J. Biol. Chem. 265, 3750.
Van Eyk, J. E., & Hodges, R. S. (1988) J. Biol. Chem. 263, 1726.
Watterson, D. M., Harrelson, W. G., Jr., Keller, P. M., Sarif, F., & Vannaman, T. C. (1976) J. Biol. Chem. 251, 4501.
Wignall, G. D., & Bates, F. S. (1987) J. Appl. Crystallogr. 20, 28.
Wu, C.-F., & Chen, S.-H. (1988) Biopolymers 27, 1065.
Yoshino, H., Minari, O., Matsushima, N., Ueki, T., Miyake, Y., Matsuo, T., & Izumi, Y. (1989) J. Biol. Chem. 264, 19706.
Zaccai, G., & Jacrot, B. (1983) Annu. Rev. Biophys. 12, 139.

High-Resolution Three-Dimensional Structure of a Single Zinc Finger from a Human Enhancer Binding Protein in Solution†

James G. Omichinski,‡ G. Marius Clore,*,‡ Ettore Appella,§ Kazuyasu Sakaguchi,§ and Angela M. Gronenborn*,‡

Laboratory of Chemical Physics, Building 2, National Institute of Diabetes and Digestive and Kidney Diseases, and Laboratory of Cell Biology, Building 37, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892

Received May 8, 1990; Revised Manuscript Received July 3, 1990

ABSTRACT: The three-dimensional structure of a 30-residue synthetic peptide containing the carboxy-terminal "zinc finger" motif of a human enhancer binding protein has been determined by two-dimensional nuclear magnetic resonance (2D NMR) spectroscopy and hybrid distance geometry–dynamical simulated annealing calculations. The structure determination is based on 487 approximate interproton distance and 63 torsion angle (ϕ, ψ, and χ_1) restraints. A total of 40 simulated annealing structures were calculated, and the atomic rms distribution about the mean coordinate positions (excluding residues 29 and 30 which are ill-defined) is 0.4 Å for the backbone atoms, 0.8 Å for all atoms, and 0.41 Å for all atoms excluding the lysine and arginine side chains, which are disordered. The solution structure of the zinc finger consists of two irregular antiparallel β-strands connected by an atypical turn (residues 3–12) and a classical α-helix (residues 14–24). The zinc is tetrahedrally coordinated to the sulfur atoms of two cysteines (Cys-5 and Cys-8) and to the $N^{\epsilon 2}$ atoms of two histidines (His-21 and His-27). The two cysteine residues are located in the turn connecting the two β-strands (residues 5–8); one of the histidine ligands (His-21) is in the α-helix, while the second histidine (His-27) is at the end of a looplike structure (formed by the end of the α-helix and a turn). The general architecture is qualitatively similar to two previously determined low-resolution Cys_2His_2 zinc finger structures, although distinct differences can be observed in the β-strands and turn and in the region around the two histidines coordinated to zinc. Comparison of the overall polypeptide fold of the enhancer binding protein zinc finger with known structures in the crystallographic data base reveals a striking similarity to one region (residues 23–44) of the X-ray structure of proteinase inhibitor domain III of Japanese quail ovomucoid [Papamokos, E., Weber, E., Bode, W., Huber, R., Empie, M. W., Kato, I., & Laskowski, M. (1982) J. Mol. Biol. 158, 515–537], which could be superimposed with a backbone atomic rms difference of 0.95 Å on residues 3–25 (excluding residue 6) of the zinc finger from the enhancer binding protein. The presence of structural homology between two proteins of very different function may indicate that the so-called zinc finger motif is not unique for a class of DNA binding proteins but may represent a general folding motif found in a variety of proteins irrespective of their function.

Recently, a number of cDNA clones coding for specific DNA binding proteins that interact with a variety of different enhancer and promotor regions of human genes have been isolated. The respective proteins have been designated human immunodeficiency virus type I enhancer binding protein (HIV-EPI;[1] Maekawa et al., 1989), major histocompatibility complex binding protein 1 (MBP-1; Singh et al., 1988), and positive regulatory domain II of the human interferon β promotor binding factor 1 (PRDII-BF1; Fan & Maniatis, 1990). HIV-EP1 was identified as a DNA binding protein that interacts with a specific DNA sequence within the long terminal repeats of the HIV-I genome, functioning as an enhancer element in HIV transcriptional regulation (Maekawa et al., 1989). MBP-1 was shown to bind specifically to the promotor of the MHC class I H-$2K^b$ gene and to a similar site within the immunoglobulin κ gene (Singh et al., 1988), and PDRII-BF1 binds to PDRII, a virus-inducible element

† This work was supported by the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health (G.M.C. and A.M.G.).

* Authors to whom correspondence should be addressed.
‡ Laboratory of Chemical Physics, NIDDK.
§ Laboratory of Cell Biology, NCI.

[1] Abbreviations: HIV-EP1, human immunodeficiency virus type I enhancer binding protein; HIV-1, human immunodeficiency virus type I; PRDII, positive regulatory domain II of the human interferon β promotor; PRDII-BF1, positive regulatory domain II of the human interferon β promotor binding factor 1; MHC, major histocompatibility complex; MBP-1, major histocompatibility complex binding protein 1; NOE, nuclear Overhauser effect; NOESY, two-dimensional nuclear Overhauser enhancement spectroscopy; PE-COSY, primitive exclusive two-dimensional correlated spectroscopy; HOHAHA, two-dimensional homonuclear Hartmann–Hahn spectroscopy; SA, simulated annealing; HPLC, high-performance liquid chromatography; rms, root mean square.

This article not subject to U.S. Copyright. Published 1990 by the American Chemical Society

of the human interferon β gene promoter (Fan & Maniatis, 1990). The full-length cDNA clone designated PRDII-BF1 was shown to comprise the sequences of the two other clones. Thus all three proteins are identical (Fan & Maniatis, 1990) and are called enhancer binding protein throughout the present paper. Two widely separated regions within the extremely long amino acid sequence of the PRDII-BF1 clone and the equivalent regions of the partial clones exhibit substantial homology to the *Xenopus* transcription factor TFIIIA (Pelham & Brown, 1980), which comprises two repeated domains of approximately 25 residues, each of which exhibits the characteristic Cys_2His_2 pattern found in the consensus sequence of the zinc finger motif (Miller et al., 1985). These two pairs of zinc fingers bind to the same DNA sequence, and it was therefore postulated that PDRII-BF1 may be involved in the assembly of a multicomponent DNA–protein complex by binding simultaneously to widely separated recognition sequences (Fan & Maniatis, 1990).

The term "zinc finger" is now used regularly to describe this highly conserved motif in a plethora of eukaryotic transcription factors [see Gibson et al. (1988) for a compilation of sequences and Klug and Rhodes (1987) and Evans and Hollenberg (1988) for reviews]. Two classes of zinc fingers can be distinguished on the basis of differences in the zinc liganding amino acids. In the classical finger, the metal atom is coordinated to two histidine and two cysteine ligands (Cys_2His_2), whereas in the second class only cysteines (Cys_4) are involved. Until recently, the only structural information known about zinc fingers was the fact that zinc was crucial for correct folding of the protein and in turn essential for specific nucleic acid recognition (Blumberg et al., 1987; Stillman et al., 1988; Freedman et al., 1988; Nagai et al., 1988). Within the last year, studies on the solution structure of two synthetic peptides comprising zinc finger sequences of the Cys_2His_2 class have been reported; these were based on nuclear magnetic resonance (NMR) data (Lee et al., 1989; Klevit et al., 1990). It was shown that the main features of the Cys_2His_2 structure consist of two β-strands in which the two cysteine ligands are located and an α-helical region containing the two histidine ligands. These results confirmed the major characteristics of previously predicted models based on amino acid homologies and known three-dimensional structures of protein domains involved in metal binding (Berg, 1988; Gibson et al., 1988).

In this paper, we present the determination of a high-resolution three-dimensional (3D) solution structure for a 30-residue synthetic peptide comprising the carboxy-terminal zinc finger of the human enhancer binding protein by means of NMR and hybrid distance geometry–dynamical simulated annealing calculations. A comparison between the previously reported structures of classical zinc finger sequences and the present one is carried out, pointing out similarities and differences. In addition, we show that the tertiary structure of the zinc finger does not represent a new and unique protein fold but already exists within the data base of protein structures solved to date.

Experimental Procedures

Peptide Synthesis. The peptide was obtained by solid-phase peptide synthesis (Merrifield et al., 1982) on an Applied Biosystems 430A peptide synthesizer. The final protected resin was cleaved by using the low/high HF cleavage method (Tam et al., 1983). The extracted peptide (5% aqueous acetic acid) was purified by gel filtration chromatography on Sephadex G-25 and by HPLC on a Vydac C_4 semipreparative reverse-phase column using an acetonitrile–water solvent gradient containing 0.1% trifluoroacetic acid. The identity of the synthetic peptide was established by amino acid analysis and by the NMR sequential assignment procedure. Samples for NMR spectroscopy contained 4.8 mM peptide in either 99.996% D_2O or 90% H_2O/10% D_2O at pH 5.8 in the presence of 50 mM $ZnCl_2$. No chemical shift differences were observed for a ratio of 1:1 versus 1:10 of peptide to zinc. Thus, the complex comprises a single zinc atom per molecule of peptide. The excess of zinc, however, was employed to safeguard complete saturation under all experimental conditions employed. The sequence of the present peptide differs in one position from the published one (Maekawa et al., 1989; Fan & Maniatis, 1990). This arises from the fact that at the time the peptide was synthesized the human sequence had not been reported. We, however, had access to the equivalent murine sequence (Nakamura et al., personal communication), which contains a Ser instead of a Thr as the amino acid adjacent to the first cysteine. This conservative change is the only one within the double zinc finger motif for both the human and murine enhancer binding protein and has no structural consequences.

