

COMMUNICATION

## Prediction of the Folding Pathways and Structure of the GCN4 Leucine Zipper

Michal Vieth<sup>1</sup>, Andrzej Kolinski<sup>1,2</sup>, Charles L. Brooks III<sup>1,3</sup> and Jeffrey Skolnick<sup>1†</sup>

<sup>1</sup>*Departments of Molecular Biology and Chemistry  
The Scripps Research Institute, 10666 North Torrey Pines Road  
La Jolla, CA 92037, U.S.A.*

<sup>2</sup>*Department of Chemistry, University of Warsaw  
02-093 Warsaw, Poland*

<sup>3</sup>*Department of Chemistry, Carnegie Mellon University  
Pittsburgh, PA 15213, U.S.A.*

A hierarchical approach is described for the prediction of the three-dimensional structure and folding pathway of the GCN4 leucine zipper. Dimer assembly is simulated by Monte Carlo dynamics. The resulting lowest energy structures undergo cooperative rearrangement of their hydrophobic core leading to side-chain fixation. The coarse-grained structures are further refined using a molecular dynamics annealing protocol. This produces full atom models with a backbone root-mean-square deviation from the crystal structure of 0.81 Å. Thus, we demonstrate the predictive ability of our approach to yield high resolution structures of small coiled coils from their sequence.

*Keywords:* GCN4 leucine zipper; *de novo* structure prediction; quaternary structure prediction; folding pathways; coiled coils

The coiled coil motif (Crick, 1953), consisting of two helices wrapped around each other, is exhibited by many biologically important proteins (Cohen & Parry, 1990; Cohen & Parry, 1986). Coiled coils are found in muscle (Phillips *et al.*, 1986), cell frameworks, bacterial cell walls, and in DNA (Ferred'Amare *et al.*, 1993; O'Shea *et al.*, 1991) and RNA binding proteins (Banner *et al.*, 1987; Cohen & Parry, 1986). Coiled coils were also the object of a number of *de novo*, synthetic design studies (Hodges *et al.*, 1981; Lovejoy *et al.*, 1993). Because of their structural simplicity, coiled coils are ideal for testing theoretical approaches to some aspects of the protein folding problem (Creighton, 1990; Levitt, 1991; Ptitsyn, 1987), i.e. understanding the rules by which a sequence folds to its three-dimensional structure. In this spirit, this communication describes the prediction of the folding pathways, including dimer assembly, and the atomic level, three-dimensional structure of a leucine zipper fragment of the transcriptional activator, GCN4 (Ellenberger *et al.*, 1992; O'Shea *et al.*, 1991). There are several reasons why this molecule was chosen for our initial study. It is important for the dimerization and DNA binding of the entire GCN4

molecule, each polypeptide chain is short (33 residues) and a high resolution crystal structure of this dimeric coiled coil has been determined (O'Shea *et al.*, 1991). Finally, in recent work by Nilges and Brünger (1991, 1993), a molecular dynamics-based protocol was used to predict the structure of the GCN4 "leucine zipper", even before the crystal structure was available. In this study by Nilges and Brünger, an initial model consisting of an idealized, parallel, coiled coil geometry for the leucine zipper was assumed. This model, with an initial 3.1 Å backbone root-mean-square (r.m.s.‡) deviation from the crystallographically determined structure, was "refined" to a level of 1.26 Å backbone r.m.s. deviation (1.75 Å for all heavy-atoms in the dimer interface) from the crystal structure. Their protocol relied on the imposition of helical backbone hydrogen bond restraints and a number of inter-helical distance restraints together with molecular dynamics "annealing" to bring the initial model to the level of agreement with the crystal structure noted above. These simulations provide a benchmark for the comparison of the accuracy of our hierarchical approach described below.

