PROTEINS: Structure, Function, and Genetics 15:133–146 (1993)

Successful Prediction of the Coiled Coil Geometry of the GCN4 Leucine Zipper Domain by Simulated Annealing: Comparison to the X-Ray Structure

Michael Nilges and Axel T. Brünger
The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, Connecticut 06511

ABSTRACT **The recently solved X-ray structure of the dimerization region ("leucine zipper") of the yeast transcriptional activator GCN4 (O'Shea, E.K., Klemm, J.D., Kim, P.S., Alber, T. Science 254:539–544, 1991) is compared to previously predicted models which had been obtained by a conformational search procedure employing simulated annealing without any knowledge of the crystal coordinates (Nilges, M., Brünger, A.T. Protein Eng. 4:649–659, 1991). During the course of the simulated annealing procedure, the models converged towards the X-ray structure. The averaged root mean square difference between the models and the X-ray structure is 1.26 and 1.75 Å for backbone atoms and all nonhydrogen atoms at the dimerization interface, respectively. The local helix–helix crossing angle of the X-ray structure falls within the range predicted by the models; a slight unwinding of the coiled coil toward the N-terminal DNA-binding end of the dimerization region has been correctly predicted. Distance maps between the helices are largely identical. The region around asparagine 20 is asymmetric in the X-structure and in the models. Surface side chain dihedrals showed a large variation in the models although the χ_1, χ_2, χ_3, χ_4 3-fold dihedrals were correctly predicted in 69, 42, 43, and 44% of the cases, respectively. Phenomenological free energies of dimerization of the models show little correlation with the root mean square difference between the models and the X-ray structure.**

Key words: coiled coil, GCN4, leucine zipper, model building, structure prediction, molecular dynamics

INTRODUCTION

The "leucine zipper" motif has been recognized to be a common dimerization domain in a novel class of transcription factor proteins.[1] Crystallographic[2,3] and NMR spectroscopic[4] structure investigations have now confirmed that the three-dimensional structure of the leucine zipper motif consists of two α-helices with the same sequential directionality forming a coiled coil. The coiled coil represents one of the most efficient packing modes of helices.[5] The dimerization interface consists of a heptad repeat of hydrophobic residues with a preference for leucine residues at the "d" position. Through this dimerization interface, the transcription factors are able to form a variety of homo- and heterodimers. The specificity of these interactions is regulated by specific interactions between the amino acid residues at the dimerization interface. As dimerization of the transcription factors is a prerequisite for strong DNA binding, the protein–protein interactions at the dimerization interface may thus indirectly affect transcriptional regulation.

The simplicity of this dimerization domain makes it a good model system for the study of protein–protein interactions. A large body of experimental data is available about the dimerization behavior of various chimeric constructs of transcription factors and point mutants (for recent reviews see 6, 7). Protein–protein recognition and specificity of dimerization are related to the protein folding problem, since the physical laws that determine the stability of complexes of proteins and those that govern the stability of monomeric proteins are the same.

As a first step toward understanding the specificity of protein–protein interactions in leucine zippers we recently applied an automated modeling procedure to GCN4, which is assisted by simulated annealing and uses only minimal assumptions about the conformation of the protein.[8] The predicted GCN4 structures were obtained without any knowledge about the X-ray structure. The procedure

Abbreviations: rmsd, root mean square differences; SA, simulated annealing; NMR, nuclear magnetic resonance; GCN4, yeast transcriptional activator.

Received March 4, 1992; revision accepted April 27, 1992.

Address reprint requests to Dr. Axel T. Brünger, Department of Molecular Biophysics and Biochemistry, Yale University, 260 Whitney Avenue, P.O. Box 6666, New Haven, CT 06511.

Present address of M. Nilges: European Molecular Biology Laboratory, Meyerhofstr. 1, D-6900 Heidelberg, Federal Republic of Germany.

TABLE I. Model Building Through Simulated Annealing

I. Initialization of C^{α} skeleton (coiled coil with zero crossing angle)	←	Various orientations
↓		
II. Generation of all atoms by simulated annealing		
↓		
III. Relaxation of the structure by simulated annealing with a simplified empirical energy function and minimal restraints		
↓		
IV. Relaxation of the structure with a full empirical energy function in explicit water	→	Family of conformations

started from a distorted coiled coil with infinite pitch, consisting of C^{α} atoms only. The remaining atoms were placed automatically and then the whole structure was allowed to relax using simulated annealing. Initially, a simplified empirical energy function with a minimal set of restraints was used to speed the convergence process. This was followed by unrestrained molecular dynamics of the solvated system with an explicit water environment. Several initial conformations were refined to produce a family of several predicted structures. The spread of this family represents a measure of the accuracy of the predicted models similar to structure determination of solvated macromolecules by nuclear magnetic resonance (NMR) spectroscopy.[9,10]

In this paper we present a detailed comparison between our predicted structures and the X-ray structure of O'Shea et al.[3] We compare several geometric and conformational properties of the structures. We also assess the accuracy of phenomenological free energies of dimerization by comparing the value obtained for each model with its rms difference to the X-ray structure.

METHODS

We give here only a brief outline of the automated modeling procedure employing simulated annealing (Tables I and II). For a detailed description of the procedure the reader is referred to our previous paper.[8] In stage I, the positions of the C^{α} atoms are initialized to form two parallel, straight, slightly overwound α-helices with a relative orientation that allows knobs-into-holes packing of the leucines. The crossing angle of these starting structures is 0. In stage II, the remaining atoms were built with a simulated annealing procedure, with the C^{α} positions held fixed. This procedure was adopted from the NMR structure determination algorithm described in ref. 11, which starts from a random array of atoms; it is as efficient and accurate as other side-chain building procedures described in the literature, e.g., the one by Holm and Sander[12] or the one by Lee and Subbiah.[13] In stage III, the whole protein was allowed to relax with an *r*-dependent dielectric screening function[14] which represents a crude approximation of solvent, a penalty function that inhibits dissociation of the helices, and harmonic restraints that "lock" the helical NH–CO hydrogen bonds while allowing some flexibility. In stage IV, unrestrained molecular dynamics was carried out on the system with an explicit water environment using a combination of stochastic and periodic boundary conditions. All calculations were carried out with X-PLOR.[15]

TABLE II. Some Details of the Model Building Procedure

Stage	I	II	III	IV
Initialization	C^{α} skeleton, infinite pitch, separation < 10.4 Å, twist 3.5 residues per turn	—	—	—
Empirical energy function	—	Geometric,[33] no electrostatic term	PARAM19,* *r*-dependent dielectric	OPLS†/Amber‡ (protein), TIP3P§ (water)
Restraints	—	C^{α} positions fixed	Helix–helix distance < 10.2 Å, NH–CO backbone hydrogen bond distance < 3.2 Å	—
Simulation time	—	1.4 psec	9 psec	5 psec

*CHARMM protein force field.[14]
†Nonbonded parameters from ref. 22.
‡Parameters for bonds, bond angles, dihedral, and improper torsion angles from ref. 34.
§TIP3P water model from ref. 21.