NMR Spectroscopy. The following spectra were recorded at 600 MHz on a Bruker AM 600 spectrometer at 6 and 25 °C in both H_2O and D_2O: PE-COSY (Mueller, 1987) with a 35° mixing pulse, P-COSY (Marion & Bax, 1988) and HOHAHA (Davis & Bax, 1985) with mixing times ranging from 26 to 53 ms, and NOESY (Jeener et al., 1979) with mixing times of 50, 150, and 300 ms. All spectra were recorded in the pure phase absorption mode by using the time-proportional incrementation method (Marion & Wüthrich, 1983). HOHAHA spectra were recorded with a $WALTZ17_y$ anisotropic mixing sequence sandwiched between 1.5-ms trim pulses (Bax, 1989). In the case of NOESY and HOHAHA spectra recorded in H_2O, the water resonance was suppressed with a semiselective jump-return read sequence (Plateau & Guéron, 1982; Bax et al., 1987), while for the P-COSY spectrum recorded in H_2O, weak coherent presaturation was used. For NOESY and HOHAHA experiments, 1024 t_1 increments of 2K data points were collected. For P-COSY and PE-COSY experiments 1024 t_1 increments of 4K data points were collected and the FIDs were zero filled to 1 Hz/point in the F_2 dimension. $^3J_{\alpha\beta}$ coupling constants were measured directly from the reduced multiplets in the PE-COSY spectrum recorded in D_2O. $^3J_{HN\alpha}$ coupling constants were obtained by best fitting the experimental peak-to-peak separation of each $NH(F_2)$–$C^\alpha H(F_1)$ cross-peak (measured from several Gaussian apodization functions) to the calculated peak-to-peak separation obtained by Fourier transformation of the time domain response for a multiplet centered at zero frequency given by the expression $(\sin \pi J_{HN\alpha} t_2) e^{-t_2/T_2}$, by optimizing the value of $^3J_{HN\alpha}$ and the transverse relaxation time T_2 (Forman-Kay et al., 1990). Stereospecific assignments and ϕ, ψ, and χ_1 torsion angle restraints were derived from $^3J_{HN\alpha}$ and $^3J_{\alpha\beta}$ coupling constants, measured from the PE-COSY spectrum, and intraresidue and sequential NOEs involving the NH, $C^\alpha H$, and $C^\beta H$ protons observed in the 50-ms mixing time NOESY spectra, employing the conformational search program STEREOSEARCH (Nilges et al., 1990).

Calculations. Structures were computed from the experimental NMR data according to the hybrid distance geometry–dynamical simulated annealing method of Nilges et al. (1988), using the programs DISGEO (Havel & Wüthrich, 1984; Havel, 1986) and XPLOR (Brünger, 1988), which is derived from the program CHARMM (Brooks et al., 1983) and has been specially adapted for dynamical simulated annealing calcu-

Table I: Proton Resonance Assignments of the Enhancer Binding Protein Zinc Finger Peptide at 6 °C and pH 5.8[a]

residue	chemical shift (ppm)			
	NH	$C^{\alpha}H$	$C^{\beta}H$	others
R1		3.84	1.58*, 1.00	$C^{\gamma}H$ 1.75, 1.65; $C^{\delta}H$ 2.69, 2.58
P2		4.28	2.04*, 1.22	$C^{\gamma}H$ 2.16, 2.04; $C^{\delta}H$ 3.70, 3.70
Y3	7.42	4.96	2.95*, 2.71	$C^{\delta}H$ 6.91; $C^{\epsilon}H$ 6.81
H4	9.56	4.47	3.26*, 2.97	$C^{\delta 2}H$ 6.76; $C^{\epsilon 1}H$ 7.91
C5	8.34	4.51	3.40*, 2.71	
S6	9.04	4.32	3.51*, 3.30	
Y7	9.83	4.52	2.72*, 1.56	$C^{\delta}H$ 6.88; $C^{\epsilon}H$ 6.74
C8	8.01	4.93	3.28*, 3.18	
N9	8.23	4.92	2.90, 2.90[b]	$N^{\delta}H_2$ 7.64*, 6.94
F10	9.03	4.30	2.70, 2.41*	$C^{\delta}H$ 7.18; $C^{\epsilon}H$ 7.37; $C^{\zeta}H$ 7.31
S11	7.91	4.71	2.63, 2.56*	
F12	8.97	4.92	3.57*, 2.64	$C^{\delta}H$ 7.26; $C^{\epsilon}H$ 6.42; $C^{\zeta}H$ 6.66
K13	9.75	4.24	2.05, 2.05	$C^{\gamma}H$ 1.48, 1.48; $C^{\delta}H$ 1.74, 1.74; $C^{\epsilon}H$ 2.97, 2.97
T14	7.31	4.79	4.38	$C^{\gamma}H_3$ 1.16
K15	8.52	2.93	1.44*, 1.16	$C^{\gamma}H$ 0.97, 0.97; $C^{\delta}H$ 1.08, 1.08; $C^{\epsilon}H$ 2.84, 2.84
G16	8.64	3.74, 3.64		
N17	7.56	4.53	2.88, 2.70*	NH_2 8.03*, 7.20
L18	7.08	3.21	2.02*, 1.17	$C^{\gamma}H$ 1.56; $C^{\delta}H$ 1.05, 1.02*
T19	8.64	3.68	4.30	$C^{\gamma}H_3$ 1.12
K20	7.75	3.80	1.75, 1.75	$C^{\gamma}H$ 1.58, 1.30; $C^{\delta}H$ 1.61, 1.61; $C^{\epsilon}H$ 2.87, 2.87
H21	7.37	4.27	3.28, 2.91*	$C^{\delta 2}H$ 7.28; $C^{\epsilon 1}H$ 7.33
M22	8.57	4.12	2.29, 2.15*	$C^{\gamma}H$ 2.92, 2.77; $C^{\epsilon}H_3$ 2.07
K23	7.68	4.24	1.90, 1.69	$C^{\gamma}H$ 1.54, 1.54; $C^{\delta}H$ 1.39, 1.39; $C^{\epsilon}H$ 2.83, 2.83
S24	7.42	4.38	4.30, 4.30	
K25	8.42	4.13	1.65, 1.65	$C^{\gamma}H$ 1.24, 1.24; $C^{\delta}H$ 1.75, 1.75; $C^{\epsilon}H$ 2.88, 2.88
A26	7.79	3.79	0.74	
H27	7.63	4.76	2.84, 2.77	$C^{\delta 2}H$ 6.49; $C^{\epsilon 1}H$ 7.75
S28	7.87	4.29	3.75, 3.75	
K29	8.42	4.27	1.81, 1.69	$C^{\gamma}H$ 1.37, 1.37; $C^{\delta}H$ 1.80, 1.80; $C^{\epsilon}H$ 2.91, 2.91
K30	8.07	4.06	1.74, 1.64	$C^{\gamma}H$ 1.32, 1.32; $C^{\delta}H$ 1.63, 1.63; $C^{\epsilon}H$ 2.89, 2.89

[a] Chemical shifts are reported with respect to 4,4-dimethyl-4-silapentane-1-sulfonate. Stereospecific assignments are denoted as follows: for the C^{β} methylene protons, the asterisk indicates the $H^{\beta 3}$ proton; for the NH_2 protons of Asn, the asterisk indicates the $H^{\delta 21}$ proton that is cis to C^{β}; for the methyl protons of Leu, the asterisk refers to the $C^{\delta 1}$ methyl group. [b] The chemical shifts of the $C^{\beta}H$ protons at 25 °C are 2.88 and 2.82 ppm, with the downfield resonance corresponding to the $H^{\beta 3}$ proton.

lations with NMR restraints (Clore et al., 1985). This involves two stages. In the first stage a set of substructures, comprising about a third of the total number of atoms, is embedded from *n*-dimensional distance space into cartesian coordinate space (Havel et al., 1983) without checking the triangle inequalities. These substructures are very crude but have approximately the correct polypeptide fold. The remaining atoms are then added with the side chains placed in an extended conformation, and the resulting structures are subjected to the protocol of dynamical simulated annealing described by Nilges et al. (1988), with a few minor modifications. This particular protocol is designed to overcome large energy barriers along the path toward the global minimum region of the target function and during the early stages of the calculation permits chains to pass through one another. The target function that is being minimized comprises quadratic harmonic terms for bonds, angles, and improper torsion angles (i.e., planes and chirality restraints), a quartic van der Waals repulsion term, and square-well quadratic potential terms for interproton distance and torsion angle restraints (Nilges et al., 1988). It should be noted that *no* electrostatic, hydrogen-bonding, or 6-12 Lennard-Jones van der Waals terms are used in the calculations.

