RESEARCH ARTICLES

Helix Packing in Proteins: Prediction and Energetic Analysis of Dimeric, Trimeric, and Tetrameric GCN4 Coiled Coil Structures

Warren L. DeLano and Axel T. Brünger
The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06520

ABSTRACT A simulated annealing method for atomic resolution structure prediction of α-helical coiled coil proteins is described which draws upon knowledge of the oligomerization state, the helix directionality, and the properties of heptad repeat sequences. Unknown structural parameters, such as the coiled coil twist angle and the side chain conformations, are heavily sampled while allowing for flexibility in the helix backbone geometry. Structures of the wild-type GCN4 dimer [O'Shea et al., Science 254:539–544, 1991] and a mutant tetramer [Harbury et al., Science 292:1401–1407, 1993] have been generated and compared with the X-ray crystal structures. The wild-type dimer model has a root mean square coordinate deviation from the crystal structure of 0.73 Å for nonhydrogen atoms in the dimerization interface. Structures of a mutant dimer and a mutant trimer have been predicted. Packing energetics were analyzed for core leucine and isoleucine side chains in dimeric and tetrameric coiled coils. Strong packing preferences were found in the dimers but not in the tetramers. Thus, packing in the dimer may be responsible for the switch from a two-stranded to a four-stranded coiled coil caused by the GCN4 leucine zipper mutations. © 1994 Wiley-Liss, Inc.

Key words: structure prediction, helix to helix packing, coiled coils, leucine zippers, heptad repeats, molecular dynamics, simulated annealing

INTRODUCTION

The ultimate goal of protein structure prediction is generation of an accurate set of atomic coordinates for a protein from the sequence alone. Yet by themselves, computerized prediction methods have difficulty achieving this aim due to the magnitude of the conformational search problem and the inadequacies of current empirical potential energy func-

tions. The first limitation arises from the finite amount of computing power available, and the second from inadequate treatment of solvation energy, entropic effects, electrostatic polarizability, and long range forces. Thus, additional information must be obtained from experiments or from prior knowledge of protein structure which can both limit the scope of the search and compensate for deficiencies in the empirical energy functions.

One way to simplify the search is to model protein structures on a discrete three-dimensional lattice.[1,2] Knowledge-based potential energy functions for use on the lattice can be derived from the geometries of existing protein structures.[3] These simplifications provide a substantial gain in performance and make it possible to explore a reasonable amount of conformational space with exhaustive or stochastic search algorithms. Structures of proteins with simple α-helical folds have been predicted with C^α atom rms coordinate deviations from the crystal structures in the range of 2–4 Å.[2] These low-resolution predictions can be used as starting points for atomic resolution computer simulations using full empirical energy functions.[4]

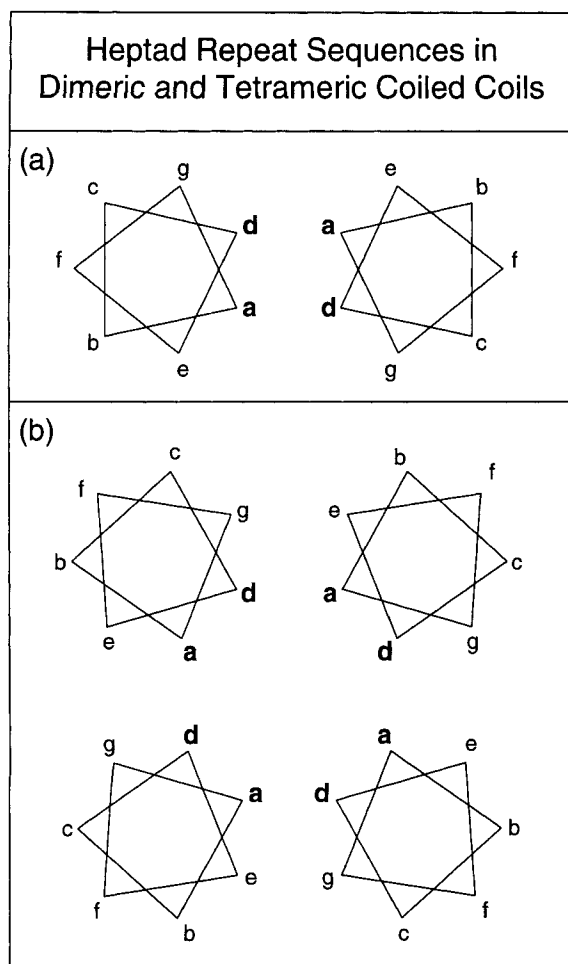Another way to reduce the scope of a conformational search is to use structural information de-

Fig. 1. Heptad repeat sequences in (a) dimeric and (b) tetrameric coiled coils. The core a and d positions are usually occupied by hydrophobic residues.

rived from sequence analysis. Heptad repeat sequences (repeats of the form [abcdefg]$_n$ where a and d are hydrophobic) often indicate the presence of interacting helices (Fig. 1).[5] Secondary structure prediction algorithms can provide clues about the locations of helices, β-sheets, and turn regions,[6,7] and close sequence homology with a protein of known structure can furnish a reliable estimate of the overall protein fold.[8]

For purposes of structure determination, deficiencies in the empirical energy functions can be compensated for by direct incorporation of experimental data.[9] Apart from high resolution X-ray crystallography and solution NMR spectroscopy, other techniques can often provide useful but limited structural information. Some examples are biochemical assays coupled with site directed mutagenesis that identify residues with structural roles,[10,11] circular dichrosim measurements that give approximate percentages of helicity,[12] and solid state NMR experiments that measure distances between a small set of

isotopically labeled atoms.[13] For proteins whose structures are difficult to solve with NMR or crystallography, such as integral membrane proteins, other biochemical and biophysical techniques coupled with judicious molecular modelling can provide a means of approximate structure determination.[11,14]

Here we present a simulated annealing method for the prediction of coiled coil helical bundle structures from experimental data. The oligomerization states and the directionality of the coiled coils are indicated from biophysical and biochemical data, and the interacting faces of the helices are determined from analysis of the heptad repeat sequences. The remaining conformational space is explicitly sampled by starting with models that have a range of twist angles and a variety of initial side chain conformations.

We apply our prediction method to the wild-type GCN4 leucine zipper peptide and to several of the mutant GCN4 peptides synthesized and characterized by Harbury et al.[16] The mutant peptides contain various substitutions at the a and d positions of the heptad repeat sequences which cause them to populate nonnative oligomerization states (Fig. 2). We produce structures for a mutant dimer (p-IL), a mutant trimer (p-II), and a mutant tetramer (p-LI), as well as for the wild-type dimer (GCN4-p1). Since the structures of both the wild-type dimer (GCN4-p1) and the mutant tetramer (p-LI) had previously been solved by X-ray crystallography,[15,16] these structures serve as test cases for our simulated annealing protocol. Future comparisons of the predicted mutant dimer (p-IL) and trimer (p-II) structures with experimentally determined structures will serve as independent tests of the method.

We also analyze the packing energetics of core residues in dimeric and tetrameric coiled coils. Two of the mutant GCN4 peptides, p-IL and p-LI, form dimers and tetramers, respectively, in solution, but they differ chemically only in the locations of core leucine and isoleucine side chains. This suggests that the specific packing of core side chains is a central aspect of coiled coil oligomerization. We generate models for a hypothetical dimeric coiled coil with the p-LI sequence and for a hypothetical tetrameric coiled coil with the p-IL sequence. By comparing the observed p-IL dimer with the hypothetical p-LI dimer, and the observed p-LI tetramer with the hypothetical p-IL tetramer, we analyze how differences in the sequences give rise to the unique oligomerization states. Past studies have identified both statistical and energetic preferences for particular residues at the a and d positions of coiled coil sequences.[17-19] We calculate the preferences of leucine and isoleucine for these positions in the mutant coiled coils. It has been proposed that the packing of heptad a and d positions is simply reversed between parallel dimers and tetramers.[16] In contrast, our en-

Peptide	Oligomer	a	d	a	d	a	d	a	d	a
GCN4-p1	Dimer	MKQLEDKVEELLSKNYHLENEVARLKKLVGQ								
p-IL	Dimer	---L---I--L---I--L---I--L---I--								
p-II	Trimer	---I---I--I---I--I---I--I---I--								
p-LI	Tetramer	---I---L--I---L--I---L--I---L--								

Fig. 2. Sequences and oligomerization states of the wild-type and mutant GCN4 peptides. The a and d positions of the heptad repeat sequences are substituted with leucine and isoleucine residues in the mutant peptides with the exception of the N-terminal methionine which remains unchanged.

ergy calculations suggest that packing at these positions is unique in both species.

METHODS

Our structure prediction method is shown schematically in Figure 3 and summarized as follows: First, a large number of independent simulated annealing minimizations are run from a variety of starting models that heavily sample the unknown conformational parameters. Solvation effects are mimicked through the use of restraints and through screening of charges. During the minimizations, the model structures converge on conformations with the lowest free energies. Clusters of converged structures are then identified using rms distance comparisons, and the lowest energy cluster is selected. A probability map refinement procedure is then used to generate a representative structure from the cluster. All of the simulated annealing, energy minimizations, refinements, and energy calculations are carried out with a developmental version of the X-PLOR program.[20]

Initial Coordinates and Restraints

To prevent the introduction of biases from our specific knowledge of the GCN4-p1 dimer and p-LI tetramer structures, we linked the decisions and assumptions made while setting up the simulations to specific biochemical and biophysical data. Thus, this work should serve as a prototype for future predictive studies when an atomic resolution structure is unavailable.

One of our assumptions was that the sequences would adopt α-helical backbone conformations. This had been suggested for GCN4-p1 by the presence of heptad repeats in the sequence and confirmed by circular dichroism measurements prior to determination of the crystal structure.[21] Similar results demonstrating helicity had been obtained for the mutant peptides.[16] The initial C^α positions for our models were those of an ideal straight α-helix (ϕ = -57, ψ = -47). Helical backbone geometry was explicitly maintained during the simulations through the imposition of one-sided distance restraints with an upper bound of 3.2 Å between i oxygens and i +

4 nitrogens along the length of the helix. Our experience has shown that these pseudohydrogen bond restraints allow for considerable flexibility but prevent helices from unraveling or forming sharp kinks during in vacuo molecular dynamics simulations.

The oligomerization preferences of both the wild-type and mutant peptides had previously been determined by analytical ultracentrifugation.[16,21] GCN4-p1 and p-IL both formed dimers in solution, while the p-II and p-LI peptides populated trimeric and tetrameric states, respectively. We modeled all of these peptides in their experimentally observed oligomerization states. For the subsequent energetic analysis, two hypothetical structures were generated with p-IL modeled as a tetramer and p-LI modeled as a dimer.

