Journal of Structural Biology **137,** 73–81 (2002)
doi:10.1006/jsbi.2002.4462

Investigating the Tolerance of Coiled-Coil Peptides to Nonheptad Sequence Inserts

Matthew R. Hicks, John Walshaw, and Derek N. Woolfson[1]

Centre for Biomolecular Design and Drug Development, School of Biological Sciences, University of Sussex, Falmer, BN1 9QG, United Kingdom

Received November 21, 2001, and in revised form March 4, 2002

Coiled-coil motifs foster a wide variety of protein–protein interactions. Canonical coiled coils are based on 7-residue repeats, which guide the folding and assembly of amphipathic α-helices. In many cases such repeats remain unbroken for tens to hundreds of residues. However, the sequences of an increasing number of putative and characterised coiled coils digress from this pattern. We probed the consequences of nonheptad inserts using a designed leucine-zipper system. The parent peptide, SKIP0, which had four contiguous heptads, was confirmed as a parallel homodimer by circular dichroism spectroscopy and analytical ultracentrifugation. Seven daughter peptides were constructed in which 1 to 7 alanine residues were inserted between the central heptads of SKIP0. Like SKIP0, SKIP7 formed a stable helical dimer, but the other peptides were highly destabilised, with the order of dimer stability SKIP4 $\gg$ SKIP5 > SKIP6 > SKIP3 > SKIP2 > SKIP1. These results are consistent with an extended theory of coiled-coil assembly in which coiled-coil-compatible motifs are based on 3- and 4-residue spacings and most notably heptad (7-residue) and hendecad (11-residue) repeats. Thus, they help explain why in natural sequences, inserts after canonical heptad repeats most commonly comprise 4 residues. Possible biological roles for nonheptad inserts are discussed. © 2002 Elsevier Science (USA)

INTRODUCTION

Coiled-coil motifs mediate a wide variety of protein–protein interactions with distinct functions and structures. For instance, coiled coils are utilised in transcription factors (O'Shea *et al.*, 1991), in structural fibre networks in muscle and the cytoskeleton (Kreis and Vale, 1999), and in the molecular machineries that mediate mitosis (Wigge *et al.*, 1998) and certain membrane associations (Skehel and Wiley, 2000). From a structural perspective, coiled-coil regions in proteins range in length from tens to hundreds of residues and facilitate assemblies with varied oligomer states and topologies (Lupas, 1996; Walshaw and Woolfson, 2001).

Despite the variety of functions and structures, the accepted hallmark of coiled-coil sequences is a 7-residue repeat also known as the heptad. The positions of this repeat are traditionally labelled *a* to *g*. Residues at the *a* and *d* sites are predominantly hydrophobic. Thus, in this pattern, the hydrophobic residues are alternately separated by three and four positions along the amino acid sequence, hence the term "3,4-hydrophobic repeat." When configured into an α-helix this pattern presents the hydrophobic *a* and *d* positions on the same face of the helix, resulting in a hydrophobic stripe. However, as the α-helix has 3.6 residues per turn, one heptad repeat falls short of two complete helical turns and the hydrophobic stripe drifts around the surface of the helix in a left-handed direction, i.e., in the opposite sense to the twist of the α-helix. As a consequence, coiled-coil helices supercoil, or wrap around one another to maintain maximum contact between their hydrophobic stripes (Crick, 1953). These structural consequences of canonical coiled-coil sequences may be seen in any of the number of coiled-coil structures in the Brookhaven Protein Data Bank (PDB) (Berman *et al.*, 2000) and those of our own recently compiled coiled-coil-structure database, available at http://www.biols.susx.ac.uk/coiledcoils (Walshaw and Woolfson, 2001).

Recently, however, a number of protein sequences and a small number of protein structures have been described for coiled coils that do not fit the heptad-based model completely (Brown *et al.*, 1996; Hicks *et al.*, 1997; Lupas, 1996). For instance, regular heptad patterns may be disrupted or broken by nonheptad inserts. Generally, these motifs and the resulting structures are described as *noncanonical* coiled

[1] To whom correspondence should be addressed. Fax: +44 1273 678433. E-mail: dek@biols.susx.ac.uk.

1047-8477/02 $35.00
© 2002 Elsevier Science (USA)
All rights reserved.

coils. One formalism for these is given by Brown *et al.* (1996), who describe such sequences in terms of deletions and insertions in the heptad repeats. For instance, a deletion of three residues, or an insertion of four, is termed a stutter, whereas the reverse is described as a stammer. In this model such inserts are accommodated in the coiled-coil structures by under- and overwinding of the α-helix.

More generally, repeats that may be compatible with coiled-coil structures are those that lead to hydrophobic residues falling on the *same face* of the helix, namely, those with apolar residues separated by combinations of 3 and 4 residues. After the heptad, the simplest solutions are 10- and 11-residue units, which have been dubbed decads and hendecads (Hicks *et al.,* 1997), i.e., those with 3,4,3- and 3,4,4-hydrophobic spacings, respectively. As Brown and colleagues describe for certain protein sequences and as observed in a small number of protein structures (see below) there are isolated examples of such inserts. However, a group of proteins from *Giardia lamblia* exhibit unusually long and regular noncanonical repeats with 3,4,3,4,4,3,4 (7-11-7; HPSR2), 3,4,4,3,4,3,4,4 (11-7-11; β-giardin), and 3,4,3,4,3,3,4 (7-10-7; median body protein) (Hicks *et al.,* 1997; Holberton *et al.,* 1988; Marshall and Holberton, 1993, 1995). Peptides corresponding to the 7-11-7 and 7-10-7 repeats have been shown to form helical dimers consistent with the coiled coil (Hicks, 2000; Hicks *et al.,* 1997). Also, a distant sequence homologue of β-giardin—namely, SF-assemblin from green flagellate algae—forms coiled-coil-like structures (Lechtreck, 1998; Lechtreck *et al.,* 1996; Lechtreck and Melkonain, 1995; Weber *et al.,* 1993).

Recently we described an algorithm, SOCKET, which recognises the knobs-into-holes packing characteristic of solved coiled-coil structures (Walshaw and Woolfson, 2001). We applied SOCKET to the PDB and collated the aforementioned database of coiled-coil structures. As SOCKET identifies knob residues and assigns them as *a* or *d* knobs, the heptad register of the sequences of the databases is automatically assigned. On this basis, we searched for noncanonical coiled-coil structures in which the canonical alternating *adad* . . . pattern was broken. As described below, this search returned a disappointingly small number of noncanonical coiled coils, most of which had been described in the original papers reporting the full protein structures.

