Buried Polar Residues in Coiled-Coil Interfaces[†,‡]

David L. Akey, Vladimir N. Malashkevich, and Peter S. Kim*

Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, Massachusetts 02142

Received December 13, 2000; Revised Manuscript Received March 20, 2001

ABSTRACT: Coiled coils, estimated to constitute 3−5% of the encoded residues in most genomes, are characterized by a heptad repeat, (***abcdefg***)$_n$, where the buried ***a*** and ***d*** positions form the interface between multiple α-helices. Although generally hydrophobic, a substantial fraction (∼20%) of these ***a***- and ***d***-position residues are polar or charged. We constructed variants of the well-characterized coiled coil GCN4-p1 with a single polar residue (Asn, Gln, Ser, or Thr) at either an ***a*** or a ***d*** position. The stability and oligomeric specificity of each variant were measured, and crystal structures of coiled-coil trimers with threonine or serine at either an ***a*** or a ***d*** position were determined. The structures show how single polar residues in the interface affect not only local packing, but also overall coiled-coil geometry as seen by changes in the Crick supercoil parameters and core cavity volumes.

Coiled coils are an oligomerization motif commonly occurring at the interface between separate protein chains. Coiled coils are found in many cytoskeletal and contractile systems (e.g., intermediate filaments, nuclear lamins, and myosin), transcription regulators (e.g., Myc and Max, Fos and Jun, GCN4), viral envelope proteins (e.g., MoMLV, HIV, SIV, influenza), and other systems (*1*). Genome database searches with coiled-coil prediction programs suggest that 3−5% of all protein residues exist as coiled coils (*2*). This implies a role for the coiled-coil motif in a wide range of biological functions. A more complete understanding of the interactions significant to coiled-coil formation will be useful in analyzing protein−protein interactions, predicting candidate partners, and assigning potential biological functions to novel proteins.

Coiled coils are comprised of two or more α-helices that wrap around each other. The hallmark of coiled coils is a heptad repeat of amino acids, (***abcdefg***)$_n$, with a predominance of hydrophobic residues at the buried first and fourth (***a*** and ***d***) positions and charged residues frequently at the fifth and seventh (***e*** and ***g***) positions (Figure 1). Despite the prevalence of hydrophobic residues, approximately 20% of the ***a*** and ***d*** residues are polar or charged (*2*−*6*).

Previous work shows that buried polar residues in coiled coils can be important determinants of structural uniqueness (*7*−*12*), influencing both oligomeric state (e.g., dimers vs trimers) and strand orientation (i.e., parallel vs antiparallel). For example, ***a***-position asparagines and lysines, both more frequent in coiled-coil dimers than trimers (Table 1), have been found to specify dimeric coiled coils (*8, 10*). In addition, it has been shown that buried asparagines can direct strand orientation (*9*).

To address the role of buried polar residues in coiled-coil assembly, we constructed and characterized a series of peptides based on the GCN4-p1 sequence (*13, 14*) (a well-studied model system for coiled coils) with a single polar residue at either the ***a*** or the ***d*** position (Figure 1). We characterized these peptides in terms of their helical content, thermal stability, and oligomeric specificity. To investigate how polar residues are accommodated in the interior of coiled coils, we determined the crystal structures of four coiled-coil trimers containing a serine or threonine at either the ***a*** or the ***d*** position.

EXPERIMENTAL PROCEDURES

Peptide Synthesis and Purification. The sequence of the GCN4-pVL variants is Ac-RMKQLEDKVEE-**L/X**-LSK-**V/X**-YHLENEVARLKKLVGER, where the ***a*** and ***d*** positions are underlined. Position 12(***d***) (boldface) is either a leucine or a polar residue (N, Q, S, or T), and position 16(***a***) (boldface) is either a valine or a polar residue. Peptides were synthesized as described previously (*15*) using FMOC chemistry, desalted over a Sephadex G-25 column in 5% acetic acid, and purified by reverse-phase HPLC using a C18 preparative column. Purity was checked by analytical HPLC with a 0.1%/min water/acetonitrile gradient with 0.1% trifluoroacetic acid, and the mass of the peptides was verified by MALDI-TOF[1] mass spectroscopy (PerSeptive Biosystems, Inc., Framingham, MA).

Circular Dichroism (CD) and Sedimentation Equilibrium. CD measurements were carried out in 150 mM NaCl, 50

[†] This research used the W. M. Keck Foundation X-ray Crystallography Facility at the Whitehead Institute and was funded by NIH Grant GM 44162.

[‡] Coordinates have been deposited in the RCSB Protein Data Bank: pVL_S, 1IJ0; pVL_T, 1IJ1; pV_TL, 1IJ2; pV_SL, 1IJ3.

* Address correspondence to this author. E-mail: kimadmin@wi.mit.edu.

[1] Abbreviations: MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight; CD, circular dichroism; CSP and LCSP, Crick supercoil parameters and local Crick supercoil parameters, respectively; T_m, midpoint of thermal unfolding transition; RMSD, root-mean-square deviation.

10.1021/bi002829w CCC: $20.00
Published on Web 05/02/2001

mM Na_2HPO_4, pH 7.0, at 4 °C with an AVIV 62A or AVIV 62A/DS CD spectrometer (AVIV Instruments, Lakewood, NJ) using a 10 mm cell. Peptide concentrations were determined from the tyrosine absorption maxima at 276 nm in 6 M guanidine hydrochloride (*16*). Peptide concentrations were 10 μM for CD experiments. Thermal melts were performed using the same buffer conditions and peptide concentrations. Data for the thermal melts were acquired by averaging the CD signal over 30 s with a 90 s equilibration time after each 2 °C step. The midpoint of the thermal unfolding transition (T_m) reported is the maximum of the first derivative of the thermal melt data.

Sedimentation equilibrium experiments were performed in the same buffer conditions as for CD measurements using a Beckman XLA-90 analytical ultracentrifuge (Beckman Coulter, Palo Alto, CA). Sedimentation equilibrium experiments were performed at peptide concentrations of 20, 50, and 100 μM. Data were collected after spinning for 18−24 h at speeds ranging from 18 000 to 35 000 rpm at 4 °C. For all speeds and time points, a second data set was acquired 3 h after the first. The two data sets were compared to ensure that the samples were at equilibrium. A single ideal species model was fit to the data. The partial specific volume was calculated from the residue-weighted average of the amino acid sequence (*17*). Solvent density was calculated from the solvent composition (*17*).

