Preorganization of molecular binding sites in designed diiron proteins

Ornella Maglio*, Flavia Nastri*, Vincenzo Pavone*, Angela Lombardi*†, and William F. DeGrado†‡

*Department of Chemistry, University of Napoli Federico II, Complesso Universitario Monte S. Angelo, I-80126 Napoli, Italy; and ‡Department of Biochemistry and Biophysics, School of Medicine, University of Pennsylvania, Philadelphia, PA 19104-6059

Contributed by William F. DeGrado, February 6, 2003

***De novo* protein design provides an attractive approach to critically test the features that are required for metalloprotein structure and function. Previously we designed and crystallographically characterized an idealized dimeric model for the four-helix bundle class of diiron and dimanganese proteins [Dueferri 1 (DF1)]. Although the protein bound metal ions in the expected manner, access to its active site was blocked by large bulky hydrophobic residues. Subsequently, a substrate-access channel was introduced proximal to the metal-binding center, resulting in a protein with properties more closely resembling those of natural enzymes. Here we delineate the energetic and structural consequences associated with the introduction of these binding sites. To determine the extent to which the binding site was preorganized in the absence of metal ions, the apo structure of DF1 in solution was solved by NMR and compared with the crystal structure of the di-Zn(II) derivative. The overall fold of the apo protein was highly similar to that of the di-Zn(II) derivative, although there was a rotation of one of the helices. We also examined the thermodynamic consequences associated with building a small molecule-binding site within the protein. The protein exists in an equilibrium between folded dimers and unfolded monomers. DF1 is a highly stable protein (K_{diss} = 0.001 fM), but the dissociation constant increases to 0.6 nM ($\Delta\Delta G$ = 5.4 kcal/mol monomer) as the active-site cavity is increased to accommodate small molecules.**

The requirements for protein stability versus function are often diametrically opposed. The folded conformations of proteins are stabilized by maximizing the burial of hydrophobic groups, minimizing voids, and forming intramolecular hydrogen bonds (1, 2). In contrast, binding and enzymatic functions generally require active-site clefts replete with solvent-exposed hydrophobic groups and hydrogen-bonding groups, which are essential for proper binding of substrates. This tradeoff between conformational stability and function presents a particularly large challenge to protein design (3), forcing the designer to skirt the waters between Scylla and Charybdis.

The stability/function tradeoff is particularly apparent in metalloproteins. Structural metal-binding sites in proteins frequently achieve stability by binding metal ions in coordinately saturated ligand environments with idealized ligand–metal bond geometries (4, 5). However, functional sites in metalloenzymes frequently contain coordinately unsaturated metal ions that are positioned appropriately for binding substrates; they also sometimes display geometries not frequently observed in simple, small-molecule metal–ligand complexes (4, 5). Further, the active sites of metalloproteins are frequently preorganized in the absence of metal ions, which requires the burial of polar ligands at the expense of folding free energy (6). The preorganization imparts tight and geometrically specific metal binding by assuring that the protein imparts its own structural preferences on the metal rather than vice versa.

Here we trace the energetic consequences of introducing an active-site access channel into a designed diiron protein, DF1 (Dueferri 1) (7–10). This protein is a small, 48-residue homodimeric model for the family of O_2-using diiron proteins that includes methane monooxygenase and the radical-forming R2 subunit of the ribonucleotide reductase from *Escherichia coli* (11–14). The DF1 dimer comprises two noncovalently associated helix–loop–helix motifs that bind the diiron cofactor near the center of the protein (8). The crystallographic structures of di-Zn(II) (8) and di-Mn(II) DF1 (unpublished results) are very similar to the di-Mn(II) and di-Fe(II) forms of other natural diiron proteins (15–18). In DF1, two Glu side chains bridge both metal ions, whereas the other two carboxylates interact with a single metal ion in a bidentate, chelating interaction (Fig. 1*a*). The two His residues form additional monodentate ligands.

Initially, DF1 was designed for maximal stability. Thus, the interior of the protein was efficiently packed with a large number of hydrophobic side chains, resulting in a high degree of conformational stability (8). Nevertheless, access to its dimetal center was blocked by a pair of symmetrically related hydrophobic Leu residues (Leu-13 and Leu-13′, Fig. 1*a*). Modeling suggested that replacement of these side chains with smaller Ala or Gly residues might open up the active site, providing a deep invagination with the metal-binding site occupying the bottom of the active-site pocket. Indeed the L13G and L13A variants of DF1 showed properties similar to those of the natural proteins. Both bound exogenous ligands in solution and catalyzed the ferroxidase reaction (unpublished results). The crystal structures of the di-Mn(II) forms of L13A and L13G have been refined to high resolution (9, 10). The L13A mutant bound a DMSO molecule with its sulfoxide oxygen bridging the two metal ions (Fig. 1*b*; ref. 9), whereas the more polar cavity of the L13G mutant engendered binding of both bridging and terminal water molecules (or hydroxide ions) (10) in manners that mimic intermediates in the mechanism of a peroxide reductase (19). Thus, as the size of the active-site cavity increased, the chemical properties of the protein increasingly resembled those of natural enzymes.

The goal of this article is to address the stability/function tradeoffs associated with the design of a metalloprotein. First, we examine the thermodynamics of folding of DF1, L13A, and L13G to determine the thermodynamic cost of carving an active-site access channel within a protein. Next, we ask what dictates the fold of DF1 by determining the structure of the protein in the apo state.