Results and Discussion

NMR Assignments and Qualitative Secondary Structure Derivation. The complete sequential assignment of the enhancer binding protein synthetic finger was accomplished by conventional 2D NMR methods [see Wüthrich (1986) and Clore and Gronenborn (1987) for reviews]. Through-bond scalar connectivities were established from HOHAHA (Bax, 1989), P-COSY (Marion & Bax, 1988) and PE-COSY (Mueller, 1987) experiments while through-space (<5 Å) connectivities were determined with NOESY (Jeener et al., 1979) spectroscopy. An example of a NOESY spectrum is shown in Figure 1 with a number of sequential NOEs indicated, a summary of the short-range NOE data involving the NH, $C^{\alpha}H$, and $C^{\beta}H$ protons is given in Figure 2, and the complete list of assignments is presented in Table I. To resolve ambiguities arising from chemical shift degeneracy, all spectra were recorded at two temperatures (6 and 25 °C).

Qualitative interpretation of the short-range NOE data given in Figure 1 reveals two elements of secondary structure. A nonclassical turn or loop comprising residues 6–10 is evidenced by the observation of medium-strength NH(i)–NH(i+1) and $C^{\alpha}H(i)$–NH(i+1) NOEs, together with a NH(i)–NH(i+2) NOE between residues 6 and 8 and a $C^{\beta}H(i)$–NH(i+3) NOE between residues 5 and 8. An 11 residue long α-helix from residues 14 to 24 is indicated by a consecutive stretch of NH(i)–NH(i+1) NOEs (see Figure 2) together with the presence of NH(i)–NH(i+2), $C^{\alpha}H(i)$–NH(i+2,3,4), and $C^{\alpha}H(i)$–$C^{\beta}H(i,i+3)$ NOEs. These qualitative features are borne out in the final complete 3D structure of the peptide.

Experimental Restraints and Structure Calculations. The NOEs were identified in the 150- and 300-ms NOESY spectra and classified on the basis of the 50- and 150-ms NOESY spectra into strong, medium, weak, and very weak, corresponding to interproton distance restraints of 1.8–2.7, 1.8–3.3, 2.3–5.0, 2.5–5.0 Å, respectively. Upper limits for distances involving methyl protons and nonstereospecifically assigned methylene protons were corrected appropriately for center averaging (Wüthrich et al., 1983), and an additional 0.5 Å was added to the upper limits for distances involving methyl protons (Clore et al., 1987). Stereospecific assignments and ϕ, ψ, and χ_1 torsion angle restraints were derived from $^3J_{HN\alpha}$

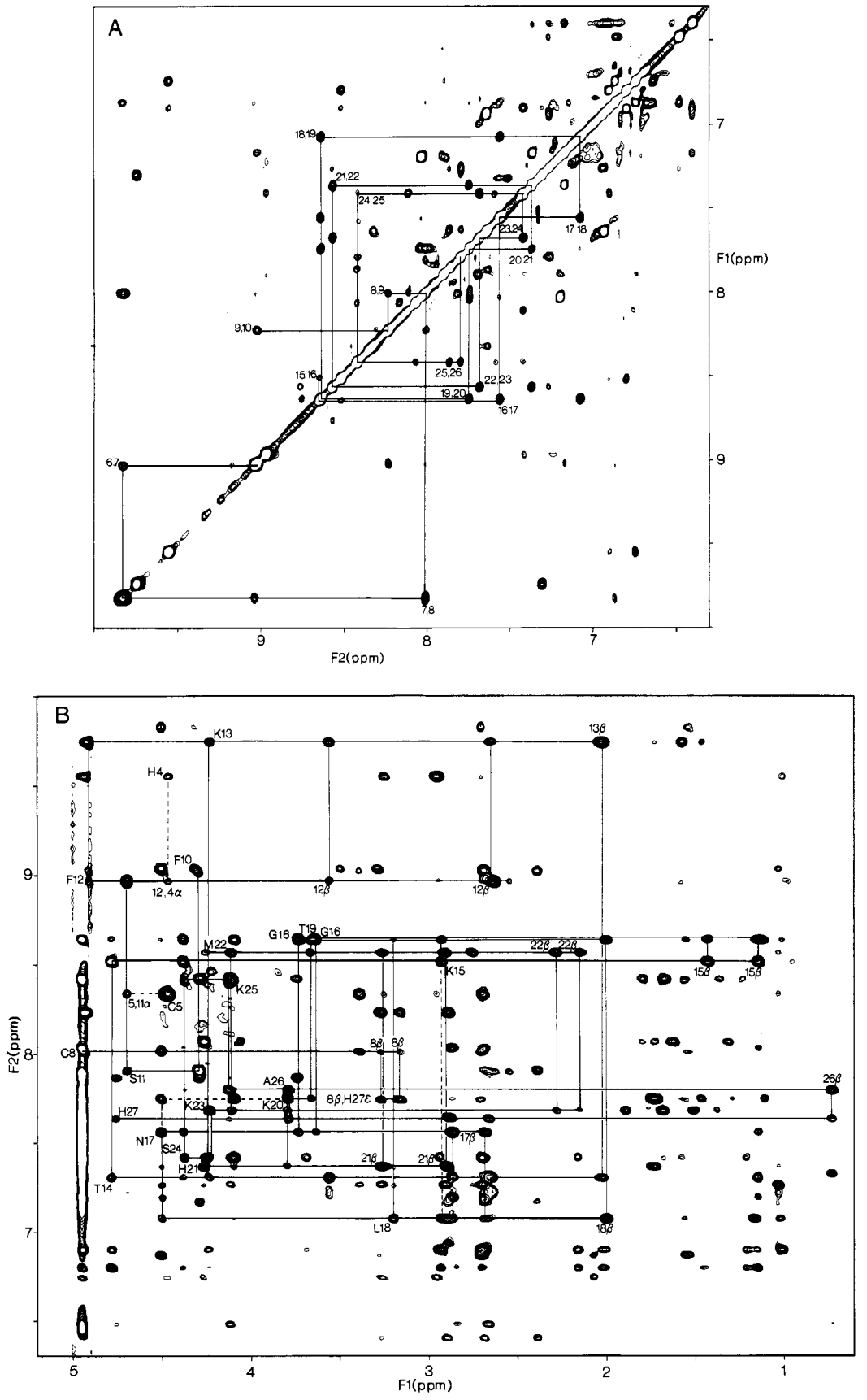

FIGURE 1: (A) NH(F_2 axis)–NH(F_1 axis) and (B) NH(F_2)–aliphatic(F_1 axis) regions of the 600-MHz NOESY spectrum (150-ms mixing time) of the zinc finger peptide (4.8 mM) in the presence of 50 mM $ZnCl_2$, pH 5.8, recorded at 6 °C in 90% H_2O/10% D_2O. A series of NH(i)–NH(i+1) and C$^{\alpha}$H(i)–NH(i+1) sequential NOE connectivities are indicated in (A) and (B), respectively. Also shown in (B) are some C$^{\beta}$H(i)–NH(i+1), C$^{\alpha}$H(i)–NH(i+3) (denoted by dashed lines), and long-range C$^{\alpha}$H(i)–NH(j) NOEs. Residue labels in (B) are at the positions of the intraresidue C$^{\alpha}$H(i)–NH(i) and C$^{\beta}$H(i)–NH(i) cross-peaks, the latter denoted by the letter β. The NH(i)–NH(i+1) connectivities shown in (A) extend from residues 6 to 10, the region of the nonclassical turn connecting the two irregular β-strands, and from residues 15 to 26, which includes the α-helix (residues 14–24). The C$^{\alpha}$H(i)–NH(i+1) connectivities in (B) extend from residues 12 to 27.

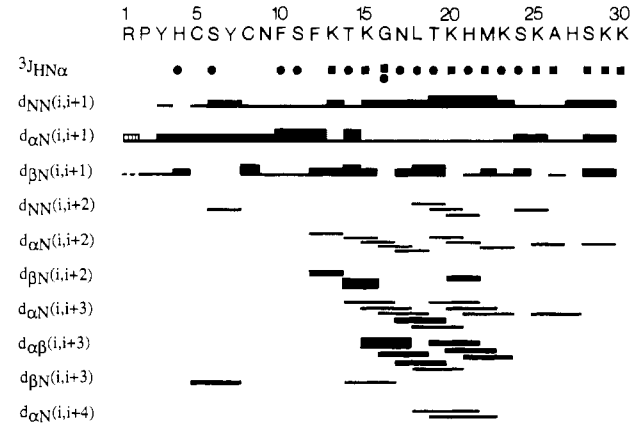

FIGURE 2: Summary of the short-range NOEs involving the NH, $C^{\alpha}H$, and $C^{\beta}H$ protons, as well as the $C^{\delta}H$ protons of proline. The NOEs are classified into strong, medium, and weak according to the thickness of the lines. The $C^{\alpha}H$(Arg-1)–$C^{\delta}H$(Pro-2) NOE is indicated by the hatched box along the same line as the $C^{\alpha}H(i)$–NH(i+1) NOEs. Also shown is a summary of the $^3J_{HN\alpha}$ coupling constant data with $^3J_{HN\alpha}$ ≤6 Hz and ≥8 Hz indicated by the solid circles and squares, respectively.

and $^3J_{\alpha\beta}$ coupling constants, measured from a PE-COSY spectrum, and intraresidue and sequential NOEs involving the NH, $C^{\alpha}H$, and $C^{\beta}H$ protons by means of a conformational grid search of ϕ, ψ, χ_1 space (Nilges et al., 1990; Kraulis et al., 1989). The minimum ranges employed for the ϕ, ψ, and χ_1 torsion angles were ±30°, ±50°, and ±20°, respectively.

