Abstract

Protein design allows sequence-to-structure relationships in proteins to be examined and, potentially, new protein structures and functions to be made to order. To succeed, however, the protein-design process requires reliable rules that link protein sequence to structure/function. Although our present understanding of coiled-coil folding and assembly is not complete, through numerous bioinformatics and experimental studies there are now sufficient rules to allow confident design attempts of naturally observed and even novel coiled-coil motifs. This review summarizes the current design rules for coiled coils, and describes some of the key successful coiled-coil designs that have been created to date. The designs range from those for relatively straightforward, naturally observed structures—including parallel and antiparallel dimers, trimers and tetramers, all of which have been made as homomers and heteromers—to more exotic structures that expand the repertoire of Nature’s coiled-coil structures. Examples in the second bracket include a probe that binds a cancer-associated coiled-coil protein; a tetramer with a right-handed supercoil; sticky-ended coiled coils that self-assemble to form fibers; coiled coils that switch conformational state; a three-component two-stranded coiled coil; and an antiparallel dimer that directs fragment complementation of larger proteins. Some of the more recent examples show an important
development in the field; namely, new designs are being created with function as well as structure in mind. This will remain one of the key challenges in coiled-coil design in the next few years. Other challenges that lie ahead include the need to discover more rules for coiled-coil prediction and design, and to implement these in prediction and design algorithms. The considerable success of coiled-coil design so far bodes well for this, however. It is likely that these challenges will be met and surpassed.

I. Introduction to Protein Design

Why bother with protein design? After all, nature has produced a wide variety of beautiful protein structures with fascinating functions. The author’s response to this question has three points. First, protein design provides the acid test of our understanding of the informational aspect of the protein-folding problem and, in particular, how protein sequence relates to the three-dimensional structure (and function) of proteins. Second, protein design attempts to capture the salient features of protein structure and function in simpler contexts. In other words, any natural protein sequence contains superimposed information about the protein’s folding, structure and stability, and its function(s). Protein design attempts to disentangle and use this information. Third, natural proteins represent only a tiny fraction of the possible protein sequence space available. Whether they also represent a fraction of the possible protein structure space is another question. It is not yet clear whether the natural protein structures observed so far, which suggest a limited number of protein folds (Liu et al., 2004a; Wolf et al., 2000), are in fact significantly greater. Whatever the case, protein design offers possibilities for exploring protein sequences, structures, and functions beyond those examined by nature. In turn, it presents a route to new structures and functions with potentially exploitable applications.

There are a number of approaches to protein engineering and design, which, for the purposes of this review, are defined here. Protein engineering in general refers to the process of making one or a relatively small number of mutations in a natural protein framework to examine sequence-to-structure/function relationships in proteins, and/or to improve their structural properties or functions. At the other extreme, de novo protein design refers to attempts to construct totally new protein sequences with prescribed structures (and functions) from first principles. By protein redesign, this author refers to the process of mutating already designed scaffolds to create new molecules with improved structural properties, stabilities, and functions.
Clearly, in de novo design the starting protein sequence is chosen by the designer. In protein engineering and redesign, there are a number of experimental routes to making mutants. These can be made one or a small number at a time, in which case the process is iterative. Alternatively, many mutations can be made simultaneously, either at specific sites in the scaffold (saturation mutagenesis) or randomly throughout the protein (random mutagenesis) to create a library of mutants from which variants with desired properties are selected. These methods are referred to as combinatorial or semirational. Finally, in both the rational and combinatorial approaches, the mutations can be guided by computational methods and searches.

Two important concepts in de novo protein design are those of positive and negative design. In positive design, sequence-to-structure rules are used to direct the formation of and stabilize the target structure, whereas negative design refers to the idea of designing against (i.e., destabilizing) alternative and often competing structures (Beasley and Hecht, 1997; Hellinga, 1997; Hill et al., 2000). The application of these principles to coiled-coil structures is particularly important because, at least at first sight, coiled coils share a straightforward repeat sequence and the possibilities for forming the wrong quaternary structure are significant.

So far, all references to designing protein functions have been bracketed. This is because the current state-of-the-art in protein design is at the structural level. In other words, we are better at designing protein structures than we are at introducing function. This is not to say that the redesign and de novo design of protein function is totally beyond us. Indeed, protein engineers have been modifying natural protein frameworks and their functions for around two decades.

This review focuses almost exclusively on the rational design and redesign of coiled-coil motifs. Many reviews have been written on protein design in general. Recent papers that, in the view of this author, reflect the current state of the art in globular protein de novo design of structure and redesign of function are from Baker (Kuhlman et al., 2003) and Hellinga (Dwyer et al., 2004; Looger et al., 2003), respectively. Discussion of the design and redesign of four-helix bundle proteins is limited. Although some four-helix bundles can also be considered as coiled coils, this is not always the case. Also, with the vagaries of design in the absence of full structure determination, this author considers it best not to group all four-helix bundles together with coiled coils at present. Finally, reviews on the design of four-helix bundles are available (Hill et al., 2000; Woolfson, 2001). Regarding the design of coiled-coil motifs there have been a number of commentaries and review papers on this over the years (Cohen and Parry, 1990, 1994; Kohn and Hodges, 1998; MacPhee and Woolfson,
The review is laid out as follows: in Part II, there is a discussion of the structural features of the coiled-coil assemblies and the current state-of-the-art of sequence-to-structure rules pertinent to design; in Part III, there is a detailed description of some of the key coiled-coil designs, including the rules used to create them and their significance in the field; and, finally, there is a discussion of the potential for new coiled-coil designs and where the field might go in the future.

II. Rules for Coiled-Coil Design: The Basics of Coiled-Coil Sequence and Structure

Any design and engineering process requires an understanding of how to assemble the basic building blocks available into the objects being targeted. Though it is not always absolutely necessary, ideally this understanding should be at as fundamental a level as possible. Protein design and assembly are no different. Here, we refer to rules that link protein sequence and structure. These rules can be general, such as “place hydrophobic amino acids alternately three and four residues apart to direct the folding and assembly of amphipathic α-helices,” or a little more specific, such as “make every second hydrophobic amino acid Leu to guide the assembly of dimers.” These are examples of rules of thumb that allow protein designers to build up a protein sequence compatible with a desired target structure. They can be used manually or built into algorithms to be implemented computationally. In many cases, these rules work very well and, if engineering was the only goal of protein design, that might be the end of it. However, it is much more powerful, and ultimately more intellectually satisfying, to understand the fundamental basis of the rules. In terms of protein design, the chemical level of understanding suffices (i.e., how the sequence-to-structure rules can be rationalized in terms of the underlying noncovalent forces). For this reason, the current state-of-the-art of the rules for coiled-coil design is presented below in terms of the noncovalent forces that direct coiled-coil folding and assembly.

For a more thorough description and discussion of coiled-coil structures, the reader is referred to Chapter 3. For the purposes of this review, coiled coils are described as canonical or noncanonical. Canonical coiled coils are those based on tandem heptad sequence repeats that form
right-handed amphipathic α-helices, which then assemble to form helical bundles with left-handed supercoils. In contrast, noncanonical α-helical coiled coils are built from nonheptad-based repeats and, as a consequence, do not necessarily form coiled coils with left-handed or even regular supercoils. For instance, 11-residue repeat sequences give rise to right-handed coiled coils. The majority of this review focuses on designs made using canonical, heptad-based coiled coils as their starting points.

