Design of a Monomeric 23-Residue Polypeptide with Defined Tertiary Structure

Mary D. Struthers, Richard P. Cheng, Barbara Imperiali*

Small proteins or protein domains generally require disulfide bridges or metal sites for their stabilization. Here it is shown that the ββα architecture of zinc fingers can be reproduced in a 23-residue polypeptide in the absence of metal ions. The sequence was obtained through an iterative design process. A key feature of the final design is the incorporation of a type II′ β turn to aid in β-hairpin formation. Nuclear magnetic resonance analysis reveals that the α helix and β hairpin are held together by a defined hydrophobic core. The availability of this structural template has implications for the development of functional polypeptides.

The design of stable, folded, and structurally characterized polypeptide motifs has been the focus of research in a number of laboratories over the past decade (*1*, *2*). A key question in protein design is the minimum size of a polypeptide required to achieve the compact, folded, and organized architecture that is the hallmark of natural proteins (*3*). Iterative modeling, synthesis, and structural characterization are facilitated for smaller motifs, allowing for the efficient testing of design principles and strategies. Although protein design has met with significant progress in recent years (*4*, *5*), most designed motifs are either relatively large (>60 residues) (*6*) or multimeric (*2*, *7*–*9*), and many require either disulfide bridges (*7*, *8*, *10*) or metal ligation (*11*, *12*) for stability. The size and aggregation state of some of these motifs have complicated detailed structural analysis. Additionally, the use of disulfides and metal cation ligation may limit the versatility of these constructs for future elaboration into functional polypeptides.

Our goal has been the construction of small (<30 residues), soluble, monomeric polypeptide motifs that fold in the absence of cross-linking sites such as disulfides or metal binding centers. The ββα motif, exemplified by the zinc finger domains, represents an ideal design target. Natural zinc finger peptides are small, independent folding units that incorporate conserved hydrophobic residues (typically Tyr, Phe, and Leu) (*13*, *14*) and adopt defined solution structures only in the presence of specific metal cations (*15*, *16*). With an iterative design and analysis procedure, sequences based on these domains and retaining the key hydrophobic cluster were evaluated for their ability to fold in the absence of metal. The circular dichroism (CD) signature of the zinc finger domains is well known (*16*, *17*); therefore, the secondary structural content of early efforts could be rapidly evaluated. The design process and spectroscopic analysis has resulted in the successful generation of a metal-independent folded structure, BBA1, with as few as 23 amino acid residues.

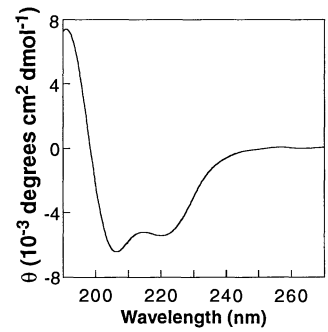

Fig. 2. The CD spectrum of BBA1 (θ, mean residue ellipticity) (686 μM peptide, 10 mM acetate buffer, pH 4.5, 7°C). We determined the peptide concentration spectroscopically using the phenanthroline chromophore (extinction coefficient at 268 nm, ε_{268} = 13,950 M^{-1} cm^{-1}).

The BBA1 motif resulted from a detailed study of five peptide sequences (*18*, *19*). In the first generation of the design process (**1**, Fig. 1), the ligand sphere of the native zinc finger Zif268 (*20*) (His_2Cys_2) was modified to include the unnatural metal-binding amino acid 3-(1,10-phenanthrol-2-yl)-L-alanine (Fen) as a reporter group. Metal binding by the construct was therefore anticipated to occur through coordination to Fen^6, His^3, and His^{21}. Additionally, the loop connecting the β strands of the hairpin was replaced with residues designed to adopt a type II β turn (*21*). The second-generation design (**2**, Fig. 1) included amino acid substitutions intended to increase the inherent secondary structure of the polypeptide (*22*). Spectroscopic analyses of

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, USA.

*To whom correspondence should be addressed.