As the pH required for zinc coordination is rather high (pH 5.8), no slowly exchanging NH protons were observed. Consequently, no explicit hydrogen-bonding restraints were derived from the experimental data.

The structures were calculated by using the hybrid distance geometry–dynamical simulated annealing method of Nilges et al. (1988). An iterative strategy to the structure determination was employed by carrying out a series of successive calculations with more and more restraints incorporated at each successive stage (Kraulis et al., 1989). By this means, many ambiguities in the assignment of NOE cross-peaks arising from chemical shift degeneracy could be resolved by analysis of the initial low-resolution structures, thereby allowing the extraction of further interproton distance restraints. Similarly, inspection of the low-resolution structures permitted additional stereospecific assignments to be obtained and the elucidation of the zinc coordination site on the histidine residues. The histidine residues can coordinate to the zinc through either the $N^{\delta 1}$ or the $N^{\epsilon 2}$ atoms. As the coordination was not known a priori, the initial calculations were carried out without zinc. These indicated unambiguously that the zinc could only coordinate in a tetrahedral geometry to the $N^{\epsilon 2}$ atoms of His-21 and His-27. Zinc was subsequently incorporated into the calculations by introducing appropriate covalent restraints to ensure that the coordination geometry was approximately tetrahedral, that the zinc lay in the plane of the two histidine imidazole rings, and that the Cys(C^{β})–Cys(S^{γ})–Zn and Zn–His($N^{\epsilon 2}$)–His($C^{\delta 2}$/$C^{\epsilon 1}$) angles were around 107° and 120°, respectively. In addition, bond lengths of 2.3 and 2.0 Å were employed for the S^{γ}(Cys)–Zn and His($N^{\epsilon 2}$)–Zn bonds (Diakun et al., 1986).

The final simulated annealing (SA) structures were calculated on the basis of 487 interproton distance restraints, made up of 121 sequential ($|i - j| = 1$), 90 medium-range ($2 \leq |i - j| \leq 5$), and 87 long-range ($|i - j| > 5$) interresidue NOEs and 189 intraresidue NOEs, and 63 torsion angle restraints comprising 28 ϕ, 14 ψ, and 21 χ_1 angles. The interproton distances included NOEs to 17 stereospecifically assigned β-methylene protons (out of a total of 26 β-methylene groups), as well as to the stereospecifically assigned methyl groups of Leu-18. The latter, as well as the stereospecific assignments for the β-methylene protons of Asn-9 and Ser-11, were only obtained after the initial rounds of calculations.

Converged Structures. A total of 40 SA structures was calculated and superpositions of the backbone atoms are shown in Figure 3A. Pertinent structural statistics are given in Table II, and the atomic rms distribution of the individual structures about the mean coordinate positions as a function of residue number is shown in Figure 4. All the structures satisfy the experimental restraints within their errors, display very small deviations from idealized covalent geometry, and have good nonbonded contacts as judged both by the low value of the quartic van der Waals repulsion term used in the target function and by the negative value of the Lennard-Jones van der Waals energy. With the exception of the last two residues (Lys-29 and Lys-30) whose positions are ill-determined by the experimental data, the backbone is very well-defined with an average angular rms difference for the ϕ and ψ backbone torsion angles of 12 ± 8°. The poor definition of Lys-29 and Lys-30 is due to the fact that, with the exception of a single nonsequential NOE between the $C^{\alpha}H$ proton of residue 28 and the NH of residue 30, the only NOEs observed involving these two residues were either intraresidue or sequential (i.e., i, i+1) in nature. The average atomic rms distribution of residues 1–28 of the 40 SA structures about the mean coordinate positions is 0.4 ± 0.008 Å for the backbone atoms, 0.78 ± 0.10 Å for all atoms, and 0.41 ± 0.07 Å for all atoms excluding lysine and arginine side chains. Thus the definition of many of the side chains is as good as that of the backbone atoms (Figure 3C). Indeed, the only poorly defined side chains are Arg-1, Ser-28, and the lysine residues (Lys-13, Lys-15, Lys-20, Lys-23, Lys-25, Lys-29, Lys-30), all of which are highly solvent accessible (Figure 3B). It is worth point out in this respect that the $^3J_{\alpha\beta}$ coupling constant data for these residues, with the exception of Arg-1 and Lys-15, are indicative of multiple χ_1 side-chain conformations (i.e., $^3J_{\alpha\beta a} \sim {}^3J_{\alpha\beta b} \sim$ 6–7 Hz). In the case of Arg-1 and Lys-15, the side-chain coordinates are defined up to the C^{γ} position, but are disordered thereafter.

Description of the Structure. The overall folding topology of the zinc finger peptide is presented in Figure 3A. The structure starts with a tight turn from residues 1–4. This is followed by two irregular antiparallel β-strands (residues 3–4 and 9–12) connected by a nonclassical loop or turn (residues 5–8) containing the two cysteine residues. In all the structures, the antiparallel arrangement of the two β-strands is stabilized by two backbone hydrogen bonds between Cys-5(NH) and Phe-10(CO) and between Phe-12(NH) and Tyr-3(CO). The second β-strand is followed by a turn (residues 12–14) that leads into a regular α-helix extending from Thr-14 to Ser-24. The helical region contains backbone NH–CO hydrogen bonds between residues 18 and 14, 19 and 15, 20 and 16, 21 and 17, 22 and 18, 23 and 19, and 24 and 20. It is important to bear in mind that these hydrogen bonds are a direct result of the experimental restraints, as neither explicit hydrogen-bonding restraints nor hydrogen-bonding or electrostatic potentials were used in the structure calculations. The α-helix terminates in a looplike structure, which positions His-27 in the correct orientation for tetrahedrally coordinating the zinc atom.

The coordination geometry of the zinc atom is almost perfectly tetrahedral, and the zinc atom lies in the plane of the two imidazole rings (Figure 3 and Table III). However, one of the six tetrahedral angles is significantly different from the value of 109° for a perfect tetrahedron. This is the angle

Table II: Structural Statistics and Atomic rms Differences[a]

(A) Structural Statistics

	⟨SA⟩	$\overline{(SA)}$r
rms deviations from exptl distance restraints (Å)[b]		
all (487)	0.044 ± 0.002	0.040
interresidue sequential ($i,i+1$) (121)	0.043 ± 0.003	0.040
interresidue medium range ($1 < \|i-j\| \leq 5$) (90)	0.041 ± 0.006	0.032
interresidue long range ($\|i-j\| > 5$) (87)	0.052 ± 0.006	0.045
intraresidue (189)	0.043 ± 0.003	0.040
rms deviations from exptl dihedral restraints (deg) (63)[b]	0.473 ± 0.101	0.383
F_{NOE} (kcal·mol^{-1})[c]	28.7 ± 2.5	22.9
F_{tor} (kcal·mol^{-1})[c]	0.89 ± 0.36	0.56
F_{repel} (kcal·mol^{-1})[c]	11.6 ± 2.0	7.5
$E_{L\text{-}J}$ (kcal·mol^{-1})[d]	−83.9 ± 55.6	−80.5
deviations from idealized geometry[e]		
bonds (Å) (512)	0.008 ± 0	0.007
angles (deg) (927)	2.283 ± 0.006	2.266
impropers (deg) (218)	0.668 ± 0.036	0.600

(B) Atomic rms Differences (Å) (residues 1–28)

	backbone atoms	all atoms	all atoms (excluding Arg and Lys side chains)
⟨SA⟩ vs $\overline{SA}$	0.40 ± 0.008	0.78 ± 0.10	0.41 ± 0.07
$\overline{(SA)}$r vs $\overline{SA}$	0.24	0.38	0.24
⟨SA⟩ vs $\overline{(SA)}$r	0.46 ± 0.09	0.87 ± 0.11	0.47 ± 0.08

[a] The notation of the structures is as follows: ⟨SA⟩ are the 40 final dynamical simulated annealing structures, $\overline{SA}$ is the mean structure obtained by averaging the coordinates of the individual 40 SA structures best fitted to each other (excluding residues 29 and 30); $\overline{(SA)}$r is the restrained minimized mean structure obtained by restrained minimization of $\overline{SA}$. The number of terms for the various restraints is given in parentheses. [b] None of the structures exhibited distance violations greater than 0.3 Å or dihedral angle violations greater than 3°. The torsion angle restraints comprise 28 ϕ, 14 ψ, and 21 χ_1 restraints. [c] The values of the square-well NOE and torsion angle potentials [cf. eqs 2 and 3 in Clore et al. (1986)] are calculated with force constants of 50 kcal·mol^{-1}·Å^{-2} and 200 kcal·mol^{-1}·rad^{-2}, respectively. The value of the quartic van der Waals repulsion term [cf. eq 5 in Nilges et al. (1988)] is calculated with a force constant of 4 kcal·mol^{-1}·Å^{-4} with the hard-sphere van der Waals radii set to 0.8 times the standard values used in the CHARMM empirical energy function (Brooks et al., 1983). [d] $E_{L\text{-}J}$ is the Lennard-Jones van der Waals energy calculated with the CHARMM empirical energy function (Brooks et al., 1983). It is not included into the target function for simulated annealing. Zn has an ionic radius of 0.74 Å (Kittel, 1966) and an atomic radius in a nonionic solid of 1.33 Å (Pauling, 1967). The closest nonbonded atom to the zinc is the NH proton of Cys-8, which is 2.5 Å away. Consequently, the contribution of the zinc to the Lennard-Jones energy is negligible. [e] The improper torsion terms serve to maintain planarity and chirality; they also maintain the peptide bond of all residues in the trans conformation. The terms describing the coordination geometry of the zinc atom are included in the bond, angle, and improper terms (see text and Table III).

