Coiled Coil Domains: Stability, Specificity, and Biological Implications

Jody M. Mason and Katja M. Arndt*[a]

Introduction

The coiled coil is a common structural motif, formed by approximately 3 – 5% of all amino acids in proteins.[1] Typically, it consists of two to five α-helices wrapped around each other into a left-handed helix to form a supercoil. Whereas regular α-helices go through 3.6 residues for each complete turn of the helix, the distortion imposed upon each helix within a left-handed coiled coil lowers this value to around 3.5. Thus a heptad repeat occurs every two turns of the helix.[2, 3] The coiled coil was first described by Crick in 1953.[4] He noted that α-helices pack together 20° away from parallel whilst wrapping around each other, with their side chains packing "in a knobs-into-holes manner". The same year, Pauling and Corey put forward a model for α-keratin.[5] It was some 20 years later that the sequence of rabbit skeletal tropomyosin was published,[6] and another twenty until the first structure of the leucine zipper motif was solved by Alber and co-workers.[7] These last discoveries pushed the coiled-coil field into the spotlight, as it became apparent that they are found in important structures that are involved in crucial interactions such as transcriptional control. The most commonly observed type of coiled coil is left-handed; here each helix has a periodicity of seven (a heptad repeat), with anywhere from two (in designed coiled coils)[8] to 200 of these repeats in a protein.[9] This repeat is usually denoted (**a-b-c-d-e-f-g**)$_n$ in one helix, and (**a′-b′-c′-d′-e′-f′-g′**)$_n$ in the other (Figure 1). In this model, **a** and **d** are typically nonpolar core residues found at the interface of the two helices, whereas **e** and **g** are solvent-exposed, polar residues that give specificity between the two helices through electrostatic interactions. Similarly in right-handed coiled coils, an eleven-residue repeat is observed (undecatad repeat).[10, 11] The apparent simplicity of the structure with its heptad periodicity has led to extensive studies. Here we aim to outline the importance of individual amino acids in maintaining α-helical structure (intramolecular interactions) within individual helices, whilst promoting specific coiled-coil interactions (intermolecular interactions) of correct oligomeric state and orientation.
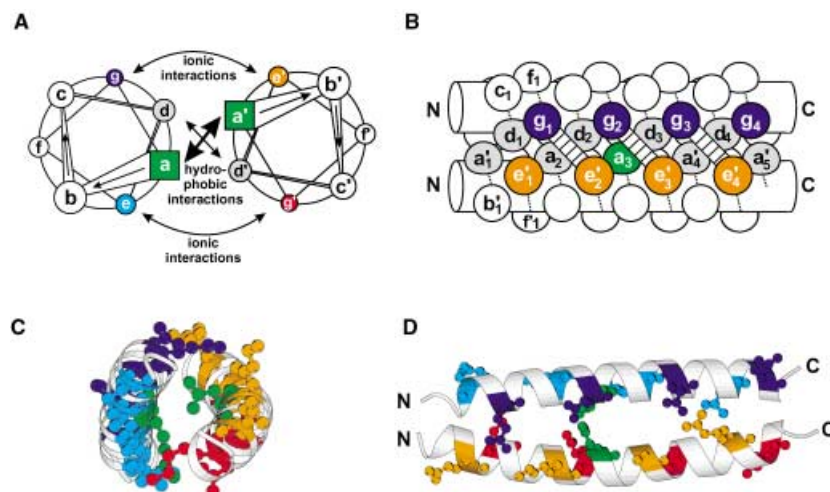

Figure 1. *A parallel dimeric coiled coil in a schematic representation (A and B) and as ribbon plot of the X-ray structure of the leucine zipper of GCN4[7] (C and D). Selected side chains are shown as balls and sticks. The helical wheel diagram in (A) and the plot in (C) look down the axis of the α-helices from N-terminus to C-terminus. Panel (B) and (C) provide a side view. The residues are labeled **a** – **g** in one helix and **a′** – **g′** in the other. The hydrophilic interactions (**g** and **g′** in blue and red, respectively; **e** and **e′** in cyan and orange, respectively) within the heptad repeat are shown. In the schematic representations, the hydrophobic core (**a/a′** and **d/d′**) is shown. For clarity, in the X-ray structure, only the middle **a** position with the exceptional charged residue is given as a green ball-and-stick model. Parts C and D were generated with molscript.[77]*

The PV Hypothesis

The PV ("Peptide Velcro") hypothesis[12] outlines three structural elements vital to the formation of a specific coiled coil. It contains one of the earliest rational design strategies for the formation of heterodimeric coiled coils and was originally used by O'Shea and co-workers.[13] Firstly, it stipulates that the **a** and **d** positions must be hydrophobic (e.g. leucine, valine, or isoleucine), thus stabilizing helix dimerization through hydrophobic and van der Waals interactions. Secondly, residues **e** and **g** must be charged (e.g. glutamate or lysine) in order to form interhelical electrostatic interactions. Such interaction patterns should be of the opposite charge in heterodimers to stabilize their interac-

[a] Dr. J. M. Mason, Dr. K. M. Arndt
    Institut für Biologie III, Albert-Ludwigs-Universität Freiburg
    Schänzlestraße 1, 79104 Freiburg (Germany)
    Fax: (+ 49) 761-203-2745
    E-mail: katja@biologie.uni-freiburg.de

tion, and of the same charge in homodimers to destabilize them. Thirdly, the remaining three positions (**b**, **c**, and **f**) must all be hydrophilic, as these will form helical surfaces that are exposed to the solvent.[14, 13] However, the PV hypothesis is only a guide for the interactions governing left-handed coiled coils, and it is subtle variations of these rules that dictate the orientation, specificity, and oligomerization state that are required for a domain to have a novel function.