Crystallization and Data Collection. Crystallization conditions were found using a sparse matrix screen (Hampton Research, Laguna Niguel, CA). The appropriate conditions were further optimized to yield crystals large enough for diffraction studies. The final crystallization conditions were as follows: pVL_T, 0.1 M sodium cacodylate, pH 6.0, 0.2 M zinc acetate, 20% PEG 8000; pVL_S, 0.1 M sodium cacodylate, pH 7.0, 0.2 M zinc acetate, 16% PEG 8000; pV_TL, 0.1 M HEPES, pH 7.0, 25 mM cadmium sulfate, 1 M sodium acetate; pV_SL, 0.1 M MES, pH 6.5, 50 mM cesium chloride, 10% PEG 20,000. Three of the crystals were soaked for a brief period in stabilizing solutions containing mother liquor and added cryoprotectant (pVL_S, 40% MPD; pV_TL, 20% MPD; pV_SL, 40% MPD) before freezing. Crystals were flash-frozen and data collected at 100 K using either an in-house source (RU-300 rotating anode with an R-AXIS IV detector), beam-line X4A at the National Synchrotron Light Source (NSLS), Brookhaven, NY, with an R-AXIS IV detector, or beam-line 5.0.2 at the Advance Light Source (ALS), Lawrence National Laboratory, Berkeley, CA, with an ADSC-4 CCD detector (Table 2). Diffraction intensities were integrated using the programs DENZO and SCALEPACK (*18*) and reduced to structure factors with the program TRUNCATE from the CCP4 program suite (*19*).

Model Building and Refinement. Molecular replacement solutions were found using the program AMORE (*19*) with either the GCN4-p1 dimer (*14*) or the GCN4-pV_QL trimer (*20*) structure as models. For molecular replacement, the model structures were truncated to alanine at position 16(***a***) for all models and also at position 12(***d***) for pVL_S and pVL_T. In all cases, AMORE identified solutions using the GCN4-pV_QL trimer with significantly higher correlation and lower *R*-factors than those found using the GCN4-p1 dimer. This confirmed that asymmetric units of the crystals contained a coiled-coil trimer.

For GCN4-pVL_S, the AMORE solution underwent simulated annealing using the program CNS (*21*) and then modified to a polyserine model, with the exception of the core residue, Ser 12(***d***), which was modified to alanine. The polyserine model was refined and water added using the program REFMAC (*19*). Water molecules were removed and side chains manually added to the model using program O (*22*) as the densities for the side-chain rotamers became apparent. The model was further refined with multiple rounds of manual improvement using O, followed by positional and *B*-factor refinement using CNS.

After finding the molecular replacement solution using AMORE, the pVL_T, pV_TL, and pV_SL structures were refined by 10 independent rounds of simulated annealing. A $2F_o-F_c$ map was computed in CNS using the structures from all 10 rounds. Density interpretation and manual improvement were done with the program O. Crystallographic refinement was done using the CNS program. The final structures were checked using simulated annealing omit maps (*23*). Density was not always apparent for residues 31−33 for some chains. When the density for these terminal residues was not observed, they were omitted from the model. In addition, the density for some side chains was not observed, and the occupancy on these side-chain atoms was reduced to zero.

RESULTS

Solution Studies. The peptide GCN4-pVL was used as the reference peptide for our studies. GCN4-pVL is equivalent to GCN4-p1 with the ***a***-position asparagine at residue 16 [denoted Asn 16(***a***)] mutated to valine. GCN4-pVL has valines at the ***a*** positions (except for a single methionine) and leucines at the ***d*** positions (Figure 1). The single methionine, Met 2(***a***), is present in wild-type GCN4-p1 and was not changed in these studies. As shown previously (*7*), GCN4-pVL is stable [T_m = 95 °C at 10 μM as determined by circular dichroism (CD)] and forms a mixture of dimers and trimers as determined by sedimentation equilibrium analysis (Table 1). Eight variants of GCN4-pVL with four different polar residues (Asn, Gln, Ser, or Thr) at either an ***a*** (residue 16) or a ***d*** (residue 12) position were synthesized (Figure 1). The resulting peptides are termed pV_XL for ***a***-position variants and pVL_X for ***d***-position variants, where X is N, Q, S, or T, denoting the polar substitution. Wild-type GCN4-p1, equivalent to pV_NL, is not renamed here.

Sedimentation equilibrium analyses of the peptides reveal only two variants with a unique oligomeric state (Table 1). pVL_T is a trimer, and as noted previously (*13*), wild-type GCN4-p1 is a dimer. These results correlate with the residue distributions observed in coiled-coil databases: asparagines at the ***a*** position are more often found in dimeric coiled coils, while threonines at the ***d*** position are more frequent in trimeric coiled coils. Other buried polar residues do not exhibit clear oligomeric preferences as judged by sedimentation equilibrium, and the slight preferences observed do not correlate with database residue distributions. Polar substitutions in the ***d*** position result in a greater variation in T_m and helical content than substitutions in the ***a*** position (Table 1). While pVL_N and pVL_Q seem to favor dimer formation, the low helical content and low T_m of these peptides suggest that there may be a significant portion of predominantly unfolded monomer in solution, skewing the data toward

Table 1: Circular Dichroism and Sedimentation Equilibrium Data for GCN4-pVL and Polar Variants

peptide	polar substitution	$-[\theta]_{222}$ ($\times 10^3$ deg cm^2 dmol^{-1})	T_m (°C)	$MW_{obs}/MW_{monomer}$[a] 20 μM	50 μM	100 μM	database distributions[b] dimer (%)	trimer (%)
pVL	none	30.0	94	2.5	2.6	2.8	–	–
pV_NL[c]	Asn 16(***a***)	30.9	54	2.0	2.1	2.1	3.9	0.7
pV_TL	Thr 16(***a***)	30.3	68	2.4	2.4	2.6	0.7	5.6
pV_SL	Ser 16(***a***)	28.5	52	2.3	2.3	2.4	2.0	5.1
pV_QL	Gln 16(***a***)	30.7	56	2.3	2.3	2.4	0.9	2.4
pVL_N	Asn 12(***d***)	15.2	22	2.0	2.2	2.3	1.3	3.1
pVL_T	Thr 12(***d***)	30.5	62	3.0	3.1	3.2	3.6	8.2
pVL_S	Ser 12(***d***)	32.0	44	2.6	2.6	2.8	4.1	2.7
pVL_Q	Gln 12(***d***)	25.8	26	2.1	2.1	2.2	1.8	1.6

[a] $MW_{obs}/MW_{monomer}$ is the molecular weight determined from sedimentation equilibrium analysis divided by the expected mass of a monomer. [b] Database distributions are the frequencies of the specified polar residue at a given heptad position in dimeric and trimeric coiled-coil databases, respectively (*2*). [c] pV_NL is equivalent to GCN4-p1.