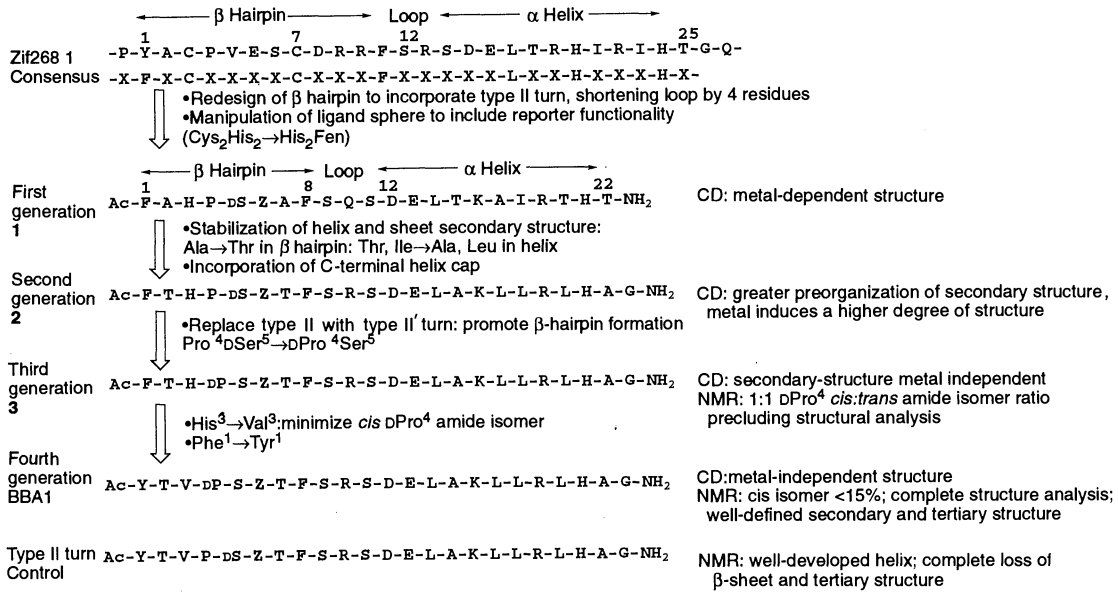

Fig. 1. Sequences and structural analyses of peptides leading to the design of BBA1. The BBA1 motif is based on the zinc finger consensus sequence (*13*) and the sequence of Zif 268, a DNA-binding protein consisting of three zinc finger domains, which has been cocrystallized with DNA (*20*). The conserved hydrophobic core residues include two aromatic residues in the hairpin (Phe or Tyr) and a hydrophobic residue in the helix (usually Leu) (*13*). Abbreviations for the amino acid residues are A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; Y, Tyr; and Z, Fen.

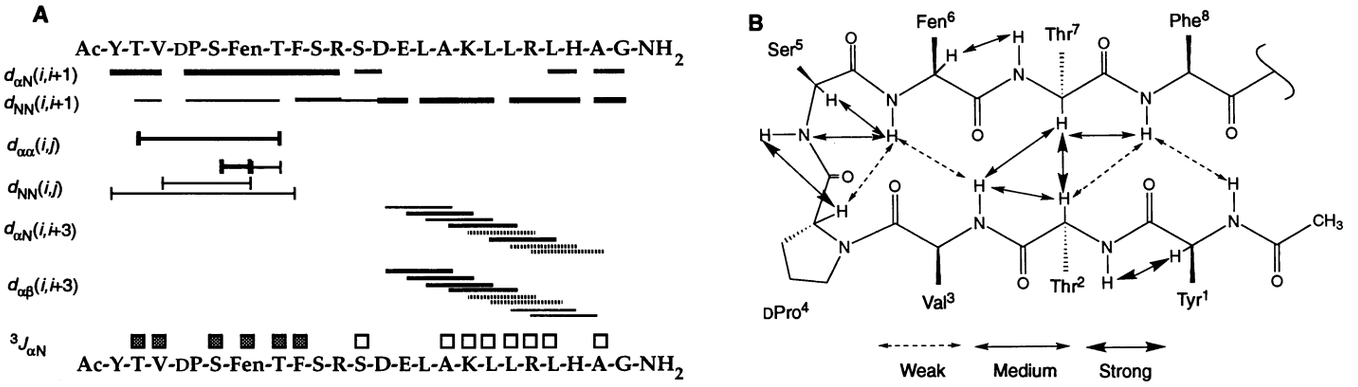

Fig. 3. (**A**) NOEs indicative of secondary structure. The width of the lines represent the relative strength of the NOEs; a dashed line indicates that the NOE could not be identified because of spectral overlap. Filled squares indicate $^3J_{\alpha N} > 8$ Hz and open squares indicate $^3J_{\alpha N} < 6.8$ Hz. (**B**) Schematic representation of key NOEs defining the β hairpin. The intensities of the NOEs were determined from measurement of cross-peak volumes and correspond to the following approximate distance ranges: weak (1.8 to 5.5 Å), medium (1.8 to 3.5 Å), and strong (1.0 to 2.5 Å).

