Rational design of a three-heptad coiled-coil protein and comparison by molecular dynamics simulation with the GCN4 coiled coil: Presence of interior three-center hydrogen bonds

JAMES E. ROZZELLE, JR.,¹ ALEXANDER TROPSHA,² AND BRUCE W. ERICKSON¹

¹ Department of Chemistry and ² Laboratory for Molecular Modeling, School of Pharmacy,
The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

(RECEIVED September 3, 1993; ACCEPTED December 1, 1993)

Abstract

α -Helical coiled coils have a 7-residue repeating pattern (**abcdefg**) where **a** and **d** are usually hydrophobic. We have designed a 2-stranded 44-residue coiled-coil protein (P44) consisting of 2 22-residue α -helices linked by 2 terminal disulfide groups to test whether the disulfide bridges could stabilize a 3-heptad coiled coil. P44 should be stabilized by intrahelical hydrogen bonds, interhelical disulfide and salt bridges, and interior hydrophobic interactions. A computer model of P44 was built and its stability was studied by molecular dynamics simulation with explicit water. This doubly crosslinked 3-heptad coiled coil did not unfold during a 300-ps simulation. But reduced P44 with 4 thiol groups did unfold. For comparison, the 62-residue crystal structure of the 4-heptad coiled coil of transcription activator GCN4 did not unfold during a 300-ps simulation. Thus P44 may be a stable folded protein in aqueous solution. These simulations revealed the presence of 2 local hydrogen bond networks involving intrahelical 3-center hydrogen bonds in the hydrophobic interior of the coiled coils of GCN4 and P44. The NH hydrogen at **d** makes a 3-center hydrogen bond whose major component is to the $i - 4$ C=O oxygen at **g** and minor component is to the solvent-inaccessible $i - 3$ C=O oxygen at **a**. Likewise, the NH hydrogen at **g** makes a 3-center hydrogen bond with the $i - 4$ C=O oxygen at **c** and the buried $i - 3$ C=O oxygen at **d**.

Keywords: coiled coils; GCN4; molecular dynamics; protein folding; rational protein design; 3-center hydrogen bonds

The amino acid sequence of a protein determines its folded 3-dimensional structure (Anfinsen, 1973). The protein-folding problem (predicting the folded protein structure that a given amino acid sequence will assume) remains unsolved despite intensive study. The inverse protein-folding problem (predicting an amino acid sequence that will assume a given folded protein structure), however, can be explored by rational protein design (Yue & Dill, 1991). The 2-stranded α -helical coiled coil is the simplest example of a folded protein because it contains only one type of secondary structure, the α -helix. Given that it has defined secondary and tertiary structure, the coiled coil is interesting both theoretically and experimentally because it can be used to study many determinants of protein stability.

Francis Crick (1953) proposed that a coiled coil is stabilized by packing of hydrophobic side chains (knobs) of 1 α -helix into

spaces (holes) between the side chains of the other α -helix. In this model, 2 right-handed α -helices cross at a dihedral angle of about 20° and twist around one another to form a loosely wound 2-stranded left-handed supercoil. Two-dimensional (2D) NMR studies showed that a synthetic coiled-coil peptide based on the transcription factor GCN4 consists of 2 α -helices aligned in parallel and in register (Oas et al., 1990). Crick's knobs-in-holes folding scheme was confirmed by the crystal structures of a GCN4 coiled-coil peptide alone (O'Shea et al., 1991) and a related peptide complexed with DNA (Ellenberger et al., 1992).

The amino acid sequence of a coiled-coil α -helix is characterized by a repeating 7-residue pattern of hydrophobic and hydrophilic residues. The positions of this heptad are labeled **abcdefg**. Symmetry-related residues are denoted by numbering the residues of chain A from 1 to n and those of chain B from 101 to 100 + n and by lettering each heptad of chain A as **abcdefg** and chain B as **a'b'c'd'e'f'g'**. Residues at **a** and **d** lie on the same side of the helix and are generally hydrophobic. In the interior of the coiled coil, the inner residues at **a** and **d** make side-to-side con-

Reprint requests to: Bruce W. Erickson, Department of Chemistry, CB 3290, University of North Carolina, Chapel Hill, North Carolina 27599-3290.

tacts with the inner residues at **a'** and **d'**, respectively, of the other helix (O'Shea et al., 1991). The flanking residues at **e** and **g** are sometimes oppositely charged and can form interhelical salt bridges (McLachlan & Stewart, 1975; Cohen & Parry, 1990). Their side-chain methylene groups also make intrahelical hydrophobic contacts. The side chain of **e** packs against the preceding **d'** of the adjacent chain and **g** against the following **a'** (see Fig. 6 of O'Shea et al., 1991; Kinemage 6). These interactions make important contributions to the stability of the coiled coil.

Two-stranded α -helical coiled coils are found in many natural proteins (Cohen & Parry, 1986, 1990) and are predicted to occur in more than 200 proteins (Lupas et al., 1991). Long coiled coils, which form extended ropes between proteins or protein subunits of tropomyosin and other classic coiled-coil proteins, favor leucine residues at both **a** and **d** (Cohen & Parry, 1986, 1990). Short coiled coils, which mediate the dimerization of the transcription activator GCN4 and other members of the bZIP class of DNA-binding proteins (Landschulz et al., 1988), favor the β -branched hydrophobic residues valine and isoleucine at **a** but have leucine residues almost exclusively at **d**. Leucines at **d** have been shown to stabilize the coiled coil more than any other hydrophobic residue (Hu et al., 1990).

Coiled coils can be used to study most types of noncovalent interactions involved in maintaining the 3-dimensional structures of proteins. They have long been attractive target structures for rational protein design (Hodges et al., 1981, 1990; Talbot & Hodges, 1982; Cohen & Parry, 1990; O'Neil & Degrad, 1990; Zhou et al., 1992b; Zhu et al., 1993). Uncrosslinked coiled coils require at least 4 heptads per helix to be stable in aqueous solution (Lau et al., 1984). Adding a single interhelical disulfide crosslink can stabilize a coiled coil (Hodges et al., 1988; Engel et al., 1991; Zhou et al., 1993a). If 1 interhelical disulfide bond can stabilize a 4-heptad coiled coil, then a second interchain disulfide bond might stabilize a 3-heptad coiled coil. To test this hypothesis, we have designed the coiled-coil protein P44. This 44-residue protein consists of 2 identical 22-residue 3-heptad α -helices kept aligned in parallel and in register by 1 disulfide bond between their N-terminal cysteine residues and another between their C-terminal cysteine residues. P44 should be stabilized by these interhelical disulfide bridges, by intrahelical $i \rightarrow i - 4$ hydrogen bonds, and by the hydrophobic and electrostatic interactions characteristic of coiled coils.