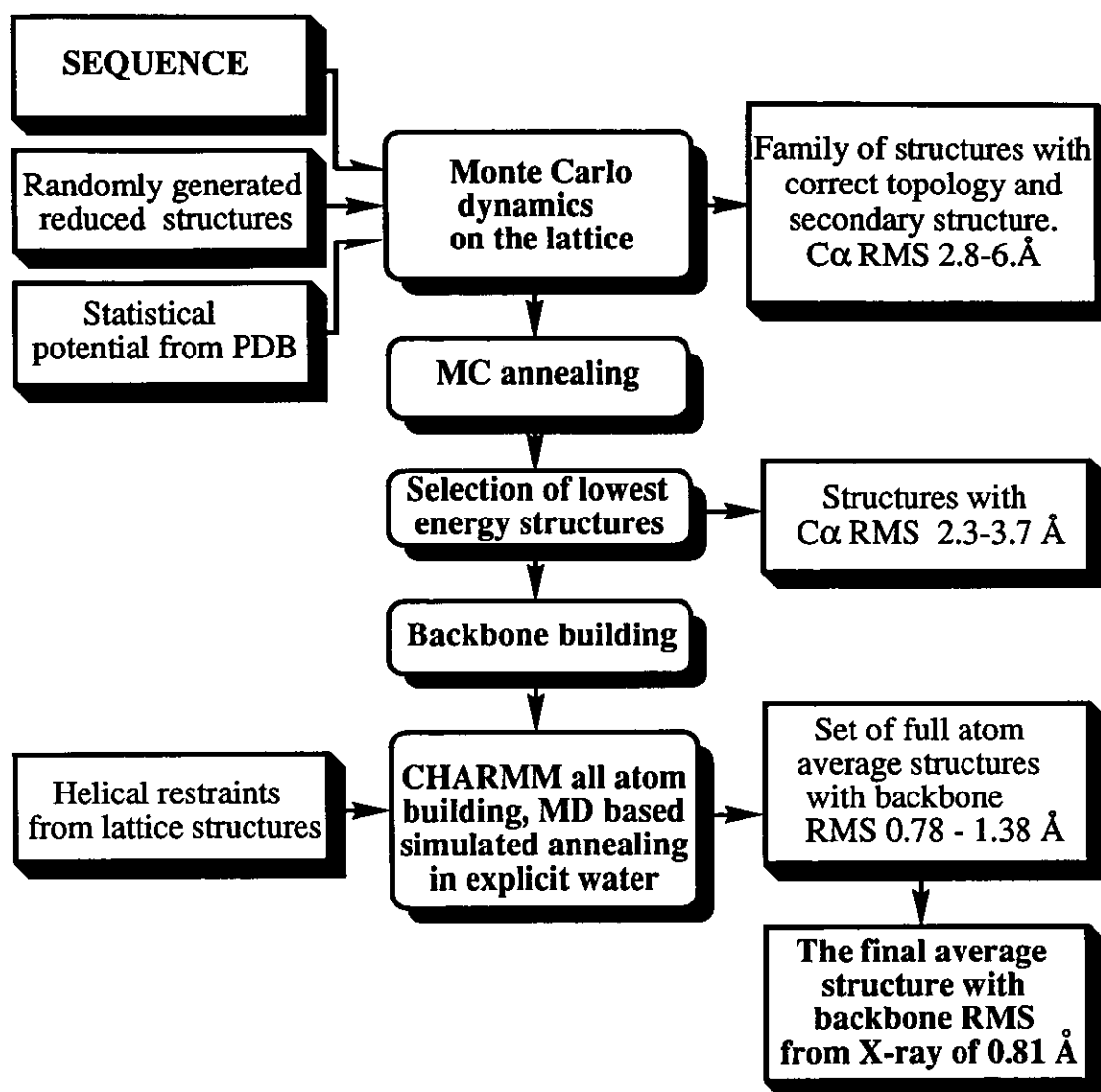
Since the time scale for protein folding is rather

† Author to whom all correspondence should be addressed.

‡ Abbreviations used: r.m.s., root-mean-square; MD, molecular dynamics.

long (ranging from  $10^{-3}$  s to  $10^2$  s), detailed molecular dynamics or Brownian dynamics algorithms cannot simulate the entire process (Jernigan, 1992; Levitt, 1991). Thus, we have developed a hierarchical approach to the protein folding problem (Kolinski & Skolnick, 1993; Skolnick *et al.*, 1993). The basic idea is to use a reduced protein model to simulate the early stages of folding and to obtain the native state topology. Restrained molecular dynamics simulations are then used to produce full atom models (Skolnick *et al.*, 1993). We have recently illustrated the success of this methodology applied to the simulation of the folding pathways and structural prediction of a number of globular proteins having simple native state topologies (Kolinski & Skolnick, 1993;

Skolnick *et al.*, 1993). The r.m.s. deviation from native for the resulting structures was in the range of 2.25 to 3.6 Å for the backbone atoms. The only information specific for a given protein in those cases (as well as in the present one) is its amino acid sequence. Here, the hierarchical approach to protein folding is extended to study the assembly of dimeric coiled coils. In addition, the full atom rebuilding protocol is improved and extended by including simulated annealing in explicit water solvation shells (Guenot & Kollman, 1992) around the structure undergoing refinement by molecular dynamics. This is done in a similar spirit to that of Nilges & Brünger (1991, 1993). An overview of the entire protocol is presented in Figure 1. This report demonstrates that the combination of the lattice