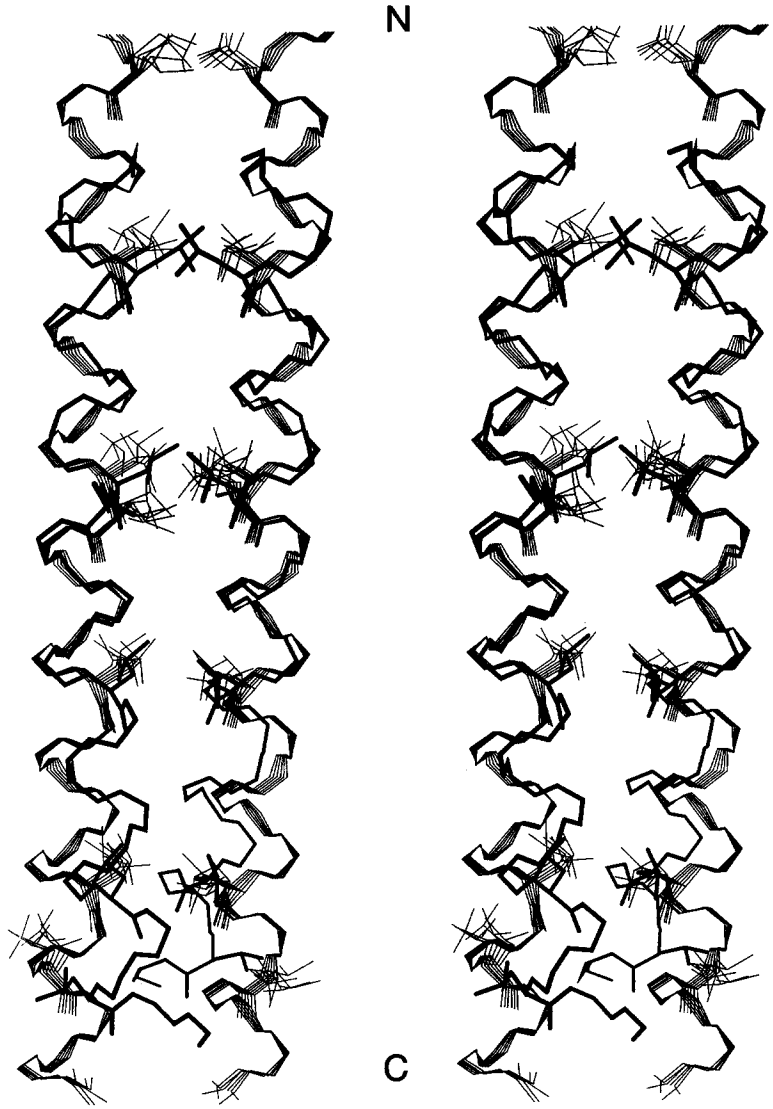

Fig. 1. Stereoplot of the C, N, C^{α} backbone atoms of the initial models m1, . . . , m6 (thin lines) after stage II (Table I) and the X-ray structure (thick lines). The model structures were superposed on the X-ray structure by least-squares fitting. Note that the thickness of the lines is modulated according to the distance of the atomic position to the plane of the paper (depth-cueing). The left helix is approximately located in the plane of the paper, while the right helix is slightly tilted with respect to the paper plane.

Six models (referred to as m1, . . . , m6) were obtained by changing the initial relative orientations of the two helices over a range of 25°. A seventh model ($\overline{m}$) was obtained by averaging the models m1, . . . , m6 after stage III (Table I) and then subjecting the average structure to stage IV. The reason for computing this average structure is based on experience with NMR structure determination: average structures appear to be closer to the X-ray structure.[16]

Our procedure is inherently suitable for parallelization on modern computer architectures at various different levels: on the lowest level, complete calculations could be carried out for each model on separate processors. At the intermediate level, the simulated annealing algorithm could be implemented for parallel processing. At the highest level, the energy functions could be parallelized.

Atomic rms differences were calculated by least squares fitting using the method of Kabsch.[17] The phenomenological free energies of dimerization F_E[18] and F_N[19] were computed. In the former case, the solvation free energy of the dissociated and extended conformation was subtracted from that of the dimer to yield F_E. The computation of both F_N and F_E required the evaluation of the accessible surface areas per atom which were computed by the method of Lee and Richards[20] using van der Waals radii calculated

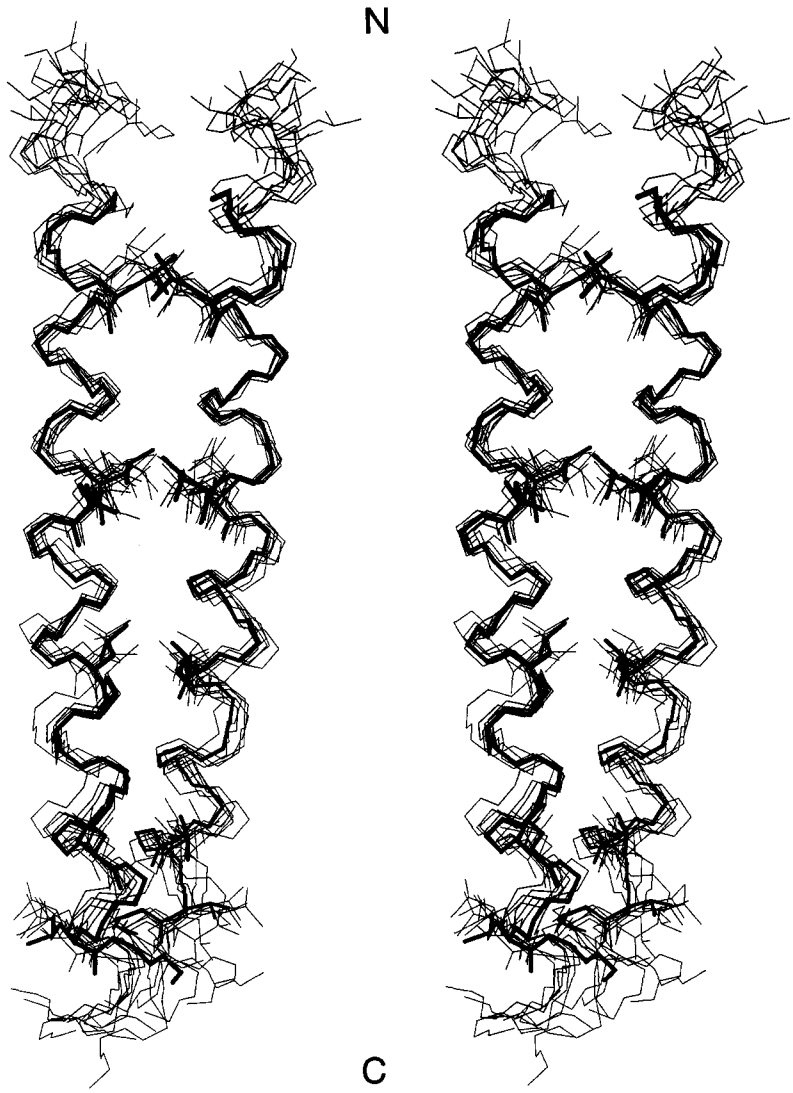

Fig. 2. Stereoplot of the C, N, C^{α} backbone atoms of the models m1, . . . , m6, $\overline{m}$ (thin lines) after stage IV (Table I) and the X-ray structure (thick lines). The model structures were superposed on the X-ray structure by least-squares fitting. Note that the thickness of the lines is modulated according to the distance of the atomic position to the plane of the paper (depth-cueing). The left helix is approximately located in the plane of the paper, while the right helix is slightly tilted with respect to the paper plane.

from the "OPLS" nonbonded parameters.[21] The "entropy loss contribution" to F_N was estimated by checking, for each dihedral degree of freedom of the side chain, if an atom of the other helix is closer than 4.5 Å. In this case, this dihedral degree of freedom is assumed to be immobilized and some entropy may be lost. The electrostatic contribution to F_N was evaluated with the OPLS charges.[22]

The residue number notation used in this paper is identical to the one used in our previous paper,[8] that is, residue 1 in our notation corresponds to residue 244 in the actual sequence of GCN4. The observed X-ray coordinates of O'Shea et al.[3] consist of Arg-5 through Gly-35 in our notation. The phenomenological free energies and atomic rms differences were computed for residues 5 to 35. The two helices of the coiled coil are referred to as "A" and "B." The helical wheel notation of the coiled coil is the same as the one used in ref. 8, i.e., the "d" positions are occupied by leucine residues. The local crossing angle between the two helices was computed as described in our previous paper,[8] i.e., it was measured as the dihedral angle between two pairs of reference points close to the local helix axes. Each reference point was defined as the geometric centers of seven consecutive C^{α} atoms.