A. Hydrophobic Interactions and Helix Formation

As with most, if not all, biological self-assembly processes, a key driving force in coiled-coil folding and assembly is the hydrophobic effect. Put simply, the hydrophobic effect is the phenomenon that hydrocarbon and water tend to phase separate. Thus, when in the context of a biological, aqueous buffer, biological macromolecules fold or self-assemble to minimize the hydrophobic surface area in contact with bulk solvent. In the case of proteins, which can be considered as essentially linear heteropolymers of hydrophobic (H: Ala, Phe, Ile, Leu, Met, Val, Trp, and Tyr) and polar (P: Asp, Glu, His, Lys, Asn, Gln, Arg, Ser, and Thr) residues, the polypeptide chain folds to bury H residues and expose P side chains. (In these terms, there is also a third category of amino acid residue, which can be designated as special [S: Cys, Gly, and Pro]. These amino acids confer unusual conformational properties on the polypeptide chain. However, for reasons outlined in the main text, these residues occur less frequently in coiled-coil sequences compared with protein sequences as a whole.) Therefore, to a first approximation, the structures arrived at through this reaction to the aqueous medium reflect the pattern of H and P residues along the primary sequence. For instance, alternating patterns of H and P residues (HHPHP...) tend to adopt β-strand conformations and lead to β-sheet-based structures. This is because alternate residues along a β-strand point in opposite directions out from the strand, so an alternating HP pattern produces an amphipathic β-strand (i.e., a strand with the H residues sequestered on one side and the P residues on the other). In turn, two or more such strands come together to form an amphipathic β-sheet, which can pack into a globular structure to bury its hydrophobic face.

As the α-helix has 3.6 residues per turn, hydrophobic side chains spaced at combinations of three and four residues apart are required to make an amphipathic structure. For helices in globular proteins, a variety of combinations of three-plus-four spacings of hydrophobic residues are observed (Chothia et al., 1981). This leads to a range of helix-helix packing angles and arrangements. However, to form more persistent fibrous coiled-coil
structures, more regular $HP$ patterns are required. The canonical heptad pattern, $HPPHPPP$, seems to predominate in nature. The full reasons for this are not clear but, as described below, it is likely that it leads to the most efficient way of packing two or more helices in a fibrous helical bundle.

The $HPPHPPP$ pattern is usually denoted $abcdefg$, with $a$ and $d$ assigned to the $H$ residues, and can be pictured in terms of a helical wheel (Fig. 1C and D). Tandem heptad repeats along a polypeptide chain give an average separation between $H$ residues of 3.5 residues. As this falls short of the 3.6-residues per turn of a regular $\alpha$-helix, the $a$ plus $d$ hydrophobic face tracks around the helix with its own helical pitch, which is slighter and in the opposite direction of the backbone $\alpha$-helix. Thus, when two or more helices pack together to form a coiled-coil oligomer, they do not pack straight as suggested by standard helical wheel diagrams, but wrap around one another in order to maximize contacts between the hydrophobic surfaces (Fig. 1A and D). The sense of this so-called supercoiling is the same as the hydrophobic stripe. Thus, in the case of a canonical coiled coil, the helices pack with a left-handed supercoil.

In terms of coiled-coil design several questions emerge from this rudimentary structural analysis. First, what is the minimum number of contiguous heptad repeats required to achieve a stably folded coiled-coil structure? The full answer to this is a little involved because for some natural coiled-coil sequences “trigger” sequences are believed to be essential for folding, even for very long sequences (Burkhard et al., 2000a; Kammerer et al., 1998; Lee et al., 2001; Steinmetz et al., 1998; Walshaw and Woolfson, 2001b). However, for design purposes, three or four heptad repeats appear to be sufficient for stable coiled-coil folding (Lumb et al., 1994; Su et al., 1994; Talbot and Hodges, 1982), though, as discussed below, a two-heptad design has been described, albeit with poor oligomer-state specificity (Burkhard et al., 2000b; Meier et al., 2002).

Second, what type of residues are best at the $a$ and $d$ sites? This is also a difficult question to answer directly because, as will be addressed in more detail over the next few sections, the nature of residues at these sites influences coiled-coil stability, oligomer state, partner selection, and helix-helix orientation (Table I). However, in general terms, natural coiled-coil sequences tend to use the aliphatic hydrophobic residues (Ala, Ile, Leu, Met, and Val) at these positions, rather than the aromatic hydrophobic side chains (Phe, Trp and Tyr) (Parry, 1982; Woolfson and Alber, 1995). The reason for this is probably a combination of bulk and steric constraints presented by the aromatic residues. However, a thorough understanding of the possible exclusion of aromatic side chains from coiled-coil
cores awaits further protein-engineering experiments in which multiple aromatic residues are introduced at \( a \) and \( d \). Note added in proof: one such experiment has now been done (see Liu et al., 2004b), and an \( a = d = \text{Trp} \) peptide forms a pentamer.

Third, what residues are best at the remaining \( b, c, e, f, \) and \( g \) sites? Again, this will be addressed in more detail below. However, in general, these positions are more permissive than the \( a \) and \( d \) sites, though polar and helix-favoring residues (Ala, Glu, Lys, and Gln) tend to be favored both by nature and by protein designers for these positions.
Furthermore, regarding the special residues Cys, Gly, and Pro, these occur with quite low frequencies in coiled-coil proteins (Conway and Parry, 1990, 1991; Lupas et al., 1991; Parry, 1982; Woolfson and Alber, 1995) and tend only to be used to perform very specific roles in de novo coiled-coil designs. Gly and Pro are regarded as \( \alpha \)-helical breakers and so are rarely chosen by designers for the central regions of coiled-coil structures. Cys can form disulphide bridges or be otherwise oxidized. It is usually avoided in design sequences unless it is specifically required for cross-linking polypeptide chains.

### B. van der Waals’ Forces, Steric Constraints, and Oligomer Specification

One property of hydrophobic interactions is that they tend not to be specific. In other words (and as evident in natural coiled-coil structures), the basic coiled coil pattern \( HPPHPPP \) is compatible with a number of helix-bundle quaternary structures. Dimeric, trimeric, tetrameric, pentameric, and dodecameric coiled coils are all known as are homomeric and heteromeric complexes and topologies with parallel, antiparallel, or mixed arrangements of helices (Burkhard et al., 2001; Gruber and Lupas, 2001).
Thus, the question in coiled-coil prediction and design is: what specific replacements are superimposed on the basic HPPHPPP pattern to direct the functional oligomerization state? This question was first tackled by Conway and Parry, who analyzed natural coiled-coil sequences that formed dimers and trimers (Conway and Parry, 1990, 1991). Woolfson and Alber (1995) advanced this approach by comparing amino-acid profiles for these two structures directly. The work that made the biggest impact on this issue, however, was the collaborative experimental study from the Kim and Alber laboratories using the GCN4 leucine-zipper peptide model system and mutants thereof.