these designs revealed that the adoption of secondary structure was still metal-dependent. In the third-generation design (3, Fig. 1), the type II β turn was replaced with a type II′ β turn, on the basis of a survey of the protein database that suggested the latter turn might more effectively promote β-hairpin formation (*23*). This replacement resulted in a metal-independent CD signature; however, nuclear magnetic resonance (NMR) analysis indicated the presence of a 1:1 ratio of *cis*:*trans* DPro[4] isomers, and therefore two distinct global conformations. Substitution of the residue immediately preceding DPro[4] (*24*) reduced this ratio to an acceptable level in the final design BBA1 (>85% trans isomer). The CD spectrum of BBA1 in the absence of metal cations strongly resembles that of completely folded, natural zinc finger peptides (Fig. 2) (*16*, *17*). Metal cation titrations monitored by CD confirmed that for this optimized design, metal binding is not a prerequisite for folding, and Fen-mediated metal binding does not result in an increase in secondary structure.

The BBA1 peptide is extremely water soluble (>5 mM) and is monomeric in solution (*25*), allowing complete structural analysis by NMR techniques (*26*). Analysis of the NMR spectra indicates that BBA1 has the desired secondary structure. Cross-strand nuclear Overhauser effects (NOEs) and large coupling constants between the alpha and amide protons ($^3J_{\alpha N}$) indicative of β-hairpin formation were observed for residues 1 through 8 (Fig. 3). Residues 13 to

