3

Considerations in the Design and Optimization of Coiled Coil Structures

Jody M. Mason, Kristian M. Müller, and Katja M. Arndt

Summary

Coiled coil motifs are, despite their apparent simplicity, highly specific, and play a significant role in the understanding of tertiary structure and its formation. The most commonly observed of the coiled coils, the parallel dimeric, is yet to be fully characterized for this structural class in general. Nonetheless, strict rules have emerged for the necessity of specific types of amino acids at specific positions. In this chapter, we discuss this system in light of existing coiled coil structures and in applying rules to coiled coils that are to be designed or optimized. Understanding and expanding on these rules is crucial in using these motifs, which play key roles in virtually every cellular process, to act as drug-delivery agents by sequestering other proteins that are not behaving natively or that have been upregulated (for example, by binding to coiled coil domains implicated in oncogenesis). The roles of the **a** and **d** "hydrophobic" core positions and the **e** and **g** "electrostatic" edge positions in directing oligomerization and pairing specificity are discussed. Also discussed is the role of these positions in concert with the **b**, **c**, and **f** positions in maintaining α-helical propensity, helix solubility, and dimer stability.

Key Words: Coiled coil; helix; heptad repeat; in vivo selection; leucine zipper; library design; protein design; protein engineering; protein fragment complementation assay; protein stability; rational design.

1. Introduction

The coiled coil is a common structural motif estimated to constitute 3 to 5% of the encoded residues in most genomes *(1)*. It consists of two to five α-helices that habitually twist around each other, typically left-handedly, to form a supercoil. Whereas regular α-helices go through 3.6 residues for each complete turn of the helix, the distortion imposed on each helix within a left-handed coiled coil lowers this value to 3.5. This equates to a seven amino acid repeat for every two turns of the helix *(2,3)*. The most frequently occurring type of coiled coil

From: *Methods in Molecular Biology, vol. 352: Protein Engineering Protocols*
Edited by: K. M. Arndt and K. M. Müller © Humana Press Inc., Totowa, NJ

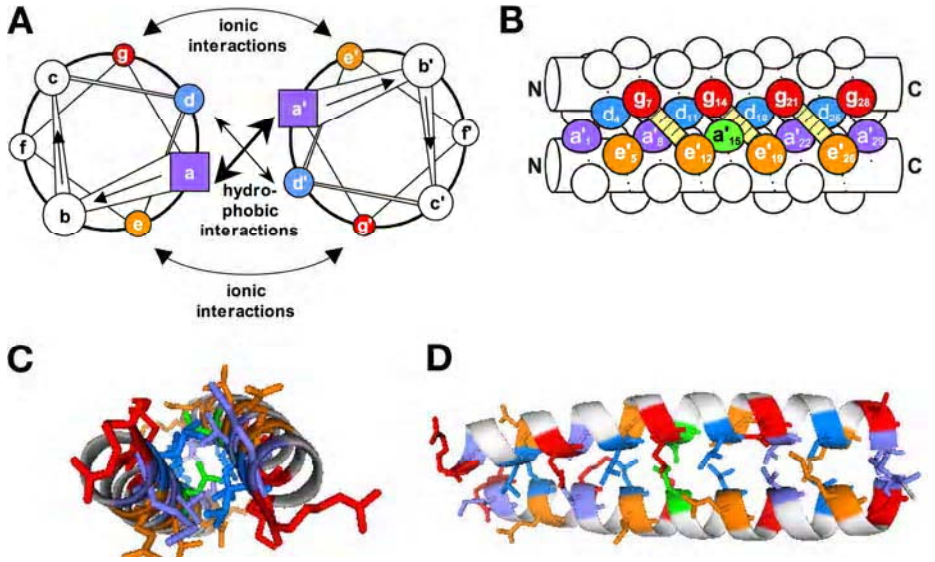

Fig. 1. Dimeric parallel coiled coil. **(A)** Helical wheel diagram looking down the helix axis from the N-terminus to the C-terminus. Heptad positions are labeled **a** to **g** and **a′** to **g′**, respectively. Positions **a**, **d**, **e**, and **g** are in different shades of gray. **(B)** Side view. The helical backbones are represented by cylinders, the side chains by knobs. The path of the polypeptide chain is indicated by a line wrapped around the cylinders. For simplicity, the supercoiling of the helices is not shown. While residues at positions **a** (dark gray) and **d** (light gray) make up the hydrophobic interface, residues at positions **e** (medium gray) and **g** (black) pack against the hydrophobic core. They can participate in interhelical electrostatic interactions between residue *i* (**g** position) of one helix and residue *i′* + 5 of the other helix (**e′** position, belonging to the next heptad), as indicated by the hatched bars. **(C,D)** Coiled-coil domain of the yeast transcription factor GCN4 (*see* **Note 1**) as ribbon plot (Protein Data Bank code: 2ETA; **ref.** *6*) to indicate supercoiling and **g/e′** interactions. The plot was made using Pymol *(7)*.

is the parallel (i.e., both helices run N to C alongside each other) dimeric, left-handed variety. In this class, the periodicity of each helix is seven, with anywhere from 2 (in designed coiled coils; **ref.** *4*) to 200 of these repeats in a protein *(5)*. In this repeat, the residues are designated $(\mathbf{a\text{-}b\text{-}c\text{-}d\text{-}e\text{-}f\text{-}g})_n$ in one helix, and $(\mathbf{a'\text{-}b'\text{-}c'\text{-}d'\text{-}e'\text{-}f'\text{-}g'})_n$ in the other (**Fig. 1**). In this model, **a** and **d** are usually nonpolar core residues found at the interface of the two helices, conversely, **e** and **g** are partially solvent exposed polar "edge" residues that give specificity between the two helices through electrostatic interactions. Finally, the remaining three residues (**b**, **c**, and **f**) are typically hydrophilic and exposed to the solvent. The apparent simplicity of the coiled coil structure, with its heptad periodicity, has led to extensive study. Remarkably, interaction between

the helices remains a highly specific process. It is this interplay of seemingly simplistic structural periodicity, combined with both high specificities and impressive affinities, which make this ubiquitous structural class of proteins so fascinating.

In both natural and designed two- and three-stranded coiled coils, complementary charge pairs on the edge of the interface that relieve repulsive pairs in alternate oligomers are sufficient to promote formation of hetero-oligomers. We call this idea the peptide velcro (PV) hypothesis *(8)*, in reference to the design by Kim and co-workers of an obligate heterodimeric coiled coil termed "Peptide Velcro" *(9)*. This pair of peptides is identical except at the **e** and **g** positions, where one sequence contains Lys and the other sequence contains Glu (*see* **Note 2**). This peptide pair (like other similar pairs) forms a stable heterodimer in vitro. A recent study tested the PV hypothesis by directly comparing rational design and genetic selection strategies *(8)*. Contrary to the PV hypothesis (but in agreement with the sequence properties of many natural coiled coils), the selected pairs neither maximized predicted attractive **g/e′** charge pairs nor eliminated predicted repulsive **g/e′** charge pairs (*see* **Subheading 2.2.3.**). A variety of factors can influence the contributions of **g/e′** ionic residues. Overall electrostatic potential, including intermolecular and intramolecular interactions, plays a major role; and interactions with the core residues, such as favorable packing or steric clashes, have also been proposed to modulate **g/e′** interactions. Other effects of the sequence context may arise from local helix flexibility or from interactions with **b**, **c**, or **f** residues. Examination of coiled coil structures also suggests that the **e** and **g** positions are structurally different, and these differences may accommodate specific charge pairs in different ways.

Here, we outline the importance of individual amino acids in maintaining α-helical structure and promoting the formation of a specific coiled coil structure of a desired oligomeric state and orientation required for a left-handed coiled coil. The aim of this chapter is to highlight the considerations required in the design or optimization of a coiled coil in this category. This chapter should, therefore, serve as a "protocol" to facilitate coiled coil design by explaining the most important aspects to obtain the desired oligomerization state (**Subheading 2.1.**), specificity (**Subheading 2.2.**), helix orientation (**Subheading 2.3.**), and stability (**Subheading 2.4.**). We discuss the influence of amino acids at the **a** and **d** hydrophobic core positions together with the **e** and **g** electrostatic edge positions, and the role of these together with the **b**, **c**, and **f** positions in maintaining α-helical propensity, helix solubility, and overall dimer stability. Additionally, N-capping and C-capping preferences are discussed. Unless specifically stated in the text, we use the dimeric parallel coiled coil motif as reference state. We also discussed stability and specificity of coiled coils in more general terms in a recent review *(10)*.

Understanding of the rules governing association of helices has already permitted these proteins to be exploited in novel ways *(11)*; for example, by fusing antibody Fv fragments to the helices to create a helix stabilized antibody *(12)*, by fusing antibody scFv fragments to the helices to create miniantibodies *(13)*, or as a thermo-sensor (in the case of TlpA, which was tagged to green fluorescent protein, with fluorescence changes monitored as a result of structural changes in TlpA during temperature changes). This could permit measurements of signal transduction processes involving coiled coil dimer formation *(14)*.

2. Materials and Methods

The several different aspects comprising the specific design of coiled coils are discussed in this section. We aim to facilitate the choice of amino acids at the core and the edge positions to achieve the desired oligomerization state (**Subheading 2.1.**), specificity (**Subheading 2.2.**), and helix orientation (**Subheading 2.3.**). Here, we also relate different design solutions to the respective stabilities. The fourth section (**Subheading 2.4.**) relates to the overall stability, focusing on the outer (**b**, **c**, and **f**) positions. In this chapter, we attempt to break down all of the aspects to be considered into the relevant subheadings. Nonetheless, by the very nature of the interplaying factors leading to the formation of the coiled coil, discussions in these sections will invariably crossover in places. Subsections dealing with different effects exerted by the same residue positions should, therefore, be regarded as companion pieces.

2.1. Oligomeric State

Any protein must achieve the desired three-dimensional structure to function properly. Likewise, a coiled coil structure must adopt the correct oligomeric (quaternary) structure. In the following section, the main factors contributing to the adoption of this correct state are discussed.

2.1.1. The Core Residues

The **a** and **d** residues are called core residues because they form a hydrophobic strip that winds around each helix. The nonpolar nature of the **a** and **d** repeat facilitates oligomerization along one face of each helix. This is analogous to a hydrophobic core, which collapses during the folding of globular proteins, and represents a dominating contribution to the overall stability of the coiled coil. Consequently, the core residues exert a major influence on defining the oligomerization state.

The hydrophobic side chains in positions **a** and **d** bury into the neighboring helix in a "knobs-into-holes" manner, first described by Crick in 1953 *(15)*. In this model, a side chain from one α-helix (the knob) packs into a space surrounded by four side chains of the opposite α-helix (the hole), and vice versa. These

Table 1
Influence of GCN4-p1 Core Mutations on Oligomerization State

Position **a**–**d**	Geometry of side chains at **a**–**d**	Oligomerization state	T_m [a]
GCN4-p1[b]	β–γ	Dimer	53°C
I–L	β–γ	Dimer	>100°C (77°C)
I–I	β–β	Trimer	>100°C (70°C)
L–I	γ–β	Tetramer	>100°C (94°C)
V–I	β–β	—[c]	73°C
L–V	γ–β	Trimers	81°C
V–L	β–γ	Mixture of dimers and trimers	95°C
L–L	γ–γ	Trimers	>100°C (76°C)

Adapted from **ref.** *16*.

[a]T_m measurements were performed in 50 m*M* phosphate buffer, pH 7.0; 150 m*M* NaCl; and 10 μ*M* peptide. In parentheses are T_m values measured in 3 *M* GuHCl.

[b]The wild-type GCN4-p1 differs from the V–L mutant by an Asn pair at the central core **a** position, which ensures dimer formation (*see* **Subheading 2.1.2.**).

[c]Species that could not be assigned.

packing geometries are defined by the angle that the Cα–Cβ bond of the knob forms with the Cα–Cα vector at the base of the hole on a projection from the end of the coiled helices. For a parallel dimer, a knob in position **a** (in heptad *i*) fits into the hole in the other helix that is lined up in a clockwise sequence (if looking into the hole) by residues in positions $\mathbf{a}'_i$, $\mathbf{d}'_i$, $\mathbf{g}'_{i-1}$, and $\mathbf{d}'_{i-1}$. Accordingly, a knob at $\mathbf{d}_i$ is in the hole lined by $\mathbf{d}'_i$, $\mathbf{a}'_i$, $\mathbf{e}'_i$, and $\mathbf{a}'_{i+1}$ (**Fig. 1**).

Exhaustive analysis of different core mutants of the coiled coil of GCN4 (a yeast homolog to the Jun transcription factor, sometimes referred to as GCN4-p1, *see* **Note 1**) revealed different packing geometries for different oligomerization states (**Table 1**; **ref.** *16*). Comparison of the side-chain packing in the X-ray structure of the GCN4-p1 dimer and a designed tetrameric GCN4 mutant showed that the local geometries of the **a** and **d** layers are reversed in the two structures. Parallel knobs-into-holes packing was found at the **a** layer of the dimer and at the **d** layer of the tetramer. In contrast, perpendicular knobs-into-holes packing was observed at the **d** layer of the dimer and the **a** layer of the tetramer. A third class of knobs-into-holes interaction appeared at the **a** and **d** positions of the parallel trimeric variant *(17)*. In both layers, the Cα–Cβ bond of each knob makes an approx 60° angle with the Cα–Cβ vector at the base of the corresponding hole. This arrangement was termed "acute" knobs-into-holes packing.

