A Switch Between Two-, Three-, and Four-Stranded Coiled Coils in GCN4 Leucine Zipper Mutants

Pehr B. Harbury, Tao Zhang,* Peter S. Kim, Tom Alber*

Coiled-coil sequences in proteins consist of heptad repeats containing two characteristic hydrophobic positions. The role of these buried hydrophobic residues in determining the structures of coiled coils was investigated by studying mutants of the GCN4 leucine zipper. When sets of buried residues were altered, two-, three-, and four-helix structures were formed. The x-ray crystal structure of the tetramer revealed a parallel, four-stranded coiled coil. In the tetramer conformation, the local packing geometry of the two hydrophobic positions in the heptad repeat is reversed relative to that in the dimer. These studies demonstrate that conserved, buried residues in the GCN4 leucine zipper direct dimer formation. In contrast to proposals that the pattern of hydrophobic and polar amino acids in a protein sequence is sufficient to determine three-dimensional structure, the shapes of buried side chains in coiled coils are essential determinants of the global fold.

Recent evidence has suggested that the three-dimensional structure of a protein is determined largely by the pattern of hydrophobic (H) and polar (P) residues in the amino acid sequence and is independent of the geometric properties of the amino acid side chains that make up the pattern (*1–3*). This simplifying hypothesis, however, fails to account for a group of proteins composed of interacting, amphipathic α helices, the coiled-coil family. Coiled-coil proteins have a characteristic seven-residue repeat, $(\mathbf{a}\cdot\mathbf{b}\cdot\mathbf{c}\cdot\mathbf{d}\cdot\mathbf{e}\cdot\mathbf{f}\cdot\mathbf{g})_n$, with hydrophobic residues at positions **a** and **d** and polar residues generally elsewhere. Despite this shared HP pattern, coiled-coil sequences adopt dimeric (*4, 5*), trimeric (*6–8*), and anti-parallel tetrameric (*9*) conformations.

In addition, parallel, dimeric coiled coils exhibit strong preferences for specific amino acids at the hydrophobic **a** and **d** positions of the heptad repeat. A striking example is provided by the leucine zipper motif, which functions to dimerize bZIP transcription factors. Unlike other coiled coils, leucine zippers contain leucine at ~80 percent of all **d** positions (*5, 10*). Multiple substitution of these leucines with similarly-sized hydrophobic residues often interferes with dimerization function (*11*).

P. B. Harbury is in the Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115 and the Howard Hughes Medical Institute, Whitehead Institute, Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, MA 02142. T. Zhang and T. Alber were in the Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, UT 84132. P. S. Kim is in the Howard Hughes Medical Institute, Whitehead Institute, Department of Biology, Massachusetts Institute of Technology, 9 Cambridge Center, Cambridge, MA 02142.

*Present address: Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720.

The structural variety of the coiled-coil family and the functional requirements of leucine zipper sequences suggest that geometric properties of buried, apolar amino acids may influence the overall structure of coiled coils. To investigate this possibility we altered the hydrophobic core of a leucine zipper molecule in a concerted fashion and characterized the structures of the derivatives.

Hydrophobic core mutants form two-, three-, and four-stranded structures. We systematically mutagenized the hydrophobic core of the dimeric leucine zipper peptide GCN4-p1 (*12*). With rare exceptions, hydrophobic residues occupy the **a** and **d** positions of the GCN4-p1 sequence, and polar residues appear at **b, c, e, f,** and **g**, generating the characteristic $(\mathrm{H}\cdot\mathrm{P}\cdot\mathrm{P}\cdot\mathrm{H}\cdot\mathrm{P}\cdot\mathrm{P}\cdot\mathrm{P})_n$ pattern. Because supercoiled α helices have approximately 3.5 residues per turn, the spacing of the **a** and **d** positions three and four residues apart places the H residues on one side of the helix. In the GCN4-p1 dimer, the hydrophobic faces of two helices pack against each other in a parallel orientation (*12, 13*). Thus, the five amino acids at position **a** and the four leucines at position **d** from each monomer of GCN4-p1 form the apolar interface of the dimer (Fig. 1).

We simultaneously changed four **a** residues (Val^9,Asn^{16},Val^{23},Val^{30}) and four **d** residues (Leu^5,Leu^{12},Leu^{19},Leu^{26}) of GCN4-p1 to leucine, valine, or isoleucine (Met^2 at the first **a** position was not changed, Fig. 1) (*14*). These peptides were named by a two-letter code, the first letter indicating the residue at the four **a** positions and the second letter indicating the residue at the four **d** positions. The sequences were otherwise identical to GCN4-p1.

On the basis of circular dichroism measurements, each peptide was >90 percent helical at 4°C, neutral pH, and a concentration of 150 μM (*15*), and each exhibited a cooperative thermal unfolding transition (Table 1). The midpoint of the thermal transition (T_m) for each variant exceeded that of the parental GCN4-p1 peptide.

Equilibrium analytical ultracentrifugation (*16*) indicated that the peptides fall into three molecular weight classes: the peptides p-IL, p-II, and p-LI sedimented as dimeric, trimeric, and tetrameric species, respectively (Table 1). The oligomerization states of these peptides (p-IL, p-II, and p-LI) were independent of peptide concentration from 20 to 200 μM. A derivative of each peptide that contained the added sequence Cys-Gly-Gly at the NH_2-terminus also was synthesized (*12*). The cysteine residue permits disulfide bond formation, and the two glycine residues provide a flexible linker. Pairing peptide monomers with a covalent disulfide bond did not change the oligomerization state of p-IL and p-LI. In contrast, disulfide-bonded p-II sedimented with the molecular mass of a hexamer, consistent with the assignment of p-II as a trimer in the absence of a disulfide linkage.

The peptides p-VI, p-VL, p-LV, and p-LL populated multiple oligomerization states (Table 1). The sizes of the complexes formed by these variants were determined by gel filtration using the peptides p-IL, p-II, p-LI, and GCN4-p1 as size standards (*17*). The peptides p-VI and p-VL exhibited concentration-dependent retention times between 5 and 50 μM, and deriva-

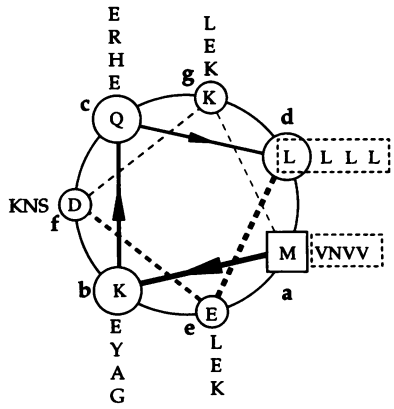

Fig. 1. Helical wheel projection of residues Met^2 to Glu^{32} of the GCN4-p1 sequence. View is from the NH_2-terminus, and residues in the first two helical turns are boxed or circled. Heptad positions are labeled **a** through **g**. In the mutant peptides described here, the residues in the dashed box at position **a** were collectively changed to I, V, or L, and, separately, the residues in the dashed box at position **d** were changed to I, V, or L (*41*). For example, p-LI contains leucine at the four boxed **a** positions and isoleucine at the four boxed **d** positions.

tives of p-VI, p-LV, and p-LL that contained an NH_2-terminal Cys-Gly-Gly disulfide linkage eluted as multiple species.

Helix orientation. The observation that the p-IL peptide remains dimeric with an NH_2-terminal disulfide linkage (Table 1) indicates that the helices are parallel. Consequently, the conformation of p-IL likely resembles the structure of GCN4-p1. The helices of the trimeric peptide p-II also are parallel because a two-dimensional double quantum-filtered correlation spectrum (DQF COSY) of the peptide showed only one magnetic environment for each residue (*18*), and the p-II peptide crystallized on a three-fold rotation axis (*19*). Thus, p-II assumes the same oligomeric structure as the trimeric stalk of influenza hemagglutinin (*8*).

