New currency for old rope: from coiled-coil assemblies to α-helical barrels
Derek N Woolfson1,2, Gail J Bartlett1, Marc Bruning1 and Andrew R Thomson1

α-Helical coiled coils are ubiquitous protein–protein-interaction domains. They share a relatively straightforward sequence repeat, which directs the folding and assembly of amphipathic α-helices. The helices can combine in a number of oligomerisation states and topologies to direct a wide variety of protein assemblies. Although in nature parallel dimers, trimers and tetramers dominate, the potential to form larger oligomers and more-complex assemblies has long been recognised. In particular, complexes above pentamer are interesting because they are barrel-like, having central channels or pores with well-defined dimensions and chemistry. Recent empirical and rational design experiments are beginning to chart this potential new territory in coiled-coil space, leading to intriguing new structures, and possibilities for functionalisation and applications.

Addresses
1 School of Chemistry, University of Bristol, Cantock’s Close, Bristol BS8 1TS, UK
2 School of Biochemistry, University of Bristol, Medical Sciences Building, University Walk, Bristol BS8 1TD, UK

Corresponding author: Woolfson, Derek N (D.N.Woolfson@bristol.ac.uk)

Current Opinion in Structural Biology 2012, 22:1–10
This review comes from a themed issue on Engineering and design
Edited by Jane Clarke and Bill Schief

Introduction: backdrop to and outlook for this review
The concept of the α-helical coiled-coil protein structure—that is, two or more helices wrapped around each other to form consolidated rope-like assemblies—has been with us for nearly 60 years [1,2], experimental validations of this for over a quarter of a century [3–5], and the rational engineering and design of such structures has been finessed considerably over the past two decades [6,7]. By any measure, coiled-coil structural biology constitutes a mature research field. Nevertheless, coiled coils continue to surprise and puzzle us. New coiled-coil assemblies are coming to light that expand the repertoire of structures observed [8**]; imaginative and ambitious design applications are being attempted with success [9**]; and new analysis, modelling and prediction methods are developing at pace [10–12,13*,14,15]. Still, there remains considerable scope for exploration in the area: it is likely that other coiled-coil structures not yet observed in nature, or realised through design are possible; the outlook for de novo design, specifically with respect to generating new protein functions is tantalising; and, with improved understanding of sequence-to-structure relationships in coiled coils, the prospects for even better prediction, design and engineering are promising. This review attempts to capture some of this recent activity in, and potential for the coiled-coil field.

Basics of coiled-coil sequence, structure and function
α-Helical coiled coils constitute on average 3% of all protein-encoding regions of the known genomes [16]. They were first recognised in certain classes of fibrous protein, and, possibly as a result of this, their primary role has mainly been considered as structural. Indeed, coiled-coil domains are largely associated with directing, specifying and cementing protein–protein interactions. However, they play oligomerisation and structural roles in virtually all intracellular and extracellular processes, including transcription, ATP synthesis, intracellular transport, the cytoskeleton and cell structure, transmembrane signalling, membrane fusion and remodelling, and the formation and remodelling of the extracellular matrix. Coiled coils also come in a wide variety of sizes, structural types, and biochemical compositions.

As stated, α-helical coiled coils comprise two or more α-helices wrapped, or supercoiled, around each other to form rope-like structures, Figure 1a. The helices are amphipathic, being encoded by a relatively straightforward sequence repeat of hydrophobic (H) and polar (P) residues, (HPPHPPP)H, known as the heptad repeat and often designated abdefg, Figure 1b; although there are variations on this pattern, which essentially combine different 3-residue and 4-residue spacings of hydrophobic side chains. The helices combine to bury their hydrophobic a/d faces, Figure 1c. However, the mismatched periodicities of the heptad repeat (with H-type residues every 3.5 residues) and the α-helix (with 3.6 residues per turn) results in this a/d seam winding slowly around the helix; and, consequently, two or more helices wrap with a left-handed supercoil to maximise burial of the H-type side chains. Partly because the hydrophobic effect is not

Please cite this article in press as: Woolfson DN et al. New currency for old rope: from coiled-coil assemblies to α-helical barrels, Curr Opin Struct Biol (2012), doi:10.1016/j.sbi.2012.03.002
2 Engineering and design

