# Heteropolymer freezing and design: Towards physical models of protein folding

Vijay S. Pande

*Chemistry Department, Stanford University, Stanford, California 94305-5080*

Alexander Yu. Grosberg

*Department of Physics, University of Minnesota, Minneapolis, Minnesota 55455*

Toyoichi Tanaka

*Department of Physics and Center for Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139*

Protein folding has become one of the most actively studied problems in modern molecular biophysics. Approaches to the problem combine ideas from the physics of disordered systems, polymer physics, and molecular biology. Much can be learned from the statistical properties of model heteropolymers, the chain molecules having different monomers in irregular sequences. Even in highly evolved proteins, there is a strong random element in the sequences, which gives rise to a statistical ensemble of sequences for a given folded shape. Simple analytic models give rise to phase transitions between random, glassy, and folded states, depending on the temperature $T$ and the design temperature $T^{\mathrm{des}}$ of the ensemble of sequences. Besides considering the analytic results obtainable in a random-energy model and in the Flory mean-field model of polymers, the article reports on confirming numerical simulations.

## CONTENTS

## I. INTRODUCTION: PROTEIN FOLDING AS A PHYSICAL PROBLEM

The subject matter of the present review is truly multidisciplinary. Although we concentrate almost exclusively on the statistical physics of protein folding we begin with a somewhat broader view of the problems and a few historical remarks. For the reader unfamiliar with proteins, we provide some basic definitions in Sec. I.B and a list of relevant terms in Appendix A.

### A. Historical background and an overview of the field

Protein folding is a field at the intersection of molecular biology, polymer physics, and the physics of disordered systems. In this Introduction, we briefly review the historical background and conceptual foundations of the questions addressed in this review.

#### 1. Molecular biology perspective

The most impressive achievement of the early years of molecular biology may well be the beautiful way in which hereditary information was found to be stored in the linear, one-dimensional sequence of the DNA base pairs. Once the main scheme had been established, with its DNA→RNA→protein concept of information processing, the emerging picture described the decoding of one linear (one-dimensional) code into another linear code.

While the details of the genetic code and the complex cell machinery involved in real protein synthesis may be complex, it is at least algorithmically simple to imagine sequential processing in which codons (triplets of nucleotides) are read one by one and the corresponding amino acids are added to the growing protein chain, each independent of the previous or subsequent parts of the sequence.

This "one-dimensional view" has permeated much of molecular biology and is manifested in various suggestions of different new "codes." Yet biology offers striking examples in which neither storage nor processing of information is to any degree one dimensional. Moreover, there are many similar examples of processes, such as reading, in which the input is a one-dimensional sequence (ordered because the process proceeds linearly in time), while the output, which is an acquired idea (or understanding), is something of a completely different nature and by no means one dimensional.

Most of the catalytic and other chemical functions in the living cell are performed by proteins. The question then becomes how genetic information, translated into the protein sequence, controls protein function. It was long known (or at least believed) that protein functions are defined by a particular three-dimensional configura-

tion uniquely characteristic for each protein species (Mirsky and Pauling, 1936), and thus the problem is reduced to a deceptively simple question: How does the sequence of amino acids in a protein chain control the fold of the chain in space? About 25 years ago, C. Anfinsen showed in a famous experiment that denatured (that is, rendered incapable of functioning by elevated temperature and/or inappropriate pH/salt conditions) protein molecules are able to restore their function upon return to the proper conditions in solution, without any assistance from cellular machinery (Anfinsen, 1973); moreover, it was later shown that denaturation is a first-order transition (see Privalov, 1979; Uversky and Ptitsyn, 1996). This phenomenon has become famous and is known as protein folding. Protein folding is perhaps the simplest example in which information processing is dramatically multidimensional in nature: although Anfinsen's experiment suggests clearly that the sequence (also called the primary structure) contains all the information necessary to find the native spatial fold (called the tertiary structure), the information processing cannot advance in a sequential symbol-by-symbol fashion and must operate simultaneously, using remote parts of the sequence.

### 2. Polymer physics perspective

Remarkably, a similar evolution of ideas can be found in another scientific discipline important to our review, namely, the theoretical physics of polymers. Indeed, polymer chains were long viewed as just one-dimensional chains of links. In the classic books (Flory, 1953, Volkenstein, 1959; Birshtein and Ptitsyn, 1964), the one-dimensional Ising model (in different incarnations) remained the main analytical tool for treating phenomena ranging from chain flexibility and rubber elasticity to the helix-coil transition. With some earlier deep insights by Flory (1953), in the 1960s Edwards, de Gennes, and Lifshitz emphasized interactions and the three-dimensional nature of polymers, thus bringing polymers into the realm of modern theoretical physics.

Protein globules have always been outstanding subjects of interest for polymer theorists. As soon as it was realized that protein denaturation was accompanied by the unfolding of the protein chain and a dramatic increase in its three-dimensional packing size (measured, say, by a diffusion coefficient), a model appeared that associated a random-walk-type coil with the denatured state and a collapsed, condensed globule with the native state (Ptitsyn and Eizner, 1965). This initiated studies of the coil-to-globule transition of (homo)polymers, viewed as a first step in the study of proteins. In particular, I. M. Lifshitz, as early as 1968, explicitly stressed the vital role of quenched disorder for proteins, although technically he examined only homopolymer globules, in which the linear chain structure is quenched (Lifshitz, 1968).

Twenty years ago, a review article by Lifshitz *et al.* (1978) appeared in *Reviews of Modern Physics*, detailing the properties of globules and coil-to-globule transitions. Most notably, it showed that the difference between the coil and globule states is qualitative; indeed, the transition between these two phases is *a phase transition*, which means that in the $N\to\infty$ limit there is a point of discontinuity of certain thermodynamic functions. Furthermore, other peculiarities of both the globule state and the coil-to-globule transition are brought about precisely by the presence of the quenched *linear memory* (to use Lifshitz's original terminology).

The main qualitative difference between the two phases is that chain connectivity provides for long-range correlations in the coil state (as well as in the dilute or semidilute solution of coils; de Gennes, 1979); moreover, attraction between monomers and/or external pressure is strong enough in the globule to suppress those correlations, making the (single-molecule) globule a thermodynamically uniform system, with an interior similar to a concentrated polymer solution or a melt. Unfortunately, this subtlety remains poorly understood and is often the source of erroneous models.

In the conclusion of the article (Lifshitz *et al.*, 1978), the authors wrote

> Since ''living matter is the most interesting subject of investigation for the living matter that is able to investigate'' (L. A. Blumenfeld) . . . it is very important to determine the degree of influence of the primary structure on the thermodynamic properties of a globule. . . . The model theory developed in this article describes with qualitative accuracy the homopolymer chain in fairly simple situations and can, apparently, be used as a basis for further, more realistic approximations. As to biopolymers, this theory can stimulate the formulation of interesting problems. [p 71]

We now are in a position to continue from where Lifshitz *et al.* (1978) left off. With regard to homopolymers, the theory of the coil-globule transition has now been developed to quantitative accuracy (Grosberg and Kuznetsov, 1992). More complicated scenarios also have been examined, including charged macromolecules, intramolecular liquid crystals, topological constraints, and several others. As for biopolymers and ''the degree of influence of the primary structure,'' in recent years protein folding theories have substantially elaborated on the original ideas of Lifshitz *et al.* (1978).

Such a theory was first attempted by Garel and Orland (1988a). An important breakthrough in the understanding of heteropolymers came with the work of Shakhnovich and Gutin. While Shakhnovich and Gutin (1989a) employed a simplified model of random sequences that assumed that the interaction energies between monomers could be taken from a Gaussian distribution, this was the first microscopic theory to include both the fundamental features of a globular state and a quenched sequence. This work bridged the gap between the insights of polymer physics and the ideas of yet another line of development important to our field, namely, the theory of spin glasses.

### 3. Disordered systems and spin glasses

A third line of development of importance for the advancement of protein folding theory is the theory of

disordered systems (Binder and Young, 1986; Mezard *et al.*, 1987). In particular, the pioneering work of Bryngelson and Wolynes (1987) suggested a phenomenological approach to protein folding based on the analogy of a protein with a spin glass.

The physical concept of frustration is important in both spin glasses and proteins. A system is said to be frustrated if the degrees of freedom cannot be optimized simultaneously, usually due to some constraints imposed by quenched disorder (see p. 1 of Mezard *et al.*, 1987 for a nice discussion of frustration). Indeed, frustration in the case of proteins is due to the chain connectivity between monomers with opposite affinities to the neighbors and/or environment. Frustration leads to a rugged energy landscape, with the consequences of a complicated topology of valleys, high barriers, strong traps, long-lived states, history dependence of kinetics, etc. The insight of Bryngelson and Wolynes (1987) was to postulate the applicability of the random-energy model, a simplified spin-glass model suggested by Derrida (1980), to the heteropolymer freezing problem. This model neglects correlations between energies of different (micro)states, and perhaps in some part due to its simplicity, it has been an exceptionally useful paradigm for understanding the basic aspects of folding thermodynamics.

Two years later, Shakhnovich and Gutin (1989a) proved, albeit for a simplified model, the applicability of the random-energy model and its insights. This paper gave rise to the development of modern views that combine the microscopic approaches of (hetero)polymer theory with the methods and results borrowed from spin-glass theory. Before we examine this picture, with its concepts of heteropolymer freezing and sequence design, we first need to establish the basic terminology and explain why protein folding is considered a problem.

### B. What is protein folding, and why is it a problem?

Proteins are long chain molecules, typically about 50 to 500 monomer units (amino acids) long. It is well established that as polymers in general (and proteins in particular) are submerged in a liquid solvent, they can be viewed as classical stringlike objects undergoing conformational fluctuations governed by interactions not only between neighboring monomers along the chain, but also throughout the volume of the system, i.e., between monomers approaching each other in the course of fluctuations.

The distinguishing feature of proteins as polymers is that they are heteropolymers: they consist of 20 different species of monomeric units (called amino acid residues). For each particular protein, monomers are connected along the chain in a strictly defined nonuniform sequence, called the primary structure. To perform biochemical functions, a protein chain must attain a particular, well-defined three-dimensional conformation. In this *native fold*, the spatial positions of heavy (nonhydrogen) atoms are determined with an accuracy on the order of 0.5 Å. Many (but not all) proteins are able to find their native fold spontaneously, without the assistance of any kinetic mechanisms of the living cell, and thus the phenomenon of correct spontaneous folding of proteins represents an outstanding challenge for physicists.

Understanding protein folding means understanding the relationship between a protein's sequence and its native-state structure. There are, however, three aspects of this relationship under vigorous study, detailed below.

#### 1. How do proteins reach their native state?

A polymer chain has many different possible conformations. Imagine that there are discrete "rotational isomers," which means that there are, on average, $\gamma$ possible orientations of each monomer with respect to its immediate neighbor along the chain ($\gamma$ is estimated to be $\approx 1.6$), and accordingly a polymer with $N$ monomers has $M=\gamma^N$ conformations (Grosberg and Khokhlov, 1994). Protein folding means finding just one of them. A naive way to find this native state would be to search exhaustively through all of the states. Unfortunately, as $M$ grows exponentially with $N$, this search would require much longer than the experimental folding time scale of about 0.1 second; even an astronomically longer time scale would be insufficient. This problem was first explicitly formulated by Levinthal (1968) and is known as the Levinthal paradox. Alternatively perhaps sequences of proteins should be selected such that they provide the necessary bias towards the native state. We know that certain sequences were selected in the course of evolution, at both the prebiological and biological stages (Volkenstein, 1994; Ptitsyn, 1995). An attempt to rationalize the evolutionary selection of sequences leads, however, to another paradox: how were proteins selected by evolution such that they fold?