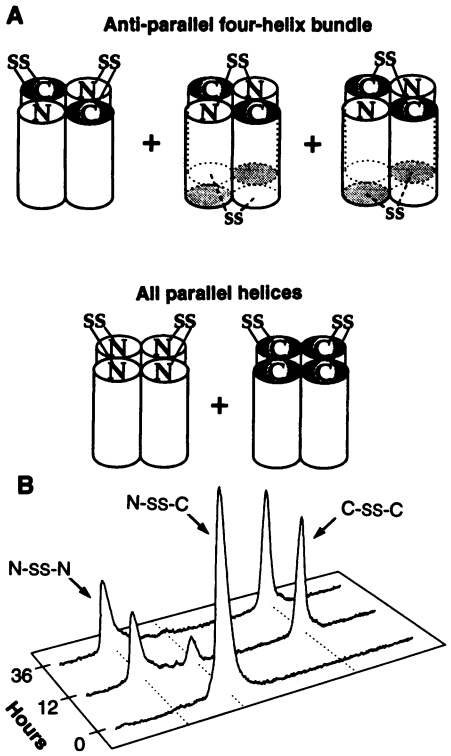

Fig. 2. The p-LI peptide forms a parallel tetramer in solution. (**A**) Given random pairing of terminal cysteine residues, the anti-parallel four-helix bundle conformation should produce a mixture of N-ss-C, N-ss-N, and C-ss-C disulfide bonds. The parallel conformation should produce only N-ss-N and C-ss-C disulfide bonds. (**B**) The disulfide-bonded p-LI-N·p-LI-C heterodimer (N-ss-C) rearranges to form homodimers (N-ss-N and C-ss-C). The disulfide-bonded heterodimer of p-LI-N (a variant of p-LI with the sequence CGG added to the NH_2-terminus) and p-LI-C (a variant of p-LI with the sequence GGC added to the COOH-teriminus) was purified and incubated in redox buffer to allow the exchange of disulfide bonds (*21*). At the indicated times a portion of the sample was removed, quenched with acid, and analyzed by HPLC. The p-LI-C peptide has a larger extinction coefficient than the p-LI-N peptide because it contains an additional tyrosine residue.

Unlike two- and three-stranded coiled coils, all well-characterized examples of four interacting α helices exist in an anti-parallel four-helix bundle conformation (Fig. 2A) (*20*). To determine the helix orientation of the p-LI tetramer, variants of the p-LI peptide with the sequence Cys-Gly-Gly at the NH_2-terminus (denoted p-LI-N) or with the sequence Gly-Gly-Cys at the COOH-terminus (denoted p-LI-C) were synthesized. The disulfide-bonded p-LI-N·p-LI-C heterodimer was purified and placed in redox buffer to allow the peptide monomers and the disulfide bonds to exchange (Fig. 2B) (*21*). Only homodimers were observed at equilibrium. Assuming that the glycyl linkers allow the terminal cysteines to assort randomly, the results indicate that the p-LI peptide assembles into a conformation containing four parallel α helices (*22*). To investigate the basis for the switch between parallel tetramer, dimer, and trimer conformations, the x-ray crystal structure of the p-LI peptide was determined at 2.1 Å resolution (Table 2) (*23*).

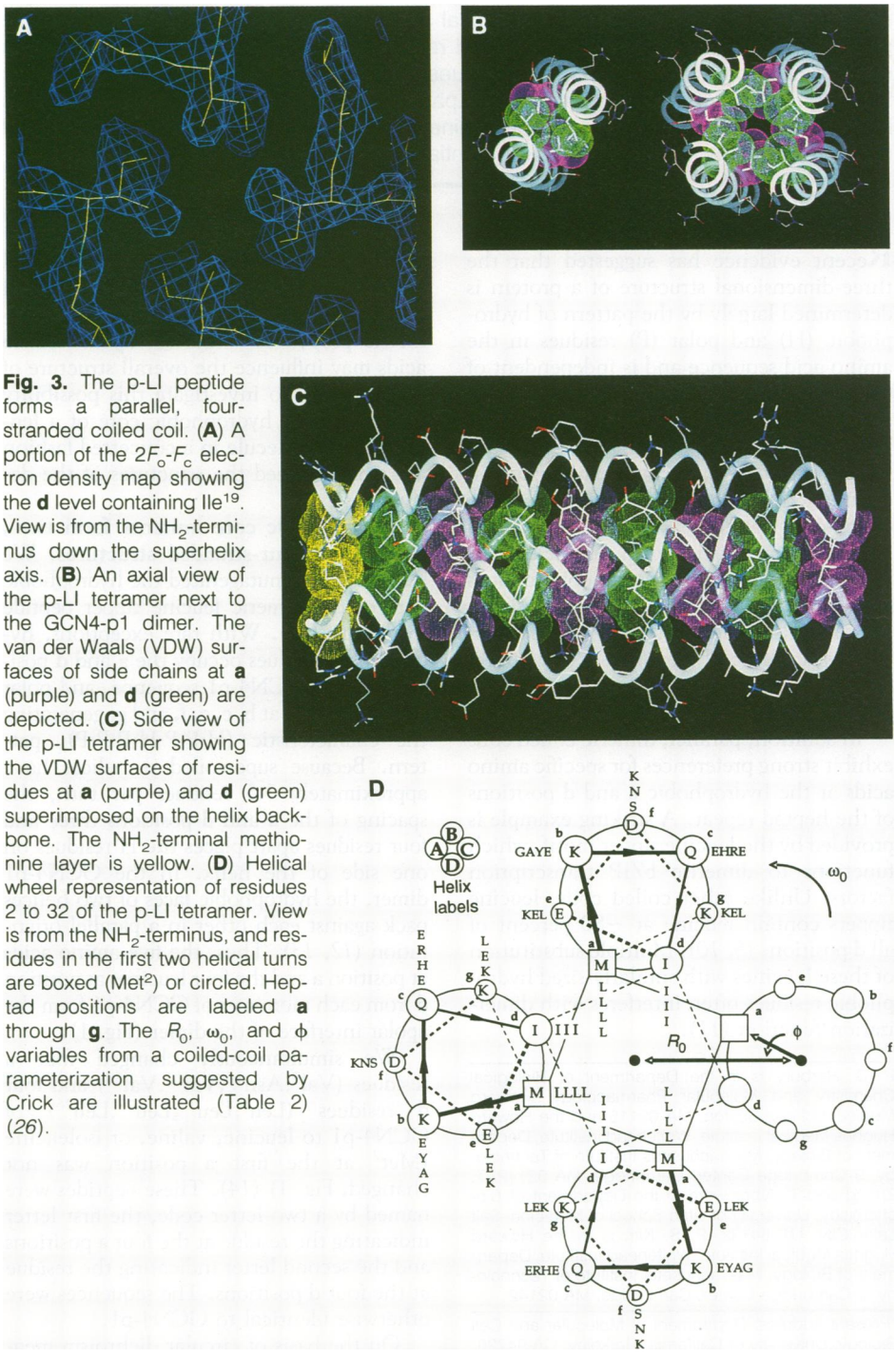

Fig. 3. The p-LI peptide forms a parallel, four-stranded coiled coil. (**A**) A portion of the $2F_o$-F_c electron density map showing the **d** level containing Ile^{19}. View is from the NH_2-terminus down the superhelix axis. (**B**) An axial view of the p-LI tetramer next to the GCN4-p1 dimer. The van der Waals (VDW) surfaces of side chains at **a** (purple) and **d** (green) are depicted. (**C**) Side view of the p-LI tetramer showing the VDW surfaces of residues at **a** (purple) and **d** (green) superimposed on the helix backbone. The NH_2-terminal methionine layer is yellow. (**D**) Helical wheel representation of residues 2 to 32 of the p-LI tetramer. View is from the NH_2-terminus, and residues in the first two helical turns are boxed (Met^2) or circled. Heptad positions are labeled **a** through **g**. The R_0, ω_0, and ϕ variables from a coiled-coil parameterization suggested by Crick are illustrated (Table 2) (*26*).

