
De Novo Predictions of the Quaternary Structure of Leucine Zippers and Other Coiled Coils

JEFFREY SKOLNICK,¹ ANDRZEJ KOLINSKI,^{1,2}
DEBASISA MOHANTY¹

¹Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 90237

²Department of Chemistry, University of Warsaw, Warsaw 02-093, Poland

Received 16 September 1998; revised 14 October 1998

ABSTRACT: Coiled coils possess a quaternary structure comprised of the side-by-side arrangement of α -helices. Due their inherent structural simplicity, they are ideal model systems for both theoretical and experimental studies. Among the coiled coils are the leucine zippers, which play an important role in the activation of DNA transcription. In contrast to the large amount of available experimental data, an overview of which is presented, there are very few theoretical studies. To address this need, the status of existing theoretical approaches to predict coiled coil quaternary structure is described. Furthermore, to treat the conformational equilibria inherent in these systems, an extension of entropy sampling Monte Carlo simulations is developed that can treat multimers. Here, the approach is applied to GCN4 leucine zippers in the context of a reduced protein model. Not only is the native conformation successfully predicted, but the model also reproduces the experimentally observed helix content in the denatured state and the observed two-state thermodynamic behavior. Such two-state behavior arises from the dissociation of highly helical dimeric chains to form monomers of low, isolated chain helix content. © 1999 John Wiley & Sons, Inc. *Int J Quant Chem* 75: 165–176, 1999

Key words: protein structure prediction; leucine zippers; coiled coil; lattice protein models; quaternary structure prediction

Correspondence to: J. Skolnick.

Introduction

The plethora of protein sequences being provided by genome sequencing projects demands the development of techniques that translate sequence information into knowledge about structure and/or function. While the native structures of many proteins can be identified based on their sequence similarity to a protein of known structure [1, 2], many proteins are not similar to any sequence having a solved protein structure [3]. For such sequences, solution of the protein folding problem is of paramount importance. One would also like to be able to understand the nature of the interactions responsible for the stability and uniqueness of the native conformation. This would allow us to design proteins having desired properties such as enhanced thermostability. Furthermore, many proteins only assume their biologically active conformation on association to multimers. For example, the leucine zipper motif present in eukaryotic gene regulatory proteins [4] is one of the three known modes by which regulatory proteins recognize and bind DNA [5–7], and a number have strong oncogenic potential [5–7]. Obviously, the ability to predict their quaternary structure is of practical importance. Surprisingly, however, in contrast to the substantial body of experimental work [4, 6, 8–28], there are few theoretical studies [14, 23, 29–35] on these important systems [4, 6, 8–28]. Thus, in this review, we describe recent progress on theoretical studies of the folding of coiled coils that are designed to address this need.

Experimental Studies of Coiled Coils

Before discussing recent theoretical studies, it is important to place such work in the context of what is known experimentally about these systems. Coiled coils are found in biological systems ranging from fibrous proteins such as hair and muscle to transcriptional activators [36–38]. They comprise a significant fraction of all known protein sequences [39] and represent an important structural motif found in a number of protein families, including the myosins and tropomyosins, the epimorphins, the kinesin motor domain family, the intermediate filaments, and the leucine zippers [4,

40–42]. All families can be identified by characteristic sequence motifs [43–47]. Their native structure consists of the side-by-side, in-register association of helices, wrapped around each other with a slight left-handed supertwist [48]. Their sequences can be described in terms of a quasirepeating heptet (*abcdefg*), where residues at the *a* and *d* positions occupy the interhelical interface and tend to be mostly hydrophobic [49]. Residues at positions *e* and *g* tend to be occupied by charged residues. It is believed that electrostatic interactions are responsible for dictating whether pairs of helices pack in the parallel or antiparallel orientation [50].

Leucine zippers contain a leucine in the *d* position of the repeating heptet [4, 43] and can form homodimers, e.g., as in GCN4 [21], or heterodimers, e.g., as in Fos-Jun [21, 51]. Ranging in length from about 14 to 45 residues, they are far shorter than other coiled coils [43]. There have been a large number of experimental studies that use this as a model system for testing various protein folding hypotheses [8, 12, 15, 16, 19–21, 52, 53]. Perhaps, the earliest designed leucine zipper is due to Hodges [54]. His sequence design is based on McLachlan's analysis of the tropomyosin sequence [49] and has the form K(LEALEGK)_{*n*}, where *n* = 3, 4, 5. Peptides greater than 28 residues long are more than 90% helical. More recently, Hodges and co-workers [55] have synthesized a series of peptides containing from 9 to 35 residues. These were designed to explore the effects of chain length on the stability of α -helical, two-chain, coiled coils. They found that the model system must contain at least three heptets for the molecule to adopt a coiled coil conformation in solution. Furthermore, these studies indicate that the conformational transition is cooperative and chain length dependent; that is, longer proteins are more stable than shorter ones.

Other studies aimed at identifying the factors that dictate the preference for parallel versus antiparallel association of the chains [50, 56]. Kim and co-workers explored the dimerization specificity of the leucine zipper fragments from c-Fos and c-Jun and found that heterodimerization is favored [20]. The formation of the heterodimer is ascribed to the low stability of dimeric Fos. They also demonstrated that dimerization is necessary for the DNA binding of the entire GCN4 molecule [57]. In 1992, the first high-resolution crystal structure of a leucine zipper coiled coil was determined and found to consist of the parallel, in-register

association of the helices [21]. This confirmed Crick's original prediction of a canonical coiled coil structure made almost 40 years earlier [48]. Additional insights into sequence structure specificity issues emerged from the work of Kim et al., who successfully designed stable heterodimeric structures [58] and studied concerted mutations in GCN4 leucine zippers [12]. Kim and co-workers claim that Asn in the dimerization interface is important because it aids in the parallel association of dimers and destabilizes higher order multimers [52]. They also showed that correlated mutations in the *a* and *d* positions of the GCN4 leucine zipper change the state of association from dimers (Ile in all *a* and Leu in all *d* positions) to trimers (Ile in both *a* and *d* positions) to tetramers (Leu in all *a* and Ile in all *d* positions) [12].

Calorimetry suggests that the GCN4 thermal unfolding transition is well described by a two-state model [59]; i.e., the molecular population is comprised either of fully native or denatured chains [60]. More recently, Holtzer and co-workers have studied the thermal unfolding of a GCN4-like leucine zipper [61] by $^{13}\text{C}\alpha$ labeling at 12 selected sites chosen to avoid overlap in the nuclear magnetic resonance (NMR) spectra. This molecule, termed GCN4-lzK, has the following substitutions from the wild type: R1K, H19K, R25K, and R33K [62]. It exhibits very similar conformational and thermal denaturation properties to the parent molecule. However, the detailed molecular picture is much richer than a simple two-state model of a fully folded molecule in equilibrium with a structureless, denatured state might suggest. Based on the observed chemical shifts, they conclude the following: (1) Urea unfolded GCN4 chains are randomly coiled. (2) In contrast, thermally unfolded chains have significant, but fluctuating, helix content, except at the chain ends. (3) The C-terminus does not have a typical coiled coil structure. (4) All sites except those interior sites flanked by salt bridges show two folded forms. Furthermore, the local unfolding curves extracted on the basis of integrated resonance intensities show that: (a) The sites differ in structural content and melting temperature; thus, there must be more than two molecular conformations in the folded state, with a substantial enthalpy difference between them [63]. GCN4-lzK is thermally less stable than wild type (A. Holtzer, personal communication). Thus, in contrast to the wild-type molecule, it might exhibit a three-state unfolding transition, but this remains to be experimentally established. (b) While signifi-

cant end fraying is evident even at the lowest temperatures, thermal unfolding does not involve progressive end unwinding. (c) Residues 9–16 denature at the lowest temperature. (d) The position of a residue in the heptet does not dictate thermal stability. (e) There is significant unfolding in the dimer prior to its dissociation into monomeric chains. As of yet, this body of experimental information has not been reproduced or rationalized by any theoretical model, but this comprises an ideal system for such studies.

