Protein Science (1997), 6:80–88. Cambridge University Press. Printed in the USA.

The crystal structure of the designed trimeric coiled coil coil-V_aL_d: Implications for engineering crystals and supramolecular assemblies

NANCY L. OGIHARA,[1] MANFRED S. WEISS,[1,3] WILLIAM F. DEGRADO,[2] AND DAVID EISENBERG[1]

[1]UCLA-DOE Laboratory of Structural Biology and Molecular Medicine and Department of Chemistry and Biochemistry, University of California at Los Angeles, Los Angeles, California 90095-1570
[2]Johnson Research Foundation, Department of Biochemistry & Biophysics, University of Pennsylvania, School of Medicine, Philadelphia, Pennsylvania 19104-6059

(RECEIVED September 4, 1996; ACCEPTED October 3, 1996)

Abstract

The three-dimensional structure of the 29-residue designed coiled coil having the amino acid sequence acetyl-E VEALEKK VAALESK VQALEKK VEALEHG-amide has been determined and refined to a crystallographic *R*-factor of 21.4% for all data from 10-Å to 2.1-Å resolution. This molecule is called coil-V_aL_d because it contains valine in the *a* heptad positions and leucine in the *d* heptad positions. In the trigonal crystal, three molecules, related by a crystallographic threefold axis, form a parallel three-helix bundle. The bundles are stacked head-to-tail to form a continuous coiled coil along the *c*-direction of the crystal. The contacts among the three helices within the coiled coil are mainly hydrophobic: four layers of valine residues alternate with four layers of leucine residues to form the core of the bundle. In contrast, mostly hydrophilic contacts mediate the interaction between trimers: here a total of two direct protein–protein hydrogen bonds are found. Based on the structure, we propose a scheme for designing crystals of peptides containing continuous two-, three-, and four-stranded coiled coils.

Keywords: coiled-coiled; de novo design; three-helix bundle

Alpha-helical coiled coils are a common motif in proteins whose roles in cell structure and function have been reviewed (Cohen & Parry, 1986, 1990; Lupas, 1996). Coiled coils described so far consist of two, three, four, and recently, five α-helical bundles (Malashkevich et al., 1996). Atomic structures are now known for several coiled coils that are either synthetic peptides or motifs of existing proteins. Beyond their common feature of being coiled coils, these structures are varied, including parallel and antiparallel orientations. Structures containing parallel two-stranded coiled coils include mainly transcription factors such as GCN4 (O'Shea et al., 1991), human proto-oncogenes *fos* and *jun* (Glover & Harrison, 1995), and the yeast transcription activator GAL4 (Marmorstein et al., 1992). An antiparallel two-stranded coiled coil is found in the structure of seryl-tRNA synthetase (Fujinaga et al., 1993). Known structures of parallel and antiparallel four-stranded coiled coils are a synthetic sequence variant peptide derived from GCN4 (Harbury et al., 1993), the transcription regulator ROP (Banner et al., 1987), and the tetramerization domain of *lac* repressor (Fairman et al., 1995; Lewis et al., 1996). Trimeric coiled coils include the oligomerization domains of hemagglutinin (Wilson et al., 1981), C-type mannose binding protein (Weis & Drickamer, 1994), a second sequence variant of the GCN4 peptide whose *a* and *d* positions within the heptad repeat (which are normally mostly valines and leucines) are replaced with leucines and isoleucines, respectively (Harbury et al., 1993), and the recent structure of the envelope domain of Moloney murine leukemia virus (Fass et al., 1996). Examples of antiparallel trimeric coiled coil structures are the repeating unit of spectrin (Yan et al., 1994) and the de novo designed peptide coil-Ser (Lovejoy et al., 1993) whose structure is the basis for the current design, coil-V_aL_d.

Coiled coils have been the subject of extensive protein design studies because of their abundance in nature and their relative structural simplicity. Hodges et al. (1972) identified a repeating pattern of hydrophobic residues extending along the protein chain of rabbit skeletal α-tropomyosin. This led to the proposal that α-helical coiled coil proteins are stabilized by the interaction of hydrophobic residues at positions *a* and *d* of a repeating heptad sequence designated $(abcdefg)_n$ (McLachlan & Stewart, 1975). It was also observed that mostly polar or charged residues occur at

Reprint requests to: D. Eisenberg, Molecular Biology Institute, Box 951570, UCLA, Los Angeles, California 90095-1570; e-mail: david@mbi.ucla.edu.

[3]Present address: Institute of Molecular Biotechnology, Department of Structural Biology and Crystallography, Box 100813, D-07708 Jena, Germany.

the *e* and *g* positions of the heptad repeat (Hodges et al., 1972) allowing further stabilization of the coiled coil via electrostatic interactions, as well as influencing parallel or antiparallel orientations (Monera et al., 1993). Based on these principles, the peptides acetyl-K(LEALEGK)$_n$-amide (n = 4 or 5) were synthesized and their α-helical structure was verified by CD spectroscopy (Lau et al., 1984) (Fig. 1).

In a second generation design (O'Neil & DeGrado, 1990), a sequence with four heptad repeats (Betz et al., 1995) was used as a template for measuring the helical propensity of the 20 naturally occurring amino acids (Fig. 1). Position 14 (an *f* position) was chosen as the guest site for substitution because in the presumed two-stranded coiled coil, this position is solvent-accessible. Also, three of the four neighboring *b* and *c* positions were changed to Ala to minimize side-chain–side-chain interactions. In addition, the initial *a* position was substituted with a Trp for use as a spectroscopic probe.

The crystal structure of one of these molecules with Ser in position 14 (therefore named coil-Ser) was determined by Lovejoy et al. (1993). Unexpectedly, the structure was found to be an antiparallel trimer, rather than the parallel dimer of tropomyosin. Furthermore, in aqueous solution, the peptide exists in a noncooperative monomer–dimer–trimer equilibrium, although, under the crystallization conditions, it is fully trimeric (Betz et al., 1995). The formation of a trimer by coil-Ser is consistent with recent structural studies (Harbury et al., 1993) of sequence variants of the coiled coil region of GCN4. In these experiments, a peptide based on GCN4 containing all leucines in both *a* and *d* positions of the heptad repeats was shown to form a parallel trimer in solution. However, the hydrophobic core does not always impart structural uniqueness, as shown in a recent study (Lumb & Kim, 1995), where peptides with leucines in all *a* and *d* positions were shown to form tetramers in solution. The antiparallel orientation of coil-Ser was designed to be electrostatically unfavorable at neutral pH, and fluorescent derivatives of coil-Ser have been reported to form parallel trimers in aqueous solution at neutral pH (Wendt et al., 1995). However, the crystal structure of coil-Ser was found to be antiparallel and was obtained from a crystal grown at pH ~5 near the pK_a of the Glu carboxylates. This antiparallel orientation may be to some extent stabilized by hydrogen bonds between protonated and deprotonated Glu carboxylates, because spin-labeled derivatives of coil-Ser form parallel coiled coils in aqueous solution (W.F. DeGrado, unpubl. results). The antiparallel orientation was also attributed to the bulky indole side chain of the Trp residue, which was unable to pack as an all *a* layer at position 2 in the hydrophobic core, thus preventing an all parallel arrangement (Lovejoy et al., 1993).