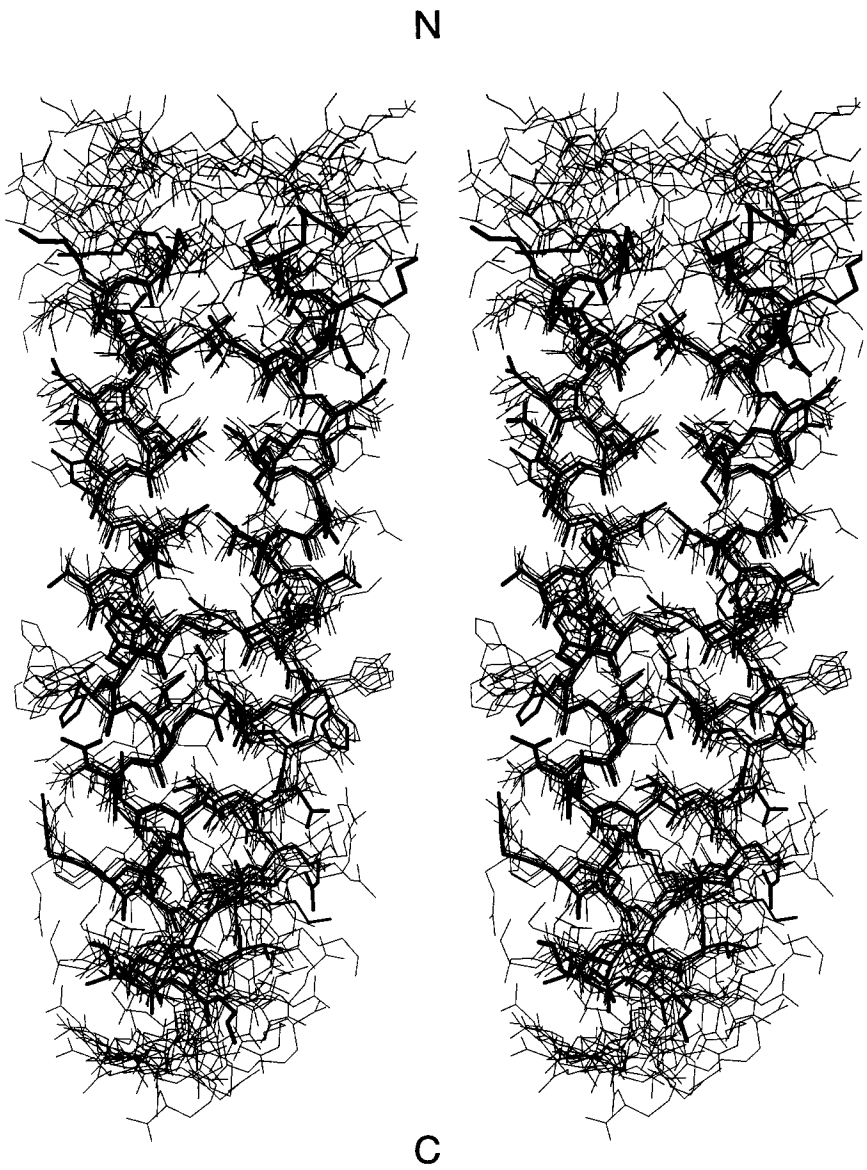

Fig. 3. Stereoplot of all nonhydrogen atoms of the models m1, . . . , m6, $\overline{m}$ (thin lines) after stage IV (Table I) and the X-ray structure (thick lines). The model structures were superposed on the X-ray structure by least-squares fitting.

RESULTS AND DISCUSSION

Convergence Towards the X-Ray Structure

Superpositions with the X-ray structure of the initial models at stage I and the final models after stage IV are reported in Figures 1 through 4. Rms difference plots between the initial models, the final models, and the X-ray structure are shown in Figure 5 and the overall rms differences between the models and the X-ray structure are listed in Table III. The models have clearly converged towards the X-ray structure; the backbone atom rms difference is reduced from 3.1 Å to an average value of 1.26 (Table III). The X-ray structure falls within the family of final models as the average pairwise rms difference between the final models is approximately of the same magnitude as the average rms difference between the final models and the X-ray structure (Fig. 5).

Crossing Angle

The local crossing angle as a function of residue number is plotted in Figure 6 for the X-ray structure, the initial, and the final models. The model's crossing angles converged toward the X-ray structure and the X-ray structure's crossing angles fall within the range predicted by the models. The mean

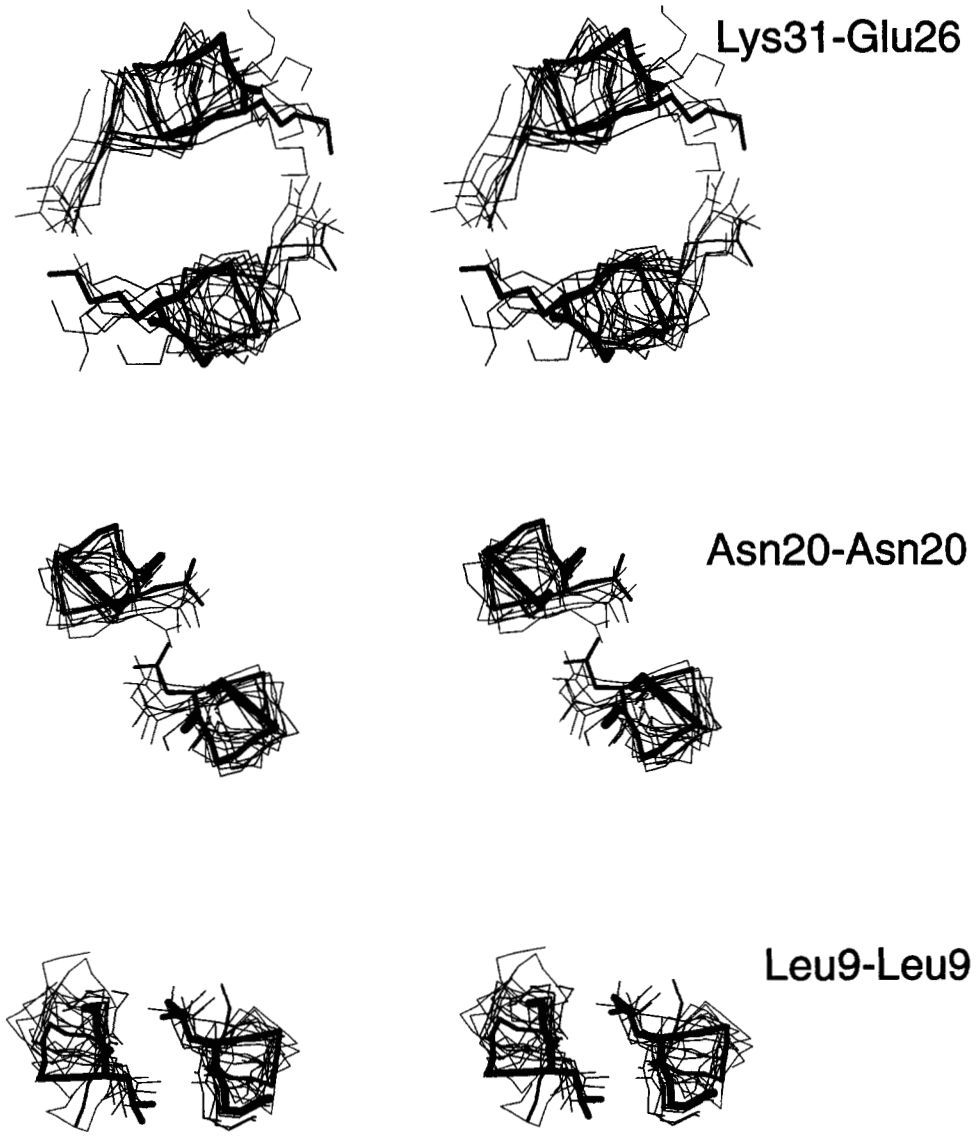

Fig. 4. Stereoplots of the Lys-31–Glu-26 salt bridge, the interaction of the Asn-20 residues (at the "a" positions), and the interaction of the Leu-9 (at the "d" positions) residues. Shown are the side chain atoms and neighboring backbone atoms. The model structures (thin lines) were superimposed on the X-ray structure (thick lines) by least-squares fitting.