Table III: Zinc Coordination Geometry[a]

Angles (deg) of the Tetrahedron Centered at the Zinc Atom	
S^{γ}(Cys-5)–Zn–S^{γ}(Cys-8)	111.8 ● 0.25
S^{γ}(Cys-5)–Zn–$N^{\epsilon 2}$(His-21)	111.3 ± 0.31
S^{γ}(Cys-5)–Zn–$N^{\epsilon 2}$(His-27)	111.1 ± 0.58
S^{γ}(Cys-8)–Zn–$N^{\epsilon 2}$(His-21)	110.4 ± 0.17
S^{γ}(Cys-8)–Zn–$N^{\epsilon 2}$(His-27)	97.9 ± 1.31
$N^{\epsilon 2}$(His-21)–Zn–$N^{\epsilon 2}$(His-27)	111.8 ± 0.49
Angles (deg) of Cysteine Coordination	
C^{β}(Cys-5)–S^{γ}(Cys-5)–Zn	111.8 ± 0.48
C^{β}(Cys-8)–S^{γ}(Cys-8)–Zn	108.9 ± 0.33
Angles (deg) of Histidine Coordination	
Zn–$N^{\epsilon 2}$(His-21)–$C^{\delta 2}$(His-21)	126.7 ± 0.32
Zn–$N^{\epsilon 2}$(His-21)–$C^{\epsilon 1}$(His-21)	128.1 ± 0.28
Zn–$N^{\epsilon 2}$(His-27)–$C^{\delta 2}$(His-27)	130.1 ± 0.26
Zn–$N^{\epsilon 2}$(His-27)–$C^{\epsilon 1}$(His-27)	124.8 ± 0.23
planarity of zinc atom and two imidazole rings[b]	180.0 ± 0.42

[a] The values given are the averages for the 40 SA structures. [b] The location of the zinc atom in the plane of the two imidazole rings is defined by two improper torsion angles for each histidine residue: Zn–$N^{\epsilon 2}$(His)–$C^{\delta 2}$(His)–C^{γ}(His) and Zn–$N^{\epsilon 2}$(His)–$C^{\epsilon 1}$(His)–$N^{\delta 1}$(His).

formed by the S^{γ} of Cys-8, the zinc atom, and the $N^{\epsilon 2}$ atom of His-27, which has a value of 97.9 ± 1.31° (Table III). In this respect, we note that exactly the same finding was observed in the recently determined NMR structure of the first zinc binding domain of the HIV *gag* protein p55 (Summers et al., 1990). In this peptide the zinc atom is coordinated to three cysteines and one histidine, and the average value for five of the six tetrahedral angles is 111.7 ± 3.9°, while that of the sixth, involving the first cysteine residue and the histidine, is 96.5 ± 0.53°. A similar coordination geometry is also observed for the noncatalytic zinc atom in the closed form of horse liver alcohol dehydrogenase: in this case the zinc atoms is coordinated to four cysteines; the average for five of the six tetrahedral angles is 111.4 ± 7.4°, while the sixth angle has a value of 97° (Eklund & Bränden, 1983).

The hydrophobic core of the zinc finger from the enhancer binding protein, which is located in close proximity to one of the zinc-coordinating histidine residues, His-21, is formed by three highly conserved residues, Tyr-3, Phe-12, and Leu-18, and Phe-10, which is not found in other zinc finger sequences. This core serves to tightly pack the antiparallel β-strands and α-helix (Figure 3C). A number of other stabilizing side-chain interactions are also found within the structure. The first turn (residues 1–4) is stabilized by a hydrogen bond between the N-terminal amide group and the $N^{\epsilon 2}$ atom in the imidazole ring of His-4. The conformation of Asn-9 is fixed by an intraresidue hydrogen bond between the side-chain amide group and the backbone carbonyl oxygen. The backbone conformation of the first four residues of the α-helix is stabilized by hydrogen bonds between the side-chain amide group of Asn-17 and the $O^{\gamma 1}$ atom of Thr-14 and between the Thr-19 $O^{\gamma 1}H$ hydroxyl group and the backbone carbonyl oxygen atom of Lys-15. Finally, there is a potential hydrogen bond between the hydroxyl group of Ser-28 and the S^{γ} atom of Met-22, which would further serve to stabilize the conformation of the last turn of the helix and the loop connecting the coordinating histidine residues, His-21 and His-27. The zinc atom is completely buried and shielded from solvent, thereby stabilizing its interaction with the histidine and cysteine ligands. This is achieved not only by the cysteine and histidine coordinating ligands but also by residues placed at the four faces of the tetrahedron. In the view shown in Figure 3C, Phe-10, Phe-12, and Leu-18 shield the top face, Ser-6, Met-22, and

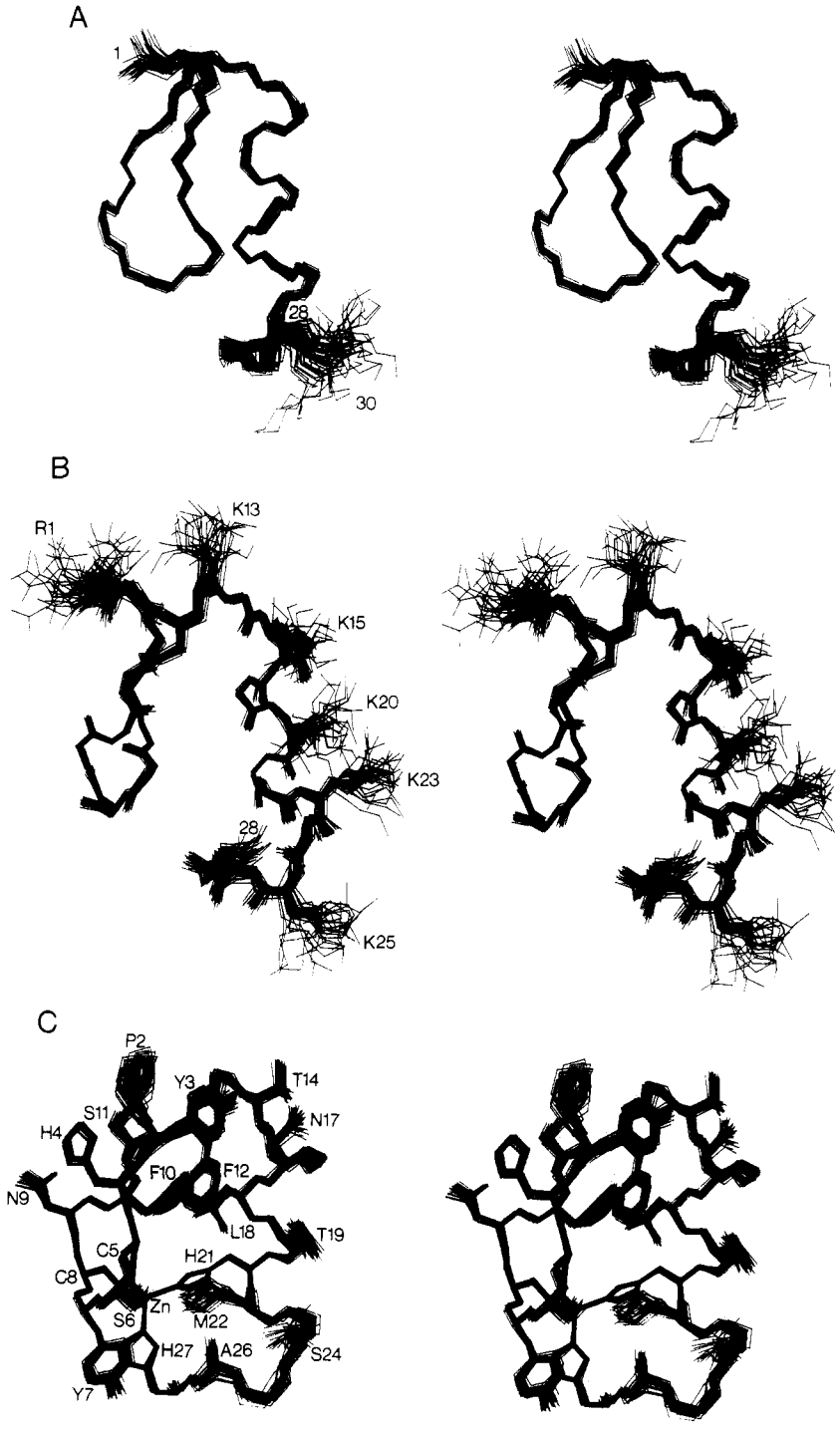

FIGURE 3: (A) Best fit superposition of the backbone (N, C^{α}, C) atoms of the 40 SA structures (residues 1–30). (B) Best fit superposition of the backbone (N, C^{α}, C, O) atoms and the arginine and lysine side chains of the 40 SA structures (residues 1–28). (C) Best fit superposition of the backbone (N, C^{α}, C) atoms and the remaining side chains of the 40 SA structures (residues 1–27).