The Role of the "a" and "d" Residues

The nonpolar nature of the **a** and **d** repeats facilitates dimerization along one face of each helix. This interaction was first suggested in Crick's 1953 paper[4] in which he proposed a "knobs-into-holes" style packing between the hydrophobic side chains. This is analogous to a hydrophobic core that collapses during the folding of globular proteins, and represents a dominating contribution to the overall stability of the coiled coil. Indeed, the most stable coiled coils are those that have the highest percentage of hydrophobic residues at the **a** and **d** positions.[15] Furthermore, variations in packing environments give different preferences for hydrophobic residues, even within the **a** and **d** positions. For example, GCN4 (a yeast transcription factor with a parallel homodimeric coiled-coil (leucine-zipper) domain, sometimes referred to as GCN4-p1) has β-branched side chains, such as valine, that pack well at position **a**, while position **d** favors a γ-branched leucine residue.[7] The insertion of a β-branched amino acid into the **d** position would require adoption of a thermodynamically unfavorable rotamer in the parallel dimer.[16] If these preferences are denied, for example, in GCN4 mutant p-LI, in which a leucine is introduced at **a** and an isoleucine at **d**, then a tetramer is formed.[17] Valine at these positions leads to a mixture of dimer and trimer, while all leucine leads to tetramer formation. Finally, in two exhaustive studies, the **a** and **d** positions were systematically changed to every amino acid to assess their effects on stability and oligomerization states.[18, 19] These changes were the first comprehensive quantitative assessment of the effect of side chain substitution within the hydrophobic core on the stability of two-stranded coiled coils, and permitted a relative thermodynamic stability scale to be constructed for the nineteen naturally occurring amino acids in the **a** and **d** positions.

Studies where the **a** and/or **d** residues have been changed to non-natural amino acids that are even more hydrophobic than naturally occurring ones (e.g. 5,5,5-trifluoroleucine) revealed a further increase in stability.[20] More recently hydrophobic burial at the **a**/**d** interface has been investigated by using mono-, di-, and trimethylated diaminopropionic acids (dap), which display increasing degrees of hydrophobic character. Addition of one methyl group to position 16 of one of the monomers (with aspartic acid at position 16 in the analogous peptide), was found to stabilize the subsequently heterodimeric fold of GCN4, possibly due to increased van der Waals interactions in the folded state and a lower desolvation penalty upon folding. However, addition of three methyl groups results in destabilization, probably because the increased steric bulk is poorly accommodated. Bizarrely, the addition of two methyl groups to

the dap causes homotrimerization. This demonstrates how small changes in hydrophobicity can alter the folding preferences.[21]

Despite the hydrophobic nature of the **a**/**d** interface, often a small percentage of polar core residues remain and add specificity to the coiled coil at the expense of stability. In GCN4, an asparagine is located at a core **a** position. When this is mutated to a valine, the coiled coil experiences a huge increase in stability at the expense of dimerization specificity. It actually leads to a trimer with increased stability compared with the wild-type dimer.[22] In another case, the core asparagine pair was again mutated to leucine; this changed the original dimeric peptide Velcro (PV) to a mixture of parallel and antiparallel tetramers.[23] Changing the core asparagine to lysine retains specificity but lowers stability further, while substituting asparagine for norleucine (lysine without the charged amino group) stabilizes, again at the expense of specificity.[24] It is when a conflict occurs between inherent secondary structure propensities and the repeat pattern of hydrophilicity/hydrophobicity that such problems arise. Usually, it appears that this change in binary pattern will dictate the overall structure of the coiled coil.[25] Nonetheless, this asparagine confers dimer specificity, possibly through interhelical hydrogen-bond formation between asparagine side chains; this is indeed observed in the crystal structure[7] and in NMR studies.[26] In addition, during selection for heterodimeric coiled coils with a protein-fragment complementation assay (PCA) by using dihydrofolate reductase (DHFR), Arndt et al. found a core asparagine pair to be favored over asparagine – valine or valine – valine combinations.[27, 28] This is in agreement with many naturally occurring coiled coils. The strategic placement of the core asparagine pair can also direct coiled-coil association from parallel to antiparallel, by changing the position of this buried polar association.[29] Finally, the Matrilins, involved in the development and homeostasis of bone, constitute a family of four oligomeric proteins that are able to form homo- and heterotypic structures of differing oligomeric states, depending on the isoforms.[30] However, all observed oligomers fold into parallel disulfide-bonded structures, and heterotypic preferences have been attributed to core changes rather than ionic interactions. It is such changes in heterotypic core contact that also permit the generation of heterospecificity in coiled coil pairings.[31]

The Role of the "e" and "g" Residues

Pairing specificity is greatly influenced by the nature of the electrostatic **e** and **g** residues (between **g** of one heptad and **e**′ of the following heptad on the other helix, termed $i \rightarrow i' + 5$). These residues are commonly found to be glutamic acid and lysine, respectively. Thus, the charge pattern on the outer contacting edges of a coiled coil will dictate its preference for homo- or heterotypic pairing, and whether the orientation of the coiled coil is to be parallel or antiparallel. Replacing attractive **g**/**e**′ pairings with repulsive pairs has been shown to destabilize the coiled-coil conformation.[32] Hodges and co-workers estimated the salt bridges between **g**/**e**′ pairs to contribute 1.5 kJ mol⁻¹ to the stability of the coiled coil.[33] Careful placing of charges within the **e** and **g** positions can permit heterodimer formation, while

additionally ensuring that formation of the homodimer is unfavorable[14, 13, 32] as implied in the PV hypothesis.