**Figure 1.** An overview of the hierarchical approach. The left-hand column shows the input information, the central column shows the protocol followed, and the right-hand column shows the resulting output. Folding of the model chains commences on the lattice from a pair of random, unfolded chains. Following dimer assembly to a parallel, left-handed, in register, coiled coil, the resulting structures are refined on the lattice to produce a family of structures whose r.m.s. from native ranges from 2.3 to 3.7 Å. Then, full atom models are built, the structures solvated and refined using molecular dynamics with the CHARMM potential. The resulting family of structures is indistinguishable from the native structure subject to the molecular dynamics refinement protocol. The average structure from this family has a backbone and heavy atom r.m.s. of 0.81 and 2.29 Å from native, respectively.

methodology with molecular dynamics refinement can give coiled coil structures whose quality is close to those obtained from experiment.

The lattice model is based on an  $\alpha$ -carbon representation (Godzik *et al.*, 1993a; Kolinski & Skolnick, 1992, 1993; Skolnick *et al.*, 1993) of the polypeptide backbone and a multiple rotamer, single sphere representation of the side groups. Consecutive  $\alpha$ -carbons, positioned on a lattice (Skolnick *et al.*, 1993) with a spacing between lattice points of 1.22 Å, are connected by a set of vectors of the type  $1.22*(3,1,1)$ ,  $1.22*(3,1,0)$ ,  $1.22*(3,0,0)$ ,  $1.22*(2,2,1)$  and  $1.22*(2,2,0)$ . This high resolution lattice represents the backbone protein geometry (Godzik *et al.*, 1993b) and permits better side-chain packing (Skolnick *et al.*, 1993) than in all previous lattice models.

The interaction scheme used in the lattice folding is a statistical potential extracted from a library of high resolution structures of globular proteins as deposited in the Brookhaven Protein Data Bank, (PDB: Bernstein *et al.*, 1977; PDB, 1992). This is described in full detail elsewhere (Kolinski & Skolnick, 1993; Skolnick *et al.*, 1993). However, for completeness we note the salient features of our potential. First, there are residue-independent interactions: an effective Ramachandran potential which acts to reproduce the average local distribution of distances; chirality and local excluded volume between  $\alpha$ -carbons. In addition, there are "effective" hydrogen bonds between main-chain  $\alpha$ -carbons. Each  $\alpha$ -carbon can participate in at most two such interactions. Proline is an exception; it can participate in only one such interaction. If residue  $i$  is hydrogen-bonded to residue  $j$  and residues  $i \pm 1$  and  $j \pm 1$  are also hydrogen-bonded, then this pair of hydrogen bonds is additionally stabilized by an explicit cooperative contribution. There is no directionality (donor-acceptor) in this scheme.

Sequence-specific, local, side group orientational correlations act between the first through fourth neighboring residues along the sequence. They reproduce the distribution found in a structural database. In addition, there is a rotamer energy which reflects the population of the particular side group rotamer in the database.

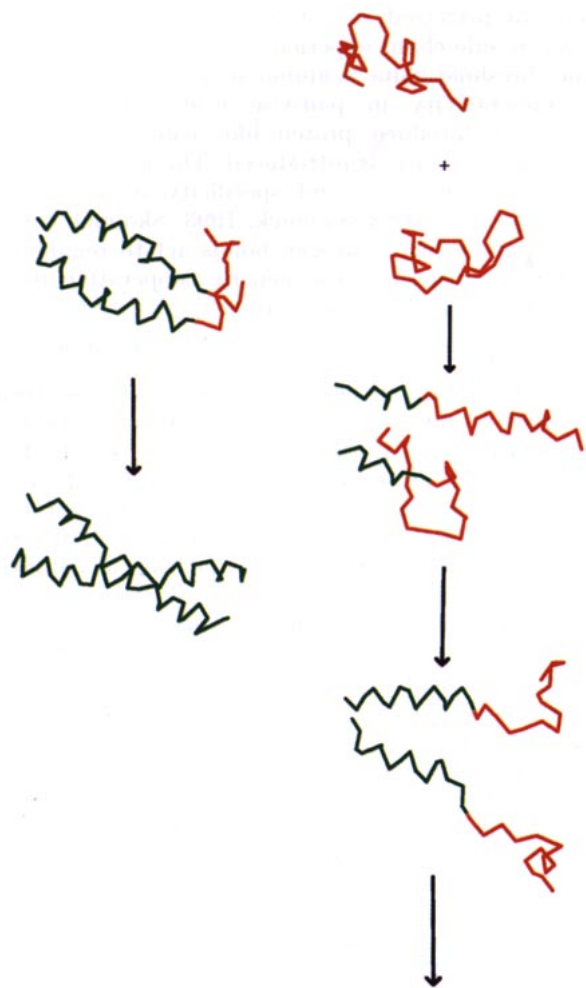
Tertiary and quaternary interactions are defined by a potential of mean force based on the statistics of interresidue contacts in our databases of protein structures from the PDB (which by construction does not contain any coiled coils). (This scale is available by anonymous ftp: Skolnick & Kolinski, 1993.) These interactions consist of a number of terms. First, there is an amino acid pair specific potential,  $E_{ij}$ , between interacting side-chains  $i$  and  $j$ . Both attractive and repulsive components are included; however, we found that the distances of hard core repulsion of the side-chains (Kolinski & Skolnick, 1993) are crucial in providing the correct, left-handed topology of coiled coils. It is interesting to note that the use of smaller values for the excluded volume of the side-chains led to incorrect, right-handed topology coiled coils which are ener-