value of the predicted crossing angles is somewhat lower compared to the X-ray structure. This is in contrast to the catabolyte gene activator protein used as a test case for our automated modelling procedure[8] where the mean crossing angle was very close to that of the X-ray structure. The average distance between the two helix axes ranges from 9.5 to 9.8 Å which is slightly larger than the value of 9.3 Å found in the X-ray structure. A decrease of the crossing angle, i.e., unwinding of the coiled coil, toward the DNA-binding end of the dimerization region is observed in the X-ray structure and in the models; however, this unwinding is more pronounced in the model structures.

Dihedral Angles

Statistics of the dihedral angle differences between the models and the X-ray structure are shown in Tables IV and V. Detailed differences are shown for selected residues in Figure 7, where dihedral angles are displayed as "dials."[23] The dihedral angles of the X-ray structure are indicated by dashed straight lines and the model's dihedral angles are displayed as "trajectories" starting at the center of the dial with model $\overline{\text{m}}$, followed by model m1, and ending at the circumference with model m6. Helices A and B are shown as thick and thin lines, respectively.

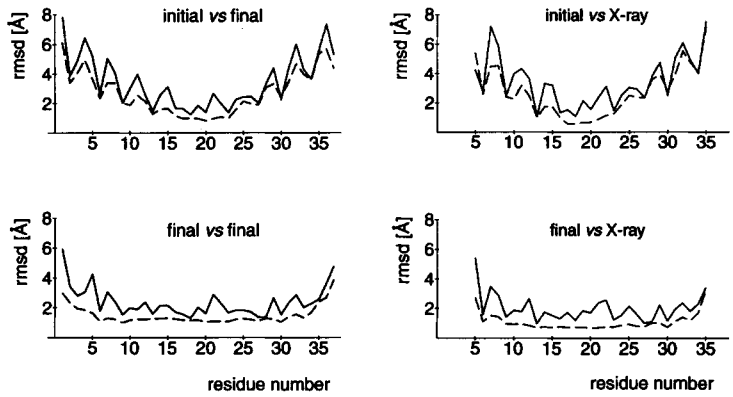

Fig. 5. Average pairwise rms differences between the model structures themselves, and between the X-ray structure and the model structures. The rms differences are computed for backbone (dashed lines) and all nonhydrogen atoms (solid lines) as a function of residue number for both the initial models m1, . . . ,m6 after stage II (Table I) and the final models m1, . . . ,m6, $\overline{m}$ after stage IV.

TABLE III. Atomic rms Differences Between the X-Ray Structure and the Initial and Final Model Structures in Å

Model	Backbone* at stage I	After stage IV	
		Backbone*	Contact†
m1	3.1	1.29	1.94
m2	3.1	1.10	1.90
m3	3.1	0.95	1.39
m4	3.1	1.23	2.14
m5	3.1	0.92	1.58
m6	3.1	2.11	2.80
$\overline{m}$	—	1.25	1.75
Average	3.1	1.26	1.93
σ	0	0.37	0.42

*Rms difference evaluated for the backbone atoms (C, N, and C^{α}).
†Rms difference evaluated for all nonhydrogen atoms at the interface between the helices.

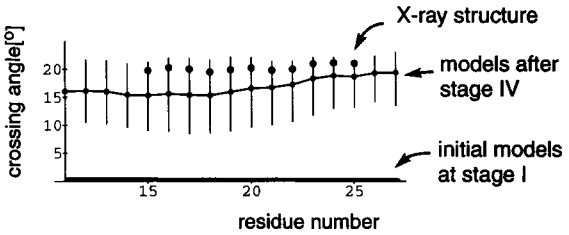

Fig. 6. Local crossing angle $\phi(i)$ between the two helices as a function of residue number i. The angle was estimated by measuring the torsion angle between neighboring points on the two helix axes as described in ref. 8. The solid line and the bars denote the mean value and the range of the model's crossing angles after stage IV. The dots indicate the crossing angles of the X-ray structure. The crossing angle (0°) for the initial models at stage I is indicated by a thick line.

TABLE IV. Rms Deviation for Backbone Dihedral Angles Between the X-Ray and Model Structures*

φ	ψ	ω
17°	18°	8°

*Excluding the first and last residues of the X-ray structure.

The backbone conformation of a peptide chain is described by the φ (C_{i-1}–N_i–C_i^{α}–C_i) and ψ (N_i–C_i–C_i^{α}–N_{i+1}), and ω (peptide plane) dihedral angles. The ω angle is restrained to *trans* during stage II, and is consequently close to 180° after stage IV (Fig. 7 and Table IV). The φ and ψ angles are never directly restrained in the energy functions employed, although the α-helical hydrogen bond restraints during stage III impose certain restrictions on these angles. Furthermore, the initial C^{α} positions are, in principle, sufficient to determine the remaining backbone atom positions.[24] However, during stage IV neither direct nor indirect restraints are imposed on the φ and ψ angles. Nevertheless, the conformation of the molecule stays α-helical with small deviations from the X-ray structure (Table IV). Significant differences are found only near the amino terminal side of the dimerization domain (e.g., φ of Met-6 in Fig. 7) where the X-ray structure deviates significantly from a standard helical conformation.

The side chain conformation of a polypeptide chain is characterized through the $\chi_1, \ldots, \chi_5$ angles. These angles clearly show a large variation among the models (Fig. 3). Nevertheless, the models predict the correct rotamer conformation of 3-fold dihedrals in 43 to 69% of the cases for χ_1, χ_2, χ_3, and χ_4 angles (Table V). This is significantly better than the random limit of 33%. Clearly, the predictability of the side chain conformation decreases with increasing side chain length.

The side chains at the interface between the two helices generally occupy a similar region in the mod-

TABLE V. Fraction of Correctly Predicted Side Chain Rotamer Conformations for 3-Fold Dihedral Angles*

χ_1	χ_2	χ_3	χ_4
69%	42%	43%	44%

*Three-fold dihedral angles were classified according to the rotamer conformations 0–120°, 120–240°, and −120–0°. A predicted side chain dihedral angle was counted as correct if both it and that of the X-ray structure fell into the same rotamer conformation.

els and the X-ray structure. This is illustrated by three examples in Figure 4, the Leu-9–Leu-9 interaction, the Asn-20–Asn-20 interaction, and the salt-bridge between residues Glu-26 and Lys-31. The average rms difference between the final models and the X-ray structure for all nonhydrogen atoms at the interface is 1.93 Å (Table III). The leucines in the "d" position of the helical wheel (Leu-9, Leu-16, Leu-23, and Leu-30) represent an interesting discrepancy: the X-ray structure exhibits symmetric conformations with respect to the molecular diad, whereas the model's leucine conformations are mostly asymmetric (see, for example, Fig. 4). This is also apparent from the nearly asymmetric pattern of the leucine χ_2 dihedral angle "trajectories" in Figure 7. The predicted conformations resemble those found in the coiled coil at the dimerization interface of the catabolyte gene activator protein.[25]

Distance and Charge–Charge Maps

Rms differences and dihedral angles are not always the best indicators to study differences in molecular conformation. Comparisons of distance matrices are intrinsically free of bias by least-squares superpositions of the molecules.[26] Figure 8 shows interhelix side chain distance maps for the X-ray structure and the models. The maps appear very similar and indicate that corresponding side chains are in similar positions relative to each other.