Structure of the tetramer. The p-LI tetramer consists of four parallel α helices wrapped in a left-handed superhelix (Fig. 3). The helices create a cylinder that is ~27 Å wide and ~48 Å long. An approximate fourfold axis of symmetry coincides with the superhelical axis (*24*).

The leucine and isoleucine side chains at the **a** and **d** positions point into the center of the tetramer. Cross-sectional layers containing leucine at the **a** positions alternate with layers containing isoleucine at the **d** positions (Fig. 3C). The dihedral angles χ_1 and χ_2 of all leucines at **a** and all isoleucines at **d** are approximately (−60,180), corresponding to the most populated rotamer for both amino acids (*25*). The side chains of Met^2 form the most NH_2-terminal **a** layer.

A cavity in the middle of each leucine and isoleucine layer forms a continuous central channel. The radius of the channel varies from 1.0 to 1.3 Å and therefore excludes a 1.4 Å radius water-molecule probe. No electron density was seen in the channel.

The gross differences between the structures of the p-LI tetramer and the GCN4-p1 dimer (*13*) may be summarized as follows: (i) Compared to the helices of the GCN4-p1 dimer, diagonally related helices of the p-LI tetramer have the same relative orientation but are 5.5 Å further apart (Fig. 3B). (ii) Adjacent helices of the p-LI tetramer are separated by approximately the same distance as the GCN4-p1 dimer helices, but each helix is rotated by ~45° around its own axis toward the center of the tetramer (Fig. 3B). Helical parameters are compared in Table 3 (*26*).

Helix interactions. On the basis of distance criteria (*27*), residues on the surface of the tetramer appear to form interhelical ion pairs. Of 12 possible interhelical salt bridges between residues at the **e** position of one heptad and the **g** position of the preceding heptad [for example, Glu^{B6}-Arg^{A1}, (*27*)], five are seen in the tetramer crystal structure. A similar frequency of interhelical **e** to **g** interactions was found in the crystal structure of GCN4-p1, which contains three of six possible **e** to **g** salt bridges (*13*). The tetramer also exhibits several types of ion pairs not present in the GCN4-p1 dimer. These include four of eight possible interhelical **g** to **b** salt bridges (for example, Lys^{D8}-Glu^{A10}, Fig. 4D) and five of eight possible interhelical **c** to **e** salt bridges (for example, His^{B18}-Glu^{C20}, Fig. 4E). Finally, a charge-stabilized hydrogen bond forms between Arg^{B25} and the main chain carbonyl oxygen of Leu^{C23}.

More surface area is buried in the tetramer (1640 $Å^2$ per helix) than in the dimer (900 $Å^2$ per helix) (*28*). Compared to the side chains of isolated helices, residues at the **a, d, e,** and **g** positions of the tetramer are substantially buried (>66 percent); residues at the **b** and **c** positions are partly buried (~15 percent); and the **f** positions remain completely exposed. The **e** and **g** residues in the tetramer are almost as buried as the **a** and **d** residues of the dimer (*28*). Similarly, the **b** and **c** residues of the tetramer are almost as buried as the **e** and **g** residues of the dimer.

The tetramer interface shows knobs-into-holes packing between helices (Fig. 4). As proposed by Crick (*29*), knobs formed by the side chains of one helix fit into holes formed by the spaces between side chains on the neighboring helix. Looking from the NH_2-terminus down the superhelix axis, each leucine knob at an **a** position packs into a hole formed by the **g** and **a** residues of the counterclockwise-related monomer (Fig. 4D) and by two **d** residues in adjacent layers along the superhelix axis. Similarly, each isoleucine knob at a **d** position packs into a hole formed by the **d** and **e** residues of the clockwise-related monomer (Fig. 4E) and by two **a** residues in adjacent layers.

The **a** and **d** layers of the tetramer exhibit two different types of knobs-into-holes packing that can be distinguished by the relative orientation of the knob to the hole. At the **a** level of the tetramer, the Cα-Cβ bond of each leucine knob makes a ~90° angle with the Cα-Cα vector at the bottom of the hole into which it packs (Fig. 4, A and B). We define this arrangement as perpendicular (⊥) knobs-into-holes packing. In contrast, at **d** levels the Cα-Cβ bond of each isoleucine knob is oriented parallel to the Cα-Cα vector at the bottom of the recipient hole (Fig. 4, A and C). We define this geometry as parallel (∥) knobs-into-holes packing.

Packing in two-, three-, and four-stranded coiled coils. Comparison of the side chain packing in the GCN4-p1 dimer and the p-LI tetramer shows that the local geometries of the **a** and **d** layers are reversed in the two structures; the dimer **d** level resembles the tetramer **a** level, and the dimer **a** level resembles the tetramer **d** level (Fig. 4, B and C). Perpendicular knobs-into-holes packing occurs at the **d** levels of

Table 1. Core mutants of GCN4-p1 form stable two-, three-, and four-helix structures.

Positions*		$-[\theta]_{222}$ (deg cm^2 $dmol^{-1}$)	T_m (°C)	T_m^{GdmCl}† (°C)	No. of helices‡	
a	**d**				Unmodified	SS
GCN4-p1		33,300	53	<0	2	2
I	L	32,400	>100	77	2	2
I	I	32,400	>100	70	3	6
L	I	30,600	>100	94	4	4
V	I	22,500§	73	<0	–	(4,6)
L	V	30,600	81	<0	3	(2,–)
V	L	32,400	95	49	(2,3)	2
L	L	31,500	>100	76	3	(2,4,6,–)

*The residues inserted at four **a** and four **d** positions of GCN4-p1 (*41*). †T_m^{GdmCl} denotes the melting temperature in 3 M GdmCl. All scans and melts were performed at 10 μM peptide concentration. ‡The number of helices in the solution complex formed by unmodified peptides and by disulfide bonded peptides. The first four peptides were assigned on the basis of equilibrium analytical ultracentrifugation data and the last four on the basis of gel filtration data. Parentheses indicate that multiple species were present; a dash indicates the presence of a species that could not be assigned. §p-VI exhibits a $-[\theta]_{222}$ value of 31,500 deg cm^2 $dmol^{-1}$ at 150 μM concentration.

Table 2. Data collection, phasing, and refinement statistics.

