Lanthanide-binding helix-turn-helix peptides: Solution structure of a designed metallonuclease

Joel T. Welch*, William R. Kearney†, and Sonya J. Franklin*‡

*Department of Chemistry and †College of Medicine NMR Facility, University of Iowa, Iowa City, IA 52242

Edited by Kenneth N. Raymond, University of California, Berkeley, CA, and approved January 3, 2003 (received for review October 28, 2002)

A designed lanthanide-binding chimeric peptide based on the strikingly similar geometries of the EF-hand and helix-turn-helix (HTH) motifs was investigated by NMR and CD spectroscopy and found to retain the same overall solution structure of the parental motifs. CD spectroscopy showed that the 33-mer peptide P3W folds on binding lanthanides, with an increase in α-helicity from 20% in the absence of metal to 38% and 35% in the presence of excess Eu(III) and La(III) ions, respectively. The conditional binding affinities of P3W for La(III) (5.9 ± 0.3 μM) and for Eu(III) (6.2 ± 0.3 μM) (pH 7.8, 5 mM Tris) were determined by tryptophan fluorescence titration. The La(III) complex of peptide P3, which differs from P3W by only one Trp-to-His substitution, has much less signal dispersion in the proton NMR spectra than LaP3W, indicating that the Trp residue is a critical hydrophobic anchor for maintaining a well-folded helix-turn-helix structure. A chemical-shift index analysis indicates the metallopeptide has a helix-loop-helix secondary structure. A structure calculated by using nuclear Overhauser effect and other NMR constraints reveals that P3W not only has a tightly folded metal-binding loop but also retains the $\alpha-\alpha$ corner supersecondary structure of the parental motifs. Although the solution structure is undefined at both the N and C termini, the NMR structure confirms the successful incorporation of a metal-binding loop into a HTH sequence.

One of the great promises of *de novo* protein design is the possibility of incorporating new reactivity or selectivity into a protein matrix for specific chemical targets or to advance biological understanding. The redesign of a known protein fold is one powerful approach to achieving new function, and the inclusion of metal sites that could confer a range of enzymatic activities is particularly intriguing. Designed metalloproteins have the potential to promote natural enzymatic reactions on unnatural substrates or even to achieve reactivity through mechanisms completely unprecedented in biology. However, to interpret and evaluate their activity and to improve subsequent designs, the resultant constructs must retain a predictable well-folded structure. A structurally well characterized robust scaffold is therefore fundamental to the design of minimalist enzymes.

In recent years, several approaches have been taken to engineer novel metalloproteins, including building a metal site and scaffold completely from first principles (1–5), grafting ligands onto known protein scaffolds by point mutations that were identified through modeling or a vector-based algorithm such as DENZYMER (5–9), or redesigning natural metalloproteins for potential new reactivity (10, 11). Alternately, native metal-binding sequences may be used in new contexts (12–14), and we have recently reported the *de novo* design (or more correctly, redesign) of metallopeptide nucleases based on this premise (12, 15).

The structural similarity in their natural protein contexts of the helix-turn-helix (HTH), a DNA-binding motif common to transcription factors, and the EF hand, a prevalent Ca-binding motif, led us to consider redesigning the DNA-binding turn to incorporate a metal site (12). These two supersecondary structures can both be classified as $\alpha-\alpha$ corners, a ubiquitous protein structural element comprising two orthogonal α-helices tethered through a turn (16). However, the rational redesign of a metallo-HTH motif is predicated on the assumption that $\alpha-\alpha$ corner motifs are essentially interchangeable and structurally autonomous. The successful design of a metallopeptide as a well folded isolated $\alpha-\alpha$ corner motif tests our understanding of the design process and verifies the HTH as a viable scaffold for further metalloenzyme design.

In addition to the challenge of protein design, the impetus to combine these two motifs is the potential to incorporate both DNA affinity and hydrolytic nuclease activity into a small system to catalyze cleavage at a targeted DNA site. We have found that peptide chimeras comprising helices α2 and α3 of the *engrailed* homeodomain and the Ca-binding loop of one *calmodulin* EF hand are hydrolytically active as the lanthanide complexes (15, 17, 18). This cleavage is both regioselective and, in the case of the well-folded 33-mer peptide P3W, moderately sequence selective (unpublished data). This level of sequence preferential hydrolysis by an underivatized peptide nuclease is unprecedented, but improving both reactivity and sequence selectivity remains a goal. In native homeodomain proteins, key sidechains in a defined 3D array are responsible for protein orientation in the major groove and for the specific contacts that allow target recognition. Therefore, the integrity of the designed fold is likely to be crucial to conferring selectivity.

Because improving both selectivity and reactivity rests on an understanding of structure, we have determined the solution structure of lanthanum(III)-loaded P3W peptide by NMR. We have investigated the overall geometry of the chimeric metallopeptide, as well as the openness and accessibility of the active site, and compared the structure to the proposed design. This study addresses whether the premise of modularly substitutable $\alpha-\alpha$ corners is achieved, thereby opening new avenues for future metalloprotein designs.

Experimental Methods

Peptide Design and Synthesis. The design of peptide P3 has been described in greater detail elsewhere, and P3W was designed in an analogous manner (12). Briefly, crystal coordinates of calmodulin (1OSA) (19) and the engrailed homeodomain (1ENH) (20) were obtained from the Protein Data Bank and visualized by using the freeware program SWISS PDB VIEWER. The proteins were docked manually to overlay the HTH and the EF hand supersecondary structures as described (12). Peptides were synthesized by using standard Fmoc chemistry, cleaved from resin, and purified by HPLC to >95% purity [P3W, New England Peptide (Fitchburg, MA); P3, California Institute of Technology Peptide Facility]. Concentrations of peptide were determined by monitoring absorbance at 280 nm (P3W ε = 7,289 $M^{-1}\cdot cm^{-1}$; P3 ε = 254 $M^{-1}\cdot cm^{-1}$).