The directionality of the helices in the GCN4-p1 and p-LI coiled coils had been deduced before the crystal structures were solved. Disulfide cross-linking experiments showed that in all of the observed coiled coils (GCN4-p1, p-IL, p-II, and p-LI), the helices were aligned in parallel (i.e., all of the N-termini were at one end of the coiled coil, and all C-termini were at the other end).[16,21] Therefore, all of the initial models were constructed with parallel directionality. The rotation of the helices about their axes was assumed to be such that the hydrophobic surfaces, evident from the a and d positions of the heptad repeats, would face inward to form a hydrophobic core. The initial separation between the helices was set to 10 Å, a typical value for the separation of neighboring helices in protein structures.[22] The two hypothetical structures, the p-LI dimer and the p-IL tetramer, were modeled in the same manner as the observed p-IL dimer and p-LI tetramer structures, respectively.

One conformational parameter not apparent from any of the biochemical or biophysical data was the twist angle of the coiled coil. We defined the twist angle to be the average dihedral angle formed between the coiled coil's symmetry axis and the axis of each individual helix (Fig. 4). The axis vector of each α-helix was defined by least-squares fitting an ideal helix to the one being measured and using the axis vector of the ideal helix. The base of this vector was
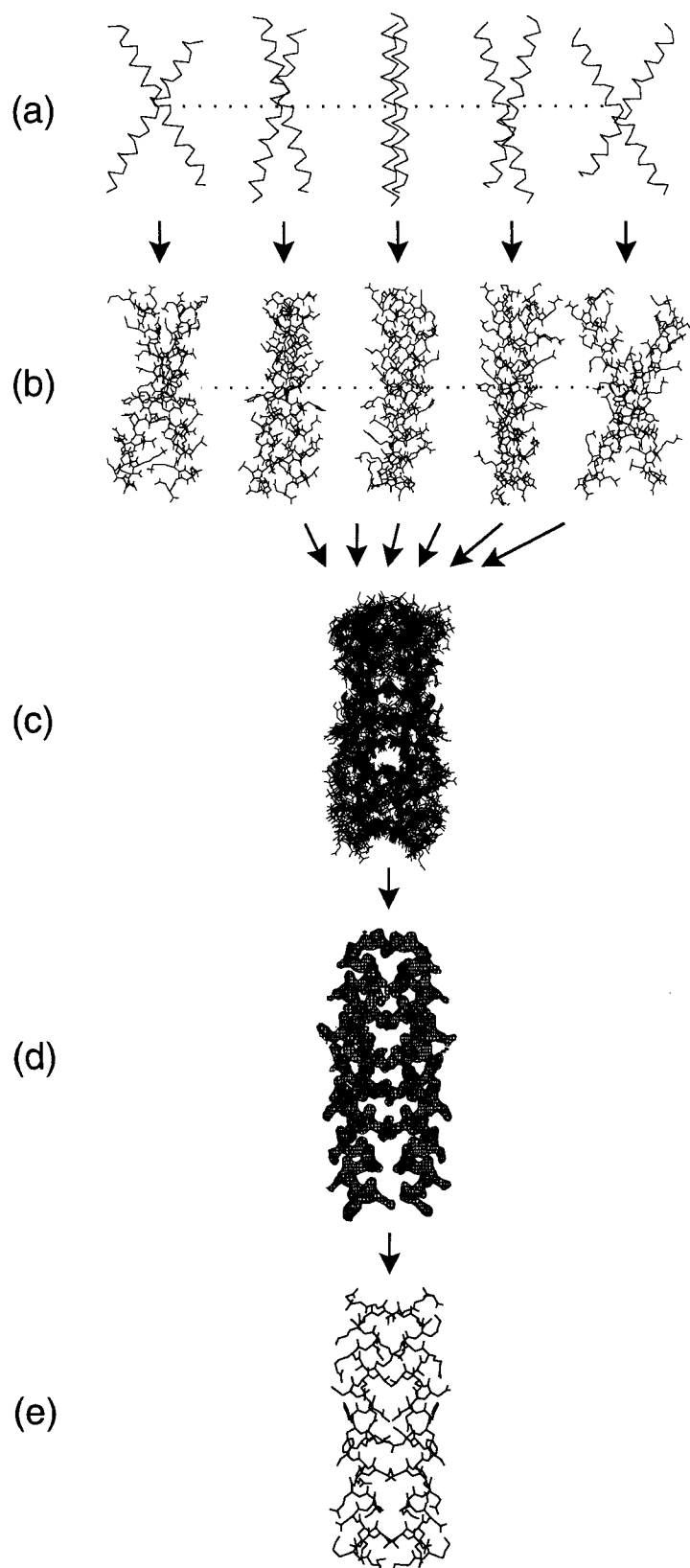
Fig. 3. Outline of the structure prediction method: (a) generation of the initial Cα atom coordinates, (b) side chain growth and conformational relaxation by simulated annealing and energy minimization, (c) cluster analysis, (d) generation of a probability map, and (e) refinement of a structure which best fits the map.
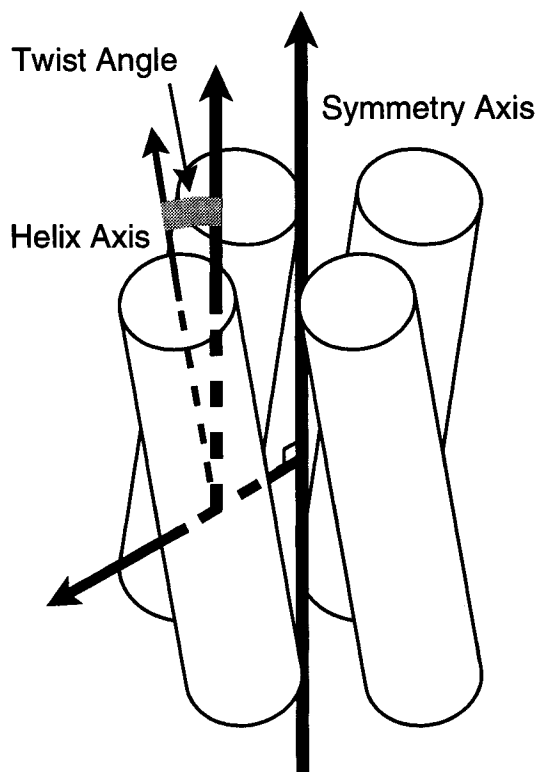
Fig. 4. The twist angle is defined to be the average dihedral angle formed between the axis of each helix and the symmetry axis of the coiled coil.

set to the geometric center of the helix. The coiled coil's symmetry axis vector was defined by adding all of the helix axis vectors head to tail, with the vector base located at the geometric center of the coil. For the dimeric coiled coils, the twist angle is equal to one-half of the helix–helix crossing angle. Our initial models were given initial twist angles ranging from −35° to 35° in one degree increments, resulting in 71 structures for each set. There were six sets of structures in all: four observed coiled coils (the GCN4-p1 and p-IL dimers, the p-II trimer, and the p-LI tetramer) and two hypothetical coiled coils (the p-LI dimer and the p-IL tetramer).

Initial atomic coordinates for the coiled coils were generated in the following manner: (1) C^α coordinates for an ideal straight α-helix were generated with the helix centered about the origin and oriented so that the helix axis coincided with the z-axis. (2) A vector was then extended into the xy-plane from the center of the helix so that it bisected the hydrophobic face of the helix. This head of this vector was to become the geometric center of the coiled coil, so the vector was given a length of 5 Å for the dimers, 5.77 Å for the trimer, and 7.07 Å for the tetramers in order to generate a uniform 10 Å sep-

aration between neighboring helices. (3) A 2-fold, 3-fold, or 4-fold axis of symmetry parallel to the z-axis and passing through the head of the vector was then used to generate the other helices and vectors in the dimeric, trimeric, or tetrameric coils, respectively. (4) The helices were then individually rotated about the described vectors in the xy-plane to generate coiled coils with specific twist angles.

For each initial structure, side chain and backbone atoms were grown from the C^α positions by applying a protocol similar to those used for the generation of initial coordinates in NMR structure determination.[23] The positions of all other atoms besides the C^α atom in each residue were set to random positions inside a sphere of radius 2 Å centered about the C^α atom position. Then, several short rounds of minimization and molecular dynamics with a soft repulsive nonbonded potential were performed with the C^α atoms fixed in order to generate a large variety of initial side chain conformations that have reasonable starting geometry.

Empirical Energy Function

For all subsequent calculations, the CHARMM PARAM19 parameter set[24] was used with a dielectric constant (ε) of 1.0. The nonbonded interaction cutoff distance was 9.5 Å with a switching function active from 5 to 9 Å as recommended by Loncharich and Brooks.[25] In order to mimic the effects of solvation and electrostatic shielding during these in vacuo simulations, formal charges on lysine, arginine, glutamine, asparagine, and histidine side chains were scaled by a factor of 0.3, and a linear distance-dependent screening function was employed. In addition, terminal charges on the helix end groups were neutralized.