Kinetic unfolding and refolding experiments on the GCN4 leucine zipper were done by Matthews et al. using stopped-flow circular dichroism (CD) as a function of peptide and final GuHCl denaturant concentrations [28]. Unfolding kinetics are characterized by single exponential behavior consistent with a unimolecular species. In contrast, folding kinetics are dependent on both peptide and denaturant concentration and are well described as simple bimolecular association. Furthermore, equilibrium unfolding in the presence of denaturant can be fit by a standard two-state model that is in consonance with the kinetic data. The kinetic studies of Wendt et al. [25] also suggest that the folding of leucine zippers requires the collision of two monomers, each of which might be partially helical, to form a dimeric intermediate which then relaxes to the native conformation. Interestingly, they also find that strand exchange between heterodimeric coiled coils is sequence dependent and occurs on time scales ranging from less than a second to over 30 min [25].

Most thermodynamic data on long coiled coils, in particular tropomyosin, are due to Holtzer and collaborators [64–78]. Tropomyosin is a coiled coil dimer, where each chain contains 284 residues [79] and whose stability is both temperature and pH dependent [74, 75]. Calorimetric studies indicate that there are two peaks associated with thermal denaturation. Even for each peak, the thermodynamic transition is not described by a two-state model [80]. Rather, the conformational transition is highly cooperative and continuous [70]. In contrast, the thermal transition as monitored by CD is monophasic. Further, the N- and C-terminal halves of the molecule exhibit different stabilities [81]. Long excised subsequences of tropomyosin that are greater than 95 residues in length form long coiled coils whose thermal stability is comparable to the entire molecule [82]. When shorter subsequences that are 20–50 residues in length are ex-

amined under identical conditions as in the parent protein, they exhibit very low helix content, little coiled coil formation, and large thermal lability of such helical structure that does form [82]. This is in strong contrast to leucine zipper and short designed peptides, where such short strands exhibit substantial secondary structure and form highly stable coiled coils [12, 54]. This length dependence could be used to calibrate theoretical models.

Theoretical Studies of Coiled Coils

DETAILED ATOMIC MODELS

Because of their inherent structural simplicity, coiled coils represent very well defined systems for elucidating key features of quaternary interactions that are important in all associating protein systems and may also provide insights applicable to globular proteins. However, as mentioned above, there are very few theoretical studies of coiled coil systems. In the 1980s, Skolnick and Holtzer generalized the Zimm–Bragg helix-coil theory [83] to include interhelical interactions [64, 67, 72, 73, 75, 76, 84, 85]. The resulting phenomenological theory was able to predict the quaternary structure of tropomyosin and its fragments over a broad range of temperature and pH. However, these studies were mainly limited to dimeric coiled coils and required that the native structure be assumed. More detailed atomic modeling of coiled coils commenced in 1991 when Nilges and Brunger developed an automated algorithm for the prediction of coiled coil structure based on the assumption that the helices are parallel and in register [29, 31]. They applied this approach to predict a structure of the GCN4 leucine zipper whose backbone-atom coordinate root-mean-square deviation, cRMSD, from the subsequently solved crystal structure was 1.2 Å. They concluded that given the correct registration, the best packing of the hydrophobic residues in the core dictates the detailed geometry. This observation was first made by Crick in 1953 [48]. Novotny et al. [14] have estimated the stability of GCN4, Fos, and Jun leucine zippers from a molecular mechanics calculation and subsequently examined sources of error [86]. Their conclusions suggest that Leu in the *d* position of the canonical coiled coil heptet makes a major contribution to the stability of dimers, whereas residues in the *a* positions are far less important. This idea

is supported by the studies of Zhang and Hermans on a model leucine zipper [87]. In 1991, Tropsha proposed an alternative assignment of residues to the heptet positions in the GCN4 leucine zipper structure, whereby the Leu are assigned to the *a* rather than to the crystallographically correct *d* positions [30]. Most recently, Harbury and Kim have developed a very simple and fast algorithm for predicting structures with assumed ideal coiled coil geometry; however, their approach will not work if the coiled coil is substantially distorted from the canonical structure [35], as has been observed experimentally in a number of cases [88].

REDUCED PROTEIN MODELS

In the case of single-domain proteins and to a far lesser extent for multimeric proteins, reduced protein models have been developed to explore issues of folding thermodynamics, kinetics, and native state structure [33, 89–140]. Here, residues are represented by one or two interaction sites, whose interactions are mainly described by knowledge-based potentials. In such models, the interaction sites either are confined to a set of lattice points [141] or may lie in continuous space [106, 107]. In many reduced protein models, sampling of conformational space is achieved by a Metropolis Monte Carlo (MMC) scheme [142], whose trajectory is equivalent to the solution of a stochastic kinetics master equation [143]. The resulting chain dynamics have been shown to be equivalent to Brownian dynamics in the low-frequency limit [143–145]. The MMC scheme is the most prevalent conformational sampling method in reduced models when an exact enumeration of compact states is not possible. In the asymmetric MMC method, the transition probability from an “old” conformation *i* to a “new” conformation *j* is controlled by the energy difference $\Delta E_{ij} = E_j - E_i$ via

$$p_{ij}^{\text{MMC}} = \min\{1, \exp(-\Delta E_{ij}/kT)\} \quad (1)$$

Here, *k* is Boltzmann’s constant and *T* is the absolute temperature. A disadvantage of this technique is its sensitivity to energy barriers. The key to a successful dynamic Monte Carlo protocol is to include a sufficiently complete move set so that no element of structure is artificially frozen in space.

To elucidate the general aspects of the folding of two-chain, coiled coils and by using a dynamic Monte Carlo algorithm, a simple off-lattice, re-

duced model capable of providing possible folding pathways was developed by Rey et al. [146]. Each amino acid is represented by two spheres, one corresponding to the α -carbon and the other located at the side-chain center of mass. The folding of sequences comprised of repetitive blocks of amino acids has been examined. These simulations assume a local conformational bias toward helical conformations. In addition, the magnitude of the quaternary interactions is coupled to the secondary structure; i.e., the strength is maximal when the pair of interacting residues are helical. Starting from initial random conformations of two separated chains, the assembly process was examined. Folding typically initiates by formation of an intermolecular contact followed by the zipping up of the helices via an on-site construction mechanism [132]. In general, the two chains are out-of-register, and the shift to correct registration occurs by a scissorslike motion of the chains. Finally, they examined the effects of intermolecular crosslinks. Experimental and theoretical studies indicate that introduction of a disulfide bond can enhance coiled coil stability [64, 67, 147]. For singly crosslinked chains, in agreement with previous theoretical conjectures, folding occurs by zipping up from the crosslinks. If disulfide bridges are introduced at both ends of the chain, due to the presence of the second crosslinks, the unfolded portions of the chain have a difficult time developing secondary structure.

