Kinking the Coiled Coil – Negatively Charged Residues at the Coiled-coil Interface

Ravid Straussman¹, Ami Ben-Ya’acov¹, Derek N. Woolfson²,³ and Shoshana Ravid¹*

¹Department of Biochemistry
Faculty of Medicine, Hebrew University, Jerusalem 91120, Israel
²School of Chemistry
University of Bristol, Cantock’s Close, Bristol BS8 1TS, UK
³Department of Biochemistry
School of Medical Sciences, University of Bristol, Bristol BS8 1TD, UK

The coiled coil is one of the most common protein-structure motifs. It is believed to be adopted by 3–5% of all amino acids in proteins. It comprises two or more α-helical chains wrapped around one another. The sequences of most coiled coils are characterized by a seven-residue (heptad) repeat, denoted \((abcdefg)_n\). Residues at the \(a\) and \(d\) positions define the helical interface (core) and are usually hydrophobic, though about 20% are polar or charged. We show that parallel coiled-coils have a unique pattern of their negatively charged residues at the core positions: aspartic acid is excluded from these positions while glutamic acid is not. In contrast the antiparallel structures are more permissive in their amino acid usage. We show further, and for the first time, that incorporation of Asp but not Glu into the \(a\) positions of a parallel coiled coil creates a flexible hinge and that the maximal hinge angle is being directly related to the number of incorporated mutations. These new computational and experimental observations will be of use in improving protein-structure predictions, and as rules to guide rational design of novel coiled-coil motifs and coiled coil-based materials.

© 2006 Elsevier Ltd. All rights reserved.

*Corresponding author

Keywords: coiled coil; amino acid profiles; fibrous proteins; myosin II; protein-structure prediction

Introduction

The α-helical coiled coil was first proposed by Crick in 1953.⁴ Crick’s model describes polypeptide chains configured into right-handed α-helices wrapping around each other slightly offset to form left-handed rope-like structures. The helix-helix interactions are cemented by tight interdigitation of side-chains at the helical interface, which Crick termed “knobs-into-holes” packing. A heptad sequence repeat, in which hydrophobic residues that fill the interface are spaced alternately three and four residues apart, specifies this structural motif. These structural and sequence features have subsequently been verified by numerous experimental studies.⁵ The coiled coil is in fact one of the most common protein-structural motifs: it is believed to be adopted by approximately 3–5% of all amino acids in proteins.⁶ In effect, the coiled coil is a motif for protein-protein interactions, or protein oligomerization. Typically α-helical coiled coils comprise two to five α-helices. The sequences of most, but not all,⁴–⁶ coiled coils are characterized by the aforementioned seven-residue (heptad) repeat, which is often denoted \((abcdefg)_n\) (Figure 1(a)). It is the \(a\) and \(d\) positions that are predominantly hydrophobic and are found at the interface between the α-helices; positions \(e\) and \(g\) flank this core, and tend to be oppositely charged residues; and the remaining positions, \(b\), \(c\) and \(f\), lie away from the coiled-coil interface and tend to be occupied by polar residues.⁷ The alternate spacing of \(a\) and \(d\) of three and four residues along the chain averages hydrophobic residues to approximately one every 3.5 residues. As this falls short of the 3.6 residues per turn of the α-helix, the resulting \(a/d\) hydrophobic
Figure 1. The structure of a coiled coil and of myosin II. (a) The helical wheel representation of the coiled-coil heptad repeat showing how the positions a-g appear when viewed from the C-terminal end of the parallel helices. (b) Schematic diagram of the myosin II molecule.

seam winds around the surface of the helix in the opposite sense to the direction of the polypeptide backbone. Thus, in order for the hydrophobic seams of the helices to remain in contact over the length of the coiled coil they must wrap or supercoil around each other, and, as predicted by Crick,1 heptad-the coiled coil they must wrap or supercoil around of the helices to remain in contact over the length of the backbone. Thus, in order for the hydrophobic seams to occur in linear sequences, there is disagreement when asking “simply” where coiled-coil motifs happen to act by destabilizing alternative folds more than the native structures.18-21