Parameter		Native	K_2PtCl_4
Space group		$P2_12_12_1$	$P2_12_12_1$
Unit cell dimensions (Å)	*a*	47.43	47.60
	b	48.63	48.94
	c	51.76	51.74
Unique reflections		5643 (2.1 Å)	1927 (2.9 Å)
R_{merge}†		0.07	
R_{iso}‡			0.20
Empirical *K*§			2.63
Number of sites			1
Mean figure of merit (20-3 Å)∥			0.58
Crystallographic *R*¶		0.15	
Rms Δbonds, Rms Δangles#		0.018 Å, 2.3°	

†$R_{merge} = \Sigma|I - \langle I\rangle|/\Sigma I$; *I*, intensity. ‡$R_{iso} = \Sigma|F_{PH} - F_P|/\Sigma F_P$; F_{PH} and F_P, derivative and native structure-factor amplitudes. §$K_{empirical} = 2\{\Sigma^*[F_{PH} - F_P]^2/\Sigma^*[F_{PH(+)} - F_{PH(-)}]^2\}^{1/2}$; $[F_{PH(+)} - F_{PH(-)}]$, anomalous difference between Friedel pairs; Σ^*, sum over acentric reflections. ∥Mean figure of merit = $\langle\|\Sigma_\alpha P(\alpha)e^{i\alpha}/\Sigma P(\alpha)\|\rangle$; α,phase; $P(\alpha)$, phase probability distribution. ¶$R_{cryst} = \Sigma|F_P - F_{calc}|/\Sigma F_P$; F_{calc}, calculated structure-factor amplitude. #Root-mean-square deviations from ideal values.

the dimer and the **a** levels of the tetramer (Fig. 4B). In contrast, parallel knobs-into-holes packing is found at the **a** levels of the dimer and the **d** levels of the tetramer (Fig. 4C). A third class of knobs-into-holes interaction appears at the **a** and **d** positions of parallel trimeric coiled coils (*8*). At both levels of the influenza hemagglutinin trimer, the Cα-Cβ bond of each knob makes a ~60° angle with the Cα-Cα vector at the base of the corresponding hole. We define this arrangement as acute knobs-into-holes packing (Fig. 4A).

Because the peptides p-IL, p-II, and p-LI differ only by volume-conserving hydrophobic substitutions at the buried **a** and **d** positions, packing interactions at **a** and **d** must mediate the switch between the dimer, trimer, and tetramer conformations. The geometric relationship between the dimer and tetramer structures suggests an explanation for the different oligomerization of the peptides p-IL and p-LI. The inversion of sequence at the **a** and **d** positions (Ile, Leu to Leu, Ile) coincides with an inversion of the packing geometry at **a** and **d** (parallel-perpendicular to perpendicular-parallel) in the two- and four-stranded conformations. Evidently, parallel packing has a geometric preference for isoleucine, or perpendicular packing has a geometric preference for leucine, or both (*30*).

Recent measurements of amino acid preferences in a dimeric coiled-coil model indicate that both types of packing exhibit the expected bias. Isoleucine is strongly favored (~0.4 kcal mol^{-1}) over leucine at the parallel geometry and leucine is more weakly favored (~0.1 kcal mol^{-1}) over isoleucine at the perpendicular geometry (*31*). If the **a** position of the tetramer is assumed to have the same residue specificity as the geometrically similar **d** position of the dimer (and similarly position **d** of the tetramer is assumed to have the same residue specificity as position **a** of the dimer), then it follows that the two- to four-stranded transition is driven primarily by preference for isoleucine at the levels with parallel packing.

However, at least two arguments support the opposite conclusion, that discrimination against isoleucine at perpendicular positions dominates the conformational switch. First, dimeric, fibrous coiled-coil sequences show a strong bias against isoleucine residues at **d** (⊥) positions and show no preference for isoleucine or leucine at **a** (∥) positions (*32*). Second, modeling studies indicate that isoleucine must adopt an infrequently observed rotamer (−,−) to occupy the perpendicular position of a dimeric coiled coil, while the most common rotamer of leucine may be accommodated at the parallel position (*33*). Poor packing of isoleucine at perpendicular positions also would explain why the p-II peptide forms a trimer. The dimer and tetramer conformations would each place four isoleucines of p-II in a perpendicular geometry, while the trimer conformation places all eight isoleucines in the acute geometry. Sequences of trimeric coiled coils show no strong bias for leucine or isoleucine at either the **a** or **d** positions (*6*).

In addition to efficient packing, a buried hydrogen bond between Asn^{16} residues specifically favors the dimer conformation of GCN4-p1, the parent molecule of the coiled coils described here (*13, 34*). The GCN4-p1 peptide forms a dimer, but substitution of Asn^{16} (at an **a** position) with valine causes the resulting peptide p-VL to populate both dimeric and trimeric conformations (Table 1). Despite this heterogeneity, p-VL exhibits a significantly higher apparent T_m than GCN4-p1. Thus, Asn^{16} imposes specificity for the dimer structure at the expense of stability (*35*).

Residues at the **b, c, e,** and **g** positions are more buried in the tetramer structure than in the dimer. Thus, the amino acid sequence at these positions can also influence oligomerization state (*31*).

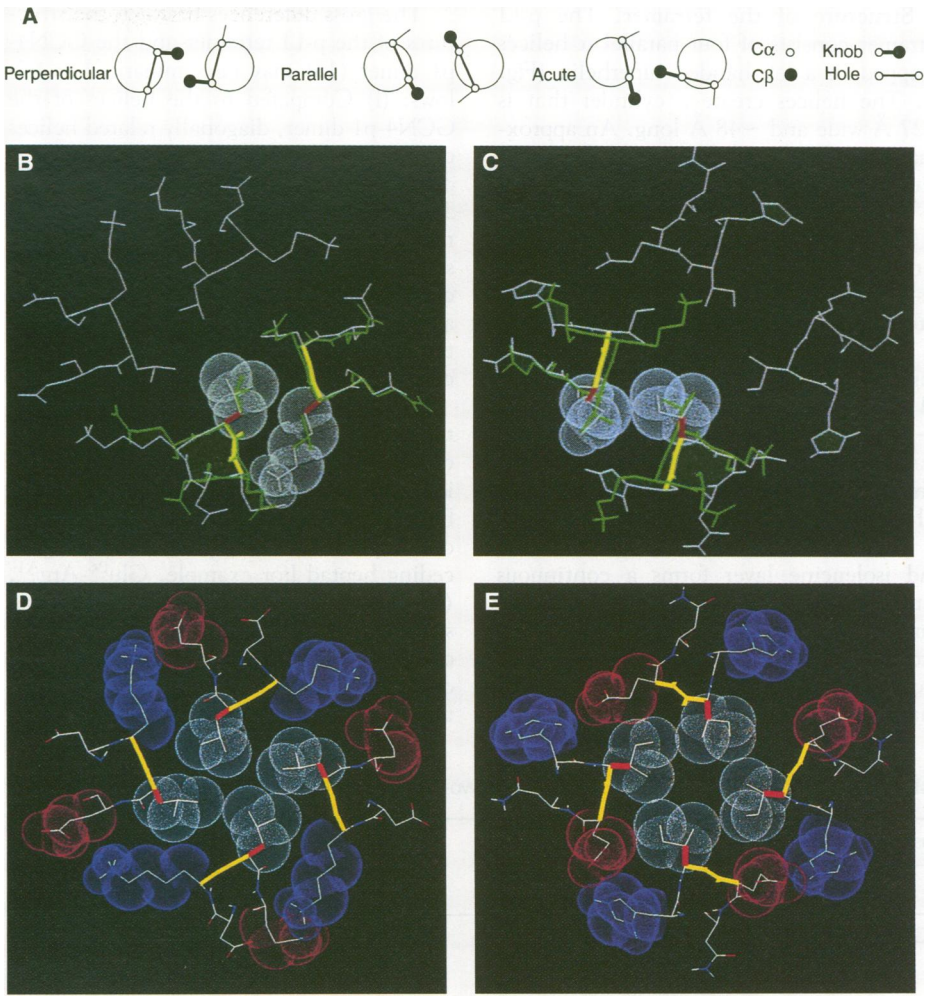

Fig. 4. Three types of knobs-into-holes packing and electrostatic interactions. (**A**) Schematic drawing (not to scale) showing the relative positions of the Cα-Cβ and Cα-Cα vectors for perpendicular, parallel, and acute knobs-into-holes packing. (**B**) Superposition of perpendicular packing in a GCN4-p1 dimer **d** level (green) and a p-LI tetramer **a** level (white). The Cα-Cβ bond of each knob (thick red line) makes a ~90° angle with the Cα-Cα vector at the base of the hole on the neighboring helix (thick yellow line). The two dimer **d** positions overlay on an **a** position and a **g** position of the tetramer (white VDW surfaces). (**C**) Superposition of parallel packing in a GCN4-p1 dimer **a** level (green) and a p-LI tetramer **d** level (white). The Cα-Cβ bond of each knob (thick red line) is parallel to the Cα-Cα vector at the base of the hole on the neighboring helix (thick yellow line). The two dimer **a** positions overlay on a **d** position and an **e** position of the tetramer (white VDW surfaces). (**D**) **g** to **b** salt bridges in the tetramer **a** level (Leu^9). Lysines at **g** positions (blue VDW surfaces) make salt bridges with glutamates at **b** positions (red VDW surfaces). White VDW surfaces idenfity leucines at **a** positions. (**E**) **c** to **e** salt bridges in the tetramer **d** level (Ile^{19}). Glutamates at **e** positions (red VDW surfaces) make salt bridges with histidines at **c** positions (blue VDW surfaces). White VDW surfaces identify isoleucines at **d** positions.

