 Å. Distance limits involving histidine side chains were corrected to 0.7 Å. In order to include $^3$J$_{HNH}$ coupling constants measured from the HNHA spectrum (14), residues with positive $^3$J$_{HNH}$ were verified using methods as described in Ref. 19.

Initially, structure calculations were performed in DYANA, and additional NOEs were assigned iteratively based on earlier sets of structures. At this stage, structures were calculated without incorporating the Zn(II) atom. Analysis of preliminary structures established that the N$^\epsilon$ atoms of both histidine residues were coordinating the Zn(II). This was later confirmed using a $^{1}$H-$^{15}$N HMBC experiment optimized to detect $^J$-couplings in histidine side chains (20). In order to include additional distance and angle constraints that maintain the tetrahedral bonding geometry and appropriate bond lengths with the zinc ion (21), as well as to incorporate ambiguous restraints, subsequent structural refinement was performed in CNS (22) using the package ARIA (23, 24). Manually assigned NOEs in combination with the remaining ambiguous NOEs were included in the ARIA structure calculations, and the latter were iteratively assigned in an automated manner.

Briefly, one cycle of 200 structures was followed by seven cycles of 20 structures each and a final cycle of 1000 structures. Manually assigned NOEs were included in iteration zero as soft restraints. In each cycle, the seven lowest energy structures from the previous iteration were used to extract additional NOE restraints with a tolerance of 0.02 ppm in the F1 dimension and 0.015 ppm in the F2 dimension. If a restraint was violated by more than a predefined target value in over 50% of the seven structures, it was discarded. The violation target value was progressively reduced from 1000 Å in iteration zero to 0.1 Å in iteration eight. 18 of 782 distance restraints were discarded in the final iteration due to these violations. Ambiguous distance restraints were treated as described previously (25), and the peak volume cut-off was gradually reduced from 1.01 in the first iteration to 0.80 in the last. The final assignments made by ARIA were checked and corrected manually where necessary. Calculations were carried out in the simplified all-hydrogen PARALLHDP2 force field with nonbonded interactions modeled by the PROLSQ force field (26); floating chirality assignment (27) was used for all methylene and isopropyl groups.

Finally, the 25 lowest energy structures were subjected to water refinement using the standard water refinement protocol supplied by ARIA1.2 (28). The structures were immersed in a 7-Å shell of water molecules and were subjected to a short molecular dynamics simulation taking into account the Lennard-Jones, van der Waals, and electrostatic interactions and based on a slightly modified OPLS force field (29). These water-refined structures were visualized and analyzed using the programs MOLMOL (30), PROCHECK (31), and WHATIF (32). The coordinates of BKLF-F3 have been deposited in the Protein Data Bank under the accession code 1P7A.

**Atomic Absorption Spectrometry—**BF3 and BF3D samples used in far UV CD analyses were dialyzed against sodium acetate buffer (10 mM, pH 5.4) containing 1 mM dithiothreitol, in order to remove excess Zn(II). The resulting samples were rechecked by CD to confirm that they had remained folded. The Zn(II) content of BF3 (28 µM) and BF3D (19 µM) samples were determined on a Varian SpectraAA 10/20 flame emission spectrometer at 213.9 nm. Concentrations were determined by reference to a standard curve constructed using atomic absorption standard zinc(II) solution (1000 mg/liter in 0.5 mol/liter nitric acid; BDH), diluted with Milli-Q® water to final concentrations of 0.01, 0.05, 0.1, 0.1, and 1 mg/liter.

**RESULTS**

**Many Variants of CCHH Zinc Fingers Exist—**A search of the Pfam data base (available on the World Wide Web at pfam...
pared that harbored substitutions of the final histidine ligand of murine BKLF (BKLF-F1–3 residues 254–344) C-terminal to We generated a fusion protein encompassing the three fingers specific to oligonucleotides containing the sequence CACCC (6). B zinc finger region of the transcriptional repressor BKLF (6). At this question, we made a series of single site mutations in the function of the parent protein.