Aside from the general requirement for a and d to be hydrophobic, the precise nature of these sites differ markedly both in terms of coiled-coil sequence and structure. Although this had been hinted at before,22 Harbury and colleagues were the first to formalize the structural differences at the two sites.23 For instance, in parallel dimers the packing geometries of the knobs-into-holes interactions at a and d are different and referred to as parallel and perpendicular, respectively; in parallel tetramers these geometries are reversed. In trimers the geometries at a and d are similar and intermediate between parallel and perpendicular packing, referred to as acute. The core-packing geometries in antiparallel coiled coils are slightly more complicated.11 These packing differences are reflected in amino acid preferences observed at the two sites in different oligomer states both of engineered23 and natural sequences.10 For example, Leu is the most favoured residue at perpendicular sites, hence the strong preference for this residue at the d sites of parallel dimers.10,11,23

Regarding the inclusion of polar residues at the interfaces of parallel coiled coils, which is the subject of this paper, there is a clear preference for certain polar residues at the a sites. It appears that the parallel packing geometry at these sites is more tolerant of different amino acid side-chains in general.11,23,24 Having said this, only certain polar amino acids, namely, Asn, Lys and Arg, appear to confer dimer specificity.10,20,21,25 High-resolution structure studies reveal that these particular residues either make specific side-chain–side-chain interactions, or simply “escape” the core at the a site; whereas, neither could be achieved by other polar residues or at the d positions.21

Firstly, here we used the SOCKET-derived database to examine the occurrence of charged amino acids at the a and d positions of known coiled-coil structures. We show that the inclusion of aspartic acid (Asp) at these sites differs between parallel and antiparallel coiled coils as it is totally excluded only from the core positions of parallel coiled coils. In contrast, glutamic acid (Glu) can be found at the core.

† http://www.biols.susx.ac.uk/coiledcoils

We used this database to investigate the occurrence of charged amino acids in the hydrophobic cores of coiled coils.

Many studies demonstrate the importance of the hydrophobic nature of the core a and d positions in coiled-coil folding and stability.12-19 Polar and charged residues are also found at these positions, however, where they account for up to 20% of the residues, and are, in many cases, highly conserved.3,8-10 In some respects, the roles of these inclusions are understood. Such inclusions help specify oligomer states (i.e. they select between dimers, trimers and tetramers), and strand orientation (i.e. parallel versus antiparallel arrangements) of coiled coils. They appear to act by destabilizing alternative folds more than the native structures.18-21

positions of both parallel and antiparallel coiled coils. We also show that both negatively charged amino acids prefer the \( d \) over the \( a \) positions.

We then used the human non-muscle myosin II heavy chain (MHC) to examine the impact of introducing a negatively charged amino acid into the \( a \) position of a long parallel homodimeric coiled coil. Myosin II is a hexamer composed of two heavy chains of ~200 kDa and two pairs of light chains. The heavy chains are organized into two globular head regions containing the force generating, actin-binding and ATPase functions, and a rod region composed of a long (>1000 residue) \( \alpha \)-helical coiled coil (Figure 1(b)). Myosin II forms filaments through self-association of its rod domain. The protein is monomeric at high ionic strengths (>300 mM NaCl), but assembles into filaments in low salt. A specific region near the C-terminal end of the rod, dubbed the “assembly competent domain” (ACD), is necessary for filament assembly.

We mutated several positions of the nematode myosin II rod (N-terminal to the ACD) replacing them with Asp, Glu or Leu. We show that all of the Asp mutations introduced flexible kinks in the coiled-coil rod at positions corresponding to the mutation site, while Glu or Leu mutations had no such effect.

Combined, our sequence analysis and experimental studies strongly indicate that Asp residues disrupt the assembly of parallel coiled-coils, when placed at the \( a \) position of the coiled-coil interface. This observation should help improve coiled-coil structure predictions from sequences, and could provide useful in de novo protein design: it could be used as a negative design rule to terminate coiled-coil motifs, and as a positive design rule for introducing kinks and hinges into novel coiled-coil-based nanostructures and biomaterials.

### Results

#### A unique pattern of charged residues at the core sites of MHC coiled coils

While working with the long, parallel homodimeric coiled-coil rods of MHC subtypes, we noticed that there were virtually no Asp at either the \( a \) or \( d \) core sites, and no Glu at the \( a \) sites, while positively charged residues (Lys and Arg) could be found at both core positions. McLachlan & Karn described the sequence of the nematode MHC rod as highly repetitive with features typical of an \( \alpha \)-helical coiled coil. They note that Asp and Glu are excluded from the \( a \) positions of the nematode MHC rod, and are rare at \( d \). Since then, many more MHC sequences have been determined. From these, we selected 14 rod sequences to represent the different subtypes of MHC from different species and, within these we assigned the coiled-coil regions (Supplementary Data, Table S1). Despite a total of 15,489 coiled-coil residues with more than 2000 \( a \) sites, there were no negatively charged amino acid residues in this position \((p<0.001)\) and there was a virtual absence of Asp from the \( d \) position \((p<0.001)\). In contrast, Glu was slightly favored at the \( d \) position \((p<0.001)\). To avoid bias resulting from including more than one representative structure of each MHC subtype, only the five known human MHC subtypes were used in the analysis presented in Table 1A and Supplementary Data, Table S2. Consistent with foregoing analyses, the positively charged amino acids, Arg and Lys, are favoured at \( a \) and disfavored at \( d \) \((p<0.001)\).