#### 2. How are protein sequences selected?

As 20 types of monomers are used to build proteins, there are $20^N$ possible sequences for proteins of length $N$. Neither the material on the earth (or, for that matter, in the universe) nor the time since the earth or the universe were formed is sufficient to exhaustively try all. Therefore "good folders" could not have been selected by an unbiased search through the space of sequences. This is not surprising, as protein sequences are the result of evolution. However, what biases evolutionary sequence selection toward "good folding" sequences? This review will describe how to select sequences of proteinlike heteropolymers that overcome the Levinthal paradox, i.e., given a desired native-state fold, we will detail methods used to design sequences to fold into that conformation. Whether nature makes an evolutionary selection in this same manner is still an open and fascinating question.

#### 3. How can we predict a protein's fold given its sequence?

An organism's genome details the protein sequences present, but tells nothing of what these proteins do. Structure prediction, i.e., predicting a protein's native-

state structure given its sequence, is a critical step towards connecting this genomic information to biological functionality. Currently the most successful attempts at such predictions are based on empirical approaches (Finkelstein, 1997). One approach assumes that small-scale, local changes in proteins can be made without disturbing the nature of the overall fold; such an approach assumes that two similar sequences will have very similar folds. Also of note is ''threading,'' in which the investigated sequence is mounted over an equally long part of a target fold (Finkelstein, 1997). We shall not discuss these ideas here, as further details of structure prediction are beyond the scope of this review, and we suggest one of the many recent reviews on this topic (Finkelstein, 1997; Levitt *et al.*, 1997).

## C. Freezing and sequence design

### 1. Freezing transition

The most notable result of the theories mentioned above (Bryngelson and Wolynes, 1987; Shakhnovich and Gutin, 1989a) is the concept of the heteropolymer freezing transition—the transition between two phases of the heteropolymer chain, of which one is dominated by very few [$\mathcal{O}(1)$] conformations, while the other is dominated by exponentially many of them [$\mathcal{O}(e^N)$]. This transition, long known for the random-energy model and other spin-glass models, is now believed to be a property of a broad class of heteropolymers. In heteropolymers, just as in spin glasses, this transition is caused by the presence of frustrations (Mezard *et al.*, 1987); the ''desire'' of attractive monomers to be in contact systematically conflicts with the ''desire'' of repelling monomers to hide from each other. When the freezing transition was predicted phenomenologically, based on the random-energy model (Bryngelson and Wolynes, 1987), and microscopically, for a model with $\mathcal{O}(N^2)$ independent interactions (Shakhnovich and Gutin, 1989a), it appeared that a satisfactory basic model for protein folding was already at hand. Indeed, these theories demonstrated the possibility of a frozen thermodynamic state of a polymer chain dominated by a single conformation. An obvious temptation was to identify this one conformation directly with the native state.

Both of these theories operated, explicitly or implicitly, with only *random-sequence* heteropolymers. This was consistent with the fact that real protein sequences statistically look very much like random sequences (Ptitsyn, 1995), which is in turn consistent with the idea that an evolutionary search could explore only a tiny part of the sequence space and thus could not pull sequences too far away from random (see also Monod, 1971). Thus it was important to learn from these models that even a heteropolymer with a random sequence could be dominated by one conformation. However, it was later shown that random sequences were not sufficiently proteinlike, as we detail below.

### 2. Evolution and design of sequences

#### *a. Nonrandomness of protein sequences*

There are three interconnected difficulties with random sequences:

(1) When a random sequence freezes, it has the proteinlike property of a unique ground state. However, the conformation that is the ground state is also random. This is unacceptable for a protein, in which particular features of the native state are of decisive importance for its function.

(2) While the lowest-energy native state for a random heteropolymer may be thermodynamically the most stable, some other states are typically only about $\sqrt{N}$ higher in energy. As the structures of these states are typically completely unrelated to the native one, and as they are only negligibly higher in energy, they serve as strong traps, making the folding kinetics hopelessly slow and unreliable.

(3) Due to the $\sqrt{N}$ difference in energy between low-energy states, a minute change ($\sim 1/\sqrt{N}$) of interaction energies (induced, for example, by a change in the surrounding solution or a mutation in the sequence) leads to a complete alteration of the native state, with a new native-state conformation that is absolutely unrelated to the old one (Bryngelson, 1994).

In light of what we have just said, random sequences do not appear to be sufficiently proteinlike. Bryngelson and Wolynes (1987) have argued that some special sequences are required to overcome these problems and have specifically conjectured that sequences should be selected such that they fold with relatively little frustration (''minimal frustration''). In hindsight, this conclusion is perhaps not surprising, as protein sequences are known to have undergone at least some (Volkenstein, 1994), although not necessarily perfect (Monod, 1971), evolutionary optimization.

#### *b. Quenched nonrandomness and evolution*

From a physical point of view we can say that protein sequences are quenched but not random. They are quenched because they are controlled by stable covalent chemical bonds between monomers, and on the time scale relevant for folding (typically milliseconds, but sometimes up to minutes) sequences certainly remain unchanged. In typical physical systems, such as spin glasses, whatever is quenched is usually random. Thus the nonrandomness of protein sequences is the major peculiarity that makes the entire field so interesting and exciting: nonrandomness reflects the biological evolutionary origin of proteins, and thus examining protein folding is possibly the first area in which we can achieve understanding of a physical process controlled by biological information.

Although quenched on a folding time scale, protein sequences are believed to change over a much longer, evolutionary time. Of course, in this case we are not speaking of a single protein molecule, but rather of an ensemble of different proteins in the evolving multitude of living species. Although not much is known about this

process, the situation with two distinct relaxation times, as described above, is reminiscent of neural networks (Amit *et al.*, 1987), in which image recognition and learning are analogous to folding and evolution, respectively (Friedrichs and Wolynes, 1989; Sasai and Wolynes, 1990; Pande, Grosberg, and Tanaka, 1994c; Ramanathan and Shakhnovich, 1994). There are attempts to simulate, both experimentally (see Sec. I.C.3.a) and computationally, the evolutionlike selection of ''good folders.'' Computational models of evolution (Gutin *et al.*, 1995; Mirny *et al.*, 1998), however exciting, require enormous computational power, even for lattice simulations—these studies simulate evolution by accepting mutations that speed the mean first passage time for folding (which is computationally intensive). Ideally, there should be some way to bypass the complexities of simulated evolution and to come up with a more effective, even if artificial, approach to designing sequences capable of robust folding. This requires a more direct means for implementing the minimal frustration principle.

#### c. Sequence design

Sequence design can also be called inverse folding. Two algorithms for achieving it using the minimal frustration principle were suggested independently by Shakhnovich and Gutin (1993) and by the present authors (Pande, Grosberg, and Tanaka, 1994b). Both schemes suggest that sequences can be designed by annealed search in the sequence space biased by the interactions between monomers as if they were in the desired target native conformation. Note that *sequence annealing* is just the opposite of heteropolymer folding, which can be considered as *conformation annealing* for a fixed sequence. Roughly speaking, sequence design selects sequences whose energy in the desirable target conformation is as low as possible. More accurately, it can be described in terms of statistical-thermodynamic equilibration in sequence space. In other words, sequence selection is modeled by drawing sequences from a Gibbs distribution $(1/Z)\exp[-E/T^{\mathrm{des}}]$, where $E$ is the energy of the polymer in the native state, the ''design'' temperature $T^{\mathrm{des}}$ characterizes the quality of design, or degree of optimization, and $1/Z$ is the normalization factor. This latter formulation gives rise to a powerful analytical theory.

It is very important to note that sequence selection based on these design algorithms (Shakhnovich and Gutin, 1993; Pande, Grosberg, and Tanaka, 1994b), as well as on others[1] does not encounter frustrations and that it occurs in a nonrugged landscape in sequence space. In general, when suggesting any kind of sequence selection approach, one has to confront the evolution paradox problem, whether related or unrelated to the real evolution. The absence of frustration for annealed sequences guarantees that this approach avoids the evolution paradox.

The idea of sequence design is fundamental and will be central to this review. We shall see that stronger design (lower $T^{\mathrm{des}}$) yields sequences that are better at resolving the Levinthal folding paradox, although they are harder to select, and vice versa. We shall even speculate that $T^{\mathrm{des}}$ can serve as a phenomenological descriptor of the stage reached by the evolution.

### 3. Applications of design concepts to proteins

#### a. Experiments on proteinlike peptide chains

One can investigate design concepts in proteinlike heteropolymers by examining the results of two recent experiments. First, in analogy to the conceptual picture of sequence selection, Robert Sauer (1996) and his colleagues created an ensemble of random amino acid sequences and then selected sequences by applying biotechnological agents that remove chains that have folded poorly. The resulting ensemble is therefore one of designed sequences. Analyses of these sequences show that a large fraction of the remaining sequences have the proteinlike properties of stability and secondary structure. (See also Kauffman and Ellington, 1999.)

Second, Michael Hecht and co-workers designed amino acid sequences to fold into a particular fold (a four-helix bundle). To accomplish this, Hecht's group first considered the set of amino acid sequences that would lead to four helices, i.e., the restricted set limited by the constraints of a fixed secondary structure. This set was further limited by the constraint that monomers on the surface of the protein should be hydrophilic and monomers in the core should be hydrophobic. The remaining set of sequences was still quite large, and sequences from this ensemble were generated and examined (Hecht *et al.*, 1990).

While the motivation behind this ''design scheme'' is formally independent of questions such as the relevance of the random-energy model, these questions lie at the heart of the problem: one selects a sequence from an ensemble that minimizes the energy of a desired conformation in such a way that the other conformations are not favored energetically. (Here this is done by the selection of hydrophobic residues in the core and hydrophilic on the surface.) The existence of secondary structure helps to simplify this process, as the resulting conformation space is greatly diminished, effectively removing undesired conformations from the phase space of the protein.

More recent experimental approaches (Regan and DeGrado, 1988; Dahiyat and Mayo, 1996; DeGrado, 1997; Lazar *et al.*, 1997) follow sequence design more closely, as they minimize the energy of the sequence in the desired conformation or some energy-inspired fitness score (Dahiyat and Mayo, 1996). These experiments have demonstrated that such design techniques can indeed yield sequences that fold to desired native-

[1]Similar and/or improved schemes include those of Kurosky and Deutsch, 1995; Sun *et al.*, 1995; Deutsch and Kurosky, 1996; Morrissey and Shakhnovich, 1996; Seno *et al.*, 1996; Khokhlov and Khalatur, 1998; Micheletti, Banavar, *et al.*, 1998.

state conformations, but with little similarity to the original, wild-type protein sequence.

*b. Correlations in protein sequences*

Design implies that the sequences are not taken at random. One may ask if real protein sequences look random, or if they bear some fingerprint of evolutionary optimization. Moreover, one can predict the qualitative character of sequence correlations to be expected if proteins are indeed designed in the sense described above. This has been done by Pande, Grosberg, and Tanaka (1994a) and Irback *et al.* (1996), and we briefly summarize the methodology and results below.

From our theoretical models, we expect that evolutionary selection of protein sequences should lead to correlations of monomer species along the chain consistent with energy minimization. For example, we expect that positively charged monomers will predominantly be followed by negatively charged monomers; also, as hydrophobicity induces an effective attraction between hydrophobic monomers, we expect hydrophobic monomers to be followed by other hydrophobic monomers along the protein chain.