There was one example of a homodimeric noncanonical coiled coil in GrpE (PDB code 1DKG (Harrison *et al.,* 1997)). One way to view residues 44–86 of this sequence is that they exhibit a 7-11-7-11-7 pattern. Thus, the canonical heptad repeat and noncanonical hendecad repeat alternate. As the predicted supercoils for the component motifs are almost exactly equal and opposite, the overall result is expected to be a straight, or nonsupercoiled, helix–helix interface. This is indeed what is observed. However, the helices are shifted relative to each other along the helix axis and, as a consequence, the knobs-into-holes packing is somewhat compromised; indeed, our recommended SOCKET parameters had to be relaxed slightly to identify this structure as a coiled coil. The hemagglutinin structures (represented by PDB code 1HTM (Bullough *et al.,* 1994)) contain a homotrimeric coiled coil in which residues 45–69 may be considered as a hendecad flanked by heptad repeats. Again this composite alters the supercoiling of the overall coiled coil, but not to the extent observed in GrpE. The Sendai virus phosphoprotein (PDB code 1EZJ (Tarbouriech *et al.,* 2000)) provides a tetrameric example of a coiled coil that contains unusual sequence motifs; residues 64–81 may be considered as a hendecad followed by a heptad. Finally, residues 57–74 in the structure of the ectodomain from the envelope glycoprotein of Ebola virus—PDB code 1EBO (Weissenhorn *et al.,* 1998)—tested positive for a hendecad–heptad composite. However, this example is a fusion between the C-terminus of a GCN4 peptide and the N-terminus of the viral protein, and, in another structure without fusion, the noncanonical repeat is not seen. The reason for this is that in the latter the helices fray and the noncanonical core position is lost as noted by the authors. Furthermore, a homologue of the Ebola protein, human respiratory syncytial virus fusion protein core (PDB code 1G2C (Zhao *et al.,* 2000)), does not show the noncanonical structure.

In addition to these examples there are examples of structures based on pure noncanonical repeats. For example, these include two all-hendecad-based structures, namely, tetrabrachion (PDB code 1FE6 (Peters *et al.,* 1996; Stetefeld *et al.,* 2000)) from the archaebacterium *Staphylothermus marinus* and a peptide of *de novo* design (1RH4 (Harbury *et al.,* 1998)). Both of these structures contain right-handed coiled coils as predicted for hendecad-based sequences.

Thus, the examples of fully characterised, noncanonical, coiled-coil structures appear to be limited to those with heptad repeats interspersed with 11-residue repeats or those based solely on the latter. However, the number of available structures for noncanonical coiled coils is limited. Thus, given that there are putative coiled-coil sequences with other nonheptad repeats (Brown *et al.,* 1996; Hicks *et al.,* 1997), a pertinent question is: are 11-residue digressions from the heptad paradigm the only ones compatible with the coiled-coil structure?

To probe the consequences of different length in-

serts into canonical coiled-coil repeats more generally, we took a protein-design approach. First, we designed a four-heptad leucine-zipper peptide, SKIP0, and confirmed that this formed parallel, helical dimers in solution. We then made a series of peptides, SKIP1–7, in which alanine residues were inserted between the second and third heptads of SKIP0. Solution-phase analysis of these revealed that all of the structures were destabilised, but that the seven- and four-residue inserts were the most tolerated. This suggests that the best building block for the coiled coil is indeed the heptad repeat followed by the hendecad. However, because the heptad-based structures were significantly stabler than those with nonheptad inserts, it raises the question why noncanonical motifs are found at all. Possible reasons are that (1) they are simply locked in during evolution; (2) they provide a mechanism for modulating stability and dynamics, particularly in long coiled coils; (3) they aid coiled-coil alignment/register and partner selection; and/or (4) they allow subtle variations in the geometries of the coiled-coil interfaces that they foster, which may have functional consequences.

MATERIALS AND METHODS

Peptide synthesis. Peptides were synthesised on an Applied Biosystems 432A automated, continuous-flow peptide synthesiser using solid-phase Fmoc chemistry. The sequences of the SKIP peptides were

$$\text{Ac-IAALERKNAALEQK (A)}_{0-7}\text{IAALEYKIAALEKK-NH}_2.$$

All peptides were N-terminally acetylated and C-terminally amidated. The peptides were cleaved from the resin using a standard mixture of thioanisole, triisopropylsilane, ethanedithiol, and trifluoroacetic acid and extracted using methyl tert-butyl ether. The peptides were then dissolved in water and freeze-dried. Peptides were purified by HPLC on a Brownlee C8 semipreparative column (1 cm × 20 cm) using a water–acetonitrile gradient and trifluoroacetic acid as the ion-pairing agent. Fractions were analysed by HPLC using a Walters Deltapak C18 analytical column (0.39 cm × 15 cm), and peptides were identified by MALDI-TOF mass spectrometry.

The concentrations of pure peptide solutions were determined from the tyrosine absorbance at 280 nm in 6 M guanidine hydrochloride according to the method of Gill and Von Hippel (1989).

Circular dichroism spectroscopy. Circular dichroism (CD) spectroscopy was carried out as described previously (Hicks *et al.*, 1997) using a Jasco J-715 spectropolarimeter fitted with a 6-cell changer Peltier temperature controller. The instrument was calibrated using a 0.06% (w/v) solution of ammonium *d*-10-camphor sulphonate, and the wavelength was calibrated using the absorbance of neodymium glass. Cuvettes of 0.1- and 1.0-cm path length were used (Hellma GmbH & Co. and Starna Ltd.). Spectra were collected in stepped mode with a resolution of 1.0 nm, averaging data for 4, 8, or 16 s at each point depending on sample concentration. The buffer system used was 25 mM potassium phosphate, 100 mM sodium chloride. Temperature dependencies of the CD signals and spectra were measured using a ramping rate of 1°C/min. The midpoint of a thermal-unfolding transition (T_M) was taken as the maximum of the first derivative of the CD signal (at 222 nm) versus the temperature curve, from a largely helical structure at low temperature to a largely random coil structure at high temperature. In all experiments baseline signals for buffer alone were measured and subtracted from the raw peptide data.

Analytical ultracentrifugation. Sedimentation equilibrium studies were carried out using a Beckman XL-I analytical ultracentrifuge fitted with a titanium An60-Ti rotor. Starting peptide concentrations were 250 μM with 100-μl loading volumes. Samples contained 25 mM potassium phosphate, 100 mM sodium chloride. Rotor speeds of 45 000 and 55 000 rpm were used. Once equilibrium had been established at each speed, absorbance data were collected at 280 nm by recording and averaging 10 replicates at stepped points using a radial step size of 0.001 cm. All experiments were performed at 20°C. Partial specific volumes (SKIP0 to SKIP7 = 0.7713 to 0.7666 liter $\cdot$ g^{-1}) were calculated from the amino acid composition of each peptide using the program Sednterp (Laue *et al.*, 1992), as were the buffer densities. Data were analysed using the programs WinNonlin (Johnson *et al.*, 1981) and Origin (OriginLab Corp.).