The first experimental study was the high-resolution X-ray crystal structure of the leucine zipper peptide, GCN4-p1, itself (O’Shea et al., 1991), which revealed that the packing geometries of residues at $a$ and $d$ were different. The significance of this finding for oligomer state selection became clearer with Harbury’s studies of hydrophobic core mutants of the same peptide (Harbury et al., 1993, 1994). As demonstrated by several protein engineering studies (Woolfson, 2001), hydrophobic residues within the cores of globular proteins tend to be very tolerant of substitution by other hydrophobic side chains. This is not the case for multichain coiled-coil structures. GCN4-p1, which forms dimers exclusively, has Leu at all four $d$ positions, and $1 \times$ Met, $1 \times$ Asn, and $3 \times$ Val at its $a$ sites. Harbury has described multiple core mutants of GCN4-p1 in which all $a$ residues (except Met-2) were simultaneously substituted for one of Ile, Leu, or Val, and all $d$ residues were simultaneously substituted for one of same subset of aliphatic side chains (Harbury et al., 1993). The peptides were generically named GCN4-p-$ad$. For example, GCN4-p-IL refers to the GCN4-p1 sequence with said $a$ and $d$ sites replaced by Ile and Leu, respectively. Seven of the nine possible peptides have been characterized. The comparison of peptides p-IL, p-II, and p-LI is interesting as they form dimers, trimers, and tetramers, respectively. Structures for the new trimeric (Harbury et al., 1994) and tetrameric (Harbury et al., 1993) forms are available and these have helped rationalize the different oligomer state selections on the basis of different packing arrangements made within the cores of the structures.

As O’Shea and colleagues (1991) have noted, the packing of side chains at the $a$ and $d$ sites of the GCN4-p1 dimer is different. At the $d$ position, the $\alpha-\beta$ bond vector of the side chain points into the interface and directly towards the neighboring helix. This type of geometry, called perpendicular packing by Harbury, precludes $\beta$-branched residues such as Ile and Val from occupying these sites and favors Leu. This is the reason that the leucine zipper is a leucine zipper: the hallmark of the bZIP transcription
factors is a run of four or so leucine residues spaced seven residues apart, and in a heptad repeat these fall at the d sites (Landschulz et al., 1988). At the a sites, however, the α–β bond vector points out from the helical interface, and is termed parallel packing. As a result, the a site of dimeric coiled coils is much more permissive of amino acid substitution (Conway and Parry, 1990; Hu et al., 1990; Woolfson and Alber, 1995). In hydrophobic side chains, the β-branched residues Ile and Val are favored at a because they contribute hydrocarbon back into the helical interface. These different geometries for the dimer case are shown in Fig. 2. In Harbury’s p-LI tetramer, the packing geometries are reversed compared with the dimer. The packing at a is perpendicular, whereas the packing at d is parallel. This fits with the swap of Ile and Leu residues at a and d between the p-IL dimer and the p-LI tetramer.

In summary, the relative order of the β-branched residues and leucine at a and d has a major influence on oligomer state selection. The rule is that Leu prefers perpendicular packing whereas the β-branched residues, Ile and Val, prefer parallel packing (Table I). On this basis, it is of little surprise that a retro-GCN4-p1 sequence (i.e., with a = Leu and d = Ile) forms a tetramer and not a dimer (Mittel et al., 2000).

Fig. 2. Schematic diagrams (left) and experimental examples (right) for different core-packing geometries. (A) Parallel packing, showing an α-layer from the GCN4 leucine zipper (O’Shea et al., 1991). (B) Perpendicular packing at a d-layer in the same structure. (C) Acute packing at a d-layer in a structure of a trimeric GCN4 mutant (Gonzalez et al., 1996c). Reproduced from Walshaw and Woolfson (2001b).
So what happens in the trimer? GCN4-p-II crystallizes as a trimer (Harbury et al., 1994). In this case, the packing at \( a \) and \( d \) is closely similar: it is intermediate between the perpendicular and parallel extremes (Fig. 2), and accordingly is referred to as acute packing. As a result, less selectivity for amino-acid type might be expected between the two sites. Indeed, this largely fits with Harbury’s experimental studies and with analyses of natural trimeric coiled-coil sequences (Conway and Parry, 1991; Woolfson and Alber, 1995) (Table I). Taking these studies together, the rule of thumb of \( a = d = \text{Ile or Leu} \) to specify trimers generally emerges. However, one further point is worth making here: the trimer state seems to be the default, and in the absence of strong sequence determinants of oligomer state, coiled-coil peptides and proteins will tend to form this structure. This is apparent from a number of studies using the GCN4-p1 background. First, Harbury’s peptides with all \( a = \text{Val} \) tend to form mixed oligomer states in solution (Harbury et al., 1993). In addition, substitutions that replace the central \( a \) site, which is Asn in the wild-type sequence, form mixtures of oligomer states (Gonzalez et al., 1996a,b,c). These results suggest that, at least for the \( a \) position, Val specifies the oligomer state very poorly compared with Ile. However, as described below, Val in combination with certain polar residues at \( a \), as in wild-type GCN4-p1, does specify the dimer (Gonzalez et al., 1996c; Lumb and Kim, 1995; Woolfson and Alber, 1995).

Walshaw and Woolfson (2001b) have analyzed the packing geometries of all coiled-coil structures identified in the Brookhaven Protein Data Bank (release #87), and have largely confirmed Harbury’s experimental analysis. Furthermore, their work provided updated and structurally validated amino-acid profiles for the main coiled-coil quaternary structures. For instance, in parallel dimers the rule of \( a = \text{Ile or Val} \) combined with \( a = \text{Ile or Val} \), together with rules for polar residues at \( a \) (see below), largely holds up (Table I). These rules are, of course, rules of thumb and considerable variation is observed in natural sequences. Nonetheless, they are extremely powerful for specifying the oligomer state of designed coiled-coil peptides with confidence. Furthermore, several protein engineering and redesign studies have reproduced these observations and put measures on the thermodynamic effects of making certain hydrophobic mutations at the core \( a \) and \( d \) sites (Moitra et al., 1997; Tripet et al., 2000; Wagschal et al., 1999). More recently, researchers have begun to use computational methods (Havranek and Harbury, 2003; Keating et al., 2001) and to incorporate nonnatural amino acids to probe, understand, specify, and stabilize coiled-coil hydrophobic interfaces (Bilgicer et al., 2001; Schnarr and Kennan, 2001, 2002, 2003; Yoder and Kumar, 2002).
Before leaving the subject of the hydrophobic component of coiled-coil interfaces, it is worth noting that hydrophobic residues at \( e \) and \( g \) and possibly at \( b \) and \( c \) also influence oligomer state selection. In coiled-coil sequences, the frequencies of hydrophobic residues at these sites increase with oligomer state (Conway and Parry, 1990, 1991; Walshaw and Woolfson, 2001b; Woolfson and Alber, 1995). Likewise, in coiled-coil structures these sites are increasingly buried and involved in the interface (Harbury et al., 1993; Kajava, 1996; Walshaw and Woolfson, 2003). Though these observations need further experimental testing and verification, potentially they provide a further route to designing coiled coils of different oligomer states.