To convert amino acid sequences into a language more convenient for analysis, Pande, Grosberg, and Tanaka (1994a) used mappings that translated the 20 amino acids into a three-digit code $-1$, 0, $+1$. For example, one such mapping translated amino acids based upon their charge. Pande, Grosberg, and Tanaka (1994a) next employed a sensitive mathematical technique to find correlations in the translated three-letter code sequences. Correlations were found in these sequences, but more interestingly, these correlations were consistent with energy minimization. For example, we found anticorrelations in the Coulomb mapping ($+1$ is typically followed by $-1$, etc.) and correlations in hydrophobic mappings.

In a recent work (Irback, *et al.*, 1996), correlations in protein sequences were reexamined, and a discrepancy was found with Pande, Grosberg, and Tanaka (1994a) regarding the type of correlations present. The most likely explanation is related to the difference in mappings used by Pande, Grosberg, and Tanaka (1994a) and by Irback *et al.* (1996). Whatever the explanation for this discrepancy, the existence of correlations based on physical ''charges'' (interaction properties) of monomers was confirmed by Irback *et al.* (1996). We stress that this is highly nontrivial, as correlations reflect the role of energies and interactions in a slow process, i.e., over an evolutionary time scale. It is also interesting to note that the search for correlations in protein sequences has a long history (see Ptitsyn, 1995 and references therein), but correlations were not found until the particular design-induced character of the correlations had been predicted (Pande, Grosberg, and Tanaka, 1994a).

## II. BUILDING MODELS

### A. Phenomenological models

We now describe models that provide a framework for discussing the microscopic details.

1. Go model

One of the first phenomenological models of proteins was suggested by Ueda *et al.* (1975). It is based on the idea that the native conformation of proteins is energetically very well optimized. If this is true, we may assume that native contacts (contacts between monomers present in the native state) are energetically favored. This leads to the conclusion that the polymer energy depends primarily on a conformation-dependent property, such as the number of native contacts $\mathcal{Q}$; thus

$$E=(1/2)\epsilon\mathcal{Q}\equiv E_{\star}q, \tag{2.1}$$

where $\epsilon<0$ is the energy brought to a monomer by one native contact, $q=\mathcal{Q}/\mathcal{Q}_{\max}\leqslant 1$, and $E_{\star}=\epsilon\mathcal{Q}_{\max}/2$ is the energy of the native state. Of course, in reality there may be some overall (homopolymeric) attraction or repulsion, and energies of some non-native contacts may be non-negligible, which modifies the above. However, as we shall see, this will not be important for many aspects of freezing, and these aspects are correctly captured by the Go model [named after its creator, N. Go (1983)].

The Go model assumes that every native contact contributes equally. We shall later derive the Go model energy from a microscopic theory in the limit of well-designed sequences. We shall show that Eq. (2.1) can be used even when non-native energies do not vanish. Moreover, corrections to it can be systematically computed. It also turns out that the Go representation (2.1) is exact for a system with independent random interactions.

The Go model is often criticized as a model that ''uses the answer to answer the question,'' i.e., uses knowledge of the native state to describe folding. This is a valid issue if one's interest is structure prediction, but it is not important for the question we address, i.e., how proteins reach the native state. If one wishes to study the statistical mechanics of proteinlike heteropolymers and, in particular, to examine the nontrivial role of the polymeric entropy, use of Eq. (2.1) clearly delineates the roles of energy and entropy. In particular, if one assumes that the energy is linear in $q$, then the cooperativity of the transition must come purely from the nature of polymeric entropy, and thus the cooperativity found may be the result of general polymeric properties and not interaction-dependent details.

2. Random-energy model

Imagine a protein folded in its native state, and then unfold it and fold it into a completely unrelated conformation. How does the energy change? In the new conformation, the set of contacting pairs of monomers is totally different from that in the original native state. As energy is mainly determined by interacting pairs of monomers, this implies that energies of different conformations could be unrelated to each other. Thinking in terms of the theory of disordered systems (Mezard *et al.*, 1987) and identifying particular protein sequences with particular random realizations of disorder, Bryngelson

and Wolynes (1987) came to the conclusion that energies of different conformations should be considered independent random variables. This idea corresponds precisely to the random-energy model, introduced by Derrida (1980). Briefly, this model can be described as follows.

Formally, to compute the partition function $Z$ of an arbitrary system, one needs only the energies $E_1, E_2, \ldots, E_{\mathcal{M}}$ of all microstates: $Z = \Sigma_{i=1}^{\mathcal{M}} \exp[-E_i/T]$, where $\mathcal{M}$ is the number of states. Generally, $\mathcal{M}$ is huge, as it scales exponentially with the number of particles (monomers) $N$, $\mathcal{M} \simeq \exp(\omega N)$, where $\omega \sim 1$ depends on the conformations available, i.e., on chain flexibility, packing conditions, lattice geometry in the case of lattice models, etc. All of these energies depend, of course, on the sequence. In the random-energy model, one says that the energy of each conformation, say, $E_1$, is distributed over the realizations of disorder in the same way as energies of all other conformations and is statistically independent of them. If we call $P(E)$ the probability distribution of the energy of some particular conformation over disorder, and $P(E_1, E_2)$ the joint probability distribution that conformations 1 and 2 have energies $E_1$ and $E_2$, respectively, then the basic assumption of the random-energy model is that

$$P(E_1, E_2) = P(E_1)P(E_2). \tag{2.2}$$

It is also usually supposed that the $P(E)$ distribution is Gaussian:

$$P(E) = (2\pi N \mathcal{E}^2)^{-1/2} \exp\left[-\frac{E^2}{2N\mathcal{E}^2}\right], \tag{2.3}$$

where $\mathcal{E}$ is the characteristic width of the distribution, but this assumption is far less important than the statistical independence of states expressed in Eq. (2.2).

We now summarize the main properties of the random-energy model (for more details, see Appendix B):

(1) The defining property is the statistical independence of states (2.2).

(2) The energy spectrum consists of a quasicontinuous part, which is independent of disorder, and a few discrete energy levels that are placed very individually for each realization of disorder.

(3) The ground state for typical realizations is of order $\sqrt{N}$ below the edge of the continuous spectrum, which, in turn, is of order $N$ below the mean energy. Also, for typical realizations discrete levels are of order $\sqrt{N}$ from each other.

(4) There is a certain temperature $T_{\text{freeze}} = \mathcal{E}/\sqrt{2\omega}$ such that at $T > T_{\text{freeze}}$ the system explores the high-entropy continuous part of its spectrum, while at $T < T_{\text{freeze}}$ it is locked with vanishing entropy into discrete individual states.

(5) The free energy is given by the equation

$$F(T) = \begin{cases} -T \ln\langle Z_{\text{seq}}(T)\rangle & \text{if } T > T_{\text{freeze}} \\ -T_{\text{freeze}} \ln\langle Z_{\text{seq}}(T_{\text{freeze}})\rangle & \text{if } T \leqslant T_{\text{freeze}}, \end{cases} \tag{2.4}$$

which means that the disorder is irrelevant above the freezing point.

Note that these properties are independent of the Gaussian form of the single energy distribution (2.3).

### 3. Modifications of the random-energy model

One modification of the random-energy model was motivated by the desire to describe lattices in which possible energy values are markedly discrete. The corresponding "discrete" version of the model (DREM) has been examined by Gutin and Shakhnovich (1993). This model retains the statistical independence of states, but it appears useful in sorting out the degeneracy of computer-simulated lattice heteropolymer models.

The generalized random-energy model (GREM) represents a much more serious departure, as it allows for some statistical interdependence between states. In particular, it was assumed in the work of Derrida (1985) that all states form an ultrametric set, and correlations exist between members of the same family, but not between different families. We shall discuss in more detail the deviations from the random-energy model and the relevance of correlations between states in Sec. VI.

The question of kinetics is, of course, very important for protein folding. The random-energy model itself, without additional assumptions, says nothing about kinetics: one needs to specify the states $1, \ldots, \mathcal{M}$ and, for example, assume which states are (mechanically, geometrically, kinetically, etc.) next to each other in the conformational space. This may seem easy for regular spin models, in which an "elementary kinetic move" is just a single spin flip; see the discussion, for instance, by Parisi (1997). However, for the random-energy model in general one must make some additional assumptions to even speak about kinetics. An example of such an assumption is discussed by Bryngelson and Wolynes (1989). Of course, these additional assumptions may contain—explicitly or implicitly—some bias toward the native state. However, without such a bias, the model itself would leave the Levinthal paradox unresolvable. Indeed, if the energy landscape is pure uncorrelated white noise, then, even if our system happens to come very close (in terms of conformations) to the desired native state, it cannot have any clue that the goal is close, or any hint of which direction to go. Since there is no bias, random-energy kinetics amount to diffusion in a multidimensional space, which is certain to have a slim chance of reaching the desired point. In fact, in this case the folding time is the Levinthal time (Bryngelson and Wolynes, 1989).

Thus, however useful the model may be for shaping our views of folding thermodynamics, folding kinetics will certainly require understanding the corrections to, and deviations from, the random-energy model. In this review, we shall concentrate on the equilibrium properties of proteins and proteinlike heteropolymers. For more information on issues relating to kinetics, see the recent reviews by Wolynes *et al.* (1995), Bryngelson *et al.* (1995), Shakhnovich (1996), Dill and Chan (1997), and Pande *et al.* (1998).

### B. Microscopic model: Compact globular heteropolymer

However useful phenomenological models may be, the ultimate challenge is to achieve a complete molecular understanding and to build up a complete microscopic theory, which starts from a microscopic, albeit approximate and coarse-grained, Hamiltonian. As protein native states are globules, it is natural to begin with a theory for compact, globular heteropolymers.

#### 1. Energy

There are three important aspects of the problem, namely, the quenched sequence of the particular set of monomer species, interactions between them, and conformations of the polymer. The simplest Hamiltonian that encompasses all three has the following form:

$$\mathcal{H}(\{s_I\},\{\mathbf{r}_I\})=\sum_{I>J}^{N} B_{s_I s_J} f(\mathbf{r}_I-\mathbf{r}_J). \tag{2.5}$$

Capital Latin indices count the monomers along the chain, $s_I \in \{1,\ldots,q\}$ is the *species* of monomer $I$ (and thus the ordered list $\{s_I\}$ represents the polymer "sequence"), $q$ is the number of species, $\mathbf{r}_I$ is the *position* of monomer $I$ (and thus $\{\mathbf{r}_I\}$ represents "conformation"), $B_{ss'}$ is the interaction free energy (i.e., including both energetic and entropic effects of side chain interactions) between monomers of species $s$ with $s'$, and $f(\mathbf{r})$ is the interaction range function.

Throughout most of this paper, we shall consider short-range interactions, where $f(\mathbf{r})=\Delta(\mathbf{r})$ is a function concentrated on the neighboring sites in space:

$$\Delta(\mathbf{r})=\begin{cases}1 & \text{when } |\mathbf{r}|\leqslant r_0\\ 0 & \text{otherwise.}\end{cases} \tag{2.6}$$

On the lattice, $r_0$ is naturally identified with the lattice spacing. Thus our model simply says that the energy of a polymer conformation is determined by the matrix of species-species energies $B_{ij}$ for the monomers in contact.

In writing the energy in the form (2.5), we implicitly assume that all the following conditions are met:

- chain connectivity: points $\mathbf{r}_I$ and $\mathbf{r}_{I+1}$ are always next to each other in space for all $I$;
- excluded volume: $|\mathbf{r}_I-\mathbf{r}_J|>v^{1/3}$ for $I\neq J$, where $v$ is the volume of monomers;
- dense packing: when structured as in their native-state form, amino acid side chains are very well packed. This can be modeled as monomers on a lattice, with each of the $N$ monomers occupying a region with exactly $N$ lattice sites. Thus the polymer "visits" every site once and only once.

The last condition describes compact globular states and means, in particular, that the density of the globule cannot fluctuate and is evenly distributed in space.