Subsequently, Skolnick and co-workers have developed a hierarchical approach to predict the three-dimensional structure and folding pathway of the GCN4 leucine zipper [32], whose crystal structure has been determined by Alber et al. [8]. Folding of the model lattice chains commences from a pair of random, unfolded chains, whose conformations are sampled using MMC simulation. Unlike the studies of Rey et al. [146], no target bias toward the native structure was included. Following spontaneous dimer assembly to a parallel, left-handed, in-register coiled coil, the resulting structures are refined on the lattice to produce a family of native structures whose C α cRMSD from native ranges from 2.3 to 3.7 Å. Full atom models are then built, and the structures are solvated and refined using molecular dynamics with the CHARMM potential [148]. The resulting family of structures is indistinguishable from the native one. The average structure from this family has a backbone heavy atom cRMSD of 0.8 Å from

native, all heavy atoms in the dimerization interface differ by 1.31 Å cRMSD from native, and all heavy atoms differ from the crystal structure by 2.29 Å cRMSD. These studies marked the first time that protein quaternary structures of this quality have been obtained from random, unfolded conformations.

With the assumed energy parameterization for the GCN4 leucine zipper, the predicted helix content of the isolated chains ranges from 30 to 35% and increases to 90% on adoption of the native conformation. The equilibrium $^{13}\text{C}\alpha$ NMR studies of Holtzer and co-workers [61] strongly suggest that the thermally unfolded chains in *aqueous solution* contain significant transient helical secondary structure, and from CD, the helix content is estimated to be about 10–15%. Furthermore, Matthews and co-workers [28], using GuHCl as the denaturant, found little or no helix structure in the monomeric state. Part of the apparent discrepancy between these two studies may reside in the fact that in the Matthews experiments, the final solvent contains GuHCl, a strong denaturant, whereas Holtzer's studies are in aqueous solution without denaturant. Subsequent work by Sosnick and collaborators [149] is largely in accord with the Matthews et al. results. However, Sosnick et al. conclude that the transition state for folding is also devoid of helical structure. Based on these experiments, we conclude that the model has overestimated the intrinsic stability of isolated helices. However, there is a severe sampling problem with MMC in that for the chains to associate and fold in a reasonable amount of computer time, a substantial denatured state helix context is required. More recent studies that eliminate this problem in the context of a solely equilibrium sampling methodology are described below.

In other studies, Sikorski and co-workers [150] have undertaken MMC simulations of proteins designed by DeGrado and co-workers [151, 152]. They studied the original chains, the α_4 tetramer, the α_2 dimer, and the α_1 monomer. At low temperatures, all possible four-helix bundle topologies were recovered. This is a prediction that is in agreement with experiment on the α_1 monomer [153], but for α_4 and α_2 , the validity of this prediction remains to be established. Finally, all folds appear to adopt a molten globule state rather than the structurally well defined state characteristic of naturally occurring globular proteins. For a given topology, the side-chain contacts are very short

lived. This could be caused by the very simple, all-Leu core design of these proteins.

PREDICTION OF COILED COIL CONFORMATIONAL EQUILIBRIA

Harbury et al. [12] have examined the shift in equilibrium among parallel, coiled coil dimers, trimers, and tetramers associated with varying the identity of the hydrophobic residues in the coiled coil heptad repeat. To treat these systems, Vieth et al. have developed two combined statistical mechanical/simulation-based approaches [23, 34]. Both treatments assume a manifold of accessible chain conformations (both parallel and antiparallel orientations of the chains are allowed), where the equilibrium constant between all allowed species is calculated. In their first approach, they were unable to treat the monomeric state [34]. Subsequently, the algorithm was generalized to also include monomeric species [23]. They then calculated the dimer, trimer, and tetramer equilibrium of GCN4 and seven of its mutants, which were studied by Kim et al. [12]. Over the entire concentration range, agreement with experiment for five of the eight sequences is found, and over a portion of the concentration range, as is agreement for another two sequences. Local, intrinsic secondary structure preferences and side-chain entropy act to favor lower order multimers, whereas quaternary interactions favor higher order multimers. They next applied the method to study the multimeric equilibria of fragments of the GCN4 leucine zipper and found that monomers should dominate for a fragment comprised of residues 11–33 in the parent molecule, whereas dimers are predicted for a fragment comprised of residues 8–30. These results are also in accord with experiment [15]. For a fragment comprised of residues 4–26, they also predicted that mostly dimers should be found at high concentrations. This prediction awaits experimental validation.

Next, they examined other leucine zippers, namely the DNA binding proteins Fos and Jun. In agreement with experiment, in solutions containing only Jun, dimers are the equilibrium conformation [20]. This indicates that the energy functions have captured some of the features of coiled coils. Encouragingly, also in agreement with experiment, in equimolar mixtures of Fos and Jun, heterodimers dominate. Finally, they examined the coil–Ser sequence designed by Eisenberg, De-Grado, and co-workers. Again, in agreement with

experiment, an antiparallel, three-helix bundle is predicted to be the native conformation [154].

Above, we described two methods that have been developed to calculate equilibrium constants: One could treat all degrees of chain association and the other was limited to dimers and higher order species [23, 33, 34]. While both were reasonably successful at reproducing the experiments, they also invoked a number of approximations of uncertain validity. Most importantly, both approaches a priori assumed the manifold of conformations that the chains can adopt. This not only included the state of association, but also restricted the chains to parallel and antiparallel arrangements of the helices. Registration is not enforced, but the implicit assumption is that a given arrangement of the helices is at least metastable. Ideally, at least for smaller systems, one would like to relax this restriction and simply assert that there is a given degree of association of the chains and then predict the conformations the system adopts, rather than starting the system in an assumed manifold of relatively specific (and native-like) conformational states. In addition, these approaches make a number of purely technical approximations that might limit their range of validity. For example, to determine the denatured state partition function, all long-range interactions beyond the fourth neighbor down the chain [23] were neglected. This approximation should work reasonably well if the denatured state is relatively expanded or if the chains are short, but otherwise, it may be severely in error. For a given multimeric state, they calculated the average energy from the Monte Carlo simulation and estimated the entropy using a partition function constructed from a series of overlapping triplets of virtual bond fragments that are observed at least once; thus, we turn to a new approach that eliminates these weaknesses.