RESULTS

Rationale of the Peptide Designs

The designed leucine-zipper peptide, SKIP0, had the following sequence:

```
acetyl-IAALERK NAALEQK IAALEYK IAALEKK-amide.
       abcdefg abcdefg abcdefg abcdefg
```

Isoleucine was placed at *a* positions and leucine at *d* positions to direct dimer formation. These placements were made on the basis of Harbury's oligomer-state selection rules, which follow from the packing considerations in the core of coiled-coil dimers (see Harbury *et al.*, 1993; Lupas, 1996; Woolfson and Alber, 1995). The asparagine at the *a* position of the second heptad was included to add dimer specificity, reduce thermal stability, and ensure parallel, in-register chain alignment (Gonzalez *et al.*, 1996; Lumb and Kim, 1995; Pandya *et al.*, 2000; Woolfson and Alber, 1995). The glutamic acid and lysine residues at *e* and *g*, respectively, were placed to specify the structure further by favouring interhelical salt bridges. The *b* and *c* positions were made alanine, which has a short side chain and a high helix propensity. To give the peptide a net charge at neutral pH to aid solubility, additional positively charged residues were placed at two *f* positions; these were in the first and fourth heptads to minimise intrahelical electrostatic repulsion. The remaining *f* positions were made tyrosine to allow concentration determinations by UV spectrophotometry and glutamine to help with solubility.

To complete the SKIP series, peptides were made with one to seven alanine residues inserted between the second and third heptads of SKIP0 to give pep-

TABLE I

Number of inserted alanines	Peptide name	Repeat	Best possible hydrophobic repeat
0	SKIP0	7-7-7-7	3, 4, 3, 4, 3, 4, 3
1	SKIP1	7-7-1-7-7	3, 4, 3, 4, 1, 3, 4, 3
2	SKIP2	7-7-2-7-7	3, 4, 3, 4, 2, 3, 4, 3
3	SKIP3	7-7-3-7-7	3, 4, 3, 4, 3, 3, 4, 3
4	SKIP4	7-7-4-7-7	3, 4, 3, 4, 4, 3, 4, 3
5	SKIP5	7-7-5-7-7	3, 4, 3, 4, 5, 3, 4, 3
6	SKIP6	7-7-6-7-7	3, 4, 3, (3, 3), 3, 4, 3
7	SKIP7	7-7-7-7-7	3, 4, 3, 4, (3, 4), 3, 4, 3

tides, SKIPX (X = 1, 2, 3, 4, 5, 6, or 7), with sequences that are summarised in Table I. Alanine was the residue chosen for the inserts for two reasons: first, it has a high helix propensity and so will not disrupt helix formation *per se*. Second, it is compatible with both the interiors and exteriors of proteins (Janin, 1979).

As illustrated in the helical-net diagrams of Fig. 1, for fully helical conformations at least, the hydrophobic stripes set up by the different SKIP peptides

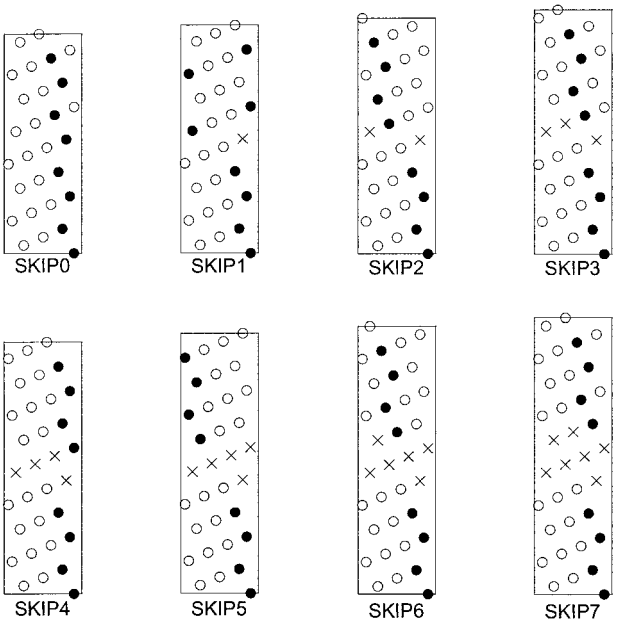

FIG. 1. Helical net representations of the surface of an α-helix. The Cα positions are plotted on a cylinder wrapped around the helix. This cylinder is then opened out flat to give a representation of the surface of the helix in two dimensions. The hydrophobic core residues are shown as black circles and the alanine inserts are shown as crosses. The hydrophobic stripe seen in SKIP0 is least disrupted in SKIP7. The hydrophobic stripes in SKIP4 and SKIP3 are also disrupted to a lesser extent than in SKIPs 1, 2, 5, and 6. This representation of course assumes completely helical conformations of the peptides, which may not be correct.

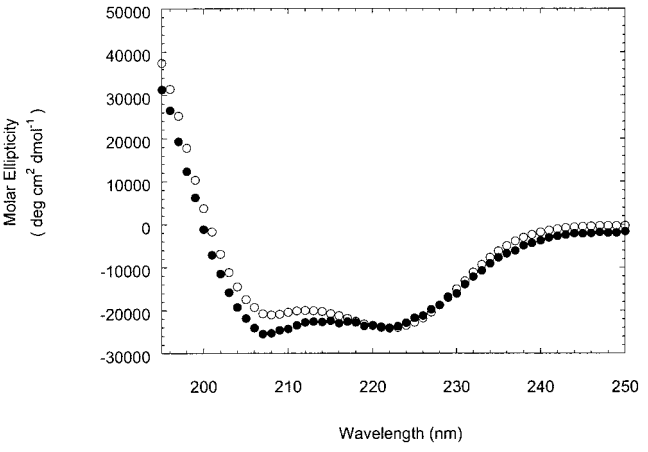

FIG. 2. CD spectra for SKIP0, the *de novo* designed leucine-zipper peptide. These were recorded at 10 μM and 0°C in 25 mM potassium phosphate buffer at pH 7.0 (open circles), and under similar conditions but in the presence of 50% TFE (closed circles).

are predicted to vary considerably; some have continuous or near-continuous stripes (0, 3, 4, and 7), whereas in others the stripe is broken (1, 2, 5, and 6). By constructing and characterising these peptides in terms of stability of the structures formed we aimed to test how a model coiled-coil system responded to nonheptad inserts.