Equation (2.5) is an approximation, since it neglects, for example, heteropolymeric three-body interactions, which may be non-negligible due to the cooperative nature of hydrophobicity. Nevertheless, it correctly captures most essential components of the problem, notably interaction-induced frustrations for different conformations. Also, the Hamiltonian (2.5) does not include any signature or even any possibility of secondary structure (such as $\alpha$ helices or $\beta$ sheets), which is very important for proteins. In considering the model (2.5) we assume a coarse-grained representation of the polymer in terms of some quasimonomers (Grosberg and Khokhlov, 1994), in which the small-scale details of the monomer's chemistry have been "renormalized out." Thus, in heteropolymer lattice models, a single site represents some arrangement of the elements of secondary structure, rather than a single amino acid.

Interactions in our Hamiltonian (2.5) are only pairwise; this is generally incorrect. Moreover, interactions depend on mutual rotations and other degrees of freedom for the monomers and not only on their coordinates, as our Hamiltonian (2.5) implies. Finally, interactions are not purely short range. We shall later address the effect of long-range (e.g., Coulomb) interactions.

#### 2. Conformations

##### a. Nature of conformation space

Various models of conformations have been motivated by the desire to capture the essential physics while keeping the model tractable. One common method of drastically simplifying conformation space and making the problem much more tractable computationally has been to use lattice models for proteins.[2] Lattice models have a long history in polymer physics (see, for example, Binder, 1995 and references therein).

Since protein native states are globules, any model of folding or design must incorporate maximally compact conformations. Examining only maximally compact conformations renders the problem much more tractable, as the constraint of filling the space greatly reduces the number of possible conformations and also allows one to neglect complications due to density fluctuations. Therefore it is not surprising that such models have become common (Lau and Dill, 1989; Shakhnovich and Gutin, 1990; Pande, Joerg, *et al.*, 1994).

The enumeration of maximally compact conformations is much more feasible than, for example, the enumeration of *all* conformations of a chain with a given length $N$. Nevertheless, exhaustive enumerations of even maximally compact conformations of chains longer than $N=27$ require a great deal of computing power (Pande, Joerg, *et al.*, 1994). For example, 36-mers can be easily enumerated on a massively parallel supercomputer. Pande, Joerg, *et al.* (1994) used a 128-node CM-5; the maximally compact 48-mer conformations have also been enumerated, but required 2 CPU weeks on the 128-node CM-5. For this reason 27-mers have become a canonical model for folding and design studies. Recently it has been conjectured that the number of contacts in a space of 27-mers is similar to that found in short pro-

[2]See, for example, Lau and Dill, 1989; Shakhnovich and Gutin, 1990; Camacho and Thirumalai, 1993; Olszewski *et al.*, 1996; Onuchic *et al.*, 1996; Shakhnovich, 1996.

teins (typically 60–80 amino acids; Luthey-Schulten *et al.*, 1995 Onuchic *et al.*, 1995) thus 27-mers may be reasonable coarse-grained models of small proteins.

Other sets of conformations have also been examined. First, polymers have far fewer conformations in two dimensions than in three, so one can enumerate much longer chains in two dimensions (Chan and Dill, 1993; Dinner *et al.*, 1994). Moreover, it has been argued that enumerable two-dimensional chains (25-mers and 36-mers, for example) have a surface-to-volume ratio more like that of a protein than that of an enumerable three-dimensional chain (a 27-mer or 36-mer; Chan and Dill, 1993). However, in general there are appreciable differences between two-dimensional and three-dimensional models in statistical mechanics, and these differences are more fundamental than, for example, the surface-to-volume ratio. Such differences are well understood for homopolymers (Grosberg and Khokhlov, 1994), and thus there are no grounds to believe that heteropolymers are an exception. Differences between two-dimensional and three-dimensional heteropolymers will be detailed in later sections.

A second approach is to enumerate in three dimensions, but restrict the set of conformations. For example, ''crumpled'' 64-mers have been enumerated (a crumpled 64-mer consists of eight $2\times2\times2$ size 8-mer cubes, strung together to make a single $4\times4\times4$ cube; Pande, Grosberg, Joerg, and Tanaka, 1996). They make an interesting model as, unlike shorter chains such as 27-mers and 36-mers, crumpled 64-mers allow small-scale rearrangements and perhaps model multidomain proteins (Panchenko *et al.*, 1995; Pande, Grosberg, Joerg, and Tanaka, 1996).

We stress that sampling just the maximally compact conformations may not be sufficient to describe the freezing transition, especially if it is accompanied by a coil-to-globule transition (i.e., the system goes directly from a coil to a frozen globule phase; Klimov and Thirumalai, 1996; Pande, Gross, and Tanaka, 1997a). To study all conformations, one must either perform Monte Carlo kinetic simulations and use sampling techniques, such as the Monte Carlo histogram technique (Socci and Onuchic, 1995), to gather information about the density of states or perform a full enumeration of all conformations. Of course, full enumeration of all conformations (compact and noncompact) is extremely computationally intensive and has been performed only for chains with $N\leqslant18$ (Pande, Grosberg, and Tanaka, 1997b). On the other hand, Monte Carlo simulations have been performed on much longer chains, up to $N=125$ (Abkevich *et al.*, 1995).

Another aspect of modeling conformations is the choice of the native state. In most of the simple models, any of the maximally compact conformations appear equally good for the role of the native state. However, it was first noticed by Finkelstein *et al.* (1993) and later stressed by Li *et al.* (1996) that some of the compact conformations may be in some sense better; for example, they may be more ''designable'' (Li *et al.*, 1996). We shall briefly return to this idea in Sec. VII.A.2.

Finally, there has also been substantial progress in off-lattice models, in which residues can take any position in real space. Off-lattice models fall into two categories: (1) realistic all-atom models, which are good at reproducing specific chemical properties of proteins,[3] and (2) coarse-grained models, which trade atomic detail for computational tractability.[4] Both models are now typically employed to examine protein folding kinetics, which is beyond the scope of this review. We refer the reader to other recent reviews on protein folding kinetics (Wolynes *et al.*, 1995; Dill and Chan, 1997; Pande *et al.*, 1998) for more details.

#### *b. Contacts of monomers and overlap of conformations*

The Hamiltonian (2.5) involves pair contacts of monomers. According to our definition (2.6), monomers $I$ and $J$ are said to be in contact[5] if they are closer in space than a certain cutoff length $r_0$: $|\mathbf{r}_I^\alpha-\mathbf{r}_J^\alpha|\leqslant r_0$. The cutoff $\Delta$ function facilitates counting these contacts, as the number of pair contacts $\Phi_\alpha$ for the conformation $\alpha$ is simply

$$\Phi_\alpha=\sum_{I\neq J}\Delta(\mathbf{r}_I^\alpha-\mathbf{r}_J^\alpha). \tag{2.7}$$

We are interested, however, not only in the total number of contacts in the current conformation $\alpha$, but also in how many of these contacts are also present in the native state $\star$, because these contacts are typically of particularly low energy. To count this number of native contacts, we first introduce the quantity

$$\mathcal{Q}_{\alpha\beta}=\sum_{I\neq J}\Delta(\mathbf{r}_I^\alpha-\mathbf{r}_J^\alpha)\Delta(\mathbf{r}_I^\beta-\mathbf{r}_J^\beta), \tag{2.8}$$

which is called the ''overlap'' of two conformations $\alpha$ and $\beta$. $\mathcal{Q}_{\alpha\beta}$ represents the number of contacts that these two conformations have in common. In terms of the overlap, the fraction of native contacts for the conformation $\alpha$ is given by

---

[3]Such models have been proposed by Caflisch and Karplus (1994); Li and Daggett (1994); Boczko and Brooks (1995); Van Gunsteren *et al.* (1995); Tirado-Rives *et al.* (1997).

[4]These include the models of (Levitt and Warshel, 1975; Thirumalai and Guo, 1995; Guo and Thirumalai, 1996; Rey and Skolnick, 1996; Pande and Rokhsar, 1997).

[5]To avoid misunderstandings, we note here that our definition of contacts is somewhat the result of convention and is model dependent. As soon as one considers more realistic models, say an all-atom model of proteins, the position of each residue cannot be specified in terms of a single vector $\mathbf{r}_I$, as there are several atoms per residue and the side group of any residue can rotate or deform. One can let the $\mathbf{r}_I$ denote the positions of the $C_\alpha$ atoms (carbon atoms of the main protein chain to which the side group of the residue is covalently attached). Then two amino acids are said to be in contact if the distance between their respective $C_\alpha$ atoms is less than $r_0$; for example, in the work of Miyazawa and Jernigan (1985), $r_0$ is taken to be $r_0=6.5$ Å. Clearly, such an approach disregards possible rotations of the residues, the fact that they are of different shapes, etc.

$$q_\alpha \equiv q_{\alpha\star} = \mathcal{Q}_{\alpha\star}/Q_{\max}, \tag{2.9}$$

where the normalization factor in the denominator $\mathcal{Q}_{\max}$ stands for the maximal possible number of contacts, which is restricted by the packing conditions and generally scales as $N$ in the thermodynamic limit, $N\to\infty$. We note that typically the native state is as compact as possible, and thus $\mathcal{Q}_{\max}$ equals the number of contacts in the native state, that is, $\mathcal{Q}_{\max}=\mathcal{Q}_{\star\star}$.

#### c. Contact maps

Another equivalent way to describe conformations in terms of pair contacts is to use the so-called contact matrix, or contact map (Mirny and Domany, 1996). The contact map of conformation $\alpha$ is defined as

$$C_{IJ}^{\alpha} = \Delta(\mathbf{r}_I^{\alpha} - \mathbf{r}_J^{\alpha}). \tag{2.10}$$

The matrix element $(I,J)$ is unity if monomers $I$ and $J$ are in contact, and vanishes otherwise. In terms of contact matrices, it is easy to rewrite both the Hamiltonian (2.5),

$$\mathcal{H}(\{s_I\},\{\mathbf{r}_I\}) = \sum_{I>J}^{N} B_{s_I s_J} C_{IJ}^{\star}, \tag{2.11}$$

and the overlap parameter

$$\mathcal{Q}_{\alpha,\beta} = \sum_{I>J} C_{IJ}^{\alpha} C_{IJ}^{\beta} = \frac{1}{2}\mathrm{Tr}\,\hat{C}^{\alpha}\hat{C}^{\beta}. \tag{2.12}$$

This representation is often useful.[6]

### 3. Interactions

Natural proteins include $q=20$ species of monomers (amino acids), and thus the interaction matrix $B_{ij}$ should be $20\times 20$. Neither the values of the matrix elements nor the characteristics aspects of models with a smaller number of species have been agreed upon in the literature. The most commonly employed $20\times 20$ matrices for simple models have been extracted from the statistics of the protein database (Miyazawa and Jernigan, 1985; Kolinski *et al.*, 1993; Mirny and Shakhnovich, 1996; Thomas and Dill, 1996, and references therein). The fact that rather different criteria are employed stresses that perhaps none of the known matrices should be trusted absolutely. Of course, the matrix Hamiltonian, Eq. (2.5), is itself an approximation, neglecting, for example, three-body and long-range interactions.

At the opposite extreme, as hydrophobicity is believed to be the main driving force of protein collapse, various models are used with just two monomeric species, hydrophobic and polar (Chan and Dill, 1993). The independent interaction model, in which the number of monomer species is as large as the total number of monomers, so that each matrix element $B_{ij}$ enters in the energy of any conformation at most once and never more than once, is somewhat special; matrix elements are then taken independently from a Gaussian distribution. This model is convenient for theorists (as we shall see in later sections). The most often used and natural interaction matrices have been recently summarized by Du *et al.* (1998).