These different geometries account for a distinct preference of amino acids for a certain oligomerization state. The following list describes the outcome of several experiments in which various amino acids were tested in the context of stability and oligomerization specificity.

1. Specific hydrophobic residues are crucial in ordaining the oligomeric state of the coiled coil. Harbury et al. systematically changed (with the exception of the **a**1 Met) all **a** and **d** residues of GCN4 to Leu, Val, or Ile (*see* **Note 2**; **ref.** ***16***). As shown in **Table 1**, this led to coiled coils of different oligomeric states. The GCN4-IL (IL referring to Ile at positions **a** and Leu at positions **d**), II, and LI mutants were dimeric, trimeric, and tetrameric, respectively, and were independent of concentration over the entire concentration range investigated. The VI, VL, LV, and LL mutants gave rise to multiple oligomeric states. Each of the L, V, and I combinations gave rise to distinct packing preferences and, thus, to distinct geometries.
   a. β-branched residues (Val and Ile) are favored in the parallel knobs-into-holes packing (the **a** layer of the dimer and the **d** layer of the tetramer), whereas the γ-branched residue Leu is favored in the perpendicular geometry (the **d** layer of the dimer and the **a** layer of the tetramer). Conversely, the insertion of a β-branched amino acid into the perpendicular position would require adoption of a thermodynamically unfavorable rotamer ***(18)***.
   b. It also seems that Ile and Val, despite similar stereochemistry, are not equivalent in establishing the oligomerization states. A Val at the **a** positions is less able to specify dimers than Ile, which gave a higher dimer specific interaction, as opposed to a dimer–trimer mixture for Val ***(17)***.
   c. In contrast to the dimer and tetramer structures, the interior packing of the trimer can accommodate β-branched residues in the most preferred rotamer at both hydrophobic positions.
2. In a study by Woolfson and Alber, the role of the core residues were also studied and used to assign dimer and trimer propensities to unambiguous heptad registers ***(19)***. Dimers and trimers were analyzed for distinguishing features, and these features were used to identify new sequences. The frequency of buried Leu, Ile, Asn, Lys, and Gln were key in this prediction algorithm, called COILER. In total, 21 different proteins known to form parallel dimeric and trimeric proteins were used, equating to some 721 heptads in the database.
   a. Of the initially considered seven amino acids (Ala, Phe, Ile, Len, Met, Val), only Ile and Leu at the **a** site and Leu at the **d** site were observed with the required statistical significance. Dimers were found to be favored by enrichments of IIe at the **a** and Leu at the **d** positions, whereas Ile was strongly depleted at the **d** positions of dimers. Clearly, the selection of the said residues at these positions gives the best packing geometries for dimer formation (*see* **item 1** and **Table 1**).
   b. Val is distributed more evenly than Ile at the core positions of coiled coil sequences. It even occurs with frequencies less than those expected by chance at the **a** and **d** positions of dimers and trimers. These results are consistent with the observation that Val at the **a** and **d** sites discriminates little between dimeric and trimeric coiled coils (*see* **item 1** and **Table 1**).
   c. Packing at the **a** and **d** sites in a trimer are comparable, and particular residue selections for these positions are consequently less specific, leading to more evenly distributed hydrophobic amino acids ***(16,19)***.

3. In studying the positional effect of alanine substitutions in the core of a designed antiparallel coiled coil (*see* **Note 3**), Monera et al. found that when the alanine residues are in register (i.e., on the same rung), a dimer forms ***(20)***. If the alanines are out of register, the helices form a tetramer. The most likely explanation for this is that the cavity formed with alanines in register in the tetramer would be highly destabilizing, and, therefore, the dimer is favored, whereas the Leu–Ala repeats are able to distribute the cavity over a larger region and minimize loss of hydrophobic burial and van der Waals interactions. This demonstrates the oligomerization specificity that is generated by core residue packing.
4. The cartilage oligomeric matrix protein (COMP), which belongs to the thrombospondin family, contains an extremely stable five-stranded parallel α-helical coiled coil. The 46-amino acid-long coiled coil region (*see* **Note 4**) includes a ring of intermolecular (i.e., helix to helix) disulfide-bonded cysteines ***(21)***. The pentameric interface displays knobs into holes packing, with the knobs formed by the **a**, **d**, **e**, and **g** positions, which pack into holes created between side chains at positions **a′–g′**, **d′–e′**, **c′–d′**, and **a′–b′** of the adjacent subunit. Only the residue at position **f** remains completely exposed; the other six positions are significantly buried. The structures of the tetrameric GCN4-pLI mutant (*see* **item 1**) and the pentameric COMP both contain a large axial cavity. The channel in the tetramer varies from 1.0 to 1.3 Å ***(16)*** and, therefore, excludes water molecules (radius 1.4 Å). In contrast, several water molecules were found along the pore of the pentamer, which is consistent with the larger diameter of the channel (2–6 Å).
5. Studies in which the **d** residues have been changed to nonnatural amino acids that are even more hydrophobic in nature (trifluoroleucine and hexafluoroleucine) revealed an increase in stability (*see* **Note 5**; **refs.** ***22*** and ***23***).
6. Hydrophobic burial at the **a/d** interface has been investigated using monomethylated, dimethylated, and trimethylated analogs of diaminopropionic acid (dap), which display increasing degrees of hydrophobicity ***(22)***. Addition of one methyl group to position 16 of one of the monomers (with aspartic acid at position 16 in the analogous peptide), stabilizes the subsequently heterodimeric fold of GCN4-p1 (*see* **Note 6**), possibly because of increased van der Waals interactions in the folded state and a lower desolvation penalty on folding. However, addition of three methyl groups results in destabilization, probably because the increased steric bulk is poorly accommodated. Curiously, addition of two methyl groups to the synthetic dap causes homotrimerization. This demonstrates how small changes in size and hydrophobicity can alter the stability and folding preferences.

In short, the best choices for amino acid at the core in dimeric coiled coils seem to be Leu at **d** positions, and β-branched Ile (or Val) at **a** positions. If Val is used, an Asn (as commonly found in natural coiled coils) should be incorporated at a central **a** position to add specificity to the interaction (*see* **Subheading 2.1.2.**). Trimers are best designed with an all Ile core, whereas tetramers favor Leu at the **a** and Ile at the **d** positions. Deviations from this β/γ side-chain branching arrangement will lead to unfavorable rotamer energies and to a lower stability of the

desired structure, and could generate coiled coils of mixed oligomeric states and antiparallel alignment and, thus, of reduced specificity.

2.1.2. Polar Core Residues

Despite core residues being, on the whole, nonpolar, some 20% of core residues are charged *(25)*. These charged residues often serve to specify correct oligomeric states, presumably by ensuring a required helix alignment in structures that would otherwise exist as a mixture of dimers, trimers, and/or tetramers, as well as parallel or antiparallel arrangements. However, this gain in specificity is usually accompanied by a decrease in stability. The following list documents the role of these residues in terms of specificity and stability.

1. In their statistical analysis of dimers and trimers (*see also* **Subheading 2.1.1.**, **item 2**) Woolfson and Alber *(19)* observed that:
   a. Lys and Asn are favored at **a** positions in dimers but depleted from trimers. Asn is three times more likely to be found at an **a** position within a dimer than at the same position in a trimer.
   b. Trimers are enriched in Gln residues at **a** sites, and Ser and Thr residues are enriched at either **a** or **d** sites.
   c. Buried lysines at **a** positions are often found in conjunction with glutamates at the flanking **e** or **g** positions. This is, for example, seen in the X-ray structures of the Jun/Fos heterodimer *(26)* as well as in the GCN4 Asn16Lys mutant (*see* **Subheading 2.1.2.**, **item 5**; **ref.** *27*).
2. The buried Asn pair confers dimer specificity (and in-register alignment), possibly through interhelical hydrogen bond formation between Asn side chains, and this is indeed observed in X-ray *(6)* and nuclear magnetic resonance studies *(28)*. Other conformations that do not satisfy the hydrogen-bonding potential of the Asn side chains are, therefore, energetically disfavored.
3. If the core **a** Asn of GCN4-p1 (*see* **Note 2**) is mutated to Val, the coiled coil experiences a huge increase in stability at the expense of dimerization specificity. Harbury et al. reported a mixture of dimers and trimers because of the lack of specificity conferred by a Val at the **a** position as compared with an Ile (*see* **Subheading 2.1.1.**, **item 1** and **Table 1**; **ref.** *16*), whereas Potekhin et al. reported trimer formation *(29)*.
4. In another case, the core Asn pair of the parallel heterodimer Peptide Velcro (*see* **Note 1**) was mutated to Leu (yielding the peptides, Acid-pLL and Base-pLL), which resulted in a mixture of parallel and antiparallel tetramers *(30)*.
5. In GCN4 (*see* **Note 2**), Alber's group exchanged the Asn pair at the core **a** residue to Gln and Lys, to investigate whether these too were able to confer oligomeric specificity. Lys formed dimers similar to the wild-type, whereas Gln formed a mixture of dimers and trimers *(27)*. They reasoned that the structural uniqueness dictated by the polar group is not merely caused by polar burial, but is also dependent on correct interaction of the side chain with the surroundings. These context effects are much more difficult to predict than mere residue frequencies within a given heptad position.

6. During selection for heterodimeric coiled coils with a protein fragment complementation assay (*see* **Subheading 2.5.3.**) using dihydrofolate reductase, Arndt et al. found a core Asn pair to be favored over Asn–Val or Val–Val combinations in an otherwise Val–Leu core (*see* **Note 7**; **ref.** ***31***). This is in good agreement with many naturally occurring coiled coils.
7. In another study, an **a** or **d** position of the dimeric GCN4-pVL (*see* **Subheading 2.1.1.**, **item 1**) was mutated to a single polar residue Asn, Gln, Ser, or Thr, respectively (*see* **Note 8**; **ref.** ***25***). Only Asn pairs at an **a** position and Thr pairs at a **d** position were capable of conferring the correct state. It is likely that the desolvation penalty on burying the residues in the core is vanquished by their interaction energy, which may also serve to ensure correct alignment.
8. Differences in packing environments yield different preferences for hydrophobic residues, even within the **a** and **d** positions. In two exhaustive studies using a model peptide with Val at **a** and Leu at **d** positions, the central **a** and **d** positions were systematically changed to every amino acid to assess their effects on stability and oligomerization state (*see* **Note 9**; **refs.** ***32*** and ***33***). These changes were the first comprehensive quantitative assessment of the effect on the stability of two-stranded coiled coils of side-chain substitution within the hydrophobic core, and permitted a relative thermodynamic stability scale to be constructed for the 19 naturally occurring amino acids in the **a** and **d** positions. **Table 2** lists those amino acids that gave rise to a well-defined oligomerization state if placed either at the central **a** or **d** position ***(32)***. Leu-, Tyr-, Gln-, and His-substituted **a** site analogs were found to be exclusively three stranded, whereas the Asn-, Lys, Orn-, Arg-, and Trp-substituted analogs formed exclusively two-stranded monomers. When substituting for the central **d** position, Ile and Val (the β-branched residues) induced the three-stranded oligomerization state (as detailed in **Subheading 2.1.1.**, **item 1**), whereas Tyr, Lys, Arg, Orn, Glu, and Asp induced the two-stranded state.
9. Ji et al. mutated gp41, a six-helix bundle envelope protein from simian immunodeficiency virus, that is, along with gp120, responsible for viral fusion with $CD4^+$ cells ***(34)***. Structurally, it consists of a trimer formed by antiparallel heterodimers. In this study, each of the four buried polar residues responsible for core hydrogen bonds and salt bridges (two Gln residues and two Thr residues) were individually mutated to Ile. Of these, three formed more-stable six-helix bundles, whereas one formed insoluble aggregates (*see* **Note 10**). These results demonstrate the role that such residues have in governing a structural stability balance and specificity. This is important because the protein undergoes a structural change before fusion and must have the correct stability balance between the two structures to render this permissible. These polar core residues aid in regulating this conformational stability and, hence, in membrane fusion itself.

In general, Asn pairs at the core **a** position clearly dominate in dimers, especially if the cores deviate from the optimal Ile–Leu **a–d** residues, because this combination appears to result in parallel dimers without the need for Asn pairs within the core. A core **a** position Gln may be a good choice for trimers, although, to confer exclusive specificity for trimers, additional factors may be required.

Table 2
Systematic Change of the Central a and d Positions to Every Amino Acid Using a Model Peptide That Otherwise Has a Val at the a and Leu at the d Positions[a]

	Position **a**		Position **d**	
	Oligomerization[b]	Normalized stability[c]	Oligomerization	Normalized stability
Leu	Trimer	100	(69% Dimer)	100
Ile	(61% Dimer)	105	Trimer	89
Val	(57% Trimer)	108	Trimer	63
Tyr	Trimer	74	Dimer	67
Trp	Dimer	55	(78% Dimer)	47
Gln	Trimer	41	(61% Dimer)	56
Asn	Dimer	56	(72% Trimer)	41
Lys	Dimer	37	Dimer	25
Orn	Dimer	10	Dimer	7
Arg	Dimer	31	Dimer	9
His	Trimer	28	(55% Dimer)	37
Glu	(54% Dimer)	10	Dimer	12
Asp	Dimer[d]	—	Dimer	24

[a]Amino acids that lead to a defined oligomerization state are shown. (From **refs. *32*** and ***33***.)
[b]Helices were disulfide bridged via an N-terminal Cys–Gly–Gly linker when assessing the oligomerization state (*see* **Note 9**). However, the reported oligomerization state relates to the number of helices.
[c]Normalized stability represents the stability of each substituted analog relative to Gly = 0 and Leu = 100.
[d]The Asp analog was 100% folded at 5°C. At room temperature, the analog was only ~20% folded.