In the current third generation design of molecules based on the original Hodges sequence, leucines in the *a* positions of the coil-Ser sequence are replaced by the smaller nonpolar residue valine to investigate effects on oligomerization (Betz et al., 1995). Valines thus have replaced the tryptophan in position 2 and the leucines at positions 9, 16, and 23, allowing formation of a trimeric coiled coil, as with coil-Ser. However, presumably because of the absence of the Trp residue, an all-parallel, rather than antiparallel structure is formed. The charged residues glutamate and lysine in the *e* and *g* positions were maintained to stabilize adjacent molecules of a parallel trimer via electrostatic interactions. The high-resolution X-ray structure of this peptide, referred to as coil-V$_a$L$_d$, has been determined and, from the results, we propose a scheme for a fourth generation of design.

Results

Quality of model

The present model of coil-V$_a$L$_d$ consists of 226 protein atoms, 18 solvent molecules, and one sulfate ion. The RMS deviation (RMSD) between the initial ideal α-helix search model and the final model is 2.3 Å. The final *R*-factor is 21.4% for all data in the resolution range 10–2.1 Å (Tables 1, 2). All atoms lie in good density of the $(2F_o - F_c)\alpha_c$ map, except for the amide group of Gln 17, the imidazole ring of His 28 and the C-terminal Gly 29, which appear to be flexible. The sulfate ion lies on a twofold axis, with two histidine and two lysine neighbors, related by crystallographic symmetry, linking together four different peptide molecules. The rather high *B*-factor of 49.7 Å^2 suggests that the sulfate position may not be fully occupied. This is in agreement with the presence of positive $(F_o - F_c)\alpha_c$ difference density at the NZ-atom of the neighboring Lys 22, indicating a possible second minor conformation away from the sulfate. With the limited resolution of 2.1 Å, we did not attempt to refine occupancies or alternative positions.

The stereochemical parameters are in excellent agreement with the parameter set derived by Engh and Huber (1991) (Table 1). A Luzzati plot (Luzzati, 1952) indicates a mean coordinate error of 0.25 Å (data not shown). Residues 1–27 are all in the most favored region in a Ramachandran diagram (Ramachandran et al., 1963). The backbone of His 28, which seems to be flexible, departs from the α-helix and is in a generously allowed region (Laskowski et al., 1993) of the Ramachandran diagram.

Main-chain conformation and H-bonding

The structure of the monomer has canonical α-helical hydrogen bonding patterns along the main chain: the carbonyl oxygen of residue (i) accepts a hydrogen bond from the amino nitrogen of residue (i + 4) throughout the sequence up to Lys 22-O. At the

	g	a	b	c	d	e	f	g	a	b	c	d	e	f	g	a	b	c	d	e	f	g	a	b	c	d	e	f	g
TM29	K	**L**	E	A	**L**	E	G	K	**L**	E	A	**L**	E	G	K	**L**	E	A	**L**	E	G	K	**L**	E	A	**L**	E	G	K
coil-Ser	E	**W**	E	A	**L**	E	K	K	**L**	A	A	**L**	E	S	K	**L**	Q	A	**L**	E	K	K	**L**	E	A	**L**	E	H	G
coil-V$_a$L$_d$	E	**V**	E	A	**L**	E	K	K	**V**	A	A	**L**	E	S	K	**V**	Q	A	**L**	E	K	K	**V**	E	A	**L**	E	H	G

Fig. 1. Sequence comparison of coil-Ser type peptides. TM29 (Lau et al., 1984) is the first generation of design based on the sequence of tropomyosin. Coil-Ser (Lovejoy et al., 1993) is the second generation of design whose sequence is based on TM29 and whose structure is the basis for the current work on coil-V$_a$L$_d$. Heptad positions are lettered above the sequences. Hydrophobic core residues in the *a* and *d* positions are shown in bold.

Table 1. *Summary of X-ray data collection for coil-V_aL_d and crystallographic refinement statistics for the final atomic model*

Data collection	
Space group	P321
Unit cell	$a = b = 33.64$ Å, $c = 40.53$ Å, $\alpha = \beta = 90.0$ degrees, $\gamma = 120.0$ degrees
Molecules per asymmetric unit	1
Resolution	2.1 Å
Total observations ($I/\sigma(I) > 0$)	35,980
Unique reflections ($I/\sigma(I) > 0$)	1,710
Completeness	99.7% (100.0%)[a]
R_{merge}[b]	0.091
Refinement	
Crystallographic R-factor	21.4%
R_{free}[c]	28.5
Number of reflections ($F/\sigma(F) > 0$)	1,691
Resolution range	10.0–2.1 Å
Completeness	99.7%
RMSD bonds	0.013 Å
RMSD angles	1.4 degrees
RMSD dihedrals	17.0 degrees
RMSD impropers	1.7 degrees
Number of protein atoms	226
Number of solvent atoms	18
Number of sulfate atoms	5
Average B-factor	
All atoms	18.6 Å^2
Protein atoms	16.8 Å^2
Main-chain atoms	13.2 Å^2
Side-chain atoms	20.7 Å^2
Water molecules	32.5 Å^2
Sulfate ion	49.7 Å^2

[a]Completeness in the outer resolution shell from 2.1–2.18 Å.
[b]$R_{merge} = \Sigma_{hkl}|I_i - \langle I\rangle|/\Sigma_{hkl}I_i$; conventional discrepancy R-factor for scaling intensities (I).
[c]$R_{free} = \Sigma_{hkl}||F_o| - |F_c||/\Sigma_{hkl}|F_o|$ for a test set of 10% of all reflections.

Table 2. *Progress of the refinement of the atomic model of coil-V_aL_d showing the round of refinement, the number of atoms and side chains (or solvent molecules) added, R-factor, R_{free} (free R-factor of Brünger, 1993), and the ratio ρ_{max}/ρ_{min} of electron density in the $(F_o - F_c)\alpha_c$ maps*[a]

Round	# Atoms	Changes	R	R_{free}	ρ_{max}/ρ_{min}
0	148	5 side chains	43.19	52.14	4.22/−3.87
1	161	5 side chains	35.79	40.24	4.38/−4.12
2	173	8 side chains	28.71	35.24	5.06/−4.22
3	196	7 side chains	26.54	34.99	5.62/−3.92
4	208	2 side chains	25.53	34.02	5.62/−4.15
5	217	2 side chains	25.48	32.78	5.88/−3.93
6	234	1 side chain, 8 waters	23.33	28.64	6.41/−3.93
7	241	7 waters	22.15	26.96	3.93/−5.20
8[b]	245	1 sulfate	21.54	28.51	4.20/−4.72
9	249	3 waters	20.64	28.46	4.25/−4.97
10[c]	249	—	21.37	—	3.91/−4.76

[a]Standard deviations for the $(F_o - F_c)\alpha_c$ maps is set to 1.0 in XPLOR.
[b]Even though R_{free} went up in this round, the density in the area where we built the sulfate ion clearly improved. We attribute this change in R_{free} at least partly to a random fluctuation because the test set contained only 166 reflections.
[c]All reflections used in the final round of refinement.