The proximity of charged side chains is shown in the "charge–charge maps" where the electrostatic interaction energy is computed on a residue-by-residue basis, averaged over the final models (Figs. 9 and 10). The patterns of intrahelix interactions (Fig. 9) and of interhelix interactions (Fig. 10) are in good agreement with the X-ray structure, although there are some differences in detail which could be caused by crystal packing contacts in the X-ray structure on the one hand and by deficiencies in the empirical energy function on the other hand. The strong interactions between helices involving Lys-31–Glu-26 and Glu-24–Lys-19 are present in both the models and the X-ray structure, whereas the model's interhelix interactions involving Glu-10–Arg-5 and Lys-12–Glu-10 are missing in the X-ray structure. In the following we define the existence of a salt-bridge between lysine and glutamic acid residues by the requirement that the distance between the center of the carboxyl O^{ϵ} and the N^{ζ} atoms is less than 6 Å. The Glu-26–Lys-31 interaction then produces two symmetric salt bridges between the helices in the X-ray structure and in 40% of the models; at least one salt bridge is formed in all the models (Table VI). The Glu-24–Lys-19 interaction produces a single, asymmetric salt bridge in the X-ray structure. Only model m3 produces this salt bridge (Table VI); the other models do not produce any salt bridge at all involving Glu-24 and Lys-19.

Asymmetry of the Asn-20 Conformation

Residue Asn-20 is buried at the dimerization interface and is conserved among other leucine zippers. It is, however, quite tolerant toward amino acid substitution and it is argued by O'Shea et al.[3] that this asparagine may play a role either in destabilizing the dimer or in maintaining proper registration of opposing heptads by establishing a unique, polar interaction at a nonpolar interface. The models are in agreement with the observation in ref. 3, that the region around Asn-20 is asymmetric (Figs. 4 and 7), and that Asn-20 sterically blocks the formation of a symmetric pair of interhelix salt bridges between Glu-24 and Lys-19 due to an intrahelical interaction of Asn-20 with Glu-24. In fact, none of the models shows the formation of a symmetric pair of salt bridges between Glu-24 and Lys-19 (Table VI). This suggests that it is difficult to pack the residues around Asn-20 in a symmetric conformation *independent of crystal packing contacts*. NMR studies, on the other hand, have shown that the average structure of the leucine zipper is symmetric on the NMR time scale.[4,27,28] The study by Saudek et al.[28] also shows that the region around Asn-20 exhibits a faster amide hydrogen exchange than other regions of the molecule, which could indicate fast exchange between two or more asymmetric conformations involving the Asn-20 residues.

Phenomenological Free Energies of Dimerization

Criteria need to be developed to assess the quality of predicted models in the absence of an X-ray or NMR structure. The use of phenomenological free energies of solvation has been suggested,[18, 19] and has been successful in discriminating correctly from incorrectly folded structures in some cases[29,30] and in studying complex formation of antibodies and hapten compounds.[19] The F_E and F_N phenomenological free energies of dimerization are listed in Table VII for the final models. The values of the two phenomenological free energies are not correlated and

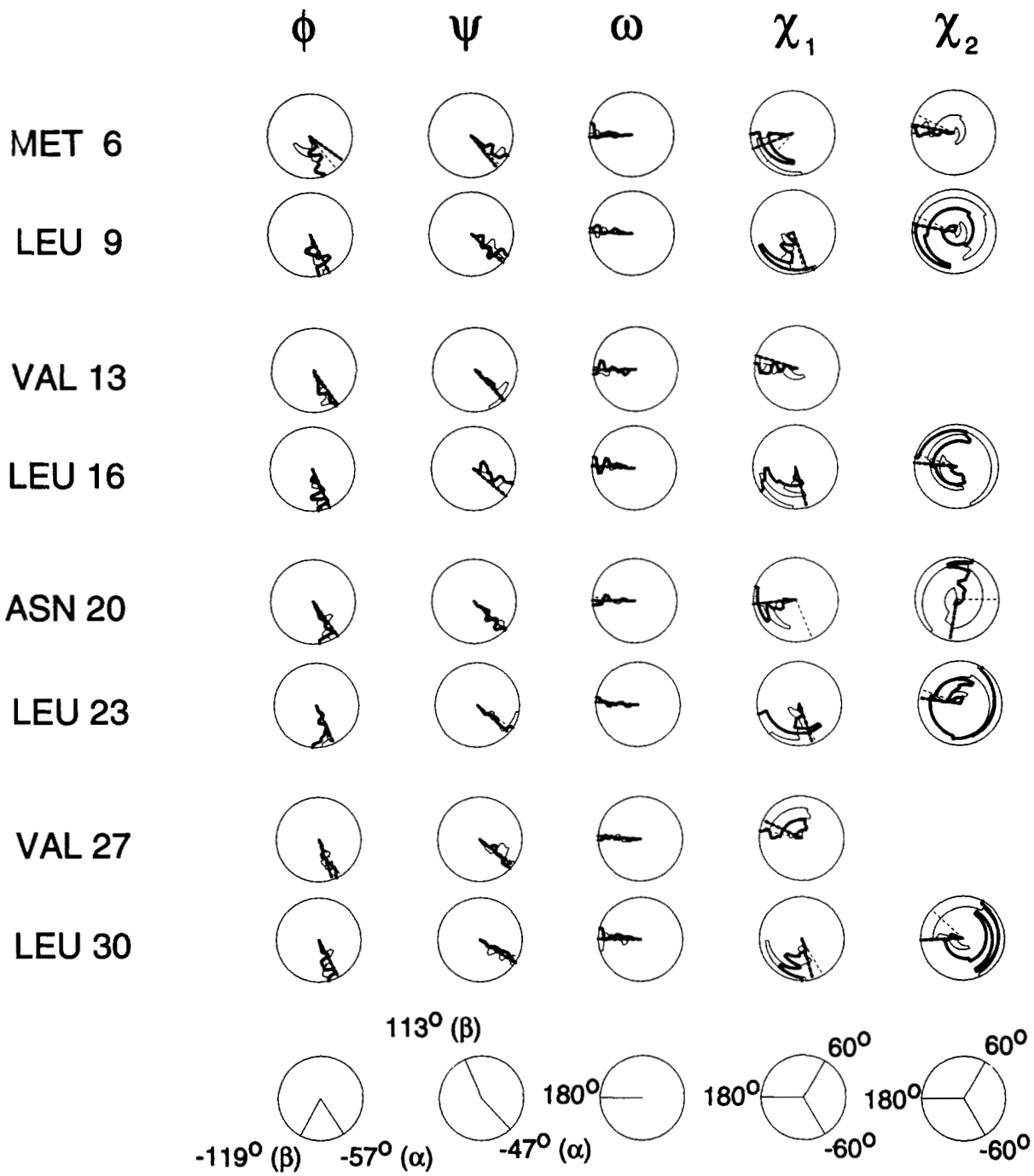

Fig. 7. "Dial" representation[23] of dihedral angles for residues at the interface between the helices (i.e., residues in positions "a" and "d" of the coiled coil). The X-ray structure is shown as dashed lines, the models m1, . . . ,m6, $\overline{m}$ are shown as a "trajectory," with model $\overline{m}$ located at the center of the dial, followed by m1, . . . , m6; m6 is located near the circumference of the dial. Dihedrals of helix A are shown as thick lines, those of helix B as thin lines. As a reference, dials for typical backbone dihedral angles of α-helical and β-strand conformations, and for preferred side chain rotamer conformations are provided at the bottom of the figure.