We report here the results of 3 molecular dynamics (MD) simulations with explicit water molecules: (1) the designed coiled-coil protein P44, (2) reduced P44 in which the 2 disulfide bonds (S-S) were replaced by 4 thiol (S-H) groups, and (3) the crystal structure of the GCN4 coiled-coil peptide. They suggest that the doubly disulfide-bridged coiled coil of P44 may be stable in aqueous solution. These MD simulations also allowed the study of dynamic structural features, such as transient hydrogen bonds, that cannot be directly observed by NMR or X-ray techniques.

Results and discussion

Comparison of the computer model of P44 with the crystal structure of the GCN4 coiled coil

The α -helical backbones of P44 and the GCN4 coiled coil were in excellent agreement. When the backbone atoms of the energy-minimized P44 structure were superimposed on the backbone

atoms of residues 2–23 of the GCN4 crystal structure, the RMS deviation between corresponding α carbons was only 0.55 Å (Kinemage 2). Aligning P44 similarly with residues 9–31 of GCN4 gave an RMS deviation of 0.60 Å. In both cases the interhelical distance for P44 was slightly less than that for the GCN4 coiled coil. The interhelical distances between the α carbons of the terminal leucines of P44 (4d–104d = 6.02 Å, 18d–118d = 5.96 Å) were less than those of the terminal leucines of GCN4 (5d–105d = 6.34 Å, 26–126d = 6.20 Å). But the C $^{\alpha}$ –C $^{\alpha}$ interhelical distance of the central leucine of P44 (11d–111d = 6.17 Å) was practically the same as those of the central leucines of GCN4 (12d–112d = 6.12 Å, 19d–119d = 6.20 Å). The shorter interhelical distances seen at the N- and C-termini of P44 are likely due to the interhelical disulfide bonds.

During construction of the P44 model, we assumed that the side-chain conformations of the leucines, valines, and isoleucines at **a** and **d** would be the same as those most often found in folded proteins (Richardson & Richardson, 1989). In energy-minimized P44, the side-chain conformations of the leucines at **d** ($\chi_1 \sim -60^\circ$, $\chi_2 \sim 180^\circ$) and Val 15a and Val 115a ($\chi_1 \sim 180^\circ$) were the same as those later reported for the crystal structure of the GCN4 coiled coil (O'Shea et al., 1991). The conformation of Ile 8a and Ile 108a ($\chi_1 \sim 180^\circ$) has no counterpart in the GCN4 coiled coil.

During building of the P44 model, the 6 Glu/Lys pairs at **e** of 1 heptad and **g'** of the adjacent heptad were initially positioned with their charged termini close together to form 6 e-to-g' (i to $i' + 2$) salt bridges: Glu 5e/Lys 107g, Glu 105e/Lys 7g, Glu 12e/Lys 114g, Glu 112e/Lys 14g, Glu 19e/Lys 121g, Glu 119e/Lys 21g (Fig. 1; Kinemage 1). But the crystal structure of the GCN4 coiled coil later showed that e-to-g' salt bridges do not form. Rather, the **g** residue bridged to the **e'** residue of the following heptad to form a g-to-following-e' (i to $i' + 5$) salt bridge. The GCN4 crystal structure revealed that the interaction of **e** with **g'** of the adjacent heptad was blocked by the bulky side chain of the leucine at **d**, as predicted by McLachlan and Stewart (1975). Therefore, instead of 6 e-to-g' interhelical salt bridges, P44 should only have 4 g-to-following-e' interhelical salt bridges (Lys 7g/Glu 112e, Lys 107g/Glu 12e, Lys 14g/Glu 119e, Lys 114g/Glu 19e), which do not involve the first glutamates (Glu 5e, Glu 105e) or the last lysines (Lys 21g, Lys 121g). In fact, during the initial stages of the MD simulation, the e-to-g' salt bridges initially imposed on the P44 model were replaced by g-to-following-e' salt bridges (see below).

The molecular mechanics procedure that we used to build our model of P44 from the polyaniline template of Parry and Suzuki (1969) is a general and relatively unbiased method for generating a high-resolution coiled-coil structure. This procedure was useful for predicting other coiled-coil structures. For example, when we used this method to model the GCN4 coiled coil, the RMS deviation between the α carbons of our computer model and those of the crystal structure was only 0.58 Å. This method is a simple alternative to the more computationally intensive ab initio method of Nilges and Brünger (1991) for predicting coiled-coil structures.

Comparison of P44, reduced P44, and the GCN4 coiled coil by MD simulation

To test our hypothesis that 2 interchain disulfide bridges could stabilize a 3-heptad coiled coil, we studied the stability of P44

Fig. 1. Computer model of P44. Side views (A) without and (B) with ribbons representing the α -helical backbones. Inner leucines, valines, and isoleucines interact side-to-side between the helices and the methylene groups of the flanking lysines and glutamates pack against them to form a continuous hydrophobic core.

by MD simulation in the presence of explicit water molecules. The energy-minimized model of P44 was equilibrated for 60 ps with a distance constraint imposed on every third $i \rightarrow i - 4$ hydrogen bond and then simulated for 300 ps of unconstrained MD. Figure 2A shows the final MD structure of P44 superimposed on the starting structure. P44 remained folded during 300

ps of unconstrained MD in water, but some changes in the starting P44 structure occurred: the backbone geometry of the first 10 N-terminal residues of chain A was changed (Fig. 2A; Kinemage 3) and the side-chain orientation of the glutamates and lysines changed (see below). Kinemage 3 contains 7 snapshots from this MD simulation. Because P44 did not undergo major unfolding during this simulation, it may be a stable folded protein in water.

The contribution of the 2 disulfide bridges to the stability of P44 was studied by performing a similar MD simulation on the model of reduced P44, which lacked both disulfide bridges. Reduced P44 was simulated for 30 ps with constraints to maintain α -helicity followed by 300 ps of unconstrained MD. Figure 2B shows the final MD structure of reduced P44 superimposed on the starting structure. Reduced P44, which lacks the 2 terminal disulfide bridges of P44, unfolded during MD simulation. Both the coiling and the supercoiling of the α -helices were substantially disrupted. Kinemage 4 contains 7 snapshots from this MD simulation showing the course of unfolding. Both α -helices unwound from the N-terminus. The interhelical hydrophobic interactions between the inner pairs Cys 1a/Cys 101a, Leu 4d/Leu 104d, and Ile 8a/Ile 108a were disrupted within 50 ps. By the end of the simulation the coiled-coil structure was completely disrupted. The first 10 residues of each helix lost all interhelical contacts. Only Val 15a/Val 115a and Leu 18d/Leu 118d remained in contact and only 18 of the initial 36 $i \rightarrow i - 4$ backbone hydrogen bonds remained intact. This result is consistent with the experimental instability of a 3-heptad coiled coil that lacks an interhelical covalent crosslink (Lau et al., 1984). Comparison of the MD simulations of P44 and reduced P44 clearly shows that the disulfide bridges of P44 make a major contribution to the sta-

Fig. 2. Partial structures of (A) the designed coiled-coil protein P44, (B) reduced P44 without the disulfide bridges, and (C) the GCN4 coiled coil. Only the backbone atoms and the side chains at a and d are shown. The starting and final structures from each MD simulation are superimposed by minimizing the RMS deviation of backbone atoms of residues Leu 111d to Leu 118d in A and B and of residues Leu 12d to Val 23a in C. The helices of P44 remained supercoiled but the helices of reduced P44 did not. The GCN4 coiled coil unfolded at the N-terminus (top) but the middle remained folded.

bility of its secondary and tertiary structure. They keep the 2 chains parallel and in register and stabilize the packing of the interior **a** and **d** residues (Zhou et al., 1992, 1993a).