Aspartic acids are excluded from the core sites of all parallel coiled coils

To test if the above observation for MHC sequences extended to coiled coils in general, we turned to the sequences of structurally verified coiled coils lodged in the SOCKET-derived CC+ database. Unlike sequence-based predictors of coiled-coil motifs, which search solely for sequence features to predict coiled-coil motifs, the SOCKET algorithm uses the characteristic knobs-into-holes side-chain packing of coiled coils to identify coiled-coil motifs in the RCSB Protein Data Bank. Table 1B gives data culled from the coiled-coil structures in the currently available version of the CC+ database. The data in Table 1B and in Supplementary Data, Tables S3–S4 show that indeed there are no Asp residues in the core sites of all identified parallel coiled coils \((p<0.001)\) while Glu can be found at both core sites, though it is much more abundant in the \( d \) position \((p<0.001)\). Indeed, the only incidence of a parallel coiled-coil with Glu at an \( a \) position (out of 205 verified parallel dimeric, trimeric and tetrameric coiled coils) was the c-myc-max heterodimer \((PDB identifier 1a93)\). The antiparallel structures were generally more permi-

| Table 1. Frequency of charged amino acids in coiled-coil “core” positions |
|--------------------------|--------|--------|--------|
|                         | Parallel | Antiparallel |
| a                        | d       | a      | d      |
| A. Five human myosin II heavy chain subtypes (Supplementary Data, Table 1) |
| Asp                      | 0       | 0.05   |
| Glu                      | 0       | 1.5    |
| Lys                      | 2.3     | 0.4    |
| Arg                      | 2.3     | 0.2    |
| B. Coiled coils that were tested positive for coiled coil motifs in the SOCKET analysis of the PDB |
| Asp                      | 0       | 0.05   |
| Glu                      | 0.08    | 0.5    |
| Lys                      | 0.2     | 0.3    |
| Arg                      | 0.4     | 0.3    |

The data were normalized for the frequency of occurrence of each amino acid in SWISSPROT.

† [http://www.biols.susx.ac.uk/coiledcoils](http://www.biols.susx.ac.uk/coiledcoils)
sive in their use of these side-chains at both positions but here again, in the a positions Asp or Glu were less abundant compared to the d positions ($p < 0.001$). Like MHC subtypes, the positively charged amino acids, Lys and Arg, occur more frequently than Asp and Glu in the cores of coiled coils as commented ($p < 0.001$). However, there are no strong preferences for Lys or Arg at these positions in any of the structures (Table 1B). Thus, the high incidences of these residues at the predicted a sites of the MHC sequences is possibly a unique and another interesting feature of these proteins (Table 1A).

**Introducing aspartic acid to a positions of the MHC rod causes local kinking**

We then used the human non-muscle MHC (NM-MHC) to test the effect of inserting a negatively charged amino acid at the a sites of a parallel homodimeric coiled coil. An expression plasmid was constructed to yield a protein similar to the C-terminal 640 residues of the NM-MHC, with an additional methionine and alanine residues at the N terminus (NM-MHC-642). Similar NM-MHC constructs have been shown to form coiled-coil filaments similar to wild-type NM-MHC. Two or four a positions along the rod were replaced by Asp or Glu. The same positions were also converted to Leu, which is the most abundant amino acid at the a sites of coiled coils. The resulting proteins were named according to the site of the mutations, counted as the number of amino acids from the N-terminal end of the construct. To verify that the expressed proteins formed rods, molecular-length measurements of rotary shadowed preparations were made taking full-length chicken-gizzard smooth-muscle myosin as a standard (Figures 2 and 3). All of the mutants formed rods of the expected length (Table 2, and data not shown). Moreover, there were clear kinks in the Asp-mutated rods but not in the Glu or Leu-mutated rods (Figures 2–4). The low baseline level of kinks observed in all mutants shown in Figure 4 is probably due to a technical artifact resulting from the force applied on the rods when spreading them on the grids. Indeed, most of the kink sites in the Asp-mutated rods correlated well with the position of the mutations (Table 2) while the rare kinks in all other rods were distributed along the rod (data not shown). Ranges of kink angles were observed for each Asp mutation (Figure 2), with about 20% of the rods showing a maximal kink. It is evident that the maximal kink is related to the number of mutations in each protein. Since the two mutations are only one heptad apart, two separate kinks cannot be resolved (Figure 3).