Figure 1

The basics of coiled-coil sequences and structures. (a) The rope-like structure formed by a relatively long coiled coil, the dimeric contractile protein tropomyosin (PDB ID: 2EFR). (b) Helical-wheel diagram showing how the heptad repeat abcd ef g tracks around a helical structure (in this case with 3.5 residues per turn). (c) One-heptad slice through the structure of the leucine-zipper region of the yeast transcriptional activator GCN4 (2ZTA). (d) A Coiled-coil Basis Set comprising design dimer (CC-Di, left), trimer (CC-Tri, middle) and tetramer (CC-Tet, right, 3R4A). Oligomer-state specifying asparagine side chains at heptad position a in CC-Di and d in CC-tri are shown in red and green, respectively. These three oligomer states represent the major ones formed by the coiled-coil motif [10]. In panels a, c and d, colouring of the heptad positions, abcd ef g, (and in C the side chains) follows the CC+ standard for heptad positions (a = red; b = orange; c = yellow; d = green; e = cyan; f = blue; g = magenta) [10]. Structural images created using PyMol (http://www.pymol.org).

specific, many different combinations of helices are possible, including parallel, antiparallel, homo-typic and hetero-typic arrangements, as well as various oligomerisation states [17–19], Figure 1d. This is the traditionally accepted, text-book model for coiled-coil sequence-to-structure relationships and assembly.

Knobs-into-holes packing and sharper sequence-to-structure relationships
Of course, there must be more to sequence-to-structure relationships in coiled coils to account for the large proportion of such sequences in the known genomes, the complexity of structures that they form, and the presumed necessity for these to be orthogonal, that is, non-promiscuous. The key to understanding this lies in the knobs-into-holes (KIH) interactions first postulated by Crick [1]. Side chains from adjacent helices in coiled-coil structures do not simply contact each other, rather they interlock in specific and intimate ways: a side chain, referred to as the knob, from one helix interdigitates within a cluster of four side chains projecting in a diamond-shaped hole formed by the partnering helix. For example in parallel structures, an a knob coordinates within a dgad hole, Figure 2a, b.
There are further subtleties. For example, the $a$ and $d$ knobs of dimers project towards their partner helices in different ways: the former projects out of the interface and the latter directly into it, Figures 1b, c and 2b. Now consider how the knob residues project from one helix relative to another as oligomer state increases, Figure 2c: they change; the angle that the knob makes with the base of the hole is not only different for each knob site ($a$ or $d$), but also in each oligomer. These angles are referred to as core-packing angles, but essentially for coiled-coil dimers, trimers and tetramers they fall into three types: the aforementioned packings at $a$ and $d$ of parallel dimers are known as parallel and perpendicular, respectively; in parallel tetramers these arrangements are swapped; and in parallel trimers the packing at the two sites is somewhere between these, and referred to as acute, Figure 2c [20]. Perhaps not surprisingly given the intimacy of KIH interactions, these different shapes lead to different amino-acid preferences [21].

This connection between KIH packing and sequence was first noted by Harbury and colleagues in two classic papers [20,22]. In these studies, variants of the leucine-zipper peptide, GCN4-p1, taken from a yeast transcriptional activator, were made in which four consecutive pairs of $a$ and $d$ sites were changed wholesale to combinations of Ile, Leu and Val. The key mutants were: $a = $ Ile, $d = $ Leu (GCN4-pIL); $a = $ d = Ile (GCN4-pII); $a = $ Ile, $d = $ Leu (GCN4-pPL). These formed parallel dimers, trimers and tetramers, respectively, and linked core-packing angles to sequence: Leu is most tolerated at perpendicular sites, and Ile (or Val) are preferred at parallel or acute packings. Largely, these have been borne out through both bioinformatics [19,21], and experimental studies [7].

Over the intervening two decades, many groups have used either these GCN4-p1 variants directly, or the principles for what to use at $a$ and $d$ in order to specify oligomer state, in an impressive range of protein engineering and de novo design studies [7,23,24]. However, nature and natural sequences are not so straightforward; they have more variations, both in terms of amino-acid usage and combinations. Moreover, a key question is, how transferable are the Harbury relations to other, non-GCN4 or even entirely de novo systems?