Interestingly, a number of these CCHX variants appear to be located at the C-terminal end of an array of CCHH fingers. Given the dogma that four ligating amino acid side chains are required for the formation of a classical zinc finger structure, the existence of these variants led us to ask whether such sequences can form functional zinc finger domains, or if instead they represent vestigial domains that no longer play a role in the function of the parent protein.

**CCHX Mutants of BKLF Bind to DNA**—In order to address this question, we made a series of single site mutations in the zinc finger region of the transcriptional repressor BKLF (6). At its C terminus, BKLF has three contiguous CCHH fingers (Fig. 1B); this three-finger array binds with high affinity and specificity to oligonucleotides containing the sequence CACCC (6). We generated a fusion protein encompassing the three fingers of murine BKLF (BKLF-F1–3 residues 254–344) C-terminal to glutathione S-transferase (GST). Six mutants were also prepared that harbored substitutions of the final histidine ligand in the third finger, namely aspartate, glutamate, asparagine, glutamine, alanine, and arginine (Fig. 1B). These constructs were termed BF1-3X (where X represents Asp, Glu, Asn, Gln, Ala, and Arg). Aspartate is known to act as a zinc ligand in LIM domains (34), a class of zinc fingers in which two zinc atoms are bound. Glutamate, asparagine, and glutamine were chosen as residues with side chains that could conceivably act as ligands to Zn(II), and alanine and arginine were selected as negative controls. We initially also attempted to make a derivative containing a cysteine, but despite several attempts, we were not successful at generating this protein by normal methods in E. coli, since the protein appears to be cytotoxic; this mutant was therefore not pursued further.

The seven proteins (wild-type BF1-3 and the six mutants) were tested for their ability to bind to a typical CACCC box motif, namely the motif in the β-globin promoter (see “Experimental Procedures”). GST alone and probe alone were also included as negative controls in this experiment (Fig. 2A). Remarkably, all of the point mutants retained near native DNA binding ability. This result was unexpected, given the generally accepted view that four ligating amino acid side chains are required for the formation of a classical zinc finger domain and the low probability that either arginine or alanine could coordinate Zn(II). Three C-terminal truncation mutants were also tested for their DNA binding ability. The mutant with four amino acids truncated at the C terminus of finger 3 (BF1-3Δ–HMLV), including the second zinc-ligating histidine, was able to bind DNA, albeit with a reduced binding affinity (Fig. 2B). In contrast, protein constructs either with 17 amino acids truncated (BF1-2½) or comprising only fingers 1 and 2 (BF1-2) were not able to bind DNA (Fig. 2B).

In order to quantify the effect of the point mutations on DNA binding, quantitative EMSAs were carried out using the wild-type and the alanine mutant BF1-3A (Fig. 2, C–E). The wild-type protein binds a typical CACCC box site with a $K_a$ of (2.5 ± 0.4) × 10^7 M^-1. Surprisingly, the His → Ala mutation had a relatively modest effect on DNA binding; the mutant BF1-3A bound DNA with a $K_a$ of (2.16 ± 0.22) × 10^6 M^-1 (Fig. 2E).

The observation that even the BF1-3A and BF1-3Δ–HMLV) mutants bind DNA is most simply explained by one of two possible models. First, it could be that BF3 folds normally in the context of the mutants, irrespective of whether a fourth
zinc ligand is present at position 341 (using the numbering in Fig. 1B). Second, it is possible that the BF1-3 mutants are either partially folded or completely unfolded but that the formation of structure and DNA binding take place concomitantly. A number of recent reports demonstrate simultaneous folding and binding events (35, 36).