Confirmation of the Dimer Design

Helicity and thermostability. The first step in characterising the SKIP-peptide series was to establish that SKIP0 folded to a helical dimer as intended. The α-helicity of SKIP0 was determined by comparing its CD spectra recorded in aqueous buffer and water–trifluoroethanol (TFE) mixtures (Hicks *et al.*, 1997). This method was employed for two reasons: first, due to the influence of side chains on the backbone CD signal, the ellipticity for 100% helix varies between systems (Chakrabartty *et al.*, 1993; Kallenbach *et al.*, 1996; Woody and Dunker, 1996). Second, we anticipated that the folded conformations of some of the SKIP peptides would be severely compromised. For both reasons a benchmark ellipticity value for 100% helix was required for each peptide. The use of TFE facilitates this because it induces helical structure (Hicks *et al.*, 1997; Jasanoff and Fersht, 1994).

At 10 μM concentration SKIP0 showed a typical α-helical spectrum with double minima at 208 and 222 nm and positive signal below 200 nm (Fig. 2). In this case, the signal at 222 nm, which is due mostly to α-helical contributions, did not increase upon addition of 50% TFE. This suggested that the peptide was fully helical at the 10 μM concentration in aqueous solution.

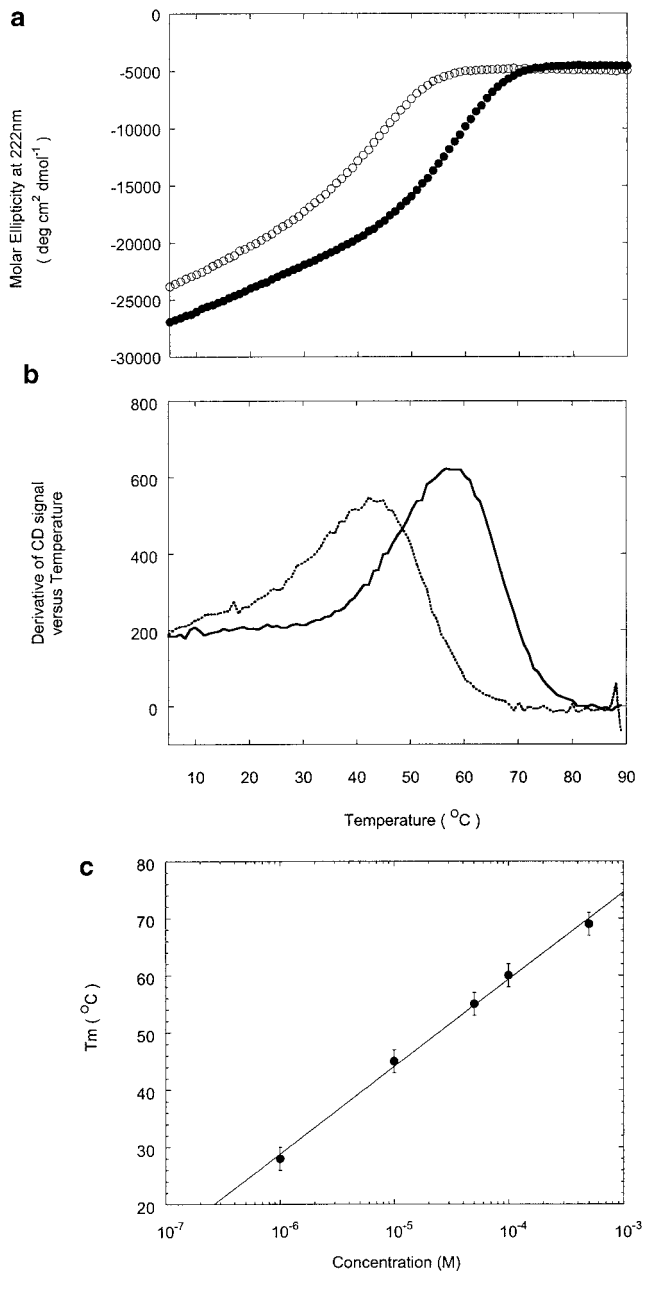

FIG. 3. Thermal stability of peptide SKIP0. (a) Thermal unfolding curves measured by the change in CD signal at 222 nm with temperature for samples of SKIP0 at 10 μM (open circles) and at 100 μM (closed circles). (b) The first derivative of the data in (a) for the 10 μM (dashed line) and 100 μM (solid line) samples. The maxima of the first derivative plots were used to determine the midpoints of the transitions (the T_M). (c) The T_M of SKIP0 at different peptide concentrations, determined using the analysis described in (a) and (b). Samples for the unfolding experiments were in 25 mM potassium phosphate buffer at pH 7.0.

The thermal stability of SKIP0 at 10 μM was also measured by following the CD signal at 222 nm with increasing temperature (Fig. 3a). The thermal unfolding transition observed was sigmoidal, characteristic of a cooperative unfolding process from a largely helical structure at low temperature to a largely random coil structure at high temperature. The steepness of these transitions is related to the enthalpy of unfolding. Usually for small peptides this is not large unless they oligomerise to form stable, uniquely folded species. Therefore, the steep thermal-unfolding transition exhibited by SKIP0 was evidence for oligomerisation. Accordingly, we found that the midpoint of the thermal unfolding curve (the T_M) for SKIP0 increased with increasing peptide concentration: the T_M increased by 41°C from 28 to 69°C between 1 and 500 μM (Fig. 3c).

Oligomerisation state. To determine the molecular weight of SKIP0 in solution, sedimentation equilibrium experiments were performed in the analytical ultracentrifuge. From the sequence and mass spectrometry, the monomer molecular weight of SKIP0 was 3140 Da. A global analysis of sedimentation-equilibrium data obtained at different rotor speeds assuming a single, ideal species gave good fits and a molecular weight of 7130 Da (95% confidence: 6725 and 7527 Da) (data not shown), which is higher than expected for a dimer. Nonetheless, at the temperature and over the concentration range of the sedimentation-equilibrium experiment (20°C and $\approx$ 150–350 μM) SKIP0 was fully folded by CD spectroscopy (Fig. 3), which discounts possibilities for equilibria between monomers and oligomer states higher than dimer. The discrepancy from the expected molecular weight value for a dimer and that determined by analytical ultracentrifugation is likely to be due to the small size of the SKIP peptides, errors in the estimation of the partial specific volume of the peptide, and nonideal behaviour of the solute.