On this point, recently we set out to create a toolkit of reliable de novo designed coiled-coil peptides that might be useful in protein engineering and synthetic biology ([25]; unpublished data). There were two primary reasons for doing this de novo: First, although it has had considerable impact, the GCN4 system suffers from structural plasticity; that is, small changes in sequence can lead to large changes in structure as demonstrated by several studies [26–29]. Second, in principle at least, the rational-design approach allows us to account for the structural influence of every amino acid in the sequence. We generated three peptides with the key Harbury combinations of Ile and Leu at $a$ and $d$ as foregrounds: that is, $a = $ Ile, $d = $ Leu (CC-pIL); $a = $ d = Ile (CC-pII); $a = $ Ile, $d = $ Leu (CC-pPL). The $g$–$f$ backgrounds were also standardised to (EHAAHKX)$_{4n}$ where: $H$ is Ile or Leu as stated, and $X$ represents the outer $f$ positions of the heptad repeat, which were made combinations of Gln, Lys or Tyr for solubility and to add a chromophore. CC-pIL and CC-pPL behaved as expected: they were trimeric and tetrameric in solution, and their structures were solved as parallel dimers and tetramers by X-ray crystallography, respectively, Figure 1d. Consequently, we dubbed these CC-Tri and CC-Tet to signify that they are characterised components of a Coiled-coil Basis Set. Subsequently, CC-Tri has been used to nucleate the assembly of an otherwise poor-folding bacterial collagen [30], and CC-Tet has been mutared to explore larger oligomer states (see below) [8**].

Surprisingly, CC-pIL did not form a dimer as anticipated, but a parallel trimer both in solution and the crystal state. Others have also observed that the $a = $ Ile, $d = $ Leu foreground does not necessarily lead to parallel dimers [31*]. These are important findings with implications for protein engineering, design and prediction, but how can they be reconciled with the earlier studies? Analysis of parallel dimers and trimers from CC+ reveals that although Leu is the most abundant of amino acids at the $a$ and $d$ sites—indeed, this holds across all oligomer states—it is not very discriminating in terms of oligomer state. Ile and Val are selected against at the perpendicular-packing sites, and are favoured, though only slightly, at the parallel-packing sites. Thus, the $a = $ Ile, $d = $ Leu foreground does not strongly favour dimers. What does appear to tip the balance further in natural sequences is a small number of other specific amino-acid placements. For example, Asn at $a$ strongly favours parallel dimers, which is a long-recognised relationship [32–34]. Indeed, placing a single Asn at the third $a$ site in CC-pIL causes a complete and robust switch to parallel dimer to give the Basis-set peptide CC-Di, Figure 1d.

**New analysis, database, prediction and modelling tools for coiled-coil structure**

Thankfully, coiled-coil structures do not have to be inspected by eye to uncover their secrets. Two programs in particular garner many of the important parameters of coiled-coil geometry. TWISTER [35] and SOCKET [19] identify important structural features from 3D coordinates of coiled-coil structures, such as those deposited in the RCSB Protein Data Bank (PDB) [36]. Whilst TWISTER calculates backbone and related parameters for coiled-coil geometries, SOCKET identifies runs of

---

*Note the order here has been changed to $gabdef$, which we find more useful because of the potential for $g_{o.o}+1$ polar–polar interactions.
4 Engineering and design

Figure 2

Knobs-into-holes (KH) interactions. (a) Crick’s original helical-net concept, which led to the idea of KH packing. On the left-hand side, the heptad repeats for two helices (coloured red and blue) are projected onto flat surfaces. These illustrate how the hydrophobic adf seams wind around the helices in the opposite sense to that of a regular, right-handed α-helix. On the right-hand side, the two nets are interlaced to show how each residue on
KIH interactions between α-helices, noting patterns within these. In this way, SOCKET assigns heptad patterns, oligomer states and topology, and it also provides quantitative measures of core-packing angles. Recently, SOCKET has been applied to the whole of the PDB to create a relational database of coiled-coil structures, CC+ [10], http://coiledcoils.chm.bris.ac.uk/ccplus/search/. CC+ may be searched in a variety of ways—based on keywords, structural properties, sequence and interactions—to return subsets of coiled-coil structures, and various data from these can be exported in a number of formats. Within these outputs, amino-acid profiles—that is, tables giving the proportions of each amino acid at the abcdedfg sites—are particularly useful as they address the aforementioned issue of more-diverse amino-acid usage, and provide a basis for developing new prediction algorithms and de novo designs. Also via the same URL, the Periodic Table of Coiled-coil Architectures organises the known structures according to oligomer state (the columns or groups of the table) and complexity of the arrangements (the rows or periods) [18]. Users should be aware, however, that whilst CC+ is regularly updated the Periodic Table of Coiled Coils is not.