Conformational Relaxation by Simulated Annealing

Once initial coordinates had been assigned, each structure was relaxed with the following simulated annealing protocol: (1) a 5 ps molecular dynamics slow-cooling stage from 500 K to 300 K, (2) a 20 ps constant temperature molecular dynamics simulation at 300 K, and (3) 1000 steps of conjugate gradient minimization. During the slow-cooling stage, C^α atoms were held in place with harmonic point restraints that were slowly reduced. The α-helical hydrogen bond restraints were active during all stages, but no other restraints were applied during the constant temperature molecular dynamics and energy minimization stages. A time step of 0.5 fs was used for temperatures above 350 K during the slow-cooling stage, otherwise a time step of 1 fs was used. Graphs of the twist angle vs. simulation time are shown in Figure 5. During simulated annealing, the
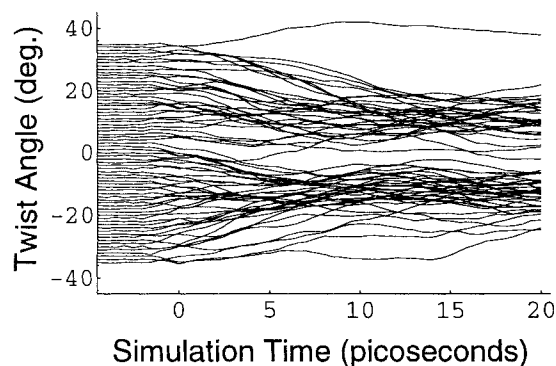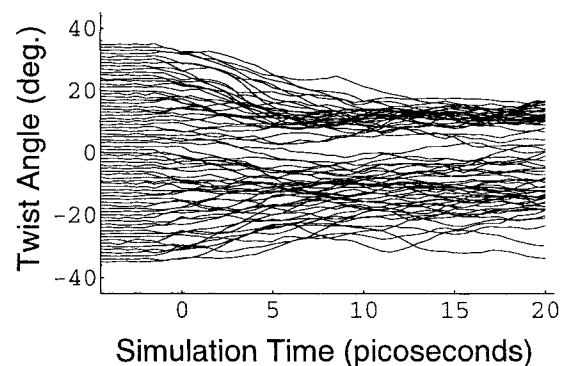
(a) GCN4-p1 (observed dimer)
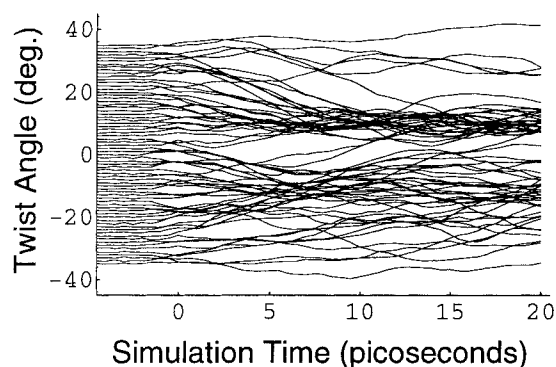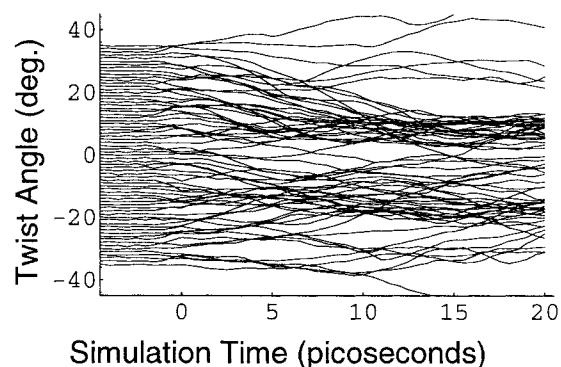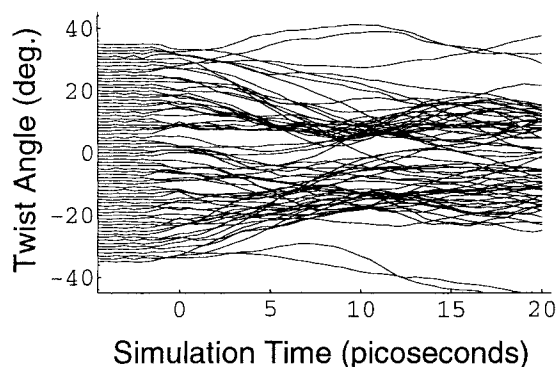

(b) p-IL (observed dimer)

(c) p-II (observed trimer)

(d) p-LI (observed tetramer)

(e) p-LI (hypothetical dimer)

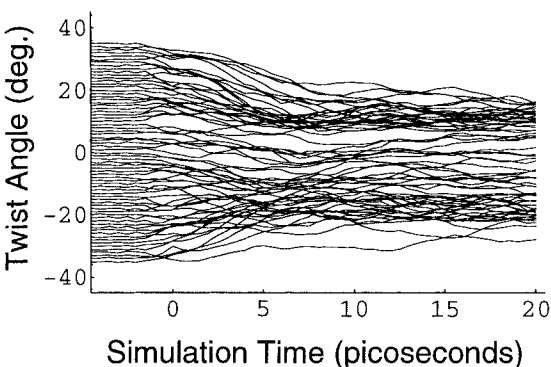## (f) p-IL (hypothetical tetramer)

Fig. 5. Graphs of twist angle vs. simulation time for the simulated annealing structures (a–f). Twist angles are shown during the slow cooling stage (−5 . . . 0 ps) and the constant temperture molecular dynamics simulation (0 . . . 20 ps).

model structures formed clusters about specific twist angles in all six cases.

Cluster Analysis

Cluster analysis was performed on each set of relaxed structures, with clustering based on a rms dis-tance criterion between all possible pairs of structures in the set. Any two structures that had a rms coordinate difference for C^α atoms below a certain cutoff value were grouped in the same cluster. A cutoff value of 1.6 Å was chosen by trial and error because it consistently produced a large cluster

while excluding structures with significant devia-
tions. The final results proved to be relatively insen-
sitive to the exact choice of the cutoff value so long
as a reasonably sized cluster (10 or more structures)
was obtained. In order to preserve any asymmetry in
the structures, rms comparisons were performed re-
peatedly as to reflect the 2-fold, 3-fold, and 4-fold
axis of near-symmetry of the dimeric, trimeric, and
tetrameric coiled coils, respectively. Once the best
match was found in each pair of structures, one
structure was then rotated (if necessary) so that the
two could eventually be superimposed without aver-
aging the asymmetry.

Generation of a Representative Structure

Once a cluster had been defined for each set of
simulated annealing structures, a representative
model was generated. First, the structures in the
cluster were used to generate a three-dimensional
probability map (analogous to an electron density
map) which combined all of the structural informa-
tion in the cluster. Then, a model was refined
against the map to produce a single representative
structure.

We found this approach to be superior to a simple
geometric average of the atomic coordinates. A
geometric average produces a poor representation of
a protein structure ensemble because it fails to
properly address the anisotropic distribution of
atomic coordinates in side chain conformations. An
extreme example of this is shown in Figure 6 for
a histidine residue which adopts two principal
conformations. The geometric average places the
side chain coordinates in an intermediate position
which has a distorted geometry and which matches
neither conformation very well. In contrast, the
calculated probability map retains information
about both conformations. Side chains in the refined
structure will usually adopt the more probable
conformations based on their better fit with the
map. Furthermore, since the refined structure has
undistorted geometry, the energy is also reasonable
and representative of the original cluster. In our
experience, this approach has been more effective
than taking the lowest energy structure in a cluster
because it consistently produced a model that was
much closer to the crystal structure than the lowest
energy structure.

The probability map was generated as follows.
First, all of the structures in the cluster were super-
imposed about the average C^α coordinates after pro-
cessing for asymmetry. A three-dimensional map
enclosing the structures was then defined with a
grid spacing of 0.5 Å and cushion of 3 Å around the
molecules. The probability density was calculated
by a sum of five Gaussians as is commonly done to
approximate atomic form and Debye–Waller fac-
tors.[26] The density at a vertex in the map with co-
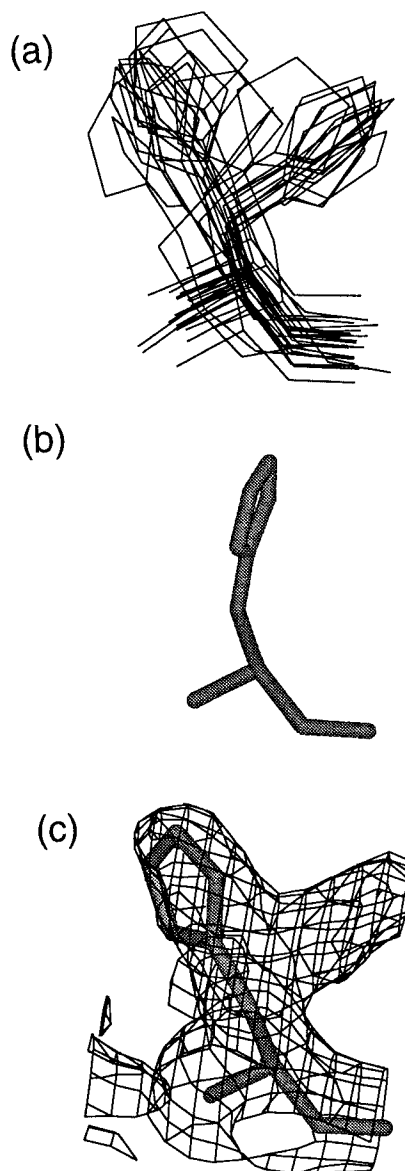ordinate $\mathbf{r}$ is therefore given by

Fig. 6. Comparison of (a) a conformational ensemble, (b) the
geometric average, and (c) the refined structure superimposed
with the probability map. The structure refined in the probability
map produces a better representation of the populated side chain
conformations than does a geometric average.

$$\rho_{cluster}(\mathbf{r}) = \frac{1}{N_M} \sum_{j \in M} \sum_{i \in A} \sum_{k=1}^{5} a_{kC}$$

$$\left(\frac{4\pi}{b_{kC} + B_i}\right)^{3/2} \exp\left[\frac{-4\pi^2(\mathbf{r} - \mathbf{r}_i)^2}{b_{kC} + B_i}\right] \quad (1)$$

where M is the set of structures for which the prob-
ability density is being computed and N_M is the
number of structures in the set, A is the set of non-
hydrogen atoms in the jth structure surrounding the

vertex, and r_i and B_i are the coordinate and effective temperature factors, respectively, of the ith atom in the set N. We arbitrarily chose a_{kC} and b_{kC} ($k = 1 \ldots$ 5) to be the atomic scattering factors for carbon, which were used uniformly for all nonhydrogen atoms in the structure (b_{5C} is zero). The set of atoms N about each vertex was chosen so that only atoms contributing significantly ($> 10^{-7}$ Å^{-3}) to the density at the vertex were included. Effective temperature factors were assigned on a per-residue basis using the formula[27]

$$B_i = \frac{8}{3} \pi^2 \sigma^2 \qquad (2)$$

where σ denotes the rms deviation of the residue's C$^\alpha$ atom from its mean coordinate.

For each cluster, a single structure was refined against the probability map by transforming the map into reciprocal space and then using crystallographic refinement techniques. Conversion of the discrete probability map into reciprocal space provided an efficient way to obtain gradients with respect to the atomic coordinates for refinement of our model structures. This procedure is essentially equivalent to directly optimizing the model structure against the probability map in real space. The target function consisted of

$$T = E_{EMPIRICAL} +$$
$$W' \sum_{\mathbf{h}} \left(\frac{[A_{\text{cluster}}(\mathbf{h}) - kA_{\text{model}}(\mathbf{h})]^2 +}{[B_{\text{cluster}}(\mathbf{h}) - kB_{\text{model}}(\mathbf{h})]^2} \right) \qquad (3)$$

where $E_{\text{EMPIRICAL}}$ is the empirical energy function and the summation comprises the vector residual.[28] W' is an overall weight, k is a scale factor, and h is a reciprocal lattice vector. The real and imaginary components of the structure factors are given by

$$A_{\text{cluster}}(\mathbf{h}) + iB_{\text{cluster}}(\mathbf{h}) = FT[\rho_{\text{cluster}}(\mathbf{r}); \mathbf{h}] \quad (4)$$

$$A_{\text{model}}(\mathbf{h}) + iB_{\text{model}}(\mathbf{h}) = FT[\rho_{\text{model}}(\mathbf{r}); \mathbf{h}] \quad (5)$$

where FT denotes a three-dimensional Fourier transform, ρ_{cluster} ($\mathbf{r}$) is the probability density of the cluster [Eq. (1)], and ρ_{model} ($\mathbf{r}$) is the probability density of the model being refined. ρ_{model} ($\mathbf{r}$) was calculated in the same manner as ρ_{cluster} ($\mathbf{r}$) except that the set M [Eq. (1)] consisted of only the single model structure. The effective temperature factors of ρ_{model} ($\mathbf{r}$) were set to those of ρ_{cluster} ($\mathbf{r}$).