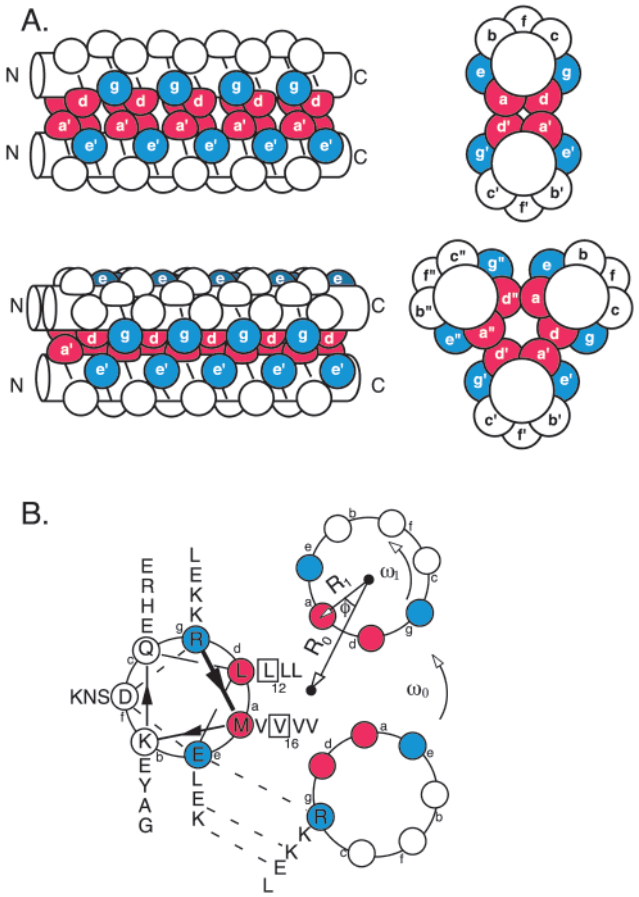

FIGURE 1: (A) Schematic view of dimeric (top) and trimeric (bottom) coiled coils. The ***a*** and ***d*** core positions are indicated in red and the ***e*** and ***g*** positions in blue. (B) Coiled-coil schematic of GCN4-pVL peptide as a trimer. Residues 16(***a***) and 12(***d***), the positions of the polar substitutions, are boxed. Interhelical ***g*** to ***e′*** salt bridges, indicated with dashed lines, are from the GCN4-pI_QI structure (*28*). The Crick supercoil parameters (CSP) are indicated as: R_0 and R_1, the radius for the supercoil and α-helix, respectively; ω_0 and ω_1, the supercoil pitch and α-helix pitch, respectively; ϕ, the C_α phase angle, defined as the angle between vectors from the α-helix center to the supercoil center and to the C_α of the ***a***-position residue.

lower molecular weight species. Thus, it is likely that the molecular weight of the folded form of these peptides is underestimated by sedimentation equilibrium analysis. In general, our data are in agreement with the oligomeric preferences seen in other model coiled-coil systems (*11, 12*). The data for pV_QL, which is found here and elsewhere (*10*) to exist as a mixture of dimers and trimers, are in conflict with other studies which suggest that ***a***-position glutamines are sufficient to specify trimeric coiled coils (*11*). These differences may illustrate the importance of context when accounting for the effects of buried polar residues on structural uniqueness.

Crystal Structures. The structures of coiled-coil trimers with threonines and serines at either the ***a*** or the ***d*** positions (pV_TL, pV_SL, pVL_T, pVL_S) were determined at a resolution of 1.9 Å or higher (Table 2). Although three of these four peptides exist as mixed oligomeric species in solution, all four peptides crystallized as a trimer. All four structures show an approximate, noncrystallographic, 3-fold symmetry. As noted, pVL_T has unique trimeric specificity and is the most stable ***d***-position polar variant (Table 1). The structure of pVL_T (Figures 2A and 3) shows symmetric packing of the Thr 12(***d***) residues with an ordered hydrogen bond network in the core. The hydroxyl group of each Thr 12(***d***) makes an identical *i* to *i* − 4 backbone hydrogen bond to the peptide carbonyl of Lys8(***g***), directing the threonine methyl groups into the core (Figure 3). Four structured water molecules, between the Val 9(***a***) and Thr 12(***d***) layers, form the vertexes of a trigonal pyramid. The three outer water molecules form potential hydrogen bonds to the Val 9(***a***) backbone carbonyl. As there are more potential hydrogen bonds than hydrogen atoms in this network (12 vs 11), partial water molecule occupancies and transient hydrogen bonds are expected.

Similar to the ***d***-position threonine residues, ***d***-position serines in pVL_S are arranged symmetrically. In contrast to the ***d***-position threonines, the hydroxyl groups of the core serines are directed toward the center to form an internal hydrogen bond network (Figures 2B and 3). No structured water molecules are observed in the core structure. In the vicinity of the serines there is a local decrease in the superhelical radius of about 0.5 Å (Figure 4A), changes in the supercoil pitch (Figure 4B), and an asymmetric rotamer distribution at Val 9(***a***) (Figure 3).

In contrast to the ***d***-position polar residues, the ***a***-position serines and threonines are not packed with 3-fold symmetry and are in different rotamers. The serines in pV_SL have the most structural variability, with one hydroxyl group directed toward the exterior, another making an *i* to *i* − 3 hydrogen bond to the peptide carbonyl of Leu 13(***e***), and the third with the partial occupancy of two alternate rotamers, either directed to the core or pointing outward toward the bulk