Arndt et al. have 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 valine and leucine (with the exception of some asparagine variants in the core to direct desired helix orientation and oligomerization state), and the **e** and **g** residues are varied by using trinucleotides to give equimolar mixtures of arginine, lysine, glutamine, and glutamate.[27, 28] Unexpectedly, even the best one, the Winzip-A2B1 heterodimer, lacks fully complementarily charged residues at **g**/**e**′ pairs despite an exhaustive selection process.[12] Rather, two of the six **g**/**e**′ pairs are predicted to be repulsive; this suggests that sequence solutions that deviate from the PV hypothesis might be tolerated in heterodimeric coiled coils, and that other factors might play a role in selection. Clearly the PV hypothesis does not represent a complete picture of the contributions of the **e** and **g** residues to dimer stability and specificity. Presumably, overall electrostatic potential (including intra- 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.[28, 12and references therein] Such observations are in agreement with naturally occurring coiled coils, which usually have a more complicated interaction pattern than implied by the PV hypothesis. These coiled coils have to fulfill a number of criteria, such as biostability and extremely high specificity within a family with almost no cross reactivity with coiled coils of other families.[34] These requirements can only be realized by more complicated networks of interaction patterns that also include the outer residues and interactions between **d** – **e**′ and **a** – **g**′ residues.[35] A buried polar **a**-position residue, such as lysine, with no preference for helix orientation, for example, can form favorable interactions with **g**′ or **e**′ glutamate in the parallel or antiparallel arrangement, respectively. Such an interaction is less destabilizing than an asparagine – asparagine (**a** – **a**′ contact), presumably because there is a greater desolvation penalty to pay for burying the latter. In this way, complementary interactions between a buried polar residue, and a surface polar group, can give structural uniqueness and incur smaller energetic penalties.[36]

Harbury's group focused on both positive (toward the desired structure) and negative design (away from undesired alternate structures) in ensuring dimer specificity.[35] Computational design is followed by experimental verification, and functional specific protein – protein recognition sequences are produced. The algorithm uses a computational double-mutant cycle to permit sequences that have an energetic preference for the target state over the negative design states. This procedure is known as "multistate" design. The benefit of such a model is that the patterning of hydrophobic and polar residues arises from simultaneous competition against unfolded and aggregated states. Consequently, polar residues are not excluded from the core of proteins. Instead, their selection is based on an energetic balance between the requirements for stability and specificity.

Finally, by placing charge pairs at **g** – **g**′ and **e** – **e**′ positions, antiparallel helix orientation can be favored.[37] In a three-helix coiled coil there is a much higher percentage of nonpolar

residues at the **e** and **g** positions. This increase in the percentage of hydrophobes at the **e** and **g** positions causes the width of the narrow hydrophobic face to increase, as well as the likelihood of such higher oligomerization states, where more nonpolar burial can occur than in a two-helix coiled coil. This demonstrates the importance of the **g**/**e**′ residues in determining homo- and heterotypic pairing, parallel or antiparallel orientation, and the oligomerization states of α-helical chains in coiled coils.

Stutters and Stammers in Coiled Coil Heptads

Breaks in the periodicity of the heptad repeat are known either as a "stutter" or a "stammer". A stutter (sometimes called a skip) corresponds to a three-residue deletion (or four-residue insert) and is compensated for by an underwinding of the supercoil. A stutter can therefore be regarded as a region that is right-handed in character (see Right-Handed Coiled Coils section), as the coil region undercoils to maintain its hydrophobic contacts. In contrast, a stammer corresponds to a deletion of four residues (or an insertion of three) and must be compensated for by an overwinding of the supercoil. Such distortions are generally confined to two α-helical turns either side of the deletion. Changes in local structure caused by under- and overwinding of the coiled coil may function to terminate the structure, or could account for flexibility of long coiled-coil domains such as myosin.[38] Woolfson and colleagues have studied a cytoskeletal coiled-coil protein, HPSR2, found in *Giardia lambia*, in which heptads are found flanking undecatads (a stutter) to give a 7-11-7 motif. Specifically, a synthetic peptide based on this consensus sequence was constructed and found to form fully helical, parallel dimers. Within the undecatad repeat, a 3,4,4 hydrophobic repeat is found that is an extension of the 3,4 heptad repeat. This combination of three- and four-residue intervals is a prerequisite for coiled coil packing.[39] It is possible that the function of 7-11-7 motifs is to give specificity to such extended coiled-coil structures during folding within the cell.