**Kinks in the coiled-coil interfere with its assembly properties**

The coiled-coil structure is well known as one of the most common motifs for protein–protein inter-
actions. Asp in the a position causes a local kink in the coiled coil, it might interfere badly with its assembly properties. To check if this is indeed the case we took advantage of the fact that our model protein, myosin, can form highly ordered filaments through interactions between its coiled-coil motifs in vivo and when dialyzed against low ionic strength buffer in vitro. We have thus carried out a sedimentation assay in which monomeric myosin II coiled coils in high ionic strength buffer are dialyzed against a range of salt concentrations and then subjected to centrifugation: monomers remain in the supernatant, while multimeric structures pellet and samples of both fractions were evaluated by SDS–PAGE and densitometry to determine the percentage of soluble myosin. Figure 5 shows the sedimentation curves of the parent rod and the mutant myosin II rods in positions 453 and 460. The parent rod NM-MHC-642 and the Leu mutant were completely insoluble below 200 mM salt, and became more soluble at salt concentrations up to 300 mM, at which point they were almost completely in the soluble fraction. In contrast, when a negatively charged amino acid was introduced to these a positions the mutated myosin II rods became soluble at much lower salt concentrations, indicating that these mutations indeed interfered badly with the assembly properties of the mutated coiled coil.

Another way to evaluate the ability of the monomeric coiled coils to form ordered filaments is to dialyze them against low salt buffer and examine the resulting samples by negative-staining electron microscopy (see Figure 6 for representative images). The parent rod, NM-MHC-642, formed long, ordered filaments with the normal striation pattern typical
for myosin filaments\textsuperscript{29,30,38} (Figure 6(a) and (f)). In contrast, all of the Asp mutated myosin II rods had aberrant morphologies. It is also interesting to note that in contrast to the parent monomeric coiled coils, which all assembled into filaments (Figure 6(a)), only fractions of the mutant coiled coils participated in filament formation and the remainder formed non-specific aggregates (Figure 6(b)–(e)), indicating their inability to foster normal protein–protein interactions.

### Discussion

In summary, by examining the sequences of all coiled coils identified with the SOCKET algorithm from the RCSB Protein Data Bank, we found that Asp is completely excluded from the $a$ and $d$ core sites of parallel coiled coils in general, while Glu can appear in both sites. We also show that the antiparallel coiled-coil structures are more permissive in their amino acid selection in both core sites and that the $d$ position is more permissive to negatively charged amino acids than the $a$ position. To test the effect of introducing a negatively charged residue into the $a$ core site of a parallel coiled coil, we made a series of mutated MHCs. The mutant MHCs had one to four mutations introduced at the $a$ sites along the rod at which the original amino acid residues were replaced with Asp, Glu or Leu. The wild-type and mutant proteins were assessed for folding and assembly using a sedimentation assay and imaging by electron microscopy. We show that all the Asp-mutated rods had clear kinks around the points of mutation, while the Glu and Leu mutants had no such kinks. We also show that the assembly properties of the Asp/Glu-mutated rods were impaired compared to the wild-type and Leu mutants, and that the Asp mutants exhibit assembly properties that were more impaired than the Glu mutants.

**Table 2.** Measured total lengths and kink sites of rotary shadowed proteins

<table>
<thead>
<tr>
<th>Protein Variant</th>
<th>Total Length in a.a. (±1 SD)</th>
<th>Kink Site in a.a. (±1 SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smooth muscle chicken gizzard rod (1130 a.a.)</td>
<td>1148.9±62.8</td>
<td></td>
</tr>
<tr>
<td>NM-MHC-642</td>
<td>641.1±46.9</td>
<td></td>
</tr>
<tr>
<td>NM-MHC-642 K(200)D</td>
<td>643.6±47.9</td>
<td>202.3±27.5</td>
</tr>
<tr>
<td>NM-MHC-642 FM(235,242)D</td>
<td>651.4±30.7</td>
<td>237±26.9</td>
</tr>
<tr>
<td>NM-MHC-642 K(390,397)D</td>
<td>634.1±46</td>
<td>409±38.5</td>
</tr>
<tr>
<td>NM-MHC-642 R(453,460)D</td>
<td>635.1±56.9</td>
<td>466.9±19.1</td>
</tr>
<tr>
<td>NM-MHC-642 K,K,R,N (390,397,453,460)D</td>
<td>638.8±47.2</td>
<td>436.5±29.6</td>
</tr>
</tbody>
</table>