Abbreviations: DF1, Dueferri 1; Gdn, guanidine; NOE, nuclear Overhauser effect; rmsd, rms deviation.

Atomic coordinates for the best 14 structures, resonance assignments, and restraints used in the structure determination have been deposited in the Protein Data Bank, www.rcsb.org (PDB ID codes 1NVO).

†To whom correspondence may be addressed at: Department of Chemistry, University of Napoli Federico II, Complesso Universitario Monte S. Angelo, Via Cynthia, I-80126 Napoli, Italy. E-mail: lombardi@chemistry.unina.it (A.L.); or Department of Biochemistry and Biophysics, 1009B Stellar-Chance Building, University of Pennsylvania, Philadelphia, PA 19104-6059. E-mail: wdegrado@mail.med.upenn.edu (W.F.D.).

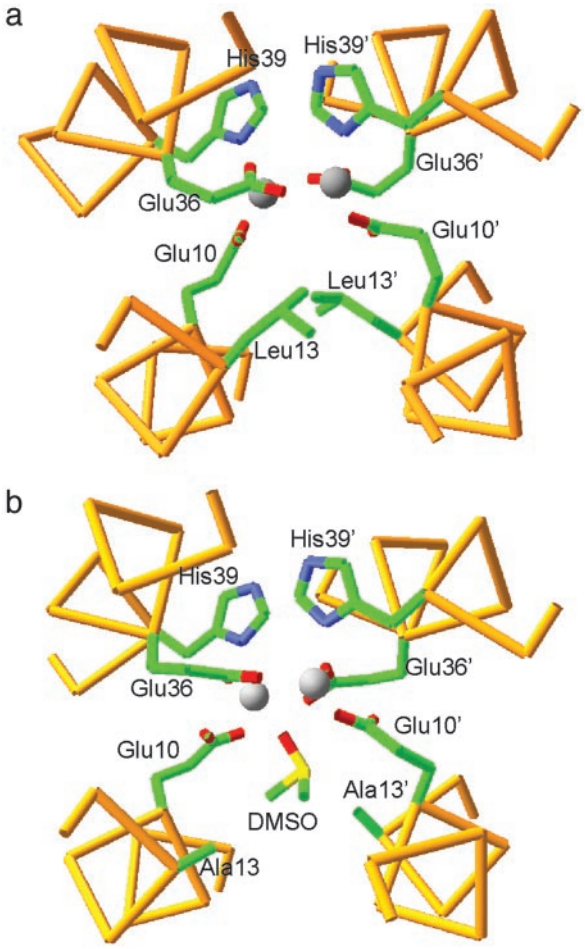

Fig. 1. (*a*) The structure of di-Zn(II) DF1: the Leu-13 and Leu-13′ residues block substrate access to the active site. The backbone trace of residues 7–15 and 33–42 is shown, along with Glu-10, Leu-13, Glu-36, and His-39 side chains. (*b*) The active site of L13A DF1. The decreased steric bulk of Leu-13 and Leu-13′ results in the formation of an active-site pocket, which accommodates a DMSO molecule, the oxygen atom of which bridges between the Mn(II) ions.

Materials and Methods

Guanidine (Gdn) Denaturation. Gdn·HCl denaturation curves were obtained by monitoring the circular dichroism (CD) at 222 nm by using a Jasco (Tokyo) J-715 dichrograph with a 60-s averaging time, and the data were analyzed as described in ref. 20 and in *Supporting Text*, which is published as supporting information on the PNAS web site, www.pnas.org.

NMR Spectroscopy and Structure Calculations. The solution conditions for NMR experiments were 1.0 mM protein concentration in 90% H_2O/10% DMSO d_6, pH 4.0. Chemical shifts were referenced to the residual protonated DMSO signal, defining it as 2.49 ppm with respect to tetramethylsilane (Table 3, which is published as supporting information on the PNAS web site). All NMR spectra were carried out at 298 K on a Bruker (Rheinstetten, Germany) Avance 600 spectrometer operating at a nominal frequency of 600.13 MHz. To obtain ^{1}H resonance assignments, a double quantum filtered correlated spectroscopy (21) spectrum, a clean-total correlation spectroscopy (22) spectrum with a mixing time of 60 ms, and two NOESY (23) spectra with mixing times of 120 and 150 ms were recorded by using standard pulse sequences and phase cycling as described in *Supporting Text*. Data were processed on an SGI Octane workstation with the program NMRPIPE (24). Peak picking in the NMR spectra, spin system identification, and volume integration of the NOESY cross peaks were performed with the interactive program XEASY (25).

Nuclear Overhauser effect (NOE) restraints were derived from the 120-ms NOESY spectrum. Upper distance limits were obtained by using the program CALIBA (26). The structure calculations were performed with the program DYANA 1.5 (27) by using the torsion angle dynamic strategy. A total of 769 distance constraints derived from NOESY spectrum were used in DYANA calculations.

To check whether DF1 existed as a symmetric homodimer in solution, preliminary DYANA runs were performed by interpreting all the distance restraints derived from the NOESY spectrum as intramonomer correlations. The structures showed a large number of distance restraints violated (185), and the average residual target function for the best 20 conformers was 42.5 ± 1.8 Å^2. Thus, some of the long-range connectivities had to be considered as intermonomer correlations. We interpreted all NOEs conservatively, and only those that were well resolved and clearly inconsistent with the designed and crystal structures of the monomer were assigned as intermonomer contacts.