The lowest energy structure in each cluster served as a starting point for simulated annealing refinement. Weighting between the empirical energy and the vector residual was obtained by balancing the gradients of the two terms.[29] The refinement was carried out as follows: (1) 200 steps of conjugate gradient minimization, (2) a slow-cooling stage from 3000 K to 300 K over 2.5 ps,[30] (3) a side chain rotamer search for core side chains (see below), and (4) 1000 steps of minimization against the empirical energy term alone. All refinements were carried out with calculated reflections between 15 and 2.5 Å resolution.

Side chain rotamer searches for the core residues were performed using the rotamer library from Tuffery et al.[31] Side chains were individually flipped through their rotamers and subjected to 50 steps of minimization against the target function, T, while the rest of the molecule was held fixed. During minimization, side chain atoms were allowed to interact only with themselves and main chain atoms. Side chain conformations producing the best match with the probability map (as measured by the vector residual) were then selected. Once all of the core side chains (**a** and **d** positions) were independently optimized, the entire molecule was subjected to 200 steps of minimization against T [Eq. (3)] to remove any bad geometries that had arisen. Plots of rms coordinate deviations from the crystal structures for GCN4-p1 and p-LI during the refinement process are shown in Figure 7. Since simulated annealing refinement was quite efficient at optimizing side chain conformations, the rotamer search mainly affected the relatively few side chains that had not yet been correctly placed.

Because a reported crystal form of the p-II trimer appears to have a 3-fold symmetry axis,[16] an additional refinement was carried out for p-II with explicit symmetry. A symmetric probability map was generated by superimposing the asymmetric simulated annealing structures repeatedly while rotating them about their axis of near-symmetry. Following the rotamer search, helices in the model were constrained to adopt identical geometry through the use of noncrystallographic least-squares restraints.[32] The symmetric structure will be referred to as p-II$_{\text{SYM}}$.

Calculation of Interaction Energies

Interaction energies between groups of atoms U and W were defined to be

$$E_{UW} = \sum_{u \in U, w \in W} E_{uw}^{\text{EMPIRICAL}} \qquad (6)$$

where $E_{uw}^{\text{EMPIRICAL}}$ represents all terms of the empirical energy function that involve both the u and w atoms. Mean interaction energies were computed from the structures in each cluster. Statistical error estimates for these means were computed from the distribution of energies present in the cluster using the formula $\sigma/\sqrt{n}$, where σ is the standard deviation of the computed energies and n is the number of structures in the cluster.

RESULTS AND DISCUSSION
Conformational Relaxation and Cluster Analysis

Graphs of the empirical energy vs. twist angle for the six sets of simulated annealing structures are

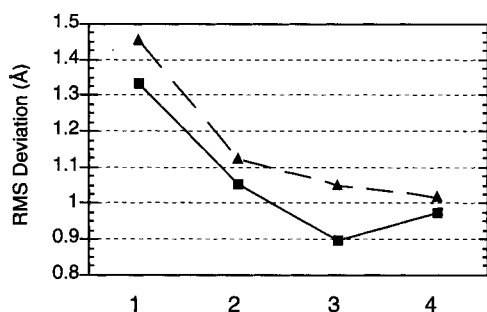## (a) GCN4-p1 (dimer)
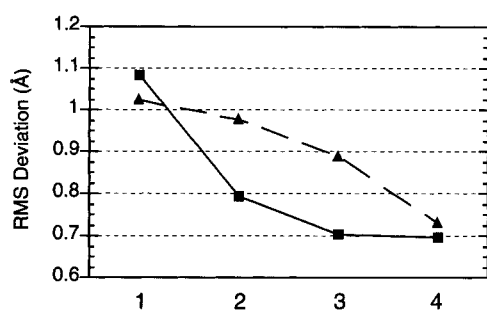
(b) p-LI (tetramer)

Fig. 7. Plots of the rms coordinate deviations from the crystal structures for the (a) GCN4-p1 and (b) p-LI models during refinement against the probability map. The rms deviations for C^α atoms are indicated by solid lines, and the rms deviation for nonhydrogen atoms of core residues are indicated by dashed lines. Column 1 is the starting model (the lowest energy structure in the cluster), column 2 is the structure following minimization against the probability map, column 3 is the structure after simulated annealing refinement against the probability map, and column 4 is the final structure following the rotamer search and energy minimization.

shown in Figure 8. Left-handed conformations (with positive twist angles) are consistently lower in energy than the right-handed conformations (with negative twist angles). The formation of minimum energy structures having twist angles of approximately 10° (corresponding to helix-helix crossing angles near 20°) is in agreement with the coiled coil packing arguments proposed by Crick.[33] The low-energy outlier in the set of p-II structures (Fig. 8c) corresponds to a simulation where the coiled coil became significantly distorted. If explicit solvent had been included in our calculations, such a structure would probably have been heavily penalized because it exposed many hydrophobic isoleucine residues. Even without the inclusion of explicit solvent molecules, formation of unphysical structures might have been prevented by the inclusion of an effective solvation energy term.[34] However, current formulations show a low degree of accuracy[35] and require a substantial amount of CPU time to evaluate since they require calculation of solvent accessible surface areas and their derivatives.

The convergence of the simulated annealing structures toward structures with specific twist an-

gles is shown in Figure 5. For each simulation set, C^α rms distance-based cluster analysis produces both left-handed and right-handed clusters that correspond to the bands of structures with positive and negative twist angles, respectively. The lowest energy cluster from each set is chosen for probability map refinement, and in each case, this turns out to be a left-handed cluster that contains a subset of structures in the left-handed bands. The existence of a bimodal distribution containing both left-handed and right-handed clusters of structures reflects the occurrence of interhelical angles found in a large number of X-ray crystal structures from helical bundle proteins.[36]

Successful Prediction of the Known Structures: The GCN4-p1 Dimer and the p-LI Tetramer

The probability map refined model of the GCN4-p1 dimer is significantly closer to the crystal structure than models from previous studies. The rms coordinate deviation for C^α atoms is 0.70 Å as compared to 0.81 Å[4] and 1.26 Å,[37] and the rms coordinate deviation for nonhydrogen atoms of residues at the dimerization interface is 0.73 Å as compared to 1.31 Å[4] and 1.75 Å.[37] Figure 9a and b shows two core residues of the predicted structure and the probability map superimposed with the crystal structure. Excellent agreement is seen not only between the structures themselves, but also between the map density and the crystal structure. By allowing structural asymmetries to persist throughout the clustering and refinement procedures, the asymmetry of asparagine-16 has been correctly captured and reproduced. All of the other core residues, valines in the a positions and leucines in the d positions, adopt symmetric conformations in agreement with the crystal structure. This differs from earlier work where asymmetric leucine conformations were obtained.[38]

The rms coordinate deviations from the crystal structure for the predicted p-LI tetramer are 0.97 Å for C^α atoms and 1.02 Å for nonhydrogen atoms in the hydrophobic core. Core residues of the tetramer are shown in Figure 9c and d. As is the case for the GCN4-p1 dimer structure, the density present in the probability map permits unambiguous determination of most side chain conformations. Although precise positioning of the side chain atoms in the p-LI model is not as accurate as in the GCN4-p1 case, the predicted rotamers are very similar to those of the crystal structure. It is encouraging that accurate predictions of the core atoms were obtained without the inclusion of explicit solvent molecules. This suggests that solvent does not specifically determine the conformations of core side chains. On the other hand, solvent, acting through the hydrophobic effect, probably provides a strong driving force behind the formation and aggregation of the helices. In our
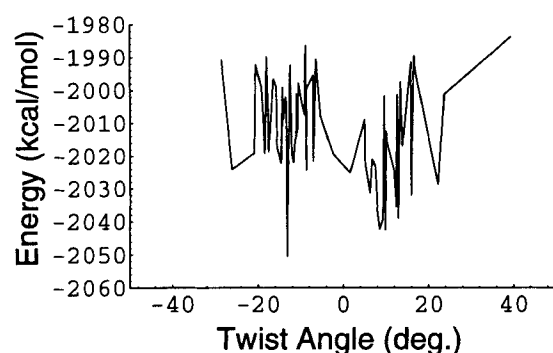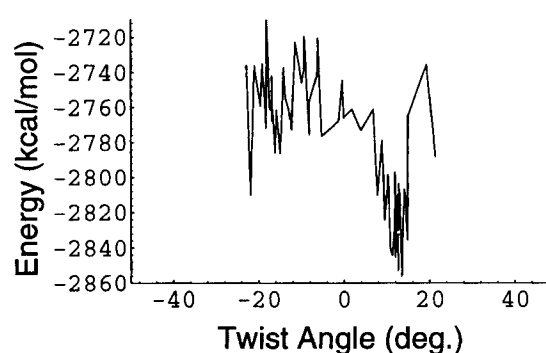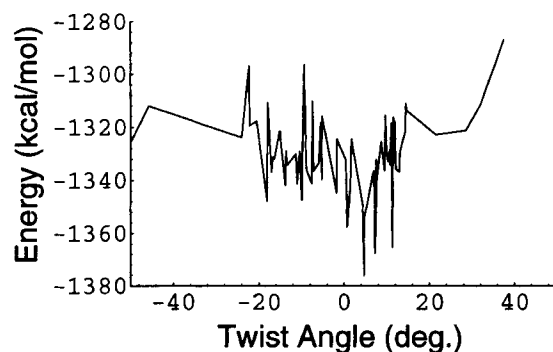
(a) GCN4-p1 (observed dimer)
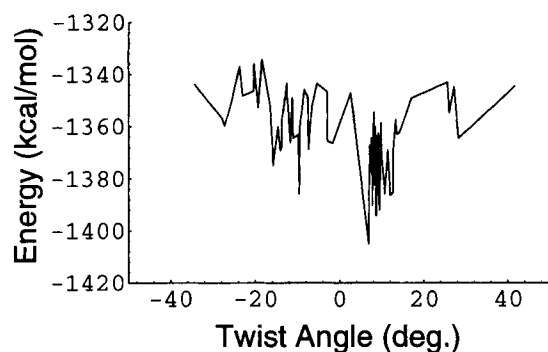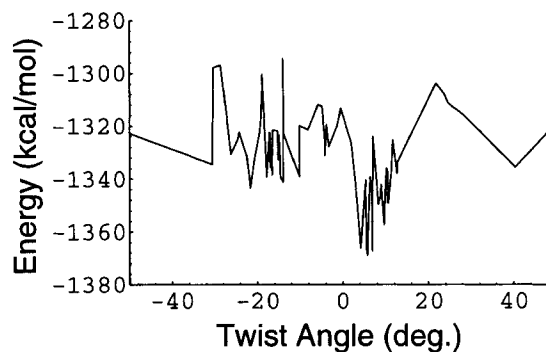
(b) p-IL (observed dimer)