2.1.3. Edge Residues

The **e** and **g** (edge) positions of the heptad repeat flank the **a** and **d** residues in coiled coil interfaces (**Fig. 1**). Burial of these positions highly depends on the oligomerization state. Consequently, the choice of amino acids at the **e** and **g** sites may be influenced by the oligomerization state. **Table 3** shows the calculated percent buried surface area expressed as the fraction of accessible side-chain surface area in the isolated helix that becomes buried in the oligomer ***(16)***.

1. Compared with the corresponding sites of dimers, the **e** and **g** positions of trimers are enriched for hydrophobic residues (Ile, Leu, Val, Phe, Tyr, and Trp) and depleted of specific hydrophilic residues (Glu, Gln, Ser, and Arg; **ref. *19***). These patterns are consistent with the extension of the hydrophobic interface of trimers, relative to that in dimers. This increase in percentage of hydrophobic residues causes the width of the narrow hydrophobic face to increase, and, with it, the likelihood of higher oligomerization states, in which more nonpolar burial can occur than in a two helix coiled coil. This can be seen in **Table 3**, in which the percentage of hydrophobic burial at the **e** and **g** positions is increased by approx 40%.

Table 3
Percentage of Buried Surface Areas for GCN4-p1 Dimer and p-LI Tetramer

Position	GCN4-p1 dimer	GCN4-pLI tetramer
a	87	92
b	0	10
c	1	19
d	87	99
e	26	72
f	0	0
g	27	66

From **ref. *14***.

The decrease in oppositely charged $\mathbf{g}_i$ to $\mathbf{e}'_{i+1}$ pairs in trimers (12%) compared with dimers (23%) is consistent with this ***(19)***.

2. Fairman et al. mutated the C-terminal homotetrameric coiled coil domain of the lac repressor to generate a heterotetramer ***(35)***. Peptides containing either all Lys or all Glu at the **b** and **c** positions, which flank the **e** and **g** positions weakly associate, but, if these are mixed, a highly stable tetramer is formed (*see* **Note 11**). This demonstrates that, at least for tetramers, the **b** and **c** residues also play a significant role in the stability of the coiled coil. This is a role akin to the **g/e′** ionic interactions found in dimeric coiled coils, but with the widened hydrophobic interface of the tetramer extending to these residues, the ionic role falls further outwards from the core to the **b** and **c** residues. By changing pH and salt levels, these ion pair interactions between Glu and Lys were shown to be responsible for the increased stability. Additionally, these charges direct against homo-oligomers, and it may be that this unfavorable charge repulsion in potential homodimeric interactions drives the heterodimer formation ***(9,35)***.

2.2. Pairing Specificity

The following section discusses the importance of pairing specificity, and some of the ways in which the coiled coil ensures that no other energetically favorable structures can be accessed. Remarkably, despite their similarity in sequence and structure, coiled coils interact preferentially with functional partners. This section analyzes the factors mediating such high selectivity.

2.2.1. Core Residues

The patterning of hydrophobic residues (mostly Leu, Ile, and Val), as outlined in **Subheading 2.1.1.**, is a dominant driving force behind the association of the helices. However, for this pattern to be observed so frequently, how can the coils, at the same time, use the core to direct against alternative structures forming? The answer is a complicated picture involving subtle changes, such as

insertions of nonstabilizing, nonhydrophobic core residues, which will select against alternative structures.

1. Sharma et al. designed a peptide (anti-APCp1) that is targeted to bind a coiled coil sequence from the adenomatous polyposis coli (APC) tumor suppressor protein, which is implicated in colorectal cancers (*see* **Note 12**; **ref. *36***). In this, they used core changes together with **g/e′** interactions, rather than Asn pairings, to ascribe specificity to the interaction. They reasoned that the low requirement discovered for core residues in driving specificity is surprising considering their dominant influence on association, and that core mutations can have large effects on stability and specificity. To address this issue, they designed a peptide to bind to the first 55 amino acids of APC (APC55) and mutated this anti-APCp1 to generate the more-frequently observed **a–a′** and **d–d′** pairings based on covariation patterns at the **a** and **d** positions of keratin type I and type II heterodimers. They made three mutations (A41I at layer **a** and A2M and M44A at layer **d**) to change the wild-type Ala–Ala and Met–Met interactions to the more-frequently found Ala–Ile, Ala–Met, and Met–Ala interactions, respectively. Two further mutations (T6G in layer **a** and N30H in layer **d**) served to destabilize the respective homodimers with Gly–Gly and His–His pairs. Additional **e–g′** pairings optimized ionic interactions while directing against anti-APCp1 homodimerization. The resulting heterodimer, APCp1/APC55, was both stable and specific.
2. Schnarr and Kennan formed heterotrimeric proteins by steric matching of core hydrophobic residues ***(37)***. In their study, unnatural residues of various side-chain lengths were used to promote specific heterotrimer formation. The authors replaced a core **a** position of GCN4 with Ala or cyclohexylalanine (*see* **Note 13**). The result was a sterically mismatched core layer in the trimer, with either three Ala or three cyclohexylalanines, generating steric void or repulsion, respectively. Two Ala and a cyclohexylalanine, however, generated a heterotrimer with good steric matching. The use of nonnatural side chains can be used in this way to generate coiled coils. The additional bulk of the cyclohexylalanine complements the Ala core layers to provide a steric match, whereas bulkier side chains only serve to destabilize the molecule.

2.2.2. Polar Core Residues

The roles of polar core residues in directing specific oligomerization states of coiled coils were discussed in **Subheading 2.1.2**. Heterotypic core contacts that permit generation of heterospecificity in coiled coil pairings were mentioned in **Subheading 2.1.1**. Further specificity can be obtained by interaction of polar core interactions with the outer residues.

1. Next to Asn, Lys at position **a** is the most common buried polar residue in natural dimeric coiled coils ***(19)***. Lys at position **a** can form an *intra*helical electrostatic interaction with an **e** position residue of the preceding heptad ***(27)*** as well as an *inter*helical **g′–a** polar interaction with a **g′** position polar residue of the preceding heptad of the opposing helix in a parallel dimer ***(26)***.

2. Campbell and Lumb placed two position **a** Lys residues into the context of the Base-pLL peptide of Peptide Velcro to enable interhelical polar interactions between these Lys residues and the **e′** or **g′** Glu residues of Acid-pLL of Peptide Velcro (*see* **Note 14**; **ref.** *38*). As expected, the dimeric state was favored, most likely because the desolvation penalty would be higher in a higher oligomeric state. In addition, such an interaction is less destabilizing than an Asn–Asn (**a–a′** contact), presumably because there is a greater desolvation penalty to pay for burying the latter. However, no discrimination between parallel and antiparallel arrangement occurred, presumably because the **a**–**g′** interactions in the parallel orientation and the **a**–**e′** interactions in the antiparallel orientation were energetically similar.
3. Harbury's group used a computational design approach (*see* **Subheading 2.5.4.**, **item 4**) to generate specificity by considering not only the desired structure but also alternate undesired structures *(39)*. The **a**, **d**, **e**, and **g** positions of the central heptad of GCN4 (*see* **Note 15**) were varied; all nonproline residues and specific homotypic and heterotypic sequences were selected *in silico* and experimentally verified. Next to volume complementarity and charge complementation at **g/e′** pairs, it was observed that Glu at **d** paired preferentially with an Arg residue at position **e′**.

2.2.3. Edge Residues

Pairing specificity is greatly influenced by the nature of the electrostatic **e** and **g** residues (in parallel dimeric helices between **g** of one heptad and **e′** of the following heptad on the other helix; this is termed $i \rightarrow i' + 5$). These residues are commonly found to be Glu and Lys, respectively. Such complementary polar interactions add specificity and consolidate the stability provided by the core hydrophobic interactions. The charge pattern on the outer contacting edges of a coiled coil will dictate whether the protein will form a homomeric or heteromeric protein, and whether the orientation of the coiled coil is to be parallel or antiparallel. However, the PV hypothesis (*see* **Heading 1**; **ref.** *8*) is an oversimplification; residues that serve little or no role in the stability of the complex serve their purpose by directing the helices away from homologs or undesirable interactions that would otherwise compromise the specificity of the molecule (negative design). Alternatively, some coiled coils are likely to have no evolutionary pressure because they are already specific and need not be any more stable than they already are.

1. As expected, replacing favorable **g/e′** Gln–Gln pairings with repulsive Glu–Glu pairs has been shown to destabilize the coiled coil conformation *(40)*.
2. Hodges and co-workers estimated the salt bridges between **g/e′** pairs to contribute 0.37 kcal/mol to the stability of the coiled coil (*see* **Note 16**; **ref.** *41*).
3. Careful placing of charges within the **e** and **g** positions can permit heterodimer formation, and additionally ensure that formation of the homodimer is unfavorable *(9,42,43)*, as implicated in the PV hypothesis.

Table 4
Coupling Energies ($\Delta\Delta\Delta G_{int}$) for g-e′ Pairings Calculated Using a Double Mutant Alanine Thermodynamic Cycle

g–e	Glu	Gln	Arg	Lys
Glu	+0.7 ± 0.2	+0.2 ± 0.1	–0.5 ± 0.1	–0.3 ± 0.15
Gln	+0.2 ± 0.1	0.0 ± 0.1	+0.3 ± 0.1	+0.3 ± 0.1
Arg	–1.1 ± 0.1	+0.4 ± 0.1	+0.8 ± 0.1	+0.8 ± 0.1
Lys	–0.9 ± 0.1	+0.3 ± 0.1	+0.6 ± 0.1	+0.6 ± 0.1

From **ref.** *44*.
Values in kcal/mol.

4. The four most common amino acids found at the **e** and **g** positions are Glu, Gln, Arg, and Lys *(44)*. These residues contain long hydrophobic side chains that are able to interact with **a** and **d** core residues, and terminate with a charged (Glu, Arg, or Lys) or polar (Gln) group. By mutating first **e**, then **g**, then both residues of two interacting heptad pairs to Ala, Krylov et al. were able to establish the coupling energies ($\Delta\Delta\Delta G_{int}$) of those contacting residues for chicken vitellogenin-binding protein *(44,45)* (*see* **Note 17**). This double mutant thermodynamic cycle was used to permit a thermodynamic scale to be generated for outer residue contact preferences (**Table 4**). At 150 m*M* KCl and pH 7.4, Glu–Arg attractions are found to be slightly more stable than Glu–Lys attractions, presumably because Arg side chains are longer and interact better with the glutamate, and the respective methylene groups shielding the core more effectively from the solvent. This may, in turn, increase the effect of the charged end groups, which give a greater contribution in less aqueous surroundings. As expected, high salt weakens these interactions, as does low pH, in which polar interactions are weakened and the hydrophobic effect is increased. The Glu–Arg interaction, followed by Glu–Lys and Gln–Gln, are the most stabilizing (regardless of orientation), with the Glu–Glu and Arg–Arg, Arg–Lys, Lys–Arg, and Lys–Lys same-charge interactions being considerably less favored.
5. Arndt et al. designed a peptide library based on the Jun-Fos heterodimer, in which the **b**, **c**, and **f** residues are from their respective wild-type proteins, the **a** and **d** positions are Val and Leu (with the exception of **a**3 Asn inserts in the core to direct desired helix orientation and oligomerization state), and the **e** and **g** residues are varied using trinucleotides to yield equimolar mixtures of Arg, Lys, Gln, and Glu (*see* **Note 7**; **ref.** *8*). Unexpectedly, even the best-selected winner, the Winzip-A2B1 heterodimer (*see* **Note 18**), lacked fully complementary charged residues at **g/e′** pairs, despite an exhaustive selection process. Rather, two of the six **g/e′** pairs are predicted to be repulsive, suggesting that sequence solutions deviate from simple charge complementarity rules (PV hypothesis). Presumably, the overall electrostatic potential (including intramolecular and intermolecular interactions) plays a major role, and interactions with core residues, such as favorable packing or steric clashes could also modulate these **g/e′** interactions (*see* **refs.** *8* and *31* and references therein). Such observations are in agreement with naturally occurring coiled coils, which

usually have a complicated interaction pattern. These coiled coils have to fulfill a number of criteria, such as biostability and extremely high specificity within a family and almost no crossreactivity with coiled coils of other families *(46)*.

2.3. Helix Orientation

The majority of coiled coils fold into a parallel alignment, however, a growing number of structurally characterized proteins contain antiparallel coiled coil domains *(47)*. Despite the growing recognition of the biological importance of antiparallel coiled coils, the study of this class of molecules has been hampered by the lack of well-behaved model systems. None of the short antiparallel coiled coil domains found in proteins such as seryl–transfer RNA (tRNA) synthetase or hepatitis delta-virus antigen have been shown to be sufficient for dimerization without undergoing further self-association.