The relevance of these 2 MD simulations to the stability of P44 in water was examined by performing a similar MD simulation of the GCN4 coiled coil, whose stability in water has been shown by 2D NMR (Oas et al., 1990). The 62-residue crystal structure of the GCN4 coiled coil was simulated by MD for 30 ps with α -helical constraints and then for 300 ps without constraints. Figure 2C shows the final MD structure of the GCN4 coiled coil superimposed on the crystal coordinates. As expected, most of the GCN4 coiled coil stayed folded during this simulation, although there was some unfolding at the helix termini. The major changes in the starting backbone geometry occurred at the N-termini of the 2 helices. One helix lost 7 of its initial 28 $i \rightarrow i - 4$ backbone hydrogen bonds and the other lost 6. Flexibility of the helix termini has also been seen in short, in vacuo MD simulations of experimentally stable model coiled coils (Zhou et al., 1992). The α -helical structure was more conserved at the C-terminus than at the N-terminus. After the MD simulation the inner pairs Met 2a/Met 102a, Leu 5d/Leu 105d, and Val 9a/Val 109a were no longer within van der Waals distance. But the following 7 hydrophobic pairs at **a** and **d** (Leu 12d/Leu 112d through Val 31a/Val 131a) remained in contact.

Further analysis of these MD simulations provided several insights into the structural features that stabilize 2-stranded α -helical coiled coils. Two features discussed below are intrahelical 3-center hydrogen bonds of the NH hydrogens at **d** and **g** and the interplay of hydrophobic and electrostatic interactions of the ionic residues at **e** and **g**.

Internal α -helical 3-center hydrogen bonds

Pauling and Corey (1953) proposed that shorter $i \rightarrow i - 4$ α -helical hydrogen bonds between inner residues at the helix-helix interface would allow the 2 α -helices to supercoil around one another. Indeed, in globular proteins the C=O groups on the hydrophobic side of an α -helix make shorter hydrogen bonds than those on the hydrophilic, solvent-exposed side (Blundell et al., 1983). A 2D NMR study of a GCN4 coiled coil found that the relative exchange rates of the amide protons of the inner residues at **a**, **d**, and **e** were slower than those of the outer residues at **c** and **f** (Goodman & Kim, 1991). Also, the chemical shifts of the amide protons of the inner residues at **a**, **d**, and **e** were shifted more downfield (relative to these residues in unstructured peptides) than those of the outer residues at **b**, **c**, and **f**, which is consistent with shorter inner hydrogen bonds and curved helices. A 2D NMR study of short amphipathic α -helices revealed that the chemical shifts of the amide protons on the hydrophobic face of the helix were similarly shifted downfield and had slower deuterium-exchange rates (Zhou et al., 1992).

Figure 3A illustrates that during our MD simulation of the GCN4 coiled coil the average $i \rightarrow i - 4$ N-to-O distances from **a** to **d** and from **d** to **g** were shorter than for the adjacent residues (**a** shorter than **g** or **b**, and **d** shorter than **c** or **e**). Interestingly, Figure 3B reveals that the average $i \rightarrow i - 3$ N-to-O distances from **d** to **a** and from **g** to **d** were 3.2 Å or less, which are within the distance of 3.4 Å thought to be stabilizing for a hydrogen bond (Burley & Petsko, 1988). The geometry of the α -helix is compatible with the simultaneous presence of both $i \rightarrow i - 4$ and $i \rightarrow i - 3$ hydrogen bonds (Stickle et al., 1992).

Fig. 3. Intrahelical N-to-O distances for chain B of the GCN4 coiled coil. **(A)** Average MD values of the $i \rightarrow i - 4$ N-to-O hydrogen bond distances. The distances from the nitrogen atoms at 116a, 119d, 123a, 126d, and 130a were shorter than those at adjacent positions. **(B)** Average MD values of the $i \rightarrow i - 3$ N-to-O distances. The distances from the nitrogen atoms at 112d, 115g, 119d, 122g, 126d, and 129g were shorter than those at adjacent positions.

These N-to-O distances are consistent with 2 local hydrogen bond networks, (1) one between the C=O groups at **g** and **a** and the NH groups at **d** and **e** and (2) the other between the C=O groups at **c** and **d** and the NH groups at **g** and **a**. These local networks of backbone hydrogen bonds, shown in Figure 4 for the last half of chain B of the GCN4 coiled coil, are found on the inner, hydrophobic side of the α -helices. They have 2 distinctive features. First, the C=O oxygens of the inner residues at **a** and **d** form bidentate hydrogen bonds. For example, the C=O oxygen at **a** forms both a normal α -helical hydrogen bond ($i \rightarrow i - 4$ closing a 13-atom ring) with the NH hydrogen at **e** and another hydrogen bond ($i \rightarrow i - 3$ closing a 10-atom ring) with the NH hydrogen at **d**. Second, the NH hydrogens at **d** and **g** each form a 3-center (bifurcated) hydrogen bond with 2 C=O oxygens of the preceding helical turn. For instance, the NH hydrogen at **d** makes a 3-center hydrogen bond whose major component is an $i \rightarrow i - 4$ interaction with the C=O oxygen at **g** and whose minor component is an $i \rightarrow i - 3$ interaction with the following C=O oxygen at **a**. Likewise, the NH hydrogen at **g** makes a 3-center hydrogen bond with the C=O oxygen at **c** and the buried C=O oxygen at **d**.