getically preferred. Continuing, there is a penalty if a given side-chain experiences too many contacts. The threshold value is amino acid-specific. Further, a cooperativity in pairwise interactions, which serves to produce protein-like contact patterns when appropriate, is introduced. These interactions mimic cooperativity and specificity in side-chain packing (Kolinski & Skolnick, 1993; Skolnick *et al.*, 1993), much like hydrogen bonds act to regularize backbone structure. The generic cooperative pairwise interactions are defined by:

$$E_{ij,template} = (E_{ij} + E_{i \pm n, j \pm n}) C_{ij} C_{i \pm n, j \pm n} \quad n = 3, 4, \quad (1)$$

where  $C_{i,j} = 1$  (0) if there is (not) a contact between residues  $i$  and  $j$ . These "templates" enhance the cooperativity for protein-like interactions in both  $\alpha$  and  $\beta$  types of secondary structure. Based on a detailed series of studies, we have demonstrated that these templates do not demand that the resulting native structure be helical; a detailed discussion regarding the effects of templates on folding is discussed elsewhere (Kolinski & Skolnick, 1993). Finally, to keep the pair of chains in the neighborhood of each other, there is an harmonic potential between the centers of masses of the chains, which is only active for distances larger than 45 Å. The effective concentration of chains is in the range of  $10^{-2}$  M, but only two chains are considered at this time. Thus, we only permit the possibility of dimer formation. Studies of the equilibria between dimers and higher-order multimers are currently underway for both wild-type and mutants. Preliminary results suggest that the wild-type dimer is preferred. Experimentally, formation of higher-order multimers is possible with changes in monomer concentration or mutations in the sequence (Harbury *et al.*, 1993).

The folding simulations on the lattice start from randomly generated conformations at random relative positions of the two chains. The system is then subjected to a slow annealing protocol beginning at temperatures above a critical "transition" temperature. A representative folding trajectory is depicted in Figure 2. Dimer formation starts from the collision of short helical stretches, typically at the chain ends. Interacting helical stretches propagate along the molecule, and after small adjustments in registration, the final parallel, in register, coiled coil, helical dimer is formed. The end point, low energy structures often contain small defects in the helical conformation, mostly in the middle of the molecule. This is consistent with our preliminary studies which indicate that there is a minimum in the predicted helix content in the vicinity of Asn16 and reflects the sensitivity of our models to such sequence-specific effects. In our model, the dimer energy in the N16V mutant is  $16 k_B T$  lower than that of the wild-type. This is to be compared to the experimental value of  $11.67 k_B T$  (Alber, 1992). This points out the destabilizing roles of Asn16 in our model that are consistent with experimental findings (Alber, 1992). In addition to parallel dimers, antiparallel dimers are sometimes



**Figure 2.** A representative folding pathway of the GCN4 leucine zipper. The folded portions of the molecule are green, the unfolded are red. The sequence of each chain is: RMKQLEDKVEELLSKNYHLENEVARLK KLVGER. The underlined residues are located in the dimerization interface.

observed. The energy of these structures is usually 20% higher ( $40 k_B T$ ) than the energy of a representative parallel dimer. Thus, they can be safely dismissed from consideration.

The helix content of the non-interacting chains is about 30 to 35%. This increases to 90% upon the assembly of the coiled coil structure. In the folding experiments described here, the correct topology was recovered in nine of the 14 independent folding simulations; six of these have almost ideal secondary structure. In the remaining five cases, the lowest energy structures are either antiparallel or hairpin-like. The lowest energy of the misfolded structures is about  $35 k_B T$  (about 17%) higher than the energy of the correctly folded structures. Long range interactions (48%), orientational coupling between neighboring side groups (35%) and hydrogen bonds (15%) make the dominant contributions to the total energy. At this stage of the folding protocol, the r.m.s. deviation of the  $C^\alpha$  trace of the

predicted structures is on the level of 2.8 to 6.0 Å from the native crystal structure. The average overlap of the interchain side group contact map with that of the native molecule is about 40 to 60% (random overlap is about 20 to 30%; 60% indicates closely homologous structures (A. Godzik & J. Skolnick, unpublished results)). The lowest energy structures from each of the 14 simulations are then subjected to a cooling protocol. Each time the temperature is lowered, the lowest energy structure from the previous (higher temperature) run is chosen as the starting conformation for subsequent refinement. At the end of this refinement stage, we found five structures whose average and minimal energies were substantially lower than all the others (20% lower than those having the next lowest energy).