One may ask how the solvent enters into our model Hamiltonian (2.5). We assume that the solvent molecules equilibrate considerably faster than the polymer and thus we integrate over all solvent degrees of freedom. This leads to an effective interaction between monomer species, i.e., our interaction matrix $B_{ij}$. For example, the hydrophobic effect explicitly details the interaction between oily molecules and water, but leads to an effective attraction between hydrophobic molecules as they come together in order to try to avoid the water molecules (Chan and Dill, 1993).

### 4. Sequences: Microcanonical and canonical design

As mentioned in Sec. I.C.2, sequences of real proteins are close to random; since only extremely sensitive statistical tools can detect the correlations present, the nonrandomness of sequences is crucially important. Therefore the question is how to find those special sequences capable of robust folding. In principle, this question has two aspects, equilibrium and kinetic. The latter (how to find sequences that lead quickly and reliably to folding to the native state) is not completely understood at present and falls beyond the scope of the present review; see other recent reviews (Bryngelson *et al.*, 1995; Wolynes *et al.*, 1995; Shakhnovich, 1996; Dill and Chan, 1997; Pande *et al.*, 1998). We shall concentrate on the former aspect, which is how to find sequences that are thermodynamically stable in the desired conformation. Throughout this review, we shall call this the desired native target conformation $\star$.

In general, when the sequence is quenched, the distribution of the polymer over its conformations obeys the standard Boltzmann formula:

$$\mathcal{P}_{\mathrm{conf}}(\mathrm{seq}) = \frac{\exp[-\mathcal{H}(\mathrm{seq},\mathrm{conf})/T]}{\Sigma_{\mathrm{conf}}\exp[-\mathcal{H}(\mathrm{seq},\mathrm{conf})/T]} = \exp[(F(\mathrm{seq}) - \mathcal{H}(\mathrm{seq},\mathrm{conf}))/T], \tag{2.13}$$

where $F(\mathrm{seq}) = -T\ln\Sigma_{\mathrm{conf}}\exp[-\mathcal{H}(\mathrm{seq},\mathrm{conf})/T]$. This is the probability distribution over *conformations* for a given sequence; this distribution is realized when a real polymer is immersed in a solvent and undergoes thermal conformational motion. In the last line of this formula, the free energy of the given sequence $F(\mathrm{seq})$ comes as a

[6] In particular, one can calculate the thermal average overlap directly from the thermal average of the contact matrix:

$$\langle \mathcal{Q}_{\alpha_1\alpha_2}\rangle = \sum_{\alpha_1\alpha_2} \mathcal{Q}_{\alpha_1\alpha_2} P_{\alpha_1} P_{\alpha_2} = \sum_{\alpha_I\alpha_2}\sum_{I\neq J} C_{IJ}^{\alpha_1} P_{\alpha_1} C_{IJ}^{\alpha_2} P_{\alpha_2} = \sum_{I\neq J}\langle C_{IJ}\rangle^2,$$

where $P_\alpha$ is the Boltzmann probability of finding conformation $\alpha$ in equilibrium. Furthermore, this relation is readily generalized for an arbitrary $m$ conformation overlap. As freezing is marked by the onset of overlap between conformations found in equilibrium and $\langle C_{IJ}\rangle$ is straightforward to calculate, this is a useful computational technique to measure freezing.

conformation-averaged energy. In terms of this distribution, the objective of sequence design can be formulated as finding a sequence for which the probability distribution (2.13) is peaked—as sharply as possible—at the desirable conformation $\star$. As Eq. (2.13) indicates, this should be achieved by minimizing $\mathcal{H}(\text{seq},\star)-F(\text{seq})$, i.e., taking the sequence that has low energy in the conformation $\star$ [$\mathcal{H}(\text{seq}, \star)$ is minimal], or, more accurately, whose energy in the conformation $\star$ is minimized compared to its energies in all other competing conformations.

This way of thinking about the sequence design is enhanced by insight gained from the random-energy approach. First of all, disordered systems usually obey the self-averaging principle, and the free energy is typically the same, or almost the same, for different samples, in which realization of disorder is different. In the case of heteropolymers, this implies that $F(\text{seq})$ is in fact independent of the sequence. Of course, self-averaging holds in the thermodynamic limit, and for any real computational model there is some sequence dependence of the free energy. However, for the analytic theory, which operates with the thermodynamic limit, we can safely ignore the sequence dependence of $F(\text{seq})$ as long as the overall monomer composition is fixed. A second and even more significant simplification comes with the idea that the energies of different compact states are independent of each other, as stipulated by the random-energy model. If this is true, selection of sequences with any given value of $\mathcal{H}(\text{seq}, \star)$ does not affect energies of other compact conformations. More accurately, the distribution of energies of any compact conformation distinct from $\star$ over the sequences is the same for both the ensemble of all (random) sequences and the (much smaller) ensemble of sequences restricted to a fixed value of $\mathcal{H}(\text{seq}, \star)$. Thus selecting sequences with small $\mathcal{H}(\text{seq}, \star)$ does not lead to sequences with low energies in other competitive conformations—pulling the native state down does not affect competitors.

This yields an extremely powerful prescription for sequence design: one selects sequences with the given value of native-state energy $E_\star=\mathcal{H}(\text{seq},\star)$, not worrying at all about other conformations. Selection of low $E_\star$ presumably yields good folder sequences, with lower $E_\star$ yielding target conformations with greater thermodynamic stability.

The ensemble of sequences with a given ground-state energy $E_\star$ is similar to the microcanonical ensemble in statistical mechanics. A design based on $E_\star$ can be called microcanonical. In statistical mechanics, it is technically more convenient to use the canonical ensemble in which temperature is fixed instead of energy. A similar idea is also valid for sequence design. We use an analog of the canonical ensemble in which the native-state energy $E_\star$ is not fixed, but rather controlled through an artificial temperature which we call the design temperature $T^{\text{des}}$. Equivalently, we constrain the energy $E_\star=\mathcal{H}(\text{seq},\star)$ with a Lagrange multiplier $1/T^{\text{des}}$. In this canonical ensemble, each sequence appears with Boltzmann-distributed probability

$$P^{\star}_{\text{seq}}=\frac{P^{(0)}_{\text{seq}}\exp[-\mathcal{H}^{\text{des}}(\text{seq},\star)/T^{\text{des}}]}{\Sigma_{\text{seq}}P^{(0)}_{\text{seq}}\exp[-\mathcal{H}^{\text{des}}(\text{seq},\star)/T^{\text{des}}]}, \tag{2.14}$$

where

$$P^{(0)}_{\text{seq}}=\prod_{I=1}^{N} p_{s_I} \tag{2.15}$$

is the probability for the sequences made randomly from independent monomer species with occurrence probabilities $p_s$,[7] and $\mathcal{H}^{\text{des}}$ is the Hamiltonian used for the design. Normally, $\mathcal{H}^{\text{des}}$ is close enough to $\mathcal{H}$ that the difference between them is frequently neglected and Eq. (2.14) is written directly in terms of $\mathcal{H}$. Nevertheless, it is important to write it in a more general form, as we shall explain later (see Secs. IV.A.2.b. and VII.B.2.c).

Note that the normalization factor in the design distribution (2.14) is quite different from that in Eq. (2.13): in Eq. (2.13), it incorporates the sum over real conformations visited by the real chain during its real heat motion; by contrast, in Eq. (2.14) it is the sum over different sequences. This is why the design prescription based on Eq. (2.14) is often referred to as *sequence annealing*. Indeed, the way to computationally realize the distribution (2.14) is to run Monte Carlo dynamics simulations in which elementary moves change monomer species on top of the frozen conformation $\star$ (Shakhnovich and Gutin, 1993). Another way to obtain essentially the same distribution (2.14), with the added advantage of hints at some experimental paths, was suggested by Pande, Grosberg, and Tanaka (1994b, 1995b) and called ''imprinting.'' This prescription works even for such complex conformations $\star$ as conformations with knots, or with a pocket imitating the enzyme active site (Pande, Grosberg, and Tanaka, 1994b).

Throughout this paper we shall mostly examine the implications of the design prescription in Eq. (2.14), but we first point out some of its obvious limitations:

- Equation (2.14) implies that overall composition of the sequence is fixed. It may be advantageous to relax this requirement, as was done by Morrissey and Shakhnovich (1996).
- Self-averaging may be inaccurate, particularly for relatively short chains of computer models. This problem may also be exacerbated for some types of interactions and conformations, such as, for instance, the two-dimensional hydrophobic-polar model. To circumvent this problem may require explicit minimization of $\mathcal{H}(\text{seq},\star)-F(\text{seq})$ instead of just $\mathcal{H}(\text{seq}, \star)$. This is computationally very intensive, because it requires summing over conformations for each sequence encountered during the search through the sequence space. Nevertheless,

[7] If the polymer is prepared under conditions of equilibrium with some ''monomer supply bath'' in which chemical potentials of the species $s$ are $\mu_s$, then $p_s=\exp(\mu_s/T^{\text{des}})\times[\Sigma_j\exp(\mu_j/T^{\text{des}})]^{-1}$, where $T^{\text{des}}$ is the design temperature.

this program has been carried out for some simple models with good results.[8]

- The idea that sequence selection based on $E_\star$ does not affect other energies is certain to be somewhat inaccurate (see Sec. V). It is sometimes said that along with designing the desirable state "in" by pulling $E_\star$ down, we have to also design other states "out" by pulling their respective energies up (Yue *et al.*, 1995). Apart from difficulties in implementation, in practice such an idea is probably needed only for some special cases, such as the hydrophobic-polar model. The design implications of Eq. (2.14) beyond the random-energy model are poorly understood at present.
- In our formulation, we did not discuss how to choose the target conformation $\star$; in most of the paper, we shall consider it to be chosen arbitrarily, with the only constraint that of being compact. However, it has been pointed out by Finkelstein *et al.* (1993) and further stressed by Li *et al.* (1996) that some of the conformations $\star$ may be much more designable than others. Again, this is particularly important for models with few monomer species. We shall touch upon this problem in Sec. VII.
- The big problem is, of course, kinetics: even if Eq. (2.14) yields sequences with very low $E_\star$, this tells us nothing about their kinetic ability to actually find the conformation $\star$ reliably and in a short time. While this problem can be overcome computationally via evolutionlike selection algorithms (Mirny *et al.*, 1998), it remains a challenge for analytical theory. Nevertheless, canonical design according to Eq. (2.14) produces sequences whose kinetic properties are much better than those of random sequences: they fold in times much shorter than the Levinthal time (Shakhnovich, 1994), even though other ways of speeding up the kinetics are possible.

In this review, we shall concentrate on the canonical ensemble of designed sequences described by the distribution (2.14). Although imperfect, this distribution appears very effective, and a close examination of it yields insights that turn out to be valuable in looking for further improvements. We repeat that the canonical ensemble of sequences is characterized by the value of $T^{\rm des}$: for lower $T^{\rm des}$, we model sequences whose native states are better optimized energetically, while for higher $T^{\rm des}$ we are left with an unaltered ensemble of random sequences. However, the ensemble with lower $T^{\rm des}$ has fewer sequences in it, and those sequences are therefore more difficult to select.

## III. DEVELOPING THE THEORY

### A. Preliminary arguments

#### 1. What is important about conformational entropy?

In this section we derive the basic equations for describing the folding thermodynamics of designed heteropolymers. We begin with a hand-waving argument to promote physical intuition and then discuss a formal replica approach.