Table 2: Data Collection, Refinement Statistics, and Crick Supercoil Parameters

crystal	pVL_T Thr 12(***d***)	pVL_S Ser 12(***d***)	pV_TL Thr 16(***a***)	pV_SL Ser 16(***a***)
space group	$P2_1$	$P2_1$	$P2_12_12_1$	*P1*
a, *b*, *c* (Å)	25.7, 42.2, 42.7	35.5, 28.6, 44.5	35.8, 46.3, 51.4	25.6, 25.8, 34.4
α, *β*, *γ* (deg)	90.0, 98.9, 90.0	90.0, 109.6, 90.0	90.0, 90.0, 90.0	86.5, 84.5, 79.4
data collection				
X-ray source	in-house	in-house	ALS	NSLS
resolution (Å)	1.86	1.86	1.70	1.80
total observations	80296	70770	63455	23291
unique reflections	7316	7255	9768	7378
completeness (%)	95.3 (94.4)[f]	94.4 (93.7)[f]	99.3 (99.2)[f]	92.6 (89.0)[f]
R_{merge}[a]	0.046 (0.306)[f]	0.065 (0.315)[f]	0.048 (0.283)[f]	0.059 (0.180)[f]
refinement				
R_{cryst}[b]	0.227	0.211	0.243	0.220
R_{free}[c]	0.278	0.258	0.279	0.269
RMSD bonds (Å)	0.010	0.014	0.014	0.009
RMSD angles (deg)	1.3	1.9	1.6	1.3
Crick supercoil parameters (CSP)[d]				
α carbon phase, ϕ (deg)	16.1	18.2	18.9	19.7
supercoil radius, R_0 (Å)	6.25	6.03	6.25	6.29
supercoil pitch (residues/turn)	101	88	98	99
supercoil deviation, RMSD (Å)[e]	0.292	0.408	0.256	0.265

[a] $R_{merge} = \sum\sum_j|I_j(hkl) - \langle I(hkl)\rangle|/\sum\sum_j|\langle I(hkl)\rangle|$, where I_j is the intensity measurement for reflection (*hkl*) and $\langle I\rangle$ is the mean intensity over *j* measurements. [b] R_{cryst} (R_{free}) $= \sum||F_{obs}(hkl)| - |F_{calc}(hkl)||/\sum|F_{obs}(hkl)|$, where F_{obs} and F_{calc} are observed and calculated structure factors, respectively. No σ-cutoff was applied. [c] Ten percent of the reflections were excluded from refinement and used to calculate R_{free}. [d] The Crick supercoil parameters (CSP) were determined by fitting an ideal coiled coil to the central three heptads of the final structure using a user-defined CHARMM function (see Results) (*24*, *25*). [e] Supercoil RMSD is between the C_α atoms of the model coiled coil and the observed structure. [f] Values in parentheses correspond to the highest resolution shell: 1.94−1.86 Å for pVL_T and pVL_S, 1.76−1.70 Å for pV_TL, and 1.86−1.80 Å for pV_SL.

solvent (Figure 2D). Two coordinated water molecules are observed in the core. The hydroxyl groups of threonine in pV_TL make intrahelical peptide carbonyl contacts, but with some variation: all three make an $i - 3$ intrahelical contact, but one also makes the alternative $i - 4$ contact observed for ***d***-position threonines (Figure 2C).

Supercoil Parameters. The geometry of a regular coiled coil is defined by a simple set of parameters [denoted Crick supercoil parameters (CSP)] that include the radius (R_0 and R_1) and pitch (ω_0 and ω_1) for both the supercoil and α-helix, respectively, and the ***a***-position C_α phase angle (ϕ) (Figure 1) (*24*, *25*). To examine how single-residue substitutions affect the coiled-coil geometry, we determined the CSP values of each coiled coil (Table 2) using the central three heptads (residues 5−25) (*24*, *25*). We fit only the central three heptads because the regions outside these residues are less structured and fit poorly to ideal supercoil parameters (Figure 4D). For each set of superhelical parameters, we determined the C_α root-mean-square deviation (RMSD) between the actual structure and a model coiled coil built using the CSP values. Three of the structures fit closely to the model coiled coil, with RMSDs less than 0.29 Å. pVL_S, however, deviates more substantially, with an RMSD of 0.41 Å (Table 2).

To investigate local effects of the buried polar residues, we utilized a heptad-specific, local Crick supercoil parameter (LCSP), determined by fitting only seven residues at a time. This local fitting of coiled-coil parameters—sliding the window in one-residue increments—allows detection of changes in structure not readily apparent by visual inspection or global fitting (Figure 4). Our method extends previously described local fitting methods (*26*) by simultaneously determining multiple parameters for the coiled coil. Polar residues in the ***a*** position seem to have a smaller effect on the LCSP than ***d***-position polar residues, as judged by the observation that LCSP values for the ***a***-position variants follow a common trend, while ***d***-position LCSP values differ more widely (Figure 4). As noted earlier, ***d***-position polar variants also vary more in stability than ***a***-position polar variants. Local structural variations include a decrease in R_0 centered around the buried serine residues in pVL_S, and variations in ϕ for both pVL_S and pVL_T. The observation that ***a***-position polar residues have a smaller effect on supercoil parameters is also supported by comparisons of GCN4-pV_QL (*10*) with the structures shown here, and between the GCN4-pII (*27*) and GCN4-pI_QI (*28*) trimer structures (data not shown).

Cavity Volumes in the Core. An analysis of coiled-coil trimers reveals sizable cavities in the cores (Figure 5). The cavities vary in volume, position, and structured water molecules, particularly around the buried polar positions. Generally, cavities are found in trimeric coiled coils between the ***a*** and ***d*** layers, although pV_SL also has a cavity that extends from the position 12(***d***) layer to the position 19(***d***) layer (Figure 5D). There are no cavities observed between the position 12(***d***) layer and the position 16(***a***) layer for either pVL_T or pVL_S (Figure 5A,B).

One might expect ***d***-position serine variants to have larger cavity volumes than the threonine variants. However, due to differences in supercoil radius and pitch (Table 2 and Figure 4), the cavities for pVL_S are in fact smaller than the corresponding cavities in pVL_T (Figure 5A,B). The largest cavity in pVL_T contains structured water molecules. The ***a***-position polar variants vary more in cavity size and position. Although pV_SL has two fewer cavities than pV_TL, the extended cavity in pV_SL accounts for space that contains two cavities in pV_TL (Figure 5C,D). This extended cavity is the largest in all four structures and contains two structured water molecules. The cavities of pV_TL are the smallest and most regular.