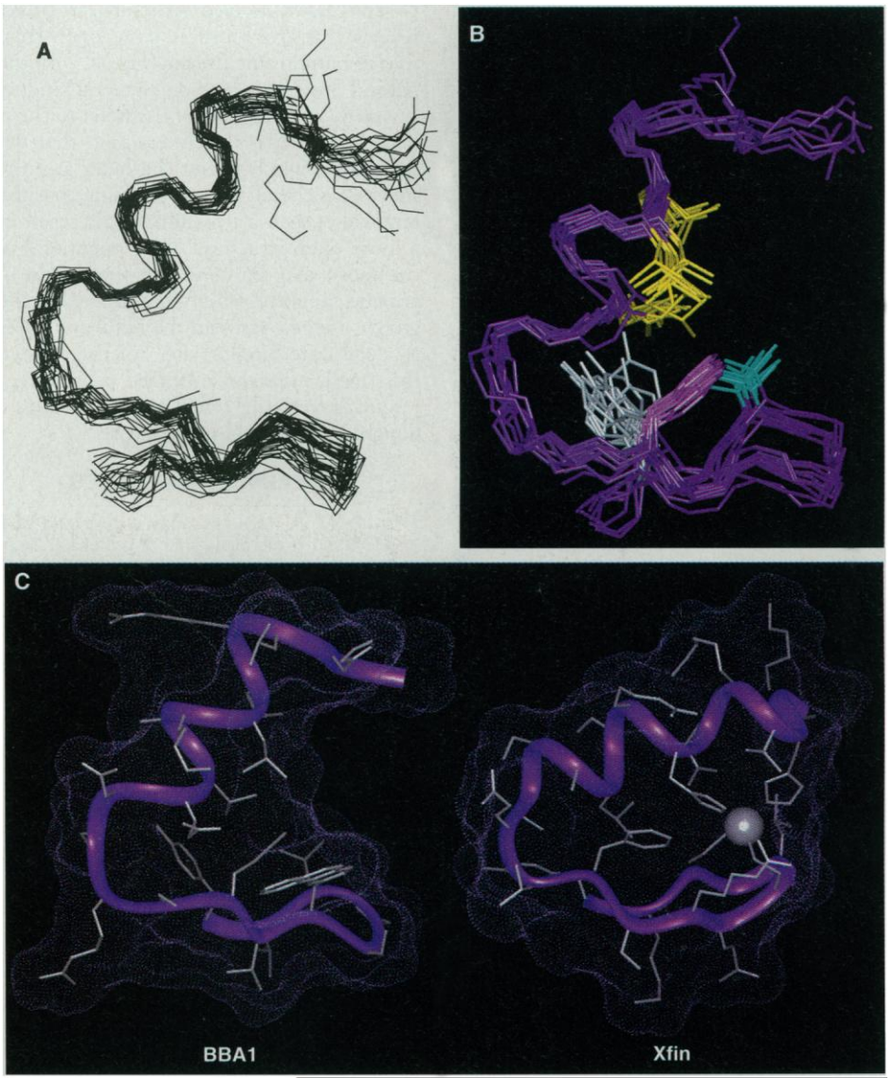

Fig. 4. (**A**) Superimposition of 29 structures of BBA1 obtained from simulated-annealing molecular dynamics using NOE-derived restraints (*29*). The backbone traces (N,Cα,C atoms) of all structures are shown. (**B**) Superimposition of the backbone traces (N,Cα,C atoms) of the 10 structures with the lowest distance and dihedral restraint energies, showing the side chains that form the hydrophobic core. (**C**) Solvent-accessible surface of a representative NMR structure of BBA1. The backbone is represented by a ribbon diagram. A similar representation from the NMR structure of a single natural zinc finger peptide (X-fin, finger 31) (*31*) is shown for comparison.

22 adopt an α-helical conformation, as demonstrated by the sequential short-range $d_{NN}(i, i + 1)$, long-range $d_{\alpha\beta}(i, i + 3)$, and $d_{\alpha N}(i, i + 3)$ NOEs and smaller values of the $^3J_{\alpha N}$ coupling constants (Fig. 3). Amide proton H-D exchange experiments at pD* 4.5 (uncorrected) are also consistent with the desired β and α secondary structure. These one-dimensional (1D) NMR experiments indicated that many amide protons of BBA1 were protected from solvent exchange (*27*). Although some protons exchanged rapidly (~6 min), several amide proton signals were still visible after 1 hour. Although not all of the protected protons could be assigned in the 1D spectra, at least one amide proton in the β sheet (Val^3), as well as several in the α helix (Leu^{17}, Leu^{18}, and Ala^{15}), exhibited slow exchange, in agreement with the predicted structure. Most importantly, a remarkable number of long-range NOEs (NOEs between spin systems separated by at least three residues in the primary sequence) were observed in the hydrophobic cluster of the motif (Tyr^1, Phe^8, Leu^{14}) (*28*), indicating the presence of a defined tertiary structure.

The 3D structure of BBA1 was determined by an NOE-restrained simulated-annealing molecular-dynamics protocol with a total of 135 intraresidue and 166 interresidue distance restraints derived from 367 NOEs (*29*). A family of 45 structures was generated, 29 of which exhibited a common fold. This subset of structures had the lowest energies and the least NOE restraint violations and was therefore selected for further analysis. These 29 structures agree well with the intended design (Fig. 4). The average root-mean-square deviation (RMSD) of the coordinates of the backbone atoms for residues 1 through 22 for these 29 structures was 0.90 ± 0.30 Å from the average structure; none of the structures had NOE restraint violations greater than 0.45 Å (*30*). The quality of the BBA1 structure compares well with that of Xfin (*31*), a natural zinc finger, under similar NMR acquisition conditions: BBA1 (6.9 mM, 7°C, pH 4.5, 29 structures) RMSD of 1.51 Å for all heavy atoms, 0.90 Å for the backbone atoms of residues 1 through 22; Xfin (5.8 mM, 5°C, pH 5.5, 1.5 equivalents of $ZnCl_2$, 37 structures) RMSD of 1.81 Å for side chain heavy atoms, 0.81 Å for backbone atoms. The α helix and the β hairpin are well defined, and a hydrophobic core is formed through interactions between the side chains of residues Tyr^1, Val^3, Phe^8, Leu^{14}, Leu^{17}, and Leu^{18}. The three residues Tyr^1, Phe^8, and Leu^{14} are involved in the majority of long-range NOE contacts between the sheet and the helix. The formation of a discrete hydrophobic core within BBA1 was further established by analysis of 8-anilino-1-naphthalenesulfonic acid (ANS) binding (*32*, *33*). Our current understanding of the 3D structure of this motif and its amenability to efficient iterative design and analysis should allow us to specifically investigate the packing of side chains within the core.