(c) p-II (observed trimer)

(d) p-LI (observed tetramer)

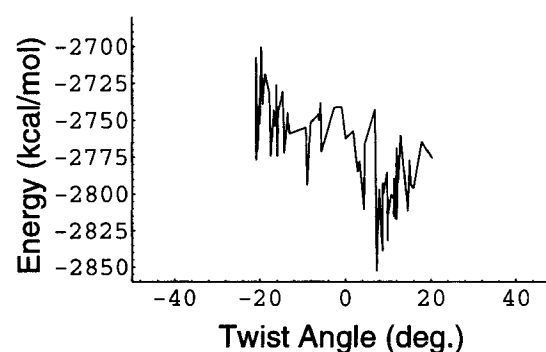## (e) p-LI (hypothetical dimer)

(f) p-IL (hypothetical tetramer)

Fig. 8.   Graphs of empirical potential energy vs. twist angle for the minimized simulated annealing structures (a–f). In all cases, the minimum energy cluster formed about a twist angle of about 10 degrees corresponding to left-handed coiled coil structures [see the text for a discussion of the low energy right-handed outlier in (c)].

simulations, these effects have been explicitly taken into account by the hydrogen bond distance restraints and in the initial placement of the helices.

In both of the predicted GCN4-p1 and p-LI structures, atoms outside the hydrophobic core have relatively high atomic rms deviations from the crystal structures (2.06 and 2.17 Å for GCN4-p1 and p-LI,

respectively). Because many of these side chains interact with symmetry related molecules in the crystal structure and are likely to have poorly defined conformations in solution, prediction of their structure is of little concern.

The geometric parameters of the predicted structures and the crystal structures are listed in Table I.
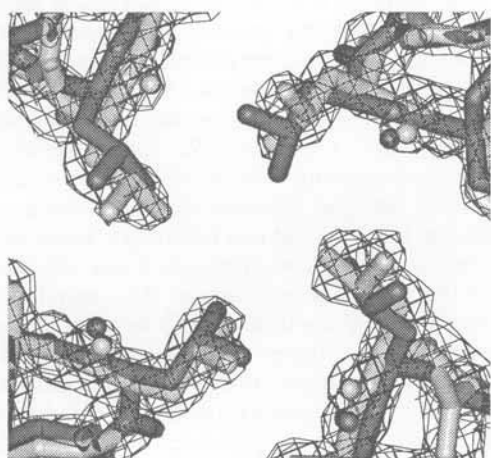
(a) GCN4-p1: ASN-16
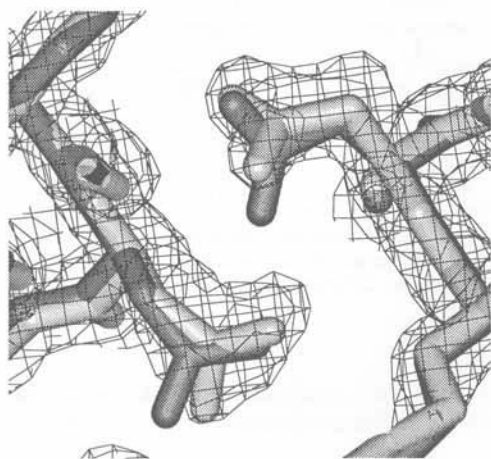
(b) GCN4-p1: LEU-19
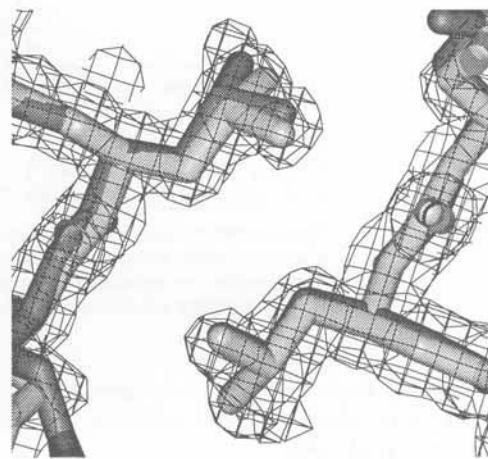
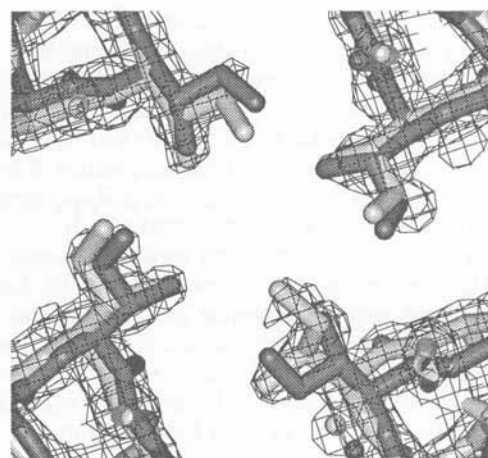## (c) p-LI: LEU-9

(d) p-LI: ILE-12

Fig. 9. Core residues in the crystal structures of GCN4-p1 and p-LI (dark gray), the predicted structures (light gray), and the probability maps (wireframe) contoured at a level of 2.5σ. Asparagine-16 side chains in the dimer (a) adopt asymmetric conformations in both the model structure and in the crystal structure. Leucine-19 side chains (b) and the other core side chains in the dimer adopt symmetric conformations in both structures. (p-LI), Good agreement with the crystal structure is seen for many side chains in the model tetramer (c,d).

While the helix separation in our models closely matches that of the crystal structures, our models are slightly underwound by 1.5–2°. Slight asymmetry is seen in both the crystal structures and our predicted models.

Prediction of the Unknown Structures: The p-IL Dimer and the p-II Trimer

The predicted p-IL dimer structure and the p-IL probability map are shown in Figure 10. As in the GCN4-p1 and p-LI models, many core side chains of p-IL have conformations that are well defined by the probability map, and most of the core residues adopt

symmetric or nearly symmetric conformations. The symmetric leucine side chains of p-IL adopt the same conformations as symmetric leucine side chains in the GCN4-p1 crystal structure. Many surface side chains do not have well-defined conformations.

The predicted p-II trimer structure contains significantly more asymmetry than do the other structures. The separation between pairs of helices is not uniform (distances are 9.98, 10.36, and 11.31 Å), and many asymmetric conformations are adopted by core side chains. As shown in Figure 11a and b, core side chain conformations are not well determined by

TABLE I. Geometric Parameters of the Refined Structures*

Peptide	Twist angle (deg.)	Helix separation (Å)	C^α asymmetry[†] (Å)
GCN4-p1 (dimer)	9.42 (11.06)	8.35 (8.46)	0.55 (0.75)
p-IL (dimer)	8.87	9.06	0.43
p-II (trimer)	10.97	10.55	1.36
p-II$_{SYM}$ (trimer)	11.16	10.56	0.03
p-LI (tetramer)	11.22 (13.15)	9.76 (9.84)	1.01 (0.91)

*Parameters of the crystal structures are shown in parentheses for comparison where available. Average distances are given for the trimers and for the tetramers.
[†]Asymmetry was measured by computing the average rms difference for C^α atoms after performing rotations about the coiled coil axis.

the probability map. Therefore, we expect the p-II prediction to be somewhat less accurate than the GCN4-p1 and p-LI models.

The predicted p-II$_{SYM}$ trimer structure possesses a nearly perfect axis of 3-fold symmetry since it was refined with explicit symmetry restraints. The symmetric probability map is much smoother than the asymmetric p-II map, but the density around core side chain atoms is still fairly ambiguous (Fig. 11c and d). As a result, core side chain conformations in the symmetric trimer structure are not any better defined than in the asymmetric trimer. However, since the coiled coil axis of the p-II crystal structure is likely to coincide with a crystallographic 3-fold axis,[16] the geometry of the predicted p-II$_{SYM}$ structure may be more accurate than that of p-II.

Pictures of the backbone and core side chains of the p-II$_{SYM}$ structure are shown in Figure 12. Layers of interacting isoleucine side chains can be identified, and staggering of these side chains between the layers is observed. The helices are slightly closer together at the N-terminal end of the coiled coil than at the C-terminal end. In each helix, all but one of the eight isoleucine residues adopt a conformation where the γ_2-carbon points inward toward the core and the γ_1-carbon points outward. This orientation may be the preferred mode of packing for isoleucine side chains in parallel trimeric coiled coils.

Interaction Energies

Figure 13 shows a thermodynamic cycle consisting of two hypothetical pathways for the conversion of two p-IL dimers into one p-LI tetramer. In one pathway, the dimers are first converted to a tetramer, and then the sequence is modified by exchanging the isoleucine and leucine residues. In the other pathway, the sequence is modified first, and then the dimers are converted to a tetramer. The conversion of two dimers into a tetramer (horizontal arrows in Fig. 13) depends heavily on changes in entropy and solvation energy, which are very difficult to compute. On the other hand, the interchanging of leucine and isoleucine residues (vertical arrows in Fig. 13) presumably affects mainly van der

Waals packing energies, which can be directly computed. The thermodynamic cycle implies

$$\Delta\Delta G = \Delta G^{IL}_{dimers \to tetramer} - \Delta G^{LI}_{dimers \to tetramer}$$
$$= \Delta G^{IL \to LI}_{dimers} - \Delta G^{IL \to IL}_{tetramer} \qquad (7)$$

In principle, calculation of $\Delta\Delta G$ could be accomplished through free energy perturbation methods. However, adequate sampling of conformational space during this rather extensive repacking of the hydrophobic core represents a formidable challenge and would be difficult to accomplish with current computational resources. We have therefore resorted to computing only the differences in average potential energies between the two dimeric coiled coils and between the two tetrameric coiled coils.