**C. Salt-Bridge Interactions and Partner Selection**

It has long been appreciated that residues that flank the main hydrophobic seam of the coiled-coil interface (i.e., residues at \( e \) and \( g \) and, more latterly, those at \( b \) and \( c \)) are likely to contribute to coiled-coil specificity and stability (McLachlan and Stewart, 1975; McLachlan et al., 1975). Specifically, in parallel coiled-coil structures, the \( g \) site of one helix is brought close to the \( e \) site of the successive heptad of a neighboring helix (Fig. 1C) in a so-called \( g_n;e_{n+1} \) interaction. As will be discussed later, in antiparallel structures the analogous interactions are \( e:e' \) and \( g:g' \) (Fig. 1D). This difference may be important in distinguishing between parallel and antiparallel structures in nature, and it also provides the protein designer with excellent positive and negative design rules.

At least from the analyses of amino-acid profiles for parallel dimers and trimers (Conway and Parry, 1990, 1991; Lupas et al., 1991; Walshaw and Woolfson, 2001b; Woolfson and Alber, 1995), it is beyond doubt that the placement of oppositely charged side chains at these sites must play a part in coiled-coil assembly, structure, and stability. Likewise, the proximity of oppositely charged side chains emanating from these sites in experimentally determined structures and the postulated resulting salt bridges add further compelling evidence for the importance of such interactions (O’Shea et al., 1991). Nonetheless, and despite many experimental studies to probe these interactions, the issue remains controversial (Lavigne et al., 1996; Lumb and Kim, 1996). Specifically, the question is whether \( g_n;e_{n+1} \) interactions stabilize and specify coiled-coil structures, or if they simply help specify partner selection and helix orientation.

Although this may seem like dodging an important issue—clearly in protein design it is important to understand how protein sequence relates to both protein structure and stability—for the purposes of this review, it is an academic question. This is for two reasons, which will become clear in
the examples of successful coiled-coil designs. First, stability tends not to be an issue when designing coiled coils, as plenty can be gained from the core interactions alone. Second, through the elegant studies of O’Shea and colleagues on the Fos:Jun system, it is clear that charge-charge complementarity in coiled-coil interfaces plays a significant role in partner selection (O’Shea et al., 1989, 1992).

On this basis, coiled-coil designers often place oppositely charged residues Lys and Glu at complementary $g$ and $e$ sites to direct the assemblies that they are targeting (positive design), and make similar potential interactions in alternative structures repulsive (negative design).

\section*{D. Buried Polar Groups: The Icing on the Cake}

Returning to the $a$ and $d$ sites of coiled-coil interfaces, they are not the exclusive province of hydrophobic residues: polar residues are found here and, in many cases, they are highly conserved. In retrospect, this is also apparent in the amino-acid profiles of dimeric and trimeric coiled coils (Conway and Parry, 1990, 1991; Lupas et al., 1991). Furthermore, inspection of these profiles reveals trends in the data; for instance, basic residues occur frequently at the $a$ sites of dimeric coiled coils (Conway and Parry, 1990). The real importance of such inclusions, however, only became apparent with the determination of the structure of the leucine-zipper peptide GCN4-p1 (O’Shea et al., 1991).

Position 16 of GCN4-p1—an $a$ site—is Asn, which is conserved across many leucine-zipper sequences. In the structure, this residue is buried and makes an asymmetric side chain-side chain hydrogen bond with the corresponding Asn in the partner strand. This Asn-Asn' pair appears to play a number of roles, some of which have been subsequently verified experimentally. First, the inclusion of the pair in the otherwise hydrophobic core is destabilizing, and replacement of the residue by the canonical Val increases coiled-coil stability considerably (Harbury et al., 1993). There may be biological significance to this in, for instance, modulating dimer stability for control and protein turnover in the cell. However, what is more dramatically obvious in the Asn16Val mutant is its loss of oligomer-state specificity: it forms mixtures of oligomers (Harbury et al., 1993). Thus, second, Asn-16 is essential for specifying the dimer state of GCN4-p1. Third, the residue may be important in destabilizing other alternative structures, such as antiparallel and out-of-register structures, both of which would result in two buried and noncomplemented Asn side chains.

The importance of Asn at $a$ in specifying the dimer is further demonstrated by the comparison of amino-acid profiles for dimeric and trimeric coiled coils (Woolfson and Alber, 1995). In the normalized profiles, Asn at
a is significant in dimers, but not in trimers. Similarly, and following on from the earlier studies of Conway and Parry (1990), the profiles from Woolfson and Alber show that the basic residues Lys and Arg occur significantly at the a sites of dimers, but not in trimers. With regard to promoting the trimer the residues Gln, Ser, and Thr are all favored at the a sites of trimer sequences, but not in dimers. A number of experimental studies lend support to these correlations, and provide structural rationales for the inclusion and accommodation of polar residues in coiled-coil interfaces (Akey et al., 2001; Campbell and Lumb, 2002; Campbell et al., 2002; Eckert et al., 1998; Gonzalez et al., 1996c; Lumb and Kim, 1995; Oakley and Kim, 1998). For instance, inclusion of Lys at a appears to specify dimers, not because of specific interhelix interactions in this state, but because the charge side chain cannot escape the core of higher-order states such as the trimer, which are therefore compromised (Gonzalez et al., 1996c).

In summary, the inclusion of the polar residues, even when hydrogen-bonding potentials are satisfied, is destabilizing. However, though the burial of a polar residue may destabilize the target structure, it may destabilize alternative structures even more (Gonzalez et al., 1996c). In this respect, the use of buried residues provides a very powerful negative-design principle for coiled-coil design.

III. KEY COILED-COIL DESIGNS

The author recognizes that some readers will find it difficult to visualize the sequences given below in the context of coiled-coil structures. For this reason, it is suggested that readers consult the original articles, nearly all of which give annotated helical-wheel diagrams for the design sequences or, preferably, construct their own annotated helical wheels using the sequences given and blank diagrams of the type shown in Fig. 1C and D.

A. Parallel Structures


To the author’s knowledge, Hodges and colleagues presented the first discrete coiled-coil peptide of de novo design (Hodges et al., 1981). This design followed earlier studies by the group on polymers of the heptapeptide repeat sequence KLESLES. The design principles, which are based on an analysis of the amino-acid composition of tropomyosin, are clear and were well ahead of their time. The peptide comprised three blocks,
A (KCAELEG), B (KLEALEG), and C (KLEALEGK), assembled in the order AB₄C to give a 43 residue peptide, which was air-oxidized through the N-terminal Cys to form an 86-residue covalent dimer. The main features of this design were the straightforward \( a = d = \text{Leu} \) core, the complementary \( g:e \) Lys:Glu pairs, and the unusual Gly residues at \( f \), which were included for synthetic reasons. Note that the heptad repeat is denoted \( gabcdedef \) in this design. The dimer exhibited considerable helicity and thermal stability consistent with a fully folded coiled coil. This original design has spawned considerable work in coiled-coil design and redesign from the Hodges group (Kohn and Hodges, 1998) and others (Schneider et al., 1998).