In natural zinc finger peptides, the conserved hydrophobic core residues are unable to drive folding without the assistance of metal coordination (*15*, *16*). In BBA1, the formation of the hydrophobic core was assisted by the inclusion of a heterochiral type II′ turn as a structural nucleation element (*34*), together with a judicious selection of residues to enhance the appropriate secondary structure throughout the peptide. The type II′ β turn is well defined in the NMR structure. The importance of this turn in the establishment of a folded motif was emphasized in the study of a control peptide that incorporated a type II rather than a type II′ reverse turn (Fig. 1). This peptide differed only in the chirality of the central two residues of the turn and yet had neither the β hairpin nor the tertiary structure observed for BBA1.

The NMR-derived structure of BBA1 reveals that the motif is more open than that of natural zinc fingers (Fig. 4C). In the natural zinc finger peptides, a metal binding site cross-links the end of the helix to the β hairpin, and it is this interaction that defines the angle between the helix and the sheet. It is therefore not surprising that the removal of such a cross-link would result in a more open structure. The "openness" of the motif may be of use in the design of enzyme mimics, which could utilize the groove created between the helix and sheet as a substrate binding site. For this reason, this small, monomeric, folded peptide represents a versatile template for the future design of functional polypeptides.

REFERENCES AND NOTES

1. L. Regan and W. F. DeGrado, *Science* **241**, 976 (1988); M. H. Hecht, J. S. Richardson, D. C. Richardson, R. C. Ogden, *ibid.* **249**, 884 (1990).
2. J. J. Osterhout *et al.*, *J. Am. Chem. Soc.* **114**, 331 (1992).
3. D. S. Goodsell and A. J. Olson, *Trends Biochem. Sci.* **18**, 65 (1993).
4. Y. Fezoui, D. L. Weaver, J. J. Osterhout, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3675 (1994); S. F. Betz, D. P. Raleigh, W. F. DeGrado, *Curr. Opin. Struct. Biol.* **3**, 601 (1993).
5. D. P. Raleigh, S. F. Betz, W. F. DeGrado, *J. Am. Chem. Soc.* **117**, 7558 (1995).
6. D. D. S. Smith, K. A. Pratt, I. G. Sumner, C. M. Henneke, *Protein Eng.* **8**, 13 (1995); A. Pessi *et al.*, *Nature* **362**, 367 (1993).
7. T. P. Quinn, N. B. Tweedy, R. W. Williams, J. S. Richardson, D. C. Richardson, *Proc. Natl. Acad. Sci. U.S.A.* **91**, 8747 (1994).
8. Y. B. Yan and B. W. Erickson, *Protein Sci.* **3**, 1069 (1994).
9. B. Lovejoy *et al.*, *Science* **259**, 1288 (1993).
10. Y. Kuroda, *Protein Eng.* **8**, 97 (1995).
11. T. M. Handel, S. A. Williams, W. F. DeGrado, *Science* **261**, 879 (1993).
12. M. Lieberman and T. Sasaki, *J. Am. Chem. Soc.* **113**, 1470 (1991); M. R. Ghadiri, C. Soares, C. Choi, *ibid.* **114**, 4000 (1992).
13. J. M. Berg, *Annu. Rev. Biophys. Biophys. Chem.* **19**, 405 (1990).
14. S. F. Michael, V. J. Kilfoil, M. H. Schmidt, B. T. Amann, J. M. Berg, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 4796 (1992).
15. P. S. Eis and J. R. Lakowicz, *Biochemistry* **32**, 7981 (1993); A. D. Frankel, J. M. Berg, C. O. Pabo, *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4841 (1987).
16. G. Párraga *et al.*, *Science* **241**, 1489 (1988).
17. M. A. Weiss, K. A. Mason, C. E. Dahl, H. T. Keutmann, *Biochemistry* **29**, 5660 (1990).
18. The peptides were synthesized by 9-fluorenylmethyloxycarbonyl chemistry on a Milligen 9050 peptide synthesizer and purified by semipreparative reversed-phase high-performance liquid chromatography. The identity of each peptide was confirmed by matrix-assisted laser desorption or plasma desorption mass spectroscopy. The synthesis of the Fen residue has been presented [S. L. Fisher and B. Imperiali, paper presented at the 33rd National Organic Symposium, Bozeman, MT, 13 to 17 June 1993].