In a recent series of papers on the origin of cooperativity in protein folding, Hao and Scheraga have employed the entropy sampling Monte Carlo (ESMC) method to explore conformational space in the chess knight lattice protein model supplemented by single, united atom side chains [155–157]. The ESMC method was originally proposed by Lee [158] in the context of a simple Ising model and is closely related to the multicanonical MC technique of Berg et al. [159]. The ESMC model generates a distribution of states that is controlled by the conformational entropy as a function of the energy of a particular conformation E_i . In ESMC simulation, the transition probability

between states i and j is

$$p_{ij}^{\text{ESMC}} = \min\{1, \exp(-\Delta S_{ij}/k)\} \quad (2)$$

with ΔS_{ij} the entropy difference between energy levels i and j . At the beginning of the simulation, the entropy is not known. However, an estimate, $J(E)$, for the entropy $S(E)$ can be iteratively generated from a density-of-states energy histogram, $H(E)$. The k th iteration consists of an ESMC simulation run, with $S(E)$ approximated by $J_{k-1}(E)$. Here,

$$J_k(E) = J_{k-1}(E) + \ln(\max\{1, H_k(E)\}) \quad (3)$$

After a sufficient number of runs, all states are sampled with the same frequency. Then, the histogram of $H(E)$ becomes flat, and $J(E) + \text{constant}$ approaches the true value, $S(E)$. The ESMC method offers the crucial advantage over MMC simulation that one knows whether or not the simulation has converged over the sampled energy range, and, if so, the free energy of the system (within a constant) is determined. However, the problem with an ESMC simulation is the computational cost required to construct the entropy-versus-energy curve. In a series of papers on the origin of the cooperativity of protein folding in designed β -barrels and other real proteins [137, 160], Kolinski et al. described a means of increasing the rate of convergence. The basic idea is to first run a standard dynamic Monte Carlo folding simulation to provide a library of "seed" structures. Then, periodically, one randomly selects one of these conformations and introduces it into the conformational pool used to construct the entropy-versus-energy curve. In practice, this technique works quite well to speed up the rate of convergence.

Until recently, ESMC has been only applied to monomeric systems [137, 155–157, 160]. In what follows, we describe a generalization of the ESMC approach that allows one to calculate the equilibrium constant among a set of assumed states of oligomerization. We first consider the partition function of an isolated single chain and separate the translational, orientational, and internal degrees of freedom [23, 33, 34, 161, 162]. To do this, we fix the first $C\alpha$ in space. Then, ESMC simulation is performed subject to the restraint that the first two $C\alpha$ virtual bond vectors of the first chain are fixed. Averaging over the set of all the allowed first two vectors of the chain provides the overall rotational contribution to the partition function,

which, in practice, is very close to $8\pi^2$ (the value obtained for a chain in continuous space). For single chains, this ESMC procedure samples the internal entropy from which the total internal monomeric, configurational partition function associated with the $3N - 6$ internal degrees of freedom of the chain, $Z_{\text{int-conf}, M}$, can be constructed to within a constant. Here, N is the total number of $C\alpha$'s and side-chain centers of mass in a monomer.

Higher order multimeric species are defined by the requirement that there be at least one intermolecular side-chain contact. For definiteness, we focus on dimers, but generalization to higher order multimers is straightforward. For chain 1, fix the first vertex and the first two $C\alpha$ virtual bonds. Again, one performs ESMC simulation on this system to obtain an estimate of the configurational entropy, $S(E)$. Later, we will average over the set of allowed vectors to obtain the overall rotational entropic contribution for the dimer. Within a constant, C_L , the total internal configurational partition function of species L is

$$Z_{\text{int-conf}, L} = \sum_E \exp(-\beta E + S_L(E)) \quad (4)$$

where $S_L(E)$ is the entropy of species L with an energy E . Here, $L = D$.

Conceptually, the entropy of a dimer, S_D , can be divided into three contributions:

1. The translational entropy of bead 1 in chain 2, which is related to the average volume accessible to this bead, $\langle v_1 \rangle$, by

$$S_{\text{trans}, D} = \ln \langle v_1 \rangle \quad (5a)$$

2. The rotational entropy of chain 2, which is related to the number of orientational conformations of the first two $C\alpha$ bond vectors of chain 2 in the dimer, Ω , by

$$S_{\text{rot}, D} = \ln \Omega \quad (5b)$$

3. The conformational entropy arising from the internal configurations of both chains in the dimer, $S_{\text{int}, D}$.

Equations (5a) and (5b) can be calculated from the manifold of states once the $S_D(E)$ -versus- E curve is calculated from ESMC simulation. Then, for any internal thermodynamic or conformational

property q ,

$$\langle q \rangle = \frac{\sum_E q(E) \exp(-\beta E + S_L(E))}{Z_{\text{int-conf}, L}} \quad (6)$$

At this juncture, to within a constant, we have calculated the total internal partition functions of the monomers and dimers. For a given degree of chain association, determination of this constant is not necessary to determine thermodynamic averages. If, however, one wishes to calculate the equilibrium constant between different multimeric species, then all must have the same reference state. Then, the equilibrium constant follows from

$$K_{MD} = \frac{(D)}{(M)^2} = \frac{\langle v_1 \rangle \Omega Z_{\text{int-conf}, D}}{\sigma_D 8\pi^2 Z_{\text{int-conf}, M}^2} \quad (7)$$

σ_D is the symmetry number of the dimer. It equals 1 if the system is a heterodimer and 2 if it is a homodimer, (L) is the concentration of species L .

The only problem remaining is to place the monomer and dimer in the same reference state.

We first observe that for the internal configurational entropy in the limit of very large energies,

$$S_D(E) - S_{\text{trans}, D} - S_{\text{rot}, D} = 2S_M(E_M) \quad (8)$$

(and for the GCN4 leucine zipper we have shown this to be true). This occurs because the rotational and translational contributions are subtracted out in the calculation of $S_M(E_M)$, the internal configurational entropy of a monomer having an energy per chain, E_M . Note here that $2E_M = E$, the energy in the dimer. By demanding that Eq. (8) hold in the large energy limit, we can shift the entropies of the monomer and dimer so that they have the same reference state. In practice, $S_D(E) - S_{\text{trans}, D} - S_{\text{rot}, D} - 2S_M(E_M)$ versus E is plotted. In the energy region where this difference approaches a constant independent of E , we simply adjust the constant so that this difference equals zero. Then, both the monomers and dimers have the same reference state and their relative internal entropy differences can be calculated.