Several cross peaks, identified as long-range intramonomer correlations, e.g., between residues 3 and 46 or 17 and 29, indicated an antiparallel alignment of the two helices and confirmed the helix–loop–helix motif characterizing the monomer. The long-range intermonomer correlations, which occurred between 14 residues, allowed definition of the tertiary fold of DF1 as an antiparallel four-helix bundle with an up-down-up-down topology as in the intended design. The long-range intermonomer correlations were confirmed by the fact that most of these contacts involved residues near either the N terminus or C terminus of the protein with those close to the loop region. A few cross peaks, e.g., between Tyr-2 and Ala-20 and Tyr-42 and Val-28, indicated an antiparallel alignment of the two monomers and were classified as arising from intermonomer interactions. In this manner, 30 intramonomer long-range and 45 intermonomer NOEs were assigned.

A total of 1,518 NOE constraints were assigned and integrated and then transformed into upper distance limits by using the program CALIBA (26). Of 1,518 constraints, 769 (16.5 NOEs per residue; 467 intraresidue, 137 sequential, 90 medium-range, and 75 long-range NOEs) were meaningful and therefore have been taken into account by DYANA. To calculate the dimeric structure, two noncovalently linked monomers were generated by using 12 "invisible" linkers consisting of pseudoatoms, which can penetrate the "real" molecule without steric repulsion or target-function violation. Typical DYANA runs were performed on 200 randomly generated starting structures with 15,000 torsion angle dynamics steps. The 40 DYANA structures with the lowest target-function values were subjected to restrained energy minimization by using the SANDER module of the AMBER 7.0 package (3,000 steps of conjugate gradient) (28). The 1991 version of the all-atom force field was used (29, 30), with a distance-dependent dielectric constant $\varepsilon = r_{ij}$. To reduce the artifacts that can arise during *in vacuo* simulation, the charge of the ionizable groups was reduced to 20% of its full value. Distance restraints were applied as a flat well, with a parabolic penalty within 0.5 Å outside the upper bond and a linear function beyond 0.5 Å, by using a force constant of 32 kcal/mol·Å^2. The best 14 structures among those with a residual restraint energy lower than −2,720 kcal/mol and with a residual distance-constraint energy <63 kcal/mol were selected to represent the solution structure. The structure analysis, in terms of Ramachandran plots (Fig. 6, which is published as

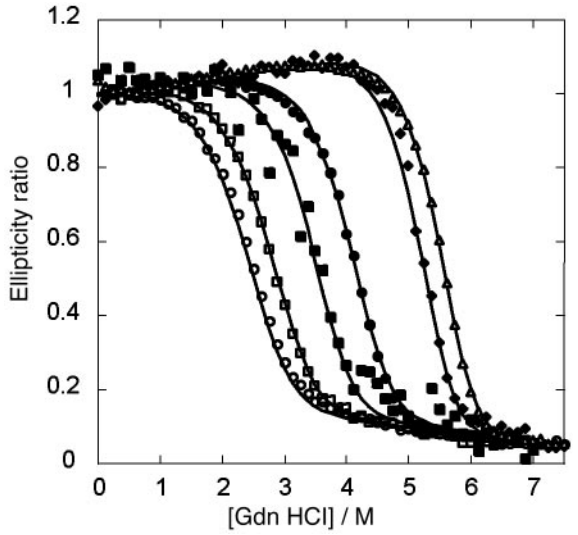

Fig. 2. Gdn denaturation curves of DF1, L13A, and L13G. The ellipticity at 222 nm was monitored as a function of the concentration of added denaturant in 10 mM phosphate buffer (pH 5.5). The smooth curves are generated by globally fitting the free energy of dimerization (ΔG_U) in the absence of Gdn, m ($\delta\Delta G_U/\delta$[Gdn]) and the baseline parameters to the data as described in *Supporting Text*. The identity and concentrations of the proteins are as follows: ○, L13G-DF1 3.76 μM; □, L13G-DF1 16 μM; ■, L13A-DF1 0.9 μM; ●, L13A-DF1 12 μM; ♦, DF1 0.78 μM; △, DF1 3.7 μM.

supporting information on the PNAS web site), deviation from ideal structural parameters, secondary elements, was performed with the PROCHECK-NMR program (see Table 2; ref. 31).

Results

Gdn·HCl-Induced Unfolding Curves. The thermodynamic stabilities of DF1, L13A-DF1, and L13G-DF1 were determined by monitoring the loss of the helical CD signal at 222 nm as a function of Gdn·HCl. The far-UV CD spectra of DF1 and its analogs L13A-DF1 and L13G-DF1 are very similar (data not shown) and show minima at 210 nm ($\approx -20{,}000$ deg·cm^2/dmol) and 222 nm ($\approx -22{,}000$ deg·cm^2/dmol) as expected for helical proteins.

The Gdn denaturation curves for DF1, L13A, and L13G depend markedly on the total protein concentration (Fig. 2). For a monomer–dimer equilibrium, the law of mass action dictates that the stability should increase with concentration, which is in fact observed for all three proteins. The free energies of dimerization of the three proteins were determined by globally fitting the baseline and thermodynamic parameters to an equilibrium between folded monomers and unfolded dimers (20).