CD. CD spectra were recorded on an Olis-DS 17 spectrophotometer (OLIS, Jefferson, GA) at 25°C under N_2 atmosphere in a 0.1-cm cell by scanning from 375 to 200 nm at 1 nm per

This paper was submitted directly (Track II) to the PNAS office.

Abbreviations: HTH, helix-turn-helix; TOCSY, total correlation spectroscopy; Ln, lanthanide.

‡To whom correspondence should be addressed. E-mail: sonya-franklin@uiowa.edu.

data point. Samples contained 50 μM peptide in 10 mM Tris, pH 7.8. Aliquots of 1 mM $LnCl_3$ (Ln, lanthanide) stock solutions were added to the sample and the solution allowed to equilibrate with stirring for 10 min before data collection. $EuCl_3 \cdot 6H_2O$ and $LaCl_3 \cdot 7H_2O$ were obtained from Aldrich and fresh stock solutions prepared by weight before use. The helical content was calculated from $[\Theta]_{222}$ as described (12).

Tryptophan Fluorescence Titrations. The metal-binding affinity of P3W for lanthanum and europium was studied by monitoring the fluorescence intensity of the single tryptophan residue (W_{24}) as a function of added lanthanide ion. Spectra were collected on an Aminco-Bowman Series 2 fluorimeter. After excitation at 280 nm, emission was observed from 300 to 450 nm (30 nm/min). Metal was added to the peptide solution (50.0 μM peptide/5 mM Tris, pH 7.8) and allowed to equilibrate as described above. The intensity as a function of metal concentration described a binding curve, which was well-fit by a single exponential 1:1 association model by using a nonlinear least-squares fitting algorithm (21, 22).

NMR Sample Preparation. Samples were prepared in 99.99% D_2O (Cambridge Isotope Laboratories, Cambridge, MA) or 90:10 H_2O:D_2O. Samples for the lanthanum titration contained 0.61 mM P3W (pH 6.3, 25 mM d^5-imidazole; 600 MHz), and for the europium titration, 0.81 mM P3W (50 mM d^{11}-Tris, pH 7.8; 400 MHz). Before titration, peptide concentrations were quantified by UV-visible spectroscopy. The lanthanides were added as the chloride salts (25 mM stock solutions in D_2O, prepared fresh by weight) and the sample allowed to equilibrate for at least 5 min before data collection. Samples for 2D spectra were 1.6 mM in P3W peptide with 1.5 equivalents of $LaCl_3$ (2.5 mM), in d^{11} Tris at pH 7.8. The LaP3 sample was made in an analogous manner (1.5 mM P3, 2.5 mM $LaCl_3$in D_2O, in d^{11} Tris at pH 7.8).

NMR Instrumentation and Experiments. The titration of P3W with lanthanides was followed by ^{1}H NMR. 1D titration data were collected at 25°C on either a 600-MHz Bruker (Billerica, MA) AMX600 Spectrospin Spectrometer [La(III)] or a 400-MHz Bruker DRX400 Spectrometer [Eu(III)]. 2D ^{1}H-NMR total correlation spectroscopy (TOCSY) and NOESY data of 1.5:1 La(III):P3W were obtained at either 600 MHz on a Bruker AMX600, or at 500 MHz on a Varian Inova Spectrometer. Initial assignments were made based on the 600-MHz TOCSY and NOESY data, which were collected at 25°C, with presaturation to remove residual water signal. The NOESY mixing time of 100 ms was chosen as it is unlikely to produce appreciable contamination by spin diffusion in a peptide as small as LaP3W.

Final assignments and integration of NOE crosspeaks for LaP3W were obtained from both D_2O and 90:10 H_2O:D_2O spectra (23). The 2D ^{1}H-TOCSY and NOESY data were collected at 25°C and 500 MHz (Varian) in 90:10 H_2O:D_2O, with 2 sec presaturation for solvent suppression. 2D NOESY spectra were recorded with t_{mix} = 100 msec, 512 t_1 transients of 2,048 t_2 points, and 16 scans per transient. The 2D TOCSY was collected similarly, with t_{mix} = 120 msec. All spectra were referenced to the residual HOD peak at 4.757 ppm at 25.0°C. The LaP3 NOESY spectrum was recorded at 600 MHz, with 2 sec presaturation of the HOD peak, t_{mix} = 100 msec, and 512 t_1 transients of 4,096 t_2 data points each, with 48 scans per transient.

Table 1. NOE distance constraints used in P3W structural calculations

Number of constraints	
Total	114
Intraresidue	71
Interesidue	43
1 residue away	12
>2 and <5 residues away	14
≥5 residues away	17
<2.50 Å	15
>2.50 and <3.50 Å	48
>3.50 Å	51

Table 2. Average distance constraint violations, Å, for the best 10 conformers generated by DYANA

Number of violations	15 ± 1
Violations per violated constraint	0.57 ± 0.03 Å
Violations per all constraints	0.07 Å

NMR Solution Structure. The conformation of the protein was determined from NMR constraints by using the program Dynamics Algorithm for NMR Applications (DYANA). NOESY peak volumes and chemical shifts of well-resolved peaks were measured by using Felix-2000 (Accelrys, Burlington, MA). One hundred fourteen NOESY constraints were used for the final structural calculation (Table 1). Peak volumes were normalized to the volume of the H_4-H_5 ring protons of Trp-24. DYANA was iteratively used to calculate and anneal the structures consistent with these constraints. A total of 100 trial structures were chosen with torsion angles selected at random. Each structure was then annealed in a series of 4,000 steps from an initial temperature of 8,000 K to 0 K, followed by 1,000 further steps of conjugate gradient minimization. For the final structure, 10 of the lowest energy conformations were written (Tables 2 and 3). Five sets of 100 structures each were calculated by using different random number seeds for the initial conformations. The backbone conformations of each subset were compared with the other subsets and to the whole ensemble and found to be invariant to within 0.10%. The final set of structures was visualized by using the program Molecular Analysis and Molecular Display (MOLMOL).