The five lowest energy lattice structures differed from the  $C^\alpha$  trace of the crystal structure of the GCN4 leucine zipper by 2.3 to 3.7 Å r.m.s. deviation. The interchain side-chain contact map overlapped 58 to 60% with that of the native molecule (based on a 5 Å heavy-atom distance cutoff), and the average distance between the helix axes is 9.5 to 10.5 Å, with a crossing angle of 25°. This is to be compared to the crystal structure, where the average distance between helical axes is 9.3 Å and the crossing angle is 18°. Starting from the  $C^\alpha$  traces of the five lowest energy structures, all atom backbones were built using SYBYL (1992). Side-chain atoms were then constructed for the first 31 residues from each chain using automated atom building procedures in CHARMM (Brooks *et al.*, 1983). The structures were minimized and surrounded by a 6 Å thick water shell (Guenot & Kollman, 1992). All minimization and dynamics procedures used the CHARMM software with the polar hydrogens, PARAM19 parameters (Brooks *et al.*, 1983) and the TIP3P water model (Jorgensen *et al.*, 1983). Each of these structures was further minimized in the presence of water. Then, the resulting minimized structures were subjected to a heating-cooling molecular dynamics refinement protocol, which employed the helical dihedral angle restraints of 20 kcal/mol rad<sup>2</sup>. These restraints were based on the predicted helical regions from the lattice simulations. A non-bonded cutoff of 10 Å was used with a distance dependent dielectric constant (Guenot & Kollman, 1993). Periodic boundary conditions were used to prevent evaporation of the water sheath but were not meant to reproduce a complete solvent environment. The solvent sheath is included to reproduce some salient features of the hydrophobic effect and hydrophilic solvation.

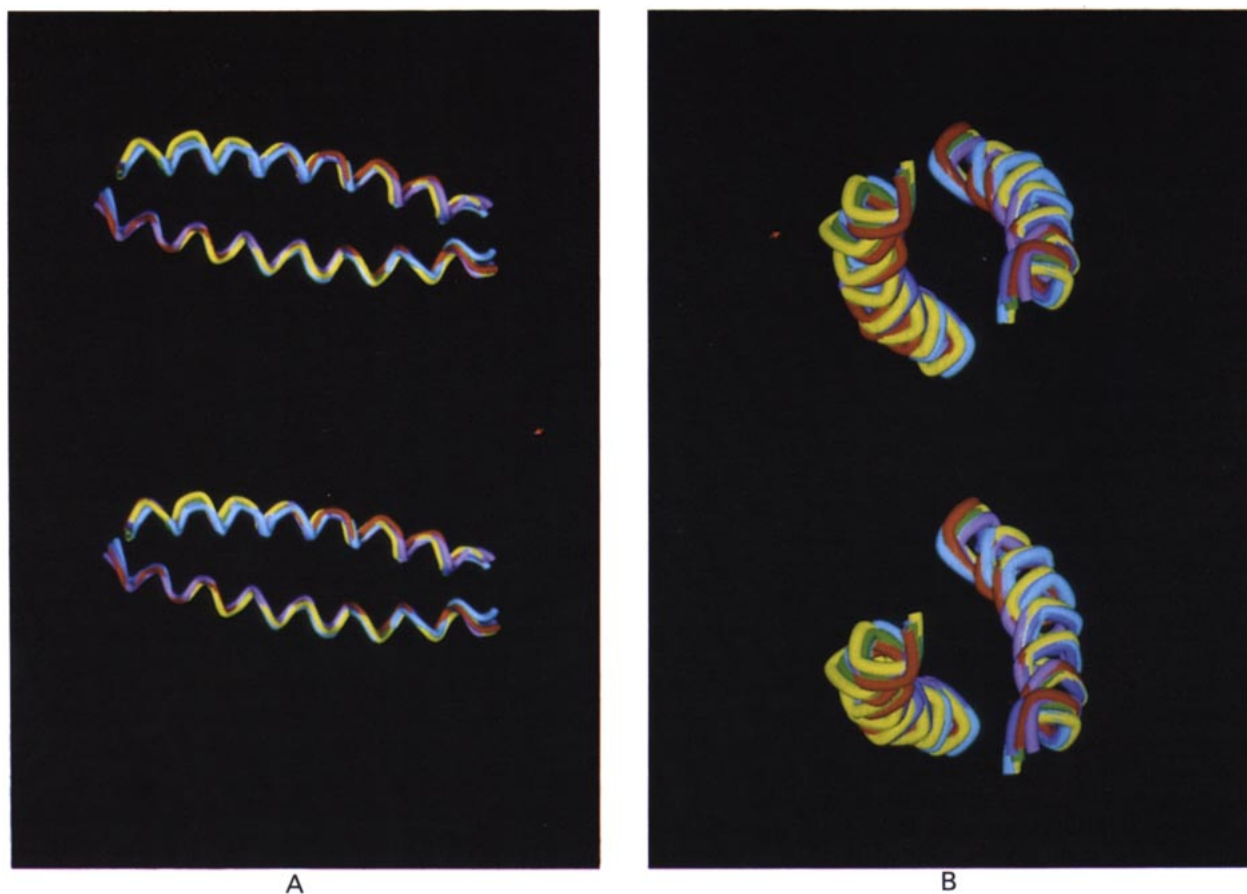
For each of the five structures derived from independent lattice simulations, ten independent solvent refinement runs were performed. Within each family of refined structures, a mean structure was computed for the protein. This average structure was minimized, surrounded by a water shell, and subjected to the identical heating-cooling MD protocol. This process was repeated until the side-chain contact maps and the backbone geometry of