Ser-28 shield the face at the back side of the figure, and Tyr-7 shields one of the side faces and Ala-26 the other.

In analyzing structures it is often useful to search the high-resolution crystallographic data base of protein structures for similar local backbone structural motifs, particularly for regions that are somewhat unusual, such as that containing the two cysteine residues. This was carried out by using the interactive graphics program FRODO (Jones, 1978) and the algorithm of Jones and Thirup (1986), which matches a fragment in the structure with those present in the data base by a comparison of C^{α}–C^{α} distances. This revealed that the region comprising the initial tight turn at the N-terminus (residues 1–5) could be superimposed within 0.3 Å on residues 121–125 of sperm whale oxyhemoglobin (Philips, 1980) and residues 285–289 of penicillopepsin (James & Sielecki, 1983). The nonclassical turn or loop region containing the two cysteine residues is well represented in the crystallographic data base with 20 four-residue fragments superimposing within 0.3 Å on residues 5–8. Further, the segment from residues 4 to 9 can be superimposed with an rms difference of 0.46 Å on residues 16–21 of cytochrome *c*-551 (Matsuura et al., 1982), while the segment from residues 5 to 10 can be superimposed on residues 109–114 of erythrocruorin (Steigemann & Weber, 1979) with an rms difference of 0.44 Å. The region connecting the second β-strand to the α-helix as well as the adjacent residues comprising the segment from residues 11 to 17 can

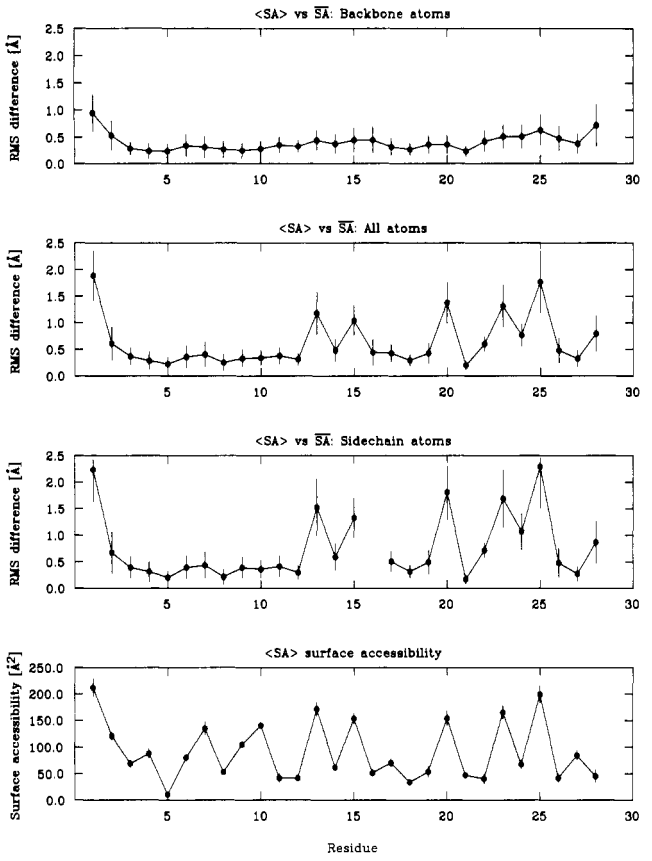

FIGURE 4: Atomic rms distribution of the 40 individual SA structures about the mean structure $\overline{SA}$, together with the variation in surface accessibility as a function of residue number (excluding residues 29 and 30). The filled-in circles represent the average values at each residue; the bars, the standard deviations in these values.

be fitted to within less than 0.4 Å on residues 3–9 and 44–50 of crambin (0.29 Å; Hendrickson & Teeter, 1981) and bovine pancreatic trypsin inhibitor (0.38 Å; Deisenhofer & Steigemann, 1975), respectively. Finally, the region (residues 21–27) between the two zinc-coordinating histidine residues can be fitted with an rms difference of 0.66 Å to residues 87–93 of poplar plastocyanin (Guss & Freeman, 1983). Thus, the structure of the zinc finger can easily be built from existing fragments present in the crystallographic data base, despite the fact that there is no sequence homology between the matching fragments.

It seems appropriate at this point to compare the high-resolution structure of the zinc finger peptide from the enhancer binding protein with the NMR structures of the two other classical zinc finger structures reported to date, *Xfin* 31 (Lee et al., 1989) and ADR1b (Klevitt et al., 1990). It is, however, important to bear in mind that the present structure has been determined to a much higher degree of precision than the other two structures due to the use of both stereospecific assignments and almost four times the number of experimental restraints (487 NOE and 63 torsion angle restraints versus 144 NOE and 12 ϕ torsion angle restraints for *Xfin* 31 and about 150 NOE restraints for ADR1b). Although qualitatively all three structures are rather similar and have essentially the same structural elements, unique differences are apparent, with respect to both the *Xfin* 31 and the ADR1b.

Like the ADR1b finger, but unlike the *Xfin* 31 structure, the enhancer binding protein finger does *not* contain a regular antiparallel β-sheet with a hairpin turn at the N-terminus (residues 3–12). We find no evidence for a hydrogen bond between the amide proton of residue 3 and the carbonyl oxygen of residue 12, as reported for the *Xfin* 31 structure; indeed the carbonyl oxygen of Phe-12 points away from the NH proton of Tyr-3. Differences are also apparent within the hydrophobic core. In particular, the Tyr-1 side chain in the *Xfin* 31 structure (which is equivalent to the Tyr-3 side chain in the present zinc finger structure) is pointing away from the hydrophobic cluster formed by the conserved Phe and Leu side chains. Whether this represents a genuine difference or arises from the fact that Tyr-1 is the N-terminal residue in *Xfin* 31 cannot be decided at the present time. The orientation, however, of Phe-12 and Leu-18 in the zinc finger from the enhancer binding protein is identical with that of the corresponding residues, Phe-10 and Leu-16, in *Xfin* 31. A comparison with the hydrophobic core of the ADR1b structure cannot be carried out in detail since the calculated ADR1b structures show a large degree of variability, even in this central portion of the structure. However, it appears from the data presented that the Phe-104 side chain of ADR1b, which is equivalent to Tyr-3 of the enhancer binding protein finger, is part of the hydrophobic core.

The α-helix of the zinc finger from the enhancer binding protein is similar in length to the helices reported for the other two zinc finger structures and starts at the same residue, namely, Thr-14 in the present case. Unlike the related *Xfin* 31 and ADR1b fingers, however, the helix in the present structure does not extend completely through the region comprising the histidine ligands. In the enhancer binding protein finger, the α-helix terminates at Ser-24, which is located midway between the two conserved histidine residues (His-21 and His-27). Thereafter, the chain continues by forming a loop which positions the imidazole ring of the second histidine for optimal interaction with the zinc atom. Further, we find no evidence for a 3_{10} helix as was reported for the C-terminal helical part of *Xfin* 31. The difference in the precise nature of the helical conformation around the two chelating histidine residues can be attributed to the difference in length of the intervening sequence between the two histidines, which comprises five residues in the enhancer binding protein finger versus three in both *Xfin* 31 and ADR1b.

A common feature to all three zinc finger structures is the fact that a large number of positively charged and polar side chains are located on the solvent-exposed face of the helical region, rendering speculations about their involvement in DNA recognition obvious. In the enhancer binding protein finger these residues are Thr-14, Lys-15, Asn-17, Thr-19, Lys-20, and Lys-23; the equivalent residues in *Xfin* 31 are Glu-12, Lys-13, Ala-15, Ser-17, Arg-18, and Arg-21, while in ADR1b they are Arg-115, Gln-116, His-118, Lys-121, Arg-120, and Arg-125.