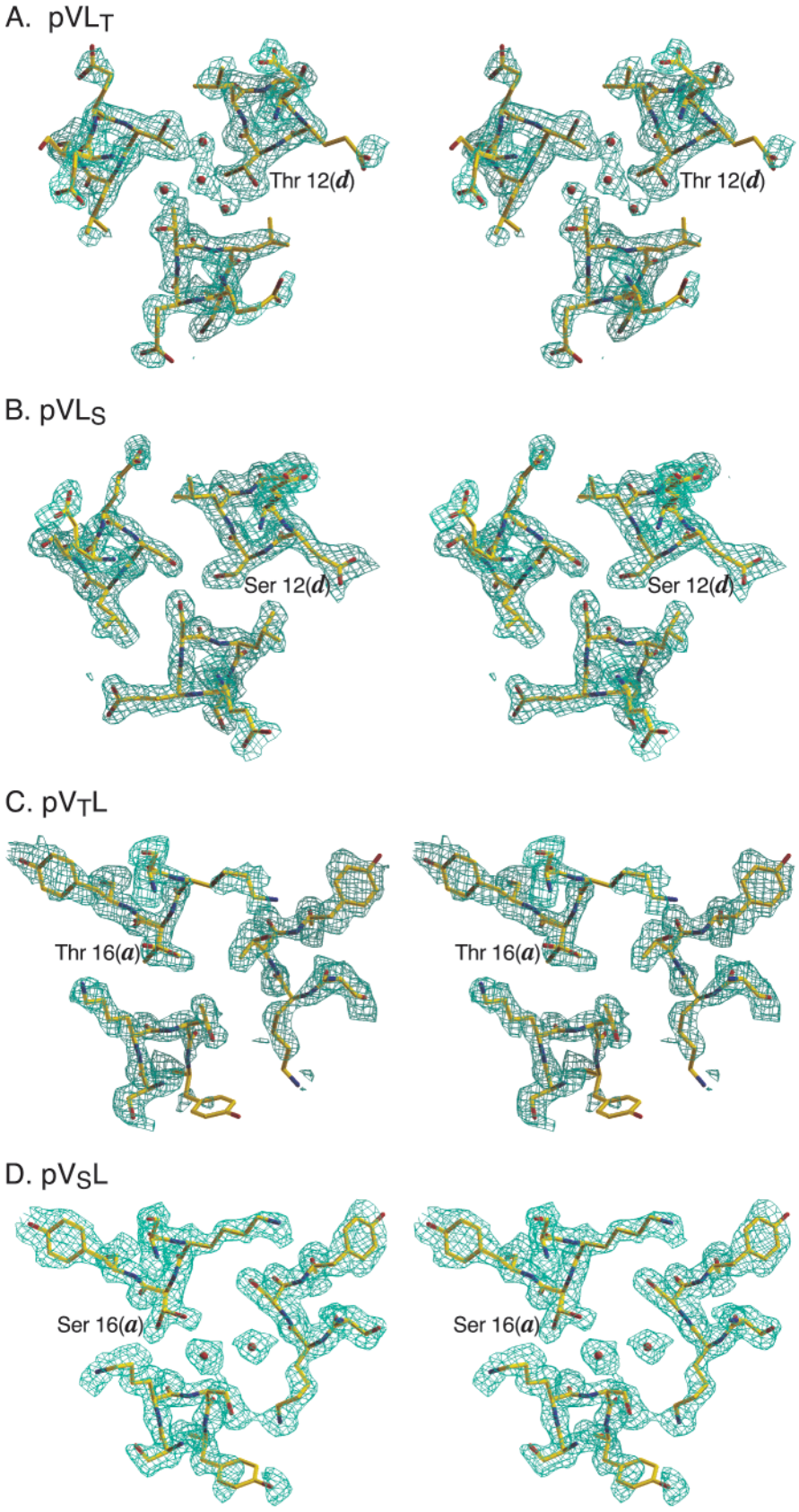

FIGURE 2: Stereodiagrams of selected area composite omit electron density maps superimposed onto the final refined structures. Maps are contoured at 1σ. Water molecules are represented as red spheres. (A) pVL$_T$; (B) pVL$_S$; (C) pV$_T$L, with alternate rotamers for the labeled Thr 16(***a***); and (D) pV$_S$L, with alternate rotamers for the labeled Ser 16(***a***).

Interhelical Surface Contacts. Ionic surface interactions between helices are important in both dimeric and trimeric coiled coils and are determinants of coiled-coil specificity (*29*, *30*). Nevertheless, it is commonly observed that not all possible interhelical interactions are formed. The four structures presented here all have identical surface residues,

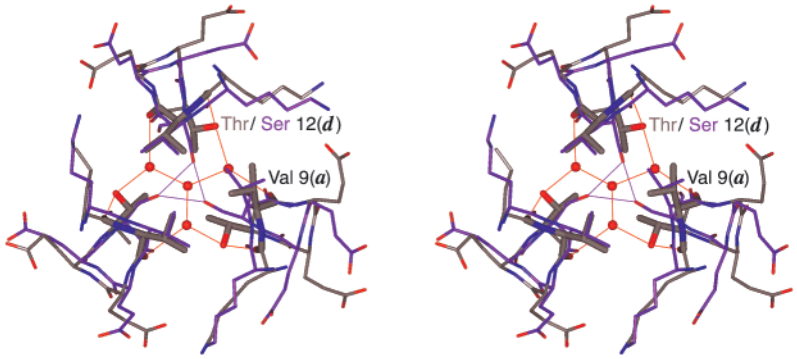

FIGURE 3: Stereodiagram of pVL_S (residues 8−12, purple backbone) superimposed on pVL_T (gray backbone) including the polar ***d*** position (residue 12). Val 9(***a***) and Thr 12(***d***) of pVL_T are shown as thicker rods. A hydrogen bond network, shown in red, is observed between water molecules (red spheres), all Thr 12(***d***) residues, and the peptide carbonyl oxygens of residues 8 and 9 in the pVL_T structure. The decreased supercoil radius of pVL_S is apparent. The labeled Val 9(***a***) of pVL_S adopts a unique rotamer not seen for any other ***a***-position valines.