Since no solvent was present in our calculations, only the core packing energies (i.e., only the energies of interactions that directly involve uncharged core side chain atoms) can be relied upon. Fortunately, given that only the core side chain atoms are modified during the conversion of a p-IL dimer to a p-LI dimer or a p-IL tetramer to a p-LI tetramer, changes in the energies of interactions involving core atoms are likely to represent a significant portion of the free energy differences. The rest of the molecule remains chemically unchanged, so to a first approximation, we can assume that the interactions not involving the core side chain atoms will remain fairly constant. If we also assume that the total entropies and the total volumes of these systems will remain nearly constant since both the number of atoms and the residue types are conserved, the following approximations should hold:

$$\Delta G^{IL \to LI}_{dimers} \approx \Delta H^{IL \to LI}_{dimers} \approx \Delta E_{COREdimers}^{IL \to LI} \qquad (8)$$

$$\Delta G^{IL \to LI}_{tetramer} \approx \Delta H^{IL \to LI}_{tetramer} \approx \Delta E_{COREtetramer}^{IL \to LI} \qquad (9)$$

where ΔE_{CORE} is the change in interaction energies [Eq. (6)] of uncharged core side chain atoms (set U) and all atoms (set W). For the clusters of dimer and tetramer structures, we obtained the following differences between the mean potential energies:
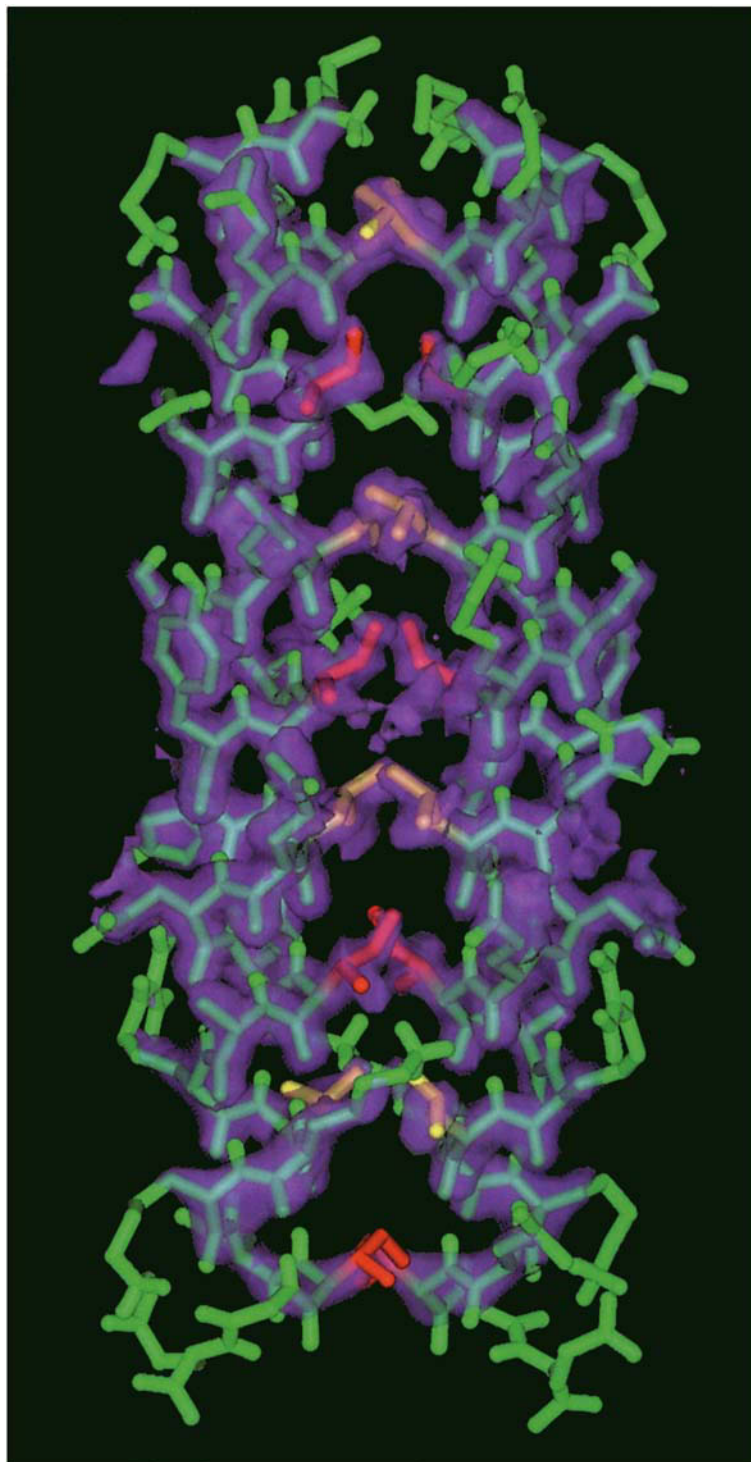
Fig. 10. The refined p-IL dimer structure in the probability map (purple) contoured at a level of 2.0σ. Side chains of isoleucines in the **a** positions are shown in red, and side chains of leucines at **d** positions are shown in yellow. All other atoms are shown in green. Six out of eight side chains adopt symmetric conformations.

$$\Delta E_{\text{COREdimers}}^{\text{IL}\rightarrow\text{LI}} = 27.0 \pm 5.1 \text{ kcal/mol} \qquad (10)$$

$$\Delta E_{\text{COREtetramer}}^{\text{IL}\rightarrow\text{LI}} = 1.3 \pm 2.0 \text{ kcal/mol} \qquad (11)$$

A large, statistically significant difference exists between the core interaction energies of the two dimer structures, and only a small and insignificant differ-
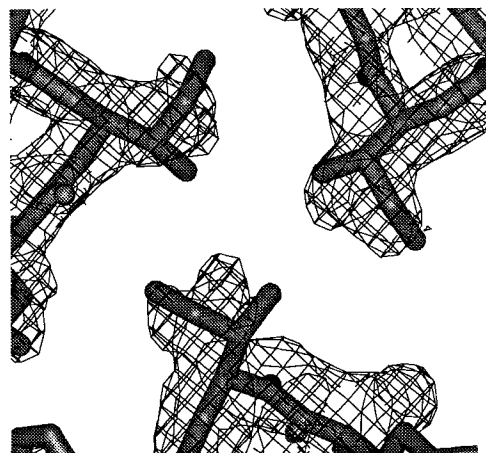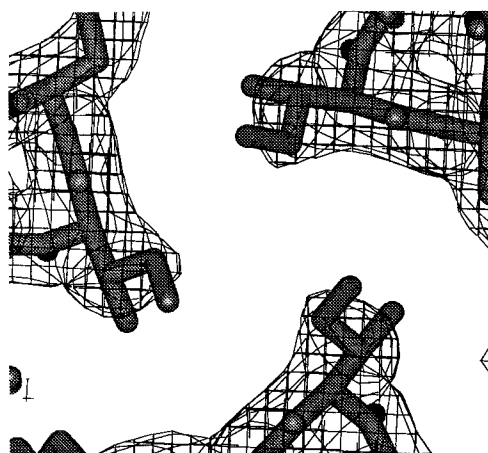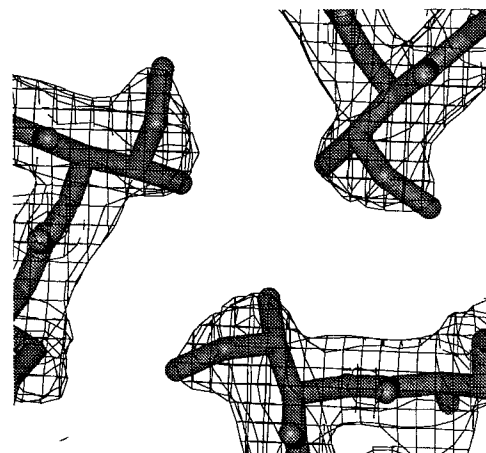
(a) p-II: ILE-16
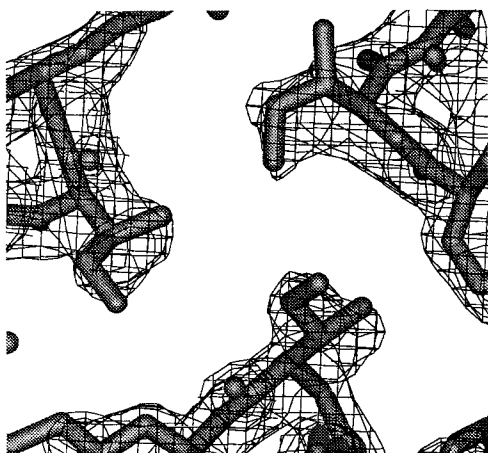
(b) p-II: ILE-19

(c) p-II$_{SYM}$: ILE-16

(d) p-II$_{SYM}$: ILE-19

Fig. 11. Core residues in the predicted p-II and p-II$_{SYM}$ trimer structures. The probability map is contoured at a level of 2.5σ. The map does not uniquely determine the conformations of most isoleucine side chains (a,b). Even after applying 3-fold averaging to generate the symmetric map, density around the side chains is still fairly ambiguous (c,d).

ence exists between the two tetramer structures. If we apply the approximations from (8) and (9) above, the free energies of the tetramer structures would be roughly equal, and the free energy difference between the dimers would strongly favor the p-IL sequence. This suggests that superior packing of the p-IL dimer as compared to p-LI dimer is the main cause of the switch in oligomerization states.

In order to qualitatively assess the validity of our assumption that the energies of interactions between atoms outside the core will remain constant, we computed the differences in these potential energies (the hydrogen bond restraint term contributed less than 0.3 kcal/mol to these differences).

$$\Delta E_{\text{NONCOREdimers}}^{\text{IL}\rightarrow\text{LI}} = -3.3 \pm 7.0 \text{ kcal/mol} \quad (12)$$

$$\Delta E_{\text{NONCOREtetramer}}^{\text{IL}\rightarrow\text{LI}} = -9.8 \pm 5.7 \text{ kcal/mol} \quad (13)$$