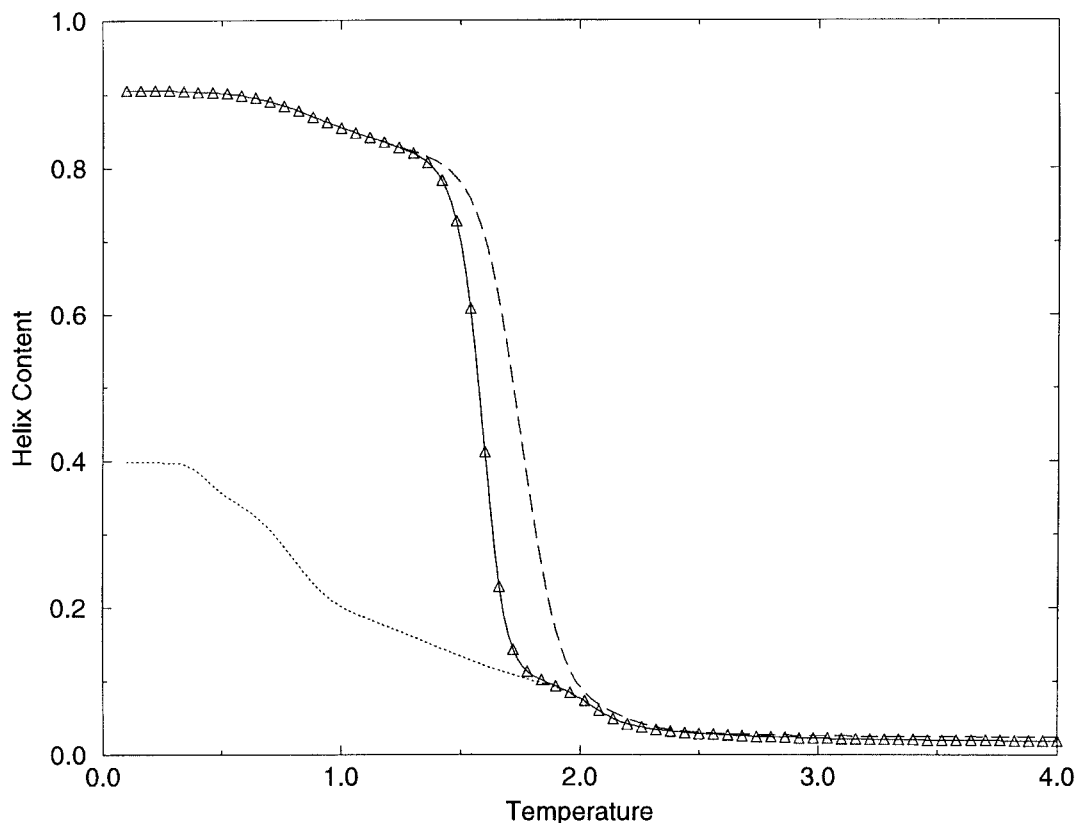


FIGURE 1. For the GCN4 leucine zipper at a concentration of $300 \mu M$, plots of calculated overall, dimer, and monomer helix contents versus reduced temperature are shown as solid, dashed, and dotted curves, respectively.

To this point, we made no restriction about whether or not the dimers associate in a parallel or antiparallel fashion. In principle, for a given degree of chain association, the ESMC protocol itself should allow us to determine the relative importance of the parallel-versus-antiparallel orientations of the chains. However, one could cluster the conformations by whatever criterion one wishes and then calculate the relative equilibrium between such subsets of species. The advantage of the proposed methodology over our previous one is that the dominant structural species, if any, will be predicted from the simulation and need not be specified in advance.

Using an ESMC approach and by simply reducing the scale factors of the terms describing secondary structure propensities and hydrogen bonding in the original treatment of Vieth et al. [23, 33, 34], we have been able to successfully fold to the native conformation of the GCN4 leucine zipper under conditions where the overall helix content of the monomers in the denatured state in the ther-

mal transition region is below 15%. A plot of the predicted monomeric, dimeric, and overall helix content of a concentration of $300 \mu\text{M}$ is shown in Figure 1. Furthermore, the calculated conformation transition is well described by a two-state model. This is confirmed in Figure 2, where we plot the ratio of the van't Hoff to calorimetric enthalpies over a temperature range that covers the transition. Since this ratio is close to unity, a two-state folding model is valid [163]. Thus, in this revised model, all qualitative features are in agreement with experiment [59]. Furthermore, the simulations suggest that the reason the overall conformational transition is two state is that the chains dissociate before the majority of the helix content is lost. However, within the dimer, the conformational transition is continuous. Finally, these simulations point out the power of the ESMC method of conformational sampling. Using classical MMC simulation, with this greatly reduced helix content, we would be unable to run the simulations long enough to obtain the native fold. Here, elucidation

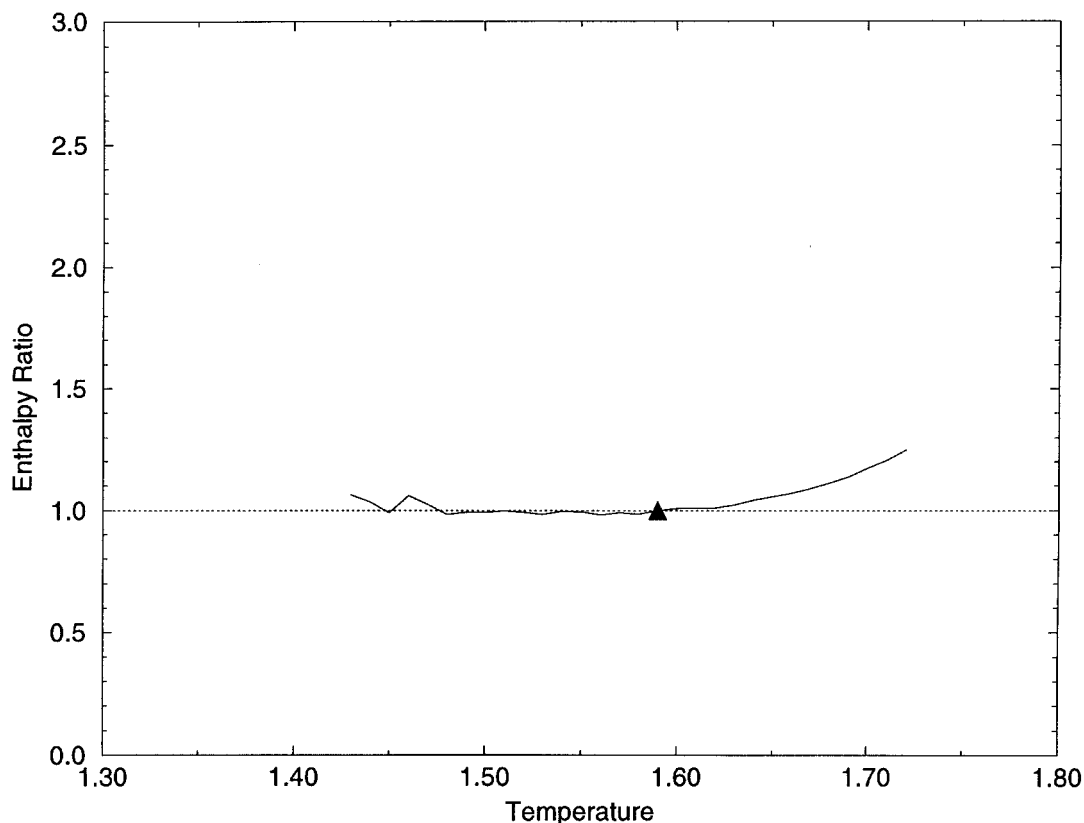


FIGURE 2. For the GCN4 leucine zipper at a concentration of $300 \mu\text{M}$, plot of the ratio of the van't Hoff to calorimetry enthalpy as a function of reduced temperature is shown. The transition midpoint is at a reduced temperature of 1.59.

of the full thermodynamic behavior in this model system has been possible, and agreement with experiment has been demonstrated.