also show little correlation with the rms difference to the X-ray structure (Table III). In fact, there is a tendency for F_N to be high when F_E is low. This may be a consequence of the numerous interactions of charged and polar side chains with other parts of the protein. F_N contains a term proportional to the electrostatic energy evaluated with a distance dependent dielectric "constant" ($\epsilon = 4R$, where R is the

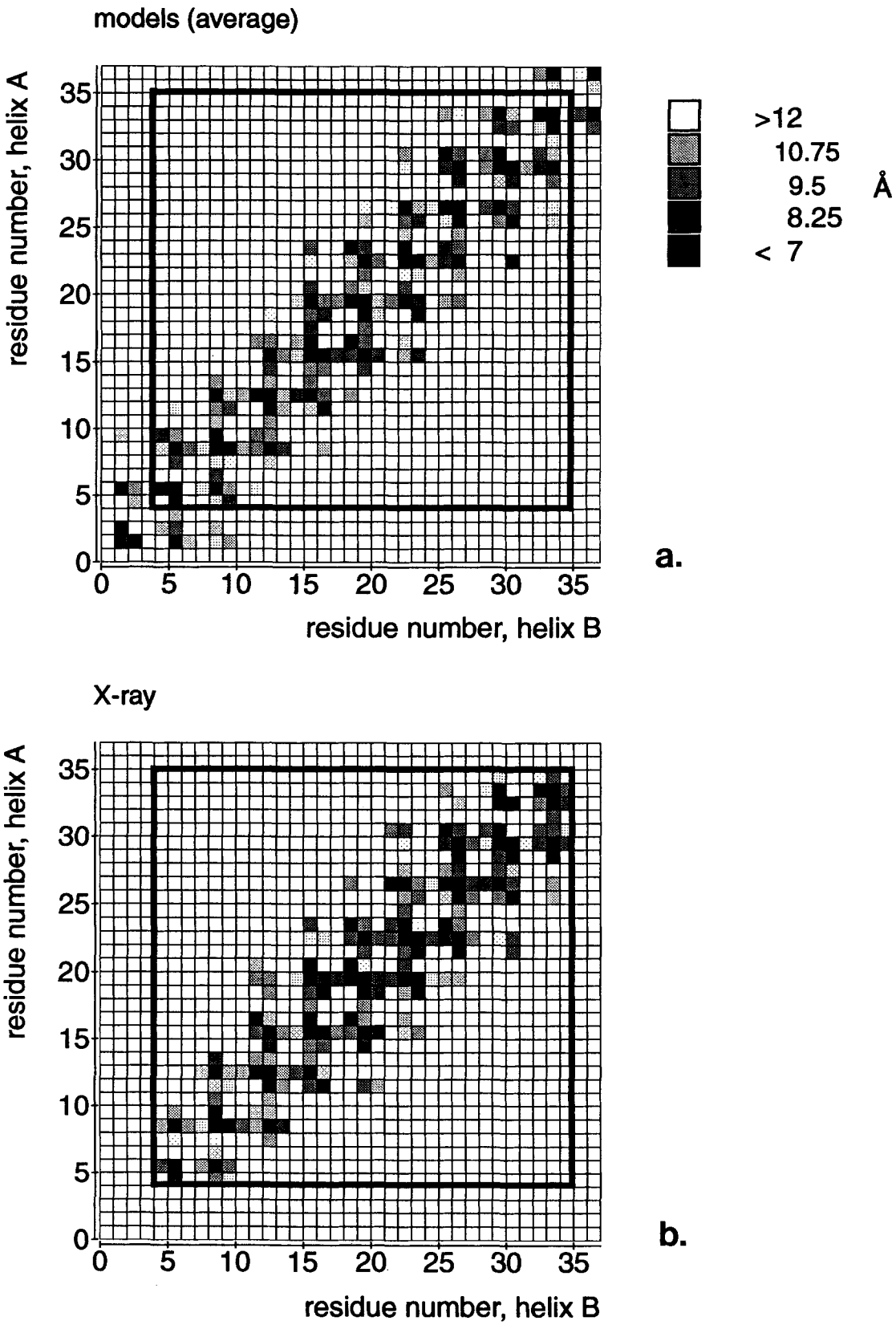

Fig. 8. Interhelix side chain distance maps. **(a)** Average distance computed from the models m1, . . . , m6, $\overline{m}$; **(b)** X-ray structure. Distances were measured from the geometric centers of the nonhydrogen side chain atoms excluding the C^{β} position.

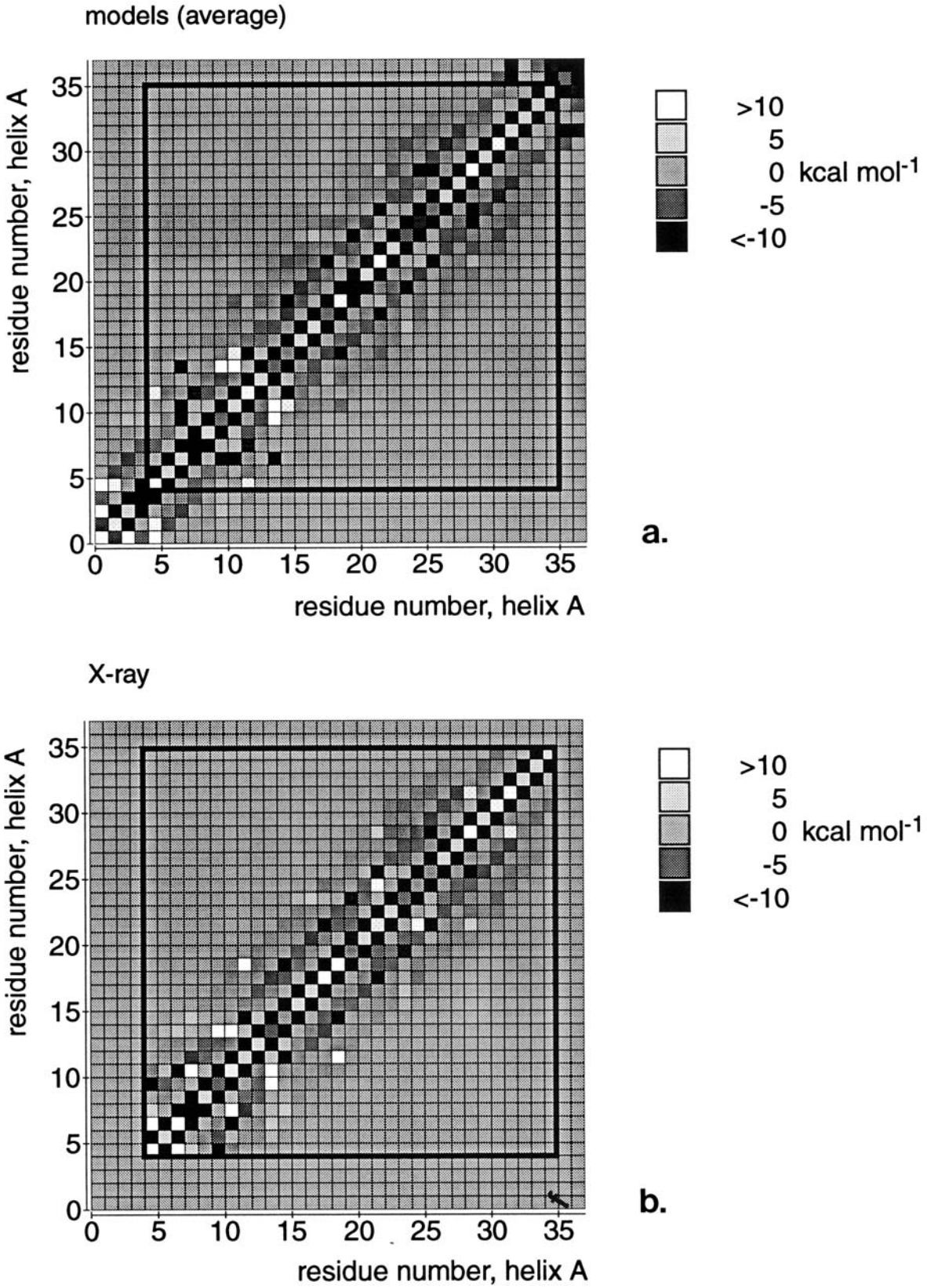

Fig. 9. Electrostatic energy between pairs of residues within helix A, on a residue by residue basis excluding solvent contributions. **(a)** Average computed from the models m1, . . . , m6, $\overline{m}$. **(b)** X-ray structure. The black square indicates the residues that are visible from the X-ray structure.