**BKL Finger 3 Is a Typical Classical ZnF**—The first step that we took in order to delineate the effect of the mutations was to determine the structure of the wild-type BF3 using NMR spectroscopy. Resonance assignment from homonuclear NMR spectra was straightforward, and the structure calculations were carried out using ARIA (23). The 20 structures with lowest overall energies (from the final refinement in water) were used to represent the solution structure of BF3 (Fig. 3A). The structures display good covalent geometry, judging from the small deviations from ideal bond lengths and angles, and good nonbonded contacts, as shown by the low value of the mean Lennard-Jones potential (Table I). The atomic coordinates for this family of conformers have been deposited with the Protein Data Bank.

The overall topology of BF3 conforms to the expected fold of classical CCHH zinc finger domains: two short strands of antiparallel β-sheet strands linked by a rubredoxin-like turn are followed by an α-helix that contains the two Zn(II)-coordinating histidine residues. The short β-sheet encompasses residues 319–321 and 326–328, and hydrogen bonds are formed that involve the backbone amide protons of Phe319, Cys321, and Phe328 and the carboxyl oxygens of Phe326, Arg326, and Phe319, respectively. The β-turn linking these two strands has hydrogen bonds and dihedral angles consistent with the rubredoxin-like turn that is commonly found in zinc-binding domains (see, for example, Refs. 37 and 38). A positive φ angle is often found in the residue following the second zinc-ligating Cys, and we confirmed its existence in BF3 using the methods described in Ref. 19. The α-helix runs from residue Ser331 to Leu343, although both $d_{\text{NH}}(i, i + 3)$ and $d_{\text{NN}}(i, i + 4)$ NOEs are observed for residues 339–343, suggesting that the conformation of this region lies somewhere between an α-helix and a 3–10 helix. This phenomenon is also seen in the last 3–4 residues of the α-helix of the third ZnF of Sp1 (39).

Overall, the structure of BF3 is very similar to other classical zinc fingers. For example, it overlays with the second ZnF of Zif268 with a root mean square deviation of 0.82 Å over the ordered backbone atoms (Fig. 3C) (40).

**BF3 Mutants Form Secondary Structure in a Zn(II)-dependent Fashion**—Having established that the BF1-3 point mutants were all capable of DNA binding and that the wild-type BF3 domain formed a normal classical zinc finger fold, we sought to determine whether the mutants were able to bind Zn(II). The six mutants were also produced as single finger GST fusion proteins (termed BF3X). These overexpressed GST-BF3X proteins were subjected to glutathione-affinity chromatography, and thrombin was used to release the BF3X domains. However, all six mutant proteins exhibited partial cleavage by thrombin at secondary sites. Useable amounts of wild-type BF3, as well as mutants BF3D, BF3E, BF3N, and BF3A but not mutants BF3R or BF3Q were isolated by reverse phase HPLC. In order to ascertain whether the mutant domains were able to bind Zn(II), far UV CD spectra were recorded. The spectra of wild-type BF3 and each of the mutants at low pH (pH 2) and in the presence of excess Zn(II) are typical of unstructured polypep-

---

**Table 1**

| Parameters | Values |
|------------|--------|
| Distance restraints | Total NOE restraints 791 |
| | Unambiguous intraresidue ($i - j = 0$) 396 |
| | Unambiguous sequential ($|i - j| = 1$) 164 |
| | Unambiguous medium range ($2 \leq |i - j| \leq 3$) 162 |
| | Unambiguous long range ($|i - j| > 3$) 40 |
| | Ambiguous 29 |
| Dihedral angle restraints | φ 31 |
| Mean root mean square deviations from idealized covalent geometry | Bonds (Å) $0.00430 \pm 0.00016$ |
| | Angles (degrees) $0.49104 \pm 0.01645$ |
| Atomic root mean square differences versus mean (Å) | Backbone atoms (319–342) 0.43 ± 0.12 |
| | Heavy atoms (319–342) 1.11 ± 0.18 |
| Quality control | PROCHECK statistics (319–342) |
| | Residues in most favored region 89.7% |
| | Residues in allowed region 10.1% |
| | Residues in generously allowed regions 0.2% |
| | Residues in disallowed regions 0.0% |
tides. Surprisingly, however, an increase of the pH to 5.5 resulted in noticeable increases in the secondary structure content of all domains (Fig. 4A, black lines). This was manifested as a red shift of the minimum and, in the case of BF3, by the presence of positive ellipticity at low wavelengths. Spectra recorded at pH 5.5 in the absence of Zn(II) were similar to the low pH spectra (data not shown).