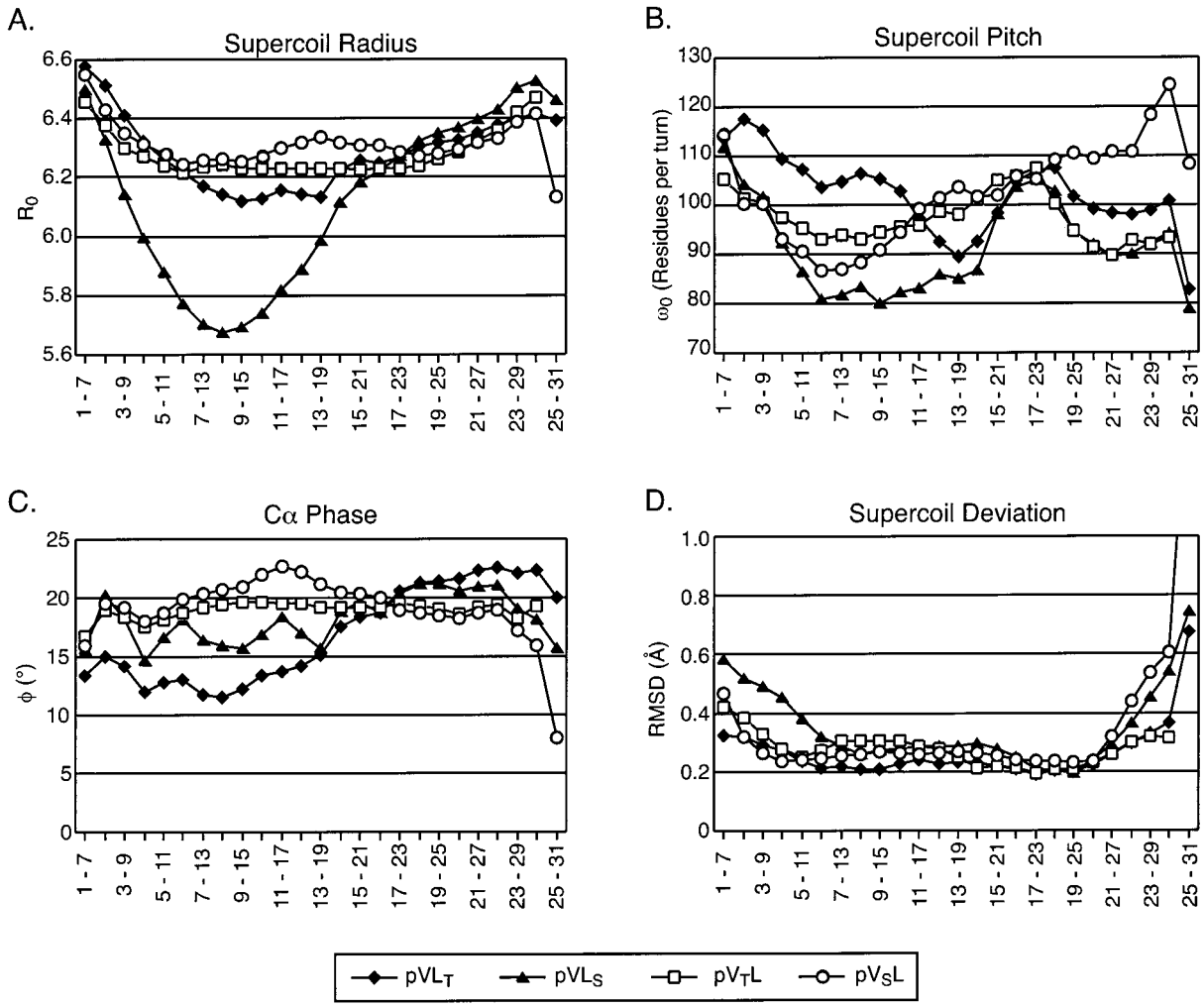

FIGURE 4: Local Crick supercoil parameters (LCSP), as defined in Figure 1, plotted for pVL_T (closed diamonds), pVL_S (closed triangles), pV_TL (open squares), and pV_SL (open circles). The indicated heptads were used to determine LCSP values. (A) Supercoil radius (R_0) in angstroms; (B) supercoil pitch (ω_0) in residues per turn; (C) ϕ in degrees; and (D) deviation from ideal supercoil.

but different surface contact patterns (Table 3). Some seemingly repulsive charge−charge interactions are allowed because of the presence of a third, mutually attractive, charged group forming a salt bridge triad (e.g., the A-chain Arg 1 to B-chain Lys-3 interaction in pVL_S is mediated by the B-chain Glu-6 carboxylate).

While the variation in surface residue contacts may be a result of crystal packing, all four trimers required different crystallization conditions and crystallized in different space groups (Table 2), suggesting that differences between the four trimers exist in solution. Differences in supercoil radius and pitch may partially account for variations in surface interhelical interactions, as subtle shifts in these parameters may be enough to bring potentially interacting residues (e.g., ***e*** and ***g*** positions) closer together or further apart. These differences illustrate that *ab initio* design and structure

