Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Structure *14*, 247–255, February 2006 ©2006 Elsevier Ltd All rights reserved DOI 10.1016/j.str.2005.10.010

Antiparallel Four-Stranded Coiled Coil Specified by a 3-3-1 Hydrophobic Heptad Repeat

Yiqun Deng,[1] Jie Liu,[1] Qi Zheng,[1] David Eliezer,[1] Neville R. Kallenbach,[2] and Min Lu[1,*]
[1] Department of Biochemistry
Weill Medical College of Cornell University
New York, New York 10021
[2] Department of Chemistry
New York University
New York, New York 10003

Summary

Coiled-coil sequences in proteins commonly share a seven-amino acid repeat with nonpolar side chains at the first (*a*) and fourth (*d*) positions. We investigate here the role of a 3-3-1 hydrophobic repeat containing nonpolar amino acids at the *a*, *d*, and *g* positions in determining the structures of coiled coils using mutants of the GCN4 leucine zipper dimerization domain. When three charged residues at the *g* positions in the parental sequence are replaced by nonpolar alanine or valine side chains, stable four-helix structures result. The X-ray crystal structures of the tetramers reveal antiparallel, four-stranded coiled coils in which the *a*, *d*, and *g* side chains interlock in a combination of knobs-into-knobs and knobs-into-holes packing. Interfacial interactions in a coiled coil can therefore be prescribed by hydrophobic-polar patterns beyond the canonical 3-4 heptad repeat. The results suggest that the conserved, charged residues at the *g* positions in the GCN4 leucine zipper can impart a negative design element to disfavor thermodynamically more stable, antiparallel tetramers.

Introduction

The coiled coil is among the most ubiquitous structural motifs in proteins consisting of interwound α helices that pack in parallel or antiparallel orientations, with a distinct preference for the parallel orientation (Gruber and Lupas, 2003; Kohn et al., 1997). Coiled-coil proteins contain a short-range seven-residue (heptad) sequence repeat, $(a\text{-}b\text{-}c\text{-}d\text{-}e\text{-}f\text{-}g)_n$, with a predominance of the aliphatic side chains Leu, Ile, Val, and Ala at the *a* and *d* positions, and polar residues generally elsewhere (Hodges et al., 1972; Lupas et al., 1991; McLachlan and Stewart, 1975; Parry, 1982). This characteristic 3-4 hydrophobic polar residue pattern is the basis of Crick's canonical "knobs-into-holes" core packing between α helices and represents an essential structural feature of the coiled-coil family (Crick, 1953; Hodges et al., 1972; Lupas et al., 1991; McLachlan and Stewart, 1975; Parry, 1982). Residues at the *e* and *g* positions are frequently charged, and have been shown to influence the number and orientation of chains in a supercoil through interhelical electrostatic interactions (Kohn et al., 1995, 1998; Krylov et al., 1994; O'Shea et al., 1992, 1993). Substantial progress has been achieved toward understanding the main factors that specify and stabilize parallel coiled-coil conformations (Bryson et al., 1995; Gonzalez et al., 1996; Harbury et al., 1993; Ogihara et al., 1997).

The determinants of antiparallel coiled-coil structure are less well understood (Kohn et al., 1997; Monera et al., 1996; Oakley and Hollenbeck, 2001; Oakley and Kim, 1998). DeGrado and coworkers have characterized a series of peptides designed to form four-stranded, antiparallel coiled coils (Bryson et al., 1995; North et al., 2001). One member of this class, α_1B, contains leucine residues at the *a*, *d*, and *g* positions and forms a helical tetramer in solution (Ho and DeGrado, 1987). In 1993, Muller-Hill and coworkers presented genetic evidence that the C-terminal heptad repeat region of the lac repressor can self-associate into an antiparallel four-helix bundle (Alberti et al., 1993). They emphasized that uncharged amino acids at the *e* and *g* positions are required to specify this tetrameric structure (Alberti et al., 1993). Fairman and coworkers subsequently showed that a 21-residue peptide corresponding to the lac repressor heptad repeat region (called Lac21) forms a stably folded, antiparallel, α-helical tetramer in solution (Fairman et al., 1995). The crystal structure of the lac repressor confirmed that residues at the *a*, *d*, and *e* positions of this tetrameric coiled coil participate in interhelical hydrophobic interactions (Friedman et al., 1995; Lewis et al., 1996). The Richardsons first identified conserved features of a tightly packed, antiparallel helical bundle structure, the Alacoil (Gernert et al., 1995), in which the backbones of two neighboring helices are offset by half a turn with alanine in position *g* (ferritin-type Alacoils) or by a full turn with alanine at position *e* (rop-type Alacoils). Experimental studies of the Lac21 peptide model of a ferritin-type Alacoil show that the presence of neutral amino acids at the *e* and *g* positions indeed contributes to folding and stability (Fairman et al., 1995, 1996; Solan et al., 2002). In connection with this, we note that introducing alanine into the *e* position of the GCN4 dimeric leucine zipper creates tetramers (Krylov et al., 1994).

Recent crystallographic studies of the tetrameric domain of the SARS coronavirus S2 protein identify an antiparallel, four-stranded coiled coil in which leucine or isoleucine residues at the *d* positions comprise the hydrophobic core, while bulky *a* and *g* side chains pack against the outer surface of the core (Y.D., J.L., Q.Z., and M.L., unpublished data). In light of the results discussed above, this interaction involving the positions *a*, *d*, and *g* of the heptad repeat raises the possibility that a distinct hydrophobic residue pattern (a 3-3-1 hydrophobic repeat) might be sufficient to specify a superhelical fold in which four α helices align in an antiparallel relative orientation. To test this hypothesis, we have engineered mutations of the *g* side chains in the GCN4 leucine zipper model which has served as an invaluable test bed for defining determinants of coiled-coil structure (Harbury et al., 1993, 1998). We simultaneously changed three charged residues at the *g* positions in the classical GCN4-pR sequence to nonpolar Ala and

*Correspondence: mlu@med.cornell.edu

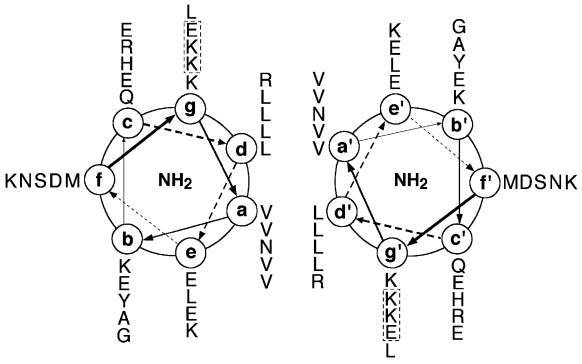


Figure 1. Helical Wheel Representation of the GCN4-pR Leucine Zipper as a Parallel, Two-Stranded Coiled Coil

The view is from the N terminus looking down the superhelical axis. Heptad repeat positions are labeled *a*–*g*. Prime (′) refers to positions from the second helix. In the mutant peptides described here, three charged residues in the dashed box at position *g* were changed to alanine and valine. The sequences (with the three mutated *g* positions underlined) are the following: GCN4-pR, MK VKQLEDK VEELLSK NYHLENE VARLKKL VGER; GCN4-pA, MK VKQLEDA VEELLSA NYHLENA VARLKKL VGER; GCN4-pV, MK VKQLEDV VEELLSV NYHLENV VARLKKL VGER. The GCN4-pR sequence contains an additional Met-Lys-Val and no Arg-Met at its N terminus, but is otherwise identical to GCN4-p1.