To assess more quantitatively whether Zn(II) binding is affected by these point mutations, the Zn(II)-binding affinities of the wild-type BF3 and BF3N mutant were determined. Zn(II) was titrated into solutions of these domains, and the change in ellipticity at wavelengths between 195 and 200 nm was recorded (Fig. 4B). Surprisingly, nonlinear least squares fitting of the titration data revealed that BF3 and BF3N bound Zn(II) with similar affinities \((4.6 \pm 1.4) \times 10^3 \text{ M}^{-1}\) (wild type, gray line) and \((1.1 \pm 0.6) \times 10^5 \text{ M}^{-1}\) (BF3N, black line).

** FIG. 4. CCHX mutants contain substantial secondary structure.** A, far UV CD spectra of both BF3 and BF3X mutants. Reported spectra are the sum of three scans collected at a rate of 20 nm min\(^{-1}\) and using a bandwidth of 1 nm and a resolution time of 1 s (25 °C). All spectra are base line-corrected. Spectra recorded at pH 5.5 are shown as black lines, and the spectrum in gray was collected at pH \sim 2. B, determination of the zinc-binding affinity for BF3 and BF3N from the observed change in ellipticity at 199 nm. The fitted values of \(K_a\) are \((4.6 \pm 1.4) \times 10^3 \text{ M}^{-1}\) (wild type, gray line) and \((1.1 \pm 0.6) \times 10^5 \text{ M}^{-1}\) (BF3N, black line).

** FIG. 5. BF3E is monomeric in solution.** Sedimentation equilibrium data for BF3E recorded at 42,000 rpm (25 °C). The lower graph displays a plot of absorbance at 230 nm versus \(r^2/2\) (cm\(^2\)), whereas the upper graph illustrates the residual deviations resulting from the fit of an ideal single species model to the data.

** FIG. 6. BF3E is probably partially folded.** Amide regions of the one-dimensional \(^1\)H NMR spectra of BF3 and BF3E. A, wild-type BF3 at pH 5.5. B, BF3E at pH 5.5. C, BF3E at pH \sim 2.0. Spectra were recorded at 25 °C on a Bruker DRX600 NMR spectrometer.

** BF3X Mutants Are Partially Folded in Solution—**The one-dimensional \(^1\)H NMR spectrum of wild-type BF3 (Fig. 6A) contained sharp and well dispersed signals, which is indicative of a folded protein with a significant degree of tertiary structure. In contrast, the one-dimensional \(^1\)H NMR spectrum of the BF3E (Fig. 6B), while exhibiting a reasonable amount of chemical shift dispersion compared with the same protein at pH 2 (Fig. 6C), was much broader. Given that BF3E is monomeric in solution (see above), the broadness indicates that BF3E is undergoing substantial chemical exchange on the microsecond–millisecond time scale.
DISCUSSION

Zinc finger domains are abundant in the genomes of eukaryotic organisms. They are known to function largely as specific binding motifs that recognize targets such as DNA, RNA, and other proteins. It has generally been observed that zinc-binding domains form compact, stable structures in which the presence of the zinc ion is essential for the formation of that structure. The functions of these domains are generally thought to be dependent on the existence of stable structure, and all known zinc-binding domains (excluding those in metalloenzymes) ligate the zinc ion through four amino acid side chains.