19. M. S. Struthers, R. P. Cheng, B. Imperiali, unpublished results.
20. N. P. Pavletich and C. O. Pabo, *Science* **252**, 809 (1991).
21. B. Imperiali, S. L. Fisher, R. A. Moats, T. J. Prins, *J. Am. Chem. Soc.* **114**, 3182 (1992).
22. G. D. Fasman, in *Prediction of Protein Structure and the Principles of Protein Conformation*, G. D. Fasman, Ed. (Plenum, New York, 1989), pp. 193–316.
23. B. L. Sibanda and J. M. Thornton, *Nature* **316**, 170 (1985).
24. C. Grathwohl and K. Wüthrich, *Biopolymers* **15**, 2025 (1976).
25. We assessed the aggregation state of BBA1 by monitoring the CD spectrum over a range of peptide concentrations. The CD signature was invariant between 10 μM and 5 mM, indicating that no aggregation occurred. A closely related peptide, differing only in the identity of the residue at position 1 ($Tyr^1 \rightarrow Phe^1$), was determined to be predominantly monomeric in equilibrium centrifugation experiments.
26. The solution structure of BBA1 was obtained from 2D ^1H NMR data (6.9 mM peptide, pH 4.5 in 9:1 H_2O:D_2O, 7°C, or 5.1 mM peptide, pD* 4.5 uncorrected in D_2O, 7°C). Complete sequence-specific assignments of backbone and side-chain protons were obtained by total correlation spectroscopy [A. Bax and D. G. Davis, *J. Magn. Reson.* **65**, 355 (1985)], double-quantum filtered correlation spectroscopy (DQFCOSY) [U. Piantini, O. W. Sørensen, R. R. Ernst, *J. Am. Chem. Soc.* **104**, 6800 (1982)], and NOE spectroscopy (NOESY) [D. J. States, R. A. Haberkorn, D. J. Ruben, *J. Magn. Reson.* **48**, 286 (1982); (*35*)]. All spectra were acquired on a Varian 600-MHz instrument. Saturation of the water signal was accomplished during the relaxation delay. The $^3J_{\alpha N}$ coupling constants were obtained from cross-peak measurements in a DQFCOSY experiment at 25°C to avoid the line broadening observed at 7°C. The structure of BBA1 was stable with respect to changes in temperature. Denaturation studies indicate that the loss of secondary structure (as assessed by CD) between 7°C and 20°C is at most 13% and more likely less than 5%. The peptide does not fully unfold at temperatures as high as 80°C, and complete denaturation requires urea concentrations in excess of 8 M. This behavior parallels that of native zinc fingers [M. A. Weiss and H. T. Keutmann, *Biochemistry* **29**, 9808 (1990)].
27. Peptide samples were freshly dissolved in D_2O (3.2 mM BBA1, 50 mM acetate-d_4, pD* 4.5, uncorrected). The H-D exchange was investigated with a 1D water-gate pulse sequence on a Bruker AMX 500-MHz instrument [M. Piotto, V. Saudek, V. Sklenár, *J. Biomol. NMR* **2**, 661 (1992)].
28. The NOEs were obtained from NOESY spectra of BBA1 taken in both H_2O (6.9 mM peptide, pH 4.5 in 9:1 H_2O:D_2O, 7°C) and D_2O (5.1 mM peptide, pD* 4.5 uncorrected, 7°C). NOEs that indicate specific interaction between the helix and the sheet are ($Leu^{14}\delta_1$,$Tyr^1\delta$), ($Leu^{14}\delta_2$,$Tyr^1\varepsilon$), ($Leu^{14}\delta_2$,$Tyr^1\delta$), ($Phe^8\beta$,$Leu^{14}\delta_2$), ($Leu^{14}\delta_1$,$Phe^8\beta$), ($Phe^8\varepsilon$,$Leu^{14}\delta_2$), ($Phe^8\varepsilon$,$Leu^{14}\alpha$), ($Leu^{14}\delta_1$,$Phe^8\delta$), ($Leu^{14}\delta_2$,$Phe^8\delta$), ($Leu^{14}\alpha$,$Phe^8\zeta$), ($Phe^8\delta$,$Leu^{14}\alpha$), ($Leu^{17}\delta_1$,$Phe^8\zeta$),

(Leu$^{17}\delta_1$,Phe$^{8}\varepsilon$), (Leu$^{14}\delta_2$,Phe$^{8}\zeta$), (Val$^{3}\gamma$,Leu$^{18}\delta_2$), (Phe$^{8}\varepsilon$,Glu$^{13}\gamma$), (Phe$^{8}\delta$,Glu$^{13}\gamma$), (Ala$^{15}\beta$,Tyr$^{1}\varepsilon$), (Phe$^{8}\alpha$,Glu$^{13}\beta$), (Tyr$^{1}\delta$,Ser$^{9}\alpha$), and (Phe$^{8}\delta$,Glu$^{13}\beta$).