We shall consider polymers whose sequences are designed, as explained and discussed in Sec. II.B.4 above. While it may seem counterintuitive, it is useful to consider a heteropolymer with a random sequence as a poorly designed heteropolymer. Although we shall explain this in more detail later (see the last paragraph of Sec. IV.A.2.a), it is worth saying here that sequence design plays the role of an external field; thus considering designed sequences is useful in the same sense as examining an Ising model with an external magnetic field sheds light on the Ising model itself. Considering designed sequences means, in particular, that the native-state conformation is assumed to be known *a priori*. Therefore there is an obvious order parameter, which is the overlap of a current conformation with the native one, or simply the number of native contacts $\mathcal{Q}$ in a current conformation. We shall use $q=\mathcal{Q}/\mathcal{Q}_{\rm max}$, where $\mathcal{Q}_{\rm max}$ is the number of monomer-monomer contacts in the native state.

The situation is particularly simple for a well-designed sequence (Sec. VI.B), in which the energy is simply linear in $q$, because each native contact has a large energy, while energies of all other non-native contacts are negligible. As we have already mentioned (Sec. II.A.1), this property of well-designed sequences is formalized in the Go model (Ueda *et al.*, 1975), in which all non-native contacts are explicitly supposed to have zero energy, and all native contacts are supposed to have the same energy $\epsilon$. For such a model, the free energy per monomer as a function of the order parameter $q$ can be written in the form

$$f(q)=\epsilon q - Ts(q), \tag{3.1}$$

where $s(q)$ is the entropy per monomer, which is related to the number of conformations with the given number of native contacts $q$. At the mean-field level, the optimal $q$ is determined by the minima of $f(q)$.

Of course, it is not surprising that a well-designed heteropolymer (large absolute value of $\epsilon$; $\epsilon<0$) at low enough temperature tends to go to the maximal possible value of $q$, which is $q=1$, and which is the native state—this is how the model was made, and Eq. (3.1) simply formalizes this fact. The remarkable property of proteins is, however, not that they attain the ground-state conformation at essentially zero temperature, but that they remain in this conformation up to some finite temperature, in the room-temperature range, with thermal fluctuations causing very small perturbations in the native conformations. This property must be included in $s(q)$, and this is why the conformational entropy $s(q)$ is of decisive importance.

#### 2. A naive estimate of conformational entropy

Equation (3.1) shows that the reliable selection of a low-energy state based on contact interaction energies between monomers can only be possible with a properly

[8]See the papers of Kurosky and Deutsch, 1995; Deutsch and Kurosky, 1996; Morrissey and Shakhnovich, 1996; Seno *et al.*, 1996; Frauenkron *et al.*, 1998; Micheletti, Seno *et al.*, 1998.

restricted set of conformations. In general, conformations are restricted by both excluded-volume and chain connectivity constraints. To estimate the effects of chain connectivity, let us imagine that each of the $\mathcal{Q}$ native contacts implies the formation of a Gaussian loop (i.e., a polymer loop subject to chain connectivity, but not excluded volume) by the polymer piece between two contacting monomers. With $\mathcal{Q}$ contacts and a total of $N$ monomers, each loop has to have a length of about $l \sim N/\mathcal{Q} \sim 1/q$; since the probability of forming a loop by an $l$-monomer-long chain scales as $l^{-3/2}$ (this is the Yacobsen-Stockmayer factor; see, for example, Grosberg and Khokhlov, 1994), the entropic price for making each loop is about $-(3/2)\ln l \simeq (3/2)\ln q$, corresponding to an entropy per monomer of

$$s_{\text{loop}} \simeq (3/2) q \ln q. \tag{3.2}$$

To a crude approximation, $s_{\text{loop}}$ takes care of the restrictions imposed by chain connectivity. As regards excluded-volume restrictions, we can crudely assume them simply to add up for each contact, thus making a linear in $q$ contribution

$$s_{\text{bond}} \simeq s q, \tag{3.3}$$

which we refer to as the bond entropy. This term simply renormalizes $\epsilon$ and makes it temperature dependent.

Finally, we have to take into account a combinatorial factor, because there may be many ways to choose $\mathcal{Q}$ native contacts out of the $\mathcal{Q}_{\max}$ possible ones. In principle, this choice is also affected by the chain connectivity and excluded-volume constraints; neglecting this effect, we can write (per monomer)

$$s_{\text{mix}} \simeq -q \ln q - (1-q)\ln(1-q). \tag{3.4}$$

Thus we arrive at an entropy estimate of the form

$$s(q) = s_{\text{mix}} + s_{\text{bond}} + s_{\text{loop}}, \tag{3.5}$$

which implies that the free energy (3.1) can be written as

$$f(q)/T \simeq (\epsilon/T - s)q - (1/2) q \ln q + (1-q)\ln(1-q). \tag{3.6}$$

### 3. The "all or nothing" minimum

In the mean-field approximation, only the minima of free energy (3.6) are relevant. For the model discussed in the previous section, there are two free-energy minima as a function of $q$, at $q$ close to one and at $q = 0$. The $q \approx 1$ minimum corresponds to the native state. The fact that it is not at $q = 1$ reflects, on the one hand, the primitive character of the model considered; specifically, mixing entropy is strongly overestimated at large $q$, where the choice of contacts to be formed is severely restricted. On the other hand, to say that the native state consists of *one* conformation is obviously an idealization, and our model captures this fact correctly. More interestingly, the $q = 0$ minimum represents a mixture of all conformations having no native contacts; obviously, there is a vast number of such conformations and they are structurally completely unrelated to the native conformation. These two minima never come close to each other in $q$, and there is a first-order transition between the two corresponding states depending on $\epsilon/T$: at low $T$, the energy bias toward the native state dominates, while at higher $T$ the entropic tendency toward a liquidlike mixture of states dominates.

Our arguments suggest the powerful conclusion that freezing is the transition between a state with $q = 0$ and one with $q \approx 1$. Conformations with intermediate $q$, which have some partial structural similarity to the native fold, never play any significant role in the thermodynamics. Thus all relevant states are structurally completely independent, since their energies are due to completely uncorrelated sets of contacts. This implies a complete lack of statistical interdependence between energies of different states, which precisely corresponds to the random-energy model of Derrida (1980).

To illustrate the internal mechanism behind the applicability of the random-energy model, it is useful to generalize Eq. (3.6) to an arbitrary spatial dimension $d$. The energy, bond entropy, and mixing entropy are all independent of dimension, while the loop entropy scales like $(d/2) q \ln q$ instead of $(3/2) q \ln q$. Thus the only change in Eq. (3.6) is that the coefficient of the $q \ln q$ term becomes $(2-d)/2$ instead of $-1/2$. Provided $d > 2$, the qualitative picture remains unaltered, and we expect the random-energy model to remain applicable. However, for $d \leqslant 2$ the situation changes dramatically and the main issue is the thermodynamic balance between the native state and a set of structurally related competitors. The corresponding phenomena are poorly understood at present, and we shall not consider them further.

For the three-dimensional case, the application of the random-energy model allows us to go far without invoking large-caliber theoretical guns such as replicas. Historically, indeed, the random-energy model was first postulated for heteropolymers (Bryngelson and Wolynes, 1987). We shall discuss later (see Sec. IV) both the model itself and the possibility of building a theory using it as an (educated) assumption. Before that, we consider a more sophisticated theory to gain further insight into why and under which conditions freezing follows a random-energy "all-or-nothing" scenario, with just one native state dominating one phase and the mixture of completely unrelated conformations representing the other phase. This theory closely follows the blueprint outlined in Eq. (3.6), with $q$ playing the role of an order parameter and with conformational entropy determined by the conditions of chain connectivity and excluded volume.

### 4. Limitations of these preliminary arguments

The argument based on Eq. (3.6) is heuristic, and our estimate for entropy is primitive. Real loops are not Gaussian, but are restricted by the excluded-volume constraints. Mixing entropy is strongly affected by both excluded volume and chain connectivity constraints. Moreover, all those constraints become increasingly severe when density increases and the conformation approaches the native state. However, one lesson that we can extract from these arguments is as follows. For rea-

sons of conformational geometry in three dimensions, there are relatively few conformations structurally similar to the native state. This makes freezing to the native state relatively easy, makes the random-energy model a useful paradigm, and allows for simple prescriptions for successful sequence design. All these circumstances clearly deserve further investigation and will be treated in the following sections.

Another comment concerns the use of $q$ as an order parameter. This is possible only for a designed sequence for which the native state is known *a priori*, since $q$ is defined as $q=\mathcal{Q}/\mathcal{Q}_{\max}$, where $\mathcal{Q}$ is the overlap with the native state. For random sequences, when the native state is unknown, a more complex order parameter must be defined. This is related to the matrix $\mathcal{Q}_{\alpha\beta}$, where Greek indices show all possible pairs of conformations, since each conformation is a potential candidate for the role of a ground state. As we shall show later, the ensemble of all random sequences is dominated by those for which freezing is a smooth, second-order transition. However, designed sequences exhibit a sharp first-order freezing transition. Considering random sequences as the limit of poorly designed ones is thus similar to considering a critical point as an end of a first-order phase-transition line.

### B. Large incompressible globule

In the previous section, we outlined a simple model for calculating the free energy of designed heteropolymers. In this section, we detail a more rigorous and systematic theory, which will be valid in the limit of a large, dense globule.

The approximations in this section are valid for globules that are large enough that one can ignore effects related to the surface of the globule, including both higher solvent exposure for monomers near the surface and the entropic contribution associated with chain loops in the fringe around the globule, as discussed by Grosberg and Khokhlov (1994; see Sec. 20). The globule should also be dense, or maximally compact, in the sense that local packing density fluctuates neither in time nor in space, remaining constant throughout the globule. In terms of the lattice model, this latter assumption means that every lattice site within the globule is visited once and only once because of the excluded-volume constraint; thus a lattice model of such a conformation is a Hamiltonian path on the sublattice.

While an analysis of the large-compact-globule extreme requires one to employ the mean-field replica approach, the outcome of the replica treatment suggests the approximate validity of the random-energy paradigm of "complete overlap or no overlap" (as discussed in the previous section). This allows one to circumvent the complexities of replicas for many purposes and obtain important results in a simpler fashion (see Sec. IV. A. 3). The reader uninterested in purely theoretical issues, such as why the random-energy model is valid in the large-globule model, may skip this section.

#### 1. Self-averaging and the replica trick

In general, each particular chain is characterized by the sequence-dependent free energy $F(\mathrm{seq})$ defined below Eq. (2.13). However, the free energy is believed to obey the principle of self-averaging (Mezard *et al.*, 1987), which says that the probability distribution of free energies for independent samples is very narrow and thus the free energy for almost all of the sequences practically coincides with the mean free energy. Thus we come to the idea of averaging the free energy over the sequences, in this case, over the set of designed sequences.

Recall that the design procedure works for a particular target conformation, which we label $\star$. Thus, if we average over the sequences designed for $\star$, we obtain a $\star$-dependent free energy. We shall assume that the self-averaging principle is obeyed at this level, too, and we shall thus average the free energy over all compact target conformations $\star$.

The question of $\star$ dependence has been examined in the literature, first by Finkelstein *et al.* (1993) and later by Govindarajan and Goldstein (1996) and Li *et al.* (1996). It has been hypothesized that some particular conformations, for instance, those with secondary structure, are in some sense better fitted for protein needs and thus have been selected by nature. This viewpoint has been particularly strongly expressed by Li *et al.* (1996). For our purposes in the present section, we shall average out these differences by formally averaging over target conformations $\star$. A theory for the $\star$ dependence of freezing remains an interesting unsolved problem.