To be sure that the dimer state of SKIP0 was the dominant preferred species we resynthesised the peptide with an N-terminal Cys-Gly-Gly linker. In the reduced form this behaved similarly to SKIP0 in thermal unfolding experiments: T_M's at 2.5 and 25 μM were 31 $\pm$ 2°C and 50 $\pm$ 2°C, respectively. However, the oxidised peptide was fully helical and the thermal denaturation experiments and the T_M's that these returned showed no concentration dependence: T_M's at 2.5 and 25 μM were 79 $\pm$ 2°C and 78 $\pm$ 1°C, respectively. Given these analyses the most plausible model for the association of SKIP0 is a parallel dimer.

Confirmation of Tolerance to a Heptad Insert of Alanine

The α-helicity, oligomer state, and stability of the SKIP7 peptide were probed as described above for SKIP0 (Fig. 4). In summary, at 100 μM concentra-

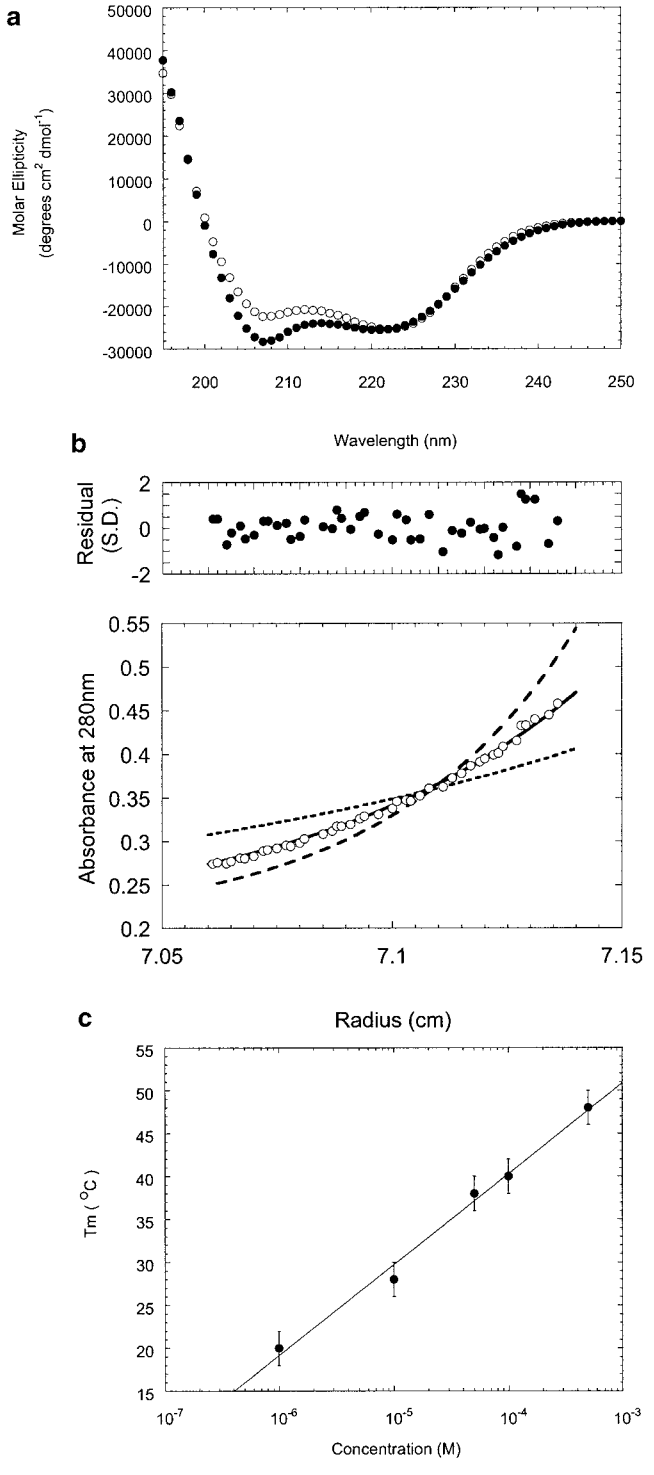

FIG. 4. Characterisation of peptide SKIP7. (a) Spectra of SKIP7 in the absence of TFE (open circles) and 67.5% TFE (closed circles). These data were recorded on 100 μM samples in 25 mM potassium phosphate buffer at pH 7.0 and at 5°C. (b) (Bottom) Sedimentation equilibrium data collected at a rotor speed of 55 000 rpm and a loading concentration of 250 μM; the buffer was 25 mM potassium phosphate and 100 mM NaCl at pH 7.0. The experimental data set is shown as open circles. Lines representing ideal monomer (short dashes), dimer (solid line), and trimer (long dashes) are shown for comparison. (b) (Top) The residuals (filled circles) give the difference between the fit and the experimental data in standard deviations. (c) The T_M of SKIP7 at different peptide concentrations determined using the analysis described in the legend to Fig. 3.

tion SKIP7 was fully helical, but less stable than the SKIP0 peptide. At 100 μM the T_M was reduced by 20°C from 60 to 40°C by the seven-alanine insert. AUC studies of SKIP7 in water gave a weight-averaged molecular weight of 7518 Da. (95% confidence: 6973, 8056 Da). The molecular weight of a SKIP7 dimer is 7274 Da. As for SKIP0, the monomer–dimer dissociation constant could not be determined from the AUC data. These results are consistent with SKIP7, like SKIP0, forming a fully helical dimer in aqueous solution at moderate (micromolar) concentrations.

The Effects of Nonheptad Inserts

Having established the validity of the design of the template, SKIP0, and SKIP7, the relative coiled-coil-forming propensities of the SKIP peptides with nonheptad inserts of alanines were determined using CD spectroscopy and AUC. CD spectra of SKIP1–6 measured in the absence and presence of 50% TFE are shown in Fig. 5, and the calculated percentage helicities are plotted *versus* SKIP number in Fig. 6a. Again, to test the relative stabilities of the peptides, thermal denaturation curves were measured using CD spectroscopy. It was not possible to assess confidently the midpoints of unfolding for some of the peptides because the peptides were too unstable and unfolded baselines were not apparent; in these cases the T_M was taken as $<$10°C. The apparent molecular weights and the monomer–dimer K_d's were assessed by AUC. Unlike SKIP0 and SKIP7, estimates of the dissociation constants were possible for SKIP1–6 because their lower stabilities meant that appreciable amounts of monomer and oligomer were accessible in the experiments. The results of the CD and AUC experiments on SKIP1–6 are summarised in Table II and Fig. 6.