Table 3. Superhelical parameters calculated from the refined dimer (*13*) and tetramer structures using a parameterization suggested by Crick (*26*). The R_0, ω_0, and ϕ variables are illustrated in Fig. 3D.

Parameter	Dimer	Tetramer
Supercoil radius, R_0 (Å)	4.9	7.6
Amino acids per supercoil turn, ω_0	100	139
Supercoil pitch (Å)	147.6	205.4
Helix crossing angle	23.4°	26.0°
Radius of curvature (Å)	118	149
Position **a** orientation angle, ϕ	21.6°	19.8°

Protein structure. Recent studies have suggested that the tertiary folds of globular proteins are insensitive to the details of packing, and are determined instead by hydrophobic-polar (HP) and secondary structure (2S) patterns in the amino acid sequence. Several globular proteins, for example, accommodate multiple hydrophobic substitutions at core positions without adopting new folds (*1*). In addition, HP-2S sequence patterns, when compared with an HP-2S template derived from a known three-dimensional structure, are more sensitive indicators of structural homology than is direct sequence comparison (*2*). Finally, bipartite lattice models of proteins, consisting of H-type and P-type residues, fold to single, compact conformations under potentials that maximize HH contacts (*3*). Within the limitations of the computer models, the drive to bury hydrophobic residues can specify unique tertiary structures.

The heptad HP-2S pattern is clearly not sufficient to guide a coiled-coil sequence to a single structure. Unlike globular proteins, however, coiled coils have a very simple HP pattern and contain unconnected units of secondary structure. Such properties allow coiled-coil monomers to assemble in multiple configurations. Thus, rearrangement of the hydrophobic core in response to mutations, which globular proteins appear to accomplish while conserving tertiary fold, can be manifested in coiled coils as dramatic changes in overall structure. Consistent with this argument, conservative core mutations often result in loss of function for coiled coils (*11*) but not for their globular counterparts (*1*).

Recent attempts to design helical proteins based on HP-2S patterns have produced molecules that lack unique packing (*36*). For example, the four-helix bundle α_4 exhibits a well-defined secondary structure and is very stable, but the leucine side chains in the core of the molecule have fluctuating conformations (*37*). Within the GCN4-p1 variant family, HP-2S design criteria predict that the p-LL sequence should form an optimal coiled coil (*38*). For the dimer and tetramer conformations, however, p-IL and p-LI exhibit equal or higher thermal stabilities than p-LL (Table 1). In the case of the trimer conformation, p-II, though less stable than p-LL, exhibits higher specificity for the trimer fold. The results suggest that by ignoring specific packing interactions, HP-2S design methods fail to produce high conformational selectivity and sometimes predict sequences with less than optimal stability.

Practical implications. The mutants of GCN4-p1 clarify the functions of conserved features of leucine zipper sequences. The leucine repeat at **d** positions and the preponderance of β-branched amino acids at **a** positions favor dimer formation due to packing considerations (Table 1), while conserved asparagines at **a** positions direct dimerization by forming buried hydrogen bonds.

Together with analyses of coiled-coil sequences (*4*, *6*), our results also suggest that predictions of coiled-coil oligomerization can be based in part on the distribution of β-branched residues at the **a** and **d** positions. The occurrence of β-branched residues at **d** positions disfavors dimers, while β-branched residues at **a** positions should disfavor tetramers. The presence of β-branched residues at both the **a** and **d** positions facilitates trimer formation.

The tetramer structure, which contains a large axial cavity, may serve as a soluble model for membrane ion channels. Related sequences that contain serine at the **a** and **d** positions have been shown to conduct small cations (*39*). Substituting serine for leucine and isoleucine residues in the crystallographic coordinates of the p-LI tetramer generates a model with an enlarged interior channel lined by the hydroxyl groups of the serine side-chains.

Finally, the short, modular peptides p-II, p-LI, and p-IL may be used to design two-, three-, and fourfold oligomers of attached sequences. Addition of p-VL, alternatively, would produce a hybrid capable of interconverting between dimer and trimer states. Such inherent structural heterogeneity could form the basis for a regulatory switch (*40*).