Val side chains. In contrast to the parent dimeric structure with the helices parallel (O'Shea et al., 1991), both mutant peptides form discrete, stable, α-helical tetramers in solution. Their X-ray crystal structures reveal antiparallel, four-stranded coiled coils with combined knobs-into-knobs and knobs-into-holes packing interactions between the hydrophobic *a*, *d*, and *g* side chains. Thus, the 3-3-1 hydrophobic pattern in the heptad repeat represents an extension of the Alacoil motif in nature for mediating protein-protein interactions via coiled coils. Our results suggest that the structural selectivity of conserved, charged residues at the *e* and *g* positions of leucine zipper dimers results not only from favorable interhelical electrostatic interactions, but also from the necessary destabilization (negative design) of alternate coiled-coil conformations.

Results and Discussion

Position *g* Mutants of the GCN4 Leucine Zipper

The extensively investigated GCN4 leucine zipper has provided a key model system for analysis of the hydrophobic heptad repeat pattern and structural selectivity in coiled coils (Harbury et al., 1993; O'Shea et al., 1991). The apolar interface of the parent double-stranded coiled coil is formed by interhelical association in a parallel orientation between residues at the *a* and *d* positions (O'Shea et al., 1991). Electrostatic interactions between oppositely charged amino acid residues at the *e* and *g* positions of different helices also can contribute to dimerization specificity (Kohn et al., 1998; O'Shea et al., 1992, 1993). The recombinant leucine zipper peptide GCN4-pR includes four *g* positions: Lys9, Lys16, Glu23, and Leu30 (Figure 1). To investigate the effect of a 3-3-1 hydrophobic residue repeat on helix association and interface specificity, we simultaneously mutated Lys9, Lys16, and Glu23 of GCN4-pR to the small aliphatic amino acids alanine and valine. The resulting mutants were designated GCN4-pA and GCN4-pV, respectively. These leucine zipper peptides were expressed in *Escherichia coli* and purified by reverse-phase HPLC.

Tetramer Formation in Solution

The circular dichroism spectra of GCN4-pA and GCN4-pV show the characteristic signature of an α helix with minima at 222 and 208 nm (Figure 2A). On the basis of the mean residue ellipticity at 222 nm at 4°C and

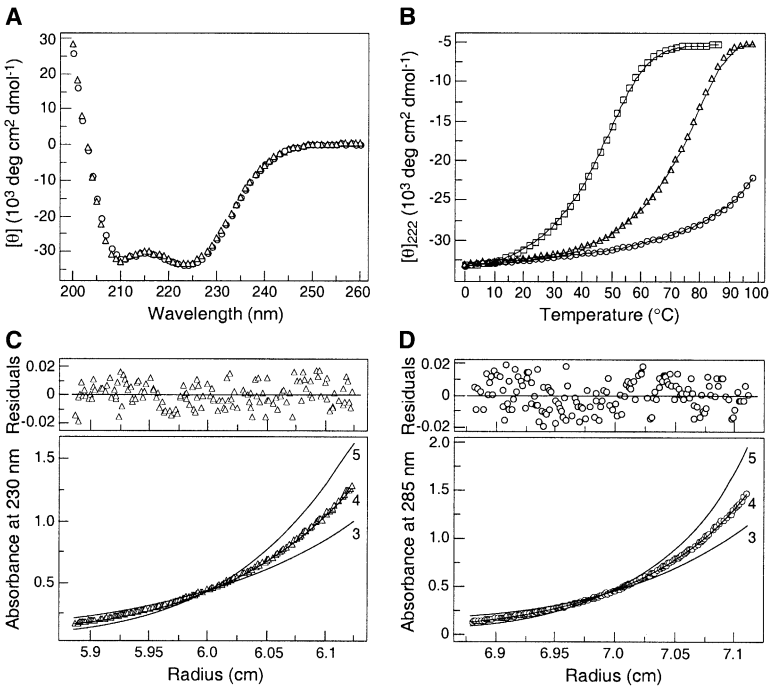


Figure 2. The GCN4-pA and GCN4-pV Peptides Form Four-Stranded Helical Bundles

GCN4-pA, triangles; GCN4-pV, circles.

(A) Circular dichroism (CD) spectra of 50 μM peptide at 4°C in PBS (pH 7.0).

(B) Thermal melts monitored by CD at 222 nm. The squares show data for the parent GCN4-pR peptide.

(C) Sedimentation equilibrium data for a 50 μM sample of GCN4-pA at 20°C and 28,000 rpm in PBS (pH 7.0). The data fit closely to a tetramer bundle. Curves expected for trimeric and pentameric models are indicated for comparison. The deviation in the data from the linear fit for a tetrameric model is plotted (upper).

(D) Sedimentation equilibrium data for a 500 μM sample of GCN4-pV at 20°C and 28,000 rpm in PBS (pH 7.0).

Table 1. Summary of Crystallographic Analysis

Diffraction Data							
Data Set	λ (Å)	Resolution (Å)	Number of Reflections (total/unique)	Completeness (%)	I/σ(I)	R_{sym} (%)[a]	Phasing Power (ano/iso)
GCN4-pA							
Native	0.9788	36.5–1.5	118,795 (20,939)[b]	98.7 (100)	18.6 (6.1)	4.2 (27.6)	
SeMet λ1	0.9793	50–1.6	92,273 (7,639)	98.8 (100)	21.4 (11.0)	4.1 (23.6)	1.5/0.7
SeMet λ2	0.9790	50–1.6	88,380 (7,462)	96.9 (99.7)	24.2 (15.2)	4.4 (17.2)	2.3/0.4
SeMet λ3	0.9675	50–1.7	76,239 (6,355)	99.1 (100)	14.9 (7.4)	4.5 (44.4)	1.9/–
GCN4-pV							
	0.9788	34.7–2.0	18,105 (2,063)	98.3 (100)	16.5 (6.5)	4.3 (39.2)	

Refinements	GCN4-pA	GCN4-pV
Resolution (Å)	36.5–1.5	34.7–2.0
Number of reflections	18,807	1,771
Number of protein atoms	980	237
Number of water molecules	117	25
R_{cryst}/R_{free} (%)[c]	23.3/28.3	22.4/27.6
Rmsd bond lengths (Å)	0.011	0.042
Rmsd bond angles (°)	1.3	3.1
Average B factor (Å^2)	23.7	54.2
Rmsd B values (Å^2)	1.8	4.7

[a] $R_{sym} = \Sigma|I - \langle I\rangle|/\Sigma I$, where I is the integrated intensity of a given reflection.
[b] Numbers in parentheses represent the statistics for the shell comprising the outer 10% (theoretical) of the data.
[c] $R_{cryst} = \Sigma|F_{obs} - F_{calc}|/\Sigma F_{obs}$, $R_{free} = R_{cryst}$ calculated using 10% of the reflection data chosen randomly and omitted from the start of refinement.