This analysis revealed that DF1 has exceptional stability, with a dissociation constant of $\approx$0.001 fM; the corresponding values for L13A and L13G were 100 fM and 0.6 nM for L13G, respectively (Table 1). The difference in stability associated with mutation of Leu-13 to Ala [$\Delta\Delta G$ (0)] is 5.6 kcal/mol dimer (2.8 kcal/mol monomer). This value is within the range expected for mutating a single buried Leu to an Ala in a native protein (32, 33). The mutation of position 13 from Ala to Gly results in a further destabilization of 5.2 kcal/mol (2.6 kcal/mol monomer), reflecting contributions from both a decrease in the helix propensity of Ala relative to Gly [$\approx$1 kcal/mol monomer (34, 35)] and a decreased hydrophobic driving force [$\approx$1.3 kcal/mol monomer (33)]. The m values decrease as the bulk of the residue at position decreases from Leu to Ala to Gly, as expected from the decrease in the extent of solvent-accessible surface area buried after folding (36).

Table 1. Globally fit parameters characterizing the Gdn-induced unfolding curves for variants of DF1

Protein	$\Delta G°$, kcal/mol	m, kcal·mol^{-1}·M^{-1}	K_{diss}, M
DF1	−23.5 ± 0.3	2.93 ± 0.1	(9.0) 10^{-18}
L13A-DF1	−17.9 ± 0.3	2.49 ± 0.1	(1.1) 10^{-13}
L13G-DF1	−12.7 ± 0.3	2.25 ± 0.1	(6.4) 10^{-10}

Structure Determination by ^{1}H NMR. The ^{1}H NMR spectra of apo-DF1 (Fig. 3) are very well dispersed in both the amide and the aliphatic regions and display narrow line widths, suggesting that DF1 folds into an unique native-like structure. Further, a single set of resonances is observed for each residue, which is consistent with the formation of a symmetrical dimer. The structure of apo-DF1 was solved by two-dimensional NMR (37) under conditions (pH 4.0, 10% DMSO) similar to those used to crystallize the protein (8). Scalar and dipolar connectivities were obtained from double quantum filtered correlated spectroscopy (21), total correlation spectroscopy (22), and NOESY (23) spectra. Main-chain assignments were obtained for all residues (except the C-terminal Leu-47 and Gly-48), and side-chain proton assignments are complete except for Leu-7 and Leu-11.

Preliminary secondary structural assignments were obtained from amide and α-proton chemical shifts, $^3J_{\alpha N}$ coupling constants, and medium-range NOEs (Figs. 7 and 8, which are published as supporting information on the PNAS web site). The good correlation between all of the experimental data defined two helices spanning residues 7–22 and 27–46, connected by a nonhelical loop, as designed.

A total of 75 long-range NOEs were assigned. NOEs between residues at the N or C terminus of the protein with those close to the loop regions are consistent with the formation of an antiparallel dimer structure in solution, as expected. For a symmetric dimer the observed long-range NOE contacts can be either inter- or intramolecular in origin. This problem was considerably simplified because NOEs were observed between protons near the N and C termini of residues within the same helix for both helices 1 and 2. These contacts, which must arise from intermolecular antiparallel interhelical interactions, agreed well with those defined by the original model as well as the crystallographic structure (8). Given this agreement, we felt justified in using the crystallographic structure to tentatively assign additional ambiguous long-range NOEs.

The structure calculation from the NMR data proceeded via two steps. First, the monomer was solved by using only the intramonomer interactions. One hundred structures were calculated, and all converged to the same overall helix–loop–helix topology. The low target-function average value (0.49 ± 0.13 Å) and the small number of residual violations (12) indicate that the constraints are well satisfied by the structures obtained.

Calculations of the DF1 dimer structure included a total of 769 intra- and intermonomer NOE correlations consisting of 467 intraresidue, 137 sequential, 90 medium-range, 45 intermonomer long-range, and 30 intramonomer long-range NOE correlations. All DYANA structures converged to the same topology, with no residual constraint violations and a good average value of the target function (0.78 ± 0.09 Å^2 for the best 40 structures) (Fig. 9, which is published as supporting information on the PNAS web site). The 40 best DYANA structures were refined further by restrained energy minimization. The