the average structures found in successive iterations converged. Convergence was reached when the relative difference between subsequent backbone r.m.s. values was smaller than 10%. Depending on the quality of the starting structure, the requisite number of iterations varied between five and eight. The mean backbone r.m.s. deviation between pairs of the resulting average structures was 0.93 Å, and was on average 1.0 Å from the crystal structure. When compared to the PDB structure, the overlap of the interchain side-chain contact maps ranges from 82 to 96%.

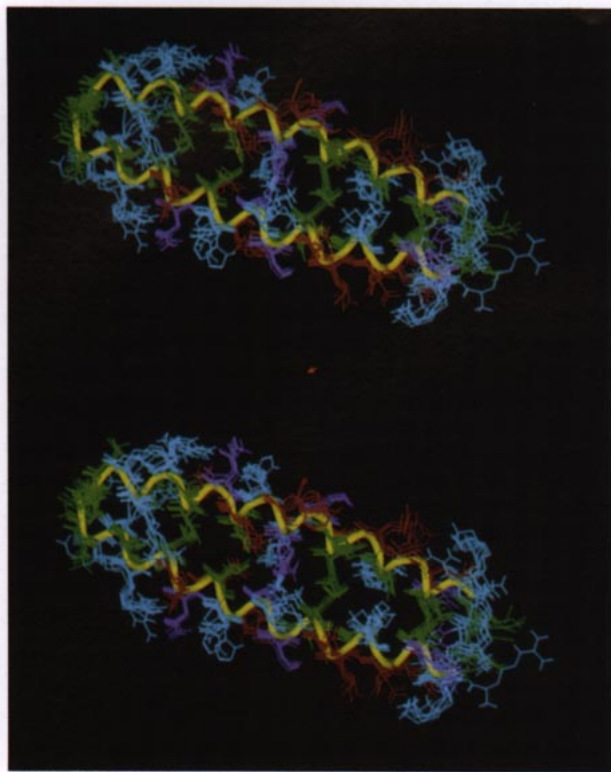
The average structure built from the entire family of refined structures has an r.m.s. deviation of 0.81 Å, 1.31 Å and 2.29 Å from the crystal structure for the backbone, heavy-atoms in the dimerization interface, and all heavy-atoms, respectively. The side-chain contact map overlap of this average structure is 96%, when compared to the PDB crystal structure. The final distance between the helical axis was 9.34 Å, with a crossing angle of 18°, i.e. virtually identical to the experimental values. To further establish the extent of agreement of the predicted structures with experiment, the crystal structure (O'Shea *et al.*, 1991) was subjected to the identical MD protocol. The resulting structure is indistinguishable from other members of the family

of refined lattice structures. Thus, based on r.m.s. and side-chain contact overlap criteria, both the refined structures and the experimentally determined structure are essentially indistinguishable. This conclusion is confirmed in Figure 3 where tube diagrams of the backbones of the refined structures are shown along with the crystal structure.