Ligand donor atoms Asp-9 $O_{\delta 1}$, Asp-11 $O_{\delta 1}$, Asp-13 $O_{\delta 1}$, Thr-15 O, and both Glu-20 O_ε were constrained during the annealing process (24) to distances from the metal representing consensus lanthanide(III)-EF-hand ligation (19, 25). However, the 3D coordination geometry (O-Ln-O angles) around the metal were not defined in any way. These distance constraints were justified by observed NOEs in the LaP3W spectra, paramagnetic shifts, line broadening of the EuP3W complex. Importantly, the inclusion of the 6 La-O distances to the DYANA model did not appreciably affect the fit or alter the model (F = 4.00 without and 4.54 with these La-O bonds, for all backbone atoms, 20 best structures; see Fig. 6, which is published as supporting information on the PNAS web site, www.pnas.org). The Asp-17 $O_{\delta 1}$–La constraint representing a possible carboxylate to coordinate-water H bonding interaction was also tested but found to be inconsistent with the structure. A description of peak assign-

Table 3. rms deviation values for the best 10 conformers generated by DYANA, based on NOE distance constraints

Residues	Mean backbone rms deviation, Å	Number of constraints
8–25	1.05 ± 0.26	103
1–33	4.27 ± 1.47	114

This tally does not include six La-O distance constraints.

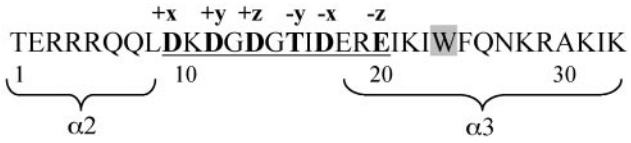

Fig. 1. The sequence of P3W, derived from the HTH motif of the engrailed homeodomain (α2, α3) and the consensus Ca-binding loop of an EF hand (underlined). The 33-mer peptide is numbered below, and the Ca-ligating residues (bold) are labeled above. Peptide P3 differs from P3W (shown) by a Trp-24→His-24 substitution (shaded).

ments and close contacts can be found in *Supporting Text*, which is published as supporting information on the PNAS web site, and the metal coordination environment is discussed further in *Results*.

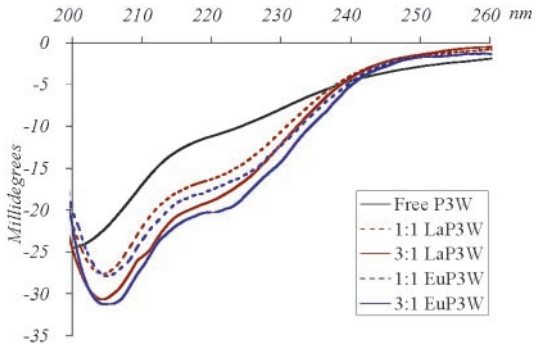

Fig. 2. CD spectra of free P3W (black), LaP3W (red), and EuP3W (blue). Samples contained 50 μM P3W, 5 mM Tris, pH 7.8, 25°C, and were titrated with one equivalent (dashed lines) or a 3-fold excess (solid lines) of the lanthanide salts.

Results and Discussion

Lanthanide-Binding Affinities. A series of peptide chimeras has been designed based on the structural resemblance of the $\alpha-\alpha$ corners of engrailed and calmodulin proteins (12). One such peptide, the 33-mer P3W, incorporates the 12-aa consensus EF-hand Ca-binding loop sequence between helices α2 and α3 of engrailed homeodomain (Fig. 1). This peptide binds one equivalent of Ca(II) or lanthanide ions with binding affinities typical of EF-hand-derived peptides (26–31). The binding affinity was determined by fluorescence titration following the change in the tryptophan fluorescence intensity as a function of added metal, after excitation at 280 nm. On the addition of La(III) or Eu(III), the fluorescence intensity at 350 nm decreases systematically (see Fig. 7, which is published as supporting information on the PNAS web site). This decrease, along with the slight red shift (≈2 nm) of the emission peak λ_{max}, suggests the single tryptophan residue of P3W moves into a slightly more solvent exposed environment caused by metal binding and peptide folding. The decrease in intensity describes a binding curve, which was fit to a 1:1 association model by using a nonlinear least-squares algorithm (21, 22). The conditional dissociation constants for La(III)P3W and Eu(III)P3W are $K_d = 5.9 \pm 0.3$ μM and 6.2 ± 0.3 μM, respectively, at pH 7.8, and 5 mM Tris buffer, in good agreement with isolated EF-hand peptides in solution (26–31) and similar to the EuP3W affinity (4.2 ± 0.3 μM) determined previously at slightly lower pH. Additionally, these values are similar to the affinity of P3W for Ce(IV) ($K_d = 3.9 \pm 0.3$ μM) and approximately an order of magnitude stronger than the affinity for Ca(II) ($K_d = 35 \pm 1.4$ μM). Although *in vivo* the EF-hand motif typically binds Ca(II), which is most similar in effective ionic radii size to the middle lanthanides [1.21 Å vs. 1.20 Å for Eu(III)] (32), the structure is known to be capable of accommodating ions with some size variation. This is reflected in the similar affinities of P3W for La(III) and Eu(III), as well as the smaller Ce(IV) ion.