The measured distances were converted to a.a.\textsuperscript{*} by the formula: 1a.a.\textsuperscript{*}=0.14 nm. SD, standard deviation.

**Figure 4.** The percent of kinked MHC rods in the different mutants. Sixty-six to 252 (average–145) MHC rods from rotary shadowing images were counted and each MHC rod was marked as straight or kinked. As can be seen, only Asp mutants have a high percentage of kinked rods that is above the noise base line. 1, NM-MHC-642; 2, 3 and 4, NM-MHC-642 M,F(235,242) to D,E, L respectively; 5, 6 and 7, NM-MHC-642 K(390,397) to D, E,L respectively; 8, 9 and 10, NM-MHC-642 R(453,460) to D,E,L respectively; 11, NM-MHC-642 K,K,R,N (390,397,453,460)D; 12, NM-MHC-642 K(200)D.

**Figure 5.** Solubility of the purified MHC rod proteins as a function of salt concentration. MHC proteins (0.5 mg/ml) were dialyzed against a range of salt concentrations (50 mM–300 mM NaCl) in 10 mM phosphate buffer (pH 7.5) and 2 mM MgCl$_2$ for 4 h at 4°C and assayed for solubility by centrifugation of 100,000 g for 1 h. The supernatant and pellet fractions were dissolved on SDS–PAGE and the percentage of soluble MHC was calculated. It is evident that Asp and Glu mutants suffer for some degree of filament destabilization and thus become soluble at lower ionic strength while Leu mutants have a solubility curve similar to the wild-type rod.

**Figure 6.** The percent of kinked MHC rods in the different mutants. Sixty-six to 252 (average–145) MHC rods from rotary shadowing images were counted and each MHC rod was marked as straight or kinked. As can be seen, only Asp mutants have a high percentage of kinked rods that is above the noise base line. 1, NM-MHC-642; 2, 3 and 4, NM-MHC-642 M,F(235,242) to D,E, L respectively; 5, 6 and 7, NM-MHC-642 K(390,397) to D, E,L respectively; 8, 9 and 10, NM-MHC-642 R(453,460) to D,E,L respectively; 11, NM-MHC-642 K,K,R,N (390,397,453,460)D; 12, NM-MHC-642 K(200)D.

### Positively versus negatively charged amino acids in the core positions

The hydrophobic interactions between non-polar residues are thought to be among the most important in determining the three-dimensional structures of proteins.\textsuperscript{16,39} Indeed, the importance of the hydrophobic nature of the core positions to the folding and stability of the coiled coil is well documented.\textsuperscript{12–19} Thus, all polar and charged residues, in general terms, would not be considered as good candidates for the $a$ and $d$ sites. This is a generalization, however, and does not apply equally to all polar residues. For instance, Asn is actually favoured at certain $a$ sites of parallel dimeric coiled coils such as the leucine zipper.\textsuperscript{20,21} The inclusion of
this residue is well understood: two Asn residues at the same α layer in a coiled-coil dimer can form a side-chain–side-chain hydrogen bond that specifies coiled-coil oligomer state, orientation and register, as this specific interaction is only possible in in-register, parallel dimers. Returning to Asp and Glu versus Lys and Arg: firstly, these charged side-chains are not accommodated within the hydrophobic cores of parallel coiled coils as readily as Asn, and they must escape the core to have their terminal charged groups solvated. In this respect Lys and Arg side-chains are amphipathic with aliphatic alkyl chains terminating with positively charged amino and guanidino groups, respectively. The aliphatic component may add to the hydrophobic interface between the two α-helical chains, or act as a spacer to help remove the charged termini to solvent.21 In contrast, the side-chains of Asp and Glu are shorter, and might be expected to be less tolerated at core sites.32,33 Indeed, Hodges and colleagues showed that introducing a negatively charged amino acid at these positions of a de novo designed coiled-coil peptide drastically destabilizes its structure.18,40,41