**Zinc-and DNA-binding Properties of the Mutants—**The transcriptional repressor BRLF contains a tandem array of three classical CCHH zinc fingers that bind with high affinity to CACC-containing DNA sequences. The affinity of this interaction ($K_d = 10^{-7}$ M$^{-1}$) is comparable with those reported for other known three-ZnF constructs binding to their cognate DNA sequence; for example, the three CCHH ZnFs of Sp1 bind DNA with a $K_d$ of $7 \times 10^{-7}$ M$^{-1}$ (41). We have shown here that the third zinc finger of BKF is capable of making a substantial contribution to DNA binding even when the most C-terminal of its four normal zinc ligands (His$^{341}$ in Fig. 1B) is mutated. The DNA binding affinity of the three-finger construct BKF-3 is only reduced 10-fold when the last zinc-ligating His mutated is Ala. Indeed, even a truncated construct lacking any residue at the fourth zinc-ligating position (BF1-3′-HMLV) retains the ability to bind DNA, demonstrating that the DNA binding contributions made by this “broken” finger do not depend on the presence of four zinc ligands. A combination of CD, atomic absorption, UV-visible, and $^1$H NMR data indicates that the point mutants are capable of binding one molar equivalent of Zn(II), thereby forming substantial secondary structure. Surprisingly, the Zn(II)-binding affinity of the single BF3 finger appears not to be affected by the $H \rightarrow N$ mutation. Sedimentation equilibrium data confirm that zinc binding is achieved by a single protein molecule (rather than by, for example, two molecules, each contributing two ligands).

Examination of the sequence of BF3 reveals that there are no other amino acids in the vicinity of His$^{341}$ that could readily substitute for that residue as the fourth zinc ligand. Whereas it is conceivable that His$^{333}$ could serve this role, such an arrangement would require a substantial structural rearrangement, and it is unlikely that the resulting structure would be capable of contributing to DNA binding. Taken together therefore, our data indicate that three ligands are sufficient (although not optimal) for Zn(II) binding in BKF. This conclusion is supported by an assortment of previous reports. For example, a truncated C$_6$H$_2$ZnF (that is missing its last zinc-ligating histidine) was found to coordinate Co(II) in a tetrahedral manner and with an affinity comparable with that of the intact ZnF (42). In a second study, it was observed that when one of the three classical ZnFs of Zf268 was mutated to a CCHA configuration, the protein retained its ability to bind DNA (43). Finally, Cook et al. (44) reported a mutant of the Saccharomyces cerevisiae transcriptional activator ADR1, in which the second histidine of the C-terminal CCHH finger (in a two-ZnF tandem array) was substituted to a tyrosine. The ability of this mutant to activate transcription was reduced but not abrogated, and it was postulated that three zinc chelators might be sufficient to bind Zn(II) and maintain the protein in its active form.

**Conformation of the Mutants—**The enhanced susceptibility of these mutants to proteolysis in *E. coli*, however, suggests that the mutants do not form compact tertiary structures. Both the CD and the $^1$H NMR data support this conclusion. Whereas considerable amounts of secondary structure were formed upon the addition of Zn(II), CD spectra of the mutants were still somewhat different from the wild-type spectrum. The broad nature of the $^1$H NMR spectrum of BF3E indicates the existence of interconverting conformers in a chemical exchange process. Given that the zinc-binding affinity for the BF3 His$\rightarrow$ Asn mutant is not significantly different from that of the wild-type domain, it is likely that the chemical exchange arises from loose packing of the mutants, much like the molten globule state often discussed in the context of protein folding. However, this partial formation of structure is obviously sufficient to allow the recognition of DNA in the context of the three-zinc finger construct BF1-3, and it is possible that the mutated third finger forms more regular structure concomitantly with DNA binding. Alternatively, the NMR broadening may be a consequence of a change in the kinetics of metal binding; distinguishing these possibilities would be problematic.