REFERENCES AND NOTES

1. W. A. Lim and R. T. Sauer, *J. Mol. Biol.* **219**, 359 (1991); J. H. Hurley, W. A. Baase, B. W. Matthews, *ibid.* **224**, 1143 (1992); S. Dao-pin, D. E. Anderson, W. A. Baase, F. W. Dahlquist, B. W. Matthews, *Biochemistry* **30**, 11521 (1991); D. Shortle, W. E. Stites, A. K. Meeker, *ibid.* **29**, 8033 (1990); W. E. Stites, A. G. Gittis, E. E. Lattman, D. Shortle, *J. Mol. Biol.* **221**, 7 (1991).
2. J. U. Bowie, R. Luthy, D. Eisenberg, *Science* **253**, 164 (1991).
3. K. F. Lau and K. A. Dill, *Macromolecules* **22**, 3986 (1989); H. S. Chan and K. A. Dill, *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6388 (1990); A. Sikorski and J. Skolnick, *Biopolymers* **28**, 1097 (1989).
4. J. F. Conway and D. A. D. Parry, *Int. J. Biol. Macromol.* **12**, 328 (1990).
5. J. C. Hu and R. T. Sauer, in *Progress in Nucleic Acids and Molecular Biology*, X. X. Eckstein and X. X. Lilley, Eds. (Springer-Verlag, Berlin, 1992), vol. 6, pp. 82.
6. J. F. Conway and D. A. D. Parry, *Int. J. Biol. Macromol.* **13**, 14 (1991).
7. P. K. Sorger and H. C. Nelson, *Cell* **59**, 807 (1989); R. Peteranderl and H. C. Nelson, *Biochemistry* **31**, 12272 (1992); B. Lovejoy *et al.*, *Science* **259**, 1288 (1993).
8. I. A. Wilson, J. J. Skehel, D. C. Wiley, *Nature* **289**, 366 (1981); W. I. Weis, A. T. Brunger, J. J. Skehel, D. C. Wiley, *J. Mol. Biol.* **212**, 737 (1990).
9. D. W. Banner, M. Kokkinidis, D. Tsernoglou, *ibid.* **196**, 657 (1987).
10. W. H. Landshultz, P. F. Johnson, S. L. McKnight, *Science* **240**, 1759 (1988).
11. J. C. Hu, E. K. O'Shea, P. S. Kim, R. T. Sauer, *ibid.* **250**, 1400 (1990); T. Kouzarides and E. Ziff, *Nature* **336**, 646 (1988); M. Neuberg, M. Schuerman, J. B. Hunter, R. Muller, *Science* **338**, 589 (1989); R. Gentz, F. Rauscher, C. Abate, T. Curran, *ibid.*, p. 1695; R. Turner and R. Tjian, *ibid.*, p. 1689; L. J. Ransone, J. Visvader, C. P. Sassone, I. M. Verma, *Genes Dev.* **3**, 770 (1989); M. Schuerman *et al.*, *Cell* **56**, 507 (1989); V. J. Dwarki, M. Montminy, I. M. Verma, *EMBO J.* **9**, 225 (1990).
12. E. K. O'Shea, R. Rutkowski, P. S. Kim, *Science* **243**, 538 (1989).
13. E. K. O'Shea, J. D. Klemm, P. S. Kim, T. Alber, *ibid.* **254**, 539 (1991).
14. Peptides were synthesized and purified as in D. J. Lockhart and P. S. Kim, *Science* **257**, 947 (1992). The identities of the peptides were confirmed by mass spectrometry (Finnigan MAT), and all molecular weights were found to be within 1 dalton of the expected mass. The peptides have acetylated NH_2-termini and free COOH-termini.
15. CD spectra were measured on an Aviv 60DS spectropolarimeter. Measurements of $[\theta]_{222}$ were made at 4°C in 50 mM phosphate (pH 7.0), 150 mM NaCl, and 10 μM peptide. Thermal melts were performed in the same buffer and also with the addition of 3 M guanidinium chloride (GdmCl) to facilitate unfolding of the extremely stable coiled coils. Values of T_m were estimated from $[\theta]_{222}$ versus temperature data (in two-degree steps) by evaluating the maximum of $d[\theta]/dT^{-1}$ [C. R. Cantor and P. R. Schimmel, *Biophysical Chemistry* (W. H. Freeman and Company, New York, 1980), vol. 3, pp. 1132]. Peptide concentrations were determined spectrophotometrically [H. Edelhoch, *Biochemistry* **6**, 1948 (1967)] with an extinction coefficient of 1280 $cm^{-1}M^1$ at 280 nm.
16. Analytical ultracentrifugation measurements were made on a Beckman XLA centrifuge equipped with an An-60Ti rotor. Ten or fifteen data sets (20 to 200 μM at two or three rotor speeds) were fit simultaneously to a single molecular weight with the program HID4000 [M. L. Johnson, J. J. Correia, D. A. Yphantis, H. R. Halvorson, *Biophys. J.* **36**, 575 (1981)], provided by M. L. Johnson and J. Lary. Molecular weights were determined under two different conditions: (i) 50 mM sodium phosphate (pH 7.1), 500 mM NaCl at 4°C and (ii) 50 mM sodium phosphate (pH 2.0), 100 mM NaCl at 0°C. Samples were dialyzed against the appropriate buffer for at least 12 hours before each experiment. Specific volumes and solvent density were calculated according to T. M. Laue, B. D. Shah, T. M. Ridgeway, S. L. Pelletier, in *Analytical Ultracentrifugation in Biochemistry and Polymer Science*, S. Harding, A. Rowe, J. Horton, Eds. (Royal

Society of Chemistry, Cambridge, 1992), pp. 90–125. At pH 7.1, molecular weights derived from the low concentration data showed no systematic deviation from the molecular weights derived from the high concentration data. The molecular weights derived from the complete pH 7.1 data sets (followed by expected molecular weight) and [the rotor speeds in krpm's at which data were collected] are: p-II: 12,500 (12,240) [27, 32]; p-II-N: 26,100 (25,782) [22, 27, 32]; p-LI: 16,300 (16,320) [27, 32]; p-LI-N: 17,600 (17,188) [27, 32]; p-IL: 8600 (8160) [27, 32]; p-IL-N: 8500 (8594) [27, 32, 37]. At pH 2.0, molecular weights derived from the low concentration data were slightly (<10 percent) but systematically larger than the molecular weights derived from the high concentration data. However, weights determined from fits with and without the inclusion of a second virial coefficient differed by less than 5 percent in all cases. The molecular weights derived from the complete pH 2.0 data sets are: p-II: 12,400 (12,240) [27, 32, 37]; p-II-N: 25,800 (25,782) [24, 29, 34]; p-LI: 16,300 (16,320) [27, 32, 37]; p-LI-N: not determined; p-IL: 8700 (8160) [27, 32, 37]; p-IL-N: 8,900 (8594) [27, 32, 37]. No systematic variation of the residuals was observed at either pH.