Outlook for Future Progress

Due to their greater complexity, theoretical studies of multimeric proteins have not yet achieved the level of sophistication of their single-chain counterparts. Nevertheless, leucine zipper multimeric systems are now within the range of contemporary theoretical approaches. Because of their geometric simplicity, they can provide numerous insights into the nature of the native state, the mechanism(s) of multimeric protein assembly, and the thermodynamics of protein conformational transitions. Among the issues that still remain are: Can reduced protein models describe the manifold of nonnative conformations, given that their knowledge-based potentials are derived from native conformations, and can these models describe the wealth of experimental data on higher order multimers as well as more complex and not necessarily helical proteins? To address these types of questions, at least in an equilibrium context, ESMC simulation is the method of choice. It provides an internal check of convergence and, with the techniques described here, a direct estimate of the partition function. Furthermore, it is well suited for coarse-grained parallel computers. Thus, the next several years should witness an increase in our understanding of the physical properties of coiled coils in particular and aspects of protein quaternary structure in general. Such an understanding is very important precisely because the function of many biological proteins is crucially dependent on the formation of correct quaternary structure.

ACKNOWLEDGMENT

This article is dedicated to the memory of Bernard Pullnam. The work was supported in part by National Institutes of Health Research Resource Grant P41 RR12255.

References

- Sali, A.; Overington, J. P.; Johnson, M. S.; Blundell, T. L. *TIBS* 1990, 15, 235–250.
- Ring, C. S.; Cohen, F. E. *FASEB* 1994, 7, 783–790.
- Holm, L.; Sander, C. *Science* 1996, 273, 595–602.
- Landshultz, W. H.; Johnson, P. F.; McKnight, S. L. *Science* 1988, 240, 1759–1764.
- Brent, R.; Ptashne, M. *Cell* 1985, 43, 729–736.
- Bush, S. J.; Sassone-Corsi, P. *Trends Genet* 1990, 6, 36–40.
- Harrison, S. *Nature* 1991, 353, 715–719.
- Alber, T. *Curr Opin Genet Devel* 1992, 2, 205–209.
- Ellenberger, T. E.; Brandl, C. J.; Struhl, K.; Harrison, S. C. *Cell* 1992, 71, 38–45.
- Goodman, E. M.; Kim, P. S. *Biochemistry* 1991, 30, 11615–11620.
- Hai, T.; Liu, F.; Coukos, W. J.; Green, M. R. *Genes Devel* 1989, 3, 2083–2090.
- Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T. *Science* 1993, 262, 1401–1407.
- Holzman, B. L.; Merritt, S. E.; Guang, F. *J Biol Chem* 1994, 269, 30808–30817.
- Krystek, Jr., S. R.; Bruccoleri, R. E.; Novotny, J. *Int J Pept Prot Res* 1991, 38, 229–236.
- Lumb, K. J.; Carr, C. M.; Kim, P. S. *Biochemistry* 1994, 33, 7361–7367.
- Lumb, K. J.; Kim, P. S. *Science* 1995, 268, 436–439.
- Mattson, J.; Soderman, E.; Svenson, M.; Borkird, C.; Engstroem, P. *Plant Mol Biol* 1992, 18, 1019–1022.
- Nishizawa, M.; Kataoka, K.; Goto, N.; Fujiwara, K. T.; Kawai, S. *Proc Natl Acad Sci USA* 1989, 86, 7711–7715.
- O'Neil, K. T.; Hoess, R. H.; DeGrado, W. F. *Science* 1990, 249, 774–778.
- O'Shea, E.; Rutkowski, R.; Stafford III, W. F.; Kim, P. S. *Science* 1989, 245, 646–648.
- O'Shea, E.; Klemm, J. D.; Kim, P. S.; Alber, T. *Science* 1991, 254, 539–544.
- Thomas, D.; Jacquemin, I.; Surdin-Kerjan, Y. *Mol Cell Biol* 1992, 12, 1719–1727.
- Vieth, M.; Kolinski, A.; Skolnick, J. *Biochemistry* 1996, 35, 955–967.
- Vinson, C. R.; Sigler, P. B.; McKnight, S. L. *Science* 1989, 246, 911–916.
- Wendt, H.; Bereger, C.; Baici, A.; Thomas, R. M.; Bosshard, H. R. *Biochemistry* 1995, 34, 4097–4107.
- White, M. J.; Hirsch, J. P.; Henry, S. A. *J Biol Chem* 1991, 266, 863–872.
- Zhu, B. Y.; Zhou, N. E.; Kay, C. M.; Hodges, R. S. *Prot Sci* 1993, 2, 383–394.
- Zitzewitz, J. A.; Bilsel, O.; Luo, J.; Jones, B. E.; Matthews, C. R. *Biochemistry* 1995, 34, 12812–12819.
- Nilges, M.; Brunger, A. T. *Prot Eng* 1991, 4, 649–659.
- Tropsha, A.; Bowen, P. J.; Brown, F. K.; Kizer, J. S. *Proc Natl Acad Sci USA* 1991, 88, 9488–9492.
- Nilges, M.; Brunger, A. T. *Proteins* 1993, 15, 133–146.
- Vieth, M.; Kolinski, A.; Brooks III, C. L.; Skolnick, J. *J Mol Biol* 1994, 237, 361–367.
- Vieth, M.; Kolinski, A.; Brooks III, C. L.; Skolnick, J. *J Mol Biol* 1995, 251, 448–467.
- Vieth, M.; Kolinski, A.; Skolnick, J. *J Chem Phys* 1995, 102, 6189–6193.
- Harbury, P. H.; Tidor, B.; Kim, P. S. *Proc Natl Acad Sci USA* 1995, 92, 8408–8412.