Specificity—Two-, Three-, Four- and Five-Stranded Coiled Coils

It seems that whilst the α-helix itself is quite resistant to conformational change by mutation, the same cannot be said of the tertiary and quaternary structures of the molecule. It is this under-specification of the orientation and oligomeric state of the coiled coils that makes in vitro design of these proteins a daunting task.[40] The nature of structural specificity is more complicated as there is no single parameter to describe it. Small changes can alter tertiary and quaternary structure, for example, changing a two-stranded helix to a three- or four-stranded one, by changing the sets of buried hydrophobic residues at the core **a** and **d** positions;[17] this indicates that the shape of the hydrophobic side chain is an important determinant. A study by Zeng and colleagues changed the oligomerization state of GCN4 by alteration of the **g**/**e**′ pairings.[41] Particularly, by varying specific **g**/**e**′ residues, they were able to increase the oligomerization state, presumably by widening the hydrophobic interface available. Alber's group has engineered a mutant of GCN4 that is

able to switch from a dimer to a trimer (known as an oligomeric switch) upon binding cyclohexane and benzene.[42] Additionally, a benzene was bound to the core of the trimer in the crystal structure. This indicates the importance of core packing in oligomeric specificity.[43] Strangely, in the Matrilin family of proteins, hetero-oligomeric coiled-coil domains consisting of two different polypeptides are able to exist in different stoichiometries.[30] It seems likely, therefore, that Matrilin chain combinations are controlled by gene expression levels, based on observed tissue-distribution levels. Finally, by changing the attraction/repulsion pattern, McClain et al. were able to change a parallel dimer into an antiparallel dimer.[44] Other more general agents, well known for affecting the properties of the solvent, such as the chaotropic agent guanidinium chloride or cosmotropic salts, can cause folding or unfolding and changes in stability of the protein. Such agents will therefore affect the specificity of the coiled coil.[45]

The Cartilage oligomeric matrix protein (COMP) belongs to the thrombospondin family and contains an extremely stable five-stranded parallel α-helical coiled coil. The 46-amino-acid-long coiled-coil region includes a ring of intermolecular (i.e. helix to helix) disulfide-bonded cysteines.[46] The pentameric interface displays "knobs-into-holes" packing, with the knobs formed by **a**, **d**, **e**, and **g** positions and packing into holes created between side chains at positions $\mathbf{a'-g'}$, $\mathbf{d'-e'}$, $\mathbf{c'-d'}$, and $\mathbf{a'-b'}$ of the adjacent subunit. Only residues at position **f** remains completely exposed, with the other six positions being significantly buried. Thus the structural stability is largely a result of hydrophobic interactions. A hydrophilic ring of hydrogen bonds formed by the amide groups of Gln54 also plays a role. They may be functioning as an "ion trap" for binding chlorine.[47] The hydrophobic core is filled with water molecules and almost completely lined with aliphatic residues. Although COMP appears not to be an ion channel, according to homology studies it does seem to have evolved from one.[46] It has also been found capable of complexing with vitamin D(3) and retinoic acid, mediated by the hydrophobic core pattern. When bound, the rotamer angles of β-branched core side chains require reorganization to adapt to this bulky ring system.[48] This could form part of a storage and delivery function within the coiled-coil domain of COMP for signaling molecules relevant in cartilage tissue. Additionally, COMP is stable from 0 to 100 °C, and thermal denaturation is only achieved with 4 – 6 M guanidine hydrochloride (extrapolation to 0 M GuHCl places the T_m at 160 °C) ranking it among the most stable of proteins.[49]

Right-Handed Coiled Coils

There are 3.6 residues per turn in an α-helix, and as explained in the Introduction, this number is reduced to 3.5 in a left-handed coiled coil. This means a repeat of seven residues every two turns of the helix because the helices form a left-handed supercoil around each other to maintain contact along the hydrophobic face. In contrast, right-handed coiled coils slightly increase the number of residues to 3.67 per turn (in the opposite direction) to give a right-handed supercoil and eleven residues every three turns of the helix. This is the undecatad repeat, and residues are

labeled **a** through to **k** accordingly. There are very few known native right-handed coiled-coil proteins, one of these, tetrabrachion from *Staphylothermus marinus*, forms a parallel tetramer. This protein differs from left-handed parallel tetramers in that the core is larger and filled with water molecules. Consequently the hydrophobic packing is very different.[11] Another study used de novo design of helical bundle proteins to give a right-handed superhelical twist.[10] In this study, the overall fold was specified by the polar hydrophobic repeat pattern (positions **a**, **d**, and **h** all fall on the same face of the helix and can specify a hydrophobic repeat pattern characteristic of a right-handed coiled coil). The oligomerization state and core packing were engineered by using computational enumerations of packing in alternate backbone structures. Main-chain flexibility was incorporated through an algebraic parameterization of the backbone. The authors were able to successfully create dimers, trimers, and tetramers, and the crystal structure of the tetramer matched the designed structure in atomic detail.

Is the Folding of Coiled Coils a Two-State Process?

One of the most dominant forces involved with protein folding is "hydrophobic collapse". That is to say that a protein will fold so as to maximize the amount of nonpolar material that is buried within its core.[50] Coiled coils have a nonpolar/polar periodicity, and it is this amphipathic nature that drives two or more to associate at their hydrophobic face. The most widely accepted model for folding is a two-state transition from the unfolded monomers to the dimeric coiled coil.[51, 52] Folding studies of most dimeric coiled coils characterized so far show the folding and dimerization to be coupled and cooperative, and best described by a two-state model.[53, 54] However, it has been shown that small changes in sequence are able to change the two-state to a three-state mechanism.[55, 56] Specifically, Dragan and Privalov observed several stages in the temperature-induced unfolding of GCN4 coiled coil using a variety of techniques.[56] The first transition at the beginning of heating corresponds to a decrease in ellipticity and is sensitive to N-terminal modifications. The second transition occurs at much higher temperatures, is more pronounced than the first, and is sensitive to modification of both termini. It is only later (at higher temperatures) that cooperative unfolding/dissociation of the two strands occurs.