Induced Secondary Structure on Metal Binding. As reflected in the change in tryptophan environment, metal binding by P3W induces global folding of the peptide. CD spectroscopy shows that in the absence of metal, P3W has little overall secondary structure. However, on addition of one equivalent of La(III) or Eu(III), the metallopeptide folds, resulting in complexes with 30 and 32% α-helicity, respectively (Fig. 2). These estimates of α-helicity are calculated on the basis of observed molar ellipticity at 222 nm, which is predicted to be ≈50% in a fully organized chimera of this design. P3W binds only one Ln(III) ion per peptide, as demonstrated by NMR titration studies. However, a higher percent helicity is observed when excess metal is added (La 35%, Eu 38% at 3:1 Ln:P3W), consistent with a somewhat flexible or labile metal-binding site. These helicity values are substantially higher than the previously reported P3 peptide (14% helical at 1:1 Eu:P3) (12), which contains the point mutation Trp-24→His-24. The more defined secondary structure of P3W suggests that the tryptophan residue plays a crucial role in stabilizing the supersecondary fold, which was confirmed by the NMR solution structure.

Structural changes on the addition of metal were investigated more directly by NMR. The addition of La(III) or Eu(III) to P3W results in a spectrum characteristic of a partially folded structure common in small well-organized peptides (see Fig. 8, which is published as supporting information on the PNAS web site). A well-folded peptide will not have the same degree of signal dispersion as a well-folded protein, because the normally disordered termini are a larger fraction of the total residues. However, consistent with the CD data, the NMR spectra support a well-organized central loop flanked by disordered ends. The HTH motif is stabilized by key hydrophobic interactions at the intersection of the two helices, which is reflected in notable shifts for these residues. The aromatic ring protons of both Trp-24 and Phe-25 shift predominately upfield and show increased signal dispersion as they are incorporated into a hydrophobic pocket. Also notable in the LaP3W spectrum are the γ-Me groups of Leu-8 that become nonequivalent and shift upfield, the δ-Me protons of Ile-21 that shift 0.10 ppm downfield, and the γ-Me of Thr-15. It is this threonine residue, the seventh residue of the metal-binding loop, that supplies the backbone carbonyl oxygen donor ligand to the metal in a typical EF hand. Consistent with the chimera using an EF-hand-like metal-binding mode, the α-proton of Thr-15 shifts downfield 0.58 ppm, and the γ-Me of Thr-15 moves 0.08 ppm upfield from their random coil positions in the LaP3W spectrum, whereas these peaks are unobservable in the EuP3W spectrum (e.g., T_{15} γ-Me; see Fig. 8).

The flexibility and kinetic lability of metal-binding results in an adduct that is in intermediate exchange on the NMR time scale. As metal is added, peptide folding causes peak shifts and results in a spectrum with greater dispersion at 1:1 stoichiometry than in the apopeptide. No further significant shifts are observed with additional metal (see Fig. 9, which is published as supporting information on the PNAS web site). However, the peaks broaden and drift rather than grow in during the titration, until the peptide is fully occupied. This is consistent with the increased structure observed by CD in the presence of excess metal to drive folding. The metallopeptide, although well organized, remains flexible.

Chemical Shifts Indicate HTH Secondary Structure. A chemical-shift index analysis (33) of the α-proton peaks was used to determine

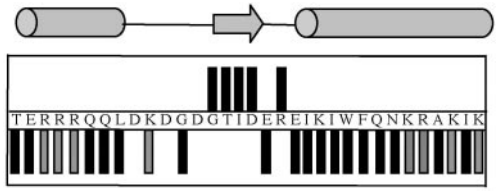

Fig. 3. A chemical-shift index analysis of LaP3W, based on α-proton chemical shifts, after the method of Wishart *et al.* (33). Peaks that shifted >0.1 ppm from tabulated values (34) were assigned values of –1 (upfield) or +1 (downfield), and those that shifted <0.1 ppm were assigned a value of 0. Although several K and R residues could not be unambiguously assigned due to overlap, these peaks were all consistent with –1 values (gray bars), confirming local secondary structure.

the predominant secondary structure adopted by LaP3W. The chemical shifts of amino acid Hα protons are very sensitive to the local environment, so comparing the chemical shift of each residue's α-proton(s) to tabulated random coil values (34) allowed the assignment of regions of secondary structure (Fig. 3). The assignments of α-protons were ambiguous for several overlapping Arg and Lys residues, but the shifts were all in a range consistent with –1 values (gray bars, Fig. 3). Thus the α-helical secondary structure assignments for these residues were confirmed despite the overlapping signals. The resultant HTH structure, including the EF-hand-like β-strand within the loop, parallels the design remarkably well.

Hydrophobic Anchor at the Turn. The importance of a cluster of hydrophobic residues anchoring the pivot of the $\alpha-\alpha$ corner is illustrated by a comparison of LaP3W and LaP3 spectra. The peptide P3 is analogous to P3W with the exception of Trp-24, for which a non-native His residue was substituted in P3 (12). This Trp residue is apparently key to the integrity of the metalloprotein fold and is in fact 100% conserved in homeodomain sequences (35, 36). The Trp residue of α3 is vital not only to the organization of the domain, thus promoting the DNA-sequence selectivity of homeodomains, but is also critical to their ability to translocate through cellular membranes (37, 38). Both of these properties depend on a well-organized, helical, and robust structure.