36. Cohen, C.; Parry, D. A. D. *Trends Biochem Sci* 1986, 11, 245–248.
37. Bairoch, A. *Nucl Acid Res* 1991, 19, 2247–2249.
38. Hodges, R. S. *Curr Biol* 1992, 2, 122–124.
39. Bairoch, A. *Nucl Acid Res* 1997, 25, 31–36.
40. McLachlan, A. D.; Karn, J. *J Mol Biol* 1983, 164, 605–626.
41. Parry, D. A.; Fraser, R. D. B. *Int J Biol Macromol* 1985, 7, 203–213.
42. Doolittle, R. F.; Cassman, K. G.; Cottrell, B. A.; Friezner, S. J.; Takagi, T. *Biochemistry* 1977, 16, 1710–1715.
43. Bairoch, A. *Department de Biochimie Medicale; Universite de Geneva: Geneva*, 1990.
44. Berger, B.; Wilson, D. B.; Wolf, E.; Tonchev, T.; Milla, M.; Kim, P. S. *Proc Natl Acad Sci USA* 1995, 92, 8259–8263.
45. Lupas, A.; Van Dyke, M.; Stock, J. *Science* 1991, 252, 1162–1164.
46. Hirst, J. D.; Vieth, M.; Skolnick, J.; Brooks III, C. L. *Prot Eng* 1996, 9, 657–662.
47. Wolf, E.; Kim, P. S.; Berger, B. *Prot Sci* 1997, 6, 1179–1189.
48. Crick, F. H. C. *Acta Cryst* 1953, 6, 689–697.
49. McLachlan, A. D.; Stewart, M. *J Mol Biol* 1975, 98, 298–304.
50. Monera, O. D.; Kay, C. M.; Hodges, R. S. *Biochemistry* 1994, 33, 3862–3871.
51. Smeal, T.; Angel, P.; Meek, J.; Karin, M. *Genes Devel* 1989, 3, 2091–2100.
52. Lumb, K. J.; Kim, P. S. *Biochemistry* 1995, 34, 8642–8648.
53. Pu, T. W.; Struhl, K. *Nucl Acid Res* 1993, 18, 4348–4355.
54. Hodges, R. S.; Saund, A. S.; Chong, P. C. S.; St-Pierre, S. A.; Reid, R. E. *J Biol Chem* 1981, 256, 1214–1224.
55. Su, J. Y.; Hodges, R. S. *Biochemistry* 1994, 33, 15501–15510.
56. Monera, O. D.; Zhou, N. E.; Lavigne, P.; Kay, C. M.; Hodges, R. S. *J Biol Chem* 1996, 271, 3995–4001.
57. Talanian, R. V.; McKnight, C. J.; Kim, P. S. *Science* 1990, 249, 769–770.
58. O'Shea, E. K.; Lumb, K. J.; Kim, P. S. *Curr Biol* 1993, 3, 658–667.
59. Kenar, K. T.; Garcia-Moreno, B.; Freire, E. *Prot Sci* 1995, 4, 1934–1938.
60. *Physical Basis of the Stability of the Folded Conformation of Proteins*, Privalov, E. I., Ed.; W. H. Freeman: New York, 1992.
61. Holtzer, M. E.; Lovett, E. G.; d'Avignon, D. A.; Holtzer, A. *Biophys J* 1997, 73, 1031–1041.
62. Lovett, E.; d'Avignon, D. A.; Holtzer, M. E.; Braswell, E. H.; Zhu, D.; Holtzer, A. *Proc Natl Acad Sci* 1996, 93, 1781–1785.
63. d'Avignon, D.; Bretthorst, G.; Holtzer, M.; Holtzer, A. *Biophys J* 1998, 74, 3190–3197.
64. Duffy, P.; Skolnick, J.; Holtzer, A. *Res Commun* 1986, 141, 394–398.
65. Holtzer, M. E.; Askins, K.; Holtzer, A. *Biochemistry* 1986, 25, 1688–1692.
66. Holtzer, A.; Skolnick, J. *Macromolecules* 1986, 19, 1769–1781.
67. Skolnick, J.; Holtzer, A. *Biochemistry* 1986, 25, 6192.
68. Skolnick, J. *Macromolecules* 1986, 19, 1153–1166.
69. Holtzer, A.; Skolnick, J. *Biopolymers* 1988, 27, 87–96.
70. Holtzer, A. M.; Holtzer, M. E.; Skolnick, J. In *Protein Folding. Deciphering the Second Half of the Genetic Code*, Gierash, L. M.; King, J., Eds.; American Association for the Advancement of Science: Washington, DC, 1990; Vol. 4, pp 2401–2411.
71. Holtzer, M. E.; Holtzer, A. *Biopolymers* 1990, 30, 985–993.
72. Skolnick, J.; Holtzer, A. *Macromolecules* 1982, 15, 303–314.
73. Skolnick, J.; Holtzer, A. *Macromolecules* 1982, 15, 812–821.
74. Holtzer, M. E.; Holtzer, A.; Skolnick, J. *Macromolecules* 1983, 16, 462–465.
75. Skolnick, J.; Holtzer, A. *Macromolecules* 1983, 16, 1548–1550.
76. Skolnick, J. *Macromolecules* 1983, 16, 1069.
77. Holtzer, M. E.; Breiner, T.; Holtzer, A. *Biopolymers* 1984, 23, 1811–1833.
78. Skolnick, J.; Holtzer, A. *Macromolecules* 1985, 18, 1549–1559.
79. Fraser, R. D. B.; MacRae, T. P. *Conformation in Fibrous Proteins and Related Synthetic Polypeptides*; Academic: London, 1973.
80. Potekhin, S.; Privalov, P. L. *J Mol Biol* 1982, 159, 519–535.
81. Pato, M.; Mak, A.; Smillie, L. B. *J Biol Chem* 1981, 256, 593–601.
82. Holtzer, M. E.; Crimmins, D. L.; Holtzer, A. M. *Biopolymers* 1995, 35, 125–136.
83. Zimm, B. H.; Bragg, J. K. *J Chem Phys* 1959, 31, 526–535.
84. Skolnick, J. *Macromolecules* 1984, 17, 645–658.
85. Skolnick, J. *Macromolecules* 1984, 17, 2153.
86. Shen, L.; Bruccoleri, R.; Krystek, S.; Novotny, J. *Biophys* 1996, 70, 1096–1104.
87. Zhang, L.; Hermans, J. *Proteins* 1993, 16, 384–392.
88. Gonzales, Jr., L.; Brown, R. A.; Richardson, D.; Alber, T. *Nat Struct Biol* 1996, 3, 1002–1009.
89. Levitt, M. *J Mol Biol* 1992, 226, 507–533.
90. Levitt, M.; Warshel, A. *Nature* 1975, 253, 694–698.
91. Levitt, M. *Nature* 1975, 253, 694–698.
92. Levitt, M. *J Mol Biol* 1975, 104, 59–107.
93. Levitt, M. *Curr Opin Struct Biol* 1991, 1, 224–229.
94. Hinds, D. A.; Levitt, M. *Proc Natl Acad Sci USA* 1992, 89, 2536–2540.
95. Park, B. H.; Levitt, M. *J Mol Biol* 1995, 493–507.
96. Go, N.; Taketomi, H. *Proc Natl Acad Sci USA* 1978, 75, 559–563.
97. Ueda, Y.; Taketomi, H.; Go, N. *Biopolymers* 1978, 17, 1531–1548.
98. *Protein Folding*, Go, N.; Abe, H.; Mizuno, H.; Taketomi, H., Eds.; Elsevier/North-Holland: Amsterdam, 1980.
99. Chan, H. S.; Dill, K. A. *Macromolecules* 1989, 22, 4559–4573.
100. Lau, K. F.; Dill, K. A. *Macromolecules* 1989, 22, 3986–3997.
101. Chan, H. S.; Dill, K. A. *Proc Natl Acad Sci USA* 1990, 87, 6388–6392.
102. Dill, K. A. *Biochemistry* 1990, 29, 7133–7155.
103. Dill, K. A. *Current Biol* 1993, 3, 99–103.
104. Dill, K. A.; Bromberg, S.; Yue, K.; Fiebig, K. M.; Yee, D. P.; Thomas, P. D.; Chan, H. S. *Prot Sci* 1995, 4, 561–602.
105. Dill, K. A.; Yue, K. *Prot Sci* 1996, 5, 254–261.