N terminus, the acetyl group is part of the α-helix with its O atom accepting a hydrogen bond from the Ala 4-N (Fig. 2). Residues 1–26 have α-helical main-chain dihedral angles with average values and standard deviations of $\phi = -63.9 \pm 5.7$ degrees and $\psi = -41.1 \pm 4.5$ degrees.

At the C terminus of the α-helix, the regular α-helical hydrogen bonding pattern breaks down in a shift toward a 3_{10}-helical conformation. As part of this, the hydrogen bond between Lys 22-O and Glu 26-N is elongated to 3.65 Å, and Leu 23-O accepts hydrogen bonds from both Leu 26-N and Glu 27-N (Fig. 2). The helical conformation terminates at Glu 27 ($\phi = -147.3$ degrees, $\psi = 151.3$ degrees). The carbonyl oxygen of Glu 24, which is not involved in any helical main-chain hydrogen bonds, instead accepts a hydrogen bond from the C-terminal amide group of Gly 29.

Stacking of helices occurs in the c direction of the crystal between the C-terminal region of one molecule and the N terminus of the next molecule to form a pseudo-continuous α-helix. In this stacking, Ala 25-O accepts a hydrogen bond from Glu 1-N of the next helix. Leu 26-O also accepts a regular α-helical hydrogen bond from Val 2-N. Additionally, a water molecule (Wat 32) mediates a third interhelical hydrogen bond between Leu 26-O and Glu 3-N, allowing the N terminus of the α-helix to be fully hydrogen bonded (Fig. 2). The oxygen atom of the N-terminal acetyl group accepts a hydrogen bond from Ala 4-N as described above.

Trimer formation

Each chain of the coil-V_aL_d trimer is related to the other two by the crystallographic threefold axis of the trigonal crystal, and is stabilized by hydrophobic interactions among triplets of valines and leucines at a and d positions (Fig. 3A,B). Trimer formation agrees with studies showing that coil-V_aL_d is a trimer in solution and that assembly is fully cooperative (Boice et al., 1996). A total of four layers of valine alternating with four layers of leucine residues (Fig. 3C) form the hydrophobic core of the trimeric protein. The four valines involved have a mean χ_1 angle of 172 degrees and are therefore in their most preferred *trans*-conformation (Ponder & Richards, 1987). The four leucines have mean (χ_1,χ_2) angles −78 degrees and 173 degrees, close to their most favored g^-t conformation (Dunbrack & Karplus, 1994).

The coiled monomer buries 40% of its surface area as it forms trimers, a fraction that is at the high end of what is found at subunit interfaces of proteins (from 9 to 40%) (Janin et al., 1988). The total accessible surface area of three isolated monomers is 9,600 Å^2. The accessible surface area of a trimer is 5,800 Å^2, leaving a total of 3,800 Å^2 of surface area buried upon trimer formation from the preformed α-helical peptides (~1,200 Å^2 buried per monomer). The value would be even larger if one were to account for the random coil to α-helix transition that accompanies trimerization (Betz et al., 1995). An additional 1,500 Å^2 per monomer is buried as the trimer forms the crystal lattice via stacking of trimers and contacts from the crystallographic twofold axes, thus burying a total of 87% of the accessible surface area of a single molecule.

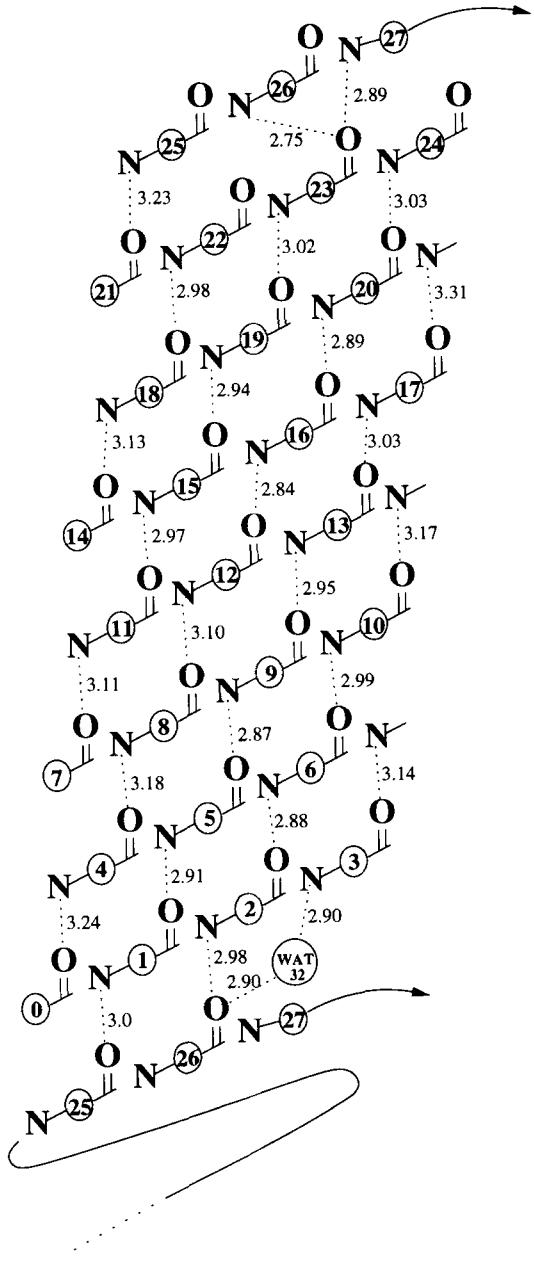

Fig. 2. Helical hydrogen bonding pattern in coil-V_aL_d. Included are all main-chain intrahelical hydrogen bonds and the two hydrogen bonds plus one water-mediated hydrogen bond that form at the interface of two stacked trimers.