29. A simulated annealing (SA) protocol using NMRchitect (Biosym Technologies, San Diego, CA) was used to generate the ensemble of conformations consistent with the NMR data. The SA protocol was derived from that developed by M. Nilges, G. M. Clore, and A. M. Gronenborn [*FEBS Lett.* **239**, 129 (1988)] and was optimized for this system. Each NOE was classified as strong, medium, weak, or very weak on the basis of a linear fit of the measured volumes from NOESY spectra with mixing times of either 50, 100, and 200 ms (H_2O spectra) or 100 and 200 ms (D_2O spectra). The volumes were fit to a linear relation with mixing time and calibrated to volume measurements for cross peaks corresponding to known interproton distances. Most distance restraints were derived from NOE cross-peak volume measurements of NOESY spectra taken in H_2O. A few weak NOEs between side-chain protons were more readily analyzed from data collected in D_2O because of both the increased sensitivity of data collection and the reduced complexity of the spectra. Each NOE category was assigned a distance range: strong (1.0 to 2.5 Å), medium (1.8 to 3.5 Å), weak (1.8 to 5.0 Å), and very weak (1.8 to 5.5 Å). A maximum force constant of 25 kcal mol^{-1} Å^{-2} was used for all NOE distance restraints. Pseudoatoms were used when necessary, and standard corrections were applied to interproton distances involving pseudoatoms (*35*). The use of pseudoatoms resulted in the generation of 301 restraints from 367 NOE volume measurements. An alternate conformation accounting for less than 15% of the total signal was visible in the NMR spectra and assigned through NOE analysis to the cis conformation of DPro^4. Cross peaks from the NOESY attributable to this conformation were omitted for the generation of restraints. Dihedral restraints were included in the SA analysis to enforce trans geometry of the amide bond between Val^3 and DPro^4.
30. The C-terminal glycine was poorly defined by the NMR data and was therefore excluded from the RMSD calculation to yield values of 0.90 Å for backbone atoms and 1.37 Å for the heavy atoms of residues 1 through 22. When Gly^{23} was included in this calculation, the RMSD of the coordinates for all residues from the average structure was 1.28 Å for backbone atoms and 1.51 Å for all heavy atoms.
31. M. S. Lee, G. P. Gippert, K. V. Soman, D. A. Case, P. E. Wright, *Science* **245**, 635 (1989).
32. G. V. Semisotnov *et al.*, *Biopolymers* **31**, 119 (1991).
33. The presence of discrete tertiary structure is often probed by studying the interaction of this hydrophobic, fluorescent dye with proteins. Although ANS will bind to the apolar regions of partially folded or "molten globule" states of proteins, native folded proteins do not bind ANS. Many designed proteins to date have demonstrated significant ANS binding, which is attributed to poor tertiary packing interaction and the adoption of "molten globule" rather than native-like states (*7, 11*). BBA1 did not enhance the fluorescence of ANS, indicating that this designed peptide does not bind ANS under conditions (*5*) previously shown to establish the presence of these poorly folded states. Experiments were performed in 20 mM acetate buffer, pH 4.5, and 30 μM BBA1 with ANS concentrations ranging from 20 to 600 μM. The results indicate that the apolar side chains of BBA1 are sufficiently buried in the folded structure to preclude interaction with the dye.
34. D. B. Wetlaufer, *Trends Biochem. Sci.* **15**, 414 (1990).
35. K. Wüthrich, *NMR of Proteins and Nucleic Acids* (Wiley, New York, 1986).
36. This work was supported by NSF grant CHE-910445, a Parsons Foundation predoctoral fellowship (M.D.S.), an NIH traineeship (R.P.C.), and a Camille and Henry Dreyfus Teacher Scholar Award (B.I.). We gratefully acknowledge the Dorothy Chandler, Camilla Chandler Frost Laboratory in Biology at the California Institute of Technology for use of the Varian Unity Plus 600 NMR Spectrometer.