In addition to an appropriate N-to-O distance, a hydrogen bond must also have an appropriate N-H...O bond angle. The

Fig. 4. Backbone hydrogen bonding of residues Lys 115g–Glu 131b of the GCN4 coiled coil. The C=O oxygens of the inner residues **a** and **d** are bidentate. The NH hydrogen at **d** forms a 3-center hydrogen bond with 2 C=O oxygens at **g** and **a** of the preceding helical turn. Likewise, the NH hydrogen at **g** forms a 3-center hydrogen bond with 2 C=O oxygens at **c** and **d**. Exclusion of water from the hydrophobic interior of the coiled coil (between the diagonal lines) enhances the electrostatic interaction of the $i \rightarrow i - 3$ NH hydrogens at **d** and **g** with the inner C=O oxygens at **a** and **d**, respectively.

energy of an $N-H \cdots O=C$ hydrogen bond changes significantly as the $N-H \cdots O$ bond angle is changed but only modestly as the $H \cdots O=C$ bond angle is changed (Hagler et al., 1974). The average $N-H \cdots O$ bond angle for the $i \rightarrow i - 4$ hydrogen bonds of α -helices in globular proteins is about 155° ; 90% lie between 140° and 180° (Baker & Hubbard, 1984). A hydrogen bond is stabilizing down to an $N-H \cdots O$ angle of 120° (Hagler et al., 1974). A study of over 300 3-center hydrogen bonds with a major $N-H \cdots O=C$ component and a minor $N-H \cdots X$ component found that the major $N-H \cdots O$ angle averaged 147° and the minor $N-H \cdots X$ angle averaged 113° (Taylor et al., 1984). A 3-center hydrogen bond in which the hydrogen atom is equally shared between 2 C=O oxygens is rarely seen.

The unequal interaction of the NH hydrogen of an inner **d** residue with the preceding C=O oxygens at **g** and **a** is reflected in the 2 $N-H \cdots O$ bond angles. Figure 5A shows the $N-H \cdots O$ angles for the major $i \rightarrow i - 4$ component and the minor $i \rightarrow i - 3$ component of the 3-center hydrogen bond involving the NH group of Leu 112d of the GCN4 coiled coil. During the MD simulation, as the $i \rightarrow i - 4$ $N-H \cdots O$ angle decreased and this component of the hydrogen bond became weaker, the $i \rightarrow i - 3$ $N-H \cdots O$ bond angle increased and this component became stronger. At several points the $i \rightarrow i - 3$ $N-H \cdots O$ angle even became greater than 120° as the $i \rightarrow i - 4$ $N-H \cdots O$ angle became less than 140° ; twice the $i \rightarrow i - 3$ $N-H \cdots O$ angle became greater than the $i \rightarrow i - 4$ $N-H \cdots O$ angle. In the GCN4 coiled coil, this correlated motion was generally seen for the NH hydrogens of residues at **d** and **g**, such as Leu 119d (Fig. 5C) and Glu 122g (Fig. 5D) but rarely seen for the NH hydrogens of the residues at **a**, **b**, **c**, **e**, or **f**, such as His 118c (Fig. 5B). This correlated motion of the NH hydrogens at **d** and **g** may be due to the inaccessibility to water of their $i - 3$ C=O oxygens at **a** and **d**, respectively. Similarly, the same correlated motion of the 2 $N-H \cdots O$ angles of the 3-center hydrogen bond was seen for the NH hydrogens at **d** and **g** in P44, such as Leu 111d (Fig. 6A), Lys 114g (Fig. 6B), and Leu 118d (Fig. 6D), but not for the NH hydrogens of other residues, such as Ser 117c (Fig. 6C).

MD simulation can provide insights into dynamic events that cannot be observed by NMR or X-ray techniques. The average N-to-O distances in Figure 3 and the $N-H \cdots O$ angles in Figures 5 and 6 provide strong computational evidence that these coiled-coil structures have a series of 3-center hydrogen bonds along the hydrophobic interface between the α -helices.

Why are these 2 local hydrogen bond networks seen on the hydrophobic interface of the helices? A solvent-accessible backbone C=O oxygen of an α -helix usually forms 1 hydrogen bond with an NH hydrogen and another to water (Blundell et al., 1983). But the backbone C=O oxygen at **a** or **d** cannot form the hydrogen bond to water because water is excluded from the interior of the coiled coil by the hydrophobic side chains of the inner residues at **a** and **d** and the methylene groups of the flanking residues at **e** and **g** of both chains (see Fig. 6 of O'Shea et al., 1991). As a result, the second lone pair of the C=O oxygen at **a** or **d** interacts with the NH hydrogen at **d** or **g**, respectively, as the minor component of a 3-center hydrogen bond. This electrostatic interaction is strengthened by the low dielectric constant of the hydrophobic interior of the coiled coil. The resulting N-to-O distances are shorter (Fig. 3) for the interior $i \rightarrow i - 4$ hydrogen bonds between the NH hydrogens at **d** and **e** and the C=O oxygens at **g** and **a** and those between the NH hydrogens

Fig. 5. N—H···O hydrogen bond angles for the $i \rightarrow i-4$ and $i \rightarrow i-3$ components of the 3-center hydrogen bonds of the NH protons of (A) Leu 112d, (B) His 118c (C), Leu 119d, and (D) Glu 122g of the GCN4 coiled coil. At several times the inner NH hydrogens of Leu 112d, Leu 119d, and Glu 122g moved so that the $i \rightarrow i-4$ N—H···O angle decreased as the $i \rightarrow i-3$ N—H···O angle increased. The outer NH hydrogen of His 118c did not show this correlation.

at **g** and **a** and the C=O oxygens at **c** and **d**. This explanation is consistent with the data of Goodman and Kim (1991) showing that the amide protons at positions **c** and **f** exchanged faster than those at **a**, **d**, **e**, and **g**.

The 2 local hydrogen bond networks described above for α -helical coiled coils are the first examples of 3-center hydrogen bonds that occur regularly in the middle of an α -helix. A related hydrogen bond network is the α_{C1} distortion, in which the C-terminus of an α -helix undergoes a transition from $i \rightarrow i-4$ to $i \rightarrow i-3$ hydrogen bonding (Nemethy et al., 1967; Baker & Hubbard, 1984). One or two 3-center hydrogen bonds are often seen at the α_{C1} transition point (see Fig. 9b of Baker & Hubbard, 1984).

Thus the slight left-handed supercoiling of the 2 right-handed α -helices of a coiled coil (Pauling & Corey, 1953) is associated with both shorter interior hydrogen bonds (Oas et al., 1990; O'Shea et al., 1991) and 3-center hydrogen bonds of the NH hydrogens at **d** or **g** with the C=O oxygens at **g** and **a** or at **c** and **d**, respectively. Because a slight curving of the axis of an α -helix is estimated to require very little energy (Parry & Suzuki, 1969), these 3-center hydrogen bonds may make a small but significant contribution to the stability of α -helical coiled coils and possibly of globular proteins containing curved amphiphilic α -helices.