50 μM peptide concentration in neutral pH phosphate-buffered saline (PBS), each peptide contains >90% helical structure. Under these conditions, GCN4-pA and GCN4-pV undergo cooperative and reversible thermal unfolding transitions with midpoints (T_ms) of 74°C and 95°C, respectively, as compared to a T_m of 47°C for the parent molecule GCN4-pR at the same concentration (Figure 2B). Sedimentation equilibrium measurements indicate that GCN4-pA and GCN4-pV sediment as tetramers and exhibit no systematic dependence of molecular weight on concentration between 50 and 500 μM (Figure 2C). Thus, both the hydrophobic *g* position mutants of the dimeric leucine zipper peptide form well-ordered and extremely stable quadruple-helical structures.

Crystal Structures of GCN4-pA and GCN4-pV

To investigate the side chain packing at the *a*, *d*, and *g* positions in the heptad repeat in atomic detail, we determined the X-ray crystal structure of GCN4-pA at 1.5 Å resolution by the method of multiwavelength anomalous diffraction (Hendrickson, 1991) using a SeMet derivative (Table 1). As hypothesized, GCN4-pA adopts a left-handed superhelix consisting of four antiparallel α-helical peptide monomers in register that cross at an angle of approximately 20° (Figure 3). The superhelix forms a cylinder with overall diameter of ~24 Å and length of ~48 Å. Diagonally related helices have identical relative orientation. The individual helices in the tetramer can be superimposed on each other with an rms deviation for α-carbon atoms of 0.30–0.51 Å and with the largest deviations occurring at their ends. The distance between the axes of parallel helices (on diagonal) is ~14 Å, whereas that between the axes of adjacent antiparallel helices is ~8 Å, an Alacoil characteristic. However, the leucine residues at the *d* positions interact between parallel helices and point into the core of the tetramer (see Discussion). Cross-sectional layers containing leucine alternate between 2-fold symmetrical pairs of parallel helices. Fourteen of the 16 leucine side chains assume χ_1 and χ_2 dihedral angles near −65°, 175° or −177°, 65°, corresponding to their most favored rotamers in α helices (Lovell et al., 2000; Ponder and Richards, 1987). Residues at positions *a* and *g* in the neighboring antiparallel helices pack against the leucines at *d* to complete the hydrophobic core.

The crystal structure of GCN4-pV at 2.0 Å resolution (Table 1) also reveals a symmetric, left-handed, antiparallel four-α helix bundle, in this case ~28 Å wide and ~46 Å long (Figure 4). An exact dyad is perpendicular to the superhelical axis. The leucine side chains at the *d* positions interact between parallel helices (on diagonal) and stagger axially to form the hydrophobic core of the tetramer. All the dihedral angles χ_1 and χ_2 of the leucine residues are approximately −65°, 175° for the N-terminal two *d* layers and −177°, 65° for the C-terminal two *d* layers in their well-populated rotamer conformations. The equally spaced *a* and *g* residues along the neighboring antiparallel helices flank the hydrophobic interface, efficiently sequestering the leucine side chains from solvent. The differential bulk of valine relative to alanine shows that interfacial interactions involving apolar residues at the *a*, *d*, and *g* positions of the heptad repeat play a major role in stabilizing the antiparallel tetramer structure. The stability differences between GCN4-pA and GCN4-pV are seen clearly in the solution data (Figure 2B).

Although the superhelix pitch, radius of curvature, and residues per supercoil turn in the parallel GCN4-pIL (with isoleucine at each *d* position and leucine at each *a*) (Harbury et al., 1993) and antiparallel GCN4-pA and GCN4-pV coiled coils differ significantly (Table 2), the fraction of solvent-accessible surface area buried in the tetramer remains about the same. For example, approximately 5470 Å^2 of accessible surface area is buried in the GCN4-pA and GCN4-pV tetramers, compared with

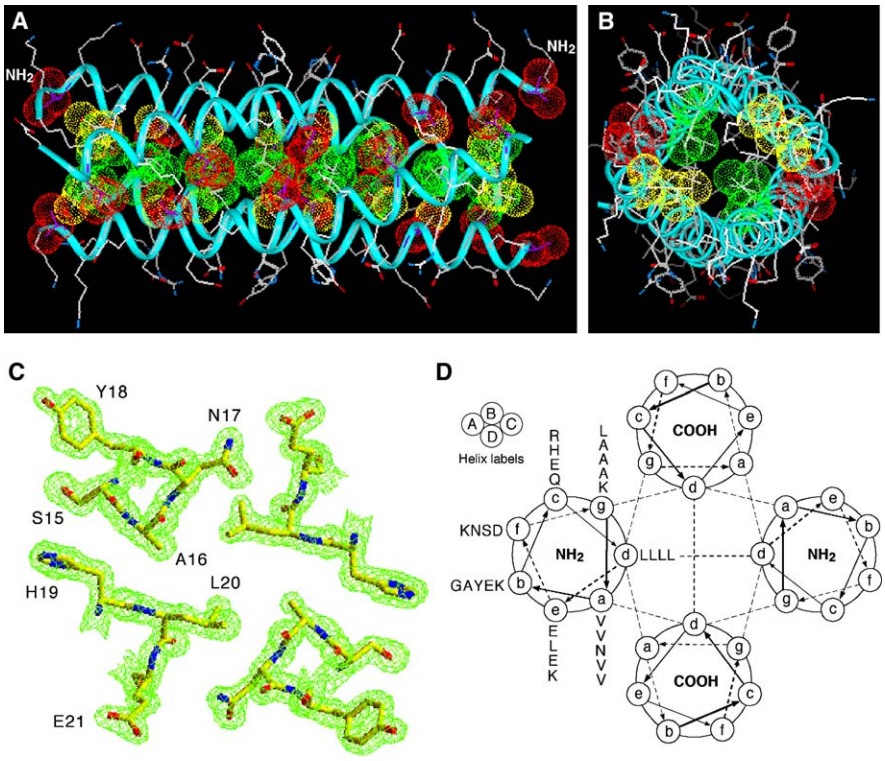


Figure 3. Crystal Structure of the GCN4-pA Tetramer

(A) Lateral view of the tetramer. Red van der Waals surfaces identify residues at the *a* positions, green surfaces identify residues at the *d* positions, and yellow surfaces identify residues at the *g* positions. The N termini of helices A and B are indicated.