Hodges and co-workers were the first to report the characterization of a *de novo* designed coiled coil that was constrained in an antiparallel orientation by an interior disulfide bond *(48)*. This and other designed antiparallel coiled coils were more stable than their respective parallel counterparts, with nearly equivalent interhelical interactions *(49,50)*. These data suggest that, assuming everything is equal, the helix–dipole interactions (*see* **Subheading 2.4.3.**) favor the antiparallel orientation.

2.3.1. Core Residues

The core residues that pack against each other are **a**–**a′** and **d**–**d′** in parallel coiled coils. Antiparallel coiled coils have **a**–**d′** and **d**–**a′** central packing, yielding identical packing layers *(51)*.

1. It has been shown that the relative position of Ala residues in the core of a *de novo* designed coiled coil can control the parallel or antiparallel orientation (*see* **Subheading 2.1.1.**, **item 3**; **ref.** *52*). By careful placement of the Ala in the middle heptad, either all-parallel or all-antiparallel tetramers are formed. This was achieved by an alternating pair of Ala and Leu residues (Ala–Leu–Ala–Leu) in each of the two planes at the central heptad core positions of the molecule. Such an alternating core of small and large residues (Ala–Leu–Ala–Leu) is the best way to accommodate these small side chains. In the parallel arrangement, the packing would be all Ala in one plane and all Leu in the other plane and would, thus, result in a large cavity that would solvate the core and destabilize the molecule (*see* **Note 19**).
2. In a similar experiment, the core Asn of the dimeric GCN4-p1 (*see* **Note 2**) was exchanged to Ala. The result was an antiparallel trimer to avoid a core cavity *(53)*. However, parallel trimers were obtained in the presence of benzene, which bound to the core cavity *(54)*.
3. In a recent design of an antiparallel homodimeric coiled coil, termed APH, steric matching of β-branched (Ile at position **d**) and truncated (Ala at the opposing **a′** position) side chains in the hydrophobic core were used, along with other features

(*see also* **Subheading 2.3.3.**, **item 3**), to promote antiparallel orientation *(55)*. The parallel arrangement should be energetically disfavored by an Ile **d** layer, which is poorly accommodated in dimeric coiled coils (*see* **Subheading 2.1.1.**) and by an Ala–Ala "hole" in the hydrophobic core (*see* **Note 20**).

2.3.2. Polar Core Residues

Similar to the parallel coiled coils, buried polar residues also play a key role in dictating the helix orientation.

1. In parallel coiled coils, **a–a′** Asn pairing is commonly observed (*see* **Subheading 2.1.2.**), however, this has not been seen for the corresponding **a–d′** interaction in an antiparallel coiled coil (for a review, *see* **ref.** *47*). Indeed, Leu–Leu packing interactions found in the antiparallel coiled coils are more stable than in parallel core-packing interactions *(49)*, and it is likely to be the Asn–Asn interactions and electrostatic interactions (*see* **Subheadings 2.2.3.** and **2.3.3.**) that drive the specificity of the parallel conformation over the antiparallel.
2. Despite no native identifications of Asn–Asn **a–d′** pairs, Oakley and Kim modified the parallel heterodimeric coiled coil Peptide Velcro such that the buried polar interaction was only expected to occur if the helices are in an antiparallel orientation (i.e., **a–d′** pairings; **ref.** *56*). It was estimated that this single buried polar interaction conferred a modest antiparallel preference of approx 2.3 kcal/mol (*see* **Note 21**). However, an exclusive formation of antiparallel coiled coils was obtained only by combining the buried polar interaction with the introduction of charge repulsions in the parallel orientation (*see* **Subheading 2.3.3.**, **item 2**).
3. Comparable to the **a**–**g′** or **d**–**e′** interactions found in parallel dimeric coiled coils (*see* **Subheading 2.2.2.**), **a**–**e′** or **d**–**g′** interactions can occur in antiparallel coiled coils, as, for example, observed in the seryl tRNA synthetase coiled coil between Arg-54 at a **d** position on one strand and Glu-74 at a **g′** position in the other *(57)*, or in an **a**–**e′** interaction in the GreA coiled coil *(58)*. It was assumed that this type of buried polar interaction might also play a role in orientation specificity. Oakley's group replaced the core Asn pair in Peptide Velcro, such that an **a** position Arg in Acid-p1 could interact with a **g′** position Glu in Base-p1 in the antiparallel orientation (*see* **Note 22**; **ref.** *59*). The parallel conformation should be destabilized by a potentially repulsive interaction between an **e′** position Lys. However, while the introduction of Arg in the core was able to promote the dimeric state, in accordance with the study of Campbell and Lumb (*see* **Subheading 2.2.2.**, **item 2** and **Note 14**; **ref.** *38*), no clear preference for the antiparallel or parallel orientation was found. The free energy difference between both states was estimated to be only 0.1 ± 0.1 kcal/mol. One possible explanation could be the formation of an interhelical interaction with a neighboring **g** position Glu.

2.3.3. Edge Residues

In parallel coiled coils, the polar **e** and **g′** positions interact favorably. In the antiparallel coiled coil, **e** interacts with **e′**, and **g** with **g′**, the result being that

one helix is effectively rotated by 180°. These changes in preference must play an important role in orientation selection.

1. Monera et al. designed parallel and antiparallel coiled coils predicted to have either interchain attractions or repulsions (*see* **Note 23**; **ref.** ***49***). It was indeed found that the major orientation found was the one resulting in electrostatic interactions between oppositely charged amino acids.
2. Oakley's group refined a previously designed antiparallel heterodimeric coiled coil (*see* **Subheading 2.3.2.**, **item 2**) further. In their first design ***(56)***, one peptide contained only Glu residues, the other only Lys at both the **e** and **g** positions (*see* **Note 21**). In their new design, they substituted a single residue at a **g** position in each peptide such that all potentially attractive interactions are expected in the antiparallel orientation (*see* **Note 24**; **ref.** ***60***). In contrast, two potentially repulsive Coulombic interactions are expected in the parallel orientation, and indeed, a strong preference for the antiparallel arrangement was found.
3. In the recently designed antiparallel homodimeric coiled coil, APH (*see also* **Subheading 2.3.1.**, **item 3**), Glu residues at the N-terminal **e** and **g** positions and Lys at the C-terminal **e** and **g** positions have been used to direct antiparallel orientation, resulting in eight potential Coulombic interactions in the antiparallel arrangement and eight potential repulsions in the parallel arrangement (*see* **Note 20**; **ref.** ***55***).

2.4. Stability

Achieving a favorable stability is the net result of large and opposing enthalpic and entropic forces. The result is a protein of modest stability that has evolved to interact with its partner protein under physiological conditions, but not to be so stable as to never dissociate. Achieving this balance is, once again, the result of core and edge residues and their interactions, helix length, helical propensity, solubility, and helical capping. These stabilizing factors are discussed in this section.

One important and delicate consideration is solubility. The participating helices must have a nonpolar core to permit a favorable interaction, but the overall helices must not be so nonpolar as to aggregate under working conditions. Residues **a** and **d** must form a hydrophobic strip that connects the two helices, and **e** and **g** should typically be polar residues involved in ensuring that only the true binding partners interact with the helix (i.e., to be destabilizing for noninteracting partners), and consolidating the stability introduced from the core. This leaves the remaining residues in the heptad, the **b**, **c**, and **f** positions, to address the charge balance and to ensure that the helix is both stable and soluble. Glu and Lys, also of reasonable helical propensity, are well suited. Charged residues may also interact favorably with the helix dipole (*see* **Subheading 2.4.3.**) and form favorable interactions with charged residues one turn away in the helix. This gives an additional advantage in the selection of these residues at solvent exposed sites away from the dimer interface. The insertion of a Tyr at a solvent-exposed position is advantageous for concentration

determination of the peptides. Crudely speaking, for the archetypal dimeric coiled coils, such as GCN4 and Jun proteins, there is a preponderance of Lys, Asp, Arg, Glu, and Asn at these positions. Also present, albeit infrequently, is Ala, which presumably adds extra stability to the helix, perhaps in cases in which residues in close spatial proximity have caused the overall α-helical propensity to lower.

2.4.1. Helical Length

Generally speaking, as the length of the coiled coil chain increases, a (nonlinear) increase in stability is observed ***(61)***. This is because the sequence of the coiled coil will play an additional major role. For example, Lau and Hodges constructed a 29-mer with greater stability than tropomyosin, a 284-residue coiled coil (*see* **Note 25**; **ref. *62***). One would expect the coiled coil structure to become more stable with length on the grounds of increased hydrophobic burial, hydrogen bonds, and polar interactions, and, again, this was proven to be the case in a length study of a designed homodimeric peptide (*see* **Note 26**; **ref. *63***), although a minimum length of two heptads (*see* **Note 27**; **ref. *64***) is also required to permit association in the case of a homotrimer. The coiled coil domain of the Lac repressor has been used as the basis to assess the effects of chain length on stability and folding ***(65)***. Unsurprisingly, the dissociation constant of the tetramers decreased as the number of heptads in each helix increased (*see* **Note 28**). Somewhat more surprising was the fact that a tetramer with as few as four heptads in each helix folded cooperatively, with no evidence for a dimeric intermediate.

Long coiled coils, such as tropomyosin and myosin heavy-chain domain are, in contrast to short coiled coils, not enriched in the core exclusively with bulky nonpolar amino acids. Rather, because of the stability afforded from the length of the protein, such proteins contain clusters of small nonpolar or charged residues ***(66)***. These residues account for approx 40% of the core. Such destabilizing clusters consist mainly of Ala, because this residue is less destabilizing to the core than a polar amino acid (being easier to pack than a polar side chain), and is able to contribute favorably to the overall helical propensity. This means that although a stability increase proportional to the gain in heptads is not observed in these long coiled coils, from an evolutionary point of view, such a vast gain in stability is not required. Instead, hydrophobic stabilizing clusters afford the coiled coil the necessary stability, whereas destabilizing clusters do not, but do maintain the helical structure. These clusters, predominant in Ala, increase the flexibility and local unfolding in such regions without affecting the overall stability of the coiled coil, presumably allowing the protein to exercise its specific biological function.

2.4.2. Helix Propensity

Litowski and Hodges have reported that increasing the α-helical propensity of noncore residues by exchanging Ser for Ala (the amino acid of highest helical

propensity), can stabilize the whole coiled coil *(67)*. In their model, the Glu/Lys coiled coil (*see* **Note 29**), this led to stabilizations of approx 0.4 to 0.5 kcal/mol per substitution. This is in good agreement with an earlier study by O'Neil and DeGrado *(68)*. This number is less than for single α-helices, presumably because of the additional stabilizing interactions involved in maintaining the coiled coil.

The outer residues, **b**, **c**, and **f** should, generally speaking, be able to form hydrogen bonds with the solvent, and, in doing so, compensate for those buried potential hydrogen bonding partners that are unable to do so *(69)*. These residues are also responsible for helping to maintain the α-helical propensity, with each residue having a distinct conformational preference that will either stabilize or destabilize the helix *(68)*. The α-helical propensity of Ala is known to be the highest of all amino acids, and, although all common core hydrophobic residues (with the exception of Asn) have good α-helical propensities, the solvent-exposed Arg and Lys also play a considerable role. This is because both Lys and Arg make good hydrogen-bonding partners and are well-suited to improving the solubility of the molecule. Indeed we found helix propensity to be an important factor in coiled coil design (*70*, http://www.molbiotech.uni-freiburgode/bCIPA).

2.4.3. Interactions With the Helix Macrodipole

Statistical analysis of the composition of α-helices in protein structures revealed that different amino acids prefer different regions of the helix. In particular, potentially negatively charged side chains (Asp or Glu) strongly prefer positions near the N-terminal end of helices, whereas potentially positively charged side chains (His, Arg, or Lys) had a less pronounced preference for the C-terminal end *(71,72)*. Explanations for the preferences of polar side chains for the ends of helices fall into two principal models. First, because the first four backbone NH groups and final four backbone CO groups of the α-helix are not able to form the $i \rightarrow i + 4$ hydrogen bond to other backbone groups, polar side chains at the end of helices can substitute as hydrogen bonding partner. This is termed helix capping and is described in **Subheading 2.4.4**. Second, electrostatic "charge–helix dipole" interactions between the charged side chains and the net dipole moment of the α-helix formed by the alignment of individual peptide backbone dipoles may also stabilize or destabilize the protein.

1. Hodges' group investigated the positional dependence of negatively charged Glu side chains on the stability of a designed homodimeric coiled coil with no intrahelical or interhelical interactions (*see* **Note 30**; **ref.** *73*). A Glu substituted for Gln near the N-terminus in each chain of the coiled coil stabilized the coiled coil at pH 7.0, consistent with the charge–helix dipole interaction model. In contrast, Glu substitution in the middle of the helix destabilized the coiled coil because of the lower helical propensity and hydrophobicity of Glu compared with Gln at pH 7.0. A Glu substitution at the C-terminus destabilized the coiled coil even more, because of

the combined effects of intrinsic destabilization and unfavorable charge–helix dipole interaction with the negative pole of the helix dipole.