106. Monge, A.; Friesner, R. A.; Honig, B. *Proc Natl Acad Sci USA* 1994, 91, 5027–5029.
107. Monge, A.; Lathrop, E. J. P.; Gunn, J. R.; Shenkin, P. S.; Friesner, R. A. *J Mol Biol* 1995, 247, 995–1012.
108. Friesner, R.; Gunn, J. R. *Ann Rev Biophys Biomol Struct* 1996, 25, 315–342.
109. Knapp, E. W. *J Comput Chem* 1992, 13, 793–798.
110. Shakhnovich, E.; Farztdinov, G.; Gutin, A. M. *Phys Rev Lett* 1991, 67, 1665–1668.
111. Karplus, M.; Shakhnovich, E. In *Protein Folding*, Creighton, T. E., Ed.; W. H. Freeman: New York, 1992; pp 127–196.
112. Shakhnovich, E. I.; Gutin, A. M. *Proc Natl Acad Sci USA* 1993, 90, 7195–7199.
113. Sali, A.; Shakhnovich, E. I.; Karplus, M. *J Mol Biol* 1994, 235, 1614–1636.
114. Shakhnovich, E. I. *Phys Rev Lett* 1994, 72, 3907–3910.
115. Walther, D.; Eisenhaber, F.; Argos, P. *J Mol Biol* 1996, 255, 536–553.
116. Dandekar, T.; Argos, P. *J Mol Biol* 1996, 256, 645–660.
117. Milik, M.; Skolnick, J. *Proc Natl Acad Sci USA* 1992, 89, 9391–9395.
118. Kolinski, A.; Skolnick, J.; Yaris, R. *Biopolymers* 1987, 26, 937–962.
119. Kolinski, A.; Skolnick, J.; Yaris, R. *J Chem Phys* 1986, 85, 3585–3597.
120. Kolinski, A.; Skolnick, J. *Proc Natl Acad Sci USA* 1986, 83, 7267–7271.
121. Skolnick, J.; Kolinski, A.; Yaris, R. *Proc Natl Acad Sci USA* 1988, 85, 5057–5061.
122. Sikorski, A.; Skolnick, J. *Proc Natl Acad Sci USA* 1989, 86, 2668–2672.
123. Sikorski, A.; Skolnick, J. *Biopolymers* 1989, 28, 1097–1113.
124. Skolnick, J.; Kolinski, A.; Yaris, R. *Biopolymers* 1989, 28, 1059–1095.
125. Skolnick, J.; Kolinski, A.; Yaris, R. *Proc Natl Acad Sci USA* 1989, 86, 1229–1233.
126. Skolnick, J.; Kolinski, A. *Annu Rev Phys Chem* 1989, 40, 207–235.
127. Sikorski, A.; Skolnick, J. *J Mol Biol* 1990, 212, 819–836.
128. Sikorski, A.; Skolnick, J. *J Mol Biol* 1990, 215, 183–198.
129. Skolnick, J.; Kolinski, A. *Science* 1990, 250, 1121–1125.
130. Skolnick, J.; Kolinski, A.; Sikorski, A. *Chem Des Automat News* 1990, 5, 1–20.
131. Kolinski, A.; Milik, M.; Skolnick, J. *J Chem Phys* 1991, 94, 3978–3985.
132. Skolnick, J.; Kolinski, A. *J Mol Biol* 1991, 221, 499–531.
133. Skolnick, J.; Kolinski, A. *J Mol Biol* 1992, 223, 499–531.
134. Kolinski, A.; Skolnick, J. *Proteins* 1994, 18, 353–366.
135. Kolinski, A.; Skolnick, J. *Proteins* 1994, 18, 338–352.
136. Kolinski, A.; Galazka, W.; Skolnick, J. *J Chem Phys* 1995, 103, 10286–10297.
137. Kolinski, A.; Galazka, W.; Skolnick, J. *Proteins* 1996, 26, 271–287.
138. Ortiz, A.; Kolinski, A.; Skolnick, J. *Proc Natl Sci USA* 1998, 95, 1020–1025.
139. Ortiz, A.; Kolinski, A.; Skolnick, J. *J Mol Biol* 1998, 277, 419–448.
140. Ortiz, A.; Kolinski, A.; Skolnick, J. *Proteins* 1998, 30, 287–294.
141. Godzik, A.; Kolinski, A.; Skolnick, J. *J Comp Chem* 1993, 14, 1194–1202.
142. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. *J Chem Phys* 1953, 51, 1087–1092.
143. Binder, K. Paper; Institut für Physik, Johannes Gutenberg-Universität: Mainz, 1991.
144. Verdier, P. H.; Stockmayer, W. H. *J Chem Phys* 1962, 36, 227–231.
145. Kolinski, A. K.; Skolnick, J. *Lattice Models of Protein Folding, Dynamics and Thermodynamics*; R. G. Landes: Austin, 1996.
146. Rey, A.; Skolnick, J. *J Chem Phys* 1994, 100, 2267–2276.
147. Skolnick, J. *Theory of the Helix-Coil Transition in Doubly Crosslinked, Two-Chain, Coiled Coils. A Globular Protein Model*; North-Holland: Amsterdam, 1986.
148. Brunger, A. T.; Clore, G. M.; Gronenborn, A. M.; Karplus, M. *Proc Natl Acad Sci USA* 1986, 83, 3801–3805.
149. Sosnick, T. R.; Jackson, S.; Wilk, R. R.; Englander, S. W.; DeGrado, W. F. *Proteins* 1996, 24, 427–432.
150. Sikorski, A.; Kolinski, A.; Skolnick, J. *Biophys J* 1998, 75, 72–105.
151. Regan, L.; DeGrado, W. F. *Science* 1988, 241, 976–978.
152. Ho, S. P.; DeGrado, W. F. *J Am Chem Soc* 1987, 109, 6751–6758.
153. Handel, T. M.; Williams, S. A.; DeGrado, W. F. *Science* 1993, 261, 879–885.
154. Lovejoy, B.; Choe, S.; McRorie, D. K.; DeGrado, W. F.; Eisenberg, D. *Science* 1993, 259, 1288–1293.
155. Hao, M.-H.; Rackovsky, S.; Liwo, A.; Pinkus, M. R.; Scheraga, H. A. *Proc Natl Acad Sci USA* 1992, 89, 6614–6618.
156. Hao, M.-H.; Scheraga, H. A. *J Phys Chem* 1994, 98, 4940–4948.
157. Hao, M.-H.; Scheraga, H. A. *J Phys Chem* 1994, 98, 9882–9893.
158. Lee, J. *Phys Rev Lett* 1993, 71, 211–214.
159. Berg, B. A.; Neuhaus, T. *Phys Rev Lett* 1991, 68, 9–12.
160. Kolinski, A.; Galazka, W.; Skolnick, J. *J Chem Phys* 1998, 108, 2608–2617.
161. Herschbach, D. R. *J Chem Phys* 1959, 31, 1652–1661.
162. Davidson, N. *Statistical Mechanics*; McGraw-Hill: New York, 1962.
163. Kidokoro, S.; Wada, A. *Biopolymers* 1987, 26, 213–229.