2. Coil-Ser: A Peptide Designed as Parallel Homodimer that Forms an Up-Up-Down Homotrimer (1990)

O’Neil and DeGrado (1990) adapted the original Hodges sequence to engineer a coiled-coil-based host-guest system for determining a thermodynamic scale of the helix-forming tendencies of the amino acids. Based on Hodges’ building block—KLEALEG—they described a four-heptad peptide. Some significant changes were made, however. The first \( a \) position and the last \( f \) site were Trp and His, respectively, to facilitate spectroscopic characterization. Also, the glycine residues at the \( f \) sites of first and third heptad were replaced by Lys to improve helical propensity, and the remaining \( f \) position of the second heptad was used as the guest site and substituted by all 20 amino acids plus Aib. The stabilities of the resulting 21 peptides were determined by urea denaturation experiments. These showed a striking correlation with statistical measures of \( \alpha \)-helix propensity.

The determination of the crystal structure of coil-Ser provided an interesting twist on this story (Lovejoy et al., 1993). Rather than the parallel dimer expected, coil-Ser formed an up-up-down homotrimer (Fig. 3). Trimerization in solution was also confirmed by analytical ultracentrifugation. The likely reasons for this were the predominantly all-Leu core, and the inclusion of Trp at the first \( a \) position. As noted above, the \( a = d \) core favors trimers, and Trp is too bulky for all three Trp residues to be accommodated within one coiled-coil layer, hence the switch of one helix to give the up-up-down arrangement. As noted by DeGrado and colleagues, this result did not affect the scale of helix-forming tendencies originally reported using the coil-Ser system (Lovejoy et al., 1993).


The “Peptide Velcro” design by O’Shea et al. (1993) was based on their experience with the GCN4-p1 and Fos/Jun systems. The core design—which preceded Harbury’s work on oligomer-state selection (Harbury et al.,
was influenced by the foregoing Hodges and DeGrado designs and was straightforward, with four \( a = d = \text{Leu} \) heptads. However, the system did not form trimers because of the GCN4-inspired inclusion of Asn at the central \( a \) position (O’Shea et al., 1991). In one partner, Acid-p1, Glu were placed at all \( e \) and \( g \) sites while its complement, Base-p1, had Lys at these positions. Thus, the idea of the design was to draw the two peptides together with a full set of complementary \( g_{a}:e_{n+1} \) pairings between the partner peptides. Consistent with the design principles, the isolated peptides did not fold in water but, when mixed, formed a stable 1:1 complex that is fully helical and unfolded cooperatively. The dissociation constant for this design is an impressive 30 nM (20 °C, pH 7). Additional experiments revealed that, as in the Fos/Jun system, repulsive electrostatic interactions in the alternative homodimers direct heterospecificity more than the favorable electrostatic interactions in the targeted heterodimer.
Interestingly, when the buried Asn was replaced by Leu, structural specificity was lost and the peptides formed a heterotetramer without fixed helix-helix orientations (Lumb and Kim, 1995). In a further report on this system, Sia and Kim (2001) showed that the two copies of Acid-p1 in the heterotetramer could be replaced by a peptide, D-Acid, in which all of the residues were made from D-amino acids. This clever redesign was guided by helical-net diagrams that considered core packing in the D/L structure. These revealed $a:d'$ rather than $a:a'$ and $d:d'$ layers in the core, which were accommodated in the new design by a half-heptad shift of a redesigned L-Base sequence.

While on the topic of heterodimer design, Vinson and coworkers have had considerable success in making heterodimerizing leucine zippers with an impressive range of properties, including very tight binding (Moll et al., 2001; Vinson et al., 1993). Hodges and colleagues have presented a series of related designs, E/K coils, which they propose for use in biotechnology as, for example, capture reagents for affinity chromatography and biosensor-based applications (Chao et al., 1998; Litowski and Hodges, 2001, 2002).


The ABC design of Nautiyal et al. (1995) represented a step up in complexity and a considerable increase in the degree of difficulty over foregoing heterodimeric coiled-coil designs. Again, this was a four-heptad design with a straightforward $a = d = $ Ile core and a single polar residue (Gln) at $a$ as a failsafe for parallel trimer formation. With only two charges to play with and three helical interfaces to specify, the selection of $e$ and $g$ residues used an algorithm to choose from more than 16 million peptide combinations. The solution maximized salt bridges in the targeted parallel heterotrimer, and minimized them in all alternative combinations. In this respect, the approach used both positive and negative-design principles. Two solutions to the problem were returned: the one synthesized, and its mirror image reversed the order of helical packing. One failing of the design was that the individual peptides did trimerize, as did the binary combinations. However, none of these combinations matched the heterotrimer in thermal stability. An important and impressive aspect of this design was that its crystal structure confirmed all features of the design, including the correct A–B–C helical packing arrangement (Nautiyal and Alber, 1999).


Based on the natural sequence of the core domain of the Lac repressor, which tetramerizes, Fairman et al. (1996) have engineered two 21-residue peptides, Lac21E and Lac21K. In isolation, the peptides did not fold
appreciably, but when combined they formed thermostable tetramers. In these short sequences, the d residues were Leu, and the three a sites were Met, Leu, and Val, respectively. Interestingly, the e and g positions were combinations of Ala, Glu, Gly, and Ser. This left specifying the structure down to charged residues at the b and c positions, which were all Glu in Lac21E and all Lys in Lac21K. This idea of extending the coiled-coil interfaces beyond the a and d seam and the e and g flanking residues to b and c has been noted elsewhere for coiled-coil structures (Harbury et al., 1993; Kajava, 1996; Walshaw and Woolfson, 2003) and sequences (Conway and Parry, 1991; Walshaw and Woolfson, 2001b; Woolfson and Alber, 1995).

6. Anti-APCp1: A Coiled-Coil Probe for the APC Tumor Suppressor Protein (1998)

Mutations of the human APC gene are associated with both sporadic and familial forms of colon cancer. The APC protein is a large, multi-domain protein that has a 55-residue, N-terminal dimeric coiled coil (APC-55). Alber and colleagues used rules of thumb and those derived from an analysis of the covariation of a:a0 and d:d0 pairs in the cytokeratins (which form obligate heterodimers) to create a mutant of APC-55, anti-APCp1, as a potential probe for the APC protein (Sharma et al., 1998).

In all, 20 mutations were described that can be grouped as follows. First, five changes were made at the a and d sites based on the analysis of the keratins. Second, eight changes were made at the e and g sites to introduce salt bridges to favor the anti-APCp1:APC-55 heterodimer and to destabilize homodimerization of anti-APCp1. Finally, seven changes were made at the e, f, and g positions to improve the charge and the helicity of the peptide and to introduce chromophores. Although anti-APCp1 did form stable oligomers, these were higher orders than the dimer. Moreover, when mixed with APC-55, a stabilized heterodimer was formed as judged by solution-phase biophysics and gel electrophoresis. In addition, the peptide probe pulled down wild-type and mutant forms of APC from extracts of cancer cell lines.