The family of refined structures exhibits many interesting features. All atom models of these structures are presented in Figure 4. Clearly, the core is uniquely packed, whereas there is higher variability in the side-chain conformations of the exposed residues as well as higher mobility in the chain ends. A number of interchain salt bridges between residues is also observed. In agreement with the crystal structure, in one out of five structures, the side-chain of Asn16 of one monomer asymmetrically binds to the side-chain of Asn16. In addition, interchain hydrogen bonds between the side-chain of Asn16 and Glu20 are predicted as well as hydrogen bonds between Asn16 and surrounding water. The variety of electrostatic interactions experienced by Asn16 perhaps implies the existence of fluctuating hydrogen bonds for this residue and rationalizes its role in the stabilization of the coiled coil structure. In contrast, we concluded that this residue acts to destabilize the dimer based on our lattice simula-



**Figure 3.** Side (A) and top (B) views of tube diagrams of the backbones of the 5 refined structures are shown in red, yellow, green, cyan and white, along with the crystal structure which is shown in magenta. The backbones of the structures are so close that they overlap in many places; the picture basically shows one fused 6-color tube.



**Figure 4.** All atom models of the 5 refined structures and the average backbone structure generated from them. The hydrophobic residues (Leu, Val, Met) are drawn in green, the acidic (Glu, Asp, Tyr, Ser) residues in red, the basic (Lys, Arg, His) residues in cyan, and the Asn residues are shown in magenta. The tube diagram of the backbone of the average structure is shown in yellow. The mobility of the core residues is minimal, whereas the exposed residues are very mobile. Salt bridges between basic and acidic residues are apparent.

tions. This could be related to the actual destabilization of the hydrophobic core by the hydrophilic Asn or could be due to the simplifications and errors in our reduced model.

Goodman & Kim (1991) have measured amide proton exchange rates as a function of position of a modified GCN4 dimer. An interesting question is what features of this experiment, if any, the present model can reproduce. We found that the lengths of the backbone hydrogen bonds computed from the final structure and averaged over the three heptets (residues 9 to 29) show a periodicity which qualitatively correlates with the experimental results. However, in the model, the  $\epsilon$  position (rather than  $\delta$  position) has the deepest minimum. In general, the proton amide exchange process is very complicated, and a detailed comparison of these models with experiment is beyond the scope of this paper.

Starting from randomly generated chains and without any assumptions about the final conformation, structures of short coiled coils, which are almost indistinguishable from experiment, have been predicted. These simulations also indicate that, in addition to the hydrophobic effect and salt bridge

interactions (Cohen & Parry, 1990; O'Shea *et al.*, 1991), hydrogen bonds may also play an important role in the stabilization of some coiled coils. Furthermore, cooperative side-chain packing interactions may substantially contribute to the stabilization of coiled coil structures. Clearly, the folding of the GCN4 leucine zipper is just a preliminary step in the study of larger coiled coils (Banner *et al.*, 1987; Jianming *et al.*, 1991; Phillips *et al.*, 1986). While much remains to be done before the protein folding problem is solved, our results for this simple folding system suggest that a hierarchical approach such as the one we have developed is a promising path towards its solution.

Valuable discussions with Drs William Beers, Reza Ghadiri, Adam Godzik, Mariusz Milik, Antonio Rey and Lee Walters are acknowledged. We thank Professor Peter Kollman for suggesting the use of a water sheath in the MD refinement. This research was supported in part by grant numbers GM-38794 and GM-37554 of the Division of General Medical Sciences of the National Institutes of Health.

#### References

- Alber, T. (1992). Structure of the leucine zipper. *Curr. Opin. Genet. Develop.* **2**, 205–210.
- Banner, D. W., Kokkinidis, M. & Tsernoglou, D. (1987). Structure of the ColE1 rop protein at 1.7 Å resolution. *J. Mol. Biol.* **196**, 657–675.
- Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer E. F., Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Simanouchi, T. & Tasumi, M. (1977). The Protein Data Bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* **112**, 535–542.
- Brooks, B. R., Bruccoleri, R., Olafson, B., States, D., Swaminathan, S. & Karplus, M. (1983). CHARMM: a program for macromolecular energy minimization, and molecular dynamics. *J. Comp. Chem.* **4**, 187–217.
- Cohen, C. & Parry, A. H. (1990).  $\alpha$ -Helical coiled coils and bundles: how to design an  $\alpha$ -helical protein. *Proteins: Struct. Funct. Genet.* **7**, 1–15.
- Cohen, C. & Parry, D. A. D. (1986).  $\alpha$ -Helical coiled coils—a widespread motif in proteins. *Trends Biochem. Sci.* **11**, 245–248.
- Creighton, T. E. (1990). Protein folding. *Biochem. J.* **270**, 1–16.
- Crick, F. H. C. (1953). The packing of  $\alpha$ -helices: simple coiled coils. *Acta Crystallogr.* **6**, 689–697.
- Ellenberger, T. E., Brandl, C. J., Struhl, K. & Harrison, S. C. (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted  $\alpha$  helices: crystal structure of the protein–DNA complex. *Cell*, **71**, 1223–1237.
- Ferre-D'Amare, A. R., Prendergast, G. C., Ziff, E. B. & Burley, S. K. (1993). Recognition by Max of its cognate DNA through a dimeric b/HLH/Z domain. *Nature (London)*, **363**, 38–45.
- Godzik, A., Kolinski, A. & Skolnick, J. (1993a). De novo and inverse folding predictions of protein structure and dynamics. *J. Comp. Aided Mol. Design*, **7**, 397–438.
- Godzik, A., Kolinski, A. & Skolnick, J. (1993b). Lattice representation of globular proteins: how good they are and how good they have to be. *J. Comp. Chem.* **14**, 1194–1202.