Structural Homology to Part of Ovomucoid Third Domain. While comparing small fragments of the zinc finger structure with fragments present in the crystallographic data base as discussed above, we found that the backbone of the enhancer binding protein zinc finger structure comprising residues 3–25 (excluding 6) could be superimposed with an atomic rms difference of 0.95 Å on residues 23–44 of the proteinase inhibitor ovomucoid third domain from Japanese quail (Papamokos et al., 1982). This striking similarity of the polypeptide fold occurs despite the lack of any significant homology between the two sequences (three matches out of 22 residues corresponding to Asn-9, Lys-15, and Asn-17). A superposition of the backbone (C, N, C^{α}, O) atoms of the two structures is shown in Figure 5 (top). In addition to the close match found for the polypeptide backbone, amino acid side chains within the hydrophobic core are packed in a nearly identical

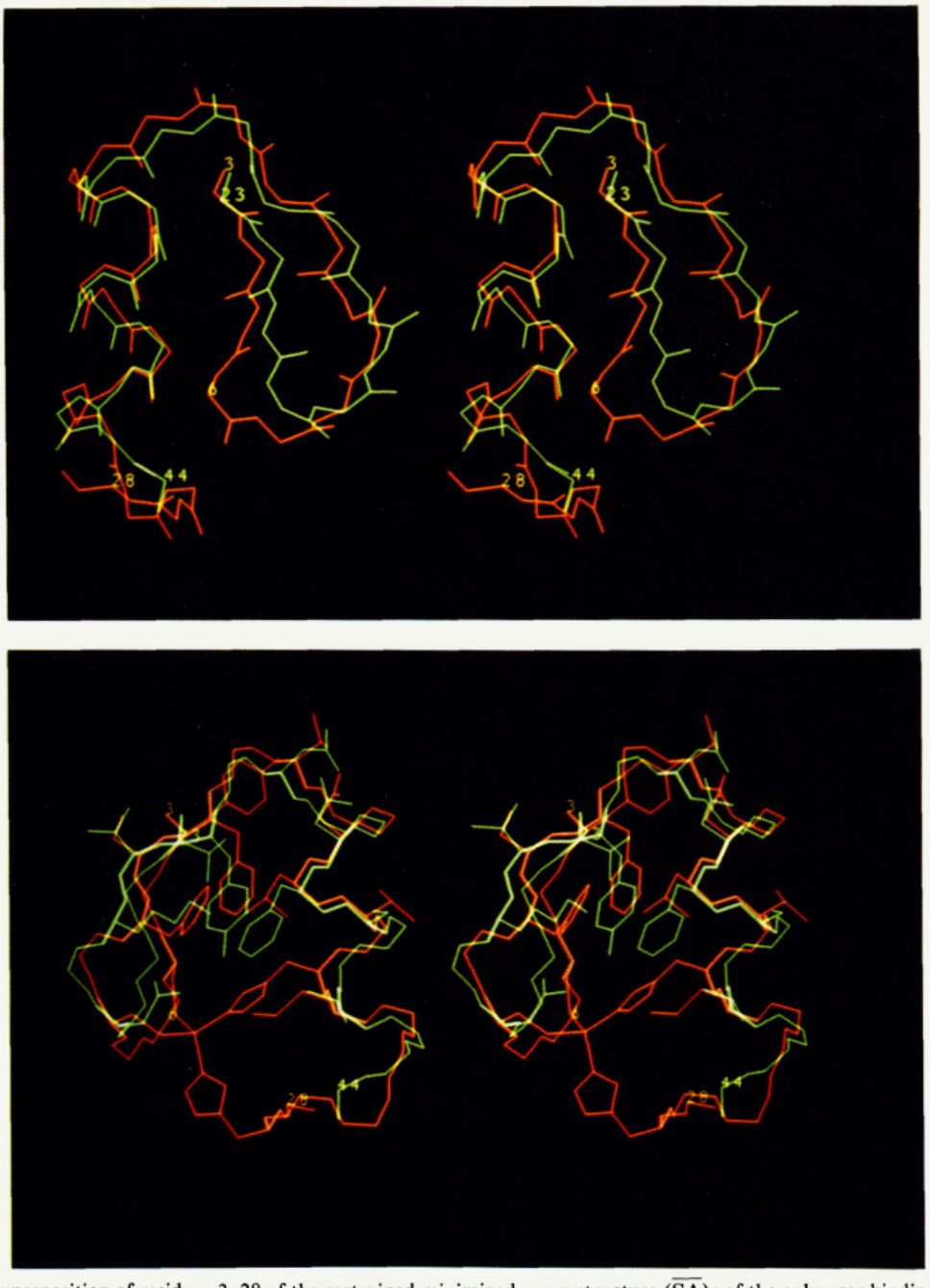

FIGURE 5: Best fit superposition of residues 3–28 of the restrained minimized mean structure $(\overline{SA})r$ of the enhancer binding protein zinc finger (red) and the X-ray structure of residues 23–24 of ovomucoid third domain (green). (Top) N, C^{α}, C, and O backbone atoms only; (bottom) N, C^{α}, and C backbone atoms together with a number of side chains. The backbone atomic rms difference for the best fit superposition of residues 3–25 (excluding residue 6) of the zinc finger and residues 23–44 of ovomucoid third domain is 0.95 Å.

arrangement as shown in Figure 5 (bottom). Thus, Phe-10, Phe-12, and Leu-18 in the enhancer binding protein zinc finger show the same orientation toward each other as the aliphatic part of the side chain of Lys-29 and the aromatic side chains of Tyr-31 and Phe-37 in ovomucoid third domain. In addition, several other side chains exhibit nearly identical conformations, in particular the pairs Ser-11/Thr-30, Thr-14/Asn-33, Asn-17/Asn-36, and Thr-19/Cys-38. Thus the $O^{\gamma 1}$ atom of Ser-11 superimposes perfectly onto the $O^{\gamma 1}$ atom of Thr-30; similarly, the $O^{\gamma 1}$ atom of Thr-19 matches the position of the S^{γ} atom of Cys-38. The similar orientation of the Thr-14/Asn-33 pair and the Asn-17/Asn-36 pair goes hand in hand with the formation of a side-chain–side-chain or side-chain–main-chain hydrogen bonds. In the case of the enhancer binding protein this involves the $N^{\delta 2}$ amino proton of Asn-17 and the $O^{\gamma 1}$ oxygen of Thr-14, whereas in ovomucoid third domain the $O^{\delta 1}$ oxygen atoms of Asn-33 and Asn-36 accept hydrogen bonds from the backbone NH of Asn-36 and the hydroxyl group of Ser-32, respectively.

This finding of close structural homology between two proteins of very different function supports the notion that tertiary folding motifs can be categorized into a limited number of classes and that these represent building blocks from which complete three-dimensional protein structures can be constructed, irrespective of the particular function of the protein. The core of the $\beta\beta\alpha$ folding motif found in both structures is formed by a set of hydrophobic side chains assembled into a tight cluster around which the polypeptide arranges itself. Even though there is no significant sequence homology, it is clear that the information content of the side chains (appropriate spacing of hydrophobic and hydrophilic side chains, relative sizes, etc.) is sufficient to specify both the formation

and relative orientations of the α-helix and the two antiparallel β-strands. Thus, although clearly stabilizing, it may well be that the complexation of the zinc atom by the two cysteine and two histidine side chains in the zinc fingers is not an essential feature of this folding motif, but could equally well be replaced by appropriate side-chain–side-chain van der Waals, hydrogen-bonding, and/or electrostatic interactions. For example, a hydrogen bond between the hydroxyl group of Tyr-31 and the $O^{\delta 1}$ atom of Asp-27 in ovomucoid third domain serves to orient the β-hairpin (residues 25–29) with respect to the long axis of the α-helix.

It is therefore a distinct possibility that the zinc finger structure does not represent a new and unique fold, but one which most likely can be found to a greater or lesser approximation in a variety of other proteins, irrespective of whether they exhibit DNA binding properties or not.

Acknowledgments

We are grateful to Dr. Paul Driscoll for help and stimulating discussions and thank Dr. Lewis Kay for assistance in measuring the coupling constants.

Supplementary Material Available

The coordinates of the 40 calculated SA structures as well as those of the restrained minimized average structure $(\overline{SA})r$, together with the complete list of experimental restraints, have been deposited in the Brookhaven Protein Data Bank.

References

Bax, A. (1989) *Methods Enzymol. 176*, 151–158.

Bax, A., Sklenar, V., Clore, G. M., & Gronenborn, A. M. (1987) *J. Am. Chem. Soc. 109*, 6511–6512.

Berg, J. M. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 91–102.

Blumberg, H., Eisen, A., Sledziewski, A., Bader, D., & Young, E. T. (1987) *Nature 328*, 443–445.

Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., & Karplus, M. (1983) *J. Comput. Chem. 4*, 187–217.

Brünger, A. T. (1988) *XPLOR Manual*, Yale University, New Haven, CT.

Clore, G. M., & Gronenborn, A. M. (1987) *Protein Eng. 1*, 275–288.

Clore, G. M., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1985) *J. Mol. Biol. 186*, 435–455.

Clore, G. M., Niles, M., Sukumaran, D. K., Brünger, A. T., Karplus, M., & Gronenborn, A. M. (1986) *EMBO J. 5*, 2729–2735.

Clore, G. M., Gronenborn, A. M., Nilges, M., & Ryan, C. A. (1987) *Biochemistry 26*, 8012–8023.

Davis, D. G., & Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2821–2822.

Deisenhofer, J., & Steigemann, W. (1975) *Acta Crystallogr., Sect. B 31*, 238–250.

Diakun, G. P., Fairall, L., & Klug, A. (1986) *Nature 324*, 698–699.

Eklund, H., & Bränden, C.-I. (1983) in *Zinc Enzymes* (Spiro, T. G., Ed.) John Wiley, New York.

Evans, R. M., & Hollenberg, S. M. (1988) *Cell 52*, 1–3.

Fang, C.-H., & Maniatis, T. (1990) *Genes Dev. 4*, 29–42.

Forman-Kay, J. D., Gronenborn, A. M., Kay, L. E., Wingfield, P. T., & Clore, G. M. (1990) *Biochemistry 29*, 1566–1572.