Harbury et al. (1998) described a series of peptides designed to form dimeric, trimeric, and tetrameric helical bundles with right-handed supercoils. To achieve right-handed coiled coils, rather than canonical left-handed structures, an HP pattern based on an 11-residue abedefghij repeat was used as a template. Combinations of the hydrophobic residues Ala, Ile, Leu, Val, allo-Ile, and nor-Val were considered for the a, d, and h sites. The
novel inclusion of the nonstandard residues was based on preliminary modeling studies. A computer algorithm was written to choose optimal combinations of these residues for the targeted dimer, trimer, and tetramer states. An additional novel feature of the design protocol, which advanced protein design in general, was the inclusion of some backbone flexibility in the modeling algorithm. Together, these novel features allowed the sequence of $a$, $d$, and $h$ sites to be optimized for the dimer, trimer, and tetramer states. Solution phase characterization of three corresponding peptides—RH2, RH3, and RH4—confirmed the designs as stable, cooperatively folded helical bundles of the correct oligomer state, though RH2 was only marginally stable. Most importantly for the confirmation of the design process, the structure of RH4 was described. This revealed a right-handed coiled-coil structure with core packing and other features that matched the design model in atomic detail. The structure of RH4 is compared with the left-handed coiled-coil tetramer GCN4-p-LI in Fig. 4.

On the theme of noncanonical coiled coils, Hicks et al. (2002) have investigated the effects of a range of inserts—of between one and seven Ala residues—at the center of an otherwise canonical designed four-heptad leucine-zipper system. All of the inserts were destabilizing, but the four-residue insert, which results in a heptad-hendecad-heptad-heptad (7-11-7-7) sequence, was the least destabilizing. Similarly, though for different reasons, Hodges and colleagues have also made and characterized canonical coiled coils as hosts for peptide inserts (Kwok et al., 2002).


Pandya et al. (2000) have described the design of a sticky-ended, or offset, leucine-zipper-based heterodimer (Fig. 5A). This is unusual because all known natural coiled coils are blunt ended, presumably as a
consequence of the hydrophobic effect, although Talbot and Hodges (1982) have commented on possible (though unintentional) promiscuous sticky-ended assembly in early coiled-coil designs. In the design from Pandya and colleagues, the free ends of the sticky-ended dimer were made complementary to promote longitudinal assembly into extended self-assembled protein fibers (SAFs). The features of the design included a core that promotes dimerization with $a = \text{Ile}$ and $d = \text{Leu}$, residues at $e$ and $g$ that promote sticky-end assembly, N-terminal halves of the peptides with $e = g = \text{Lys}$, and corresponding residues in the C-terminal halves that were Glu. However, the key novel feature of this design is the offset placing of complementary Asn residues at a C-terminal $a$ site in one peptide and an N-terminal $a$ site in the other. In the absence of a high-resolution structure, the design was confirmed using a combination of spectroscopy (circular and linear dichroism), electron microscopy, and X-ray fiber diffraction (Pandya et al., 2000). The peptides assembled into linear, nonbranched fibers tens of microns long and approximately 50 nm thick (Fig. 5B). The thickness of the fibers was unexpected—coiled-coil dimers are typically $\approx 2$ nm thick—and reflected assembly of 2 nm protofibrils into the matured fibers, which in turn indicated that the protofibrils with helically repeated building blocks were inherently sticky.

The SAF design has been built upon considerably over the past two years (Ryadnov and Woolfson, 2003a,b, 2004). Others are pursuing coiled-coil designs as a route to self-assembling fibrous structures. Indeed, the first example dates back to 1997 (Kojima et al., 1997). More recently, Kajava and colleagues have described the design and characterization of self-assembling fibers based on a novel pentameric coiled-coil building block (Kajava et al., 2004; Melnik et al., 2003; Potekhin et al., 2001), and Conticello and colleagues have described fibrous structures based on a

Fig. 5. (A) Molecular model for the designed SAF sticky ended heterodimer. The Lys and Glu residues at the $e$ and $g$ sites are blue and red, respectively; the buried, offset Asn residues at $a$ are green. Adapted from Pandya et al. (2000). (B) Negative-stain transmission electron micrograph image for fibers assembled from the sticky ended building blocks. Adapted from MacPhee and Woolfson (2004).
single, 42-residue leucine zipper-based helix (Zimenkov et al., 2004). On a related theme, Ghosh and colleagues have reported fibers produced by combining basic (EZ) and acidic (KZ) \( a = d = \text{Leu} \) coiled-coil peptides (Zhou et al., 2004). The novelty in this work was that each peptide was tethered to a core dendrimer (D) to give constructs D-EZ\(_4\) and D-EK\(_4\), which were then used in assembly. Along with other work—for example, in the development of coiled coil-based hydrogel systems (Petka et al., 1998; Wang et al., 1999)—these papers represent an exciting beginning for the application of coiled-coil design in the area of the rational design and exploitation of nanostructured biomaterials. Some of these papers and this field have been reviewed recently (MacPhee and Woolfson, 2004; Yeates and Padilla, 2002).


In three recent publications (Burkhard et al., 2000b, 2002; Meier et al., 2002), Burkhard and colleagues have described the design, redesign, and characterization of a two-heptad coiled-coil system, Suc-DELEARIRELARIK-NH\(_2\). This unit was stabilized by hydrophobic interactions (\( a = \text{Ile} \), \( d = \text{Leu} \)) and a network of salt bridge interactions. Initial characterization of the peptide showed that it was helical and dimeric. However, at increased ionic strength, the oligomer state switched to trimer. A structure for this form is available. Encouragingly, the introduction of improved saltbridge interactions, Suc-EELRRRIEELERRIR-NH\(_2\), did further specify the dimer state in solution, though the structure deposited in the Protein Data Bank for this peptide is still for the trimer. Likewise, removal of salt bridges, Suc-DELEARIRELARIK-NH\(_2\), resulted in the loss of coiled-coil specificity and the formation of a noncoiled-coil octamer. It is probable that the shortness of this design contributes to the change in oligomer state and overrides the influence core residues, (\( a = \text{Ile} \), \( d = \text{Leu} \)), which would normally specify dimer. In other words the relative stabilities of alternative structural states of the unit is probably marginal and, thus, small changes in the noncovalent interactions affect the equilibria between these states significantly.


Coiled-coil motifs have been known to play roles in conformational switching in natural proteins for some time (Oas and Endow, 1994). The key examples are influenza hemagglutinin (Bullough et al., 1994; Carr and Kim, 1993; Carr et al., 1997), and the heat shock transcription factor (Rabindran et al., 1993). Furthermore, an engineered form of GCN4-p1, with Asn-16 replaced by Ala, switches from dimer to trimer upon addition of
cyclohexane or benzene (Gonzalez et al., 1996b). In the presence of benzene, the mutant peptide crystallizes as a parallel trimer with a single benzene molecule bound at a cavity bordered by Leu-12, Ala-16, and Leu-19.