Inspection of these data revealed (1) that peptides SKIP1–6 were all destabilised relative to SKIP0 and SKIP7. (2) There was a general increase in helicity in water with SKIP number (Fig. 6a), which possibly reflected the increased helical propensity with increased alanine content. (3) The one peptide with a nonheptad insert that exhibited increased helicity and thermal stability above the trend was SKIP4; furthermore, this peptide was also the only one, other than SKIP0 and SKIP7, with a submillimolar dissociation constant (Fig. 6b). The outlying SKIP4

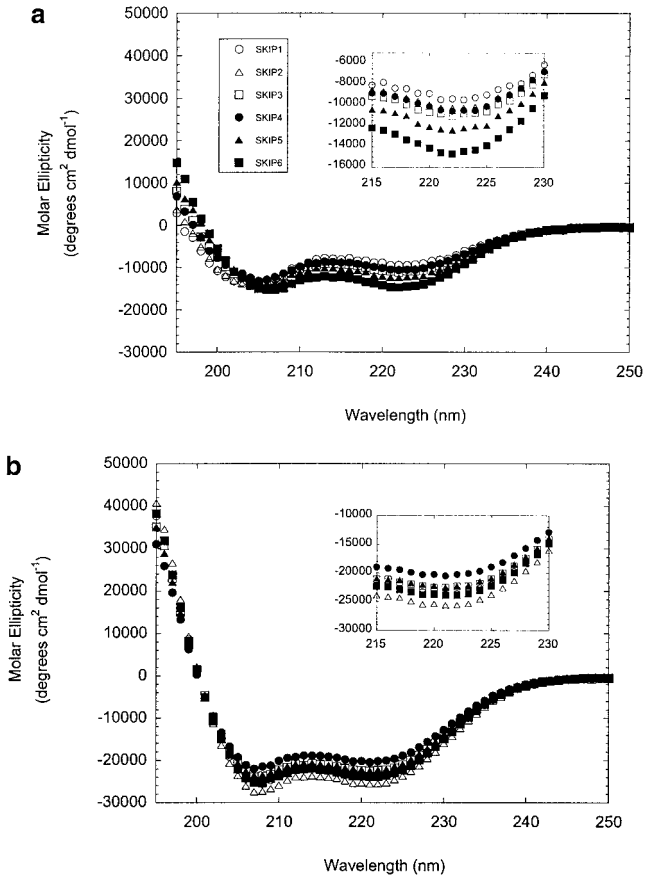

FIG. 5. CD spectra of SKIP1–6. (a) Spectra were measured in 25 mM potassium phosphate, 100 mM NaCl at pH 7.0 and at 5°C. An inset showing the signal around 222 nm is included for clarity. (b) Spectra of SKIP1–6 recorded under conditions similar to those in (a) but in the presence of 50% (v/v) TFE. Peptide concentration was 50 μM.

is interesting as it may be described as having an 11-residue repeat. The significance of this is discussed below.

DISCUSSION

The discovery of coiled-coil sequences with nonheptad (noncanonical) sequence motifs prompted us to investigate the effects of hydrophobic-seam displacement in the context of a simple leucine-zipper system. To achieve this, we designed a canonical, four-heptad, dimeric coiled coil (SKIP0) and systematically inserted blocks of one to seven alanine residues between the central heptad repeats to create SKIP1–7. The thermal stabilities and dissociation constants for the resulting eight peptides were determined using a combination of CD spectroscopy and sedimentation-equilibrium studies in the analytical ultracentrifuge.

SKIP0 was confirmed as a parallel, dimeric coiled coil. SKIP7 also folded to a stable, dimeric coiled coil, but with a slightly lower thermal stability. The reduced stability relative to SKIP0 is possibly due to reduced packing by the alanine residues at the two potential core positions of the insert. By contrast, the folded states of SKIP1–6 were significantly less stable than those for the purely heptad-based se-

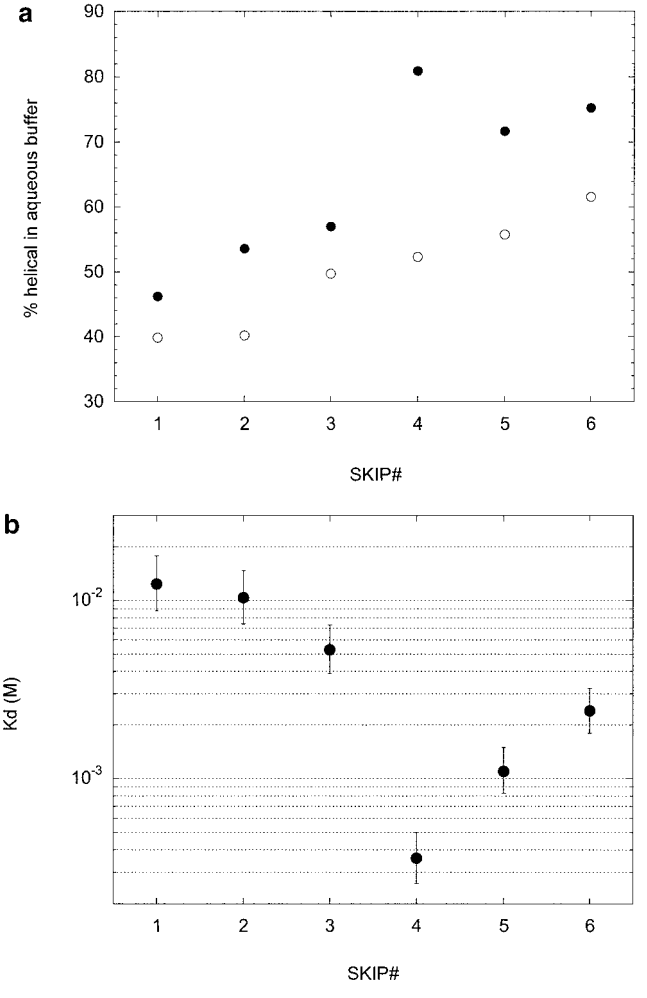

FIG. 6. Percentage helix and dissociation constants for SKIP1–6. (a) Percentage helix of nonheptad SKIP peptides in aqueous buffer versus SKIP number. The percentage helix was calculated using the signal at 222 nm in aqueous buffer and taking the corresponding value in TFE to be 100% helix. Open circles are at 50 μM and closed circles are at 500 μM. All measurements were performed in 25 mM potassium phosphate and 100 mM NaCl at pH 7.0 and 5°C. (b) Dissociation constant at 20°C versus number of SKIP residues. Dissociation constants were determined by simultaneously fitting data sets obtained at 45 000 and 55 000 rpm to a monomer–dimer model. The resulting association constant in units of absorbance at 280 nm was converted to a dissociation constant in units of mol/liter using an extinction coefficient of 1490 $M^{-1}cm^{-1}$ (Pace *et al.*, 1995). Samples were all in 25 mM potassium phosphate and 100 mM NaCl at pH 7.0 and 20°C.