A portion of the 2D NOESY spectrum in D_2O is presented in Fig. 4 for both LaP3W and LaP3. Because these metallopeptides are diamagnetic systems, the greater signal dispersion observed for LaP3W reflects the enhanced diversity of chemical environments observed in folded structures. In addition to the difference in signal dispersion, it is notable that both His-24 and Phe-25 of LaP3 have two distinct chemical environments, and Leu-8 has only one averaged methyl position (Leu-8 indicated by a green arrow in Fig. 4). In contrast, Phe-25 and Trp-24 (which replaces His-24) occur in only one environment in LaP3W, and the two Leu-8 δ-Me groups are distinct and shifted upfield. These data indicate that LaP3W has a well-organized hydrophobic pocket, in which Leu-8-Trp-24 contacts brace side-chain rotations that are allowed in the analogous His-containing LaP3. The Trp-24 HN indole peak occurs in two similar but distinct chemical environments (10.00 and 10.05 ppm in the 90:10 H_2O:D_2O spectrum) and is in slow exchange on the NMR time scale. In contrast, the analogous imidazole HN protons of His-24 (LaP3) are exchanging too rapidly to observe in the water spectrum. The slow exchanging indole peak(s) lend additional support to the role of Trp in bracing the hydrophobic pivot of the $\alpha-\alpha$ turn. Furthermore, a collection of NOE close contacts between Leu-8, Ile-21, and Trp-24 in LaP3W verifies the importance of these hydrophobic residues in organizing and stabilizing the turn (for example, red boxes, Fig. 4), and these contacts

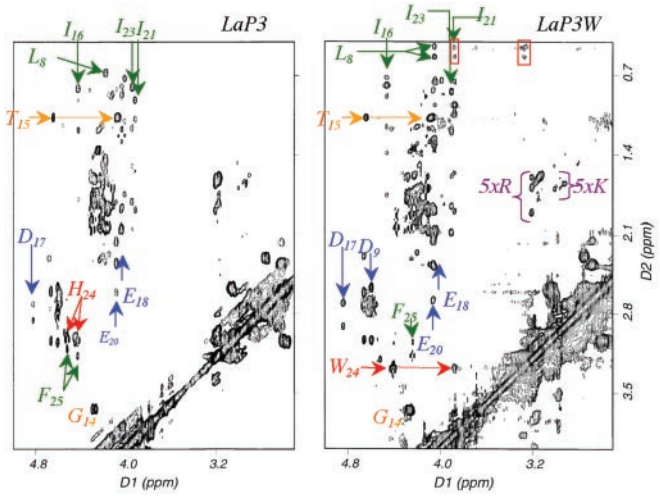

Fig. 4. A comparison of 2D NOESY spectra (expanded regions) of LaP3 (*Left*) and LaP3W (*Right*) under similar experimental conditions (50 mM d^{11}-Tris, pH 7.8, 25°C, 600 MHz, t_{mix} = 100 msec). LaP3W has greater signal dispersion and notable long-range crosspeaks (for example, red boxes: L_8-I_{21} and L_8-W_{24}), indicating this metallopeptide is better folded in solution. These peptides differ by only one residue (red; P3: His-24; P3W Trp-24). Selected peak assignments are noted (acidic D, E in blue; basic K, R in purple; hydrophobic L, I, F in green; T, G of the loop in gold).

serve to preserve the relative distances and orientations of the two helices.

Paramagnetic Broadening in the EuP3W Spectrum. The similarity of the coordination chemistry across the lanthanide series makes these ions convenient and varied spectroscopic probes for Ca sites in proteins. With the exception of a small size variation, the series of Ln(III) ions are virtually interchangeable with Ca(II) and with each other in protein environments. Thus Eu(III), a paramagnetic ion, will bind to the chimeric peptide in a manner analogous to the La(III) ion and will report on metal-proton distance by through-space pseudocontact shifts and broadening (39).

The 1D NMR spectra of the uncomplexed peptide were compared with that of the 1:1 EuP3W and LaP3W (Fig. 8). Although small differences in structure between the two metallopeptides are likely, the spectral shifts are predominantly quite similar, consistent with the similarity of the absorption peak intensities in their CD spectra (Fig. 2). Signal broadening and dramatic peak shifts due to the paramagnetic ion were more readily identified in the 2D TOCSY spectra (Fig. 10, which is published as supporting information on the PNAS web site). This allowed the assignments of several shifted or missing peaks, such as Gly-12 (Hα), Ile-16 (Hα, Hβ, Hγ), and Thr-15 (Hα, Hβ). Importantly, α-proton and side-chain protons of the putative ligating residues Asp-9, Asp-11, Asp-13, and Glu-20 are among those residues whose signals broaden significantly or into the baseline. The unresolvable peaks in the EuP3W spectrum therefore confirm the proximity of the metal ion to these residues (within 10 Å), and the expected EF-hand-loop-binding site. At temperatures from 4 to 40°C, some broad peaks are observed from –5 to 0 ppm and from 15 to 20 ppm but could not be unambiguously assigned. Therefore, the nonisotropic pseudocontact shift axes were not defined, and H-Eu distances were based on isotropic estimates and consistent with the crystal structures of Ln-EF-hand proteins. These residues were therefore explicitly defined as metal-donor ligands in the model (see below).