2. During selection for heterodimeric coiled coils from two designed libraries with an equimolar mixture of Glu, Gln, Arg, and Lys residues at four **e** and four **g** positions (*see* **Note 7**), Arndt et al. observed an enrichment of negatively charged Glu and neutral Gln at the N-terminal part and positively charged Lys and Arg at the C-terminal part of selected coiled coil sequences *(31)*. This bias is in good agreement with the proposed charge–helix dipole interactions.

2.4.4. Helix Capping

A helix can be labeled as N′-Ncap-N1-N2-N3-N4-mid-C4-C3-C2-C1-Ccap-C′. Of these positions, N′, Ncap, Ccap, and C′ have nonhelical ψ and ϕ angles, and only N1-N2-N3...C3-C2-C1 participate in the $i \rightarrow i + 4$ hydrogen bonding that is characteristic of the α-helix. The N1, N2, N3, C1, C2, and C3 residues are unique because their amide groups participate in $i \rightarrow i + 4$ backbone–backbone hydrogen bonds using either only their CO (at the N terminus) or their NH (at the C terminus) groups (*see also* **Subheading 2.4.3.**). The need for these groups to form hydrogen bonds has powerful effects on helix structure and stability *(74)*. From N4 (or C4) upward, the residues can satisfy both NH and OH backbone hydrogen bonds.

1. In helix design, the most selective position for the stability at the N-terminus is the Ncap position. The six best amino acids for this position are Ser, Asp, Thr, Asn, Gly, and Pro, with another 11 (Val, Ile, Phe, Ala, Lys, Leu, Tyr, Arg, Glu, Met, and Gln) being strongly avoided *(75)*. Of the six preferred residues for Ncap, Ser, Asp, and Thr are the best.
2. A good example of an Ncap motif is Ser-Xaa-Xaa-Glu (Ncap-N1-N2-N3), in which a reciprocal side chain/main chain interaction pattern (OH of the Ncap Ser to NH of Glu, and the carboxyl group of Glu to NH of Ser) stabilizes the helix. Further stabilization can be achieved by hydrophobic residues before and after the Ser-Xaa-Xaa-Glu motif *(76)*. Lu et al. introduced this capping motif in the GCN4 sequence and were thereby able to stabilize the coiled coil by 1.2 kcal/mol (*see* **Note 31**; **ref.** *77*).
3. The N1 and C′ positions strongly favor Pro (it is sterically compatible because the proceeding residues have nonhelical dihedral backbone angles), which is indeed a common helix termination motif, but should be avoided both in the main body of the helix and in the C3, C2, and Ccap positions. Pro, being the most water soluble of all amino acids *(78)*, is compatible with solvent-exposed positions at the helix ends, and also requires no hydrogen bonding acceptor because it lacks a backbone NH group *(76)*.
4. C-terminal capping motifs involve backbone–backbone hydrogen bonds, rather than the side chain to backbone hydrogen bonding that is observed between Ncap and N3. At the C-terminus backbone, hydrogen bonds are satisfied by posthelical backbone groups (e.g., C′ and the following C″ in the Schellman motif *(76)*. This means that the C-terminus need only select for C′ residues that can adopt positive ϕ angles, for example, Gly.

2.5. Rational Design vs Selection Strategies

The complex nature of protein design means that confidence in designing or preselecting a sequence that is to be of greatest stability is a daunting task. Rather, if procedures permit, it can be more fruitful to generate a library of interactions; a collection of changes to the molecule that can then be assayed for the most successful interacting sequences. This can be regarded as a semi-rational approach, whereby balanced changes at specific positions are introduced into the library (and often also keeping wild-type residues) and the resulting large combination of molecules can be screened. Normally, in this type of design, to change all heptad positions to all amino acids would generate a large and unrealistic library size (e.g., a four heptad coiled coil with only two amino acid options at each position would generate $2^{28} = 2.68 \times 10^8$ library members). However, to generate only one sequence that is to be synthesized and tested for binding strength would be time consuming, cumbersome, and, often, disappointing. Using a semi-rational design, one is afforded a reasonable compromise, whereby a molecule can be included within a library to test whether it is indeed a good candidate for binding, and, at the same time, affording the versatility to generate new, unintuitive, and, often, novel binding partners. Selection in a cell-based system has the added bonus of concomitantly screening out sequences that are susceptible to proteases found within that host organism.

2.5.1. Degenerate Codons

By using degenerate codons, one is able to encode a mixture of amino acids at positions at which a change is desired. Again, by carefully choosing the corresponding nucleotide positions that are to be randomized, one is able to generate a codon that is degenerate, but only for the desired bases and, hence, desired amino acids. As has been discussed, the coiled coil has a preference for "types" of amino acids at various positions. For example, the **e** and **g** residues are commonly found to be polar but complementary (*see* **Table 4**), therefore, an Glu–Lys interaction could be exchanged for an Glu–Arg interaction with little perturbation to the structure of the molecule (because Lys and Arg are rather similar in terms of both bulk and charge). In this situation, Arg could be introduced and Lys could be retained in the library to observe which interaction is more favorable in context of the whole molecule. Likewise, the hydrophobic core's preference for β-branched amino acids at the **a** position means that although a Val may be found in the wild-type molecule, an Ile might be preferred. It should be noted, however, that adhering to these preferences is an oversimplification, and apparent non-favored residues are required to design against nonspecific structures or even overtly stable structures (*see* **ref.** *8*). Generally speaking, however, keeping the overall "binary patterning" will ensure that all resulting library members have a

reasonable chance of forming a coiled coil structure rather than alternate structures being energetically favored. The coiled coil library members will be able to bury their hydrophobic residues and surround them with the polar **e** and **g** residues, in the same way that the wild-type structure was able to. For a coiled coil, the maintenance of the hydrophobic core, termed the "3-4 repeat" (**a**-to-**d** and **d**-to-**a** is a spacing of three and four residues, respectively), is essential. Finally, it is possible to polymerize specific synthetic trinucleotides such that the final gene sequence codes for those, and only those, amino acid sequences that are desired in the final peptide *(79)*. This will also facilitate in keeping the library size low, allowing more desired changes to be screened. This is not true of degenerate codon usage using regular degenerate oligonucleotides, which, depending on the codon usage table, may imply that to have two desired amino acids, another undesirable amino acid must be included at that position (*see* **Subheading 2.5.2.** and **Fig. 2**). It is also possible that two amino acids from opposite ends of the codon table are desired, but realistically can only be included by using such presynthesized trinucleotide codons. Such trinucleotide building blocks have recently become commercially available (Glen Research, Sterling, VA; Metkinen Oy, Finland).

2.5.2. Codon Bias

Codon bias is the probability that one codon rather than another will code for a particular amino acid (**Fig. 2A**). For example, in the case of *Escherichia coli,* CGT is found to code for Arg approximately five times more frequently than CGA. Obviously, when transforming cells (e.g., *E. coli*) with plasmids containing a gene coding for library peptides, one must select triplets that are frequently observed in the host organism over infrequently used codons, which may lead to frame shifts *(80–82)* or poorly represented sequences in the library. In addition, when designing a library that is to be screened, one may, for example, wish to introduce all combinations of β-branched amino acids into the **a** position, but depending on the codon usage, this may or may not be possible. **Fig. 2** shows some of our most favored codon combinations for different positions in coiled coils. Most of the combinations include only a few options to keep the libraries to a reasonable size. By examining **Fig. 2** (the codon usage table for *E. coli*; *see* www.kazusa.or.jp/codon for this and other organisms), we see that to include Val, Ile, and Thr at an **a** position of the coiled coil would require the codons GTN, ATH (if Met should be avoided), or ACN. This would mean requiring G or A at position one of the codon, T or C at position two, and any nucleotide other than G at the third position, because ATG would code for Met, resulting in the degenerate codon RYH (all resulting codons in this case are used significantly by *E. coli*). For such a Val, Ile, and Thr combination, however, it is not possible to rule out amino acids that are not required in the library,

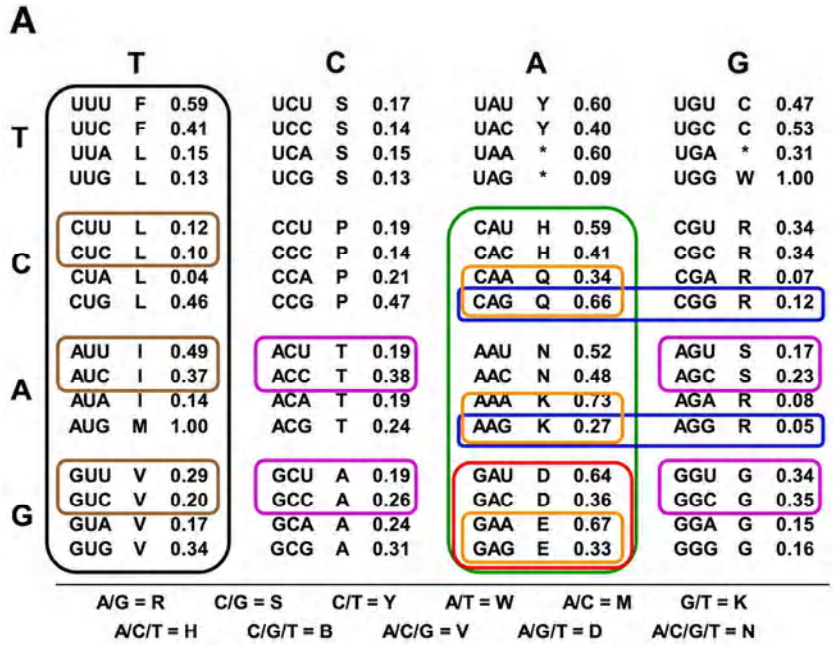

Fig. 2. Codons and possible combinations for the design of coiled coil libraries. Codons and their bias as represented in the *E. coli.* Shown is the standard textbook triplet code, followed by the amino acid represented and the fraction of this codon for this organism (http://www.kazusa.or.jp/codon; **ref.** *83*). Also shown are the letters that represent a mixture of oligonucleotides, and will subsequently lead to the degenerate base positions required in the generation of the library. The frames give examples of possible combinations for coiled-coil library design.

because a mixture of above mentioned oligonucleotides would also include GCH, which codes for Ala. Another example is that if one wanted to include the amino acids Gln, Asn, and Lys, one must also include His, by default. This is a problem that can be overcome by polymerizing synthetic trinucleotides that code only for those amino acids desired *(31,79)*. In addition, one must be careful to rule out codons that are poorly represented, e.g., when representing Arg, because amino-acyl tRNA synthetases for the AGG and AGA codons are in short supply for *E. coli*. Finally, it should be noted that if a position is to be represented without bias in the library, then a 1:1 ratio between all amino acids at that position (and likewise for the respective codons) should be selected. Representing an amino acid more than once over another at a degenerate position will naturally put a bias toward that change into the system, because it becomes overrepresented.

codon	NNK	NTN	NTK	DTK	VTY	RSY	VAN	VAR	GAN	MRG	
Gly	2	–	–	–	–	2	–	–	–	–	aliphatic side chains
Ala	2	–	–	–	–	2	–	–	–	–	
Val	2	4	2	2	2	–	–	–	–	–	
Leu	3	6	3	1	2	–	–	–	–	–	
Ile	1	3	1	1	2	–	–	–	–	–	
Met	1	1	1	1	–	–	–	–	–	–	
Pro	2	–	–	–	–	–	–	–	–	–	
Phe	1	2	1	1	–	–	–	–	–	–	
Trp	1	–	–	–	–	–	–	–	–	–	
Ser	3	–	–	–	–	2	–	–	–	–	polar side chains
Thr	2	–	–	–	–	2	–	–	–	–	
Asn	1	–	–	–	–	–	2	–	–	–	
Gln	1	–	–	–	–	–	2	2	–	1	
Tyr	1	–	–	–	–	–	–	–	–	–	
Cys	1	–	–	–	–	–	–	–	–	–	
Lys	1	–	–	–	–	–	2	2	–	1	charged side chains
Arg	3	–	–	–	–	–	–	–	–	2	
His	1	–	–	–	–	–	2	–	–	–	
Asp	1	–	–	–	–	–	2	–	2	–	
Glu	1	–	–	–	–	–	2	2	2	–	
stop (TAG)	1	–	–	–	–	–	–	–	–	–	
total	32	16	8	6	3	4	6	3	2	4	

Fig. 3. Listed are amino acid frequencies corresponding to the codon randomization scheme shown in **Fig. 2**. Additionally, the distribution for NNK is given. This combination is widely used when all 20 amino acids should be included and stop codons should be minimized. "Total" gives the number of amino acids to be used when calculating the library size. Hydrophobic amino acid combinations encoded by NTN (black line in **Fig. 2**) or VTY (gray long dashes), respectively. To minimize overrepresentation of Ile and Leu, NTN can be reduced to NTK or DTK. Different polar and charged combinations are encoded by VAN (gray line), VAR (black long dashes), GAN (black dots), or MRG (gray short dashes), respectively. Possible loop regions, if desired, can be obtained by the codon RSY (black short dashes). Please note that, in our examples, the third base is also mixed to increase variety in codon usage, which might be important for repetitive sequences. However, if desired, the last base can be kept constant in most instances. This is especially important in the case of the Gln, Arg, and Lys mixture (short gray dashes), in which the codon MRG yields only one rare Arg codon (AGG), whereas the combination MRR yields three rare Arg codons (CGG, AGA, AGG). Combinations NTN and VAN have been applied previously in the binary design pattern developed in the Hecht group (*see* **refs.** ***84*** and ***85*** and Chapter 9).