17. Gel filtration experiments were performed with a TosoHaas G2000SW_{xl} column (7.8 mm by 30 cm) in 50 mM sodium phosphate (pH 7.1), 500 mM NaCl at 4°C on a Waters 625 microbore HPLC. Samples (10 μl) with initial peptide concentrations of 1 mM and 100 μM were chromatographed at 0.5 ml min^{-1} with absorbance monitored at 205, 214, and 275 nm. Concentrations reported in the text were calculated from the peak absorbance at 275 nm. The peptides p-IL, p-II, p-LI, and GCN4-p1 were used as size standards. The standards eluted from the column as single peaks with retention times that were independent of peptide concentration between 5 and 50 μM. Essentially identical results were obtained with a Synchropak GPC 100 column (4.6 mm by 25 cm) in 50 mM sodium phosphate (pH 7.1), 150 mM NaCl, 2 M GdmCl at 25°C.
18. Phase sensitive double quantum-filtered (DQF) COSY spectra of p-II were collected (37°C, 5 mM peptide, pH 5) and analyzed as for the GCN4-p1 peptide in T. G. Oas *et al.*, *Biochemistry* **29**, 289 (1990). The H^N to H^α crosspeak region contained 33 peaks, consistent with a single conformation for the 33 residue peptide or for multiple conformations that are in rapid exchange on the NMR timescale. The chemical shift dispersion was comparable to that seen in the GCN4-p1 COSY spectrum.
19. Crystals of peptide p-II were grown by vapor diffusion from 10 mg ml^{-1} peptide mixed 1:1 with 100 mM sodium acetate (pH 3.0), 85 mM Na_2SO_4, 10 percent polyethylene glycol (PEG) 1400, 10 percent isopropanol and equilibrated against a reservoir of mother liquor. The crystals had the symmetry of space group $P4_132$ with unit cell dimensions of 61.21 Å and an asymmetric volume of 9556 $Å^3$, corresponding to one monomer. Because p-II was found to be a trimer in solution under similar conditions, crystallographic symmetry likely relates the three members of the oligomer. The symmetry operations of $P4_132$ could only generate a parallel structure. Strong reflections at a 5.2 Å resolution, which arise from the axial repeat of a coiled coil, were observed along the threefold axes of the crystal.
20. C. Cohen and D. A. D. Parry, *Proteins* **7**, 1 (1990).
21. The p-LI-N (Ac-CGG R MKQIEDK LEEILSK LYHIENE LARIKKL LGER) (*41*) and p-LI-C (Ac-Y R MKQIEDK LEEILSK LYHIENE LARIKKL LGER GGC) peptides were air-oxidized in 6 M GdmCl. The heterodimer was purified by preparative HPLC and incubated (25 μM concentration, ~25°C) in redox buffer containing 50 mM tris (pH 9.0), 100 mM NaCl, 2 M GdmCl, 500 μM reduced glutathione, 125 μM oxidized glutathione in an anaerobic chamber (Coy Laboratory Products Inc.) Reactions were quenched with concentrated acetic acid (10 percent final concentration by volume) and analyzed by reversed-phase HPLC (Vydac C-18 column, stock #218TP54) with a linear gradient of acetonitrile-water in the presence of 0.1 percent trifluoroacetic acid.
22. Fiber diffraction data suggest that honeybee silk consists of a parallel four-stranded coiled coil. [see (*20*) and E. D. T. Atkins *J. Mol. Biol.* **24**, 139 (1967)].
23. The p-LI peptide, Ac-R MKQIEDK LEEILSK LYHIENE LARIKKL LGER (*41*), was crystallized by vapor diffusion from 10 mg ml^{-1} peptide mixed 2:1 with 100 mM sodium acetate (pH 5.1), 350 mM NaCl, 5 percent polyethylene glycol 1400 and equilibrated against a reservoir of mother liquor. The crystals had the symmetry of space group $P2_12_12_1$ with a = 47.43 Å, b = 48.63 Å, c = 51.77 Å, and four monomers in the asymmetric unit. Data to 2.9 Å resolution were collected from both native and K_2PtCl_4 derivative crystals with a Rigaku AFC6R diffractometer. The isomorphous and anomalous difference Patterson maps indicated the presence of a single heavy atom site per asymmetric unit. An ideal four-stranded coiled-coil model with the four methionines pointing radially inward toward the single bound Pt was fit into the resulting single-isomorphous-replacement map. The last two residues of each peptide were left out of the model because of the absence of interpretable electron density for these atoms. After rigid-body, least-squares, and simulated-annealing refinement with XPLOR [A. T. Brunger, J. Kuriyan, M. Karplus, *Science* **235**, 458 (1987)], the model exhibited reasonable geometry and gave an R factor of 0.20 with a single B factor of 25 $Å^2$ for all atoms. FRODO [T. A. Jones, *J. Appl. Crystallogr.* **11**, 268 (1978)] and TNT [D. E. Tronrud, L. F. TenEyck, B. W. Matthews, *Acta Crystallogr.* **A43**, 489 (1987)] were used for manual rebuilding and refinement of atomic positions and individual isotopic B factors against 6–2.1 Å native data collected using an R-Axis image plate detector. The present model includes 124 amino acids of the tetramer and 81 water molecules. The following observations support the accuracy of this model: (i) The R factor is 0.15 for the strongest 95 percent of the data between 6.0 and 2.1 Å resolution. These data represent 73 percent of all possible reflections in this resolution range. (ii) Bond lengths and bond angles of the model have root mean square deviations from ideality of 0.018 Å and 2.3°, respectively. (iii) All main chain dihedral angles fall within allowed regions of the Ramachandran plot, and most side chains assume well populated rotamer conformations. (iv) The model fits the $2F_o-F_c$ map well, and the F_o-F_c map contoured at $\pm 3\sigma$ has no interpretable features. (v) The heavy atom difference map calculated with model phases contains a single 10σ peak that is within 3 Å of each of the four methionine sulfur atoms. Pertinent statistics are listed in Table 2. The coordinates of the refined model have been deposited in the Brookhaven Protein Data Bank.
24. The tetramer lacks perfect fourfold symmetry. Two helices on one diagonal (A and C; Fig. 3A) are closer than the two helices on the other diagonal (B and D). For the methionines at the NH_2-terminal **a** level, the diagonal Cα-Cα distances differ by approximately 3 Å. The difference in the diagonal Cα-Cα distances falls to 0.5 Å at Ile^{12}. From Ile^{12} to Ile^{26} a 0.5 Å difference persists. At the most COOH-terminal level **a** (Leu^{30}), the difference in diagonal distances increases to 1.3 Å.
25. J. W. Ponder and F. M. Richards, *J. Mol. Biol.* **193**, 775 (1987). The rotamer conventions in this reference are followed here.
26. Superhelical parameters were obtained by fitting the Cα supercoil backbone to a parameterization suggested by Crick [F. H. C. Crick, *Acta Crystallogr.* **6**, 685 (1953)]. The radius (R_0), frequency (ω_0) and pitch (P) of the superhelix and the radius (R_1) and phase (ϕ) of the α helix were treated as variables while the frequency (ω_1) of the α helix was fixed at 4π per 7 residues to preserve the periodicity of the heptad structure. The values for the supercoil pitch and crossing angle reported here for the GCN4-p1 helices differ from those in (*13*) because the superhelical parameters were calculated previously with Crick's assumed boundary condition $\omega_0 + \omega_1 = 100°$ per residue.
27. Charged heavy atoms separated by less than four angstroms were treated as possible ion pairs. Glu^{B6} denotes glutamate, monomer B, residue number six (Fig. 3D).
28. Surface areas were calculated with a 1.4-Å radius probe with the program CHARMM [B. R. Brooks *et al.*, *J. Comput. Chem.* **4**, 187 (1983)]. Percent buried surface area is expressed as the fraction of accessible side chain surface area in the isolated helix that becomes buried in the oligomer. The NH_2-terminal **g** and **a** positions and the COOH-terminal **a** and **b** positions (that cap the ends of the helices) were omitted from the calculation. Buried surface area for the GCN4-p1 dimer (percent): **a**, 87; **b**, 0; **c**, 1; **d**, 87; **e**, 26; **f**, 0; **g**, 27. Buried surface area for the p-LI tetramer (percent): **a**, 92; **b**, 10; **c**, 19; **d**, 99; **e**, 72; **f**, 0; **g**, 66. The side chains of residues at **a**, **d**, **e**, and **g** account for 90 percent of the total area buried on oligomer formation.
29. F. H. C. Crick, *Acta Crystallogr.* **6**, 689 (1953).
30. Because perpendicular and parallel packing are defined by the positions of Cα-Cα and Cα-Cβ vectors, the packing geometries at the **a** and **d** positions depend on the relative backbone orientation of the interacting helices and not on the identity or the conformation of the side chain.
31. B. Y. Zhu, N. E. Zhou, C. M. Kay, R. S. Hodges, *Protein Science* **2**, 383 (1993); K. J. Lumb and P. S. Kim, unpublished results.
32. Among more than 4800 heptads from dimeric fibrous coiled coils, isoleucine and valine together occur at only 10 percent of all **d** positions (perpendicular) while leucine appears at over 41 percent of all **d** positions. At **a** positions (parallel), valine and isoleucine together occur with the same frequency as leucine (~27 percent) (*4*). The expected values at fully buried α-helical positions are ~30 percent for valine and isoleucine together and ~32 percent for leucine, on the basis of (*2*) and amino acid usage [T. E. Creighton, *Proteins* (W. H. Freeman, New York, 1993)].
33. The χ_1 and χ_2 dihedral angles were enumerated for p-II and p-LL and built into a dimeric coiled coil backbone. The potential of the system was minimized with the adopted basis Newton-Raphson algorithm and the PARAM19 potential of CHARMM (*28*). The lowest energy dihedral angles for leucine at an **a** position corresponded to (−,t), the most common leucine rotamer, while the lowest potential dihedral angles for isoleucine at position **d** corresponded to (−,−), an infrequently observed isoleucine rotamer (*25*).
34. C. Chothia, *J. Mol. Biol.* **105**, 1 (1976).
35. In 2 M GdmCl [50 mM sodium phosphate (pH 7.0), 150 mM NaCl, 34 μM peptide], the apparent T_m of p-VL exceeds that of GCN4-p1 by 45°C (C. M. Carr and P. S. Kim, unpublished results).
36. W. F. DeGrado, D. P. Raleigh, T. Handel, *Current Op. in Struc. Biol.* **1**, 984 (1991).
37. D. P. Raleigh and W. F. DeGrado, *JACS* **114**, 10079 (1992); T. M. Handel, S. L. Williams, W. F. DeGrado, *Science* **261**, 879 (1993).
38. Measurement of helix-promoting ability in a coiled coil places leucine fourth among the natural amino acids (behind alanine, arginine and lysine) and places isoleucine and valine twelfth and fifteenth, respectively [K. T. O'Neil and W. F. DeGrado, *Science* **250**, 646 (1990)]. A similar rank order is obtained when helix propensities are measured in isolated helices [S. Padmanabhan, S. Marqusee, T. Ridgeway, T. M. Laue, R. L. Baldwin, *Nature* **344**, 268 (1990); P. C. Lyu, M. I. Liff, L. A. Marky, N. R. Kallenbach, *Science* **250**, 669 (1990); G. Merutka and E. Stellwagen, *Biochemistry* **30**, 1591 (1991)]. According to a recent structure-based analysis (*2*), leucine ranks second only to phenylalanine in frequency of appearance at fully buried, α-helical positions [3D-1D scores of 1.32(F), 1.27(L), 1.17(I), and 0.66(V) in the $B_1\alpha$ environment].
39. W. F. DeGrado and J. D. Lear, *Biopolymers* **29**, 205 (1990).
40. S. K. Rabindran, R. I. Haroun, J. Clos, J.