In contrast, in the two-state model (a parallel dimeric leucine-zipper model), the rate-limiting step involves an electrostatically stabilized, dimeric intermediate that is not well structured despite a large amount of hydrophobic surface burial. Any subsequent folding to the dimer is rapid and follows a downhill free-energy profile.[57] That is, the folding is a result of collisions between unstructured monomers with helix formation occurring only after collapse. This folding process is enthalpic in origin, and is opposed by an entropic loss as the structure becomes ordered.[58] Others report that partial helix formation precedes dimerization,[59] and that initiation sequences within these helices are indispensable for correct association and folding. In this diffusion – collision model, peptides with partial secondary-

structure formation can collide to productively fold to the native state, whilst those less helical diffuse apart as monomers. In this model, the rate-limiting step would be dependent on both peptide concentration and intrinsic helicity. Initiation sequences within secondary structures in this model would clearly be important in determining the folding kinetics. A different study by Moran and colleagues looked at the folding of the leucin-zipper region of GCN4 both as monomers (reduced), and tethered by a disulfide bridge (oxidized).[60] The authors found minimal helical content prior to productive collision of the two chains by multiple routes within the reduced sample. Conversely, the monomeric sample folded by one robust pathway.

Are Trigger Sequences Required for the Initiation of a Coiled Coil?

It has been reported that sequences within the coiled coil could be responsible for initiation of its formation. Within GCN4, cortexillin I, myosin, kinesin, and tropomyosin, a thirteen-residue sequence has been found to be required for initiation of the coiled-coil assembly.[61] The authors propose that this thirteen-residue sequence is an autonomous helical folding unit that can mediate coiled-coil formation, and that favorable intrahelical interactions within this sequence play an important role. They claim that these "trigger sequences" are necessary to mediate proper assembly and are of particular importance in the light of de novo design. More recently the same group has looked at "trigger sequences" within three-stranded coiled coils. Deletion mapping identified a seven-residue sequence that was necessary for proper coiled-coil formation, that is to say, heptad repeats alone may not be enough to promote oligomerization.[62] It appears that the "trigger sequence" forming an α-helix early within the folding process is able to act as a seeding event, by limiting the number of possible conformations available to the chain and, further, by acting as a scaffold around which the remainder of the coiled coil structure can "zip up". In the case of dimer formation, two initiation sites may again be able to interact, and the formation of the remainder of the coiled-coil will follow. Despite this, some known "trigger sequences" have a low helical content; this means they cannot be identified on the basis of helical propensity alone and shows that the helical content of a given heptad repeat is not sufficient to mediate specific coiled-coil formation.[62]

Despite such evidence for coiled-coil initiation sites, their existence, or at least their absolute necessity for folding, remains controversial. Many designed coiled-coil peptides do not have such a postulated trigger sequence. Moran and co-workers, using monomeric (crosslinked) and dimeric (non-crosslinked) samples of GCN4-p1, concluded that the major fraction of nucleation sites need not occur in the most-helical region of the molecule. On the contrary, the highest helical propensity region of the molecule is the last to fold.[60] In another paper by Arndt et al.,[12] coiled-coils were selected with a protein-fragment complementation assay (WinZip-A1, -B1, -A2 and -B2) and characterized together with rationally designed peptides VelA1 and VelB1. Interestingly, of all six sequences, only WinZip-A2 has

a potential trigger sequence that follows the scheme proposed by Kammerer et al[61]:

L(d)-E(e)-x(f)-c(g)-h(a)-x(b)-c(c)-x(d)-c(e)-c(f)-x(g)

namely, WinZip-A2: L-E-S-E-V-Q-R-L-R-E-Q (here x, c, and h are random, charged, and hydrophobic residues, respectively). However, the postulated intrahelical interaction between the first **e** and the second **f** position ($i \rightarrow i+8$) is also missing. Nevertheless, all peptides folded rapidly and reversibly, and formed homo- and heterodimeric coiled coils with dissociation constants (K_D) as low as 2.3 nM. Finally, Lee et al. designed a 31-residue coiled-coil domain hybrid sequence based on GCN4 and cortexillin I, both of which contain no consensus trigger sequence and no appreciable secondary structure, but do contain stable residues in the core **a** and **d** positions.[63] Changes were introduced to positions other than **a** and **d** to affect α-helical propensity, electrostatics, and hydrophobicity. None of these changes brought the peptide in closer agreement to the consensus trigger sequence, but did increase coiled-coil folding and stability. The authors suggested, therefore, that the combination of stabilizing effects along a protein is a more general indicator of folding and stability than identification of any specific trigger sequences. Finally, it is possible that known consensus sequences are too restrictive to generalize upon. That is to say that trigger sequences in proteins currently thought to have no trigger sequence are being missed, or are not critical for initiation.