17 August 1995; accepted 13 November 1995

Assembly of a Ribonucleoprotein Catalyst by Tertiary Structure Capture

Kevin M. Weeks and Thomas R. Cech

CBP2 is an RNA tertiary structure binding protein required for efficient splicing of a yeast mitochondrial group I intron. CBP2 must wait for folding of the two RNA domains that make up the catalytic core before it can bind. In a subsequent step, association of the 5′ domain of the RNA is stabilized by additional interactions with the protein. Thus, CBP2 functions primarily to capture otherwise transient RNA tertiary structures. This simple one-RNA, one-protein system has revealed how the kinetic pathway of RNA folding can direct the assembly of a specific ribonucleoprotein complex. There are parallels to steps in the formation of a much more complex ribonucleoprotein, the 30S ribosomal subunit.

Several processes essential for gene expression are carried out by ancient machinery whose function requires the interplay of RNA and protein components. These ribonucleoprotein (RNP) enzymes include ribonuclease P, the spliceosome, and the ribosome (*1*). Although much is known about the order of assembly of complex RNPs from their RNA and protein components, the kinetic pathways of assembly are poorly understood.

Excision of the group I intron bI5, the fifth intron in the cytochrome b pre–messenger RNA in yeast mitochondria, is carried out by a simple RNP composed of the RNA intron and the splicing factor CBP2 (*2*). The intron RNA consists of three domains of ~100 nt each (Fig. 1A) (*3*). Under near-physiological Mg^{2+} concentrations (7 mM), the RNA is in a state (termed 2°) in which the secondary structure is formed but higher order structure is largely absent. Studies of the RNA at equilibrium (*3, 4*) have revealed a pathway for formation of the active tertiary structure in the absence of CBP2. First, the P5-P4-P6 and P7-P3-P8 domains associate to form the catalytic core (termed the core state). The 5′ domain, which includes P1, the site of guanosine addition in the first step of splicing, then associates with the core to form the active, assembled ribozyme (E):

$$\text{RNA } 2° \rightleftarrows \text{core} \rightleftarrows \text{E} \qquad (1)$$

Mg^{2+} promotes formation of the core state in a cooperative transition that is complete at 40 mM divalent ion; however, Mg^{2+} does not efficiently promote formation of the E state. Thus, in experiments at either 7 or 40 mM Mg^{2+}, the RNA can be forced to be predominantly in the 2° or core state, respectively.

CBP2 is an RNA tertiary structure binding protein that binds preferentially to the folded RNA rather than to isolated elements of the secondary structure. Upon formation of a complex with CBP2, the RNA is in the reactive E state. Thus, whereas the RNA component contains the active site for catalysis of splicing (*3–5*), the protein enhances the rate of splicing 1000 times by holding the RNA in its active conformation. CBP2 binds at a complex site on one face of the folded RNA (Fig. 1B).

That CBP2 recognizes the tertiary structure of the intron presents a mechanistic dilemma. In most encounters between the

Department of Chemistry and Biochemistry, Howard Hughes Medical Institute, University of Colorado, Boulder, CO 80309–0215, USA.

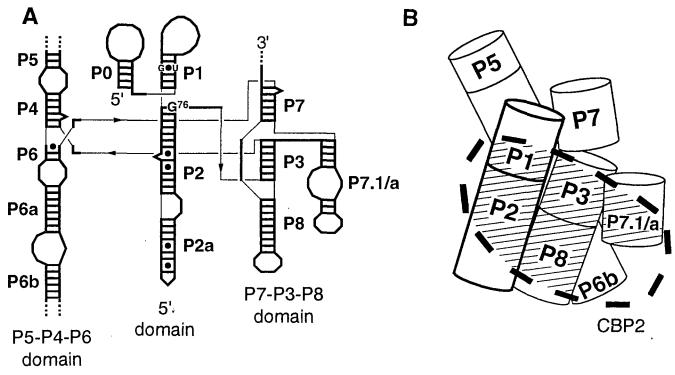

Fig. 1. Domain structure and CBP2 interaction site of bI5 RNA. (**A**) The secondary structure and connections between secondary structure elements are shown as heavy and light lines, respectively. The G·U base pair at the 5′ splice site is shown explicitly. For simplicity, some peripheral structures are omitted (dashed lines). The RNA precursor used in these experiments spans 458 nt including 5′ and 3′ exons 35 and 55 nt in length, respectively (*4*). (**B**) Schematic of bI5 RNA tertiary structure (*25*). Helices shown in A are represented as cylinders, except P0 is omitted. The CBP2 contact site on the RNA as inferred from hydroxyl radical footprinting (*3*) is shown by cross-hatching.