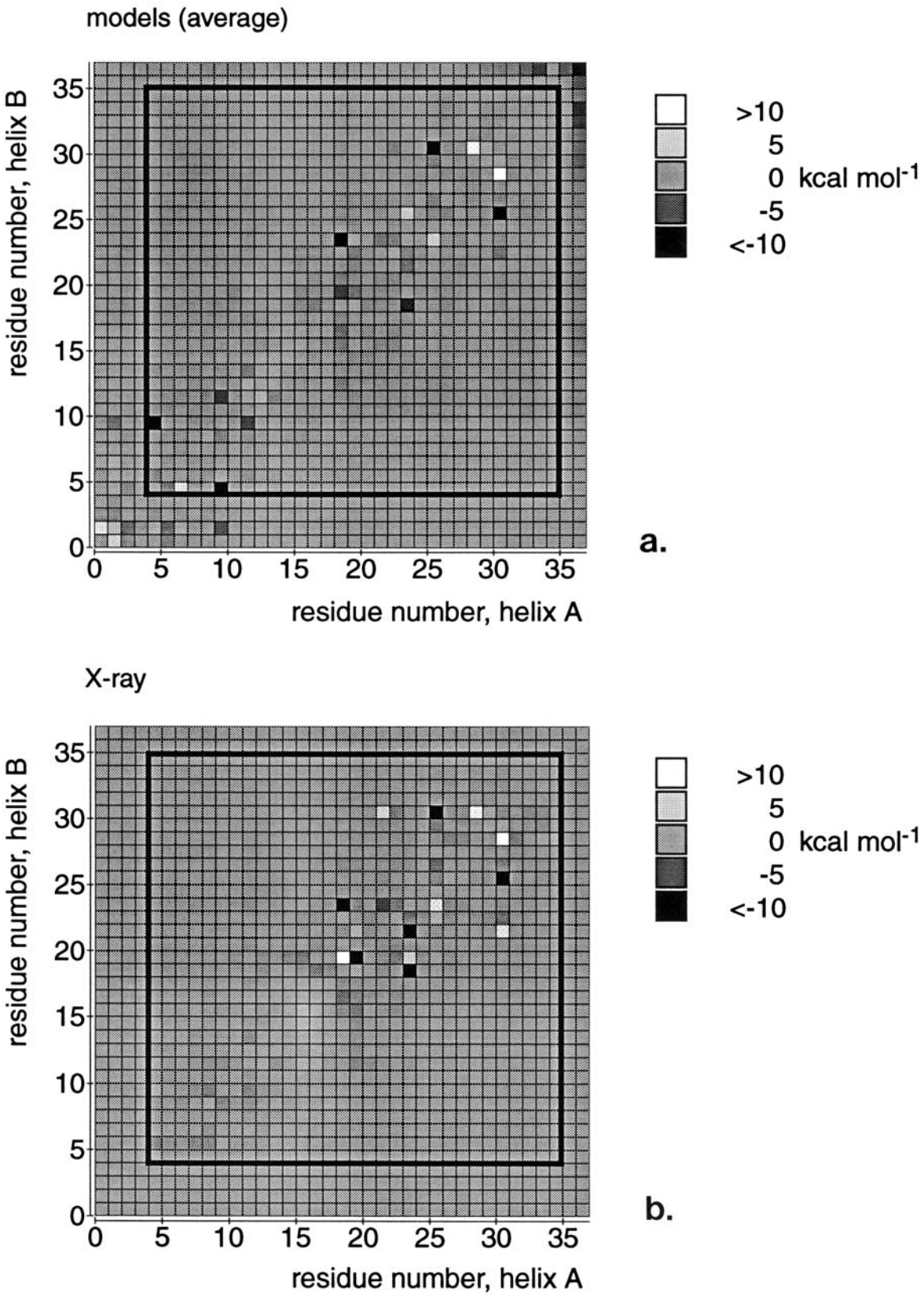

Fig. 10. Electrostatic energy between pairs of residues on different helices, on a residue by residue basis. **(a)** Average computed from the models m1, . . . , m6, $\overline{m}$. **(b)** X-ray structure. The black square indicates the residues that are visible in the X-ray structure.

TABLE VI. Interhelix Salt Bridge Formation Between Glu-26, Lys-31, Lys-19, and Glu-24*

Structure	Glu-24 (A) –Lys-19 (B)	Lys-19 (A) –Glu-24 (B)	Glu-26 (A) –Lys-31 (B)	Lys-31 (A) –Glu-26 (B)
X-ray	4.3	7.3	4.7	4.4
m1	7.5	8.1	7.1	5.4
m2	8.2	10.7	4.8	10.2
m3	4.1	8.9	3.7	3.5
m4	11.1	10.3	3.6	4.9
m5	8.9	7.3	4.6	3.5
m6	6.6	6.4	5.5	7.5
m7	6.3	11.8	11.1	5.1

*Distances in Å between the geometric center of the O^{ϵ} carboxyl atoms of the glutamic acid residues and the N^{ζ} atom of the lysine residues.

TABLE VII. Energies of Model Structures (kcal/mol)

Model	$E_{empirical}$*	F_E†	F_N‡
m1	−26202	−34.5	−14.0
m2	−26083	−27.6	−3.7
m3	−26134	−32.8	−20.0
m4	−26351	−33.5	−14.8
m5	−26202	−38.3	−8.5
m6	−26018	−36.3	−13.2
$\overline{m}$	−26613	−37.4	+0.1
Average	−26229	−34.3	−10.6
σ	185	3.3	6.5

*Energy of the whole system (protein and waters), averaged over the last 50 steps of the simulation.
†$F_E = G_E$(extended)$-G_E$(dimer); G_E is the phenomenological free energy of solvation as described in ref. 18.
‡Phenomenological free energy of dimerization as described in ref. 19.

distance between the two charges). The variation of F_N is largely due to the variation in the electrostatic interaction, the two other contributions that depend on the conformation (a term proportional to the contact surface, and a term depending on the number of side chain degrees of freedom inhibited in the complex) are very similar in all seven models (data not shown). The empirical energy of the whole system (protein and water) shows no correlation with the rms difference to the X-ray structure (Tables III and VII).

CONCLUDING REMARKS

Our successful modeling of the dimerization region of the leucine zipper GCN4 shows that with the current state of methodology and availability of computing power it is possible to make "low-resolution" predictions for structures when the folding motif is known. The coiled coil represents a particularly successful case as the molecular dynamics calculations actually *converge* toward the X-ray structure when starting from a distorted conformation. This is in contrast to molecular dynamics calculations of globular proteins where the models tend to diverge from the X-ray structure after a molecular dynamics calculation *when starting in the X-ray conformation.* Presumably this is related to the fact that the coiled coil represents one of the most efficient packing modes between helices.[5] The remaining differences between models and crystal structure are probably due to limitations of the employed empirical energy function, the multiple minimum problem, and crystal packing effects. Work is in progress to investigate these problems.