Potential Uses of Coiled-Coil Domains and Biological Implications

Coiled coils are abundant structures found in a diverse array of proteins, from transcription factors such as Jun and Fos,[64] involved in cell growth and proliferation, to Matrilins, involved in the development of cartilage and bone.[30] In order to identify coiled-coil structures, several computational programs have been developed. These are aimed at predicting coiled-coil regions, the likelihood that helices will form a coiled coil, and the oligomeric state. "SOCKET"[65] defines the beginning and end of coiled-coil motifs and assigns a heptad register to the sequence (http://www.biols.susx.ac.uk/Biochem/Woolfson/html/coiledcoils/socket/). "COILS"[66] compares a protein sequence to a database of known two-stranded parallel coiled-coil structures and then computes the probability that the sequence will adopt a coiled coil structure (http://www.ch.embnet.org/software/COILS_form.html). "PAIRCOIL"[67] predicts the location of coiled-coil regions in amino acid sequences (http://paircoil.lcs.mit.edu/cgi-bin/paircoil), and "MULTICOIL"[1] locates dimeric and trimeric coiled-coil sequences based upon the paircoil algorithm (http://multicoil.lcs.mit.edu/cgi-bin/multicoil). Such programs can enable coiled-coil domains to be screened for within protein sequences, and inferences can be made about their oligomeric state and function. Whilst the COILS and MULTICOIL programs are accurate at predicting the oligomeric state of coiled coils, eligible methods for partner prediction, orientation within the bundle, and register relative to other helices, for true tertiary and

quarternary structure predictions are perhaps missing from them.[68] The following section looks at some of the ways in which coiled coils are being exploited to function as temperature regulators, antibody stabilizers, anticancer drugs, purification tags, hydrogels, and linker systems. These are but a few of the ways in which the structural uniqueness of coiled coils has been harnessed for therapeutic, biological, or nanotechnological benefit.

TlpA from *Salmonella* contains an elongated homodimeric coiled coil, and has been shown to function as a temperature-sensing gene regulator. Naik and colleagues have demonstrated the thermal folding to be rapid, reversible, and sensitive to changes in temperature.[69] By coupling the protein to green fluorescent protein (GFP), the authors were able to use GFP fluorescence changes as an indicator of the folding and unfolding of the TlpA fusion protein. Specifically, the GFP acts as a fluorescent indicator of the structural transitions that occur within the TlpA coiled-coil homodimer in response to temperature changes. This has important implications for the measurement of signal transduction processes involving dimerization of coiled-coil domains.

By fusing an in vivo-selected coiled coil (WinZip-A2B1)[12] to the Fv fragment of an antibody, a helix-stabilized antibody fragment (hsFv) was constructed. This hsFv fragment had similar expression, purification, stability and oligomerization properties to other Fv constructs.[70] Additionally, it is postulated that through design coupled with in vivo selection, it will be possible to create coiled coils that have the correct expression, protease sensitivity, localization, pairing kinetics, and interactions with targeted cellular proteins, to exert in vivo function.[12] For example, Sharma and co-workers have designed a peptide (anti-APCp1) that is targeted to bind with a coiled-coil sequence from the adenomatous polyposis coli (APC) tumor-suppressor protein that is implicated in colorectal cancers.[31] Another publication described the design a of a dominant negative coiled-coil peptide that heterodimerized with a bZIP transcription factor to prevent DNA binding. The transgenic expression of this construct in mice demonstrated its functionality in vivo.[71]

Coiled coils have been used as tags in the expression and purification of other proteins. In such systems, the target protein is expressed as a fusion with one of the coiled-coil strands and purified by using an affinity chromatography column derivatized with the complementary coiled coil.[72, 73] However, a limitation of this technique is the high stability of the coiled-coil heterodimer; this means that the elution buffer needs to be of low pH (to disrupt electrostatic interactions) and contain acetonitrile (to disrupt hydrophobic interactions), both of which can be damaging to the protein. With that in mind, Litowski and Hodges modified the E/K coiled coil to lower its stability, by reducing the length of the coiled coil, and wisely changing only the **g/e**′ charge pattern to repulse homodimers, without the compromise of lowered heterospecificity.[45]

Wang et al. have used coiled coils, covalently bound to water-soluble synthetic polymers, that can undergo temperature-induced collapse, thus acting as a hydrogel with engineered volume-changing properties.[74] This is due to the cooperative conformational transition inherent to the coiled coil. The authors go on to suggest that different types of coiled coils could be used in the same system, to result in gels capable of stepwise transitions in volume, with each step triggered at a different temperature (or possibly pH change/change in ionic strength). The idea is that these can then be used as a potential drug-delivery system.[75]

Finally, Ryadnov et al. have used coiled coils as a peptide-based linker system.[76] Three leucine-zipper sequences, known as "belt and braces", contain one free peptide (the "belt") that can template the assembly of the other two peptides (the "braces"), to give a 1:1:1 specific self-assembling and thermally stable complex. This could potentially be utilized in peptide-directed immobilization at surfaces, or in guiding nanoparticle assembly.

Clearly coiled coils are important structural motifs involved in a variety of important interactions that have the potential to be biochemically and therapeutically exploited. However, it appears that a better understanding of the energetics and interactions that drive specific associations of coiled coils (together with design options for delivery to the site of action) are necessary before such strategies become realistic. However, with the fast-paced nature of progress in this field, designed coiled-coil peptides are likely to be found in a wide variety of in vitro and in vivo biochemical applications within the near future.