TABLE II

Peptide skip No.	T_M determined by CD		Measured mass, M_w	M_w/monomer mass	K_d (mM)
	50 μM	500 μM			
1	<10	<10	3506 (3255, 3762)	1.09 (1.01, 1.17)	12.4 (8.8, 17.8)
2	<10	<10	3625 (3364, 3890)	1.10 (1.02, 1.19)	10.4 (7.4, 14.7)
3	<10	<10	3978 (3665, 4298)	1.19 (1.09, 1.28)	5.3 (3.9, 7.3)
4	12	17	5465 (4981, 5949)	1.60 (1.45, 1.74)	0.36 (0.26, 0.50)
5	12	13	4977 (4553, 5401)	1.42 (1.30, 1.55)	1.1 (0.83, 1.5)
6	12	12	4644 (4278, 5015)	1.30 (1.20, 1.41)	2.4 (1.8, 3.2)

quences. Furthermore, dissociation constants (from sedimentation equilibrium) for SKIPs 1, 2, 3, 5, and 6 were measurable and were in the range 1–12 mM (Table II). However, SKIP4, which effectively had a 7-11-7 repeat, was stabler with a dissociation constant of 360 μM, which was in agreement with the CD thermal stability data that gave a T_M of 17°C at 500 μM. This result is consistent with and supports the prevalence of hendecad inserts in natural coiled coils noted in the Introduction.

The fact that SKIP3, which had a potential 7-10-7 repeat, did not show a greater coiled-coil-forming propensity than SKIPs 1, 2, 5, and 6 appears somewhat contradictory to the idea that 7-10-7-based sequences can form stable coiled coils (Hicks, 2000). However, we have noted previously that peptides based on a naturally occurring 7-10-7 repeat from the median body protein of *G. lamblia* are less stable than peptides based on the sequence of another protein from the same organism that contains 7-11-7 repeats (Hicks, 2000). A possible feature of the natural 7-10-7-based sequence that may rescue stability is its length; the median body protein from *G. lamblia* contains twenty-one 7-10-7 repeats. Thus, SKIP3 may simply be too short to exhibit complete coiled-coil formation at the concentrations accessible in this study.

It is, of course, possible that the inserted residues of noncanonical repeats in coiled coils loop out of the structures as seen in fibritin (PDB code 1AA0 (Efimov *et al.*, 1994)). We consider that this is unlikely for our peptides, although without high-resolution structures it is impossible to rule out this possibility completely. It is also possible that noncanonical repeats are quirks "locked in" during evolution. Again, this seems unlikely given repeats of the type reported for the *G. lamblia* sequences (see Hicks *et al.*, 1997) and that observed in tetrabrachion (Peters *et al.*, 1996; Stetefeld *et al.*, 2000). Nonetheless, it is interesting to speculate on the possible functional roles of nonheptad inserts in coiled coils. Continuing the theme above, one might imagine that long coiled coils comprising many residues and based only on canonical heptad repeats might be very stable. This may present problems if the structure that the coiled coil is part of needs to be dynamic, for example, the cytoskeleton or the spindle pole body. The inclusion of 11-residue (and possibly 10-residue) units in such sequences may provide a mechanism by which coiled-coil stability is modified. The stability of canonical dimeric coiled coils can also be modified by substitution of polar residues at *a* positions. However, this can also change the oligomerisation order (Gonzalez *et al.*, 1996; Harbury *et al.*, 1993; Woolfson and Alber, 1995). The use of noncanonical sequence repeats may be an alternative that, in combination with polar core substitutions, can modify stability whilst maintaining oligomer order. Another possible benefit of using noncanonical inserts may be to help maintain in-register alignment of α-helices in long coiled-coil structures. For instance, a long heptad-only coiled coil could be out of alignment by 7 residues and only lose two core positions (i.e. one heptad worth), whereas the "best" out-of-register alignment for a long 7-11-7-based sequence, for example, would result in the loss of seven core positions (one 7-11-7 motif worth). This mechanism may be particularly valuable in proteins that rely on a coiled-coil domain to place other domains in the correct alignment so that they may perform their function.

We thank the MRC and the BBSRC of the UK for financial support.

REFERENCES

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res.* **28,** 235–242.

Brown, J. H., Cohen, C., and Parry, D. A. D. (1996) Heptad breaks in alpha-helical coiled coils: Stutters and stammers. *Proteins Struct. Function Genet.* **26,** 134–145.

Bullough, P. A., Hughson, F. M., Skehel, J. J., and Wiley, D. C. (1994) Structure of influenza hemagglutinin at the pH of membrane-fusion. *Nature* **371,** 37–43.

Chakrabartty, A., Kortemme, T., Padmanabhan, S., and Baldwin, R. L. (1993) Aromatic side-chain contribution of far-ultraviolet circular-dichroism of helical peptides and its effects on measurement of helix propensities. *Biochemistry* **32,** 5560–5565.

Crick, F. H. C. (1953) The packing of alpha-helices: Simple coiled-coils. *Acta. Crystallogr.* **6,** 689–697.

Efimov, V. P., Nepluev, I. V., Sobolev, B. N., Zurabishvili, T. G., Schulthess, T., Lustig, A., Engel, J., Haener, M., Aebi, U., Venyaminov, S. Y., Potekhin, S. A., and Mesyanzhinov, V. V. (1994) Fibritin encoded by bacteriophage-T4 gene wac has a parallel triple-stranded alpha-helical coiled-coil structure. *J. Mol. Biol.* **242,** 470–486.

Gill, S. C., and Vonhippel, P. H. (1989) Calculation of protein extinction coefficients from amino-acid sequence data. *Anal. Biochem.* **182,** 319–326.

Gonzalez, L. J., Woolfson, D. N., and Alber, T. (1996) Buried polar residues and structural specificity in the Gen4 leucine-zipper. *Nat. Struct. Biol.* **3,** 1011–1018.

Harbury, P. B., Plecs, J. J., Tidor, B., Alber, T., and Kim, P. S. (1998) High-resolution protein design with backbone freedom. *Science* **282,** 1462–1467.

Harbury, P. B., Zhang, T., Kim, P. S., and Alber, T. (1993) A switch between 2-stranded, 3-stranded and 4-stranded coiled coils in Gen4 leucine-zipper mutants. *Science* **262,** 1401–1407.

Harrison, C. J., HayerHartl, M., DiLiberto, M., Hartl, F. U., and Kuriyan, J. (1997) Crystal structure of the nucleotide exchange factor GrpE bound to the ATPase domain of the molecular chaperone DnaK. *Science* **276,** 431–435.