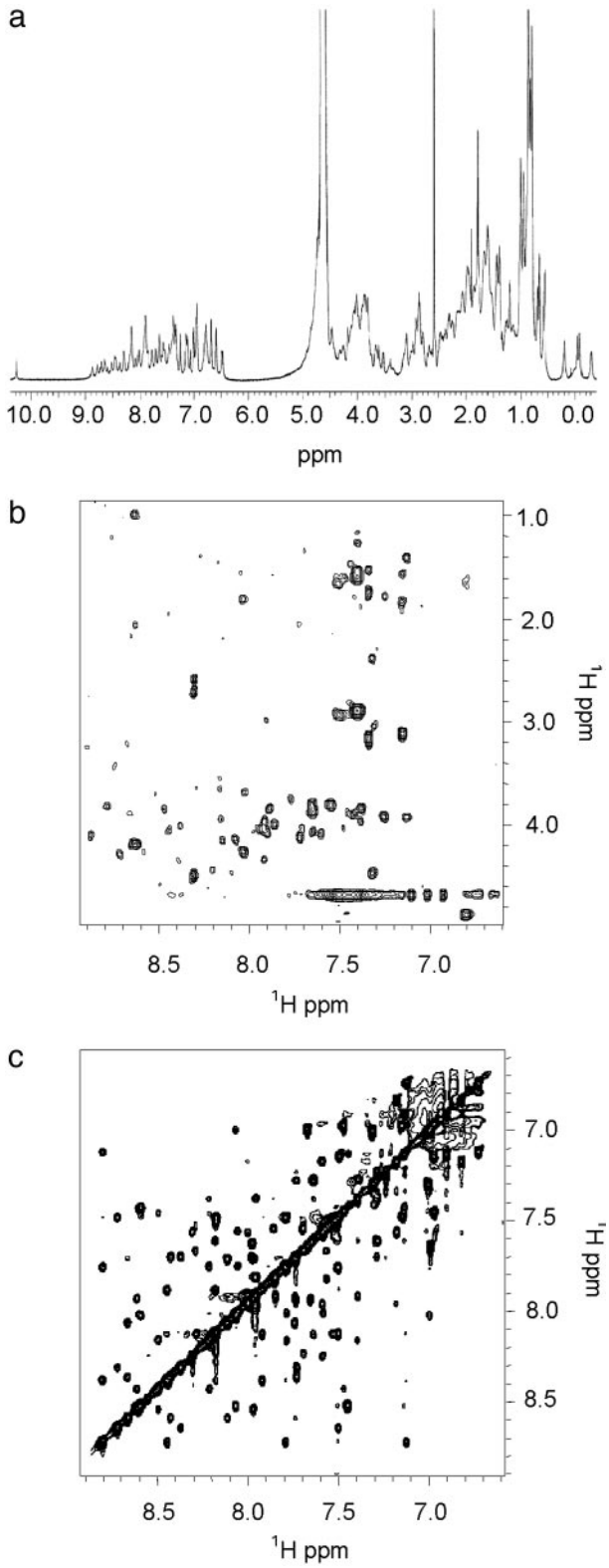

Fig. 3. (*a*) One-dimensional NMR spectrum of DF1 [H_2O/DMSO 90:10 (vol/vol), pH 4.0] at 600 MHz and 298 K. (*b*) Total correlation spectroscopy in the NH-aliphatic region. Of 47 possible NH–$^{\alpha}$CH cross peaks, 45 are observed. (*c*) NOESY spectrum (120-ms mixing time) in the NH–NH region.

best 14 refined structures (Table 2), with no violations in interproton distances >0.5 Å, were selected to represent the solution structure (Fig. 4). The overall stereochemical quality of the set of structures, assessed by PROCHECK-NMR (31), indicates that the DF1 dimer structure is comparable to that of a 1.9-Å well defined x-ray structure.

The solution conformation of the dimer is well defined except for the N- and C-terminal regions (Asp-1–Leu-7 and Ile-46–Gly-48). The structural disorder in 1–7 appears to be intrinsic to the protein, whereas the structure of 46–48 was underdetermined because of a paucity of NOEs. Superposi-

Table 2. Summary of the number of constraints used in the calculations and structural statistics for the 14 final NMR structures of apo-DF1

Parameter	Value
Assigned NOE cross peaks	1,518
Nonredundant NOE upper distance limits	769
Intraresidue	467
Interresidue sequential ($\|i-j\| = 1$)	137
Interresidue medium range ($1 < \|i-j\| \leq 1$)	90
Interresidue long range ($\|i-j\| > 5$)	75
Residual target function,* Å^2	0.75 ± 0.09
Residual distance constraint violations, Å	
$0.1 < d \leq 0.2$	33.5 ± 3.7
$0.2 < d \leq 0.3$	4.8 ± 1.6
$0.3 < d \leq 0.4$	1.3 ± 0.9
$0.4 < d \leq 0.5$	0.7 ± 0.8
Maximum violation	0.37 ± 0.08
AMBER energies, kcal·mol^{-1}	
Distance constraint	49.6 ± 7.3
van der Waals	−535.5 ± 11.6
Total	−2,878.1 ± 80.2
rmsd to mean coordinates,† Å	
N, C$^{\alpha}$, C′ (8–44)	0.47 ± 0.06
All heavy atoms (8–44)	0.90 ± 0.09
Ramachandran statistics from PROCHECK-NMR‡	
Most favored regions, %	85.2
Additional allowed regions, %	13.3
Generously allowed regions, %	1.3
Disallowed regions, %	0.2

*Residual target function value for the best 40 DYANA conformers before energy minimization.
†Average coordinates of the 14 energy-minimized conformers after superposition for the best fit of the atoms of the residues indicated in parentheses.
‡The program PROCHECK-NMR was used to check the overall quality of the structures.

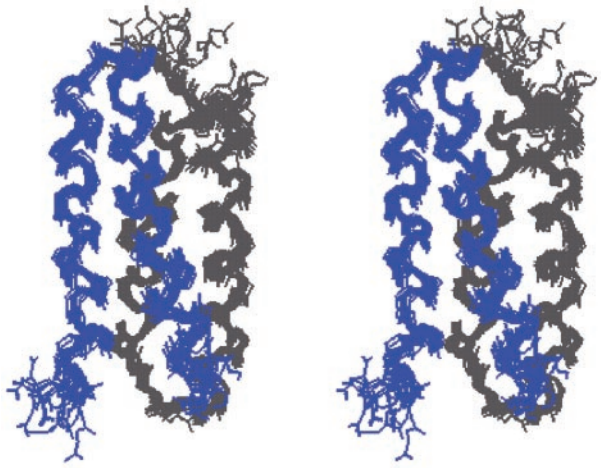

Fig. 4. Stereoview of the superposition of the best 14 minimized structures for apo-DF1.

tion of the best 14 calculated structures on the mean coordinates gave an average rms deviation (rmsd) of 0.47 ± 0.06 Å for all backbone atoms and 0.90 ± 0.09 Å for all heavy atoms in the well defined regions (8–44). This deviation is similar to that observed between dimers within the asymmetric unit of DF1 in the crystal structure of di-Zn(II) DF1.