Along with the large buried surface area, the molecules within the trimer form salt bridges between glutamate residues in the *e* positions and lysine residues in the *g* positions of adjacent monomers. Two salt bridges are observed between symmetry-related molecules: those between Glu 6-OE1 and Lys 8-NZ, and between Lys 15-NZ and Glu 20-OE1 (Fig. 4). The remaining unbridged *e* and *g* positions are Glu 13, which hydrogen bonds to Wat 41 and Wat 33, and Lys 22, whose NZ bonds to the sulfate anion. The remaining glutamates, Glu 1 and Glu 27, in *g* positions, are presumably hydrogen bonded to water molecules, although there are no ordered waters visible in the electron density.

α-Helix stacking

As shown in Figure 3D, the crystal is a stack of α-helices along its *c* direction. Between each two stacked α-helices, two of the three possible interhelical main-chain hydrogen bonds are formed and the third is mediated by a water molecule. Of all the polar main-chain atoms of the α-helix, only Glu 24-O is not involved in an α-helical hydrogen bond. Furthermore, separation of the hydrophobic layers between stacked trimers (5.59 Å between Leu 26-CA and Val 2-CA) is nearly the same as the separation distance between hydrophobic layers within one trimer (average distance between CAs is 5.45 A). Hence, there is no discontinuity of the hydrophobic core along the superhelix throughout the crystal. The solvent-accessible surface area buried upon stacking of two trimers is 1,000 $Å^2$ or 500 $Å^2$ per trimer, or ~170 $Å^2$ per monomer. Because each monomer is involved in two stacking interactions, approximately 11% of the surface of one monomer is buried upon helix stacking.

A stack of three trimers along the *c* axis of the crystal (Fig. 3D) forms one repeat of a superhelix. The pitch of the superhelix is the spacing of 84 residues (3 × 28 residues) spanning the length of three unit cells in the *c* direction. In other words, one trimer consisting of 26 α-helical residues, plus an acetyl group, and the space of approximately one residue between α-helices, forms one-third of the superhelical repeat. Thus, when stacked, each trimer makes a 120 degree rotation around the superhelical axis. The average rise per residue is 1.45 Å and the total repeat of the superhelix is 121.59 Å (exactly three times the length of the *c* axis of the unit cell).

Contacts between trimers related by a twofold axis

The area buried in the contact between trimers, related by a crystallographic twofold axis, is ~1,100 $Å^2$ per monomer. These contacts are mainly polar and include ionic interactions between Lys 15-NZ (*g*-position) and Glu 3-OE1 (*b*-position), between Glu 24-OE1 (*b*-position) and Lys 21-NZ (*f*-position), and between Gln 17-NE2 (*b*-position) and Ala 25-O (*c*-position). Because of the twofold axes at $z = 0$ and $z = 0.5$, these salt bridges are duplicated by symmetry. These six salt links bridge one monomer of a trimer to two monomers within two adjacent trimers.

Solvent molecules

All of the 18 visible solvent molecules are listed in Table 3, along with their protein–hydrogen bonding partners. Glu 20-OE1 (both *e*-positions) and Lys 22-NZ (*g*-position) hydrogen bond via waters 33, 35, and 37, respectively, to main-chain carbonyl oxygens to stabilize trimer formation. No water-mediated interactions between trimers across the twofold axis are observed.

Discussion

Design and structure

The antiparallel structure of coil-Ser (Lovejoy et al., 1993) led us to look at substituting hydrophobic residues at *a* and *d* positions of the heptad repeat of coiled coils and their effect on strand polarity. The new structure of coil-V_aL_d forms a parallel three-stranded coiled coil. This three-stranded structure is in agreement with predictions based on substitution of hydrophobic residues at *a* and *d* positions in peptides of GCN4 (Harbury et al., 1993). In the GCN4