(B) Axial view of the tetramer. The green, yellow, and red van der Waals surfaces of the L6 (*d*), L30 (*g*), and V31 (*a*) side chains are depicted.

(C) Cross-section of the superhelix in the L20 (*d*) layer. The 1.5 Å $2F_o - F_c$ electron density map (contoured at 1.5σ) is shown with the refined molecular model.

(D) Helical wheel projection of residues 2–32 of the GCN4-pA tetramer. Heptad repeat positions are labeled *a*–*g*. The leucines at the *d* positions form the apolar interface of the tetramer.

~5930 Å^2 of buried surface area in the GCN4-pIL tetramer (Harbury et al., 1993). Relative to the side chains of isolated helices, the leucine side chains at the *d* positions of GCN4-pA and GCN4-pV are completely buried; residues at the *a* and *g* positions are substantially buried; those at the *e* position are partly buried; whereas the *b* position remains completely exposed (see below). Due to the smaller size of the alanine side chain, GCN4-pA has a compressed *g*-*d*′ interface, and therefore the resulting structure is less symmetric than that of GCN4-pV. More surface area at the *c* and *f* positions is buried in GCN4-pA than GCN4-pV (58% versus 35% at position *c* and 20% versus 1% at position *f*).

Knobs-into-Knobs and Knobs-into-Holes Packing Interactions

The tetramer interfaces of GCN4-pA and GCN4-pV show nonclassical packing of side chains in the core of interacting helices. The leucine side chains at the *d* positions directly face each other between parallel helices (on diagonal), giving rise to a so-called knobs-into-knobs packing in these layers (Gottschalk, 2005; Lupas and Gruber, 2005). By contrast, adjacent antiparallel helices adopt a new kind of knobs-into-holes interaction in which each leucine knob at the *d* position of one helix packs into a hole formed by the *a* and *g* residues of one neighboring helix and by two *d* residues in adjacent layers along the superhelical axis (Figure 5). This geometry results in a similar placement of atoms around the side chains at positions *a* and *g*. Knobs formed by *a* residues of one helix fit into holes formed by the spaces between the *d* and *e* residues on the neighboring helix and by two *a* residues in adjacent layers along the superhelical axis. Similarly, knobs at *g* positions pack into holes formed by the *c* and *d* residues of the neighboring helix and by two *g* residues in adjacent layers. Thus, the *a*, *d*, and *g* residues of the heptad repeat segregate into four geometrically distinct helix-helix interfaces. Consequently, the GCN4-pA and GCN4-pV superhelices have alternating rectangular cross-sections.

Comparison with Other Antiparallel Tetramers

We begin by considering the classification of antiparallel helices due to the Richardsons (Gernert et al., 1995). They distinguished two fundamental types of their "Alacoils": the ferritin type, with the helices offset by 0.25 heptad, and the rop type (Banner et al., 1987), with the helices shifted by 0.5 heptad (see Table 2). Our GCN4-pA and GCN4-pV structures apparently conform to the rop type. On the other hand, the lac repressor tetramerization domain (Kercher et al., 1997) and the WSPLB 21–52 model (Slovic et al., 2005) are of the ferritin type. There are substantial differences in the interior packing of the tetramer among each of these structures (Figure 6).

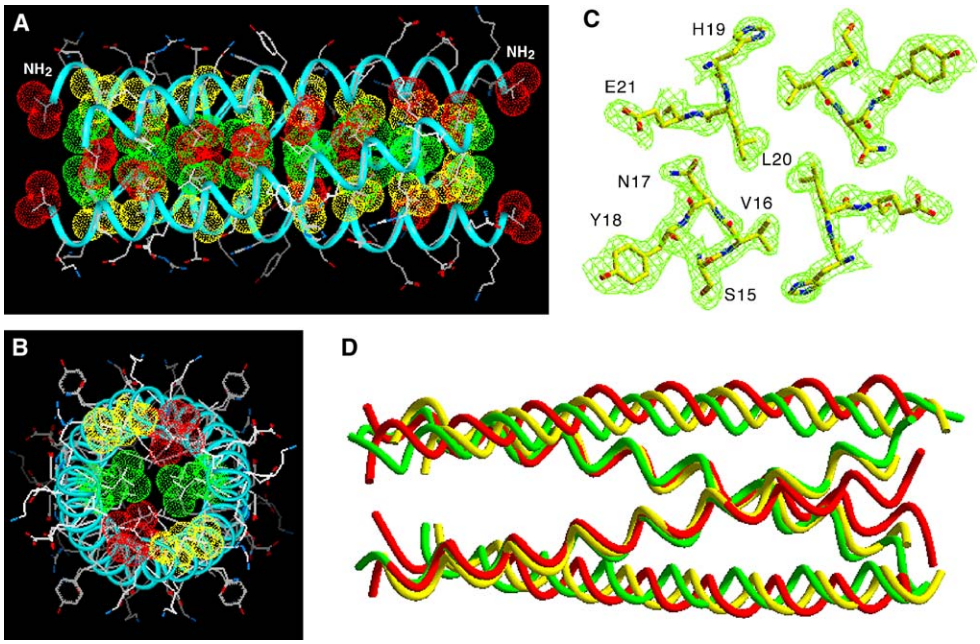


Figure 4. Crystal Structure of the GCN4-pV Tetramer

(A) Lateral view of the tetramer. Red van der Waals surfaces identify residues at the *a* positions, green surfaces identify residues at the *d* positions, and yellow surfaces identify residues at the *g* positions. The N termini of helices A and B are indicated.

(B) Axial view of the tetramer. The green, yellow, and red van der Waals surfaces of the L6 (*d*), L30 (*g*), and V31 (*a*) side chains are depicted.

(C) Cross-section of the superhelix in the L20 (*d*) layer. The $2F_o - F_c$ electron density map at 1.2σ contour is shown with the refined molecular model.

(D) Superposition of the backbone conformations of the parallel GCN4-pIL tetramer (red) and the antiparallel GCN4-pA (green) and GCN4-pV (yellow) tetramers.