Bidentate hydrogen bonds at the N-terminus of P44

Bidentate hydrogen bonds, which can occur at the N-terminus of an α -helix (Presta & Rose, 1988), were also observed for the C=O oxygens of the N-terminal cysteines of P44. The disulfide bridge between Cys 1a and Cys 101a constrained Cys 1a of chain A so that its C=O oxygen was close enough to form both an $i \rightarrow i-4$ hydrogen bond with the NH hydrogen of Leu 5e and an $i \rightarrow i-3$ hydrogen bond with the NH hydrogen of Leu 4d, and similarly for Cys 101a of chain B. These bidentate hydrogen bonds were maintained throughout the MD simulation even though several backbone $i \rightarrow i-4$ hydrogen bonds after Glu 5e were broken in chain A of P44. Interestingly, a structural domain consisting of the first 8 N-terminal residues of each chain moved in a concerted fashion during the MD simulation (Kine-mage 3). This structural domain was stabilized (1) by the inter-helical disulfide bridge between Cys 1a and Cys 101a, (2) by hydrophobic interactions of the side chains of Leu 4d and Leu 104d, (3) by the intrahelical bidentate hydrogen bonds of Cys 1a and Cys 101a, and (4) by the intrahelical hydrogen bonds and salt bridges between the terminal NH_3^+ groups of Cys 1a and Cys 101a and the side-chain CO_2^- groups of Asp 2b and Asp 102b, respectively.

Fig. 6. N-H...O hydrogen bond angles for the $i \rightarrow i-4$ and $i \rightarrow i-3$ components of the 3-center hydrogen bonds of the NH protons of (A) Leu 111d, (B) Lys 114g, (C), Ser 117c, and (D) Leu 118d of P44. At several times the inner NH hydrogens of Leu 111d, Lys 114g, and Leu 118d moved so that the $i \rightarrow i-4$ N-H...O angle decreased as the $i \rightarrow i-3$ N-H...O angle increased. The outer NH hydrogen of Ser 117c did not show this correlation.

Interhelical salt bridges and the hydrophobic core of P44

The starting model of P44 had the 6 lysine and the 6 glutamate side chains positioned to form 6 e-to-g' (i to $i' + 2$) interhelical salt bridges (Fig. 1; Kinemage 3). But during the first 10 ps of the MD simulation, 8 of the 12 residues moved from their starting e-g' pairings to form 4 g-to-following-e' (i to $i' + 5$) interhelical salt bridges (Fig. 7; Kinemage 3). The side chains of Glu 5e, Glu 105e, Lys 21g, and Lys 121g became unpaired. These 4 new g-to-following-e' salt bridges (Lys 7g-Glu 112e, Lys 107g-Glu 12e, Lys 14g-Glu 119e, Lys 114g-Glu 19e) were located in the C-terminal two-thirds of P44. After about 240 ps, the Lys 7g-Glu 112e salt bridge was broken and a new salt bridge between Lys 7g and Asp 9b was formed.

The X-ray structure of the GCN4 peptide showed that the flanking lysine and glutamate residues at e and g stabilize the coiled coil by packing against the hydrophobic inner residues at a and d, so the methylene groups of the flanking residues at e and g are part of the hydrophobic core (O'Shea et al., 1991). The 4 g-to-following-e' salt bridges of P44 similarly stabilized the hydrophobic core located at the C-terminal two-thirds of P44. These 8 lysine and glutamate side chains buttressed the hydrophobic interactions of the 8 inner residues at a and d (Ile 8a/Ile 108a, Leu 11d/Leu 111d, Val 15a/Val 115a, Leu 18d/Leu 118d). But Leu 4d and Leu 104d lacked these stabilizing interactions,

so the N-terminal third of P44 was less stable than the C-terminal two-thirds.

The difference in stability of the N- and C-termini of the P44 coiled coil was manifest in several ways. (1) During the MD sim-

Fig. 7. Plots of the distance between the Lys NH_3^+ nitrogen (N^6) and the Glu CO_2^- carbon (C^5) versus time for the Lys/Glu salt bridges of P44. After about 240 ps the Lys 7g/Glu 112e salt bridge was broken due to the unfolding of the A chain near the N-terminus and the formation of a new salt bridge between Lys 7g and Asp 9b.

ulation of P44, 7 of the initial 18 α -helical hydrogen bonds at the N-terminus of chain A were broken. (2) Two new side-chain-side-chain salt bridges were formed, one between the NH_3^+ group of Lys 7g and the CO_2^- group of Asp 9b and another between the N-terminal NH_3^+ group of Cys 1a and the CO_2^- group of Glu 4e. Neither of these new salt bridges could have formed if this part of chain A had remained α -helical. (3) The N-terminal pair of cysteines was farther apart during the MD simulation than the C-terminal pair. Specifically, at the beginning of the MD simulation the C^α - C^α distance between the N-terminal cysteines increased until it was about 1.5–2 Å longer than that of the C-terminal cysteines (Fig. 8). Also, because the β carbons of the side chains in an α -helix are oriented toward the N-terminus, the disulfide-bridged side chains of the C-terminal cysteines were buried but those of the N-terminal cysteines were exposed to water (Fig. 1; Kinemage 3).

Roles of residues at e and g in the GCN4 coiled coil

Both intrahelical side-chain-side-chain hydrogen bonds (Lys 8g/Glu 11e, Arg 25c/Glu 22e) seen in the crystal structure of the GCN4 coiled-coil peptide (O'Shea et al., 1991) were quickly broken and did not reform during the MD simulation. Three interhelical salt bridges (Lys 27e/Glu 122g, Lys 127e/Glu 22g, Lys 115g/Glu 20e) were seen in the GCN4 crystal structure, but the Lys 15g-Glu 120e salt bridge was absent because of steric hindrance from the side chain of Asn 16a. During the MD simulation, however, all 4 of these Lys/Glu pairs behaved similarly. The distance between the NH_3^+ nitrogen (N^6) and the CO_2^- carbon (C^5) was 5–8 Å for all 4 Lys/Glu pairs (Fig. 9). Also, the methylene groups of Lys 15g and Glu 120e were packed against the next inner hydrophobic pair (Leu 19d/Leu 119d) in the same manner as those of Lys 115g and Glu 20e (Kinemage 6). All 4 Lys/Glu salt bridges evidently contribute to the stabilization of the hydrophobic core located at the C-terminal two-thirds of the GCN4 coiled coil. During the MD simulation, the structure of the GCN4 coiled coil, like that of P44, changed more at the N-terminus than at the C-terminus (Kinemage 5). Among the factors contributing to this result are (1) the C-terminal location of the Lys/Glu interactions discussed above, which stabilize the hydrophobic core, and (2) the linear side chains

Fig. 8. Plots of the distance between the cysteine α carbons (C^2) versus time for the terminal cysteines of P44.

Fig. 9. Plots of the distance between the Lys NH_3^+ nitrogen (N^6) and the Glu CO_2^- carbon (C^5) versus time for the interhelical Lys/Glu pairs of the GCN4 coiled-coil peptide.

of the N-terminal hydrophobic pair at a (Met 2a/Met 102a), which are less stabilizing than branched side chains.