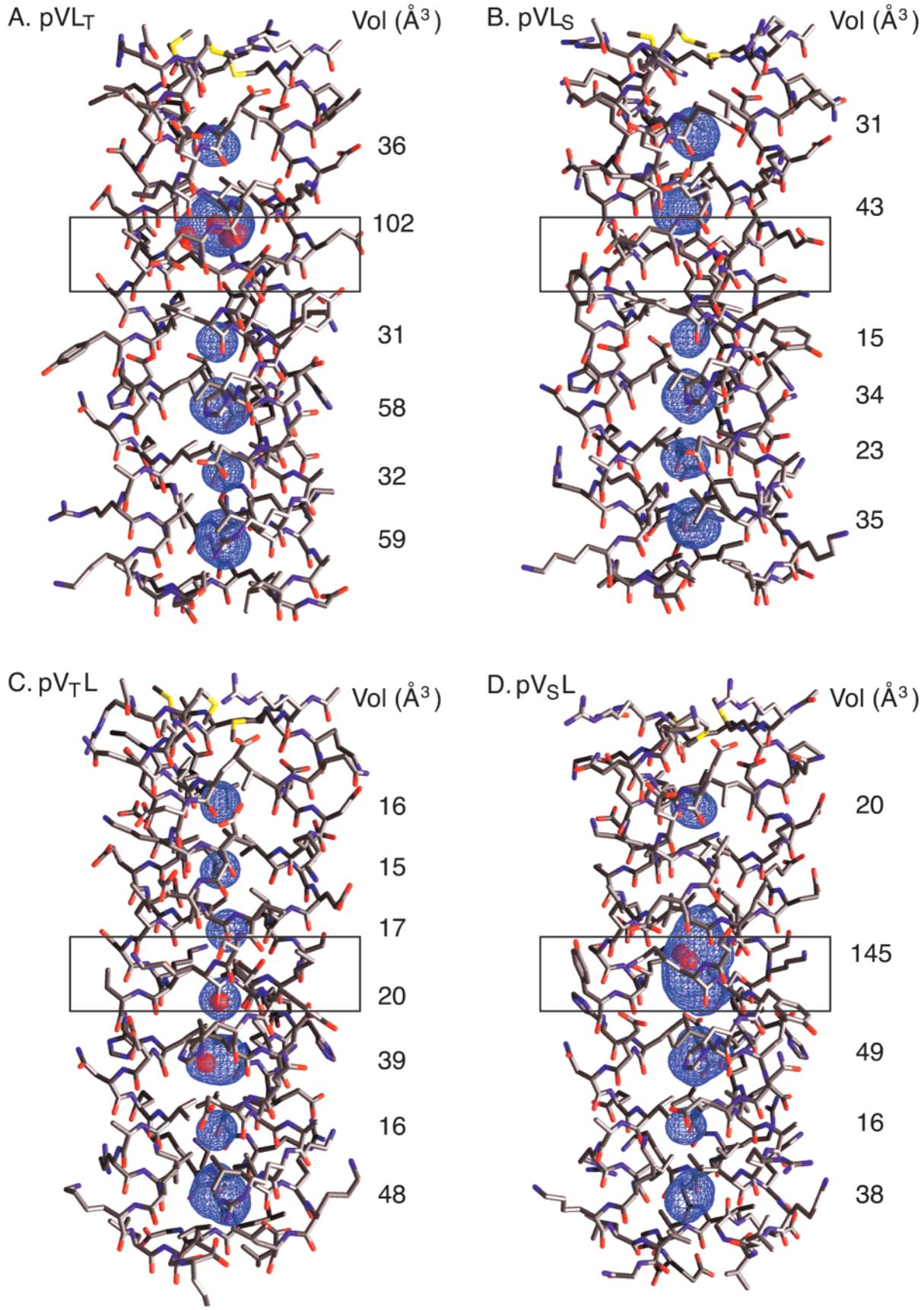

FIGURE 5: Cavities in the cores of coiled-coil trimers: (A) pVL$_T$, (B) pVL$_S$, (C) pV$_T$L, and (D) pV$_S$L. Vol, cavity volume (Å^3). Cavities were from molecular surface calculated with the program GRASP (*36*). Water molecules were omitted for these calculations. The cross-sections containing the core polar residues (see also Figure 2) are indicated with boxes. There are four water molecules (red spheres) in the cavity of pVL$_T$ (A), two in pV$_T$L (C), and two in pV$_S$L (D).

Table 3: Interhelical Interactions between Charged and Polar Groups[a]

interhelical contacts		interaction distances (Å)			
type	residues	pVL_T	pVL_S	pV_TL	pV_SL
g*−*e′	Arg A1($N_{\eta1}$)−Glu B6 ($O_{\epsilon1}$)	2.8	3.6	3.0	
g*−*e′	Arg B1 ($N_{\eta2}$)−Glu C6 ($O_{\epsilon1}$)	2.7	3.0	3.3	2.7
g*−*e′	Arg C1 ($N_{\eta2}$)−Glu A6 ($O_{\epsilon2}$)	W		3.1	2.7
g*−*e′	Lys A15 (N_ζ)−Glu B20 ($O_{\epsilon1}$)	2.6			3.7
g*−*e′	Lys C15−Glu A20	W	W		W
g*−*e′	Glu A22 ($O_{\epsilon1}$)−Lys B27 (N_ζ)	2.9	2.9	3.6	2.7
g*−*e′	Glu B22 ($O_{\epsilon2}$)−Lys C27 (N_ζ)	2.7	3.0	2.7	3.5
g*−*e′	Glu C22 ($O_{\epsilon1}$)−Lys A27 (N_ζ)	2.7	2.7	2.8	
g*−*b′	Arg A1 ($N_{\eta2}$)−Lys B3 (N_ζ)		3.0		
g*−*b′	Lys A8 (N_ζ)−Glu B10 ($O_{\epsilon2}$)	2.9		2.9	
g*−*b′	Lys B8 (N_ζ)−Glu C10 ($O_{\epsilon2}$)	2.5	W		
g*−*b′	Lys C8 (N_ζ)−Glu A10 ($O_{\epsilon2}$)	3.7			
g*−*a′	Lys A15 (N_ζ)−Ser B16 (O_γ)				3.4
g*−*a′	Lys C15 (N_ζ)−Thr A16 ($O_{\gamma1}$)			2.8	
e*−*g′	Glu A6 ($O_{\epsilon1}$)−Lys C8 (N_ζ)		3.8		
c*−*e′	Gln A4 ($N_{\epsilon2}$)−Glu B6 ($O_{\epsilon2}$)		3.9	W	
c*−*e′	Arg A25 ($N_{\eta2}$)−Lys B27 (N_ζ)	2.8	W		
c*−*e′	Arg B25 ($N_{\eta2}$)−Lys C27 (N_ζ)				2.6
c*−*b′	Glu A11 ($O_{\epsilon2}$)−Tyr B17 (O_η)		2.8		

intrahelical contacts		interaction distances (Å)			
type	residues	pVL_T	pVL_S	pV_TL	pV_SL
b*−*e	Lys B3 (N_ζ)−Glu B6 ($O_{\epsilon2}$)		3.0		
g*−*c	Glu A22 ($O_{\epsilon2}$)−Arg A25 ($N_{\eta2}$)	2.9	W		

[a] Interactions listed as ***g*−*e′*** indicate an interaction between the ***g***-position residue of one chain and the ***e***-position residue of an adjacent chain. Interacting residues are identified by amino acid type, chain designation (A, B, and C for the three chains), and residue number. Interaction distances are shown. A "W" indicates that a potential salt bridge interaction is mediated through a structured water molecule, but that the residues are too far apart to directly interact. The intrahelical contacts listed mediate an otherwise repulsive interhelical contact.

prediction of coiled coils may need to account for alternative ionic interactions among the surface residues.

DISCUSSION

There are several possible reasons for the presence of buried polar residues in natural coiled coils. Buried polar residues may help specify the oligomeric state in sequences that would otherwise exist as a mixture of dimers, trimers, or higher order structures (*6*, *8*, *10*, *20*). We find that determining oligomeric specificity may be a significant role of only a few buried polar residues, as only an ***a***-position asparagine and a ***d***-position threonine are found to specify unique oligomeric states. In contrast, other polar residues do not specify a unique oligomeric preference. As hydrophobic core sequences are an important determinant of oligomeric state (*7*), the effects of polar residues on oligomeric specificity may be limited and depend on the identities of the other core residues.