Hicks, M. R. (2000) Coiled-Coil Assembly by Proteins and Peptides With Unusual Sequence Motifs, D. Phil. Thesis, University of Sussex, United Kingdom.

Hicks, M. R., Holberton, D. V., Kowalczyk, C., and Woolfson, D. N. (1997) Coiled-coil assembly by peptides with non-heptad sequence motifs. *Folding Design* **2,** 149–158.

Holberton, D., Baker, D. A., and Marshall, J. (1988) Segmented alpha-helical coiled-coil structure of the protein giardin from the *Giardia* cytoskeleton. *J. Mol. Biol.* **204,** 789–795.

Janin, J. (1979) Surface and inside volumes in globular proteins. *Nature* **277,** 491–492.

Jasanoff, A., and Fersht, A. R. (1994) Quantitative-determination of helical propensities From trifluoroethanol titration curves. *Biochemistry* **33,** 2129–2135.

Johnson, M. L., Correia, J. J., Yphantis, D. A., and Halvorson, H. R. (1981) Analysis of data from the analytical ultra-centrifuge by non-linear least-squares techniques. *Biophys. J.* **36,** 575–588.

Kallenbach, N. R., Pingchiang, L., and Hongxing, Z. (1996) CD spectroscopy and the helix-coil transition in peptides and polypeptides. *In* Fasman, G. D. (Ed.), Circular Dichroism and the Conformational Analysis of Biomolecules, pp. 201–259, Plenum, New York.

Kreis, T., and Vale, R. (1999) Guidebook to the Cytoskeletal and Motor Proteins, 2-nd ed., Oxford Univ. Press, Oxford.

Laue, T. M., Shah, B. D., Ridgeway, T. M., and Pelletier, S. L. (1992) Computer-aided interpretation of analytical sedimentation data for proteins. *In* Harding, S., and Rowe, A. (Eds.), Analytical Ultracentrifugation in Biochemistry and Polymer Science, pp. 90–125, Royal Society of Chemistry, London.

Lechtreck, K. F. (1998) Analysis of striated fiber formation by recombinant SF-assemblin *in vitro*. *J. Mol. Biol.* **279,** 423–438.

Lechtreck, K. F., Frins, S., Bilski, J., Teltenkotter, A., Weber, K., and Melkonian, M. (1996) The cruciated microtubule-associated fibers of the green alga *Dunaliella bioculata* consist of a 31 kDa SF-assemblin. *J. Cell Sci.* **109,** 827–835.

Lechtreck, K. F., and Melkonain, M. (1995) Sf-assemblin, the structural protein of the striated microtubule-associated fibers of algal flagellar roots forms a segmented coiled coil. *Mol. Biol. Cell* **6,** 235–235.

Lumb, K. J., and Kim, P. S. (1995) A buried polar interaction imparts structural uniqueness in a designed heterodimeric coiled-coil. *Biochemistry* **34,** 8642–8648.

Lupas, A. (1996) Prediction and analysis of coiled-coil structures. *Methods Enzymol.* **266,** 513–525.

Marshall, J., and Holberton, D. V. (1995) *Giardia* gene predicts a 183 kDa nucleotide-binding head-stalk protein. *J. Cell Sci.* **108,** 2683–2692.

Marshall, J., and Holberton, D. V. (1993) Sequence and structure of a new coiled-coil protein from a microtubule bundle in *Giardia*. *J. Mol. Biol.* **231,** 521–530.

O'Shea, E. K., Klemm, J. D., Kim, P. S., and Alber, T. (1991) X-ray structure of the Gen4 leucine zipper, a 2-stranded, parallel coiled coil. *Science* **254,** 539–544.

Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995) How to measure and predict the molar absorption-coefficient of a protein. *Protein Sci.* **4,** 2411–2423.

Pandya, M. J., Spooner, G. M., Sunde, M., Thorpe, J. R., Rodger, A., and Woolfson, D. N. (2000) Sticky-end assembly of a designed peptide fibre provides insight into protein fibrillogenesis. *Biochemistry* **39,** 8728–8734.

Peters, J., Baumeister, W., and Lupas, A. (1996) Hyperthermostable surface layer protein tetrabrachion from the archaebacterium *Staphylothermus marinus*: Evidence for the presence of a right-handed coiled coil derived from the primary structure. *J. Mol. Biol.* **257,** 1031–1041.

Skehel, J. J., and Wiley, D. C. (2000) Receptor binding and membrane fusion in virus entry: The influenza hemagglutinin. *Annu. Rev. Biochem.* **69,** 531–569.

Stetefeld, J., Jenny, M., Schulthess, T., Landwehr, R., Engel, J., and Kammerer, R. A. (2000) Crystal structure of a naturally occurring parallel right-handed coiled coil tetramer. *Nat. Struct. Biol.* **7,** 772–776.

Tarbouriech, N., Curran, J., Ruigrok, R. W. H., and Burmeister, W. P. (2000) Tetrameric coiled coil domain of Sandal virus phosphoprotein. *Nat. Struct. Biol.* **7,** 777–781.

Walshaw, J., and Woolfson, D. N. (2001) SOCKET: A program for identifying and analysing coiled-coil motifs within protein structures. *J. Mol. Biol.* **307,** 1427–1450.

Weber, K., Geisler, N., Plessmann, U., Bremerich, A., Lechtreck, K. F., and Melkonian, M. (1993) Sf-Assemblin, the structural protein of the 2-nm filaments from striated microtubule-associated fibers of algal flagellar roots, forms a segmented coiled-coil. *J. Cell Biol.* **121,** 837–845.

Weissenhorn, W., Carfi, A., Lee, K. H., Skehel, J. J., and Wiley, D. C. (1998) Crystal structure of the Ebola virus membrane fusion subunit, GP2, from the envelope glycoprotein ectodomain. *Mol. Cell* **2,** 605–616.

Wigge, P. A., Jensen, O. N., Holmes, S., Soues, S., Mann, M., and Kilmartin, J. V. (1998) Analysis of the *Saccharomyces* spindle pole by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry. *J. Cell Biol.* **141,** 967–977.

Woody, R. W., and Dunker, A. K. (1996) Aromatic and cystine side-chain circular dichroism in proteins. *In* Fasman, G. D. (Ed.), Circular Dichroism and the Conformational Analysis of Biomolecules, pp. 109–157, Plenum, New York.

Woolfson, D. N., and Alber, T. (1995) Predicting oligomerization states of coiled coils. *Protein Sci.* **4,** 1596–1607.

Zhao, X., Singh, M., Malashkevich, V. N., and Kim, P. S. (2000) Structural characterization of the human respiratory syncytial virus fusion protein core. *Proc. Natl. Acad. Sci. USA* **97,** 14172–14177.