Ciani et al. (2002) have described the de novo design of Template-α, a 29-residue peptide with a canonical dimeric leucine zipper repeat, αAALEQK (where $a =$ Val, Lys and Ile). As expected, the peptide folded to a helical dimer that unfolded reversibly and in a concentration-dependent manner upon heating in solution. The peptide design was also compatible with an antiparallel $\beta$-hairpin structure. To further enhance $\beta$-sheet propensity, a mutant peptide, Template-αT (in which the Gln residues at the $f$ sites of Template-α were replaced by Thr) was made. Consistent with this dual-sequence design principle, thermal unfolding of the mutant resulted in the formation of $\beta$-structure as judged by CD spectroscopy, and amyloid-like structures as observed by transmission electron microscopy. On a similar note, Kammerer et al. (2004) have described a shorter designed coiled-coil peptide, ccβ, which crystallized as a trimer and transformed to amyloid-like structures upon heating. These systems present possibilities for assessing how $\alpha$-helical structures switch to $\beta$-sheet-based amyloid-like structures, and this may shed light on how natural proteins transform to amyloid in certain diseases.

Related to these ideas, Pandya et al. (2004) have described the design of an antiparallel coiled-coil (helix-loop-helix) peptide, which is stabilized by a disulfide bridge between the termini of the peptide. Reduction of the disulfide triggered a switch to a dimeric leucine zipper.


The design and engineering of systems to control the assembly of structures and functional modules with nanometer precision is one goal of the burgeoning discipline of nanotechnology. Designed coiled-coil systems offer great possibilities in this area for several reasons. First, coiled coils can be considered as relatively stiff nanoscale rods. Second, there is a precise relationship between peptide length and coiled-coil length, as each heptad meters out $\approx 1$ nm. Finally, coiled coils fold and self-assemble to stable structures at $\mu$M concentrations, or lower in water.

Ryadnov et al. (2003) have designed a coiled coil-based nanoscale linker system dubbed Belt-and-Braces. The system was novel in a number of respects. Though based on a leucine-zipper dimer design, it was a ternary system in which one peptide (the “Belt”) templated the assembly of two half-sized peptides (the “Braces”); thus, the system was the first and simplest example of a coiled-coil vernier assembly (Kelly et al., 1998). The Belt-and-Braces design employed all of the key design rules for a
parallel coiled-coil dimer: that is, \( a = \text{Ile} \), \( d = \text{Leu} \), and the \( g_n; e_{n+1} \)' pairs are Lys:Glu or Glu:Lys. Ternary complexation was directed by the Belt, which was a six-heptad acid peptide with the general \((g-f)\) heptad repeat, EIAALEQ, that templated the two three-heptad Braces \( B_N \) and \( B_C \), with the repeat sequence KIAALKQ. Specificity was achieved by complementary Asn residues at the fifth \( a \) site of the Belt and the central \( a \) site of \( B_C \). Finally, the N- and C-termini of \( B_N \) and \( B_C \), respectively, were extended by Cys-Gly-Gly- and -Gly-Gly-Cys units. The construct was expected to span 6 to 7 nm. The folding and assembly of the naked Belt-and-Braces peptides was confirmed by a combination of CD spectroscopy, analytical ultracentrifugation (AUC), and surface plasmon resonance (SPR) using Biacore. Moreover, the assembly of bound cargo was demonstrated as follows: the Cys-termini of the Brace peptides were labeled with 15 nm colloidal gold particles and mixed with the Belt. Transmission electron microscopy (TEM) visualization of the resulting assemblies revealed networks of nanoparticles separated by approximately 7 nm consistent with the design.

Stevens et al. (2004) reported a similar leucine-zipper-like linker system. This comprised two six-heptad peptides \((d = \text{Leu}, a = \text{Val}, \text{Ala}, \text{and Ile})\), one of which was basic with predominantly Lys at \( e \) and \( g \), while the other was acid with predominantly Glu at \( e \) and \( g \). Assembly was confirmed by CD spectroscopy and visualized by coupling the peptides to gold nanoparticles followed by TEM. A nice touch in this work was the use of different size nanoparticles to create binary assemblies including satellite structures, in which 8.5 nm particles (derivatized with the acid peptide) were organized around 53 nm particles (derivatized with the basic peptide).

The designs by Hodges and colleagues in the section on heterodimeric coiled-coil systems (Chao et al., 1998; Litowski and Hodges, 2001, 2002), and by Ghosh et al. (2000) in the next section, provide other examples of coiled-coil systems as potential peptide linkers.

B. Antiparallel Structures

With the exception of the discussion centered on the crystal structure of coil-Ser, which forms a mixed parallel/antiparallel trimer (Lovejoy et al., 1993), this review has so far ignored antiparallel coiled-coil structures. This is for a number of reasons. First, our understanding of antiparallel coiled coils is not as advanced as that for the parallel structures (Oakley and Hollenbeck, 2001; Walshaw and Woolfson, 2001b). Second, and as a result of the first point, fewer design attempts have targeted antiparallel coiled coils. Finally, there is a very good recent review on antiparallel coiled-coil structures and design (Oakley and Hollenbeck, 2001). Nonetheless, this Chapter would not be complete without the inclusion of these design
attempts. Beforehand, however, the main differences between antiparallel and parallel coiled-coil structures are summarized. Because of the relative dearth of experimental data on other antiparallel structures, discussion is limited to heptad-based antiparallel dimers that, like the parallel structures, have left-handed supercoils.

In order to maintain the register of hydrophobic residues between two antiparallel coiled-coil strands, $a$ residues from one strand must pair with $d$ residues of the other (Fig. 1). Thus, whereas in parallel dimers alternating core layers comprise $a:a'$ and $d:d'$ residues, antiparallel dimers have alternating $a:d'$ and $d:a'$ layers. On this basis, the stereochemical packing of mixed $a:d$ layers might be expected to differ from those in the symmetric layers of parallel structures. However, analysis of the known antiparallel dimer structures show that packing geometries of the $a$ and $d$ side chains can still be considered as roughly parallel and perpendicular, respectively (Walshaw and Woolfson, 2001b), although both distributions of angles are shifted slightly towards acute (trimer) packing. This is reflected in the amino-acid profile for antiparallel dimers (Walshaw and Woolfson, 2001b) (Table I). Briefly, this profile is a somewhat dampened version of the parallel dimer profile, with a preference for Leu at $d$ and Leu/Ile at $a$. The interchain interactions between the core-flanking $e$ and $g$ residues are, however, very different: in parallel structures $g_n$ and $e_{n+1}$ interactions are made across the interface, whereas in antiparallel structures it is $g:g'$ and $e:e'$. This provides the protein designer with a negative-design rule and a route to designing antiparallel coiled coils, avoiding alternative parallel structures.