2.5.3. Selection Systems

The most common selection system open to the coiled coil is the protein-fragment complementation assay selection. In this assay, interacting proteins, e.g., two coiled coil fragments, are tethered to two halves of a reporter protein that only becomes active after association of the two fused proteins or peptides. This has been used for dihydrofolate reductase ***(8,31,70,86–88)***, ubiquitin ***(89)***,

β-galactosidase *(90)*, β-lactamase *(91,92)*, and green fluorescent protein *(93)*. Such intracellular assays have the additional benefit of concomitantly selecting against protease susceptible or toxic peptides.

Another selection system for coiled coil interactions is the yeast two-hybrid system. In this system, one helix is fused to the binding domain (which binds to a promoter) and the other to the activating domain (which interacts with the polymerase) of the Gal4p transcription factor. Only interaction between the two helices will bring the two chimeric proteins into close proximity and permit the transcriptional activator to function, thus, switching on a reporter gene (e.g., β-galactosidase) by which the experiment can then be assayed for activity. Furthermore, novel interacting partners can be found by screening a single protein or domain against a library of other proteins using this system (for a recent review, *see* **ref. *94***).

The λ repressor system is based on reconstituting the activity of the *E. coli* λ repressor by replacing the C-terminal domain with a heterologous oligomerization domain. The interaction is detected when the C-terminal domain forms a dimer (or higher-order oligomer) with itself (homotypic interaction) or with a different domain from another fusion (heterotypic interaction). Functional repressors can be detected either by monitoring immunity to phage infection or by assaying repression of reporters, e.g., β-galactosidase *(95,96)*.

Finally, using phage display, proteins can be displayed on the surface coat proteins of filamentous bacteriophage. Peptide libraries displayed on the phage can be selected for binding and enriched by several rounds of selection. This technique also permits selection for binding to nonnatural compounds *(97)*.

2.5.4. Calculations Concerning Designs

Instead of using in vivo selection systems, some groups also used an *in silico* approach to generate promising sequences for coiled coil design. Some aspects will be discussed in this section.

1. Keating et al. developed a computational method to predict hydrophobic core mutation effects on interaction specificity *(98)*. This is achieved using an algorithm that considers van der Waals packing interaction energies, a solvation term, and α-helical propensities. From this, partnering preferences arising from core packing was predicted. Coiled coils were designed with a core **a** and **d** residue mutated to Leu, Ile, and Val to yield six different heterodimers with a range of stabilities (*see* **Note 32**). The algorithm was able to predict stability with good agreement to the experimental data.
2. Mayo's group used a design algorithm to assess the surface position interactions. They tested three scoring functions: a hydrogen-bond potential, a hydrogen-bond potential in conjunction with a penalty for uncompensated burial of polar hydrogens, and a hydrogen-bond potential in combination with helix propensity *(69)*.

The algorithm was used to find the optimal amino acid sequence for each of the three scoring functions, using GCN4-p1 (*see* **Note 2**), and the corresponding peptides were consequently synthesized. All resultant peptides were dimeric, close to 100% helical at 1°C and had melting temperatures of 69°C to 72°C compared with 15°C for a GCN4 peptide with random hydrophilic surface residues. The data suggest that helix propensity is the key factor in sequence design for the surface residues of the coiled coil peptides.

3. For the generation of coiled coils with a right-handed superhelical twist, an algorithm was developed that incorporated main-chain flexibility *(99)*. For main-chain fixing to be used, a naturally existing example is needed to supply the coordinates, and even then, these values may not be valid for close homology-adjusted sequences. By allowing backbone flexibility, a small subset of main-chain conformations could be sampled. These samplings were then coupled with side-chain packing sampling. Consequently, right-handed dimeric, trimeric, and tetrameric-coiled coils were computationally designed (*see* **Note 33**). The overall protein fold was specified by hydrophobic polar residue patterning. The oligomerization state, main-chain conformation, and side-chain rotamers were computationally selected by best packing in alternate backbone structures. The resulting designed peptides formed the correct oligomeric state ensembles in accord with the design goals, and the X-ray structure of the tetramer matched the design structure in atomic detail.
4. A paper by Harbury used a combination of positive (toward the desired structure) and negative design (away from undesired alternate structures) to optimize interaction specificity (*see* **Subheading 2.2.2., item 3**; **ref.** *39*). The **a**, **d**, **e**, and **g** positions of the central heptad of GCN4 were varied with all nonproline residues to generate approx 8×10^9 possible sequences (*see* **Note 15**). The algorithm used was akin to a computational double mutant cycle in which peptides are sequence optimized rather than structurally optimized. Competing energetic states considered were:
   a. The folded homodimeric desired structure.
   b. Energetically favored heterodimers.
   c. Unfolded energetically destabilized states.
   d. Aggregated states of poor solubility.

   The free energies were evaluated in each of these four competing states for candidate sequences. This so-called "multistate" design has an advantage over single state design systems, which only look for the lowest free energy states of the target. Unlike these designs, the multistate design structures often deviate from the PV hypothesis, and, at the same time, take into consideration factors such as too much hydrophobic exposure causing aggregation and too much polar burial being destabilizing. Rather, all of these factors are considered in the design, which only requires change when these competing forces dominate over the selection of the target state. Selection is, therefore, the result of a balance between stability and specificity, and not of target stability alone.

3. Notes

1. Sequence of GCN4-p1: Ac-R MKQ**L**EDK **V**EE**L**LSK **N**YH**L**ENE **V**AR**L**KKL **V**GER-COOH. In boldface are the **a** and **d** residues that have been mutated in the studies by Harbury et al. *(16,17)*. Underlined is the core Asn16, which has been subjected to many mutations.
2. The designed heterodimeric Peptide Velcro *(9)* consists of the synthetic peptides Acid-p1 (Ac-AQLEKE LQALEKE NAQLEWE LQALEKE LAQ-NH_2) and Base-p1 (Ac-AQLKKK LQALKKK NAQLKWK LQALKKK LAQ-NH_2). In this and the following notes, the individual heptads of the respective sequence are separated by a space. The core Asn residue, which has been mutated as described, e.g., in **Subheading 2.1.2.**, **item 4**, is underlined.
3. The sequence of the two antiparallel cysteine disulfide-bridged peptides 2H (Ac-K **C**EALEGK LEALEGK LEAAEGK LEALEGK LEALEG-NH_2 and Ac-E LAELKGE LAELKGE AAELKGE LAELKGE LAE**C**KG-NH_2) and 4H (Ac-K **C**EALEGK LEALEGK LEAAEGK LEALEGK LEALEG-NH_2 and Ac-E LAELKGE LAELKGE LAEAKGE LAELKGE LAE**C**KG-NH_2). Cysteines are shown in bold to indicate the point of covalent bonding between the two chains in 2H and 4H, respectively, and the Ala residue that specifies the oligomerization state is underlined *(20)*.
4. Sequence of the coiled coil domain (amino acids 27–72) of the rat COMP protein *(21)*: GDL APQMLRE LQETNAA LQDVREL LRQQVKE ITFLKNT VMEC-DAC G. In the expressed fragment of rat COMP, Gly 27 was replaced by Met.
5. Studies were based on the peptide A1 (MRGSHHHHHHGSMA SGD**L**ENE YAQ**L**ERE VRS**L**EDE AAE**L**EQK VSR**L**KNE IED**L**AEI GD**L**NNTSGIRRPAA K**L**N). Incorporation of trifluoroleucine and hexafluoroleucine for Leucine (in boldface) was achieved by induction of gene expression in leucine-free culture media supplemented either with trifluoroleucine or with hexafluoroleucine *(22,23)*.
6. Kretsinger et al. based their study on the GCN4-p1 peptide as given in **Note 2** with the differences that the C-terminus was amidated, and that they reported a sequence with an additional Ser between the Asp and Lys in the first full heptad *(24)*. Because this addition would shift the heptad repeat, we presume that it is an error in the figure. The underlined Asn (*see* **Note 2**) was subjected to exchange to Asp dap as well as monomethylated, dimethylated, and trimethylated analogs of dap.
7. In this selection, heterodimers were selected from two designed coiled coil libraries: LibA: VAQL#E# VKTL#A# §YEL#S# VQRL#E# VAQL and LibB: VDEL#A# VDQL#D# §YAL#T# VAQL#K# VEKL, where # denotes an equimolar mixture of E, Q, K, and R, and § denotes an equimolar mixture of V and N *(31)*. The sequences for the core **a** and **d** positions were taken from GCN4 (*see* **Note 2**) and for the **b**, **c**, and **f** positions (underlined) were taken from the coiled coil domains of c-Jun (IARLEEK VKTLKAQ NYELAST ANMLREQ VAQL) and c-Fos (TDTLQAE TDQLEDE KYALQTE IANLLKE KEKL), respectively.
8. The sequence of the GCN4-pVL variant is Ac-R **M**KQ**L**EDK **V**EE**#**LSK **§**YH**L**ENE **V**AR**L**KKL **V**GER, where the **a** and **d** positions are in boldface.

Position 12(**d**), indicated by #, is either a Leu or a polar residue (N, Q, S, or T), and position 16(**a**), indicated by §, is either a Val or a polar residue *(25)*.

9. Both studies used a disulfide-bridged coiled coil based on the model sequence VGALKKE, with some modification to avoid intrachain and interchain charge–charge interactions with the site of substitution (**X**) and to adjust the overall charge. The sequences were Ac-CGGE VGALKAQ VGALQAQ **X**GALQKE VGALKKE VGALKK-NH_2 *(33)* and Ac-CGGE VGALKAE VGALKAQ IGA**X**-QKQ IGALQKE VGALKK-NH_2 *(32)*, respectively.
10. Ji et al. used a recombinant model of the simian immunodeficiency virus gp41 core, designated N36(L6)C34, where the amino-terminal helices (N36) form a central, trimeric coiled coil, while the carboxyl-terminal helices (C34) pack in an antiparallel orientation into hydrophobic grooves on the surface of this coiled coil trimer. N36 and C34 are separated by a short linker (L6; **ref.** *34*). Mutations of polar core residues (in boldface) to Ile were made in the N36 domain: AGIVQQ **Q**QQLLDV VKR**Q**QEL LRL**T**VWG **T**KNLQTR VT. The Q → I mutation that formed insoluble aggregates is underlined.
11. The sequence of the parent peptide, Lac21 is: Ac-MKQLADS LMQLARQ VSR-LESA-NH_2 (*see also* **Note 28**). Underlined residues were mutated to E or K to result in the peptides Lac21E and Lac21K, respectively, which formed a heterotetramer *(35)*.
12. Sequence of APC-55: AAAS YDQLLKQ VEALKME NSNLRQE LEDNSNH LTKLETE ASNMKEV LKQLQGS I and of anti-APCp1: **M**AAK **G**DQLKKE VEALE*Y*E NSNLR*K*K LED**H**K*K*K LTKLKTE **I**SNA*K**KM*** LKQL*Y*AS I *(36)*. Core changes in anti-APCp1 compared with APC-55 are marked in boldface; changes at the **e** and **g** positions are underlined; and changes to increase stability, to increase the net charge to facilitate purification, and to add chromophores are marked in italics.
13. The three peptides were T_9: Ac-R MKQLEKK **X**EELLSK **A**QQLEKE **A**AQLKKL VG-NH_2, T_{16}: Ac-R MKQLEKK **A**EELLSK **X**QQLEKE **A**AQLKKL VG-NH_2, T_{23}: Ac-R MKQLEKK **A**EELLSK **A**QQLEKE **X**AQLKKL VG-NH_2 *(37)*. Residues that were different in all three peptides are marked in boldface, **X** denotes the cyclohexylalanine residue.
14. Sequences were based on Acid-pLL and Base-pLL, which are identical to Acid-p1 and Base-p1 (*see* **Note 1**) but with the core Asn mutated to Leu (*see also* **Subheading 2.1.2., item 4**). Two L → K mutations (in boldface) were made to Base-pLL to yield Base-pK: Ac-AQLKKK LQALKKK **K**AQLKWK **K**QALKKK LAQ-NH_2 *(38)*.
15. Studies were based on an N-terminally capped variant of GCN4 (*see* **Note 31**; **ref.** *7*) with the Asn16 shifted by one heptad level to position 9, yielding the peptide p-CAP: **S** VK**E**LEDK **N**EELLS**X** **X**YH**XX**NE VARLKKL VGER. Changes compared with GCN4-p1 (*see* **Note 2**) are marked in boldface. **X** denotes positions allowed to vary in the design calculations *(39)*.
16. Studies were based on the designed homodimeric coiled coil EK: Ac-K CGALEKK LGALEKK AGALEKK LGALEKK LGALEK-NH_2. Three mutant coiled coils were made in which:

    a. Five Glu residues at **e** positions in EK (underlined) were mutated to Gln residues (peptide QK).

b. Five Lys residues at **g** positions (underlined) were mutated to Gln residues (peptide EQ).
c. Both were combined (peptide QQ).

Using a double-mutant cycle analysis, the energetic contribution of interhelical ionic attractions to coiled coil stability was calculated ***(41)***.