Resources such as SOCKET and CC+ provide unprecedented links between coiled-coil sequence and structure, and thus open possibilities for protein-structure prediction, engineering and design. For example, the SCORER algorithm [21], which predicts propensities of sequences to form coiled-coil dimers and trimers has recently been retrained using structurally verified coiled coils in CC+ [12]; similarly PairCoil [37] and MultiCoil [38], sequence-based coiled-coil assignment and oligomeric-state predictors respectively, have been retrained on sets including sequences derived from SOCKET analyses of the PDB [15,39]. In addition, PrOCoil [14] uses SOCKET-derived information to train a support-vector-machine classifier of dimeric and trimeric coiled-coil sequences. A significant improvement on coiled-coil coverage in prediction has been achieved with a new four-state predictor for distinguishing antiparallel and parallel dimers, and parallel trimers and tetramers, LOGICOIL (unpublished method), which covers >90% of known coiled-coil structures. Using data from CC+ and hidden Markov models implemented in SUPERFAMILY [40], an annotation of coiled coils in all completely sequenced genomes has led to a genome-wide evaluation of coiled-coil evolution and homology based prediction called SPIRICOIL [16].

Related to the prediction problem, Gellman and colleagues have compared experimental data for a model antiparallel coiled-coil system and correlations culled from CC+. Specifically for this class of structure, they find that high-order interactions in d–a–d' vertical triads contain information that favours certain helix-helix pairings [41], notably hetero-combinations such as Ile–Leu–Ile are favoured; whereas, residue combinations at d–d–d are less discriminating [42]. These studies are important, demonstrating that coiled-coil prediction, modelling and design needs to move past amino-acid profiles and pairwise correlations, and tackle the problem more holistically including larger constellations of side chains. Keating’s group and others are also developing methods that bridge the computational and experimental aspects of coiled-coil prediction and design, particularly in the areas of antiparallel coiled-coil formation [43], and the definition of hetero-specific parallel coiled-coil dimers [44–46]; both of which are important areas that are far from fully understood.

New modelling methods, which build on foregoing work [1,47–49], are being developed for coiled-coil structures with both prediction and design in mind. For example, Grigoryan and DeGrado present a method of parameterizing coiled coils using modified Crick equations that allow for parallel, antiparallel and mixed orientations as well as helical sliding along the interfacial axis [13*]. Using this approach, the vast majority of natural coiled-coil structures were found to be within 1 Å RMSD of idealised ‘Crick’ backbones. Parameters derived from known coiled coils demonstrate that the geometrical space occupied by coiled coils is severely restricted: parallel and antiparallel structures have distinct and limited axial offset; likewise superhelical frequency and pitch angle are limited among all natural coiled coils regardless of orientation and oligomeric state; there is a linear relationship between the superhelical radius and oligomerisation state, with residues at the a and d positions showing distinct propensities at different superhelical radii in parallel coiled coils. This work is important

(Figure 2 Legend Continued) one helix effectively interdigitates between four residues of the partnering helix. This leads naturally to the left-handed supercoil of the coiled coil. (b) Examples from the X-ray crystal structure of GGN4-p1, a parallel, homodimeric coiled coil (PDB ID: 2ZTA); top, an a knob (red) coordinating within a dgad hole (blue); bottom, a d knob and an adea hole. (c) Typical core-packing angles made by a and d knobs in parallel dimers (uppermost, 2ZTA), trimers (middle, 1AQ5) and tetramers (lowermost, 2GUS). (d) From top to bottom: the standard N-type heptad repeat, and its structural realisation at a parallel dimer interface (2ZTA); type I repeat partially observed in trimers (1AQ5); type III repeat, tetramers–hexamers (BR3K); and the type III repeat anticipated above heptamer (1EK9). The angles shaded green in the bottom three helical-wheel diagrams are idealised and represent those subtended between the two hydrophobic seams of the types I–III repeats. In both type II and type III the interface is extendable: the residues involved in forming this interface are picked out in dots, and the region of the interface is shown on the structure and helical-wheel diagrams with a solid black line. Note that the structural example of the type III repeat comes from an antiparallel coiled-coil assembly: the central helix follows an N-to-C-terminal path into the page, whilst the two outer helices follow a C-to-N terminal path into the page; and the juxtaposition of heptad positions making knobs–into–holes interactions is altered. Colouring follows the CC+ standard for heptad positions (a = red; b = orange; c = yellow; d = green; e = cyan; f = blue; g = magenta) [10]. Structural images created using PyMol (http://www.pymol.org).
in narrowing down parameter and, therefore, structural space available to coiled coils, and, so, paves the way to better model building for prediction and design applications. A web site, CCCP (http://arteni.cs.dartmouth.edu/cccpc), is available for generating idealised Crick backbones using these methods.