The Implications for Other Studies—The identification of ZnFs from genomic sequence data relies in part on the presence of predictably spaced cysteine and histidine residues. Whereas a number of variant CCHX zinc fingers are picked up by automated methods, it is possible that others are not, given the apparent plasticity of the requirements for a functional ZnF. Results such as those presented here may assume increasing importance as the use of sequence data alone to infer protein function becomes more prevalent. Thus, our results indicate that care should be taken before presuming that CCHX zinc fingers found in putative proteins are nonfunctional or vestigial.

These data also cast a question on the routine use of alanine substitution to create null ZnFs in functional studies. Because it had been assumed that the structure and, hence, function of a ZnF depends on the ligation of Zn(II) by four ligands, alanine-substituted ZnFs were generally expected to be nonfunctional. This study indicates that, at least in the context of some sequences, alanine substitution mutants may retain significant residual ZnF structure and activity.

The CCHH zinc finger fold is a common scaffold from which proteins with different DNA binding specificities have been generated. It is the simple structure and small number of residues required to structurally stabilize the domains that makes them particularly versatile and adaptable. The high number of CCHX zinc fingers genes in eukaryotic genomes suggests that they may have evolved early in evolution, and an intermediate containing only three zinc ligands that exhibited suboptimal function may have played a role in their evolution. Another interesting possibility is that the presence of an amino acid other than histidine and cysteine as a zinc ligand or the absence of one ligand may constitute a means by which ZnF proteins could be regulated. Whereas the CCHX ZnFs may still recognize DNA, their greater susceptibility to proteolytic degradation might reduce their cellular half-life. This idea might also be pertinent in the design of novel ZnFs as a method in which the bioavailability of designed ZnFs might be controlled.

Acknowledgment—We thank Dr. Bill Bubb for expert maintenance of the DRX600 NMR spectrometer at the University of Sydney and for valuable advice.

REFERENCES

1. Dawid, I. B., Breen, J. J., and Toyama, R. (1998) Trends Genet. 14, 156–162
2. Luscombe, N. N., Austin, S. E., Berman, H. M., and Thornton, J. M. (2000) Gen. Biol. 1, 001.001–001.037
3. Mackay, J. P., and Crossley, M. (1998) Trends Biochem. Sci. 265, 1–4
4. Wolfe, S. A., Grant, R. A., Elrod-Erickson, M., and Pabo, C. O. (2001) Structure 9, 717–723
5. Pabo, O. C., Pfeilschifter, J., and Grant, R. A. (2001) Annu. Rev. Biochem. 70, 313–349
6. Crossley, M., Whiteclaw, E., Perkins, A., Williams, G., Fujiwara, Y., and Orkin, S. H. (1996) Mol. Cell. Biol. 16, 1695–1705
7. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., pp. 10.59–10.61, Cold Spring Harbor Labo-
CCHX Zinc Fingers Can Bind DNA

8. Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) Biophys. J. 36, 575–588
9. Perkins, S. J. (1986) J. Biol. Chem. 261, 169–180
10. Hayes, D. B., Laue, T., and Philo, J. (1995) SEDNTERP, 1.05, University of New Hampshire, Durham, NH
11. Bax, A., and Davis, D. G. (1985) J. Magn. Reson. 65, 355–360
12. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wuthrich, K. (1985) Biochem. Biophys. Res. Commun. 157, 169–180
13. Kumar, A., Ernst, R. R., and Wuthrich, K. (1980) Biochem. Biophys. Res. Commun. 95, 1–6
14. Vuister, G., and Bax, A. (1993) J. Am. Chem. Soc. 115, 7772–7777
15. Marion, D., Driscoll, P. C., Kay, L. E., Wingfield, P. T., Bax, A., Gronenborn, A. M., and Clore, G. M. (1995) Proteins 1, 1–6
16. Kowalski, K., Czolij, R., King, G. F., Crossley, M., and Mackay, J. P. (1999) J. Biomol. NMR 13, 249–261
17. Wuthrich, K. (1986) NMR of Proteins and Nucleic Acids, Wiley Interscience, New York
18. Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) J. Mol. Biol. 273, 203–208
19. Ludovigsen, S., and Poulsen, F. M. (1992) J. Biol. Chem. 267, 227–233
20. Pelton, J. G., Torchia, D. A., Meadow, N. D., and Roseman, S. (1993) Protein Sci. 2, 543–558
21. Neuhaus, D., Nakaseko, Y., Schwabe, J. W., and Klug, A. (1992) J. Mol. Biol. 228, 637–651
22. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Acta Crystallogr. D Biol. Crystallogr. 54, 905–921
23. Nilges, M. (1995) J. Mol. Biol. 245, 645–660
24. Nilges, M., Macias, M. J., O’Donoghue, S. I., and Oshkinat, H. (1997) J. Mol. Biol. 269, 408–422
25. Nilges, M. (1997) Protein Sci. 6, 589–593
26. Hendrickson, W. A. (1985) Methods Enzymol. 15, 252–270
27. Felmer, R. H., Hilbers, C. W., Kunings, R. N., and Nilges, M. (1997) J. Biomol. NMR 9, 245–258
28. Linge, J. P., Williams, M. A., Spronk, C. A., Bonvin, A. M., and Nilges, M. (2003) Proteins 50, 496–506
29. Nilges, M., and Nilges, M. (1999) J. Biomol. NMR 13, 51–55
30. Koradi, R., Billeter, M., and Wuthrich, K. (1998) J. Mol. Graph. 14, 51–55
31. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) J. Biomol. NMR 8, 477–486
32. Vriend, G., and Sandor, C. (1995) J. Appl. Crystallogr. 28, 47–60
33. Liew, C. K., Kowalski, K., Fox, A. H., Newton, A., Sharpe, B. K., Crossley, M., and Mackay, J. P. (2000) Structure 8, 1157–1166
34. Bach, I. (2000) Mechanisms of Development 91, 5–17
35. Demarest, S. J., Martinez-Yamout, M., Chang, J., Chen, H., Xu, W., Dyson, H. J., Evans, R. M., and Wright, P. E. (2002) Nature 415, 549–553
36. Radakrishnan, I., Perez-Alvarado, G. C., Parker, D., Dyson, H. J., Montminy, M. R., and Wright, P. E. (1997) Cell 91, 741–752
37. Schwabe, J., and Klug, A. (1994) Nat. Struct. Biol. 1, 345–349
38. Summers, M. F., South, T. L., Kim, B., Hare, D. R. (1990) Biochemistry 29, 329–340
39. Narayan, V. A., Kriwacki, R. W., and Caradonna, J. P. (1997) J. Biol. Chem. 272, 7801–7809
40. Wang, B., Grant, R., and Pabo, C. (2001) Nat. Struct. Biol. 8, 589–593
41. Nagaoka, M., Kondo, Y., Uno, Y., and Sugiyama, Y. (2002) Biochem. Biophys. Res. Commun. 296, 553–559
42. Merkle, D. L., Schmidt, M. H., and Berg, J. M. (1991) J. Am. Chem. Soc. 113, 5439–5451
43. Green, A., and Sarkar, B. (1998) Biochem. J. 333, 85–90
44. Cook, W. J., Mosley, S. P., Audino, D. C., Mullaney, D. L., Rovelli, A., Stewart, G., and Deniz, C. L. (1994) J. Mol. Biol. 229, 9374–9379
45. Pavletich, N. P., and Pabo, C. O. (1991) Science 252, 809–817
CCHX Zinc Finger Derivatives Retain the Ability to Bind Zn(II) and Mediate Protein-DNA Interactions
Raina J. Y. Simpson, Edward D. Cram, Robert Czolij, Jacqueline M. Matthews, Merlin Crossley and Joel P. Mackay

J. Biol. Chem. 2003, 278:28011-28018.
doi: 10.1074/jbc.M211146200 originally published online May 7, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M211146200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 40 references, 4 of which can be accessed free at
http://www.jbc.org/content/278/30/28011.full.html#ref-list-1