Aspartic versus glutamic acid in the core positions

As shown in Table 1 and Figures 2, 3 and 4, Asp is by far less tolerated in the core positions as compared to Glu. This is probably due to the fact that Asp has the shortest side-chain of all of the polar charged residues.33 Furthermore, Asp, like the other short polar side-chains, Asn, Ser and Thr, is considered an α-helix breaker. This is related to the ability of their side-chains to interact with the main-chain and compete with the normal α-helical hydrogen-bonding pattern.42 These two points, namely, the requirement to solvate the short side-chain of Asp, and the ability of the residue to form side-chain–main-chain hydrogen bonds, are the likely causes of the exclusion of Asp from the hydrophobic core sites of parallel coiled coils (Table 1B), and the sharp kinking in the rods that we observe in the Asp-containing myosin mutants (Figures 2–4).

Note, that as the crystallographic structures of the MHC rods have not been determined, we cannot be sure that the two rare incidences of Asp in position d of the MHC rods (Table 1A; NM-MHC-B, heptad 78 and skeletal MHC, heptad 146) do not cause some local disruption of the coiled-coil structure.

The α versus d core positions

As shown, side-chains in the α position show a structural geometry that is different from that of side-chains at d.22 For two stranded coiled-coils these structural geometries were described as parallel and perpendicular, respectively. This is the basis for the work of Harbury and colleagues regarding the influence of these different packing angles on the amino acid selection at the core sites.11,23 Indeed, even though both α and d sites lie in the hydrophobic core of the coiled coil and both are considered to be mainly hydrophobic, there is a
clear difference in the amino acid selection between these two positions (Table 1). 10,11,33,43

A computer model for the packing of side-chains in the inner core of coiled coils has shown that Glu at the d position can change its conformation by changes in the plane of the COOH group and thus contact different partners, and a hydrogen bond may form between the COOH group and a carbonyl group from the main-chain of the other strand. 44 As proposed, 33 based on sequence analysis of representative coiled coils, the virtual absence of Glu from the a positions suggests that the local structure of the coiled coil around this position does not in fact allow Glu to be solvated, whereas in the d position this apparently does become possible. Since this analysis was based on the sequences of only four representative coiled coils, all of them of having parallel conformation, we can now state more accurately that this rule does not generally apply to antiparallel coiled coils (Table 1B). Indeed, it has been shown that the parallel structures show greater discrimination of amino acid residues between the two core sites as compared to antiparallel structures. 11

Aspartic acid in an a position kinks the coiled coil and affects filament assembly

Introducing Asp into several a positions along the MHC coiled coil caused a clear kink at the mutation site (Figures 2–4; Table 2). The maximal angle of the kink was directly related to the number of mutations that were introduced (Figure 3). However, there was a great variation of the kink angles between rods with the same mutation, with only about 20% of the rods showing the maximal kink angle. We suggest that the mutations cause local disruptions of the coiled coils, which then serve as flexible linkers between the coiled-coil regions on either side. The number of these local disruptions probably determines the maximal kink angle. To our knowledge this is the first report of such an effect, resulting in lower salt concentrations compared to wild-type MHC (Figure 5, and data not shown). The maximal kink angle was directly related to the number of mutations introduced to the a position of a homodimer. Indeed, we have shown that Asp, and not Glu, is absent from the core positions of all parallel coiled coils gathered from the Protein Data Bank using the SOCKET algorithm; while Glu can be found at these positions. We also show for the first time that introducing Asp, but not Glu, into the a position of a parallel coiled coil, leads to the formation of kinks in the coiled-coil structure, and that the maximal kink angle is directly related to the number of mutations that were introduced. Furthermore, we demonstrate that the introduction of a kink along the coiled coil interferes with its ability to play its critical role in fostering protein–protein interactions and forming higher-ordered structures.

These new observations should help achieve better predictions of coiled-coil motifs from sequence analysis: for instance, in helping to delimit coiled-coil regions and possibly even to help resolve conflicts in oligomer-state and topology predictions of coiled-coil structure. They could also be used as a strong negative-design rule when planning de novo coiled-coil designs. We also believe that the described flexible hinge/kink in the coiled-coil can be used as a positive design rule. Considerable effort in protein engineering has been devoted to introducing variations into otherwise natural protein sequences to achieve new or altered functions or properties. 44,45 Several research groups have begun to take protein design to the next level of organization by designing self-assembling systems. Most of these studies have focused on particularly simple protein-folding patterns, especially, some simple coiled-coil arrangements. Recently, new strategies for designing more complex protein assemblies, such as filaments, cages and extended ordered arrays have been developed. 24,31,44 The strategic introduction of Asp at an a site of an otherwise canonical coiled coil that assembled to a fibrous biomaterial could be used to kink the structures.