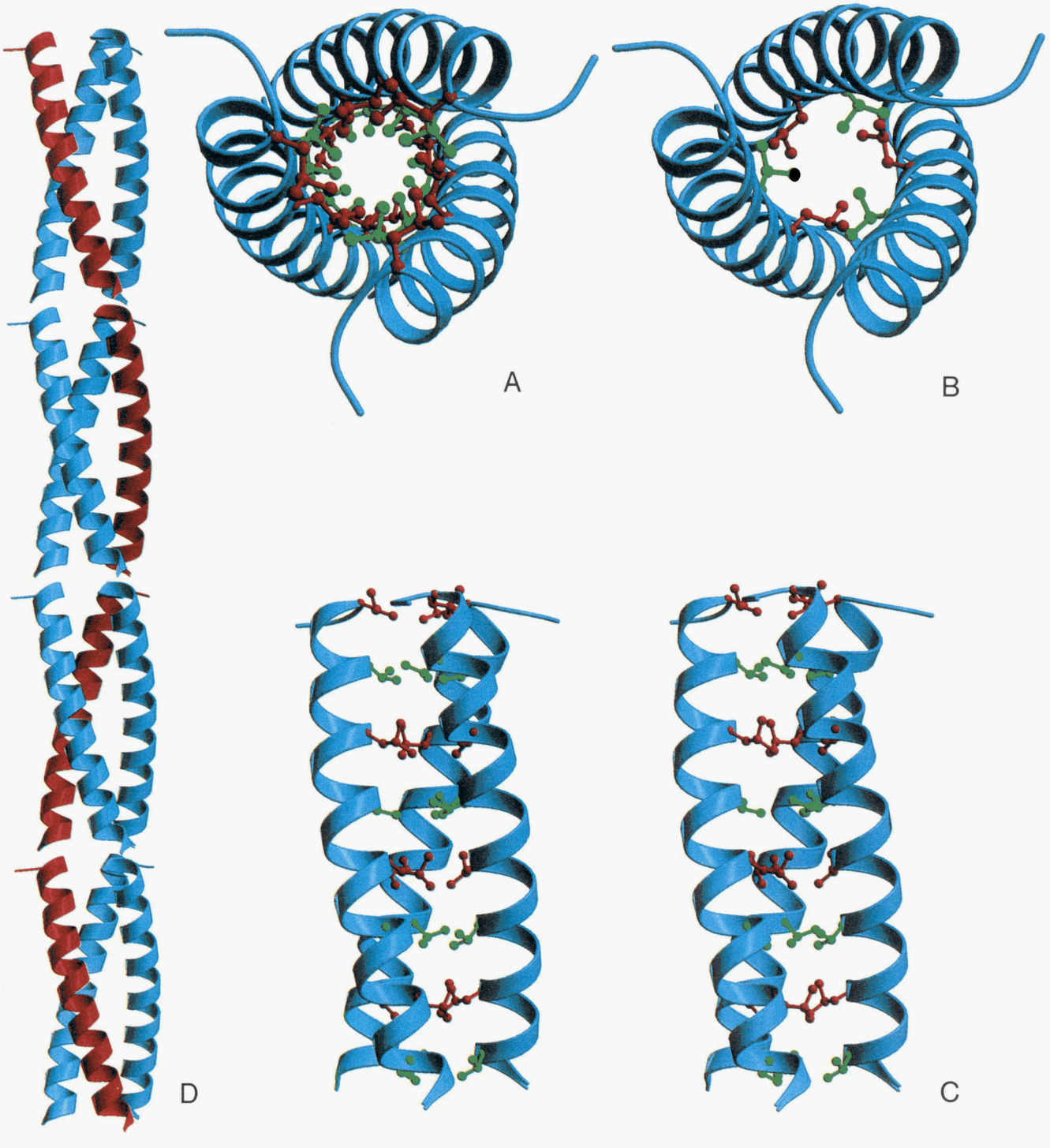

Fig. 3. Ribbon diagrams of the coil-V_aL_d trimer. Shown are the main-chain atoms represented as helical ribbons (blue) and the side-chain atoms of the valine residues (green) and the leucine residues (red) forming the hydrophobic core, shown in ball-and-stick representation. The figure was prepared using MOLSCRIPT (Kraulis, 1991) and rendered through Raster3D (Bacon & Anderson, 1988; Merritt & Murphy, 1994). **A:** View down the *c* axis showing the *a* layers (Val) and *d* layers (Leu) in ball-and-stick form. **B:** View down the *c* axis; only one *a* layer (Val 16) and one *d* layer (Leu 12) are shown in ball-and-stick. **C:** Stereo side view of coil-V_aL_d trimer. **D:** Ribbon diagram of four stacked trimers showing the pseudo-continuous left-handed superhelix. Note that three stacked trimers, 84 residue-equivalents in length, form one repeat of the superhelix with a pitch of 121.6 Å.

system, β-branched residues, such as valine, at *a* positions favor dimeric or trimeric coiled coils and disfavor formation of tetrameric coiled coils. With the formation of the trimer, the structure of coil-V_aL_d has all of the expected features of the hydrophobic core with all valines and leucines in their most favored side-chain conformations.

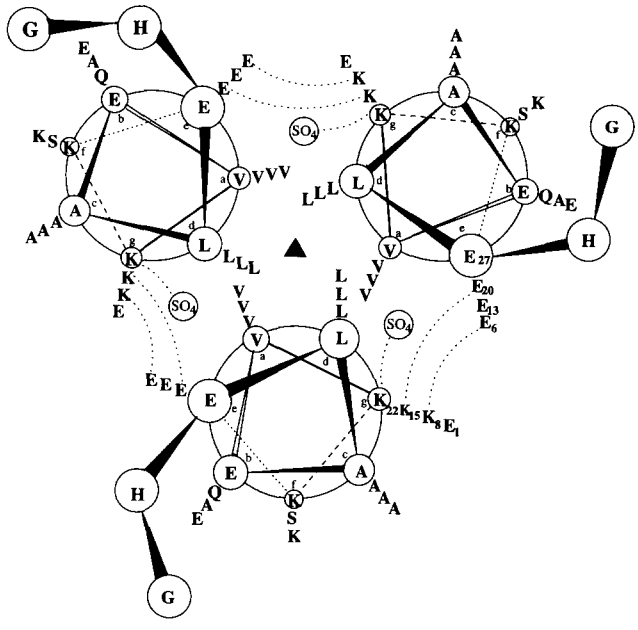

Fig. 4. Helical wheel diagram of a trimer of coil-V_aL_d with amino acid residues represented by one-letter code. Also shown are the intratrimer electrostatic interactions between residues in the *e* and *g* positions of the heptad repeat (dashed lines).

Deviation of the structure of coil-V_aL_d from that of an all α-helical coiled coil occurs only at the C terminus. The three C-terminal residues are nonhelical because they are excluded from the helix to permit the α-helices in the crystal to stack into a pseudo-continuous superhelix (Fig. 3D). The stacking allows the hydrophobic layers to form a single hydrophobic core running along the *c* axis of the crystal. Thus, maintaining an unperturbed hydrophobic core seems to be preferred over extending the α-helix throughout the entire length of the molecule.

In the design of the parallel trimer, the *e* and *g* positions within the same layer were designed to form salt bridges with charged residues in adjacent monomers; that is, glutamate in an *e* position was expected to form a salt bridge to lysine in position *g* within the same heptad of an adjacent molecule of the trimer. Two of three possible *e–g* salt bridges are formed in the crystal structure, but only the Glu 6–Lys 8 contact is a "designed" salt bridge as described above (Fig. 4). Lys 15, which is expected to form a salt bridge to Glu 13, instead interacts with Glu 20, possibly due to crystal packing; Lys 15-NZ moves away from Glu 13-OE1 slightly by forming a hydrogen bond to Glu 3-OE1 (*b*-position) on a molecule within a different trimer, related by a crystallographic twofold axis. Glu 13-OE1 is left to form a salt bridge with Wat 33 and Wat 41 in the absence of symmetry related polar side chains nearby.

Formation of a third *e–g* salt bridge between Glu 27 and Lys 22 would be possible were it not for a sulfate ion lying on the twofold axis near Lys 22, because Glu 27-OE1 is only 4.1 Å from Lys 22-NZ. It appears that the closeness of a symmetry-related molecule related by a twofold axis causes a competition for the salt bridge at Lys 15, resulting in a misregister of *e–g* salt bridges.

Comparison of coil-V_aL_d to coil-Ser

A comparison of coil-V_aL_d with its parent molecule, coil-Ser, shows a few minor differences in structure. The only sequence differences between the two nearly identical peptides is that coil-V_aL_d contains valines at the four *a* positions, replacing leucines and the tryptophan at position 2. Structurally, one coil-V_aL_d monomer is very similar to one monomer of coil-Ser: superimposing 146 atoms

Table 3. *Characteristics of the 18 solvent molecules within 3.5 Å of any protein atom in coil-V_aL_d*[a]

Water no.	*B* [Å^2]	Dist (Å); atom			
31	31.3	2.98; Ace 0-O	—	—	—
32	27.7	2.90; Glu 3-N	2.85; Wat 40-O	2.90; Leu 26-O	—
33	22.1	2.67; Lys 8-O	3.06; Wat 42-O	3.10; Glu 13-OE1	—
34	19.4	2.84; Glu 6-OE1 2.68; Glu 6-O	3.14; Wat 41-O	—	—
35	21.4	3.04; Lys 15-O	—	2.95; Glu 20-OE1	—
36	35.9	3.49; Glu 3-OE2 3.44; Lys 7-NZ	—	—	—
37	12.9	2.59; Glu 20-O	—	3.24; Lys 22-NZ	—
38	22.1	3.04; Ala 4-O	—	—	—
39	46.7	3.48; Lys 8-NZ	—	—	—
40	22.9	3.50; Glu 3-OE1	2.85; Wat 32-O	—	3.22; Wat 43
41	30.0	3.03; Glu 13-OE2	3.14; Wat 34-O	—	3.01; Wat 41
42	42.1	—	3.06; Wat 33-O	—	—
43	46.7	3.59; Ser 14-OG	—	—	3.22; Wat 40
44	31.1	2.87; Glu 20-OE2	—	—	—
45	35.0	2.61; Ala 18-O	—	—	—
46	44.7	3.31; Gln 17-O	—	—	—
47	60.7	—	—	—	—
48	46.4	3.23; Gln 17-OE1	—	—	—