17. Sequence of parental vitellogenin-binding protein ER_{34}: ITIR AAFLEKE NTAL-RTE VAELRKE VGRCRNI VSKYETR YGPL. Underlined **e** and **g** residues are altered in subsequent peptides ***(45)***.
18. The best winner WinZip-A2B1 from the selection (*see* **Note 7**) consists of peptides WinZip-A2 (Ac-STT VAQLRER VKTLRAQ NYELESE VQRLREQ VAQL AS-NH_2) and WinZip-B1 (Ac-STS VDELQAE VDQLQDE NYALKTK VAQLRKK VEKL SE-NH_2) ***(8)***.
19. Studies were based on designed E- and K-peptides. E-Peptide: Ac-E LGALEKE LGALEKE **L**GA**L**EKE LGALEKE LGALEK-NH_2; K-Peptide: Ac-K LGALKEK LGALKEK **L**GA**L**KEK LGALKEK LGALKE-NH_2. Positions 16 or 19 (bold) were changed to Ala to create different Leu–Ala core combinations, and positions 2 or 33 (underlined) were exchanged to Cys to allow disulfide bridging in the parallel or antiparallel state ***(52)***.
20. Sequence of APH: MKQLEKE LKQLEKE LQA**I**EKQ LAQLQWK **A**QA*R*KKK LAQLKKK LQA ***(55)***. Sterically matched core residues (Ile and Ala) are shown in bold; designed Coulombic interactions between N-terminal glutamines and C-terminal lysines are underlined. In addition, a single Arg residue was incorporated at a **d** position (in italics) to promote dimer formation.
21. The resulting peptides were termed Acid-a1 (Ac-AQLEKE LQALEKE **L**AQLEWE **N**QALEKE LAQ-NH_2) and Base-a1 (Ac-AQLKKK LQA**N**KKK **L**AQLKWK LQALKKK LAQ-NH_2) ***(56)***. Changes compared with Acid-p1 and Base-p1 (*see* **Note 1**) are in boldface.
22. Peptides Acid-R**d**L (Ac-AQLEKE LQALEKE LAQ**R**EWE LQALEKE LAQ-NH_2) and Base-E**g**L (Ac-AQLKKK LQALKK**E** LAQLKWK LQALKKK LAQ-NH_2) were based on Acid-a1 and Base-a1 (*see* **Note 21**), with the designed polar interaction in boldface ***(59)***.
23. Five different peptides were synthesized with an N- or C-terminal Cys and a core Ala residue. Three peptides (C2A16, C33A16, and C33A19) were based on the heptad repeat, LEALEGK: Ac-K LEALEGK LEALEGK **L**EA**L**EGK LEALEGK LEA**L**EG-NH_2, with the Cys either at position 2 or 33 (underlined) and the Ala either at position 16 or 19 (bold). Two peptides (C33A16 and C33A19) were based on the heptad repeat, LAELKGE: Ac-E LAELKGE LAELKGE **L**AE**L**KGE LAELKGE LAE**C**KG-NH_2, with the Cys at position 33 and the Ala either at position 16 or 19 ***(49)***.
24. Peptide Acid-K**g** (Ac-AQLEKE LQALEK**K** LAQLEWE NQALEKE LAQ-NH_2) was based on Acid-a1 (*see* **Note 21**), and peptide Base-E**g** (Ac-AQLKKK LQANKK**E** LAQLKWK LQALKKK LAQ-NH_2) was based on Base-a1 with the changes marked in boldface.

25. Synthetic peptides with the following sequence were constructed: Ac-(K LEA-LEG)$_n$-K-NH$_2$, with n = 1 to 5 ***(62)*** and compared with carboxamidomethylated α-tropomyosin at cysteine 190 (CM-tropomyosin).
26. A series of polypeptides containing 9, 12, 16, 19, 23, 26, 30, 33, and 35 amino acid residues were designed with the sequence: Ac-E iealkae iealkae iealkae iealkae ieacka-NH$_2$ for the 35-mer peptide ***(63)***. Shorter peptides were comprised of the respective number of amino acids counted from the C-terminus.
27. The peptide Succ-DELERR IRELEAR IK-NH$_2$ was used in this study ***(64)***, Succ indicated the succinylated N-terminus.
28. Investigated peptides were Lac 21: Ac-MKQLADS LMQLARQ VSRLESA-NH$_2$, Lac 28: Ac-LMQLARQ MKQLADS LMQLARQ VSRLESA-NH$_2$, and Lac 35: Ac-LMQLARQ LMQLARQ MKQLADS LMQLARQ VSRLESA-NH$_2$ ***(65)***.
29. Studies were based on the mutants of the E/K heterodimer comprised of the E-peptide with the sequence Ac-(E **#§**ALEK)$_n$-NH$_2$ and the K-peptide with the sequence Ac-(K **#§**ALKE)$_n$-NH$_2$ with n = 3 or 4. **#** indicates I or V; **§** indicates A or S ***(67)***.
30. The peptide sequence was Ac-Q CGALQKQ VGALQKQ VGALQKQ VGALQKQ VGALQK-NH$_2$. Positions 1, 6, 15, 20, and 34 that were mutated to Gln are underlined ***(73)***.
31. This study worked with recombinant GCN4-pMSE peptide (**MS** VK**E**LEDK VEELLSK NYHLENE VARLKKL VGER). The capping motif is in boldface and mutations compared with GCN4-p1 (*see* **Note 2**) are underlined. Other peptides from this study were GCN4-pSE, which lacked the initiator methionine, and GCN4-pAA, in which the Ser and Glu of GCN4-pSE were mutated to Ala ***(77)***. Stabilities of GCN4-pAA were comparable to GCN4-p1, whereas the GCN4-pSE and GCN4-pMSE variants were stabilized by 0.5 kcal/mol and 1.2 kcal/mol, respectively, relative to GCN4-pAA. Thus, the hydrophobic contribution of the terminal Met residue is 0.7 kcal/mol.
32. Heterodimeric coiled coils, denoted GABH, for GCN4 (*see* **Note 2**) Acid/Base Heterodimer were designed with the acidic sequence A (Ac-E VKQLEAE VEE#ESE #WHLENE VARLEKE NAECEA-NH$_2$) and the basic sequence B (Ac-K VKQLKAK VEE#KSK #WHLKNK VARLKKK NAECKA-NH$_2$) ***(98)***. Positions **d**12 and **a**16 (**#**) were mutated to Val, Ile, and Leu, respectively, to yield the peptides A_{LL}, A_{IV}, and A_{LI}, and B_{LL} and B_{LV}.
33. Designed sequences were dimeric RH2 (Ac-AE **I**EQ**L**KKE**§**AYL **I**KK**L**KAE**K**LAE **I**KK**L**KQE**K**A-NH$_2$), trimeric RH3 (Ac-AE **#**EQ**#**KKE**I**AYL **#**KK**#**KAE**I**LAE**#**K K**#**KQE**I**A-NH$_2$), and tetrameric RH4 (Ac-AE **L**EQ**#**KKE**I**AYL **L**KK**#**KAE**I**L AE **L**KK**#**KQE**I**A-NH$_2$) ***(99)***. The hydrophobic residues (**a**, **d**, and **h**) of the undecatad repeat (**a** to **k**) are marked in boldface; **§** indicates norvaline; and **#** indicates alloisoleucine residues.

Acknowledgment

This work was funded in the Emmy Noether program of the Deutsche Forschungsgemeinschaft (grant Ar 373).

References

1. 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.
2. Lupas, A. (1996) Coiled coils: new structures and new functions. *Trends Biochem. Sci.* **21**, 375–382.
3. Landschulz, W. H., Johnson, P. F., and McKnight, S. L. (1988) The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. *Science* **240**, 1759–1764.
4. Burkhard, P., Stetefeld, J., and Strelkov, S. V. (2001) Coiled coils: a highly versatile protein folding motif. *Trends Cell Biol.* **11**, 82–88.
5. Kohn, W. D., Mant, C. T., and Hodges, R. S. (1997) Alpha-helical protein assembly motifs. *J. Biol. Chem.* **272**, 2583–2586.
6. 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.
7. DeLano, W. L. (2002) The PyMOL molecular graphics system. DeLano Scientific, San Carlos, CA; http://www.pymol.org.
8. Arndt, K. M., Pelletier, J. N., Müller, K. M., Plückthun, A., and Alber, T. (2002) Comparison of in vivo selection and rational design of heterodimeric coiled coils. *Structure* **10**, 1235–1248.
9. 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.
10. Mason, J. M. and Arndt, K. M. (2004) Coiled coil domains: stability, specificity, and biological implications. *Chem. Biochem.* **5**, 170–176.
11. Müller, K. M., Arndt, K. M., and Alber, T. (2000) Protein fusions to coiled-coil domains. *Methods Enzymol.* **328**, 261–282.
12. Arndt, K. M., Müller, K. M., and Plückthun, A. (2001) Helix-stabilized Fv (hsFv) antibody fragments: substituting the constant domains of a Fab fragment for a heterodimeric coiled-coil domain. *J. Mol. Biol.* **312**, 221–228.
13. Pack, P., Müller, K. M., Zahn, R., and Plückthun, A. (1995) Tetravalent miniantibodies with high avidity assembling in Escherichia coli. *J. Mol. Biol.* **246**, 28–34.
14. Naik, R. R., Kirkpatrick, S. M., and Stone, M. O. (2001) The thermostability of an alpha-helical coiled-coil protein and its potential use in sensor applications. *Biosens. Bioelectron.* **16**, 1051–1057.
15. Crick, F. H. S. (1953) The packing of α-helices: simple Coiled Coils. *Acta Crystallogr.* **6**, 689–697.
16. 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.
17. Harbury, P. B., Kim, P. S., and Alber, T. (1994) Crystal structure of an isoleucine-zipper trimer. *Nature* **371**, 80–83.
18. Betz, S. F., Bryson, J. W., and DeGrado, W. F. (1995) Native-like and structurally characterized designed alpha-helical bundles. *Curr. Opin. Struct. Biol.* **5**, 457–463.
19. Woolfson, D. N. and Alber, T. (1995) Predicting oligomerization states of coiled coils. *Protein Sci.* **4**, 1596–1607.

20. Monera, O. D., Sonnichsen, F. D., Hicks, L., Kay, C. M., and Hodges, R. S. (1996) The relative positions of alanine residues in the hydrophobic core control the formation of two-stranded or four-stranded alpha-helical coiled-coils. *Protein Eng.* **9**, 353–363.
21. Malashkevich, V. N., Kammerer, R. A., Efimov, V. P., Schulthess, T., and Engel, J. (1996) The crystal structure of a five-stranded coiled coil in COMP: a prototype ion channel? *Science* **274**, 761–765.
22. Tang, Y. and Tirrell, D. A. (2001) Biosynthesis of a highly stable coiled-coil protein containing hexafluoroleucine in an engineered bacterial host. *J. Am. Chem. Soc.* **123**, 11,089–11,090.
23. Tang, Y., Ghirlanda, G., Petka, W. A., Nakajima, T., DeGrado, W. F., and Tirrell, D. A. (2001) Fluorinated coiled-coil proteins prepared in vivo display enhanced thermal and chemical stability. *Angew. Chem. Int. Ed.* **40**, 1494–1496.
24. Kretsinger, J. K. and Schneider, J. P. (2003) Design and application of basic amino acids displaying enhanced hydrophobicity. *J. Am. Chem. Soc.* **125**, 7907–7913.
25. Akey, D. L., Malashkevich, V. N., and Kim, P. S. (2001) Buried polar residues in coiled-coil interfaces. *Biochemistry* **40**, 6352–6360.
26. Glover, J. N. and Harrison, S. C. (1995) Crystal structure of the heterodimeric bZIP transcription factor c-Fos-c-Jun bound to DNA. *Nature* **373**, 257–261.
27. Gonzalez, L., Jr., Woolfson, D. N., and Alber, T. (1996) Buried polar residues and structural specificity in the GCN4 leucine zipper. *Nat. Struct. Biol.* **3**, 1011–1018.
28. Junius, F. K., Mackay, J. P., Bubb, W. A., Jensen, S. A., Weiss, A. S., and King, G. F. (1995) Nuclear magnetic resonance characterization of the Jun leucine zipper domain: unusual properties of coiled-coil interfacial polar residues. *Biochemistry* **34**, 6164–6174.
29. Potekhin, S. A., Medvedkin, V. N., Kashparov, I. A., and Venyaminov, S. (1994) Synthesis and properties of the peptide corresponding to the mutant form of the leucine zipper of the transcriptional activator GCN4 from yeast. *Protein Eng.* **7**, 1097–1101.
30. Lumb, K. J. and Kim, P. S. (1995) A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled coil. *Biochemistry* **34**, 8642–8648.
31. Arndt, K. M., Pelletier, J. N., Müller, K. M., Alber, T., Michnick, S. W., and Plückthun, A. (2000) A heterodimeric coiled-coil peptide pair selected in vivo from a designed library-versus-library ensemble. *J. Mol. Biol.* **295**, 627–639.
32. Tripet, B., Wagschal, K., Lavigne, P., Mant, C. T., and Hodges, R. S. (2000) Effects of side-chain characteristics on stability and oligomerization state of a de novo-designed model coiled-coil: 20 amino acid substitutions in position "d". *J. Mol. Biol.* **300**, 377–402.
33. Wagschal, K., Tripet, B., Lavigne, P., Mant, C., and Hodges, R. S. (1999) The role of position a in determining the stability and oligomerization state of alpha-helical coiled coils: 20 amino acid stability coefficients in the hydrophobic core of proteins. *Protein Sci.* **8**, 2312–2329.