Materials and Methods

Myosin heavy chain sequences

We have selected 14 MHC rods (15,489 amino acid residue(s) (a.a.)) that represent the different subtypes of MHC in different species and assigned them to their correct heptad register (Supplementary Data, Table S1).
The first and last heptads of the rod region were not included in the database because of uncertainties as to the coiled-coil boundaries. The ten residues before and after each skip residue were not included because of uncertainties. Methionine and the skips were removed, and the entire helical sequence of the coil can be mapped out as a cyclic pattern of 28 amino acid residues that is interrupted at three or four positions by the insertion of one extra residue that is termed a “skip residue”. To reduce bias resulting from including more than one representative structure of each MHC subtype, only the five human MHC subtypes were used for the statistical analysis (Table 1A). We compiled a table for each myosin rod that represents the number of times that each amino acid is found at each of the heptad positions. Then we summed all five tables and compiled a table that represents the frequency of each of the 20 amino acid residues at each heptad position (20/7 frequency table). The resulting profile was then normalized by dividing that frequency by the relative frequency of each amino acid in SWISSPROT (Table 1B; Supplementary Data, Table S2).

**SOCKET derived coiled-coil database**

The SOCKET database contains all 561 structures that were identified by SOCKET to have a coiled-coil region out of 9255 structures in release 87 of the RCSB Protein Data Bank. To reduce bias resulting from duplication of proteins or their homologues in the RCSB Protein Data Bank, the positive coiled-coil structures were grouped into 92 sequence-based homologous families and only representative structures of each family were included in the statistical analysis. We created an Excel macro program that uses the database to compile a table that represents the frequency of each of the 20 amino acids at each heptad position (20/7 frequency table). The resulting profile was then normalized by dividing that frequency by the relative frequency of each amino acid in SWISSPROT (Table 1B; Supplementary Data, Tables S3–4).

**Cloning and site directed mutagenesis of NM-MHC-B 642 a.a. proteins**

To create the NM-MHC-642 protein we used the described PET21-Rod vector that contains the C-terminal 657 a.a. of human NM-MHC-B in PET21c. We inserted a stop codon just after the last NM-MHC-B amino acid using the primers 5'-GAT CCC GGG CCC GCG and 5'-GAA CTT GAA AAA TCC. All site-directed mutageneses were done with ExSite™ PCR-Based Site-Directed Mutagenesis Kit (20052; Stratagene). We then cut the rod sequence using BamHI-BamHI restriction enzyme and prepared an empty PET21c vector by cutting it with Nhel-BamHI restriction enzymes. Both insert and vector were treated with Klenow polymerase to fill-in the overhang arms and the insert was then sub-cloned into the empty PET21c vector. The resulting expression vector was fully sequenced and was shown to code for a protein with methionine and alanine at both 5′ and then the C-terminal 640 a.a. of human non-muscle MHC-B (accession no. A99252). We then used the site-directed mutagenesis kit to change several amino acids at α position into Asp, Glu or Leu. Here are the primers that were used to create the Asp mutants (mutagenesis sites are underlined): For NM-MHC-642 K200D: 5'-C GAA CTT GAA AAA TCC GAC CGG CCA CTA GAG CAG GGT G and 5'-C CAG CTG CTC TAG GGC CCG GGT GCA AGT TTT TTC AAG TTC G. For NM-MHC-642 R.N(453,460): 5'-GT GAC AAT GCA GAC CAA CTG GAG CGG CAG GAC AAG GTC GGT and CAAGCTC GGACGAGCTC GGAGTCGACG CGC TGCTGCTG ATG TGC TGC ATT GTC A. For NM-MHC-642 K(390,397): 5'-CG GCC TCT GGC GAC TCC GGG CTG CTG GAT GAC GAG CCG CGT CTG G and 5'- C CAG AGC GGG CTG CTC ATC CAG CAG CGG CGA GTC GCC AGA GGC GC. NM-MHC-642 M.F(235,242): 5'-G CAG GTC ATC GAC GAC GCC ATG AAG GCC GAG GAG AGA GAC C and 5'-G TGC TCT CTC GTC CGC TTT CAT GAC CTG GTC GTT GAC CTC.