[a]Column 1 gives the identifier; column 2 gives the *B*-factor; columns 3 and 4 give protein and solvent atoms within the same unit cell; columns 5 and 6 give protein and solvent atoms related by crystallographic symmetry.

containing the main chain and CBs of residues Ac-1–29 of the structure of coil-V_aL_d with those of each of the three distinct molecules of the coil-Ser yields RMSDs of 1.85 Å, 1.41 Å, and 1.51 Å for molecules A, B, and C, respectively. Because the largest differences are found at the C terminus where coil-V_aL_d is non-helical from residues 27 to 29, superposition of residues Ac-1–26 yields smaller RMSDs. Superposition of the main-chain and CB atoms (133 atoms) from residues Ac-1–26 of coil-V_aL_d with those in each molecule of coil-Ser yields RMSDs of 0.91 Å, 0.91 Å, and 0.99 Å for A, B, and C, respectively.

Despite minor changes in the sequence and crystallization conditions, major changes are observed in the molecular organization and the crystal packing of coil-V_aL_d relative to coil-Ser. Coil-V_aL_d crystallizes under slightly different conditions, yet grows in the trigonal space group P321, with one molecule per asymmetric unit, and the threefold axis of the trimer coincident with the crystallographic threefold axis along *c*. Coil-V_aL_d thus forms parallel three-helix bundles that are stacked C terminus to N terminus along the *c* axis of the crystal creating a pseudo-continuous superhelical structure. Coil-Ser, in contrast, crystallizes in the orthorhombic space group $P2_12_12$, with three molecules of the antiparallel trimer per asymmetric unit and no stacking of α-helices.

Designer crystals or fourth generation design

Based on the stacking of the α-helices that we observe in coil-V_aL_d, we propose a general scheme for designing coiled coils of various numbers of strands that all form continuous superhelices in a crystal, as does coil-V_aL_d. The ability of a peptide to stack may lead to crystals easily, thus short-circuiting the usual rate-limiting step in structure determination of such molecules. In analyzing the structure of coil-V_aL_d and the high-resolution structure of another designed peptide, α1 (G.G. Privé, N.L. Ogihara, L. Wesson, D.H. Anderson, D. Cascio, & D. Eisenberg, in prep.), which also forms stacked α-helices, efficient stacking appears to require the space of approximately one residue between helices, along the helical axis. For the following discussion, we introduce the term residue-equivalent. We define a residue-equivalent as either an amino acid residue or an N-terminal acetyl group, or the approximately one-residue spacer that is needed between two stacked helices.

For α-helical stacking to occur, the peptide chain must contain an integral number of heptad repeats. In the case of coil-V_aL_d, there are 29 residues, three of which unwind and extend into the solvent region of the crystal (in order to be accommodated in the structure and are not part of the superhelix), plus one acetyl group, plus one spacer: 29 − 3 + 1 + 1 = 28 residue-equivalents = 4 heptads. A stack of three such helices then produces one repeat of one strand of the superhelix as shown in Figure 3D. In other words, each helix contributes a 120 degree rotation around the superhelical axis, or each heptad contributes 30 degrees. The total repeat or pitch of the superhelix is then equal to 12 heptads or 84 residue-equivalents. Included in these are 3 acetyl groups, 3 × 26 amino acid residues, and 3 one-residue spacers, as explained above. Thus, in a designer crystal for a trimeric coiled coil, we expect a superhelix to form from an acetylated peptide of 26 residues (Table 4).

Assuming the repeat of the superhelix is the essential element of the crystal, we propose a design scheme for coiled coil peptides that crystallize as continuous stacked dimers, trimers, or tetramers, each containing integral numbers of heptad repeats. In a situation analogous to coil-V_aL_d, where one molecule of the trimer contributes 120 degrees of rotation around the superhelical axis, a dimeric coiled coil with six heptad repeats would be expected to contribute 180 degrees of rotation. Thus, an N-terminally acetylated peptide of 40 residues (corresponding to six heptad repeats or 42 residue-equivalents corresponding to Ac-40 in Table 4), forming a dimeric coiled coil, could be expected to crystallize as stacked superhelices, where two stacked dimers form one repeat of the superhelix and the dimer axis coincides with a crystallographic twofold axis.

In the case of a tetrameric coiled coil, a rotation of only 90 degrees (or 180 degrees) would be needed. A peptide length of three heptads (21 residue equivalents corresponding to Ac-19 in Table 4) would be expected to form a tetrameric coiled coil where four stacked tetramers form one repeat of the continuous superhelix within a crystal. Alternatively, six heptads repeats contributing 180 degrees of rotation around the superhelical axis (42 residue equivalents corresponding to Ac-40 in Table 4) could also be expected to form a tetrameric stack of continuous α-helices in a crystal. Although this tetrameric stack has the same number of residue equivalents as the dimeric stack, its heptad repeats have a different amino acid sequence (Table 4).

For the stacked trimer, as explained above, an ideal length for the peptide is four heptad repeats or 28 residue equivalents (corresponding to Ac-26 in Table 4). In the case of coil-V_aL_d, the three extra residues were able to extend into the solvent region without disturbing the stacking of α-helices significantly. Energetically and entropically, however, it might be advantageous to use a slightly shorter peptide consisting of the acetyl group and 26 residues with a C-terminal amide group. This simple scheme does not include the possibility that the pitch and the rise per residue of the superhelix can change between different oligomerization states. Further experiments will show if this design scheme holds or whether it needs refinement and extension.

Table 4. *General scheme for designer crystals, giving the number of heptad repeats, the peptide length, the number of residue equivalents, and the residues in the a and d positions of the heptad repeat (Woolfson & Alber, 1995)*

Oligomerization	# Heptads	Peptide length	Residue equivalents	*a* position	*d* position
Dimer	6	Ac-40	42	I	L
Trimer	4	Ac-26	28	V	L
Tetramer	3	Ac-19	21	L	I
	6	Ac-40	42	L	I

In conclusion, we have described the structure of a designed three-stranded coiled coil that forms a continuous superhelix along the c axis of the crystal. The stoichiometry of the peptide is in accord with the ideas of Woolfson and Alber (1995), in which the stoichiometry of coiled coils is determined by the nature of the hydrophobic residues in the a and d positions of the heptad repeat. Based on the structure of coil-V_aL_d, we propose a general scheme for designing peptides made of integral numbers of heptad repeats to crystallize as dimeric, trimeric, and tetrameric stacked coiled coils by forming continuous superhelices throughout the crystal.