Description of the Structure. The solution structure of apo-DF1 is highly similar to the crystal structure of the di-Zn(II) derivative (8) as well as the original model for the dimetal form of the protein. The overall rmsd for the backbone atoms of the best 14 structures of DF1 (residues 8–22 and 27–44) versus the protein model was 1.53 ± 0.05 Å. The protein consists of a pair of symmetrical helix–loop–helix motifs with the loop on opposite sides of the bundle. The helices within each monomer span residues 8–22 and 27–44 as defined by the pattern of hydrogen bonding as well as the backbone torsion angles ϕ and ψ. The loop region comprises residues 24–26 and is characterized by a $\gamma\alpha_L\beta$ conformation (7, 38). In this turn, the first helix is capped in a conformation that lies between an α_L and a Schellman motif (39) in which Lys-25 adopts a left-handed helix conformation. The side chain of Leu-26 caps the hydrophobic core of the structure.

In the dimeric structure, helix 1 and helix 2 pack in an antiparallel manner against their symmetry-related helix 1′ and helix 2′, respectively. As in native proteins, the apolar side chains are located in the core of the dimeric structure and generally have unique conformations. The computed structure is also consistent with the chemical shifts observed for the interior side chains. For example, the Leu-6 (0.61 and −0.04 ppm) and Val-28 (0.74 and 0.25 ppm) methyl groups were significantly upfield-shifted via ring-current effects arising from Tyr-17 and Trp-42. Similarly, Ile-33 showed upfield-shifted methyl proton chemical shifts (0.06 and −0.26 ppm) arising from their proximity to Tyr-17 and His-39.

Discussion

One goal of this study was to determine the extent to which the protein was folded and its active site organized in the absence of metal ions. The metal-binding site was partially preorganized in the absence of metal ions, as assessed from a comparison of the geometries of the liganding residues in the di-Zn(II) versus the apo structures. Glu-10 adopts the g^+, g^+ conformation (40) in almost all crystal structures of DF1 and its mutants ($\chi_1 = -65 \pm 5°$, $\chi_2 = -78 \pm 12°$) (8–10). A g^+, g^- conformation was found only in one monomer of the asymmetric unit in di-Zn-DF1 (8). Analogously, in the solution structure of the protein, Glu-10 adopts a g^+, g^+ conformation in 90% of the structures ($\chi_1 = -69 \pm 12°$, $\chi_2 = -67 \pm 11°$), with a g^+, g^- conformer in the remaining conformers. Glu-36 is also found to adopt a single conformer with the g^+, g^+ rotamer in the crystal structures of di-Zn(II)-DF1 (8), di-Mn(II)-L13A-DF1 (9), and di-Mn(II)-L13G-DF1 (10) ($\chi_1 = -50 \pm 4°$, $\chi_2 = -57 \pm 6°$). A very similar conformation is found for this residue in solution for most of the structures of the bundle (68%, $\chi_1 = -59 \pm 3°$, $\chi_2 = -80 \pm 15°$). Also, the χ_1 torsion angle of His-39 is g^+ both in solution (−79 ± 8°) and in the crystal structures (−60 ± 4°), although the orientation of the ring is frequently flipped by a rotation about χ_2 in the solution structures.

The metal-binding site in di-Zn(II)-DF1 is stabilized by second-shell hydrogen bonds, which are largely retained in the apo structure. The primary ligand Glu-10 accepts second-shell hydrogen bonds from Tyr-17′ of a neighboring monomer in approximately half of the solution structures. Similarly, Asp-35 accepts a second-shell hydrogen-bonded interac tion from the primary ligand His-39 in half of the solution structures.

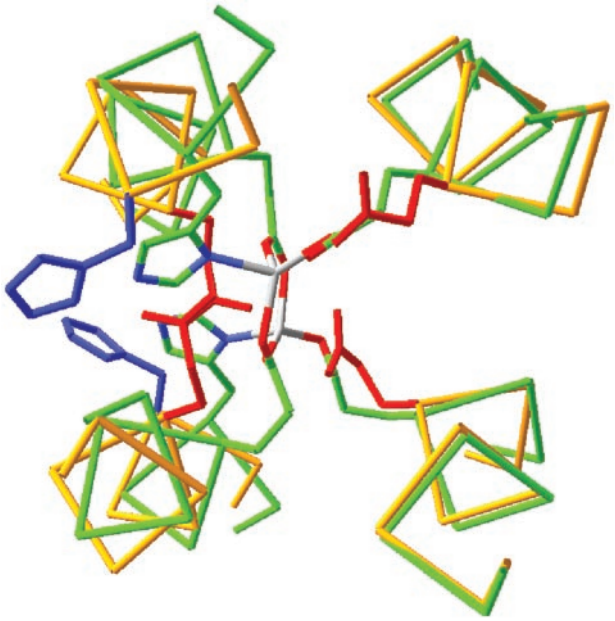

Fig. 5. Superposition of the di-Zn(II) and apo structures of DF1. The backbone trace of the di-Zn(II) structure is in green, and the side chains are plotted with Corey–Pauling–Koltun colors (C, green; N, blue; O, red). The backbone of apo-DF1 is in yellow, the Glu side chain is in red, and the His residue is in blue. (*Right*) Helices 1 and 1′, which superpose well between the structures. (*Left*) A much poorer superposition of helices 2 and 2′ arises from a rotation of the helices about their axes, which increases the exposure of the His and Glu side chains.