**Bacterial expression and purification of NM-MHC-B 642 a.a. proteins**

Expression and purification of myosin proteins was carried out essentially as described with only a few changes. Bacteria containing the expressed vectors were grown at 37 °C in 500 ml of 2xYT + Amp to an A595 nm of 0.5, then IPTG was added to 0.5 mM and the cells were grown for additional 2 h. The bacteria were pelleted, washed with cold phosphate-buffered saline (PBS), pelleted again and frozen at −20 °C overnight. The bacterial pellet was suspended in 15 ml of lysis buffer B (50 mM Tris-HCl (pH 7.5), 0.8 M NaCl, 10 mM EDTA, 10 mM EGTA, 1 mM DTT, 1 mg/ml of lysozyme, 0.5 mM PMSF), lysed using sonication and the cell debris removed by centrifugation. Polyethyleneimine (PEI) (250 μl) (Sigma, P-3143) was slowly added from a 10% stock solution in water and incubated for 10 min of stirring on ice before being centrifuged (Beckman Ti-60 rotor, 45,000 rpm at 4 °C for 1 h). To get rid of any trace of PEI, the purified myosin rods were precipitated twice with ammonium sulphate (1.5 ml saturated ammonium sulphate/ml myosin solution), suspended with 8 ml of buffer G (20 mM Tris-HCl (pH 7.5), 600 mM NaCl, 5 mM EDTA, 1 mM DTT) and dialyzed overnight against buffer G. The purified myosin rods were stored at −20 °C at concentrations of 0.5–4 mg/ml. The concentrations of the myosin proteins were determined by comparing the densitometry of the band of a sample of each protein on Coomassie-stained PAGE to a known series of bovine serum albumin (BSA) samples on the same gel.

**Sedimentation analysis**

Samples of the expressed myosin rod proteins were diluted with buffer G to 0.5 mg/ml. A 100 μl of the diluted samples of each protein were dialyzed in small dialysis tubes (mini GEBAflex tubes; catalog no. D070-6-30) against a range of salt concentrations (50 mM–300 mM NaCl) in 10 mM phosphate buffer (pH 7.5) and 2 mM MgCl2 for 4 h at 4 °C. Then 80 μl of each dialyzed sample were transferred to centrifugation tubes and centrifuged at 100,000 g for 1 h (Beckman TLA-55 rotor, 45,000 rpm at 4 °C). The top 50% of each supernatant was taken for analysis by SDS–PAGE. The reminder of the supernatant was removed from the tube and the pellet was dissolved in 360 μl of buffer G. Samples of the dissolved pellet were taken for SDS–PAGE. The solubility of the myosin rod proteins (expressed as percentage of the fragment in the supernatant) were determined by measuring the relative
amount of each fragment in the pellet and in the supernatant fractions by densitometry of the Coomassie-stained PAGE bands using Fujifilm image analyser LAS-1000plus and the densitometry program Fujifilm Image Gauge version 3.4.

Rotary shadowing

Samples were diluted to 50 μg/ml in 600 mM ammonium acetate, 50% (v/v) glycerol (pH 7.0) and sprayed onto freshly cleaved mica, then rotary shadowed with platinum at a 6° angle and the replicas picked up on 400 mesh hex grids. The replicas were imaged on a Philips CM 10 transmission electron microscope using an accelerating voltage of 80 kV.

Negative staining

Expressed myosin rods in buffer G were diluted with buffer G to 0.5 mg/ml and then dialyzed against buffer Q (10 mM Tris-HCl (pH 7.5), 50 mM NaCl, 20 mM MgCl2, 25 mM CaCl2) for 16 h at 4 °C in small dialysis tubes (mini GelBlast tubes; catalog no. DO70-6-30). The 7 μl samples were applied onto formvar-carbon coated 400 mesh grids (SPI Supplies, Westchester, PA, USA; 3440c) and left for 2 min and then negatively stained with 2% (w/v) uranyl acetate for 15 s. The negatively stained samples were examined on a transmission electron microscope (Philips CM12). Striation periods were measured at X88K magnification using Fujifilm Image Gauge version 3.4.

Acknowledgements

We thank Dr Ezra Rachamin (Hebrew University, Faculty of Medicine), Mr Oliver Testa and Dr John Walshaw for advice on using the CC+ database, and Dr Howard Cedar for critical reading of the manuscript. We also thank S.R. laboratory members for their support and assistance.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2006.11.083

References