Methods

Synthesis, purification, and crystallization

The peptide coil-V_aL_d was synthesized on a Milligen 9050 peptide synthesizer using Fmoc chemistry (Choma et al., 1994) and purified by reverse phase HPLC. Crystals were grown at room temperature by the hanging drop vapor diffusion method (McPherson, 1985). Lyophilized samples were dissolved in water to 5 mg/mL and mixed with an equal volume (4 μL) of reservoir solution containing 3.0 M ammonium sulfate, 0.1–3.5 mM NaOH, and 0.05 M KH_2PO_4, pH ~5.5. Small crystal rods grew within one week and subsequently were macroseeded by transferring single crystals into hanging drops containing 2.5 mg/mL protein mixed with an equal volume of the same reservoir solution. Crystals grew in space group P321 with unit cell parameters $a = b = 33.6$ Å and $c = 40.5$ Å. Based on a molecular weight of 3,213, and one molecule per asymmetric unit, the Matthews parameter (Matthews, 1968) is 2.06 Å^3/Da, with a solvent content of 40%.

Data collection and processing

Diffraction data were collected at room temperature using a RAXIS IIC imaging plate detector system with a Rigaku RU-200B generator and double focusing mirrors operated at 50 kV and 100 mA. The crystal to detector distance was 100 mm, and crystals were rotated around the spindle axis with 3 degree oscillation images collected to a resolution of 2.1 Å. Three sweeps of 96, 111, and 75 degrees were collected and processed using DENZO (Minor, 1993; Otwinowski, 1993) yielding 35,980 independent observations. These were reduced to a 99.7% complete data set of 1,710 unique reflections in the resolution range ∞–2.1 Å with an R_{sym} on intensities of 9.1%. In the resolution range from 2.18 to 2.1 Å, the signal to noise ratio, as expressed in $I/\sigma(I)$ is approximately 13. Intensities were then converted to structure factors using the method of French and Wilson (1978) as implemented in the program TRUNCATE.

Structure solution and refinement

The structure of coil-V_aL_d was solved by molecular replacement using the XPLOR program package (Brünger, 1988). An ideal α-helical model of N-acetyl-$(Ala)_{28}$-Gly was built and oriented with the α-helical axis parallel to the crystallographic c axis in a unit cell with dimensions $a = b = 50$ Å, $c = 100$ Å, $\alpha = \beta = \gamma = 90$ degrees. The temperature factors of all 148 nonhydrogen atoms of the idealized α-helical model were set to 18 Å^2, the value determined from a Wilson plot (Wilson, 1949) of the native data set. The rotation function was calculated using data from 15 to 3.5 Å, with an inner and outer radius of integration of 3.5–20 Å. All of the 42 peaks found from the 10,000 grid points that had the highest values in the rotation function were on θ_2 sections of 14–18 degrees, indicating that this was probably the tilt angle of the α-helix with respect to the c axis. In the geometry used, θ_1 denotes a rotation around the α-helical axis, θ_2 denotes a rotation of the α-helical axis away from the c axis, and θ_3 denotes a rotation around the oriented α-helical axis. The peak that led to the solution of the structure had Eulerian angles of (314.0, 18.0, and 20.0 degrees) and was sixth from the top at 1.80 σ above the mean, with the top peak being 1.81 σ above the mean.

Patterson correlation refinement was performed subsequently on the highest peaks from the rotation function in the resolution range 10–3.0 Å. Ten cycles of rigid body refinement with the whole chain, followed by ten cycles with the chain broken into two pieces (between residues 15 and 16), followed by ten cycles with the chain broken into six pieces (every fifth residue) were carried out. Again, no distinction between the correct and any incorrect solutions was found. Therefore, all 42 peaks determined from the rotation function were subjected to Patterson correlation refinement, and then used for calculating a translation function. For computing the translation function, we also used all data from 10 to 3.0 Å. Again, when comparing all 42 translation function peaks, it was not possible to distinguish between the correct and any incorrect solutions. However, orientation number 6 from the rotation function, which produced the highest translation function peak at 3.97 σ above the mean, finally led to the solution of the structure. The second highest peak in the translation function of orientation number 6 was at 3.71 σ. The top peaks of all 42 translation functions were then subjected to 200 cycles of least-squares refinement in XPLOR using, as an indicator, the free R-factor based on 10% of the reflections in the resolution range. In only four cases did the free R-factor remain below "random" values (i.e., below 55%), and among those four, one led to the solution of the structure.

After translating the model to (0.394, 0.212, 0.225) in fractional coordinates, the R-factor was 56.6% (R_{free} = 48.4%). The R_{free} value could be lower because of the small size of the TEST set, which contained only 57 reflections. Twenty cycles of rigid-body refinement of this model, using data from 10–3 Å and 200 cycles of least-squares refinement reduced the R-factor to 43.2% (R_{free} = 52.1%). Model-phased electron density maps, displayed using FRODO (Jones, 1982), clearly showed side-chain density for Val 2, Leu 5, Ser 14, Val 16, and Val 23, and allowed unambiguous assignment of the sequence to the model. As a result, we found that the acetyl group of this model was lying at the position of residue 1; that is, the rotation function solution was off by approximately 100 degrees. However, because of the helical symmetry, the overlap with the correct solution was sufficient and the model phases good enough to reveal the correct orientation and position of the structure.

In the next nine rounds of refinement (Table 2), the complete sequence was built into the electron density and 18 ordered water molecules as well as one sulfate ion were identified. In each round, R_{free} was monitored to avoid overfitting. The last round of refinement was repeated with all data included in the working set.

Quality of model

The quality of the model was verified using RMSDs of the geometry with respect to the parameters derived by Engh and Huber (1991) in the program XPLOR (Brünger, 1988) and the program PROCHECK (Laskowski et al., 1993). The PDB ID code for the coordinates of coil V_aL_d is 1COI.

Acknowledgments

This research was supported by the National Institutes of Health Chemistry/Biology Interface Predoctoral Training at UCLA (N.L.O) and by NSF grant MCB 94-20769 (D.E.).

References

Bacon DJ, Anderson WF. 1988. A fast algorithm for rendering space-filling molecule pictures. *J Mol Graphics 6*:219–220.

Banner DW, Kokkinidis M, Tsernoglou D. 1987. Structure of ColE1 Rop protein at 1.7 Å resolution. *J Mol Biol 196*:657–675.

Betz S, Fairman R, O'Neil K, Lear J, DeGrado W. 1995. Design of two-stranded and three-stranded coiled-coil peptides. *Phil Trans R Soc London B 348*:81–88.

Boice J, Dieckmann GR, DeGrado WF, Fairman R. 1996. Thermodynamic analysis of a designed three-stranded coiled coil. *Biochemistry 35*:14480–14485.