Freedman, L. P., Luisi, B. F., Korscun, Z. R., Basavappa, R., Sigler, P. B., & Yamamoto, K. R. (1988) *Nature 334*, 543–545.

Havel, T. F. (1986) *DISGEO*, Quantum Chemistry Program Exchange No. 507, Indiana University, Bloomington, IN.

Havel, T., & Wüthrich, K. (1984) *Bull. Math. Biol. 46*, 673–698.

Havel, T. F., Kuntz, I. D., & Crippen, G. M. (1983) *Bull. Math. Biol. 45*, 665–720.

Gibson, T. J., Postma, J. P. M., Brown, R. S., & Agros, P. (1988) *Protein Eng. 2*, 209–219.

Guss, J. M., & Freeman, H. C. (1983) *J. Mol. Biol. 169*, 521–563.

Hendrickson, W. A., & Teeter, M. M. (1981) *Nature 290*, 107–112.

James, M. N. G., & Sielecki, A. R. (1983) *J. Mol. Biol. 163*, 299–361.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.

Jones, T. A., & Thirup, S. (1986) *EMBO J. 5*, 819–822.

Kittel, C. (1966) *Introduction to solid states physics*, 3rd ed., John Wiley, New York.

Klevit, R. E., Herriott, J. R., & Horvath, S. J. (1990) *Proteins: Struct. Funct., Genet. 7*, 215–226.

Klug, A., & Rhodes, D. (1987) *Trends Biochem. Sci. 12*, 464–469.

Kraulis, P. J., Clore, G. M., Nilges, M., Jones, T. A., Pettersson, G., Knowles, J., & Gronenborn, A. M. (1989) *Biochemistry 28*, 7241–7257.

Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A., & Wright, P. E. (1989) *Science 245*, 635–637.

Maekawa, T., Sakura, H., Sudo, T., & Ishii, S. (1989) *J. Biol. Chem. 264*, 14951–14593.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.

Marion, D., & Bax, A. (1988) *J. Magn. Reson. 80*, 528–533.

Matsuura, Y., Takano, T., & Dickerson, R. E. (1982) *J. Mol. Biol. 156*, 389–409.

Merrifield, R. B., Vizili, L. D., & Boman, H. G. (1982) *Biochemistry 21*, 5020–5031.

Miller, J., McLachlan, A. D., & Klug, A. (1985) *EMBO J. 4*, 1609–1614.

Mueller, L. (1987) *J. Magn. Reson. 72*, 191–196.

Nagai, K., Nakaseko, Y., Nasmyth, K., & Rhodes, D. (1988) *Nature 332*, 284–286.

Nilges, M., Gronenborn, A. M., & Clore, G. M. (1988) *FEBS Lett. 229*, 317–324.

Nilges, M., Clore, G. M., & Gronenborn, A. M. (1990) *Biopolymers 29*, 813–822.

Papamokos, E., Weber, E., Bode, W., Huber, R., Empie, M. W., Kato, I., & Laskowski, M., Jr. (1982) *J. Mol. Biol. 158*, 515–537.

Pauling, L. (1967) *The nature of the chemical bond*, Oxford University Press, Cambridge.

Pelham, H. R. B., & Brown, D. D. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 4170–4175.

Phillips, S. E. V. (1980) *J. Mol. Biol. 142*, 531–554.

Plateau, P., & Guêron, M. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.

Singh, H., LeBowitz, J. H., Baldwin, A. S., & Sharp, P. A. (1988) *Cell 52*, 415–423.

Steigemann, W., & Weber, E. (1979) *J. Mol. Biol. 127*, 309–339.

Stillman, D. J., Bankier, A. T., Seddon, A., Groenhout, E. G., & Nasmyth, K. A. (1988) *EMBO J. 7*, 485–494.

Summers, M. F., South, T. L., Kim, B., & Hare, D. R. (1990) *Biochemistry 29*, 329-340.
Tam, J. P., Heath, W. F., & Merrifield, R. B. (1983) *J. Am. Chem. Soc. 105*, 6442-6455.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949-961.

Fluorescence Investigation of the Sex Steroid Binding Protein of Rabbit Serum: Steroid Binding and Subunit Dissociation†

Emanuela Casali,‡ Philip H. Petra,§,‖ and J. B. Alexander Ross*,‡
Department of Biochemistry, Mount Sinai School of Medicine, One Gustave L. Levy Place, New York, New York 10029, and Departments of Obstetrics and Gynecology and of Biochemistry, University of Washington, Seattle, Washington 98195
Received February 13, 1990; Revised Manuscript Received June 15, 1990

ABSTRACT: The relationship between steroid binding and protein subunit interactions of rabbit sex steroid binding protein (rSBP) has been studied by steady-state and time-resolved fluorescence spectroscopy. The high-affinity ($K_a \sim 10^8$ M^{-1} at 4 °C), fluorescent estrogen *d*-1,3,5(10),6,8-estrapentaene-3,17β-diol [dihydroequilenin (DHE)] was used as a fluorescent probe of the steroid-binding site. Perturbation of the binding site with guanidinium chloride (Gdm·Cl) was monitored by changes in the steady-state fluorescence anisotropy of DHE as well as by changes in fluorescence quenching of DHE with acrylamide. The results of acrylamide quenching at 11 °C show that, while between 0 and 1 M Gdm·Cl the steroid-binding site is completely shielded from bulk solvent, there is decreased DHE binding. To study the subunit–subunit interactions, rSBP was covalently labeled with dansyl chloride in the presence of saturating 5α-dihydrotestosterone (DHT), which yielded a dansyl-conjugated protein that retained full steroid-binding activity. The protein subunit perturbation was monitored by changes in the steady-state fluorescence anisotropy of the dansyl group. At 11 °C, the dansyl anisotropy perturbation, reflecting changes in global and segmental motions of the dimer protein, occurs at concentrations of Gdm·Cl above 1 M. The Gdm·Cl titration in the presence of steroids with equilibrium association constants less than 10^8 M^{-1} shows a plateau near 3 M Gdm·Cl at 11 °C; at this Gdm·Cl concentration, no DHE is bound. No plateau is observed at 21 °C. At higher Gdm·Cl concentrations, the dansyl fluorescence anisotropy decreases further and shows no steroid dependence. Recovery of steroid-binding activity (assayed by saturation binding with [^{3}H]DHT), under renaturation conditions, is dependent on both steroid concentration and affinity. Both unlabeled and dansyl-labeled protein recovery the same amount of activity, and according to fluorescence anisotropy, dansyl-labeled rSBP re-forms a dimer upon dilution below 1 M or removal of Gdm·Cl. From the steroid requirement for recovery of steroid-binding activity, it appears that a conformational template is required for the dimeric protein to re-form a steroid-binding site with native-like properties.

The blood of most vertebrates contains a high-affinity sex steroid binding protein, generally referred to as SBP[1] (also called sex hormone binding globulin, abbreviated SHBG), which binds both estrogens and androgens [for reviews, see Westphal (1986) and Moore and Bulbrook (1988)]. The relative affinities for estrogens and androgens vary among different species; androgens are generally bound with higher affinity. For example, compared with primate SBPs, rabbit SBP has a significantly weaker affinity for 17β-estradiol, but these species have similar affinities for testosterone and 5α-dihydrotestosterone (DHT) (Mickelson & Petra, 1978; Kotite & Muto, 1982). It has been proposed that SBP is a blood carrier for the sex steroids and that its biological function is to maintain the proper balance in the free steroid concentrations. This hypothesis is consistent with the observation that increased SBP concentrations reduce the clearance rate of specific steroids in the blood (Vermeulen et al., 1969; Petra et al., 1985; Plymate et al., 1990). This hypothesis is also consistent with the view held by many endocrinologists that steroids enter target cells by free diffusion. Recent data from several laboratories, however, have provided suggestive evidence for the existence of specific SBP receptors on the membranes of target cells (Sakiyama et al., 1988; Hryb et al., 1985). These data could mean that the biological function of SBP involves more than control of the free steroid concentration in plasma and that an additional function of SBP is to target the bound steroid to specific cells.

Our laboratories have been investigating the mechanism of steroid binding by rabbit and human SBP in an effort to understand the basis for their differences in specificity and affinity for the different sex steroids (Örstan et al., 1986).

†Supported by National Institutes of Health Grants HD-17542 and GM-39750 (J.B.A.R.), by National Science Foundation Grant DMB-8516318 (J.B.A.R.), and in part by National Institutes of Health Grant HD-13956 (P.H.P.). Preliminary aspects of this work were presented at the Joint Meeting of the American Society for Biochemistry and Molecular Biology and the American Society for Cell Biology, San Francisco, CA, Jan 29 to Feb 2, 1989.
*Author to whom correspondence should be addressed.
‡Department of Biochemistry, Mount Sinai School of Medicine.
§Department of Obstetrics and Gynecology, University of Washington.
‖Department of Biochemistry, University of Washington.

[1] Abbreviations: SBP, sex steroid binding protein; rSBP, rabbit SBP; hSBP, human SBP; DHT, 5α-dihydrotestosterone; DHE, 17β-dihydroequilenin; DNS, dansyl or 5-(dimethylamino)naphthalene-1-sulfonyl; Gdm·Cl, guanidinium chloride.

0006-2960/90/0429-9334$02.50/0 © 1990 American Chemical Society