Acknowledgements

[1] E. Wolf, P. S. Kim, B. Berger, *Protein Sci.* **1997**, *6*, 1179–1189.
[2] W. H. Landschulz, P. F. Johnson, S. L. McKnight, *Science*. **1988**, *240*, 1759–1764.
[3] A. Lupas, *Trends Biochem. Sci.* **1996**, *21*, 375–382.
[4] F. H. S. Crick, *Acta Crystallogr.* **1953**, *6*, 689–697.
[5] L. Pauling, R. B. Corey, *Nature*. **1953**, *171*, 59.
[6] J. Sodek, R. S. Hodges, L. B. Smillie, L. Jurasek, *Proc. Natl. Acad. Sci. USA* **1972**, *69*, 3800–3804.
[7] E. K. O'Shea, J. D. Klemm, P. S. Kim, T. Alber, *Science* **1991**, *254*, 539–544.
[8] P. Burkhard, M. Meier, A. Lustig, *Protein Sci.* **2000**, *9*, 2294–2301.
[9] W. D. Kohn, C. T. Mant, R. S. Hodges, *J. Biol. Chem.* **1997**, *272*, 2583–2586.
[10] P. B. Harbury, J. J. Plecs, B. Tidor, T. Alber, P. S. Kim, *Science* **1998**, *282*, 1462–1467.
[11] J. Stetefeld, M. Jenny, T. Schulthess, R. Landwehr, J. Engel, R. A. Kammerer, *Nat. Struct. Biol.* **2000**, *7*, 772–776.
[12] K. M. Arndt, J. N. Pelletier, K. M. Müller, A. Plückthun, T. Alber, *Structure* **2002**, *10*, 1235–1248.
[13] E. K. O'Shea, K. J. Lumb, P. S. Kim, *Curr. Biol.* **1993**, *3*, 658–667.
[14] T. J. Graddis, D. G. Myszka, I. M. Chaiken, *Biochemistry* **1993**, *32*, 12664–12671.
[15] S. Y. Lau, A. K. Taneja, R. S. Hodges, *J. Biol. Chem.* **1984**, *259*, 13253–13261.
[16] S. F. Betz, J. W. Bryson, W. F. DeGrado, *Curr. Opin. Struct. Biol.* **1995**, *5*, 457–463.
[17] P. B. Harbury, T. Zhang, P. S. Kim, T. Alber, *Science* **1993**, *262*, 1401–1407.
[18] K. Wagschal, B. Tripet, P. Lavigne, C. Mant, R. S. Hodges, *Protein Sci.* **1999**, *8*, 2312–2329.

[19] B. Tripet, K. Wagschal, P. Lavigne, C. T. Mant, R. S. Hodges, *J. Mol. Biol.* **2000**, *300*, 377 – 402.

[20] Y. Tang, D. A. Tirrell, *J. Am. Chem. Soc.* **2001**, *123*, 11 089 – 11 090.

[21] J.K. Kretsinger, J. P. Schneider, *J. Am. Chem. Soc.* **2003**, *125*, 7907 – 7913.

[22] S. A. Potekhin, V. N. Medvedkin, I. A. Kashparov, S. Venyaminov, *Protein Eng.* **1994**, *7*, 1097 – 1101.

[23] K. J. Lumb, P. S. Kim, *Biochemistry* **1995**, 34, 8642 – 8648.

[24] L. Gonzalez, Jr., D. N. Woolfson, T. Alber, *Nat. Struct. Biol.* **1996**, *3*, 1011 – 1018.

[25] W. D. Kohn, R. S. Hodges, *Trends Biotechnol.* **1998**, *16*, 379 – 389.

[26] F. K. Junius, J. P. Mackay, W. A. Bubb, S. A. Jensen, A. S. Weiss, G. F. King, *Biochemistry* **1995**, *34*, 6164 – 6174.

[27] J. N. Pelletier, K. M. Arndt, A. Plückthun, S. W. Michnick, *Nat. Biotechnol.* **1999**, *17*, 683 – 690.

[28] K. M. Arndt, J. N. Pelletier, K. M. Müller, T. Alber, S. W. Michnick, A. Plückthun, *J. Mol. Biol.* **2000**, *295*, 627 – 639.

[29] M. G. Oakley, P. S. Kim, *Biochemistry* **1998**, *37*, 12 603 – 12 610.

[30] S. Frank, T. Schulthess, R. Landwehr, A. Lustig, T. Mini, P. Jenö, J. Engel, R. A. Kammerer, *J. Biol. Chem.* **2002**.

[31] V. A. Sharma, J. Logan, D. S. King, R. White, T. Alber, *Curr. Biol.* **1998**, *8*, 823 – 830.

[32] W. D. Kohn, C. M. Kay, R. S. Hodges, *Protein Sci.* **1995**, *4*, 237 – 250.

[33] N. E. Zhou, C. M. Kay, R. S. Hodges, *J. Mol. Biol.* **1994**, *237*, 500 – 512.

[34] J. R. Newman, A. E. Keating, *Science* **2003**, *300*, 2097 – 2101.

[35] J. J. Havranek, P. B. Harbury, *Nat. Struct. Biol.* **2003**, *10*, 45 – 52.

[36] K. M. Campbell, K. J. Lumb, *Biochemistry* **2002**, *41*, 7169 – 7175.

[37] O. D. Monera, C. M. Kay, R. S. Hodges, *Biochemistry* **1994**, *33*, 3862 – 3871.

[38] J. H. Brown, C. Cohen, D. A. Parry, *Proteins* **1996**, *26*, 134 – 145.

[39] M. R. Hicks, D. V. Holberton, C. Kowalczyk, D. N. Woolfson, *Fold. Des.* **1997**, *2*, 149 – 158, and references therein.

[40] Y. B. Yu, *Adv. Drug Deliv. Rev.* **2002**, *54*, 1113 – 1129.

[41] X. Zeng, H. Zhu, H. A. Lashuel, J. C. Hu, *Protein Sci.* **1997**, *6*, 2218 – 2226.