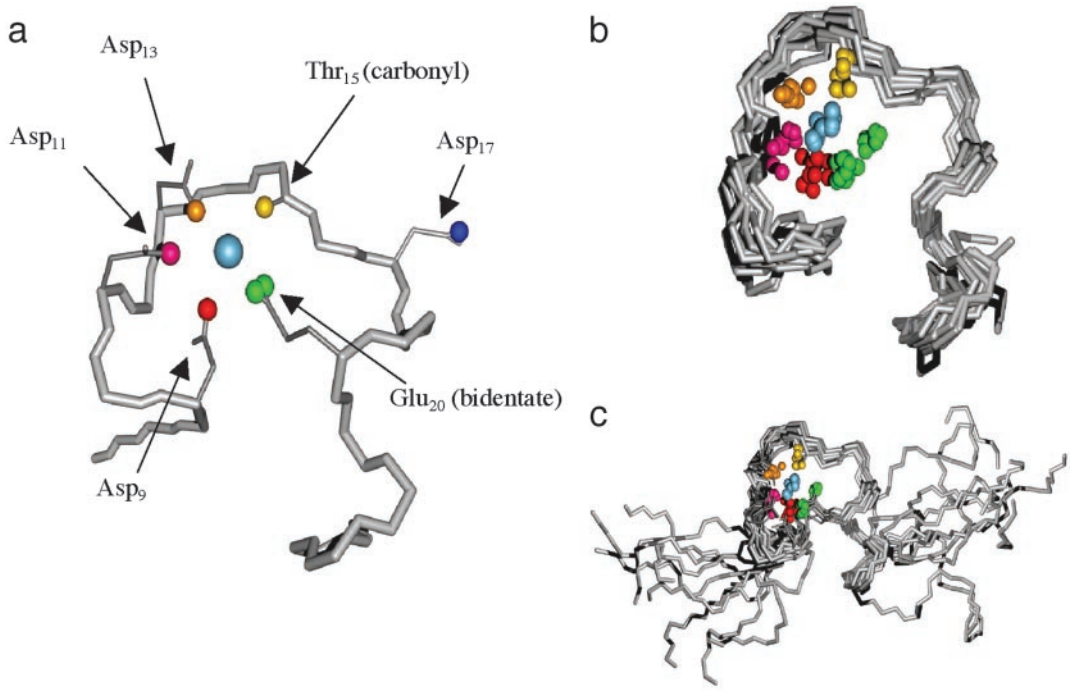

Fig. 5. Solution structure of LaP3W. (*a*) Backbone representation of residues 7–26 for the lowest energy conformer (lowest size; 5qDYANA *f*-value). La(III) is shown in light blue and other ligating oxygen donor atoms in smaller colored spheres, as labeled. Asp-17, which in EF hands often stabilizes an inner-sphere water, point away from the metal in this structure, leaving a very open coordination environment. (*b*) The 10 lowest energy conformers (DYANA *f* value) for the well defined region from residues 7 to 26. (*c*) Depiction of the full metallopeptide (residues 1–33) for the same subset of 10 calculated structures, showing the disorder in the N and C termini. Both flexibility and spectral overlap contribute to the poorly defined termini, which are likely more helical than our model based on a chemical-shift index analysis.

Comparison of the Structure with the Parental Design. P3W was designed to incorporate a Ca-binding loop into a new context, thus retaining the HTH structure of a DNA-binding motif, although including an accessible reactive metal site. The NMR solution structure of LaP3W shows that the chimeric metallopeptide does indeed fold into the intended 3D structure. Fig. 5 shows the superposition of the 10 best calculated structures, with ligating oxygens and the metal ion depicted as colored spheres. As predicted, the lanthanides bind to the EF-hand loop, nucleating the start of helices on either side. The side chains of the ligating Asp and Glu residues are oriented with the donor atoms toward the metal site.

The prediction that the binding mode is similar to an EF hand and the justification of including distance (but not geometric) constraints to specific donor oxygens are based on several observations. It was previously shown that P3 and P3W fold as a function of added metal, with binding affinities typical of EF hands and metal/peptide stoichiometry of 1:1. A control peptide with the EF-hand-binding sequence in reverse order did not bind metals or fold in their presence, despite having the same set of donor residues (12). Additionally, the coordination environment of the lanthanide was probed by Eu-luminescence spectroscopy. The emission peak positions and lifetimes confirmed the binding affinities and were consistent with two inner-sphere water molecules and six additional oxygen ligands (W. DeW. Horrocks and S.J.F., unpublished results). These data suggest that the binding mode used by our peptide chimeras is an EF-hand-type site, which is also consistent with the paramagnetic NMR data.

The expected ligand contacts at positions +x, +y, +z, −y, −z of the loop (Fig. 1) were confirmed by the close proximity of these residues to the paramagnetic Eu(III) ions and, on the basis of these paramagnetic shift data and the observations mentioned above, metal–oxygen constraints to the consensus EF-hand ligands were included in the model. However, even in the absence of these La-O distance constraints, a reproducible structured central loop is observed, albeit with somewhat lower resolution in the model. An overlay of the 15 best structures without the six La-O constraints shows that the overall fold of the EF hand was maintained. The first residue of the metal-binding loop (Asp-9) is notably disordered, but the last ligand (Glu-20) anchoring the start of the second helix is nearly unchanged in its position, well placed toward the center of the binding pocket (see supporting information on the PNAS web site).