The GCN4 coiled-coil peptide contains a leucine at the second e position (Leu 13e, Leu 113e) and another at the fourth g position (Leu 29g, Leu 129g). During the MD simulation, the interior residues at d and a interacted with the side chains of these leucines in the same fashion as they did with the methylene groups of a glutamate or lysine at e or g (Kinemage 6). For example, the side chain of Leu 13e packed against Leu 112d of the adjacent chain and Leu 29g packed against Val 130a.

Hydrogen bonding of the inner GCN4 residues Asn 16a and Asn 116a

Because polar residues are not uncommon at the inner positions a and d of a coiled coil (Landschulz et al., 1988), they may have important functions. Destabilization of the GCN4 coiled coil by the polar side chains of Asn 16a and Asn 116a could facilitate its reversible dimerization in vivo and/or modulate its affinity of binding to DNA by controlling the dimer concentration (O'Shea et al., 1991). These asparagine residues may also favor the un-staggered alignment of the 2 parallel α -helices because an asparagine side chain should pack better against another asparagine side chain than against a nonpolar side chain (O'Shea et al., 1991). In the crystal structure of the GCN4 coiled coil, the side chains of Asn 16a and Asn 116a are positioned asymmetrically (Fig. 10A). This general asymmetry was maintained during the MD simulation. Hydrogen bond (2) between a side-chain NH hydrogen of Asn 116a and the $\text{C}=\text{O}$ oxygen of Leu 12d was still present in the final MD structure (Fig. 10B). But hydrogen bond (3) between the side-chain NH hydrogen of Asn 16a and the side-chain oxygen of Asn 116a was lost after about 60 ps (Fig. 10C). A hydrogen bond between the polar side chain of an inner residue, such as asparagine, lysine, or arginine, and a backbone $\text{C}=\text{O}$ oxygen of the other helix might stabilize a coiled coil (Tropsha et al., 1991). Polar side chain backbone-hydrogen bonds are often seen in folded proteins (Baker & Hubbard, 1984). The dynamic stability of the interhelical side-chain backbone-hydrogen bond (2) between Asn 116a and Leu 12d supports the importance of such hydrogen bonds in protein folding and association.

Fig. 10. Different hydrogen bonding of the side chains of Asn 16a and Asn 116a in the GCN4 coiled coil during MD simulation. **A:** Starting asymmetric orientation of their side chains from the crystal structure (O'Shea et al., 1991). Hydrogen bond (1) is the $i \rightarrow i - 4$ α -helical bond between the NH hydrogen of Asn 16a and the C=O oxygen of Leu 12d, (2) is between the same C=O oxygen and an N⁴ hydrogen of Asn 116a, and (3) is between oxygen O⁴ of Asn 116a and an N⁴ hydrogen of Asn 16a. **B:** Final MD structure with the Asn/Asn hydrogen bond (3) broken and the Asn/Leu hydrogen bonds (1) and (2) intact. **C:** Plots of the N-to-O distance for hydrogen bonds (2) and (3) during the MD simulation.

Methods

Design of P44

Each α -helix of P44 contains 3 nonidentical heptads plus 1 more residue (Fig. 11). Cysteine is most often found at **a** (Cohen & Parry, 1990; Lupas et al., 1991), so cysteine residues were placed at both the terminal **a** positions. One disulfide bridge between Cys 1a and Cys 101a and another between Cys 22a and Cys 122a form the 2 terminal crosslinks (Fig. 12). Disulfide bonds between cysteines in terminal **a** positions are stabilizing (Zhou et al., 1993a). Because leucine is found almost exclusively at **d** in shorter coiled coils (Landschulz et al., 1988), leucines were placed in the 3 inner **d** positions. β -Branched hydrophobic residues often occur at **a** (Landschulz et al., 1988; Lupas et al., 1991), so a valine and an isoleucine were placed in the 2 nonterminal **a** positions. In recent studies using model coiled coils, leucine provided maximum stability at **d**, isoleucine at **a** provided more stabilization than leucine, and valine provided more stabilization at **a** than at **d** (Zhu et al., 1993). Glutamates and lysines were placed at the flanking positions **e** and **g** to form interhelical salt bridges (Cohen & Parry, 1990). Due to favorable interactions with the helix dipole, negatively charged glutamates are frequently found near the N-terminus of an α -helix and positively

Fig. 11. Amino acid sequences of the peptide chains of P44 and GCN4-p1. The inner residues at **a** and **d** are shown in boldface. P44 has 3 heptads per chain and a disulfide bridge at each end. The GCN4 coiled coil contains 4 heptads per chain and no disulfide bridges. Glu 32 and Arg 33 of GCN4-p1 were not visible in the crystal structure.

charged lysines found near the C-terminus (Richardson & Richardson, 1988). In order to provide favorable interactions with the helix dipoles of P44, the glutamates were placed at **e** and the lysines at **g**, so the first glutamate was Glu 5e near the N-terminus and the last lysine was Lys 21g near the C-terminus.

Several choices of amino acid residues were made to facilitate future experimental studies. Except for Tyr 13f, the outer amino acids at **b**, **c**, and **f** were mainly selected based on their hydrophilicity and propensity to form an α -helix. Although tyrosine is not highly favored in helices or on protein surfaces, it absorbs ultraviolet light and its phenolic hydroxyl group can form hydrogen bonds with water. For ease of synthesis the amino acids arginine, asparagine, glutamine, histidine, methionine, and tryptophan were not used. In order to insure good solubility in water, phenylalanine was excluded. In order to avoid the formation of loops or turns, glycine and proline were not used. Amino acids were selected for the remaining positions in order to make the amino acid sequence of P44 resemble those of naturally occurring proteins in 3 ways: relative frequency of occurrence at specific positions of the α -helix (Richardson & Richardson, 1989), relative frequencies of occurrence at specific positions of the coiled coil (Lupas et al., 1991), and a nonrepetitive sequence. The final design of P44 placed alanine and aspartate at **b**; alanine, serine, and threonine at **c**; and alanine, serine, and tyrosine at **f** (Fig. 11). P44 contains 11 types of amino acids. Their relative positions on an α -helical wheel are shown in Figure 12. Each α -helix contains 20 unique tripeptide segments, 19 unique dipeptide segments, and 3 Leu-Glu dipeptide segments.

Fig. 12. Helical-wheel representation of P44. View is from the N-termini looking down the α -helical axes. Heptad positions are labeled a-g and a'-g'. Residues of the first heptad in the first 2 helical turns are circled. The N-terminal disulfide bridge is shown and the Lys/Glu salt bridges are connected with dashed lines.