The use of phenomenological free energies to assess the quality of the models as measured by their rms difference to the X-ray structure proved unsuccessful. This is presumably a consequence of inappropriate approximations made when evaluating the phenomenological free energies for a system such as GCN4 with many interacting charged side chains. Thus, it is unlikely that phenomenological free energies can be used to obtain absolute stabilities of different homo- and heterodimers of leucine zippers. The experience of others[31] and ourself (unpublished data) for the *fos–jun* system supports this conclusion. However, as our simulated annealing approach appears to sample conformations that are reasonably close to the X-ray structure, microscopic approaches to assess relative stability, such as the free energy perturbation technique (see ref. 32, for a review), should be successful. A combination of these approaches might improve our understanding of the large body of experimental data on various combinations of leucine zippers and mutants thereof.[7] Clearly, proper sampling of conformational space will require substantial computational resources. However, as the problem is intrinsically parallel, the arrival of modern massively parallel computers will probably alleviate this problem.

ACKNOWLEDGMENTS

We thank Dushyant Pathak and Paul Sigler for discussions, Tom Alber for providing us with the GCN4 X-ray structure, Thomas Simonson for writing a dials-type plotting program, and the Pitts-

burgh Supercomputer Center for support (ATB, Grant DMB870007P).

REFERENCES

1. Landschulz, W.H., Johnson, P.F., McKnight, S.L. The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. Science 240:1759–1764, 1988.
2. Rasmussen, R., Benvegnu, D., O'Shea, E.K., Kim, P.S., Alber, T. X-ray scattering indicates that the leucine zipper is a coiled coil. Proc. Natl. Acad. Sci. U.S.A. 88:561–564, 1991.
3. 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.
4. Oas, T.G., McIntosh, L.P., O'Shea, E.K., Dahlquist, F.W., Kim, P.S. Secondary structure of a leucine zipper determined by nuclear magnetic resonance spectroscopy. Biochemistry 29:2891–2894, 1990.
5. Crick, F.H.C. The packing of α-helices: Simple coiled coils. Nature (London) 170:882–883, 1952.
6. Lamb, P., McKnight, S.L. Diversity and specificity in transcriptional regulation: the benefits of heterotypic dimerization. TIBS 16:417–422, 1991.
7. Pathak, D., Sigler, P.B. Updating structure-function relationships in the b Zip family of transcription factors. Curr. Opinion Struct. Biol. 2:116–123, 1992.
8. Nilges, M., Brünger, A.T. Automated modeling of coiled coils. Application to the GCN4 dimerization region. Protein Eng. 4:649–659, 1991.
9. Havel, T.F., Wüthrich, K. An evaluation of the combined use of nuclear magnetic resonance and distance geometry for determination of protein conformations in solution. J. Mol. Biol. 182:281–294, 1985.
10. Clore, G.M., Gronenborn, A.M. Structures of larger proteins in solution: Three- and four-dimensional heteronuclear NMR spectroscopy. Science 252:1390–1399, 1991.
11. Nilges, M., Clore, M.C., Gronenborn, A.M. Determination of three-dimensional structures of proteins from interproton distance data by dynamical simulated annealing from a random array of atoms. FEBS Lett. 239:129–136, 1988.
12. Holm, L., Sander, C. Database algorithm for generating protein backbone and side chain coordinates from a C_α trace. Application to model building and detection of coordinate errors. J. Mol. Biol. 218:183–194, 1991.
13. Lee, C., Subbiah, S. Prediction of protein side-chain conformation by packing optimization. J. Mol. Biol. 217:373–388, 1991.
14. 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. 79:187–217, 1983.
15. Brünger, A.T. "X-PLOR Manual," version 2.1, Yale University, 1990.
16. Clore, G.M., Brünger, A.T., Karplus, M., Gronenborn, A.M. Application of molecular dynamics with interproton distance restraints to three-dimensional protein structure determinations: A model study of crambin. J. Mol. Biol. 191:523–551, 1986.
17. Kabsch, W. A solution for the best rotation to relate two sets of vectors. Acta Crystallogr. A32:922–923, 1976.
18. Eisenberg, D., McLachlan, A.D. Solvation energy in protein folding and binding. Nature (London) 319:199–203, 1986.
19. Novotný, J., Bruccoleri, R.E., Saul, F.A. On the attribution of binding energy an antigen-antibody complexes McPC 603, D1.3, and HyHEL-5. Biochemistry 28:4735–4749, 1989.
20. Lee, B., Richards, F.M. The interpretation of protein structures: Estimation of static accessibility. J. Mol. Biol. 55: 379–400, 1971.
21. Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79:926–935, 1983.
22. Jorgensen, W.L., Tirado-Rives, J. The OPLS potential function for proteins. Energy minimizations for crystals of cyclic peptides and crambin. J. Am. Chem. Soc. 110:1657–1666, 1988.
23. Swaminathan, S., Ravishankar, G., Beveridge, D.L., Lavery, R., Etchebest, C., Sklenar, H. Conformational and helicoidal analysis of the molecular dynamics of proteins: "Curves", dials and windows for a 50 psec dynamic trajectory of BPTI. Proteins 8:179–193, 1990.
24. Purisma, E.O., Scheraga, H.A. Conversion from a virtual-bond chain to a complete polypeptide backbone. Biopolymers 23:1207–1224, 1984.
25. Weber, I.T., Steitz, T.A. Structure of a complex of catabolite gene activator protein and cyclic AMP refined at 2.5 Å resolution. J. Mol. Biol. 198:311–326, 1987.
26. Richards, F.M., Kundrot, C.E. Identification of structural motifs from protein coordinate data: Secondary structure and first-level supersecondary structure. Proteins 3: 71–84, 1988.
27. Saudek, V., Pastore, A., Castiglione Morelli, M.A., Frank, R., Gausepohl, H., Gibson, T., Weih, F., Roesch, P. Solution structure of the DNA-binding domain of the yeast transcriptional activator protein GCN4. Protein Eng. 4:3–10, 1990.
28. Saudek, V., Pastore, A., Castiglione Morelli, M.A., Frank, R., Gausepohl, H., Gibson, T. The solution structure of a leucine zipper motif peptide. Protein Eng. 4:519–529, 1991.
29. Novotny, J., Rashin, A.A., Bruccoleri, R.E. Criteria that discriminate between native proteins and incorrectly folded models. Proteins 4:19–30, 1988.
30. Chiche, L., Gregoret, L.M., Cohen, F.E., Kollman, P.A. Protein model structure evaluation using the solvation free energy of folding. Proc. Natl. Acad. Sci. U.S.A. 87: 3240–3243, 1990.
31. Krystek, S.R., Jr., Bruccoleri, R.E., Novotný, J. Stabilities of leucine zipper dimers estimated by an empirical free energy method. Int. J. Peptide Protein Res. 38:229–236, 1991.
32. Beveridge, D., DiCapua, F. Free energy via molecular simulation: Applications to chemical and biomolecular systems. Annu. Rev. Biophys. Biophys. Chem. 18:431–492, 1988.
33. Brünger, A.T. Refinement of three-dimensional structures of proteins and nucleic acids. In: "Topics in Molecular Biology." J.M. Goodfellow, J.M., ed. London: Macmillan, 1991:137–178.
34. Weiner, S.J., Kollman, P.A., Case, D.A., Singh, U.C., Ghio, C., Alagona, G., Profeta, S., Jr., Weiner, P. A new force field for molecular mechanical simulation of nucleic acids and proteins. J. Am. Chem. Soc. 106:765–784, 1984.