Each type of Alacoil can exhibit at least two modes of core packing arrangements: the well-known *a*,*d* pattern (WSPLB and Rop), and a triad *a*,*d*,*e* (the lac repressor tetramerization domain) or *a*,*d*,*g* as in GCN4-pA and GCN4-pV. The packing properties in each of these cases are clearly revealed by calculating the extent of buried surface areas of the side chains at these heptad positions (Figure 6D). We see that each packing mode has a signature in surface areas. For the *a*,*d* packing, for example, both *a* and *d* side chains show nearly complete burial at these sites. Conversely, the *a*,*d*,*e* and *a*,*d*,*g* patterns show lower burial values at either of the *a* and *d* sites. Thus, surface burial provides a useful diagnostic for each core packing interaction.

3-3-1 Hydrophobic Residue Repeat

Our results show that heptad sequence repeats with a 3-3-1 hydrophobic residue pattern can encode a stable, antiparallel four-helix structure. In this tetramer conformation, the *a*, *d*, and *g* residues form a left-handed apolar strip along the surface of the right-handed α helices, while their side chains mesh between interacting helices by the winding of antiparallel chains around one another. This motif adds an additional element to the current repertoire of coiled-coil structures: the hydrophobic core is formed by interlocking of residues at positions *a* and *d* in rop (Banner et al., 1987) but at positions *a*, *d*, and *g* in both GCN4-pA and GCN4-pV. The 3-3-1 hydrophobic residue repeat retains two structural features of coiled

Table 2. Helix-Helix Interactions in Antiparallel, Four-Stranded Coiled Coils

	WSPLB	Lac	Rop	GCN4-pA	GCN4-pV	GCN4-pIL[a]
α helix parameter						
α helix radius, r_1 (Å)	2.28	2.28	2.30	2.29	2.29	2.25
Angular frequency, ω_1 (°/residue)	−100.1	−100.3	−99.7	−99.7	−99.3	−100.8
Rise per residue, h (Å)	1.51	1.52	1.51	1.51	1.51	1.53
Residues per α-helical turn, n	3.60	3.59	3.61	3.61	3.63	3.58
Superhelix parameter						
Supercoil radius, r_0 (Å)	8.0	6.8	7.4	6.9	7.4	7.1
Supercoil pitch, P (Å)	252	213	187	161	156	198
Superhelix phase, ϕ_0 (°)	10.1	71.9	82.6	93.5	89.0	–
Translation of the supercoiled helix along the superhelical axis, Z_{trans} (Å)	−3.3	1.81	−2.60	−0.46	−1.31	–
Residues per superhelix turn, n	170.5	143.0	127.9	110.9	107.3	133.1

[a] GCN4-pIL forms a parallel coiled-coil tetramer (Harbury et al., 1993).

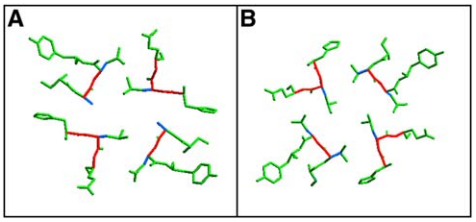


Figure 5. Knobs-into-Holes Packing in the Antiparallel GCN4-pA and GCN4-pV Tetramers

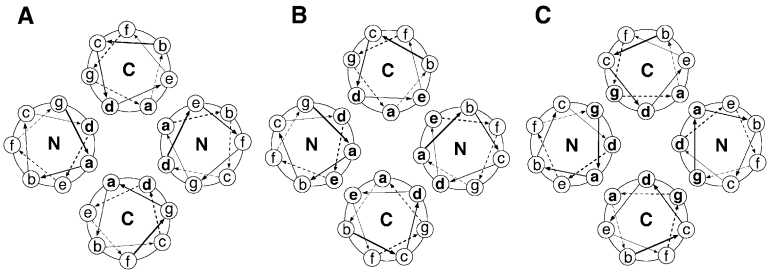
Helix cross-sectional layers centered on the *a*, *c*, *d*, *e*, and *g* positions of GCN4-pA (A) and GCN4-pV (B) are depicted. Knobs formed by the side chains of one helix fit into holes formed by the spaces between side chains on a neighboring helix. The C_α-C_β bond of each knob (thick blue line) and the C_α-C_α vector at the base of the recipient hole on the neighboring helix (thick red line) are indicated.

coils predicted by Crick from helical net diagrams (Crick, 1953): an interhelix packing angle near 20° and knobs-into-holes packing of side chains between adjacent helices. Our results support the notion that the stoichiometry and geometry of coiled coils can be determined by distinctive hydrophobic patterns within the classical heptad repeat (Lupas and Gruber, 2005).

Structural Selectivity

One major challenge in protein design is to engineer conformational specificity in addition to thermodynamic stability (Bryson et al., 1995). Buried polar interactions of appropriate geometry have been shown to impart specificity (structural uniqueness) in protein folding and design even if they do so at the expense of stability (Hendsch et al., 1996). For example, a buried hydrogen bond between Asn residues in the GCN4 leucine zipper is responsible for maintaining the register of the two parallel helices in the dimer structure (Harbury et al., 1993). It is tempting to apply a similar argument to explain why the 3-3-1 hydrophobic repeat is not typically found in coiled-coil proteins. Crick pointed out in 1953 that coiled coils contain fewer hydrophobic side chains than globular proteins (Crick, 1953). Here we find that increasing the hydrophobic side chain content of the GCN4 leucine zipper can lead to enhanced stabilization of the resulting structure. The highly stable antiparallel, symmetric 3-3-1 tetramer might present a functionally compromised conformational trap, as we argue to be the case in the SARS S2 protein (Y.D., J.L., Q.Z., and M.L., unpublished data). Consistent with this notion, there is growing evidence that polymorphic interactions between two heptad repeat regions in viral envelope proteins provide a reservoir of free energy that can be released to drive structural transitions in a precise temporal sequence that is required for membrane fusion (Y.D., J.L., Q.Z., and M.L., unpublished data).