Buried polar residues in coiled coils likely serve important roles in determining structural uniqueness. Interactions between polar atoms may compensate for the unfavorable desolvation energy of buried polar residues. These interactions are commonly observed in protein cores and have been proposed to be an important source of structural specificity (*8*, *31*). In the case of coiled coils, the selection of polar−polar pairings could ensure that composite α-helices are aligned correctly in homotypic coiled coils (*8*, *14*). In an analogous manner, polar−polar pairings may be a discriminating factor in partner selection in heterotypic coiled coils. A preference for polar−polar pairing has been confirmed for ***a***-position asparagines in a number of genetic selection experiments (*32*−*34*). In the structures presented here, the core-position threonines and serines satisfy their hydrogen bond requirements either with residues on adjacent strands (often meditated though water molecules), with the peptide backbone, with external residues and solvent, or combinations of the above. This ability to satisfy hydrogen bond requirements through several means may allow these polar residues flexibility in forming heterotypic coiled coils as they may pair with both polar and hydrophobic residues. The potential dual nature of polar residues suggests an avenue of further research in determining factors involved in heterotypic coiled-coil assembly.

Finally, a comparison of the four closely related peptide structures presented here illustrates the plasticity of coiled coils in response to polar substitutions. These structures show two distinct mechanisms for accommodating polar residues in the coiled-coil core. The first method is the filling of cavities with structured water molecules, and the second involves changes in the supercoil parameters, particularly the supercoil radius. Additionally, single polar residue core substitutions lead to variation in surface contacts. These variations may need to be accounted for in computational methods that have been successful at predicting and designing coiled coils, but that used cores that lacked polar residues (*24*, *35*). In addition, these computational methods have assumed constant supercoil parameters along the length of the coiled coil. Our results show that superhelical parameters can be fit using local Crick supercoil parameters (LCSP), even as they vary along the length of the coiled coil. Utilization of such local fitting methods could play an important role in future coiled-coil design efforts.

ACKNOWLEDGMENT

We thank Dr. Amy Keating for useful discussion and intellectual contributions, Chris Liu and other members of the Kim lab for comments on the manuscript, Dr. Craig Ogata and the staff at the X4A beamline at the National Synchrotron Light Source (Brookhaven National Laboratory, Upton, NY) for beamtime allocation and help, and Dr. Thomas Earnest for beamtime allocation and help at the 5.0.2 beamline at the Advanced Light Source (Berkeley, CA). We are grateful to Leslie Gaffney for assistance in preparation of the figures and manuscript, and Michael Burgess and James Pang for peptide synthesis.

REFERENCES

1. Cohen, C., and Parry, D. (1994) *Science 263*, 488−489.
2. Wolf, E., Kim, P. S., and Berger, B. (1997) *Protein Sci. 6*, 1179−1189.
3. Parry, D. A. (1982) *Biosci. Rep. 2*, 1017−1024.
4. Berger, B., Wilson, D. B., Wolf, E., Tonchev, T., Milla, M., and Kim, P. S. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 8259−8263.
5. Lupas, A., van Dyke, M., and Stock, J. (1991) *Science 252*, 1162−1164.
6. Woolfson, D. N., and Alber, T. (1995) *Protein Sci. 4*, 1596−1607.
7. Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993) *Science 262*, 1401−1407.
8. Lumb, K. J., and Kim, P. S. (1995) *Biochemistry 34*, 8642−8648.

9. Oakley, M. G., and Kim, P. S. (1998) *Biochemistry 37*, 12603−12610.
10. Gonzalez, L., Jr., Woolfson, D. N., and Alber, T. (1996) *Nat. Struct. Biol. 3*, 1011−1018.
11. Wagschal, K., Tripet, B., Lavigne, P., Mant, C., and Hodges, R. (1999) *Protein Sci. 8*, 2312−2329.
12. Tripet, B., Wagschal, K., Lavigne, P., Mant, C., and Hodges, R. (2000) *J. Mol. Biol. 300*, 377−402.
13. O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1989) *Science 243*, 538−542.
14. O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991) *Science 254*, 539−544.
15. Lockhart, D. J., and Kim, P. S. (1992) *Science 257*, 947−951.
16. Edelhoch, H. (1967) *Biochemistry 6*, 1948−1954.
17. Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) in *Analytical Ultracentrifugation in Biochemistry and Polymer Science* (Harding, S. E., Rowe, A. J., and Horton, J. C., Eds.) pp 90−125, Royal Society of Chemistry, Cambridge, U.K.
18. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol. 276*, 307−326.
19. Collaborative Computational Project (1994) *Acta Crystallogr. 4*, 760−763.
20. Gonzalez, L., Jr., Brown, R. A., Richardson, D., and Alber, T. (1996) *Nat. Struct. Biol. 3*, 1002−1009.
21. Brunger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, N., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. D54*, 905−921.
22. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
23. Brunger, A. T., Adams, P. D., and Rice, L. M. (1997) *Structure 5*, 325−336.
24. Harbury, P. B., Tidor, B., and Kim, P. S. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 8408−8412.
25. Crick, F. H. C. (1953) *Acta Crystallogr. 6*, 689−697.
26. Seo, J., and Cohen, C. (1993) *Proteins: Struct., Funct., Genet. 15*, 223−234.
27. Harbury, P. B., Kim, P. S., and Alber, T. (1994) *Nature 371*, 80−83.
28. Eckert, D. M., Malashkevich, V. M., and Kim, P. S. (1998) *J. Mol. Biol. 284*, 859−865.
29. O'Shea, E. K., Rutkowski, R., and Kim, P. S. (1992) *Cell 68*, 699−708.
30. Nautiyal, S., Woolfson, D. N., King, D. S., and Alber, T. (1995) *Biochemistry 34*, 11645−11651.
31. Cordes, M. H. J., Davidson, A. R., and Sauer, R. T. (1996) *Curr. Opin. Struct. Biol. 6*, 3−10.
32. Zeng, X., Herndon, A. M., and Hu, J. C. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 3673−3678.
33. Arndt, K., Pelletier, J., Muller, K., Alber, T., Michnick, S., and Pluckthun, A. (2000) *J. Mol. Biol. 295*, 627−639.
34. Pelletier, J., Arndt, K., Pluckthun, A., and Michnick, S. (1999) *Nat. Biotechnol. 17*, 683−690.
35. Harbury, P. B., Plecs, J., Tidor, B., Alber, T., and Kim, P. S. (1998) *Science 282*, 1462−1467.
36. Nicholls, A., Bharadwaj, R., and Honig, B. (1003) *Biophys. J. 64*, A166.

BI002829W