No statistically significant difference exists between the dimeric coiled coils, but a significant difference (1.7σ above zero) is found between the tetrameric coiled coils. Therefore, the assumption that interactions of noncore atoms would remain constant may not be valid for the tetrameric coiled coils. It is necessary to reiterate that these interaction energies of noncore atoms may not have physically relevant magnitudes since our model systems are incomplete. The significant difference only suggests that there
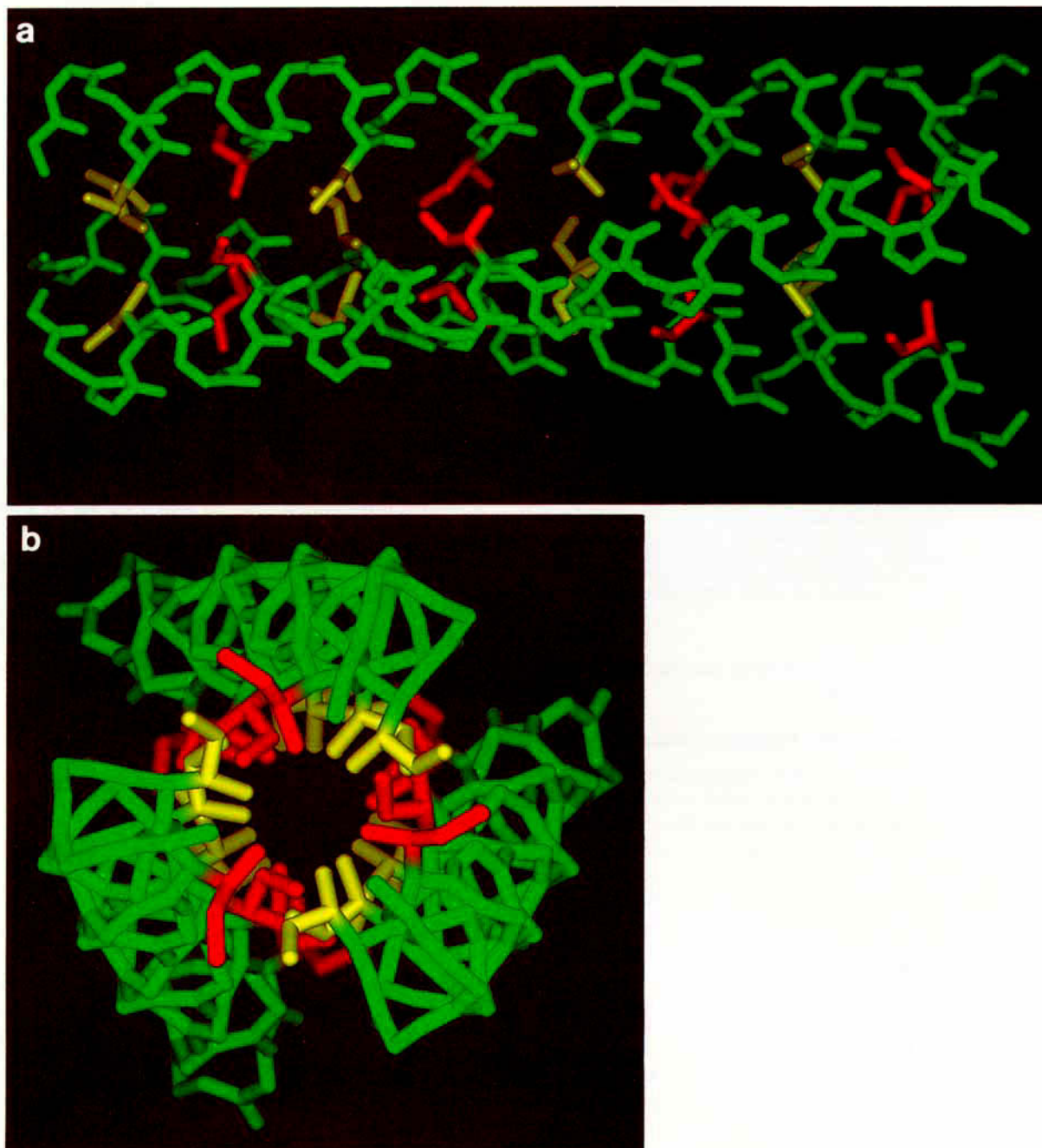
Fig. 12. Two views of the predicted p-II$_{SYM}$ structure. Side chains of isoleucines in the **a** positions are shown in red, side chains of isoleucines at **d** positions are shown in yellow, and backbone atoms are shown in green. Layers of interacting residues at the **a** and **d** positions can be seen (**a**), and staggering of side chains between layers is observable (**b**).

may be important changes in the energies of interactions that do not directly involve core atoms. Some possible causes for these differences might be changes in the packing energies of side chains that surround the core residues, changes in the total surface area of contact between the molecules, or disruption of electrostatic interactions between side chains on adjacent helices.

Allowing for the possibility of other energetic factors, what remains is a large difference in interaction energies involving core atoms between the dimeric coiled coils with no corresponding difference between the tetrameric coiled coils. Other differences may be important, but their existence cannot be confirmed or denied with current methodology. Extremely time-consuming simulations with ex-
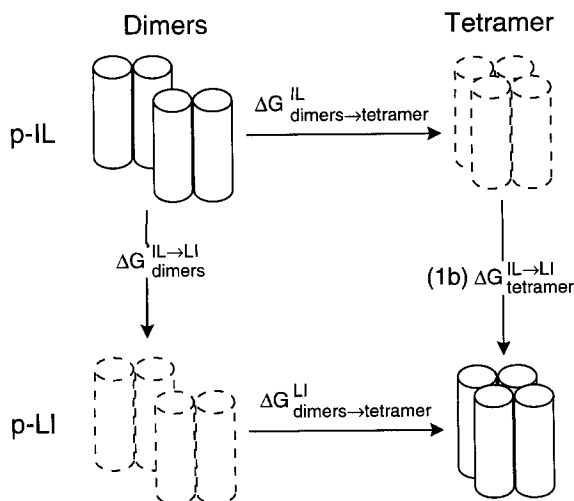
Fig. 13. A thermodynamic cycle showing two hypothetical pathways for the conversion of two p-IL dimers into a single p-LI tetramer. In pathway 1, the dimers are first converted to a tetramer, and then the sequence is switched. In pathway 2, first the sequence is switched, and then the dimers are converted to a tetramer. The observed coiled coils are drawn with solid lines, and the hypothetical intermediates are drawn with dashed lines.

plicit solvent molecules would be required in order to address this possibility.

Leucine and Isoleucine Packing Preferences

The large difference in core side chain interaction energies between the dimeric coiled coils is almost completely accounted for by van der Waals interactions [other interactions account for less than 0.4 kcal/mol of the differences in Eqs. (10) and (11)]. In order to understand why such a large difference exists between the two dimers and not between the two tetramers, we compute the differences in average interaction energies of core leucine and isoleucine side chains in the dimers and in the tetramers. These differences are then further broken down to show the relative importance of interactions with specific subsets of atoms (Tables II, III, IV, and V).

Harbury et al. classified the packing environments of dimeric and tetrameric coiled coils based on the angle formed between the $C^\alpha \rightarrow C^\beta$ vector of a side chain and the $C^\alpha \rightarrow C^\alpha$ vector on the adjacent helix (Fig. 14).[16] The **a** positions of dimers and **d** positions of tetramers have parallel geometry, and the **d** positions of dimers and the **a** positions of tetramers have perpendicular geometry. Since isoleucines occupied parallel positions and leucines occupied perpendicular positions in the observed p-IL dimer and in the observed p-LI tetramer, it was suggested that the switch in oligomerization states of the peptides might simply be due to a preference of leucine for perpendicular environments and of isoleucine for parallel packing environments. This proposal can now be evaluated against the interaction

TABLE II. Leucine Packing in Dimeric Coiled Coils (average interaction energy differences* in kcal/mol·residue)

Interacting atoms		d → a ΔE
Set U	Set W	p-IL → p-LI
	Backbone Helix A	0.54 ± .06
	Backbone Helix B	0.46 ± .09
Leu side chain Helix A	Leu side chain Helix B (same layer)	−0.22 ± .03
	Other side chains Helix A	0.06 ± .10
	Other side chains Helix B	0.33 ± .10
	All subsets	1.17 ± .20

*Interaction energies E_{UW} [Eq. (6)] were computed for side chains at the three innermost core positions of every structure in each cluster. These energies were then summed over all structures, and the total energy was divided by the number of side chains sampled. Differences between these mean energies were then computed. Error estimates for the differences were calculated from the distribution of energies present in each cluster.

Table III. Isoleucine Packing in Dimeric Coiled Coils (average interaction energy differences* in kcal/mol·residue)

Interacting atoms		a → d ΔE
Set U	Set W	p-IL → p-LI
	Backbone Helix A	−0.11 ± .07
	Backbone Helix B	0.51 ± .10
Ile side chain Helix A	Ile side chain Helix B (same layer)	−0.15 ± .05
	Other side chains Helix A	−0.08 ± .08
	Other side chains Helix B	0.43 ± .11
	All subsets	0.59 ± .15

*See Table II for a description of the averaging procedure.

energies of the leucine and isoleucine side chains in dimeric and tetrameric packing environments.

In the dimeric coiled coils, leucine exhibits a strong energetic preference (1.17 ± 0.20 kcal/mol·residue) for perpendicular packing at the **d** positions of p-IL over parallel packing at the **a** positions of p-LI (Table II). Isoleucine exhibits a moderate and reversed preference (0.59 ± 0.15 kcal/mol·residue) for parallel packing at the **a** positions of p-IL over perpendicular packing at the **d** positions of

TABLE IV. Leucine Packing in Tetrameric Coiled Coils (average interaction energy differences* in kcal/mol·residue)

Interacting atoms		$d \rightarrow a$ ΔE
Set U	Set W	p-IL $\rightarrow$ p-LI
Leu side chain Helix A	Backbone Helix A	$-0.11 \pm .04$
	Backbone Helices B, C, D	$-0.14 \pm .04$
	Leu side chain Helices B, C, D (same layer)	$0.12 \pm .05$
	Other side chains Helix A	$-0.22 \pm .05$
	Other side chains Helix B	$-0.02 \pm .05$
	All subsets	$-0.38 \pm .09$

*See Table II for a description of the averaging procedure.

TABLE V. Isoleucine Packing in Tetrameric Coiled Coils (average interaction energy differences* in kcal/mol·residue)

Interacting atoms		$a \rightarrow d$ ΔE
Set U	Set W	p-IL $\rightarrow$ p-LI
Ile side chain Helix A	Backbone Helix A	$-0.03 \pm .05$
	Backbone Helices B, C, D	$0.39 \pm .04$
	Ile side chain Helices B, C, D (same layer)	$-0.10 \pm .06$
	Other side chains Helix A	$0.03 \pm .04$
	Other side chains Helix B	$0.16 \pm .05$
	All subsets	$0.45 \pm .08$

*See Table II for a description of the averaging procedure.

p-LI (Table III). Since both residue types occupy their preferred environments in the p-IL dimer, the packing preferences combine to give the p-IL dimer a high relative stability over the p-LI dimer. These results qualitatively agree with in vitro studies of leucine and isoleucine stabilities in synthetic peptides,[19] and they mirror the high occurrence of leucine and the low occurrence of isoleucine in d positions of heptad repeat sequences.[17,18] However, the less dramatic relative occurrence of leucine in d positions over a positions (3.9% d vs. 3.2% a)[18] in a large number of dimeric coiled coil sequences suggests that a much weaker preference exists in general.