Specificity in coiled coils (in terms of oligomeric state and strand orientation) stems in part from the precise complementarity of hydrophobic residues at the *a* and *d* positions (Bryson et al., 1995; Harbury et al., 1993), and to a larger extent from buried polar interactions (hydrogen bonds and ion pairs) that mildly destabilize the native structure yet strongly destabilize alternate conformations (Hendsch et al., 1996). While interactions between charged side chains at the *e* and *g* positions have previously been implicated in determination of helix orientation, number, and heterogeneity, their precise role has been the subject of some controversy (Lavigne et al., 1996; Lumb and Kim, 1995). Recent studies have revealed that electrostatic effects are complex, including both stabilizing and destabilizing interactions (Garcia-Mayoral et al., 2003; Marti and Bosshard, 2003, 2004). Changes in the free energy of the unfolded state may account for the effects of electrostatic interactions rather than differences in the stability of the native state per se. In all cases, however, the net free energy contributions of a pair of opposite charges at the *e* and *g* positions are small (~1 kcal/mol or less), although there can be a stronger repulsive interaction between similar charged side chains at these positions (Jelesarov and Bosshard, 1996; Kohn et al., 1995; Lavigne et al., 1996; O'Shea et al., 1993). The finding that replacing *g* residues by nonpolar side chains in a dimeric coiled coil



D

	Buried surface area (%)				
	a-**d** core		**a**-**d**-**e** core	**a**-**d**-**g** core	
Positions	WSPLP	Rop	Lac repressor	GCN4–pA	GCN4–pV
a	93	100	100	80	89
b	8	21	28	0	0
c	6	15	0	58	35
d	97	98	76	100	94
e	71	76	84	26	28
f	0	1	11	20	1
g	51	68	28	85	90

Figure 6. Core Packing in Antiparallel, Four-Stranded Coiled Coils

(A–C) Helical wheel representation of the antiparallel tetramer showing the hydrophobic interface formed by side chains from positions *a*,*d* (A), *a*,*d*,*e* (B), and *a*,*d*,*g* (C). Heptad repeat positions are labeled *a*–*g*.

(D) Buried surface areas in WSPLB 21–52 (*a*,*d* packing), the lac repressor tetramerization domain (*a*,*d*,*e* packing), rop (*a*,*d* packing), GCN4-pA (*a*,*d*,*g* packing), and GCN4-pV (*a*,*d*,*g* packing). Percent buried surface area is expressed as the fraction of accessible side chain surface area in the isolated helix that becomes buried in the antiparallel tetramer.

can switch both stoichiometry and helix orientation to produce a tetramer structure of enhanced stability suggests that the selection of charged or polar side chains at the *g* positions might provide an additional mechanism to regulate structural features of interfacial interaction specificity. For example, to restrict a given interaction exclusively to a parallel coiled coil might necessitate extensive charges at the *e* and *g* positions in order to avoid the potential trap represented by the antiparallel 3-3-1 tetramer. Indeed, a significant fraction of these positions in a protein such as myosin contains charged or polar residues (Li et al., 2003). It should be noted that while the 3-4 heptad sequence regularity has led to the successful development of prediction algorithms to identify potential coiled-coil domains from primary amino acid sequence (Lupas, 1997; Wolf et al., 1997), the 3-3-1 hydrophobic pattern is poorly accommodated in these algorithms. Accurate prediction of diverse coiled-coil interaction motifs will require an expanded understanding of the principles governing hydrophobic heptad repeat patterns.

Experimental Procedures

Peptide Expression and Purification

The peptides GCN4-pA and GCN4-pV were expressed in *E. coli* BL21(DE3)/pLysS using a modified pET3a vector. Cells were lysed by glacial acetic acid and centrifuged to separate the soluble fraction from inclusion bodies. The soluble fraction was subsequently dialyzed into 5% acetic acid overnight at 4ºC. Peptides from the soluble fraction were purified to homogeneity by reverse-phase HPLC on a C18 preparative column using a water-acetonitrile gradient in the presence of 0.1% trifluoroacetic acid and lyophilized. The identity of each peptide was confirmed by mass spectrometry (PerSeptive Biosystems Voyager Elite, Cambridge, MA); all molecular masses were found to be within 1 dalton of the expected mass. Selenomethionine (SeMet)-substituted GCN4-pA was produced for multiwavelength anomalous diffraction analysis in amino acid-supplemented minimal media and purified as described above (Doublie, 1997).

Biophysical Analysis

Circular dichroism (CD) spectra were measured on an Aviv 62A/DS CD spectrometer (Lakewood, NJ) at 4ºC in PBS (50 mM sodium phosphate/150 mM NaCl [pH 7.0]) and 50 μM peptide. Thermal stability was assessed by monitoring $[\theta]_{222}$ as a function of temperature under the same conditions. A $[\theta]_{222}$ value of –33,000 deg cm^2 dmol^{-1} was taken to correspond to 100% helix (Chen et al., 1974). Values of T_m were estimated by evaluating the maximum of the first derivative of $[\theta]_{222}$ versus temperature data (Cantor and Schimmel, 1980). Peptide concentrations were determined by absorbance at 280 nm in 6 M guanidine hydrochloride, using an extinction coefficient of 1280 cm^{-1} M^{-1} for tyrosine (Edelhoch, 1967). Sedimentation equilibrium experiments were performed on a Beckman XL-A analytical ultracentrifuge (Fullerton, CA) at 20ºC as described (Shu et al., 1999). Peptide solutions were dialyzed overnight against PBS (pH 7.0), loaded at initial concentrations of 50, 150, and 500 μM, and analyzed at rotor speeds of 25,000 and 28,000 rpm. Data sets were fitted to a single-species model. Random residuals were observed in all cases.

Crystallization

GCN4-pA was crystallized at room temperature using the hanging drop vapor diffusion method against 0.1 M Tris-HCl (pH 8.2), 16% ethanol. Crystals belong to space group $P2_12_12_1$ (a = 47.66 Å, b = 47.70 Å, c = 56.52 Å) and contain one tetramer in the asymmetric unit. The crystals were transferred into cryosolution containing 0.1 M Tris-HCl (pH 8.2), 22% ethanol, 15% glycerol and frozen in liquid nitrogen. Crystals of GCN4-pA/SeMet were obtained from 2 M $(NH_4)_2SO_4$, 5% isopropanol. Crystals belong to space group $I4_122$ (a = b = 46.38 Å, c = 55.82 Å) and contain one monomer in the asymmetric unit. The crystals were harvested in the reservoir buffer supplemented with 15% glycerol and frozen in liquid nitrogen. GCN4-pV was crystallized from 0.1 M sodium acetate (pH 4.6), 0.3 M $CdCl_2$, 12% PEG 400. Crystals belong to space group $I4_122$ (a = b = 45.98 Å, c = 52.99 Å) and contain one monomer in the asymmetric unit. The crystals were transferred to 15% glycerol in the reservoir buffer and frozen in liquid nitrogen.

Data Collection and Structure Determination

All data sets were collected on beamline X4A at the National Synchrotron Light Source. Data for all crystals were indexed and reduced with DENZO/SCALEPACK (Otwinowski and Minor, 1997). The CCP4 set of programs was used for truncating and scaling the data sets (CCP4, 1994). One selenium site (corresponding to the GCN4-pA M14 position) in the asymmetric unit was located with Solve (Terwilliger and Berendzen, 1999), and approximately 80% of the polypeptide chain was traced automatically into the electron density maps. Iterative rounds of model building with O (Jones et al., 1991), refinement with Refmac (Murshudov et al., 1997), and addition of ordered solvent clarified the trace except for a few residues at the helix termini. The structures of GCN4-pA and GCN4-pV were solved by molecular replacement with Phaser (Storoni et al., 2004) using the GCN4-pA/SeMet structure as a search model. All main chain torsional angles fall within the helical regions of the Ramachandran plot. Figures were generated using SETOR (Evans, 1993) and Insight II (Accelrys, San Diego, CA).