Extended KIH packing and the potential for higher-order states
Another consequence of increasing oligomer state is that residues other than those at a and d become more involved in the helix-helix interfaces, Figure 2c, d. Specifically, the e and g sites become progressively buried. The extreme is

Orthogonal views of α-helical barrels. (a) The CC-Hex hexamer with a 5–6 Å channel (PDB ID: 3R3K). (b) The L24D:L24H heterohexamer (3R48). (c) The heptameric mutant of GCN4-p1, which has an ~8 Å channel (2HY6). (d) The TolC structure, which has a periplasmic 12-helix α-barrel with a ~25 Å pore (1EK9). (e) The Wza assembly with an 8-helix transmembrane barrel and a 17 Å internal diameter (2J58). (f) The cytotoxin ClyA with a >50 Å pore contributed to by 24 helices (2WCD). (g) The open state of the heptameric mechanosensitive channel MscS with a ~13 Å channel (2VV5). The α-helical barrel structures are shown in colour.
that residues at these positions become knobs, and part of what have been termed peripheral KIH interactions [50]. These trends can be seen through SOCKET analyses of the GCN4 variants, the Basis-set peptides and natural coiled coils. The natural extremes are provided by a small number of pentamers in the CC+ database, which include COMP, phospholamban, a rotavirus enterotoxin, CorA, and some engineered structures with aromatic cores [51–56]. In these structures many of the g, a, d and e positions act as knobs, and some of the sequences approximate to HHxxHHx repeats, rather than more traditional sHxxHxx heptad repeats. (Note the gabcde order.) This relates back to an idea of offset double heptad repeats [50], which had been considered previously, at least in theory, by several groups [57–59].

The above HHxxHHx pattern can be split up to give two heptad repeats, sHxxHxx plus HxxHxx, both of which are 3,4-repeats of H-type residues. This is one of three possible offset double-heptad repeats, the others being HHxxHxx and sHxsHHsH, Figure 2d [50]. For simplicity, we refer to these and the original heptad pattern as: xHxxHxx, N; HHxxHxx, I; HHxxHHx, II, and sHxsHHsH, III. These are arranged in the order of increasing oligomer-state that they give rise to: dimers best fit N; I is observed in trimers; II in tetramers and pentamers; and III opens up possibilities for a range of complex coiled coils, including α-helical sheets and α-helical barrels, Figure 2d [50,57–60]. These are not hard and fast rules, but we find the nomenclature and concepts useful. What’s more, they question what is meant and understood by ‘coiled coil’ both in terms of sequence and structure, and what other structures might be accessible to this often-considered straightforward domain.

New higher-order assemblies, a coiled-coil hexamer (CC-Hex)