34. Ji, H., Bracken, C., and Lu, M. (2000) Buried polar interactions and conformational stability in the simian immunodeficiency virus (SIV) gp41 core. *Biochemistry* **39**, 676–685.
35. 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.
36. Sharma, V. A., Logan, J., King, D. S., White, R., and Alber, T. (1998) Sequence-based design of a peptide probe for the APC tumor suppressor protein. *Curr. Biol.* **8**, 823–830.
37. Schnarr, N. A. and Kennan, A. J. (2002) Peptide tic-tac-toe: heterotrimeric coiled-coil specificity from steric matching of multiple hydrophobic side chains. *J. Am. Chem. Soc.* **124**, 9779–9783.
38. Campbell, K. M. and Lumb, K. J. (2002) Complementation of buried lysine and surface polar residues in a designed heterodimeric coiled coil. *Biochemistry* **41**, 7169–7175.
39. Havranek, J. J. and Harbury, P. B. (2003) Automated design of specificity in molecular recognition. *Nat. Struct. Biol.* **10**, 45–52.
40. Kohn, W. D., Kay, C. M., and Hodges, R. S. (1995) Protein destabilization by electrostatic repulsions in the two-stranded alpha-helical coiled-coil/leucine zipper. *Protein Sci.* **4**, 237–250.
41. Zhou, N. E., Kay, C. M., and Hodges, R. S. (1994) The net energetic contribution of interhelical electrostatic attractions to coiled-coil stability. *Protein Eng.* **7**, 1365–1372.
42. Graddis, T. J., Myszka, D. G., and Chaiken, I. M. (1993) Controlled formation of model homo- and heterodimer coiled coil polypeptides. *Biochemistry* **32**, 12,664–12,671.
43. Moll, J. R., Olive, M., and Vinson, C. (2000) Attractive interhelical electrostatic interactions in the proline- and acidic-rich region (PAR) leucine zipper subfamily preclude heterodimerization with other basic leucine zipper subfamilies. *J. Biol. Chem.* **275**, 34,826–34,832.
44. 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.
45. Krylov, D., Barchi, J., and Vinson, C. (1998) Inter-helical interactions in the leucine zipper coiled coil dimer: pH and salt dependence of coupling energy between charged amino acids. *J. Mol. Biol.* **279**, 959–972.
46. Newman, J. R. and Keating, A. E. (2003) Comprehensive identification of human bZIP interactions with coiled-coil arrays. *Science* **300**, 2097–2101.
47. Oakley, M. G. and Hollenbeck, J. J. (2001) The design of antiparallel coiled coils. *Curr. Opin. Struct. Biol.* **11**, 450–457.
48. Monera, O. D., Zhou, N. E., Kay, C. M., and Hodges, R. S. (1993) Comparison of antiparallel and parallel two-stranded alpha-helical coiled-coils. Design, synthesis, and characterization. *J. Biol. Chem.* **268**, 19,218–19,227.

49. Monera, O. D., Kay, C. M., and Hodges, R. S. (1994) Electrostatic interactions control the parallel and antiparallel orientation of alpha-helical chains in two-stranded alpha-helical coiled-coils. *Biochemistry* **33**, 3862–3871.
50. Oakley, M. G. and Kim, P. S. (1997) Protein dissection of the antiparallel coiled coil from Escherichia coli seryl tRNA synthetase. *Biochemistry* **36**, 2544–2549.
51. Kohn, W. D. and Hodges, R. S. (1998) De novo design of α-helical coiled coils and bundles: models for the development of protein-design principles. *Trends Biotechnol.* **16**, 379–389.
52. 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.
53. Holton, J. and Alber, T. (2004) Automated protein crystal structure determination using ELVES. *Proc. Natl. Acad. Sci. USA* **101**, 1537–1542.
54. Gonzalez, L., Jr., Plecs, J. J., and Alber, T. (1996) An engineered allosteric switch in leucine-zipper oligomerization. *Nat. Struct. Biol.* **3**, 510–515.
55. Gurnon, D. G., Whitaker, J. A., and Oakley, M. G. (2003) Design and characterization of a homodimeric antiparallel coiled coil. *J. Am. Chem. Soc.* **125**, 7518–7519.
56. 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**, 12,603–12,610.
57. Cusack, S., Berthet-Colominas, C., Hartlein, M., Nassar, N., and Leberman, R. (1990) A second class of synthetase structure revealed by X-ray analysis of Escherichia coli seryl-tRNA synthetase at 2.5 A. *Nature* **347**, 249–255.
58. Stebbins, C. E., Borukhov, S., Orlova, M., Polyakov, A., Goldfarb, A., and Darst, S. A. (1995) Crystal structure of the GreA transcript cleavage factor from Escherichia coli. *Nature* **373**, 636–640.
59. McClain, D. L., Gurnon, D. G., and Oakley, M. G. (2002) Importance of potential interhelical salt-bridges involving interior residues for coiled-coil stability and quaternary structure. *J. Mol. Biol.* **324**, 257–270.
60. McClain, D. L., Woods, H. L., and Oakley, M. G. (2001) Design and characterization of a heterodimeric coiled coil that forms exclusively with an antiparallel relative helix orientation. *J. Am. Chem. Soc.* **123**, 3151–3152.
61. Litowski, J. R. and Hodges, R. S. (2001) Designing heterodimeric two-stranded alpha-helical coiled-coils: the effect of chain length on protein folding, stability and specificity. *J. Pept. Res.* **58**, 477–492.
62. Lau, S. Y., Taneja, A. K., and Hodges, R. S. (1984) Synthesis of a model protein of defined secondary and quaternary structure. Effect of chain length on the stabilization and formation of two-stranded alpha-helical coiled-coils. *J. Biol. Chem.* **259**, 13,253–13,261.
63. Su, J. Y., Hodges, R. S., and Kay, C. M. (1994) Effect of chain length on the formation and stability of synthetic alpha-helical coiled coils. *Biochemistry* **33**, 15,501–15,510.
64. Burkhard, P., Meier, M., and Lustig, A. (2000) Design of a minimal protein oligomerization domain by a structural approach. *Protein Sci.* **9**, 2294–2301.
65. 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.

66. Kwok, S. C. and Hodges, R. S. (2004) Stabilizing and destabilizing clusters in the hydrophobic core of long two-stranded α-helical coiled-coils. *J. Biol. Chem.* **279**, 21,576–21,588.
67. Litowski, J. R. and Hodges, R. S. (2002) Designing heterodimeric two-stranded alpha-helical coiled-coils. Effects of hydrophobicity and alpha-helical propensity on protein folding, stability, and specificity. *J. Biol. Chem.* **277**, 37,272–37,279.
68. O'Neil, K. T. and DeGrado, W. F. (1990) A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science* **250**, 646–651.
69. Dahiyat, B. I., Gordon, D. B., and Mayo, S. L. (1997) Automated design of the surface positions of protein helices. *Protein Sci.* **6**, 1333–1337.
70. Mason, J. M., Schmitz, M. A., Müller, K. M., and Arndt, K. M. (2006) Semirational design of Jon-Fos coiled coils with increased affinity: universal implications for leucine zipper prediction and design. *Proc. Natl. Acad. Sci. USA*, in press.
71. Richardson, J. S. and Richardson, D. C. (1988) Amino acid preferences for specific locations at the ends of alpha helices. *Science* **240**, 1648–1652.
72. Dasgupta, S. and Bell, J. A. (1993) Design of helix ends. Amino acid preferences, hydrogen bonding and electrostatic interactions. *Int. J. Pept. Protein Res.* **41**, 499–511.
73. Kohn, W. D., Kay, C. M., and Hodges, R. S. (1997) Positional dependence of the effects of negatively charged Glu side chains on the stability of two-stranded alpha-helical coiled-coils. *J. Pept. Sci.* **3**, 209–223.
74. Doig, A. J. (2002) Recent advances in helix-coil theory. *Biophys. Chem.* **101-102**, 281–293.
75. Kumar, S. and Bansal, M. (1998) Dissecting alpha-helices: position-specific analysis of alpha-helices in globular proteins. *Proteins* **31**, 460–476.
76. Aurora, R. and Rose, G. D. (1998) Helix capping. *Protein Sci.* **7**, 21–38.
77. Lu, M., Shu, W., Ji, H., Spek, E., Wang, L., and Kallenbach, N. R. (1999) Helix capping in the GCN4 leucine zipper. *J. Mol. Biol.* **288**, 743–752.
78. Sober, H. A. (1977) *CRC Handbook of Biochemistry and Molecular Biology*. 3rd ed., The Chemical Rubber Co, Cleveland, OH.
79. Virnekäs, B., Ge, L., Plückthun, A., Schneider, K. C., Wellnhofer, G., and Moroney, S. E. (1994) Trinucleotide phosphoramidites: ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis. *Nucleic Acids Res.* **22**, 5600–5607.
80. Spanjaard, R. A. and van Duin, J. (1988) Translation of the sequence AGG-AGG yields 50% ribosomal frameshift. *Proc. Natl. Acad. Sci. USA* **85**, 7967–7971.
81. Jung, S., Arndt, K. M., Muller, K. M., and Pluckthun, A. (1999) Selectively infective phage (SIP) technology: scope and limitations. *J. Immunol. Methods* **231**, 93–104.
82. Arndt, K. M., Jung, S., Krebber, C., and Pluckthun, A. (2000) Selectively infective phage technology. *Methods Enzymol.* **328**, 364–388.
83. Nakamura, Y., Gojobori, T., and Ikemura, T. (2000) Codon usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res.* **28**, 292; http://www.kazusa.or.jp/codon.
84. Kamtekar, S., Schiffer, J. M., Xiong, H., Babik, J. M., and Hecht, M. H. (1993) Protein design by binary patterning of polar and nonpolar amino acids. *Science* **262**, 1680–1685.

85. West, M. W. and Hecht, M. H. (1995) Binary patterning of polar and nonpolar amino acids in the sequences and structures of native proteins. *Protein Sci.* **4**, 2032–2039.
86. Pelletier, J. N., Campbell-Valois, F. X., and Michnick, S. W. (1998) Oligomerization domain-directed reassembly of active dihydrofolate reductase from rationally designed fragments. *Proc. Natl. Acad. Sci. USA* **95**, 12,141–12,146.
87. Pelletier, J. N., Arndt, K. M., Plückthun, A., and Michnick, S. W. (1999) An in vivo library-versus-library selection of optimized protein-protein interactions. *Nat. Biotechnol.* **17**, 683–690.
88. Arndt, K. M., Jouaux, E. M., and Willemsen, T. (2004) Der richtige Dreh—Coiled Coils auf dem Weg zur Anwendung. *BioForum* **10**, 48–49.
89. Johnsson, N. and Varshavsky, A. (1994) Split ubiquitin as a sensor of protein interactions in vivo. *Proc. Natl. Acad. Sci. USA* **91**, 10,340–10,344.
90. Rossi, F., Charlton, C. A., and Blau, H. M. (1997) Monitoring protein-protein interactions in intact eukaryotic cells by beta-galactosidase complementation. *Proc. Natl. Acad. Sci. USA* **94**, 8405–8410.
91. Wehrman, T., Kleaveland, B., Her, J. H., Balint, R. F., and Blau, H. M. (2002) Protein-protein interactions monitored in mammalian cells via complementation of beta -lactamase enzyme fragments. *Proc. Natl. Acad. Sci. USA* **99**, 3469–3474.
92. Galarneau, A., Primeau, M., Trudeau, L. E., and Michnick, S. W. (2002) Beta-lactamase protein fragment complementation assays as in vivo and in vitro sensors of protein protein interactions. *Nat. Biotechnol.* **20**, 619–622.
93. Ghosh, I., Hamilton, A. D., and Regan, L. (2000) Antiparallel leucine zipper-directed protein reassembly: application to the green fluorescent protein. *J. Am. Chem. Soc.* **122**, 5658–5659.
94. Miller, J. and Stagljar, I. (2004) Using the yeast two-hybrid system to identify interacting proteins. *Methods Mol. Biol.* **261**, 247–262.
95. Marino-Ramirez, L., Campbell, L., and Hu, J. C. (2003) Screening peptide/protein libraries fused to the lambda repressor DNA-binding domain in E. coli cells. *Methods Mol. Biol.* **205**, 235–250.
96. Hu, J. C., O'Shea, E. K., Kim, P. S., and Sauer, R. T. (1990) Sequence requirements for coiled-coils: analysis with lambda repressor-GCN4 leucine zipper fusions. *Science* **250**, 1400–1403.
97. Willats, W. G. (2002) Phage display: practicalities and prospects. *Plant Mol. Biol.* **50**, 837–854.
98. Keating, A. E., Malashkevich, V. N., Tidor, B., and Kim, P. S. (2001) Side-chain repacking calculations for predicting structures and stabilities of heterodimeric coiled coils. *Proc. Natl. Acad. Sci. USA* **98**, 14,825–14,830.
99. 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.