Structure Analysis

Superhelix phase (ϕ_0) and translation of the supercoiled helix along the superhelical axis (Z_{trans}) were obtained using an analytical equation for an antiparallel coiled coil (North et al., 2001; Slovic et al., 2005), whereas other coiled-coil parameters were calculated with TWISTER (Strelkov and Burkhard, 2002). Residues 26–51 of WSPLB 21–52 (PDB code: 1YOD), residues 341–355 of lac repressor (PDB code: 1LBI), residues 7–23 of rop (PDB code: 1ROP), residues 3–31 of GCN4-pA and GCN4-pV, and residues 2–30 of GCN4-pIL (PDB code: 1GCL) were used in the calculations. Buried surface areas were calculated from the difference of the accessible side chain surface areas of the tetramer structure and of the individual helical monomers using CNS 1.0 (Brunger et al., 1998). Residues 340–357 of helices A and B of lac repressor (PDB code: 1LBI), residues 31–51 of the A and B chains of WSPLB 21–52 (PDB code: 1YOD), and residues A6–A28 and A33–A54 of rop (PDB code: 1ROP) were used in the calculations. The three most N-terminal residues and the three most C-terminal residues of GCN4-pA and GCN4-pV were omitted from the calculation to minimize end effects.

Acknowledgments

We thank John Schwanof at the National Synchrotron Light Source for support at beamline X4A, and Benjamin North and William DeGrado for help with the superhelical parameter calculations. This work was supported by NIH grant AI511151 and by the Irma T. Hirschl Trust.

Received: July 13, 2005
Revised: September 30, 2005
Accepted: October 12, 2005
Published: February 10, 2006

References

Alberti, S., Oehler, S., von Wilcken-Bergmann, B., and Muller-Hill, B. (1993). Genetic analysis of the leucine heptad repeats of Lac repressor: evidence for a 4-helical bundle. EMBO J. *12*, 3227–3236.

Banner, D.W., Kokkinidis, M., and Tsernoglou, D. (1987). Structure of the ColE1 rop protein at 1.7 Å resolution. J. Mol. Biol. *196*, 657–675.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Bryson, J.W., Betz, S.F., Lu, H.S., Suich, D.J., Zhou, H.X., O'Neil, K.T., and DeGrado, W.F. (1995). Protein design: a hierarchic approach. Science *270*, 935–941.

Cantor, C., and Schimmel, P. (1980). Biophysical Chemistry, Volume III (New York: W.H. Freeman).

CCP4 (Collaborative Computational Project, Number 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Chen, Y.H., Yang, J.T., and Chau, K.H. (1974). Determination of the helix and β form of proteins in aqueous solution by circular dichroism. Biochemistry *13*, 3350–3359.

Crick, F.H.C. (1953). The packing of α-helices: simple coiled-coils. Acta Crystallogr. *6*, 689–697.

Doublie, S. (1997). Preparation of selenomethionyl proteins for phase determination. Methods Enzymol. *276*, 523–530.

Edelhoch, H. (1967). Spectroscopic determination of tryptophan and tyrosine in proteins. Biochemistry *6*, 1948–1954.

Evans, S.V. (1993). SETOR: hardware-lighted three-dimensional solid model representations of macromolecules. J. Mol. Graph. *11*, 134–138, 127–128.

Fairman, R., Chao, H.G., Mueller, L., Lavoie, T.B., Shen, L., Novotny, J., and Matsueda, G.R. (1995). Characterization of a new four-chain coiled-coil: influence of chain length on stability. Protein Sci. *4*, 1457–1469.

Fairman, R., Chao, H.G., Lavoie, T.B., Villafranca, J.J., Matsueda, G.R., and Novotny, J. (1996). Design of heterotetrameric coiled coils: evidence for increased stabilization by Glu(−)-Lys(+) ion pair interactions. Biochemistry *35*, 2824–2829.

Friedman, A.M., Fischmann, T.O., and Steitz, T.A. (1995). Crystal structure of lac repressor core tetramer and its implications for DNA looping. Science *268*, 1721–1727.

Garcia-Mayoral, M.F., Perez-Canadillas, J.M., Santoro, J., Ibarra-Molero, B., Sanchez-Ruiz, J.M., Lacadena, J., Martinez del Pozo, A., Gavilanes, J.G., Rico, M., and Bruix, M. (2003). Dissecting structural and electrostatic interactions of charged groups in α-sarcin. An NMR study of some mutants involving the catalytic residues. Biochemistry *42*, 13122–13133.

Gernert, K.M., Surles, M.C., Labean, T.H., Richardson, J.S., and Richardson, D.C. (1995). The Alacoil: a very tight, antiparallel coiled-coil of helices. Protein Sci. *4*, 2252–2260.

Gonzalez, L., Jr., Brown, R.A., Richardson, D., and Alber, T. (1996). Crystal structures of a single coiled-coil peptide in two oligomeric states reveal the basis for structural polymorphism. Nat. Struct. Biol. *3*, 1002–1009.

Gottschalk, K.E. (2005). A coiled-coil structure of the αIIbβ3 integrin transmembrane and cytoplasmic domains in its resting state. Structure *13*, 703–712.

Gruber, M., and Lupas, A.N. (2003). Historical review: another 50th anniversary—new periodicities in coiled coils. Trends Biochem. Sci. *28*, 679–685.

Harbury, P.B., Zhang, T., Kim, P.S., and Alber, T. (1993). A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. Science *262*, 1401–1407.

Harbury, P.B., Plecs, J.J., Tidor, B., Alber, T., and Kim, P.S. (1998). High-resolution protein design with backbone freedom. Science *282*, 1462–1467.

Hendrickson, W.A. (1991). Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. Science *254*, 51–58.

Hendsch, Z.S., Jonsson, T., Sauer, R.T., and Tidor, B. (1996). Protein stabilization by removal of unsatisfied polar groups: computational approaches and experimental tests. Biochemistry *35*, 7621–7625.

Ho, S.P., and DeGrado, W.F. (1987). Design of a 4-helix bundle protein: synthesis of peptides which self-associate into a helical protein. J. Am. Chem. Soc. *109*, 6751–6758.

Hodges, R.S., Sodek, J., Smillie, L.B., and Jurasek, L. (1972). Tropomyosin: amino acid sequence and coiled-coil structure. Cold Spring Harb. Symp. Quant. Biol. *37*, 299–310.