Three-dimensional modeling

Two research groups (Parry & Suzuki, 1969; McLachlan, 1978) have produced models of idealized coiled coils based on the mathematical description of a coiled coil given by Crick (1953). The model of P44 used in our MD simulations was built before publication of the crystal structure of the GCN4 coiled coil (O'Shea et al., 1991). Because this crystal structure was not available to serve as a template, our model was built starting with the idealized polyaniline coiled coil of Parry and Suzuki (1969). The model of P44 was built using SYBYL 5.4 (Tripos Associates, St. Louis, Missouri).

Step 1 of model building began with the 56-residue polyaniline coiled coil (Parry & Suzuki, 1969). Each α -helix of this structure contained 28 alanines, so the 6 C-terminal residues of each helix were deleted to yield a 44-residue polyaniline coiled coil. Then the cysteines, leucines, valines, and isoleucines were substituted for the alanines at **a** and **d** (Fig. 11). Next, 2 disulfide bonds were formed between the terminal cysteines. The sulfur atoms of the cysteines at 1a–101a and 22a–122a were close enough so that disulfide bond formation required only small changes in the conformations of the cystine side chains. The disulfide bridges were left-handed spirals, a common disulfide conformation (Richardson & Richardson, 1989).

Step 2 of model building involved increasing the interhelical distance to accommodate the larger side chains of the new hydrophobic residues at **a** and **d**. Systematically varying the interhelical distance as well as the side-chain conformations of these residues and energy minimization of the intermediate structures would have been prohibitively time consuming. Instead, the side-chain conformations of the **a** and **d** residues were constrained, whereas the interhelical distance was allowed to vary. Because the coiled coil is the simplest folded protein structure, we assumed that the side-chain dihedral angles of the leucines, valines, and isoleucines would resemble the angles most commonly found in naturally occurring proteins. Therefore, the 6 leucines, 2 valines, and 2 isoleucines were initially set at their most common rotamer conformations: $\chi_1 = -60^\circ$ and $\chi_2 = 180^\circ$ for leucine, $\chi_1 = 180^\circ$ for valine and isoleucine (Richardson & Richardson, 1989). These rotamer conformations were maintained during energy minimization of the whole molecule by applying torsional constraints with a dihedral force constant of 10 kcal/deg. After 200 steps of minimization using the Powell algorithm, the average interhelical distance had increased by approximately 1 Å and the disulfide bridges remained left-handed spirals.

Step 3 of model building involved replacing the alanines at **e** and **g** with glutamate and lysine residues, respectively. We assumed that it would be possible for all 6 Glu/Lys pairs to form interhelical ion pairs by the glutamate at **e** forming an interhelical salt bridge with the lysine at **g'** of the adjacent heptad of the other helix. These interhelical salt bridges were introduced into the model by setting the equilibrium distance between glutamate CO_2^- carbons (C^5) at **e** and the lysine NH_3^+ nitrogens (N^6) at **g'** to 3.2 Å and minimizing.

Step 4 of model building involved replacing most of the remaining alanines at **b**, **c**, and **f** with the appropriate P44 residues in their most common rotamer conformations. The energy of this structure was minimized without constraints until the convergence criterion ($\Delta E = 0.01$ kcal/mol) was met. The final structure (Fig. 1; Kinemage 1) was used in subsequent MD simulations.

MD simulations

MD calculations were performed with the CEDAR program (Carson & Hermans, 1985) using (1) all-atom representations for CH , CH_2 , and CH_3 groups, (2) a force field with the nonbonded parameters of the GROMOS program (Van Gunsteren & Berendsen, 1987) and the geometry and geometric deformations described earlier (Ferro et al., 1980), (3) simple point charge (SPC) water molecules (Berendsen et al., 1981), (4) periodic boundary conditions, (5) an 8-Å cutoff on nonbonded interactions between formally neutral groups (CH_x , CO , NH , H_2O), (6) the Shake algorithm (Ryckaert et al., 1977) to keep bond lengths rigorously constant, (7) a time step of 2 fs, and (8) a mean temperature (300 K) and pressure (1 atm) kept constant by small adjustments at each time step of the kinetic energy and the dimensions of the periodic box (Berendsen et al., 1981).

MD simulations were performed on 3 structures: (1) the P44 model containing 2 disulfide bridges, (2) reduced P44 in which the 2 S-S groups were replaced by 4 S-H groups, and (3) the crystal structure of the 62-residue GCN4 coiled coil (O'Shea et al., 1991). For reduced P44, after replacement of the disulfide bonds the structure was reminimized, which optimized the thiol–thiol interactions. These 3 structures were solvated with 1,827, 1,784, and 3,875 water molecules, respectively, in rectangular boxes with dimensions $35 \times 36 \times 51$ Å, $34 \times 36 \times 52$ Å, and $39 \times 55 \times 62$ Å, respectively. The crystallographic water molecules were included in the GCN4 simulation. In each simulation the conformationally rigid and translationally static structure was first equilibrated with dynamic water molecules for 20 ps at a temperature of 300 K and a pressure of 1 atm followed by 10 steps of minimization of the whole system. Next the entire system was heated to 300 K stepwise over 2 ps followed by a short constrained MD simulation during which the distance between the NH nitrogen and C=O oxygen of every third $i \rightarrow i - 4$ backbone hydrogen bond was not allowed to exceed 3.2 Å in order to maintain helicity. Finally, the MD simulation was continued for 300 ps without this constraint.

Acknowledgments

We thank Jan Hermans for use of his CEDAR program, Tom Alber for providing the X-ray coordinates of the GCN4 coiled coil before publication, and David and Jane Richardson for their help with the kinemages. This work was supported by a fellowship to J.E.R. from the North Carolina Supercomputing Center and Cray Research, Inc., grants to A.T. from Tripos Associates and the North Carolina Supercomputing Center, and by NIH grant GM 42031 to B.W.E.

References

- Anfinsen CB. 1973. Principles that govern the folding of protein chains. *Science* 181:223–230.
- Baker EN, Hubbard RE. 1984. Hydrogen bonding in globular proteins. *Prog Biophys Mol Biol* 44:97–179.
- Berendsen HJ, Postma JPM, van Gunsteren WF, Dinola A, Haak JR. 1981. Interaction models for water in relation to protein hydration. In: Pullman B, ed. *Intermolecular forces*. Holland: Reidel. pp 331–342.
- Blundell T, Barlow D, Borkakoti N, Thornton J. 1983. Solvent-induced distortions and the curvature of α -helices. *Nature* 306:281–283.
- Burley SK, Petsko GA. 1988. Weakly polar interactions in proteins. *Adv Protein Chem* 39:125–189.
- Carson M, Hermans J. 1985. The molecular dynamics workshop laboratory. In: Hermans J, ed. *Molecular dynamics and protein structure*. Chapel Hill: University of North Carolina. pp 165–166.
- Cohen C, Parry DAD. 1986. α -Helical coiled coils – A widespread motif in proteins. *Trends Biochem Sci* 11:245–248.