As we adopt a statistical approach, we analyze properties of an ensemble, in this case the ensemble of designed sequences. Thus we average over both sequences and target conformations:

$$F\equiv\langle F(\mathrm{seq})\rangle=-T\langle\ln Z(\mathrm{seq})\rangle, \tag{3.7}$$

where $\langle\cdots\rangle$ denotes averaging over target conformations and $F(\mathrm{seq})$ already averages sequences for a given target conformation. To technically perform the average of $\ln Z$ we employ the replica trick, that is, we employ the identity

$$F=-T\langle\ln Z(\mathrm{seq})\rangle=-T\lim_{n\to 0}\frac{\langle Z^n(\mathrm{seq})\rangle-1}{n}. \tag{3.8}$$

The probability of the appearance of a certain sequence in the design procedure is given by the distribution (2.14). Noting that the chain partition function is given by the sum over all (compact) conformations $C$,

$$Z(\mathrm{seq})=\sum_C \exp\left[-\frac{1}{T}\mathcal{H}(C,\mathrm{seq})\right], \tag{3.9}$$

we arrive at the conclusion that the average of $Z^n$ over $P_{\mathrm{seq}}$ is naturally treated as an additional replica:

$$\langle Z^n(\mathrm{seq})\rangle=\sum_{\mathrm{sequence}}\prod_{I=1}^{N} p_{s_I}\sum_{C_0,C_1,\ldots,C_n}\exp[-\mathcal{K}_{\mathrm{eff}}],$$

$$\mathcal{K}_{\rm eff}=\sum_{\alpha=0}^{n}\mathcal{H}_\alpha(C_\alpha,{\rm seq})/T_\alpha, \tag{3.10}$$

where the following notations are used: $\alpha=0,1,\ldots,n$ are the numbers of replicas: $C_\alpha=C_1,\ldots,C_n$ stand for conformations of replica number $\alpha$; replica $\alpha=0$ is attributed to the target conformation $\star$, that is, $C_0=\star$; $T_\alpha=T^{\rm des}$, the design temperature, for $\alpha=0$ and $T_\alpha=T$ otherwise. Similarly, $\mathcal{H}_\alpha=\mathcal{H}$ for $\alpha\neq 0$ and $\mathcal{H}_0=\mathcal{H}^{\rm des}$. Note that for brevity we skip all normalization factors. We shall take care of all of them at the end.

Recalling the structure of the Hamiltonian for our model, we can rewrite Eq. (3.10) as

$$\begin{aligned}\langle Z^n\rangle=&\sum_{\rm sequence}\prod_{I=1}^{N}p_{s_I}\sum_{\{\mathbf{r}_I^\alpha\}}\\&\times\exp\left[\frac{1}{2}\sum_{\alpha=0}^{n}\sum_{I\neq J=1}^{N}B^{\alpha}_{s_Is_J}\delta(\mathbf{r}_I^\alpha-\mathbf{r}_J^\alpha)/T_\alpha\right]\\=&\sum_{\rm sequence}\prod_{I=1}^{N}p_{s_I}\sum_{\{\mathbf{r}_I^\alpha\}}\\&\times\exp\left[\frac{1}{2}\sum_{\alpha,\beta=0}^{n}\int d\mathbf{R}_1\,d\mathbf{R}_2\right.\\&\left.\times\sum_{i,j}^{q}\tilde\rho_i^\alpha(\mathbf{R}_1)\mathcal{B}_{i,j}^{\alpha\beta}\delta(\mathbf{R}_1-\mathbf{R}_2)\tilde\rho_j^\beta(\mathbf{R}_2)\right],\end{aligned} \tag{3.11}$$

where

$$\mathcal{B}_{ij}^{\alpha\beta}\equiv\frac{B_{ij}^{\rm des}}{T^{\rm des}}\delta_{\alpha0}\delta_{0\beta}+\frac{B_{ij}}{T}(\delta_{\alpha\beta}-\delta_{\alpha0}\delta_{0\beta}) \tag{3.12}$$

is a matrix that expresses the interactions used for both chain preparation (i.e., replica $\alpha=0$) and chain folding (other replicas). In the latter transformation we have used the definition of densities for monomers of a given species, position, and replica:

$$\tilde\rho_i^\alpha(\mathbf{R})=\sum_{I=1}^{N}\delta(s_I,i)\,\delta(\mathbf{r}_I^\alpha-\mathbf{R}). \tag{3.13}$$

Hereafter conformations are given in terms of position vectors $\mathbf{r}_I^\alpha$ for each monomer number $I$ in each replica $\alpha$. It is important that the sum over $\mathbf{r}_I^\alpha$ run over conformations. This means that the set $\mathbf{r}_I^\alpha$ for each replica $\alpha$ satisfies the conditions of chain connectivity ($|\mathbf{r}_I^\alpha-\mathbf{r}_{I+1}^\alpha|\leqslant a$) and excluded volume ($|\mathbf{r}_I^\alpha-\mathbf{r}_J^\alpha|\geqslant r_0$). Moreover, as we are dealing with a maximally compact globule, we assume that the density is spatially uniform:

$$\rho\equiv\rho_\alpha(\mathbf{R})=\sum_{I=1}^{N}\delta(\mathbf{r}_I^\alpha-\mathbf{R})=\sum_{i=1}^{q}\tilde\rho_i^\alpha(\mathbf{R}). \tag{3.14}$$

To facilitate the summation in which the condition of chain connectivity is strictly obeyed, Edwards (Doi and Edwards, 1986) introduced a path-integral representation, and Lifshitz (1968) introduced a similar discrete representation. In many later studies, additional terms in the Hamiltonian have been introduced to enforce chain connectivity, excluded volume, and compactness. We do not introduce these, considering all three conditions as *quenched* properties of the system, which define the set of microstates, or conformations over which we perform the partition sum.

### 2. Average over sequences

Inspection of the structure of Eq. (3.11) shows that the exponent is a quadratic form with respect to density (since we assume two-body interactions in the Hamiltonian). For simplicity we use vector, matrix, and scalar product notations to rewrite

$$\begin{aligned}\langle Z^n\rangle=&\sum_{\rm sequence}\prod_{I=1}^{N}p_{s_I}\sum_{\{\mathbf{r}_I^\alpha\}}\exp[-\mathcal{K}_{\rm eff}],\\ \mathcal{K}_{\rm eff}=&\langle\vec{\tilde\rho}|\hat{\mathcal{B}}^{(q(n+1))}\otimes\hat I^{(\infty)}|\vec{\tilde\rho}\rangle^{(q(n+1)^\infty)},\end{aligned} \tag{3.15}$$

where we use superscripts as a shorthand notation to denote the dimensionalities of the space involved. Specifically, superscript $n+1$ represents replica space (index $\alpha$); $q$ represents species space ($i$); and $\infty$ represents real space ($\mathbf{r}$) or, more precisely, the Hilbert space of *functions* of $\mathbf{r}$. We use here the operation of direct product $\otimes$, in the following sense: if there are two matrices (or operators) of different dimensionalities $r$ and $s$, say $\hat A^{(r)}$ and $\hat B^{(s)}$, then $\hat A^{(r)}\otimes\hat B^{(s)}$ is the matrix of dimensionality $rs$ obtained by mapping the matrix $A_{uv}\hat B^{(s)}$ onto each matrix element $(u,v)$ of $\hat A^{(r)}$ matrix. The properties of this operation are summarized by Pande *et al.* (1995a). Operator $\hat I^{(\infty)}$ is the identity operator with respect to real coordinate space, meaning that it has the kernel $\delta(\mathbf{R}_1-\mathbf{R}_2)$.

The most serious problem with Eq. (3.15) is that summation over the sequences, or over $s_I$ variables, is difficult, because monomers interact with each other and the corresponding monomer variables are coupled. One approach to this problem is to trade coupling of monomers for coupling of replicas by performing the Hubbard-Stratonovich transformation on the quantity $\tilde\rho_i^\alpha(\mathbf{R})$, thus introducing the conjugate fields $\phi_i^\alpha(\mathbf{R})$[9]:

$$\begin{aligned}\langle Z^n({\rm seq})\rangle_{\rm seq}=&\sum_{\{\mathbf{r}_I^\alpha\}}\int\mathcal{D}\{\phi\}\exp\left\{\frac{1}{2}\langle\vec\phi|(\hat{\mathcal{B}}^{-1})^{(q(n+1))}\right.\\&\left.\otimes\hat I^{(\infty)}|\vec\phi\rangle^{(q(n+1)^\infty)}\right\}\\&\times\sum_{\rm sequence}\prod_{I=1}^{N}p_{s_I}\exp\{\langle\vec\phi|\vec{\tilde\rho}\rangle^{(q(n+1)^\infty)}\}.\end{aligned} \tag{3.16}$$

[9]The Hubbard-Stratonovich transformation is the equality

$$\exp[\langle\vec{\tilde\rho}|\hat{\mathcal{A}}|\vec{\tilde\rho}\rangle]=(4\pi\det\hat{\mathcal{A}})^{-1/2}\int d\{\phi\}\exp\left[\frac{1}{4}\langle\vec\phi|\hat{\mathcal{A}}^{-1}|\vec\phi\rangle+\langle\vec\phi|\vec{\tilde\rho}\rangle\right].$$

In this section, we skip the normalization factor generated by the transformation; this factor will be canceled while performing the inverse transformation. Thus the sum over sequences involves only uncoupled monomers in the last factor (the ''source'' term) of the partition function above. This facilitates the summation over the sequences:

$$\begin{aligned}&\exp\{\text{source term}\}\\&=\sum_{\text{sequence}}\prod_{I=1}^{N}p_{s_I}\exp\{\langle\vec{\phi}|\vec{\tilde{\rho}}\rangle^{(q(n+1)^\infty)}\}\\&=\prod_{I=1}^{N}\sum_{i=1}^{q}p_i\exp\left\{\sum_{\alpha=0}^{n}\int d\mathbf{R}\,\phi_i^\alpha(\mathbf{R})\,\delta(\mathbf{r}_I^\alpha-\mathbf{R})\right\}.\end{aligned}\tag{3.17}$$

We now make the single most important approximation of the theory: we expand the source term in powers of $\phi$ and truncate the expansion to $\mathcal{O}(\phi^2)$. As we discuss later (see Sec. VI.B.3), this approximation is equivalent to a high-temperature expansion, or, in other words, an expansion over the disordered part of the interactions. We thus arrive at the following expression for the source term:

$$\begin{aligned}&\sum_{i=1}^{q}\sum_{\alpha=0}^{n}\int d\mathbf{R}\,\rho^\alpha(\mathbf{R})p_i\phi_i^\alpha(\mathbf{R})\\&+\frac{1}{2}\sum_{i,j=1}^{q}[p_i\delta_{ij}-p_ip_j]\\&\times\sum_{\alpha,\beta=0}^{n}\int d\mathbf{R}_1\int d\mathbf{R}_2\,\phi_i^\alpha(\mathbf{R}_1)\\&\times Q_{\alpha\beta}(\mathbf{R}_1,\mathbf{R}_2)\,\phi_j^\beta(\mathbf{R}_2),\end{aligned}\tag{3.18}$$

where the replica overlap order parameter[10] is defined as

$$Q_{\alpha\beta}(\mathbf{R}_1,\mathbf{R}_2)=\sum_{I=1}^{N}\delta(\mathbf{r}_I^\alpha-\mathbf{R}_1)\,\delta(\mathbf{r}_I^\beta-\mathbf{R}_2)\tag{3.20}$$

and obeys the relations

$$\begin{aligned}&\int d\mathbf{R}_1\,d\mathbf{R}_2\,Q_{\alpha\beta}(\mathbf{R}_1,\mathbf{R}_2)=N;\\&\int d\mathbf{R}_2\,Q_{\alpha\beta}(\mathbf{R}_1,\mathbf{R}_2)=\rho_\alpha(\mathbf{R}_1);\\&\int d\mathbf{R}_1\,d\mathbf{R}_2\,Q_{\alpha\beta}^2(\mathbf{R}_1,\mathbf{R}_2)=\mathcal{Q}_{\alpha\beta},\end{aligned}\tag{3.21}$$

where $\mathcal{Q}_{\alpha\beta}$ is the contact overlap (2.8).