Brünger AT. 1988. Crystallographic refinement by simulated annealing. In: Isaacs NW, Taylor MR, eds. *Crystallographic computing 4: Techniques and new technologies*. Oxford: Clarenden Press. pp 126–140.

Brünger AT. 1993. Assessment of phase accuracy by cross validation—the free R-value—methods and applications. *Acta Crystallogr D 49*:24–36.

Choma CT, Lear JD, Nelson MJ, Dutton PL, Robertson DE, DeGrado WF. 1994. Design of a heme-binding four-helix bundle. *J Am Chem Soc 116*:856–865.

Cohen C, Parry DAD. 1986. α-Helical coiled coils—A widespread motif in proteins. *Trends Biochem Sci 11*:245-248.

Cohen C, Parry DAD. 1990. α-Helical coiled coils and bundles: How to design an α-helical protein. *Proteins Struct Funct Genet 7*:1–15.

Dunbrack RL Jr, Karplus M. 1994. Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains. *Nature Structural Biology 1*:334–339.

Engh RA, Huber R. 1991. Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr A 47*:392–400.

Fairman R, Chao HG, Mueller L, Lavoie TB, Shen L, Novotny J, Matsueda GR. 1995. Characterization of a new four-chain coiled coil: Influence of chain length on stability. *Protein Sci 4*:1457–1469.

Fass D, Harrison SC, Kim PS. 1996. Retrovirus envelope domain at 1.7 Å resolution. *Nature Structural Biology 3*:465–469.

French S, Wilson K. 1978. On the treatment of negative intensity observations. *Acta Crystallogr A 34*:517–525.

Fujinaga M, Berthet-Colominas C, Yaremchuk AD, Tukalo MA, Cusack S. 1993. Refined crystal structure of the seryl-tRNA synthetase from *Thermus thermophilus* at 2.5 Å. *J Mol Biol 234*:222–233.

Glover JN, Harrison SC. 1995. Crystal structure of the heterodimeric bZIP transcription factor c-fos-c-jun bound to DNA. *Nature 373*:257–261.

Harbury PB, Zhang T, Kim PS, Alber T. 1993. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science 262*:1401–1406.

Hodges RS, Sodek J, Smillie LB, Jurasek L. 1972. Tropomyosin: Amino acid sequence and coiled-coil structure. *Cold Spring Harbor Symp Quant Biol 37*:299–310.

Janin JJ, Miller S, Chothia C. 1988. Surface subunit interfaces and interior of oligomeric proteins. *J Mol Biol 204*:155–164.

Jones TA. 1982. In: FRODO: A graphics fitting program for macromolecules. Sayre D, ed. *Computational crystallography*. Oxford: Oxford University Press. pp 303–317.

Kraulis PJ. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr 24*:946–950.

Laskowski RA, MacArthur MW, Moss DS, Thornton JM. 1993. PROCHECK—A program to check the stereochemical quality of protein structures. *J Appl Crystallogr 26*:283–291.

Lau SYM, Taneja AK, Hodges RS. 1984. Synthesis of a model protein of defined secondary and quaternary structure. *J Biol Chem 259*:13253-13261.

Lewis M, Chang G, Horton NC, Kercher MA, Pace HC, Schumacher MA, Brennan RG, Lu P. 1996. Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science 271*:1247–1254.

Lovejoy B, Choe S, Cascio D, McRorie DK, DeGrado WF, Eisenberg D. 1993. Crystal structure of a synthetic triple-stranded α-helical bundle. *Science 259*:1288–1293.

Lumb KL, Kim PS. 1995. A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. *Biochemistry 34*:8642–8648.

Lupas A. 1996. Coiled coils: New structures and new functions. *TIBS 21*:375–382.

Luzzati V. 1952. Traitement statistique des erreurs dans la determination des structures cristallines. *Acta Crystallogr 5*:802–810.

Malashkevich VN, Kammerer RA, Efimov VP, Schulthess T, Engel J. 1996. The crystal structure of a five-stranded coiled coil in COMP: A prototype ion channel? *Science 274*:761–765.

Marmorstein R, Carey M, Harrison SC. 1992. DNA recognition by GAL4: Structure of a protein–DNA complex. *Nature 356*:408–411.

Matthews BW. 1968. Solvent content of protein crystals. *J Mol Biol 33*:491–497.

McLachlan AD, Stewart M. 1975. Tropomyosin coiled-coil interactions: Evidence for an unstaggered structure. *J Mol Biol 98*:293–304.

McPherson A. 1985. Crystallization of macromolecules: General principles. *Methods Enzymol 114*:112-120.

Merritt EA, Murphy MEP. 1994. Raster3D version 2.0—A program for photorealistic molecular graphics. *Acta Crystallogr D 50*:869–873.

Minor W. 1993. *XDISPLAYF program*. West Lafayette, Indiana: Purdue University.

Monera OD, Zhou NE, Kay CM, Hodges RS. 1993. Comparison of antiparallel and parallel two-stranded α-helical coiled-coils. *J Biol Chem 268*:19218–19227.

O'Neil KT, DeGrado WF. 1990. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science 250*:646–651.

O'Shea EK, Klemm JD, Kim PS, Alber T. 1991. X-ray structure of the GCN4 leucine zipper, a 2-stranded, parallel coiled coil. *Science 254*:539–544.

Otwinowski Z. 1993. Oscillation data reduction program. In: Sawyer L, Isaacs N, Bailey S, eds. *Proceedings of the CCP4 study weekend: Data collection and processing. 29–30 Jan 1993*. Warrington, UK: SERC Daresbury Laboratory. pp 56–62.

Ponder JW, Richards FM. 1987. Tertiary templates for proteins—Use of packing criteria in the enumeration of allowed sequences for different structural classes. *J Mol Biol 193*:775–791.

Ramachandran GN, Ramakrishnan C, Sasisekharan V. 1963. Stereochemistry of polypeptide chain configurations *J Mol Biol 7*:95–99.

Weis WI, Drickamer K. 1994. Trimeric structure of a C-type mannose-binding protein. *Structure 2*:1227–1240.

Wendt H, Berger C, Baici A, Thomas RM, Bosshard HR. 1995. Kinetics of folding of leucine zipper domains. *Biochemistry 34*:4097–4107.

Wilson AJC. 1949. The probability distribution of X-ray intensities. *Acta Crystallogr 2*:318–321.

Wilson IA, Skehel JJ, Wiley DC. 1981. Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 Å resolution. *Nature 289*:366–373.

Woolfson DN, Alber T. 1995. Predicting oligomerization states of coiled coils. *Protein Sci 4*:1596–1607.

Yan Y, Winograd E, Viel A, Cronin T, Harrison SC, Branton D. 1994. Crystal structure of the repetitive segments of spectrin. *Science 262*:2027–2030.