- Cohen C, Parry DAD. 1990. α -Helical coiled coils: How to design an α -helical protein. *Proteins Struct Funct Genet* 7:1-15.
- Crick FHC. 1953. The packing of α -helices: Simple coiled-coils. *Acta Crystallogr* 6:689-697.
- Ellenberger TR, Brandl CJ, Struhl K, Harrison SC. 1992. The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted α helices: Crystal structure of the protein-DNA complex. *Cell* 71:1223-1237.
- Engel M, Williams RW, Erickson BW. 1991. Designed coiled-coil proteins: Synthesis and spectroscopy of two 78-residue α -helical dimers. *Biochemistry* 30:3161-3169.
- Ferro DR, McQueen JE, McCrown JT, Hermans J. 1980. Energy minimizations of rubredoxin. *J Mol Biol* 136:1-18.
- Goodman EM, Kim PS. 1991. Periodicity of amide proton exchange rates in a coiled-coil leucine zipper peptide. *Biochemistry* 30:11615-11620.
- Hagler AT, Lifson S, Huler E. 1974. The amide hydrogen bond in energy functions for peptides and proteins. In: Blount ER, Bovey FA, Goodman M, Lotan N, eds. *Peptides, polypeptides and proteins*. New York: Wiley. pp 35-48.
- Hodges RS, Saund AK, Chong PCS, St. Pierre SA, Reid RE. 1981. Synthetic model for two-stranded α -helical coiled-coils: Design, synthesis, and characterization of an 86-residue analog of tropomyosin. *J Biol Chem* 256:1214-1224.
- Hodges RS, Semchuk PD, Taneja AK, Kay CM, Parker JMR, Mant CT. 1988. Protein design using model synthetic peptides. *Pept Res* 1:19-30.
- Hodges RS, Zhou NE, Kay CM, Semchuk PD. 1990. Synthetic model proteins: Contribution of hydrophobic residues and disulfide bonds to protein stability. *Pept Res* 3:123-137.
- Hu JC, O'Shea EK, Kim PS, Sauer RT. 1990. Sequence requirements for coiled-coils: Analysis with 1 repressor-GCN4 leucine zipper fusions. *Science* 250:1400-1403.
- Landschulz WH, Johnson PF, McKnight SL. 1988. The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science* 240:1759-1764.
- Lau SYM, Taneja AK, Hodges RS. 1984. Synthesis of a model protein of defined secondary and quaternary structure: Effect of chain length on the stabilization and formation of two-stranded α -helical coiled-coils. *J Biol Chem* 259:13253-13261.
- Lupas A, Van Dyke M, Stock, J. 1991. Predicting coiled coils from protein sequences. *Science* 252:1162-1164.
- McLachlan AD. 1978. The double helix coiled coil structure of murein lipoprotein from *Escherichia coli*. *J Mol Biol* 122:493-506.
- McLachlan AD, Stewart M. 1975. Tropomyosin coiled-coil interactions: Evidence for an unstaggered structure. *J Mol Biol* 98:293-304.
- Nemethy G, Phillips DC, Leach SJ, Scheraga HA. 1967. A second right-handed helical structure with the parameters of the Pauling-Corey α -helix. *Nature* 214:363-365.
- Nilges M, Brünger AT. 1991. Automated modeling of coiled coils: Application to the GCN4 dimerization region. *Protein Eng* 4:649-659.
- Oas TG, McIntosh LP, O'Shea EK, Dahlquist FW, Kim PS. 1990. Secondary structure of a leucine zipper determined by nuclear magnetic resonance spectroscopy. *Biochemistry* 29:2891-2894.
- O'Neil KT, DeGrado WF. 1990. A thermodynamic scale for the helix-forming tendencies of the commonly occurring amino acids. *Science* 250:646-651.
- O'Shea EK, Klemm JD, Kim PS, Alber T. 1991. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* 254:539-544.
- Parry DAD, Suzuki E. 1969. Intrachain potential energy of the α -helix and of a coiled coil strand. *Biopolymers* 7:189-197.
- Pauling L, Corey RB. 1953. Compound helical configuration of polypeptide chains: Structure of proteins of the α -keratin type. *Nature* 171:59-61.
- Presta LG, Rose GD. 1988. Helix signals in proteins. *Science* 240:1632-1641.
- Richardson JS, Richardson DC. 1988. Amino acid preferences for specific locations at the ends of α helices. *Science* 240:1648-1652.
- Richardson JS, Richardson DC. 1989. Principles and patterns of protein conformation. In: Fasman GD, ed. *Prediction of protein structure and principles of protein conformation*. New York: Plenum Press. pp 1-98.
- Ryckaert JP, Ciccoliti G, Berendsen HJC. 1977. Numerical integration of a cartesian system with constraints: Molecular dynamics of n-alkanes. *J Comput Phys* 23:327-341.
- Stickle FD, Presta LG, Dill KA, Rose GD. 1992. Hydrogen bonding in globular proteins. *J Mol Biol* 226:1143-1159.
- Talbot JA, Hodges RS. 1982. Tropomyosin: A model protein for studying coiled-coil and α -helix stabilization. *Acc Chem Res* 15:224-230.
- Taylor R, Kennard O, Versichel W. 1984. Geometry of the N-H...O=C hydrogen bond. 2. Three-center ("bifurcated") and four-center ("trifurcated") bonds. *J Am Chem Soc* 106:244-248.
- Tropsha A, Bowen JP, Brown F, Kizer JS. 1991. Do interhelical side chain-backbone hydrogen bonds participate in formation of leucine zipper coiled coils? *Proc Natl Acad Sci USA* 88:9488-9492.
- Van Gunsteren WF, Berendsen HJC. 1987. *Groningen molecular simulation manual*. Groningen: Biomos.
- Yue K, Dill KA. 1991. Inverse protein folding: Designing polymer sequences. *Proc Natl Acad Sci USA* 89:4163-4167.
- Zhou NE, Kay CM, Hodges RS. 1992. Synthetic model proteins: Positional effects of interchain hydrophobic interactions on stability of two-stranded α -helical coiled coils. *J Biol Chem* 267:2664-2670.
- Zhou NE, Kay CM, Hodges RS. 1993a. Disulfide bond contribution to protein stability: Positional effects of substitution in the hydrophobic core of the two-stranded α -helical coiled-coil. *Biochemistry* 32:3178-3187.
- Zhou NE, Zhu BY, Sykes JD, Hodges RS. 1993b. Relationship between amide proton chemical shifts and hydrogen bonding in amphipathic α -helical peptides. *J Am Chem Soc* 114:4320-4326.
- Zhu BY, Zhou NE, Kay CM, Hodges RS. 1993. Packing and hydrophobicity effects on protein folding and stability: Effects of β -branched amino acids, valine and isoleucine, on the formation and stability of two-stranded α -helical coiled coils/leucine zippers. *Protein Sci* 2:383-394.