[42] L. Gonzalez, Jr., J. J. Plecs, T. Alber, *Nat. Struct. Biol.* **1996**, *3*, 510 – 515.

[43] L. Gonzalez, Jr., R. A. Brown, D. Richardson, T. Alber, *Nat. Struct. Biol.* **1996**, *3*, 1002 – 1009.

[44] D. L. McClain, H. L. Woods, M. G. Oakley, *J. Am. Chem. Soc.* **2001**, *123*, 3151 – 3152.

[45] J. R. Litowski, R. S. Hodges, *J. Biol. Chem.* **2002**, *277*, 37 272 – 37 279.

[46] R. A. Kammerer, *Matrix Biol.* **1997**, *15*, 555 – 565.

[47] V. N. Malashkevich, R. A. Kammerer, V. P. Efimov, T. Schulthess, J. Engel, *Science* **1996**, *274*, 761 – 765.

[48] S. Özbek, J. Engel, J. Stetefeld, *EMBO J.* **2002**, *21*, 5960 – 5968.

[49] Y. Guo, R. A. Kammerer, J. Engel, *Biophys. Chem.* **2000**, *85*, 179 – 186.

[50] K. A. Dill, *Biochemistry* **1990**, *29*, 7133 – 7155.

[51] J. A. Zitzewitz, O. Bilsel, J. Luo, B. E. Jones, C. R. Matthews, *Biochemistry* **1995**, *34*, 12 812 – 12 819.

[52] H. Wendt, L. Leder, H. Harma, I. Jelesarov, A. Baici, H. R. Bosshard, *Biochemistry* **1997**, *36*, 204 – 213.

[53] K. S. Thompson, C. R. Vinson, E. Freire, *Biochemistry* **1993**, *32*, 5491 – 5496.

[54] H. Qian, *Biophys. J.* **1994**, *67*, 349 – 355.

[55] H. Zhu, S. A. Celinski, J. M. Scholtz, J. C. Hu, *Protein Sci.* **2001**, *10*, 24 – 33.

[56] A. I. Dragan, P. L. Privalov, *J. Mol. Biol.* **2002**, *321*, 891 – 908.

[57] T. R. Sosnick, S. Jackson, R. R. Wilk, S. W. Englander, W. F. DeGrado, *Proteins* **1996**, *24*, 427 – 432.

[58] H. R. Bosshard, E. Durr, T. Hitz, I. Jelesarov, *Biochemistry* **2001**, *40*, 3544 – 3552.

[59] J. K. Myers, T. G. Oas, *J. Mol. Biol.* **1999**, *289*, 205 – 209.

[60] L. B. Moran, J. P. Schneider, A. Kentsis, G. A. Reddy, T. R. Sosnick, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10 699 – 10 704.

[61] R. A. Kammerer, T. Schulthess, R. Landwehr, A. Lustig, J. Engel, U. Aebi, M. O. Steinmetz, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 13 419 – 13 424.

[62] S. Frank, A. Lustig, T. Schulthess, J. Engel, R. A. Kammerer, *J. Biol. Chem.* **2000**, *275*, 11 672 – 11 677.

[63] D. L. Lee, P. Lavigne, R. S. Hodges, *J. Mol. Biol.* **2001**, *306*, 539 – 553.

[64] J. N. Glover, S. C. Harrison, *Nature* **1995**, *373*, 257 – 261.

[65] J. Walshaw, D. N. Woolfson, *J. Mol. Biol.* **2001**, *307*, 1427 – 1450.

[66] A. Lupas, M. Van Dyke, J. Stock, *Science* **1991**, *252*, 1162 – 1164.

[67] B. Berger, D. B. Wilson, E. Wolf, T. Tonchev, M. Milla, P. S. Kim, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 8259 – 8263.

[68] A. Lupas, *Curr. Opin. Struct. Biol.* **1997**, *7*, 388 – 393.

[69] R. R. Naik, S. M. Kirkpatrick, M. O. Stone, *Biosens. Bioelectron.* **2001**, *16*, 1051 – 1057.

[70] K. M. Arndt, K. M. Müller, A. Plückthun, *J. Mol. Biol.* **2001**, *312*, 221 – 228.

[71] J. Moitra, M. M. Mason, M. Olive, D. Krylov, O. Gavrilova, B. Marcus-Samuels, L. Feigenbaum, E. Lee, T. Aoyama, M. Eckhaus, M. L. Reitman, C. Vinson, *Genes Dev.* **1998**, *12*, 3168 – 3181.

[72] B. Tripet, L. Yu, D. L. Bautista, W. Y. Wong, R. T. Irvin, R. S. Hodges, *Protein Eng.* **1996**, *9*, 1029 – 1042.

[73] K. M. Müller, K. M. Arndt, T. Alber, *Methods Enzymol.* **2000**, *328*, 261 – 282.

[74] C. Wang, R. J. Stewart, J. Kopecek, *Nature* **1999**, *397*, 417 – 420.

[75] N. A. Peppas, *Curr. Opin. Colloid Interface Sci.* **1997**, *2*, 531 – 537.

[76] M. G. Ryadnov, B. Ceyhan, C. M. Niemeyer, D. N. Woolfson, *J. Am. Chem. Soc.* **2003**, *125*, 9388 – 9394.

[77] P. J. Kraulis, *J. Appl. Crystallogr.* **1991**, *24*, 946 – 950.