Wisniewski, C. Wu, *Science* **259**, 230 (1993).
41. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr.
42. We thank R. Rutkowski and M. Burgess for peptide synthesis, B. Santarsiero of Molecular Structures Corporation for collecting high resolution data from the p-LI crystals, and C. Carr, A. Cochran, D. Lockhart, Z. Y. Peng, B. Schulman, J. Staley, B. Tidor, and J. Weissman for discussions and critical reading of the manuscript. P.S.K. is a Pew Scholar in the Biomedical Sciences. This research was supported by grants from the National Institutes of Health (GM44162 to P.S.K. and GM48958 to T.A.) and by grants from the Lucille P. Markey Charitable Trust and the American Cancer Society (T.A.).

22 June 1993; accepted 26 October 1993

A Third Recognition Element in Bacterial Promoters: DNA Binding by the α Subunit of RNA Polymerase

Wilma Ross, Khoosheh K. Gosink, Julia Salomon, Kazuhiko Igarashi, Chao Zou, Akira Ishihama, Konstantin Severinov, Richard L. Gourse*

A DNA sequence rich in (A + T), located upstream of the −10, −35 region of the *Escherichia coli* ribosomal RNA promoter *rrnB* P1 and called the UP element, stimulates transcription by a factor of 30 in vivo, as well as in vitro in the absence of protein factors other than RNA polymerase (RNAP). When fused to other promoters, such as *lacUV5*, the UP element also stimulates transcription, indicating that it is a separable promoter module. Mutations in the carboxyl-terminal region of the α subunit of RNAP prevent stimulation of these promoters by the UP element although the mutant enzymes are effective in transcribing the "core" promoters (those lacking the UP element). Protection of UP element DNA by the mutant RNAPs is severely reduced in footprinting experiments, suggesting that the selective decrease in transcription might result from defective interactions between α and the UP element. Purified α binds specifically to the UP element, confirming that α acts directly in promoter recognition. Transcription of three other promoters was also reduced by the COOH-terminal α mutations. These results suggest that UP elements comprise a third promoter recognition region (in addition to the −10, −35 recognition hexamers, which interact with the σ subunit) and may account for the presence of (A + T)–rich DNA upstream of many prokaryotic promoters. Since the same α mutations also block activation by some transcription factors, mechanisms of promoter stimulation by upstream DNA elements and positive control by certain transcription factors may be related.

The strength of promoters recognized by $E\sigma^{70}$, the most abundant of the *E. coli* RNAP holoenzymes, can be correlated to a considerable extent with their similarity to consensus recognition hexamers in the core promoter region, centered approximately 10 and 35 bp upstream of the start site of transcription, and the spacing between these hexamers (*1*). Nevertheless, it has been proposed that sequences outside of the core promoter region can modulate promoter activity (*2*). Upstream sequences have been shown to increase the activities of several *Escherichia coli* or *Bacillus subtilis* promoters in vitro in the absence of protein factors other than RNAP (*3–9*), and regions rich in (A + T) have been noted upstream of many promoters (*2, 10, 11*).

The *rrnB* P1 promoter is representative of a class of seven rRNA promoters in *E. coli* that together account for more than half of the transcription in the cell at high growth rates (*12*). Although the *rrnB* P1 core promoter is sufficient for specific initiation and for response to the two systems known to regulate transcription under different nutritional conditions—namely, growth rate–dependent control and stringent control (*13, 14*)—the region upstream of the core promoter is largely responsible for its high activity (*4, 9, 13, 15, 16*). A 20–base pair (bp) region rich in (A + T), to which we refer as the upstream (UP) element, is located immediately upstream of the core promoter (Fig. 1). The UP element increases *rrnB* P1 activity by a factor of at least 30 in the absence of protein factors other than RNAP (*9, 17*). The UP element is protected by RNAP in footprinting experiments, and replacement of the UP element with non-*rrnB* DNA results in severe reduction of protection in the upstream region (*8, 18, 19*). Therefore, the core and UP element together can be considered an extended promoter (Fig. 1) (*9*). The region adjacent to the UP element (between bp −60 and bp −150) contains binding sites for the activator protein Fis, which results in increasing the activity of the promoter by a factor of 10 (*8, 16, 20, 21*). Fis is not required for stimulation of transcription by the UP element (*9*).

The σ^{70} subunit of RNAP holoenzyme ($\alpha_2\beta\beta'\sigma$) interacts with the −10 and −35 hexamers (*22*). However, the region or regions of RNAP required for UP element recognition have not been defined. On the basis of studies with mutant derivatives of the α subunit, it has been proposed that α interacts directly with certain transcription factors, leading to stimulation of promoter activity (*23–26*). We therefore used mutants of α to investigate the role of this subunit in UP element function although transcription activation in this case is achieved by a DNA element rather than by a trans-acting protein. Two mutant forms of the 329 amino acid α subunit, COOH-terminal truncations of 73 or 94 amino acids (α-256 or α-235, respectively), are stable in vivo and assemble into holoenzyme (*27*). Furthermore, reconstituted α-235 or α-256 core enzymes ($\alpha_2\beta\beta'$), prepared in vitro from purified subunits,

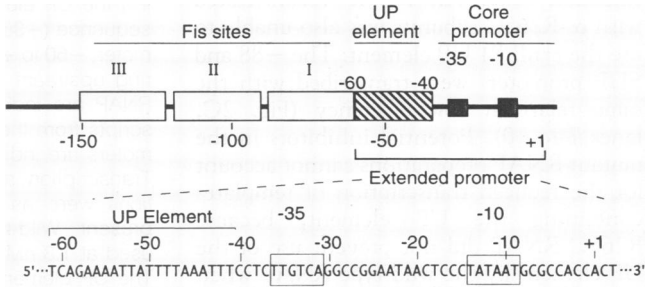

Fig. 1. The *rrnB* P1 promoter. The extended promoter region includes elements recognized by RNA polymerase: the core promoter, which consists of −10 and −35 consensus hexamers (filled boxes), and the UP element, which consists of bp −40 to −60 (shaded box) (*9*). Three binding sites for the transcriptional activator protein Fis (open boxes) occur upstream of the UP element. Site I is responsible for most of the activation by Fis at this promoter (*20, 21*). The sequence of the extended promoter is indicated below the diagram.

W. Ross, K. K. Gosink, J. Salomon, and R. L. Gourse are in the Department of Bacteriology, University of Wisconsin–Madison, 1550 Linden Drive, Madison, WI 53706. K. Igarashi, C. Zou, and A. Ishihama are in the Department of Molecular Genetics, National Institute of Genetics, Mishima, Shizuoka 411, Japan. K. Severinov is at the Public Health Research Institute, 455 First Avenue, New York, NY 10016.

*To whom correspondence should be addressed.