- Goodman, E., M. & Kim, P. S. (1991). Periodicity of amide proton exchange rates in a coiled-coil leucine zipper peptide. *Biochemistry*, **30**, 11615–11620.
- Guenot, J. & Kollman, P. (1992). Molecular dynamics studies of DNA-binding protein: 2. An evaluation of implicit and explicit solvent models for the molecular dynamics simulation of the *Escherichia coli trp* repressor. *Protein Sci.* **1**, 1185–1205.
- Guenot, J. & Kollman, P. A. (1993). Conformational and energetic effects of truncating nonbonded interactions in an aqueous protein dynamics simulation. *J. Comp. Chem.* **14**, 295–311.
- Harbury, P. B., Zhang, T., Kim, P. S. & Alber, T. (1993). A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science*, **262**, 1401–1407.
- Hodges, R. S., Saund, A. S., Chong, P. C. S., St -Pierre, S. A. & Reid, R. E. (1981). Synthetic model for two-stranded,  $\alpha$ -helical coiled coils. *J. Biol. Chem.* **256**, 1214–1224.
- Jernigan, R. L. (1992). Protein folds. *Curr. Opin. Struct. Biol.* **2**, 248–256.
- Jianming, M., Holtzer, M. E. & Holtzer, A. (1991). Kinetics of self-assembly of  $\alpha\alpha$ -tropomyosin coiled coils from unfolded chains. *Proc. Nat. Acad. Sci., U.S.A.* **88**, 916–920.
- Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935.
- Kolinski, A. & Skolnick, J. (1992). Discretized model of proteins. I. Monte Carlo study of cooperativity in homopolypeptides. *J. Phys. Chem.* **97**, 9412–9426.
- Kolinski, A. & Skolnick, J. (1993). A general method for the prediction of the three dimensional structure and folding pathway of globular proteins: application to designed proteins. *J. Chem. Phys.* **98**, 7420–7433.
- Levitt, M. (1991). Protein folding. *Curr. Opin. Struct. Biol.* **1**, 224–229.
- Lovejoy, B., Seunghyon, C., Cascio, D., McRorie, D. K., DeGrado, W. F. & Eisenberg, D. (1993). Crystal structure of a synthetic triple-stranded  $\alpha$ -helical bundle. *Science*, **259**, 1288–1293.
- Nilges, M. & Brunger, A. T. (1991). Automated modeling of coiled coils: application to the GCN4 dimerization region. *Protein Eng.* **4**, 649–659.
- Nilges, M. & Brunger, A. T. (1993). Successful prediction of the coiled coil geometry of the GCN4 leucine zipper domain by simulated annealing: comparison to the X-ray structure. *Proteins: Struct. Funct. Genet.* **15**, 133–146.
- O'Shea, E. K., Klemm, J. D., Kim, P. S. & Alber, T. (1991). X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science*, **254**, 539–544.
- PDB (1992). *Quarterly Newsletter*, no. 61.
- Phillips, G. N., Jr, F., J. P. & Cohen, C. (1986). Tropomyosin crystal structure and muscle regulation. *J. Mol. Biol.* **192**, 111–131.
- Ptitsyn, O. B. (1987). Protein folding: hypotheses and experiments. *J. Protein Chem.* **6**, 273–293.
- Skolnick, J. & Kolinski, A. (1993). Parameters of statistical potentials. Available by ftp from the public directory: scripps.edu (pub/MCDP).
- Skolnick, J., Kolinski, A., Brooks, C., III, Godzik, A. & Rey, A. (1993). A method for prediction of protein structure from sequence. *Curr. Biol.* **3**, 414–423.
- SYBYL, 6.0 (1992). Tripos Associates, Inc., St Louis, MO.

Edited by F. Cohen

(Received 4 November 1993; accepted 5 January 1994)