It is notable that the Asp-17 β protons (loop position −x, or the ninth residue) also disappear in the EuP3W spectrum (assignments based on TOCSY spectra), indicating that the protons of this side chain were also likely within 10 Å of the metal. Because in solid-state structures the ninth residue of EF-hand loops is often, although not always, found to be involved in stabilizing an inner sphere water through hydrogen bonding, we explored this interaction by including a longer 7-Å distance restraint from the Asp-17 oxygen to La that could accommodate this type of hydrogen bond to coordinated water. Although some reasonable structures were generated with this constraint included, the model was overall less well-behaved than the absence of this constraint. However, if the Asp-17 side chain points away from the metal, the β-protons are still within 9 Å of the metal, consistent with the observed small paramagnetic broadening. The refined structure shows that LaP3W does not appreciably involve the ninth residue of the loop in rigidly stabilizing a bound water molecule, leaving the metal-binding site very exposed in solution (Fig. 5, navy sphere). The open face of the coordination site could easily accommodate two inner sphere water molecules, nicely consistent with the luminescence data.

The open coordination environment of the metal, perhaps in part a result of the lack of an EF-hand dimer interface present in native proteins, presumably enhances nuclease reactivity by

allowing the approach of DNA substrate, but likely also decreases the kinetic inertness toward metal dissociation. The typical mechanism for Ln-mediated DNA hydrolysis involves the metal site donating a deprotonated inner-sphere water nucleophile as well as activating the substrate P-O bond through phosphate coordination to the metal. The NMR structure of LaP3W shows that the metallopeptide has several open coordination sites for substrate binding and for labile water exchange. Unlike native EF-hand motifs, which nearly universally occur as back-to-back dimers, LaP3W occurs as a monomer in solution, on the basis of a determination of the solution molecular weight by static light scattering experiments (J. M. Weincek, J.T.W., and S.J.F., unpublished data). Although P3 has some propensity to dimerize, as determined by NMR dilution studies and isothermal titration calorimetry, this is apparently not the case for P3W, a better-organized metallopeptide. An isolated EF-hand-like structure apparently has enough accessibility of the active site metal to promote reactivity. Thus it appears that a single EF hand is not quite the same as half of an EF-hand pair, in structure or reactivity.

The NMR model shows that the structure of LaP3W is a HTH. The metallopeptide is well defined from residues 8–25, with a rms deviation of only 1.05 Å for the 10 best structures (backbone atoms; see also Fig. 11, which is published as supporting information on the PNAS web site). Both N and C termini are frayed in the model (Fig. 5), which may in part reflect the inherent flexibility of these regions. Furthermore, unlike self-assembling helical structures such as leucine-zipper-derived peptides or *de novo* α-helical bundles (40), the chimeric HTH metallopeptides are designed to have orthogonal rather than parallel (or antiparallel) helices. Thus, the helices of these chimeras are not stabilized by adjacent secondary structures along their full lengths as are helical bundles, so the termini may be less organized in solution. This does not preclude the possibility that metallopeptide interactions with a complementary interface, such as the DNA major groove, would increase chimera secondary structure throughout the length of the helices.

Although the chimera is certainly disordered at either end, the NMR α-proton chemical-shift data suggest that the peptide termini are more helical than our model indicates. The apparent lack of structure in these regions is exacerbated in the model by the difficulty with Lys and Arg assignments. Arg and Lys residues comprise 10 of 33 residues and occur in similar chemical environments, preventing full assignments and extensive sequential NOE walks along the helices. However, the chemical shifts of these residues indicate they predominantly occur within α-helical secondary structure.

Conclusion

In designing P3W, we proposed that the turn of the HTH motif could be modularly substituted by a metal-binding loop of another $\alpha-\alpha$ corner. The solution structure of LaP3W shows that this premise is sound, and that the metallopeptides adopt a well-defined turn structure consistent with the overall HTH DNA-binding motif. Although the termini are disordered in our model, even these regions have a propensity to be helical. The structure of LaP3W illustrates the similarity of the chimeric metallopeptide to the EF-hand and HTH parental structures, including initial helical directions consistent with roughly orthogonal helices, rather than the antiparallel helices of an α-helix bundle. The metal-binding turn is anchored by a cluster of hydrophobic residues at the pivot that are required for a well-ordered motif. The open metal-coordination site, along with the well-organized $\alpha-\alpha$ corner turn of this small chimeric peptide, validates the approach of engineering selective nucleases by incorporating metal-binding turns into transcription factors.

The retention of the turn configuration suggests the metallopeptide could interact with DNA as a unit, with tertiary structure complementary to the major groove, and thus maintain many of the contacts that allow the parental homeodomains to recognize target sequences. We postulate that groove binding by the recognition helix (residues 20–33) of these chimeric peptides could deliver the hydrolytically active metal to a specific DNA location, as we have described in a recent review (41). In keeping with this design premise, we have found that CeP3W cleaves supercoiled DNA with both apparent regioselectivity and a moderate degree of specificity despite its small size (unpublished data). Using this or any motif to explore metalloenzyme design is predicated on a well-organized well-characterized protein scaffold. The solution structure of LaP3W presented here suggests that the HTH is one such promising template for further metalloenzyme design, including tuning reactivity or exploring selectivity by combinatorial changes to the sequence.

This work was supported by National Science Foundation CAREER Award CHE-0093000 (S.J.F.), Center for Biocatalysis and Bioprocessing National Institutes of Health Predoctoral Training Grant T32 GM-08365 (J.T.W.), and the University of Iowa College of Medicine (W.R.K.).