Jelesarov, I., and Bosshard, H.R. (1996). Thermodynamic characterization of the coupled folding and association of heterodimeric coiled coils (leucine zippers). J. Mol. Biol. *263*, 344–358.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Kercher, M.A., Lu, P., and Lewis, M. (1997). Lac repressor-operator complex. Curr. Opin. Struct. Biol. *7*, 76–85.

Kohn, W.D., Monera, O.D., Kay, C.M., and Hodges, R.S. (1995). The effects of interhelical electrostatic repulsions between glutamic acid residues in controlling the dimerization and stability of two-stranded α-helical coiled-coils. J. Biol. Chem. *270*, 25495–25506.

Kohn, W.D., Mant, C.T., and Hodges, R.S. (1997). α-helical protein assembly motifs. J. Biol. Chem. *272*, 2583–2586.

Kohn, W.D., Kay, C.M., and Hodges, R.S. (1998). Orientation, positional, additivity, and oligomerization-state effects of interhelical ion pairs in α-helical coiled-coils. J. Mol. Biol. *283*, 993–1012.

Krylov, D., Mikhailenko, I., and Vinson, C. (1994). A thermodynamic scale for leucine zipper stability and dimerization specificity: e and g interhelical interactions. EMBO J. *13*, 2849–2861.

Lavigne, P., Sonnichsen, F.D., Kay, C.M., and Hodges, R.S. (1996). Interhelical salt bridges, coiled-coil stability, and specificity of dimerization. Science *271*, 1136–1138.

Lewis, M., Chang, G., Horton, N.C., Kercher, M.A., Pace, H.C., Schumacher, M.A., Brennan, R.G., and Lu, P. (1996). Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. Science *271*, 1247–1254.

Li, Y., Brown, J.H., Reshetnikova, L., Blazsek, A., Farkas, L., Nyitray, L., and Cohen, C. (2003). Visualization of an unstable coiled coil from the scallop myosin rod. Nature *424*, 341–345.

Lovell, S.C., Word, J.M., Richardson, J.S., and Richardson, D.C. (2000). The penultimate rotamer library. Proteins *40*, 389–408.

Lumb, K.J., and Kim, P.S. (1995). Measurement of interhelical electrostatic interactions in the GCN4 leucine zipper. Science *268*, 436–439.

Lupas, A. (1997). Predicting coiled-coil regions in proteins. Curr. Opin. Struct. Biol. *7*, 388–393.

Lupas, A.N., and Gruber, M. (2005). The structure of α-helical coiled coils. Adv. Protein Chem. *70*, 37–78.

Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. Science *252*, 1162–1164.

Marti, D.N., and Bosshard, H.R. (2003). Electrostatic interactions in leucine zippers: thermodynamic analysis of the contributions of Glu and His residues and the effect of mutating salt bridges. J. Mol. Biol. *330*, 621–637.

Marti, D.N., and Bosshard, H.R. (2004). Inverse electrostatic effect: electrostatic repulsion in the unfolded state stabilizes a leucine zipper. Biochemistry *43*, 12436–12447.

McLachlan, A.D., and Stewart, M. (1975). Tropomyosin coiled-coil interactions: evidence for an unstaggered structure. J. Mol. Biol. *98*, 293–304.

Monera, O.D., Zhou, N.E., Lavigne, P., Kay, C.M., and Hodges, R.S. (1996). Formation of parallel and antiparallel coiled-coils controlled by the relative positions of alanine residues in the hydrophobic core. J. Biol. Chem. *271*, 3995–4001.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likehood method. Acta Crystallogr. D *53*, 240–255.

North, B., Summa, C.M., Ghirlanda, G., and DeGrado, W.F. (2001). D(n)-symmetrical tertiary templates for the design of tubular proteins. J. Mol. Biol. *311*, 1081–1090.

Oakley, M.G., and Hollenbeck, J.J. (2001). The design of antiparallel coiled coils. Curr. Opin. Struct. Biol. *11*, 450–457.

Oakley, M.G., and Kim, P.S. (1998). A buried polar interaction can direct the relative orientation of helices in a coiled coil. Biochemistry *37*, 12603–12610.

Ogihara, N.L., Weiss, M.S., DeGrado, W.F., and Eisenberg, D. (1997). The crystal structure of the designed trimeric coiled coil

coil-VaLd: implications for engineering crystals and supramolecular assemblies. Protein Sci. *6*, 80–88.

O'Shea, E.K., Klemm, J.D., Kim, P.S., and Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science *254*, 539–544.

O'Shea, E.K., Rutkowski, R., and Kim, P.S. (1992). Mechanism of specificity in the Fos-Jun oncoprotein heterodimer. Cell *68*, 699–708.

O'Shea, E.K., Lumb, K.J., and Kim, P.S. (1993). Peptide 'Velcro': design of a heterodimeric coiled coil. Curr. Biol. *3*, 658–667.

Otwinowski, Z., and Minor, W. (1997). Processing X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Parry, D.A. (1982). Coiled-coils in α-helix-containing proteins: analysis of the residue types within the heptad repeat and the use of these data in the prediction of coiled-coils in other proteins. Biosci. Rep. *2*, 1017–1024.

Ponder, J.W., and Richards, F.M. (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.

Shu, W., Ji, H., and Lu, M. (1999). Trimerization specificity in HIV-1 gp41: analysis with a GCN4 leucine zipper model. Biochemistry *38*, 5378–5385.

Slovic, A.M., Stayrook, S.E., North, B., and DeGrado, W.F. (2005). X-ray structure of a water-soluble analog of the membrane protein phospholamban: sequence determinants defining the topology of tetrameric and pentameric coiled coils. J. Mol. Biol. *348*, 777–787.

Solan, A., Ratia, K., and Fairman, R. (2002). Exploring the role of alanine in the structure of the Lac repressor tetramerization domain, a ferritin-like Alacoil. J. Mol. Biol. *317*, 601–612.

Storoni, L.C., McCoy, A.J., and Read, R.J. (2004). Likelihood-enhanced fast rotation functions. Acta Crystallogr. D Biol. Crystallogr. *60*, 432–438.

Strelkov, S.V., and Burkhard, P. (2002). Analysis of α-helical coiled coils with the program TWISTER reveals a structural mechanism for stutter compensation. J. Struct. Biol. *137*, 54–64.

Terwilliger, T.C., and Berendzen, J. (1999). Automated MAD and MIR structure solution. Acta Crystallogr. D Biol. Crystallogr. *55*, 849–861.

Wolf, E., Kim, P.S., and Berger, B. (1997). MultiCoil: a program for predicting two- and three-stranded coiled coils. Protein Sci. *6*, 1179–1189.

Accession Numbers

Coordinates and structure factors have been deposited in the Protein Data Bank under ID codes 2B1F (GCN4-pA) and 2B22 (GCN4-pV).