There are also small differences in the overall orientations of helices 2 and 2′ in the NMR structure of the protein. The rmsd between the individual members in the ensemble of solution structures is on the order of 0.5 Å for the backbone atoms. A significantly greater rmsd (1.6 Å) was observed when the ordered regions of the solution structure were superimposed on di-Zn(II) DF1 (8).

To identify the structural differences in the apo versus the di-Zn(II) structures, they were examined in a common Cartesian coordinate system in which the central axis of the bundle is aligned along the z axis and the approximate C_2 axis of symmetry of the metal-binding site is aligned along the x axis (41). The N-terminal helices, which contain a single Glu liganding site, are nearly invariant between the structures. In contrast, the C-terminal helices are rotated by ≈30° in the apo versus the di-Zn(II) structures, which allows the primary ligands, Glu-36 and His-39 to become more exposed in the apo versus the di-Zn(II) structures (Fig. 5). Thus, in the absence of stabilizing metal ion–ligand interactions, the His and Glu ligands prefer a more exposed location near the surface of the protein.

In a related crystallographic investigation (10), we recently observed a limited shift in the geometry of the helix 1–helix 1′ interface that occurs concomitantly with a change in the coordination geometry of the di-Mn(II) derivative of L13G-DF1. Similar changes in tertiary structure have not been observed yet in the crystal structures of diiron proteins, which tend to use more local conformational motions to adjust to changes in the redox state of the metal ion and other molecular events within their active sites (17, 19, 42). It will now be interesting to determine how these differences affect the chemical and functional properties of this class of proteins.

The degree of preorganization of metalloproteins has been shown to vary significantly depending on the functional re-

quirements of the sites. For example, zinc fingers feature a structural metal ion-binding site, and in these structures the binding of Zn(II) is required to fold the structure (43). At the other extreme the crystal structure of the R2 subunit of ribonucleotide reductase has been determined in the apo and various dimetalated forms (16, 17, 44). In this case, there is a significantly smaller difference in the geometry of the metallo site in the apo forms versus the dimetalated structures than is observed in the DF1 structures. The comparison for DF1 is complicated somewhat by the fact that the apo structure was solved in solution, whereas the other di-Zn(II) structure was solved crystallographically. Nevertheless, the structural differences were significantly larger than expected from the experimental error. This finding suggests that DF1 is more malleable than ribonucleotide reductase, possibly because DF1 is a homodimer of covalently unconnected helix–loop–helix motifs. By comparison, the diiron-binding site in R2 is housed in a single-chain four-helix bundle that is embedded within a much more complex and protein fold, which might provide additional structural restraints on the binding site.

A second goal of this work was to determine the thermodynamic consequences of introducing a substrate-access channel into the structure of DF1. The thermodynamic data show that the destabilization is extreme; the Leu-to-Gly mutation destabilized the dimer by >10 kcal/mol. Thus, although this functional mutation extracts a large thermodynamic price, the extreme stability of DF1 provides adequate stability to compensate for the mutation.

In summary, the *de novo* design of metalloproteins provides a critical test of the features required for the folding and function of this class of proteins. To obtain the most complete knowledge, however, it is essential to characterize the thermodynamic, structural, dynamic and functional properties of the targets thoroughly and to compare them to natural counterparts. Only in this way can the requirements for function be appreciated fully. This article, which details the solution properties of apo-DF1 and its active-site variants, can now provide a basis for understanding the metal-binding, redox, and chemical properties of metallo derivatives of the DF1 family of proteins.

We thank Professors Bertini, Banci, and Luchinat for kindly allowing us to collect NMR data at the Large Scale Facility Magnetic Resonance European Center (University of Florence, Florence, Italy) and Dr. Crescenzi (University of Naples, Federico II, Naples, Italy) for helping with computations. This work was supported by the Italian Ministry of University and Scientific Research Grants PRIN 2000 and 03185591 and National Institutes of Health Grant GM54616.