1. Hill, R. B., Raleigh, D. P., Lombardi, A. & DeGrado, W. F. (2000) *Acc. Chem. Res.* **33,** 745–754.
2. Lombardi, A., Summa, C. M., Geremia, S., Randaccio, L., Pavone, V. & DeGrado, W. F. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 6298–6305.
3. Matzapetakis, M., Farrer, B. T., Weng, T.-C., Hemmingsen, L., Penner-Hahn, J. E. & Pecoraro, V. L. (2002) *J. Am. Chem. Soc.* **124,** 8042–8054.
4. Gibney, B. R., Isogai, Y., Rabanal, F., Reddy, K. S., Grosset, A. M., Moser, C. C. & Dutton, P. L. (2000) *Biochemistry* **39,** 11041–11049.
5. Lu, Y., Berry, S. M. & Pfister, T. D. (2001) *Chem. Rev.* **101,** 3047–3080.
6. Hellinga, H. & Richards, F. M. (1991) *J. Mol. Biol.* **222,** 763–785.
7. Benson, D. E., Wisz, M. S. & Hellinga, H. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 6292–6297.
8. Farinas, E. & Regan, L. (1998) *Protein Sci.* **7,** 1939–1946.
9. Wilcox, C. S., Putnam, C. D., Sastry, M., Blankenship, J., Chazin, W. J., McRee, D. E. & Goodin, D. B. (1998) *Biochemistry* **37,** 16853–16862.
10. Merkle, D. L., Schmidt, M. H. & Berg, J. M. (1991) *J. Am. Chem. Soc.* **113,** 5450–5451.
11. Conte, D., Narindrasorasak, S. & Sarkar, B. (1996) *J. Biol. Chem.* **271,** 5125–5130.
12. Kim, Y., Welch, J. T., Lindstrom, K. M. & Franklin, S. J. (2001) *J. Biol. Inorg. Chem.* **6,** 173–181.
13. Yang, W., Lee, H.-W., Hellinga, H. & Yang, J. T. (2002) *Proteins Struct. Funct. Genet.* **47,** 344–356.
14. McColl, D. J., Honchell, C. D. & Frankel, A. D. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 9521–9526.
15. Welch, J. T., Sirish, M., Lindstrom, K. M. & Franklin, S. J. (2001) *Inorg. Chem.* **40,** 1982–1984.
16. Efimov, A. V. (1996) *FEBS Lett.* **391,** 167–170.
17. Kim, Y. & Franklin, S. J. (2002) *Inorg. Chim. Acta* **341,** 107–112.
18. Sirish, M. & Franklin, S. J. (2002) *J. Inorg. Biochem.* **91,** 253–258.
19. Chattopadhyaya, R., Meador, W. E., Means, A. R. & Quiocho, F. A. (1992) *J. Mol. Biol.* **228,** 1177–1192.
20. Tucker-Kellogg, L., Rould, M. A., Chambers, K. A., Ades, S. E., Sauer, R. T. & Pabo, C. O. (1997) *Structure (London)* **5,** 1047–1054.
21. Wilcox, C. S. (1991) in *Frontiers in Supramolecular Organic Chemistry and Photochemistry*, eds. Schneider, H.-J. & Durr, H. (VCH, Weinheim, Germany), pp. 125–139.
22. Sirish, M. & Schneider, H.-J. (1999) *Chem. Commun.* 907–908.
23. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids* (Wiley, New York).
24. Lytle, B. L., Volkman, B. F., Westler, W. M., Heckman, M. P. & Wu, J. H. D. (2001) *J. Mol. Biol.* **307,** 745–753.
25. Bentrop, D., Bertini, I., Cremonini, M. A., Forsén, S., Luchinat, C. & Malmendal, A. (1997) *Biochemistry* **36,** 11605–11618.
26. Marsden, B. J., Shaw, G. S. & Sykes, B. D. (1990) *Biochem. Cell Biol.* **68,** 587–601.
27. Procyshyn, R. M. & Reid, R. E. (1994) *J. Biol. Chem.* **269,** 1641–1647.
28. Gariépy, J., Sykes, B. D. & Hodges, R. S. (1983) *Biochemistry* **22,** 1765–1772.
29. Reid, R. E., Gariépy, J., Saund, A. K. & Hodges, R. S. (1981) *J. Biol. Chem.* **256,** 2742–2751.
30. Borin, G., Ruzza, P., Rossi, M., Calderan, A., Marchiori, F. & Peggion, E. (1989) *Biopolymers* **28,** 353–369.
31. Vázquez-Ibar, J. L., Weinglass, A. B. & Kaback, H. R. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 3487–3492.
32. Huheey, J. E., Keiter, E. A. & Keiter, R. L. (1993) *Inorganic Chemistry* (HarperCollins, New York).
33. Wishart, D. S., Sykes, B. D. & Richards, F. M. (1992) *Biochemistry* **31,** 1647–1651.
34. Merutka, G., Dyson, H. J. & Wright, P. E. (1995) *J. Biomol. NMR* **5,** 14–24.
35. Gehring, W. J., Affolter, M. & Bürglin, T. (1994) *Annu. Rev. Biochem* **63,** 487–526.
36. Laughon, A. (1991) *Biochemistry* **30,** 11357–11367.
37. Prochiantz, A. (2000) *Curr. Opin. Cell Biol.* **12,** 400–406.
38. Derossi, D., Calvet, S., Trembleau, A., Brunissen, A., Chassaing, G. & Prochiantz, A. (1996) *J. Biol. Chem.* **271,** 18188–18193.
39. Bertini, I., Luchinat, C. & Parigi, G. (2001) *Solution NMR of Paramagnetic Molecules* (Elsevier Science, Amsterdam).
40. DeGrado, W. F., Summa, C. M., Pavone, V., Nastri, F. & Lombardi, A. (1999) *Annu. Rev. Biochem.* **68,** 779–819.
41. Franklin, S. J. (2001) *Curr. Opin. Chem. Biol.* **5,** 201–208.