1. The First Hodges Construct (1993)

As with the first design for a parallel coiled coil, Hodges and colleagues were also the first to describe a construct for an antiparallel coiled coil (Monera et al., 1993). This was based on the $(g-f)$ heptad repeat KLEA-LEG (Hodges et al., 1981). To construct the antiparallel system, several changes were made to the design. First, Leu-16, an $a$ site, was replaced by Ala. Second, two peptides were used, incorporating Cys at the first $a$ site (peptide C2A16) and the last $d$ site (peptide C33A16). Both peptides were five heptads long. The two peptides were mixed under denaturing conditions and air-oxidized to give a 1:1:2 mixture of the two forced-parallel homodimers and the forced-antiparallel heterodimer. Note that this construct did not consider electrostatic pairings between the $e$ and $g$ sites: the $g:g'$ and $e:e'$ pairs in the antiparallel construct are all repulsive, whereas the $ge'$ pairs in the parallel forms are attractive. Characterization of the forced-parallel and antiparallel constructs reflected this. With respect to heat and urea-induced denaturation, the forced-antiparallel dimer was less stable than either of the covalent parallel dimers; although the relative
stabilities with respect to guanidinium hydrochloride-induced denaturation were reversed, this is explained by a salt effect. Furthermore, when the covalent dimers were made under benign (folding) conditions, only the parallel dimers were formed; equilibration of the forced-antiparallel dimer with disulfide-exchange reagents returned the parallel dimers. Nonetheless, the antiparallel construct made was helical, cooperatively folded, and represented the first engineered antiparallel coiled coil. Furthermore, in the same paper, Hodges and colleagues recognized the error in their design and described a final construct in which charged pairs were optimized to favor the antiparallel heterodimer. They used these constructs in several subsequent papers to probe sequence-to-structure/stability relationships in parallel and antiparallel coiled coils (Monera et al., 1994, 1996a,b).


Myszka and Chaiken (1994) described the rational design of a single-chain antiparallel coiled-coil structure. This design was very well conceived. The peptide was 56-residues long and it had two 25-residue coiled-coil sequences separated by a flexible and functional (RGD-containing) six-residue loop. Except for the first \( a:d' \) layer (which is Cys:Cys), the \( d:a' \) and \( a:d' \) layers were Leu:Leu and Val:Val, respectively. Furthermore, the \( e \) and \( g \) sites in the N-terminal half were all Glu, whereas those in the C-terminal half were Lys to facilitate favorable \( g:g' \) and \( e:e' \) interactions. The unit was stapled by a disulfide bond between Cys-1 and Cys-56 at \( a \) and \( d' \) sites, respectively, and the remaining \( b, c, \) and \( f \) sites were Ala and Ser. CCSL was soluble, monomeric, approximately 80% helical, and unfolded cooperatively. Finally, the peptide was active, competing with fibrinogen for the GPIIbIIIa receptor.

Based on very similar principles, though more recently, Suzuki and Fujii (1999) designed a helix-loop-helix peptide. On a related theme, the aforementioned design from Pandya et al. (2004) for a helix-loop-helix peptide stabilized by a disulfide bridge used similar ideas, although the final sequence was very different from the Myszka and Chaiken and the Suzuki and Fujii peptides, as it was also made compatible with a parallel coiled-coil dimer to promote conformational switching.

3. An Antiparallel Variant of Peptide Velcro (1998)

Oakley and Kim (1998) have described a simple and elegant experiment to switch the aforementioned Peptide Velcro from a parallel heterodimer to an antiparallel structure. The strategy was straightforward and resulted in a powerful design rule for specifying antiparallel structures. The Acid-p1
and Base-p1 sequences of Peptide Velcro were redesigned so as to mismatch the Asn residues at their central $a$ sites; in Acid-a1 Asn is placed at the third $a$ site, while in Base-a1 it is at the second $d$ site. In the targeted antiparallel structure, these sites form an $acd$ layer. The solution phase characterization indicated that the favored complex was the antiparallel heterodimer. This study added to the rules for coiled-coil assembly and specification because an Asn:Asn pair at $acd$ was unprecedented in natural coiled-coil structures (Oakley and Hollenbeck, 2001). Oakley and colleagues have extended this work to improve the designs of Acid-a1 and Base-a1, using both positive and negative design principles to improve the $gg'$ interactions to favor the antiparallel orientation over the competing parallel arrangement (McClain et al., 2001).


A number of proteins are known that can be split into two fragments, which when recombined form functional, though usually less stable, folded structures (Michnick, 2001). This is known as fragment complementation. For proteins that cannot be split and reconstituted so straightforwardly, there is the possibility that reassembly might be directed by additional reagents such as ligands. Such systems present a potential route to new biosensors.

Ghosh et al. (2000) have described an elegant system in which two fragments of the green fluorescent protein (GFP) from Aequorea victoria could be recombined only in the presence of helical tags engineered at the new termini created by the split. The helices were two halves, NZ and CZ, of a designed antiparallel coiled-coil dimer. The peptides were 29 and 30 residues long respectively. The cores were all Leu, except the second $d$ site of NZ and the third $a$ of CZ, which were Asn. All of the potential $ee'$ and $gg'$ pairs were Lys:Glu, and the remaining $b$, $c$, and $f$ sites were combinations of Ala, Gln, and Lys with a single Trp chromophore at an $f$ site in each peptide. Reconstitution of GFP was demonstrated both in vitro using the purified protein fragments, and in vivo by co-expression of the two components in E. coli.


The designs outlined above are effective all for heterodimers. The design of a homodimeric antiparallel coiled-coil is more challenging. For instance, specification via the burial of two Asn residues, one at $a$ and the other at the corresponding $d'$ site, is not possible. In a thoughtful design, Oakley and colleagues tackled this problem to generate APH (Gurnon et al., 2003). APH is a 45-residue recombinant peptide
overexpressed in *E. coli*. The design had the following key features: although the core was predominantly Leu, there were two Ala:Ile pairs and two Leu:Arg pairs at \( a:d \) layers. The argument for using the former pair was that the small Ala side chain helps accommodate Ile in the antiparallel form. The inclusion of Arg within the core followed work by the groups of Oakley and of Lumb indicating that this provided specificity at a smaller energetic price than buried Asn residues (Campbell et al., 2002; McClain et al., 2002). On the basis of amino-acid profiles (Walshaw and Woolfson, 2001b; Woolfson and Alber, 1995) and experimental studies (Gonzalez et al., 1996c), these inclusions probably also destabilize potential parallel dimers and trimers, respectively. Finally, to favor the antiparallel form over alternative parallel structures, all potential \( ee' \) and \( gg' \) interactions were made into attractive Glu:Lys pairs, whereas potential \( gn:en+1' \) interactions were made repulsive. The design held up to scrutiny by solution-phase biophysics.

IV. Summary

This review has attempted to build on the preceding Chapters that convey the richness of coiled-coil sequences and structures. With this backdrop of many potential coiled-coil structures and very many potential coiled-coil sequences, one aim of this Chapter has been to illustrate that sequence-to-structure relationships in coiled coils can be understood sufficiently to attempt with some confidence and, hence, with some degree of success, the rational design of coiled-coil structures. At present, this process is largely done on the backs of envelopes guided by good rules of thumb, intuition, and imagination although, and as noted, computer-aided designs are increasingly being reported. Nonetheless, the back-of-the-envelope approach has had many successes from designs for basic, naturally observed structures to more-imaginative ones that might be used as probes, tethers, molecule rulers, affinity matrices, and components of new materials. In these respects, coiled-coil design is arguably the most successful of all protein-design areas. As well as continuing such efforts, one challenge in the near future will be to formalize the design rules by linking them to thermodynamic measurements and employing them further in prediction and computer-aided design algorithms. The future of coiled-coil design is very exciting indeed.

Acknowledgments

The author thanks Andrei Lupas, Markus Gruber, David Parry, and members of my research group for their critical reading of and constructive comments on this manuscript.
References