In the tetrameric coiled coils, leucine shows a weak preference (0.38 ± 0.09 kcal/mol·residue) for perpendicular packing at the a position of p-LI over parallel packing at the d position of p-IL (Table IV). Isoleucine exhibits a similar preference (0.45 ± 0.08 kcal/mol·residue) for perpendicular packing at the a position of p-IL over parallel packing at the d position of p-LI (Table V). Since both leucine and isoleucine energetically prefer perpendicular packing in the tetrameric coiled coils, only one residue type can occupy the preferred environment, and the other will have to occupy the disfavored environment. As a result, the packing preferences of isoleucine and leucine side chains effectively cancel between the p-IL and p-LI tetramers. These results contest the proposed model of constant isoleucine and leucine packing preferences in parallel and perpendicular packing environments that are independent of the oligomerization state.[16] In the dimeric coiled coils, leucine strongly prefers perpendicular packing and isoleucine moderately prefers parallel packing. However, both residues weakly prefer perpendicular packing in the tetrameric coiled coils. Apparently, the packing environments are not simply reversed between the dimeric and tetrameric coiled coils.

The breakdown of specific van der Waals interactions of core side chains shows that many separate interactions contribute to the relative stability of leucine and isoleucine side chains in a particular environment. In all cases, a large part of the difference arises from interactions between side chain atoms and backbone atoms, but interactions between side chain atoms are important as well. Surprisingly, interactions between corresponding leucine and isoleucine side chains in a particular layer are not a major determinant of the packing preferences (see the third row of Tables II, III, IV, and V). In fact, interactions between these atoms are consistently worse in the favored packing environment than in the disfavored environment. In contrast, interactions with other side chain atoms are consistently better in the preferred environment. It is also apparent that van der Waals interactions between side chain atoms and atoms on the same helix can be as important as are interactions with atoms on other helices. Since the energetics of side chain packing are dependent on a number of complicated and variable interactions, packing models such as the parallel and perpendicular packing classification scheme appear somewhat oversimplified.

CONCLUSIONS

We have shown that simulated annealing can be highly effective at predicting the structures of coiled coil proteins when combined with sufficient experimental data to narrow the scope of the conformational search. In our approach, unknown conformational parameters are explicitly sampled with a large number of initial models. During simulated
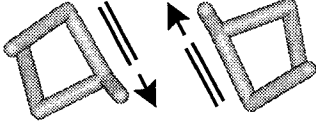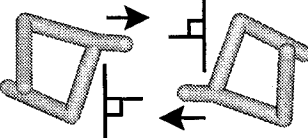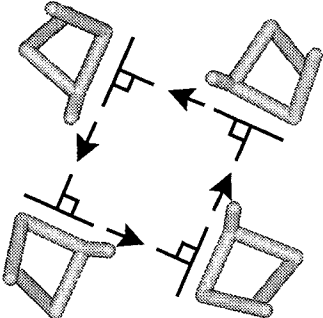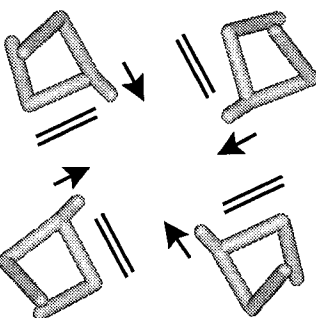
Fig. 14.   Packing environments of **a** and **d** positions of dimeric and tetrameric coiled coils shown with C^α and C^β traces. The packing environments are classified as parallel or perpendicular as described by Harbury et al.[16]

annealing, the models relax into their low energy conformations, and then cluster analysis and probability map refinement are used to produce a single representative structure. Models of the GCN4-p1 and p-LI peptides have been generated with unprecedented accuracy, and structures of the dimeric p-IL and trimeric p-II peptides have been predicted without knowledge of the crystal structures.

In the dimeric coiled coils, leucine and isoleucine side chains exhibit strong packing preferences which, in combination, stabilize the p-IL dimer relative to the p-LI dimer. In the tetrameric coiled coils, isoleucine and leucine show weak packing preferences which cancel to make core packing energetics roughly equal in the p-IL and p-LI tetramers. In light of these results, side chain packing discrimination at core positions in the dimeric coiled coil appears to be a major factor behind the switch in oligomerization preferences of the p-IL and p-LI GCN4 leucine zipper mutants.

The packing environments of dimeric and tet-rameric coiled coils are not simply reversed. Each core heptad position in the dimeric and tetrameric coiled coils appears to have a unique profile of interactions within and between layers of the coiled coil. A breakdown of the packing energetics of isoleucine and leucine residues in the coiled coils demonstrates the cooperative nature of packing interactions.

REFERENCES

1. Hinds, D.A., Levitt, M. A lattice model for protein structure prediction at low resolution. Proc. Natl. Acad. Sci. U.S.A. 89:2536–2540, 1992.
2. Skolnick, J., Kolinski, A., Brooks, C.L., Godzik, A., Rey, A. A method for predicting protein structure from sequence. Curr. Biol. 3:414–423, 1993.

3. Godzik, A., Kolinski A., Skolnick, J. De novo and inverse folding predictions of protein structure and dynamics. J. Comp. Aided Mol. Design 7:397–438, 1993.

4. Vieth, M., Kolinski, A., Brooks, C.L., Skolnick, J. Prediction of the folding pathways and structure of the GCN4 leucine zipper. J. Mol. Biol. 237:361–367, 1994.

5. Conway, J.F., Parry, D.A.D. Structural features in the heptad substructure and longer range repeats of two-stranded alpha-fibrous proteins. Int. J. Biol. Macromol. 12: 328–334, 1990.

6. Chou, P.Y., Fasman, G.D. Secondary structural prediction of proteins from their amino acid sequence. TIBS 2:128–131, 1977.

7. Kneller, D.G., Cohen, F.E., Langridge, R. Improvements in protein secondary structure prediction by an enhanced neural network. J. Mol. Biol. 214:171–182, 1990.

8. Schiffer, C.A., Caldwell, J.W., Kollman, P.A., Stroud, R.M. Prediction of homologous protein structures based on conformational searches and energetics. Proteins 8:30–43, 1990.

9. Brünger, A.T., Kuriyan, J., Karplus, M. Crystallographic R factor refinement by molecular dynamics. Science 235: 458–460, 1987.

10. Lemmon, M.A., Flanaga, J.M., Treutlin, H.R., Zhang, J., Engelman, D.M. Sequence specificity in the dimerization of transmembrane alpha-helices. Biochemistry 31:12719–12725, 1992.

11. Treutlein, H.R., Lemmon, M.A., Engelman, D.M., Brünger, A.T. The glycophorin A transmembrane domain dimer: Sequence-specific propensity for a right-handed supercoil of helices. Biochemistry 31:12726–12733, 1992.

12. Chen, Y.H., Yang, J.T., Chau, K.H. Determination of the helix and β form of proteins in aqueous solution by circular dichroism. Biochemistry 13:3350–3359, 1974.

13. Tycko, R., Smith, S.O. Symmetry principals in the design of pulse sequences for structural measurements in magic angle spinning nuclear magnetic resonance. J. Chem. Phys. 98:932–943, 1993.

14. Lemmon, M.A., Treutlin, H.R., Adams, P.D., Brünger, A.T., Engelman, D.M. A dimerization motif for transmembrane α-helices. Nature Stru. Biol. 1:157–163, 1994.

15. O'Shea, E.K., Klemm, J.D., Kim, P.S., Alber, T. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science 254:539–544, 1991.

16. Harbury, P.B., Zhang, T., Kim, P.S., Alber, T. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. Science 292:1401–1407, 1993.

17. Parry, D.A.D. Coiled-coils in α-helix-containing proteins: Analysis of the residue types within the heptad repeat and the use of these data in the prediction of coiled-coils in other proteins. Biosci. Rep. 2:1017–1024, 1982.

18. Lupas, A., van Dyke, M., Stock, J. Predicting coiled coils from protein sequences. Science 252:1162–1164, 1991.

19. Zhu, B.Y., Zhou N.E., Kay, C.M., Hodges, R.S. 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. Prot. Sci. 2:383–394, 1993.

20. Brünger, A.T. X-PLOR, Version 3.1. A System for X-ray Crystallography and NMR. New Haven: Yale University Press, 1992.

21. O'Shea, E.K., Rutkowski, R., Kim, P.S. Evidence that the leucine zipper is a coiled coil. Science 243:538–542, 1989.

22. Chothia, C., Levitt, M., Richardson, D. Helix to helix packing in proteins. J. Mol. Biol. 145:215–250, 1981.

23. Nilges, M., Clore, M., Gronenborn, A.M. Determination of three-dimensional structure of proteins from interproton distance data by dynamical simulated annealing from a random array of atoms. FEBS Lett. 239:129–136, 1988.

24. Brooks, B.R., Bruccoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S., Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. J. Comp. Chem. 4:187–217, 1983.

25. Loncharich, R.J., Brooks, B.R. The effects of truncating long-range forces on protein dynamics. Proteins 6:32–45, 1989.

26. Brünger, A.T. A memory-efficient fast fourier transformation algorithm for crystallographic refinement on supercomputers. Acta Cryst A45:42–50, 1989.

27. Brooks, C.L., Karplus, M., Pettitt, B.M. "Proteins: A Theoretical Perspective of Dynamics, Structure, and Thermodynamics." New York: John Wiley, 1988.

28. Arnold, E., Rossmann, M.G. The use of molecular-replacement phases for the refinement of the human rhinovirus 14 structure. Acta Cryst. A44:270–282, 1988.

29. Brünger, A.T., Karplus, M., Petsko, G.A. Crystallographic refinement by simulated annealing: Application to crambin. Acta Cryst. A45:50–61, 1989.

30. Brünger, A.T., Krukowski, A. Slow-cooling for crystallographic refinement by simulated annealing. Acta Cryst. A46:585–593, 1990.

31. Tuffery, P., Etchebest, C., Haxout, S., Lavery, R. A new approach to the rapid determination of protein side chain conformations. J. Biomol. Struc. Dynam. 8:1267–1289, 1991.

32. Hendrickson, W.A. Stereochemically restrained refinement of macromolecular structures. Meth. Enzymol. 115: 252–271, 1985.

33. Crick, F.H.C. The packing of α-helices: simple coiled-coils. Acta Cryst. 6:689–697, 1953.

34. Wesson, L., Eisenberg, D. Atomic solvation parameters applied to molecular dynamics of proteins in solution. Protein Sci. 1:227–235, 1992.

35. Simonson, T., Brünger, A.T. Solvation free energies estimated from macroscopic continuum theory: an accuracy assessment. J. Phys. Chem. (in press) 1994.

36. Harris, N.L., Presnell, S.R., Cohen, F.E. Four helix bundle diversity in globular proteins. J. Mol. Biol. 236:1357–1368, 1994.

37. Nilges, M., Brünger, A.T. Successful prediction of the coiled coil geometry of the GCN4 leucine zipper domain by simulated annealing: comparison to the x-ray structure. Proteins 15:133–146, 1993.

38. Nilges, M., Brünger, A.T. Automated modeling of coiled coils: application to the GCN4 dimerization region. Protein Eng. 4:649–659, 1991.