Considering these possibilities of offset double-heptad repeats opens up a whole new space in coiled-coil assembly that is just beginning to be realised and explored. For example, recently, the Basis-set tetramer, CC-Tet, has been used to create a parallel coiled-coil hexamer, CC-Hex [8**]. The details of this study are illuminating. First, CC-Tet has a traditional N-type heptad repeat that specifies tetramer using the Harbury relationships, that is (ELAAIKX)4. The X-ray structure reveals a parallel tetramer with classical perpendicular and parallel KIH interactions at a and d, respectively. However, and consistent with a type-II pattern: ≈1/4 of the total number of KIH interactions identified by SOCKET fall at the g and e sites. The sequence was mutated to exchange all of Lys at e with the Ala at b, to give the following repeat, (ELKAILAX)4, which potentially broadens the hydrophobic seam to a, d and e. The new peptide forms a very stable α-helical hexameric complex in solution, which crystallises as a parallel hexamer, Figure 3a, [8**]. This structure had not been observed before; indeed, on the basis of likely overstretched KIH interactions, it had been predicted that it should not exist [50]. Nonetheless, the new structure, CC-Hex, tests positive as a 6-helix parallel coiled coil in both TWISTER and SOCKET, and all of the residues at g, a, d and e act as knobs in classical KIH interactions. However, and somewhat consistent with the earlier prediction, the core-packing angles at the g and a sites fall on the periphery of the distribution observed for all perpendicular sites of parallel dimers tetramers in CC+; n.b., those at d and e fall in the middle of the range observed for parallel packing in the same structures.

CC-Hex is intriguing for another reason, which is likely to have a much broader impact [61]: though the six helices come together by largely hydrophobic contacts they leave a central well-defined channel of 5–6 Å. Figure 3a. This is lined exclusively by methyl groups of the Leu and Ile side chains at a and d. Nonetheless, there is electron density within the channel, which is adequately modelled as a disrupted chain of water molecules. The side chains of Leu-24—the a site of the third heptad, and which point directly into the channel—are mutable, accepting residues as diverse as aspartic acid and histidine. These L24D and L24H mutants, although destabilised, form parallel hexamers in solution and the crystal state. Moreover, they combine to form an intriguing heterohexamer with alternating Asp-containing and His-containing chains, Figure 3b.

Towards α-helical barrels

CC-Hex raises the bar of oligomer states accessible to coiled-coil motifs. So how far can the concept of extended KIH patterns be pushed to model, predict and design new coiled-coil oligomers? CC-Hex is a true and classical coiled coil; that is, it has a contiguous ring of KIH interactions, albeit surrounding an internal channel. There is heptameric coiled coil in the CC+ database—a mutant of GCN4-p1 with g = e = Ala [28]—however we note that this is far from a classical coiled coil: the SOCKET assignments of KIH interactions and heptad register are not straightforward; and the helices spiral rather than packing blunt-ended, with the first and seventh helices offset by a whole heptad, Figure 3c. Moreover, an ideal sequence pattern for a heptameric coiled coil falls midway between the type-II and III patterns. In these respects, the heptamer might be regarded as the tipping point between classical and more-complex coiled coils [8**]. Nevertheless, the structure is barrel-like, with the internal channel is fully occupied by hexane-1,6-diol co-solvent, and the structure is interesting further and potentially useful for that.

The next oligomer state represented in the PDB and CC+ is a dodecamer that is part of the multidrug efflux protein TolC from E. coli [60], Figure 3d. This is a beautiful, multifunctional, but complicated structure. The central,
periplasm-spanning domain comprises 12 antiparallel α-helices arranged in a barrel; indeed, to our knowledge, this was where the term α-barrel was first coined [60]. The antiparallel topology leads to some unwinding of the coiled coil [58]. Nevertheless, the helices do show some KIH interactions in SOCKET, and these fit the type-III pattern, Figure 2; effectively the structure is a ring of helices connected via two dimer-like hydrophobic seams per helix.

Interestingly, recent X-ray and cryoEM studies of the viral genome-delivery portal from bacteriophage P22 reveal a long 12-helix barrel, proposed to be responsible genome ejection into host cells [62,63]. At this stage the resolution of this part of the structure is insufficient to test if the helices are cemented by KIH interactions. Nonetheless, it would appear the 12-helix barrel of TolC is not unique, that is, a singleton in structural terms.