1. Honig, B. & Yang, A.-S. (1995) *Adv. Protein Chem.* **46,** 27–58.
2. Baldwin, E. P. & Matthews, B. W. (1994) *Curr. Opin. Biotechnol.* **5,** 396–402.
3. Soichet, B. K., Baase, W. A., Kuroki, R. & Matthews, B. W. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 452–456.
4. Lippard, S. J. & Berg, J. M. (1994) *Principles of Bioinorganic Chemistry* (Univ. Sci. Books, Mill Valley, CA).
5. Holm, R. H., Kennepohl, P. & Solomon, E. I. (1996) *Chem. Rev. (Washington, D.C.)* **96,** 2239–2314.
6. Pasternak, A., Kaplan, J., Lear, J. D. & DeGrado, W. F. (2001) *Protein Sci.* **10,** 958–969.
7. DeGrado, W. F., Summa, C. M., Pavone, V., Nastri, F. & Lombardi, A. (1999) *Annu. Rev. Biochem.* **68,** 779–819.
8. Lombardi, A., Summa, C. M., Geremia, S., Randaccio, L., Pavone, V. & DeGrado, W. F. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 6298–6305.
9. Di Costanzo, L., Wade, H., Geremia, S., Randaccio, L., Pavone, V., DeGrado, W. F. & Lombardi, A. (2001) *J. Am. Chem. Soc.* **123,** 12749–12757.
10. DeGrado, W. F., Di Costanzo, L., Geremia, S., Lombardi, A., Pavone, V. & Randaccio, L. (2003) *Angew. Chem. Int. Ed. Engl.* **42,** 417–420.
11. Merkx, M., Kopp, D. A., Sazinsky, M. H., Blazyk, J. L., Muller, J. & Lippard, S. J. (2001) *Angew. Chem. Int. Ed. Engl.* **40,** 2782–2807.
12. Lange, S. J. & Que, L., Jr. (1998) *Curr. Opin. Chem. Biol.* **2,** 159–172.
13. Waller, B. J. & Lipscomb, J. D. (1996) *Chem. Rev. (Washington, D.C.)* **96,** 2625–2657.
14. Stubbe, J. & Riggs-Gelasco, P. (1998) *Trends Biochem. Sci.* **23,** 438–443.
15. Frolow, F., Kalb, A. J. & Yariv, J. (1994) *Nat. Struct. Biol.* **1,** 453–460.
16. Atta, M., Nordlund, P., Aberg, A., Eklund, H. & Fontecave, M. (1992) *J. Biol. Chem.* **267,** 20682–20688.
17. Andersson, M. E., Högbom, M., Rinaldo-Matthis, A., Andersson, K. K., Sjöberg, B.-M. & Nordlund, P. (1999) *J. Am. Chem. Soc.* **121,** 2346–2352.
18. Lindqvist, Y., Huang, W., Schneider, G. & Shanklin, J. (1996) *EMBO J.* **15,** 4081–4092.
19. Jin, S., Kurtz, D. M., Jr., Liu, Z. J., Rose, J. & Wang, B. C. (2002) *J. Am. Chem. Soc.* **124,** 9845–9855.
20. Ghirlanda, G., Lear, J. D., Lombardi, A. & DeGrado, W. F. (1998) *J. Mol. Biol.* **281,** 379–391.
21. Rance, M., Sørensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R. & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun.* **117**, 479–485.
22. Griesinger, C., Otting, G., Wüthrich, K. & Ernst, R. R. (1988) *J. Am. Chem. Soc.* **110,** 7870–7872.
23. Jeener, J., Meier, B. H., Bachman, P. & Ernst, R. R. (1979) *J. Chem. Phys.* **71,** 4546–4553.
24. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995) *J. Biomol. NMR* **6,** 277–293.
25. Bartels, C., Xia, T., Billeter, M., Güntert, P. & Wüthrich, K. (1995) *J. Biomol. NMR* **5,** 1–10.
26. Güntert, P., Braun, W. & Wüthrich, K. (1991) *J. Mol. Biol.* **217,** 517–530.
27. Güntert, P., Mumenthaler, C. & Wüthrich, K. (1997) *J. Mol. Biol.* **273,** 283–298.
28. Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E., III, Wang, J., Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Stanton, R. V., *et al.* (2002) AMBER (Univ. of California, San Francisco), Version 7.
29. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr., & Weiner, P. (1984) *J. Am. Chem. Soc.* **106,** 765–784.
30. Weiner, S. J., Kollman, P. A., Nguyen, D. T. & Case, D. A. (1986) *J. Comput. Chem.* **7,** 230–252.
31. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477–486.
32. Xu, J., Baase, W. A., Baldwin, E. & Matthews, B. W. (1998) *Protein Sci.* **7,** 158–177.
33. Pace, C. N. (1992) *J. Mol. Biol.* **226,** 29–35.
34. Bryson, J. W., Betz, S. F., Lu, H. S., Suich, D. J., Zhou, H. X., O'Neil, K. T. & DeGrado, W. F. (1995) *Science* **270,** 935–941.
35. Myers, J. K., Pace, C. N. & Scholtz, J. M. (1997) *Biochemistry* **36,** 10923–10929.
36. Schellman, J. A. (1987) *Annu. Rev. Biophys. Biophys. Chem.* **16,** 115–137.
37. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids* (Wiley, New York).
38. Esimov, A. V. (1991) *Protein Eng.* **4,** 245–250.
39. Aurora, R. & Rose, G. D. (1998) *Protein Sci.* **7,** 21–38.
40. Schrauber, H., Eisenhaber, F. & Argos, P. (1993) *J. Mol. Biol.* **230,** 592–612.
41. Summa, C. M., Lombardi, A., Lewis, M. & DeGrado, W. F. (1999) *Curr. Opin. Struct. Biol.* **9,** 500–508.
42. Whittington, D. A. & Lippard, S. J. (2001) *J. Am. Chem. Soc.* **123,** 827–838.
43. Berg, J. M. (1990) *Annu. Rev. Biophys. Biophys. Chem.* **19,** 405–421.
44. Logan, D. T., deMare, F., Persson, B. O., Slaby, A., Sjoberg, B. M. & Nordlund, P. (1998) *Biochemistry* **37,** 10798–10807.