In addition, to rings of 12 helices, there is a growing class of membrane-spanning and membrane-associated α-helical barrels that suggest that various pore and channel sizes can be achieved, Figure 3e–g. For example, the first membrane-spanning α-helical barrel was reported in the structure of Wza [64]. Wza is an outer-membrane lipoprotein from E. coli. It forms part of the polysaccharide export machinery essential to capsule formation around the bacterium. Domain 4 of the protein is an amphipathic α-helix that octomerises in the membrane to produce a polar channel with an internal diameter of ∼17 Å. A much larger barrel has been observed in the α-pore-forming toxin ClyA from E. coli and Salmonella [65]. The active form results from a large structural rearrangement of the protein in which 12 three-helix bundles combine to produce the barrel. Each bundle contributes two helices to the innerface of barrel, leading to a negatively charged, cation-selective pore with an internal diameter of 70 Å. Finally, the transmembrane region of the mechanosensitive channel, MscS, again from E. coli, comprises a heptamer of three-helix bundles [66,67]. In this case the first two helices buttress the third, which forms a much tighter and controllable channel. It is proposed [67] that the conformations of the first and second helices respond to changes in osmotic pressure on cells by twisting and thus opening up the channel from ∼5 Å to 13 Å in its closed [66] and open states [67], respectively. In terms of coiled-coil structures, however, only part of the large ClyA barrel tests positive in SOCKET, Figure 3f.

Conclusions: prospects for the future discoveries and designs
The leap from 6-membered and 7-membered barrels to those with 12 helices raises the questions: are true coiled-coil octomers, nonomers, decamers and hendecamers possible, and what about oligomers above that? If any of these were to be observed in nature, or could be designed de novo, they would not only be of interest in coiled-coil ‘stamp collecting’, but could have an impact in the design of ion-channels, binding proteins, sensors and even enzymes [8,61]. To illustrate this potential, an elegant computational redesign—using the aforementioned methods from Grigoryan and DeGrado [13]—of a previous de novo coiled-coil bundle has successfully produced a peptide that binds and solubilises carbon nanotubes [9]; interestingly, the designed bundle is for an antiparallel 6-helix coiled coil, though confirmation of this awaits high-resolution structural studies.

Accessing the various other ‘largermers’, however, presents a number of challenges, including: the apparent requirement for very hydrophobic peptides to satisfy the xHxHHxH, type-III pattern; residue selection at the H sites; and the aforementioned problem with unwinding the coiled coils. Even with these tackled it may be that specifying particular oligomer states will be difficult, as the difference in energy between alternate states is expected to become smaller as oligomer size increases. Indeed, on purely geometric grounds, the type-III pattern should lead to a 14-mer, which has yet to be observed [57,59], rather than the dodecamer that is realised in TolC [60]; though of course, other forces are almost certainly at play in the folding of this complex protein structure. Further on this point, the TolC structure has been analysed theoretically to understand its physics and sequence-to-structure relationships [58], but, to our knowledge, these have not led to a successful de novo design for a dodecamer. Thus, it is almost certain that further developments in our current theory, sequence-to-structure relationships and methods for modelling coiled-coil proteins will be required to designed versions of the dodecamer, and, indeed any other largermer states above hexamer. Nonetheless, with the new activities in all three of these areas, and the new structures that have been realised experimentally, the outlook for coiled-coil research leading to better understanding and functional designs is promising.

Acknowledgements
We thank the Woolfson group, Leo Brady, Vince Conticello, Stuart Conway, Noah Linden and Richard Sessions for helpful discussions, and the BBSRC and EPSRC of the UK for funding (grant numbers: BB/G008533/1 and EP/J01440/1).

References and recommended reading
Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest


1. The first reported structure for a classical, parallel coiled-coil hexamer. This was achieved by engineering a coiled-coil tetramer of de novo design. The structure has a central channel, which can be mutated to residues including charged side chains, and, in turn, these can be used to direct heterohexameric assemblies.


8. The rational computational redesign of a de novo coiled-coils bundle to engineer peptides that assemble around carbon nanotubes. The peptides are used subsequently to solubilise and decorate the carbon nanotubes.


14. The development of computational tools to explore the structural space available to coiled-coil assemblies. The authors provide a web site for generating model backbones for many coiled-coil architectures and topologies to order. Methods such as this will improve our ability to model and design coiled-coil structures.


32. Contributions to the mounting evidence that engineered and designed coiled-coil sequences with a = Ile and d = Leu sequence signatures do not necessarily form parallel dimers, but can also form trimeric coiled coils.


42. Steinkruger JD, Bartlett GJ, Hadley EB, Fay L, Woolfson DN, Gellman SH: The d-d-d Vertical line is less discriminating than the a-a-a Vertical line or the antiparallel coiled-coil dimer motif. J Am Chem Soc 2012, 134:2626-2633.

10 Engineering and design


64. X-ray crystal structures for parts of the genome-delivery machinery described in reference [62].