We define a new matrix $\hat{\Delta}^{(q)}$ with elements

$$(\hat{\Delta}^{(q)})_{ij}\equiv p_i\delta_{ij}-p_ip_j\tag{3.22}$$

and new vectors

$$\vec{\rho}^{(q(n+1)^\infty)}\equiv\rho_i^\alpha(\mathbf{R})=p_i\sum_{I=1}^{N}\delta(\mathbf{r}_I^\alpha-\mathbf{R})=p_i\sum_{j=1}^{q}\tilde{\rho}_j^\alpha(\mathbf{R}).\tag{3.23}$$

Note that $\vec{\rho}^{(q(n+1)^\infty)}=\vec{\rho}^{((n+1)^\infty)}\otimes\vec{p}^{(q)}$. Using these definitions, we rewrite the $(n+1)$-replica partition function in the form

$$\begin{aligned}\langle Z^n\rangle=\sum_{\{\mathbf{r}_I^\alpha\}}\int\mathcal{D}\{\phi\}\exp\Bigg\{&\left\langle\vec{\phi}\left|\frac{1}{4}(\hat{\mathcal{B}}^{-1})^{(q(n+1))}\otimes\hat{I}^{(\infty)}\right.\right.\\&\left.\left.+\frac{1}{2}\hat{Q}^{((n+1)^\infty)}\otimes\hat{\Delta}^{(q)}\right|\vec{\phi}\right\rangle^{(q(n+1)^\infty)}\\&+\langle\vec{\rho}|\vec{\phi}\rangle^{(q(n+1)^\infty)}\Bigg\}.\end{aligned}\tag{3.24}$$

We evaluate this Gaussian integral over $\phi$ and arrive at[11]

$$\langle Z^n\rangle=\frac{1}{2}\sum_{\{\mathbf{r}_I^\alpha\}}\exp[-\mathcal{K}_{\text{eff}}],$$

$$\begin{aligned}&\mathcal{K}_{\text{eff}}=\langle\vec{\rho}|\hat{\mathcal{B}}_{\text{eff}}|\vec{\rho}\rangle+\frac{1}{2}\ln\det\hat{\mathcal{B}}_{\text{eff}},\\&\hat{\mathcal{B}}_{\text{eff}}=(\hat{\mathcal{B}}^{1/2})^{(q(n+1))}\otimes\hat{I}^{(\infty)}[\hat{I}^{(q(n+1)^\infty)}+(\hat{Q}^{((n+1)^\infty)}\otimes\hat{\Delta}^{(q)})\\&\quad\times(\hat{\mathcal{B}}^{(q(n+1))}\otimes\hat{I}^{(\infty)})]^{-1}(\hat{\mathcal{B}}^{1/2})^{(q(n+1))}\otimes\hat{I}^{(\infty)}.\end{aligned}\tag{3.25}$$

### 3. Saddle-point approximation and replica symmetry breaking

#### a. Order parameter

Before going forward with the calculation, it is useful to stop and examine the physics of what we are trying to calculate. Equation (3.25) represents a standard polymer globule partition function, with summation over compact conformations and with *pairwise* interactions (because the effective Hamiltonian is still quadratic in densities). The only unusual aspect is that, in addition to the interaction between monomers as described by the matrix $\hat{\mathcal{B}}$, there is an interaction between replicas, which is introduced by the matrix $\hat{Q}$. Indeed, $\hat{\mathcal{B}}$ is diagonal in replica space (3.12), while the matrix $\hat{Q}$ is not. It is sometimes helpful to imagine this situation in terms of a braid of polymers, each folded in its own space, possibly with different conformations. Importantly, however, they are all *homopolymers*; there is no quenched disordered sequence in the partition function (3.25). This al-

[10]One can also define the $k$-replica overlap

$$Q_{\alpha_1,\ldots,\alpha_k}(\mathbf{R}_1,\ldots,\mathbf{R}_k)=\sum_{I=1}^{N}\prod_{\kappa=1}^{k}\delta(\mathbf{r}_I^{\alpha_\kappa}-\mathbf{R}_\kappa),\tag{3.19}$$

which enters into higher-order terms of the expansion.

[11]We use a symbolic notation $\ln\det\hat{\mathcal{B}}_{\text{eff}}$, which is poorly defined for an infinite-dimensional matrix—if understood literally. The precise meaning of that symbol will become apparent later. For now, we can formally imagine that our system is constrained to a lattice within a finite box.

lows, following Shakhnovich and Gutin (1989a), application of the general Lifshitz consideration of globules (Lifshitz *et al.*, 1978), which amounts to the use of a generalized density distribution as an order parameter. In this case, it is the matrix $\hat{Q}$ that generalizes the density distribution and has to serve as the order parameter. Indeed, the dependence of $\mathcal{K}_{\text{eff}}$ (3.25) on the conformation comes through $Q$ dependence only (similar to the usual case, in which the dependence on conformations comes through density, for example, in the virial terms $B\rho^2+C\rho^3$). We therefore have to pass from summation over conformations to integration over $\hat{Q}$. First, let us write this purely formally: as Eq. (3.25) is of the form

$$\langle Z^n\rangle=\sum_{\{\mathbf{r}_I^\alpha\}} \exp[-\mathcal{K}_{\text{eff}}\{Q\{\mathbf{r}_I^\alpha\}\}], \tag{3.26}$$

we have

$$\langle Z^n\rangle=\int \mathcal{D}Q \exp[-\mathcal{F}\{Q\}];$$

$$\mathcal{F}\{Q\}=\mathcal{K}_{\text{eff}}\{Q\}-S\{Q\}, \tag{3.27}$$

where the entropy is formally defined as

$$e^{S\{Q_{\alpha\beta}\}}=\sum_{\{\mathbf{r}_I^\alpha\}} \delta\left(Q_{\alpha\beta}-\sum_{I=1}^{N} \delta(\mathbf{r}_I^\alpha-\mathbf{R}_1)\,\delta(\mathbf{r}_I^\beta-\mathbf{R}_2)\right). \tag{3.28}$$

It is possible to avoid calculating this entropy explicitly by resorting to the following set of arguments, yielding at the end the simple expression (3.34).

### *b. Maximizing free energy in replicas*

The mean-field evaluation of the partition function implies a saddle-point approximation for the integral over $\hat{Q}$, Eq. (3.27). Normally this means taking the maximal value of the integrand. It is commonly believed (Mezard *et al.*, 1987), however, that in order to find the correct analytic continuation in the $n\to 0$ limit, one has to take the minimal value of the integrand, or, in other words, the maximal rather than the minimal value of the relevant free energy $\mathcal{F}$. As there are $n(n-1)/2$ independent off-diagonal elements in the $Q_{\alpha\beta}$ matrix, for the $0<n<1$ case, the integral over $Q_{\alpha\beta}$ formally represents the summation over a negative number of variables. Following this principle, we write

$$\langle Z^n\rangle=\exp[-\text{Max}_{\{Q\}}\mathcal{F}\{Q\}], \tag{3.29}$$

i.e., we have to maximize the effective free-energy functional (3.27).

### *c. One-step replica permutation symmetry breaking*

To maximize the free energy, we employ a variational approach. For a large spatially uniform globule we have

$$Q_{\alpha\beta}(\mathbf{R}_1,\mathbf{R}_2)=Q_{\alpha\beta}(\mathbf{R}_1-\mathbf{R}_2);\int d\mathbf{R}\, Q_{\alpha\beta}(\mathbf{R})=\rho. \tag{3.30}$$

We adopt a trial function $\varphi(\mathbf{x})$ normalized such that $\int d\mathbf{x}\,\varphi(\mathbf{x})=1$, and write

$$Q_{\alpha\beta}(\mathbf{R}_1-\mathbf{R}_2)=\frac{\rho}{(R_t^{\alpha\beta})^d}\,\varphi\left(\frac{\mathbf{R}_1-\mathbf{R}_2}{R_t^{\alpha\beta}}\right), \tag{3.31}$$

where $d$ is the dimensionality of space and the normalization condition defines the coefficient. The details about the form of $\varphi$ are not important, and for simplicity we have chosen $\varphi$ to be Gaussian.

$R_t^{\alpha\beta}$ can be interpreted as the diameter of the tube in which replicas $\alpha$ and $\beta$ coincide. Thus when $R_t^{\alpha\beta}$ is small, i.e., on a microscopic scale such as the distance between monomers along the chain $R_t^{\alpha\beta}\sim a$, then replicas $\alpha$ and $\beta$ overlap completely and are therefore essentially the same conformation. When $R_t^{\alpha\beta}$ is of a macroscopic length scale, such as the size of the globule $R_t^{\alpha\beta}\sim aN^{1/d}$, then the two conformations are completely different, i.e., they do not overlap at all.

To determine the nature of replica symmetry breaking, we must consider the possible equilibrium values of the overlap. To examine this, we now repeat the arguments of Shkhnovich and Gutin (1989a) and examine the free energy as a function of $R_t^{\alpha\beta}$, which is a measure of the overlap between replicas $\alpha$ and $\beta$. We assume that the chain in the large globule behaves as a Gaussian on scales smaller than the overall globule size. If that is the case, confinement of a polymer in a tube is associated with an entropic price (de Gennes, 1979; Grosberg and Khokhlov, 1994) that scales as $1/R^2$. This is independent of the overall shape of the tube and should be applicable for confinement of, say, replica $\alpha$ to within $R_t^{\alpha\beta}$ from replica $\beta$—if we believe that conformations inside the large globule obey the Flory theorem and follow Gaussian statistics. This argument, however, should be modified for replicas in the $n\to 0$ limit: for $n<1$ we expect that confinement will be associated with an entropic bonus and thus we expect the entropy contribution to scale as $-(R_t^{\alpha\beta})^{-2}$. As to the energy, it is seen from the most simple normalization consideration that it scales as $Q$ and thus as $(R_t^{\alpha\beta})^{-d}$ (3.31). Therefore every $\alpha\neq\beta$ term of the free energy has the functional form

$$\mathcal{F}_{\alpha\beta}=-\frac{A_1}{(R_t^{\alpha\beta})^2}+\frac{A_2}{(R_t^{\alpha\beta})^d}, \tag{3.32}$$

where $A_1$ and $A_2$ are positive numbers.

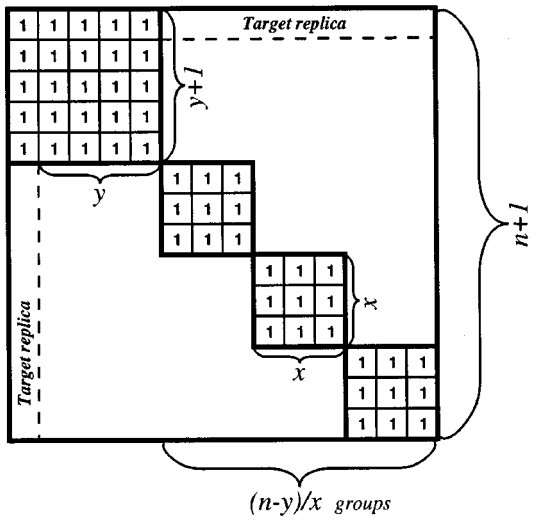
For $d>2$ we find two maxima (we have to maximize free energy at $n<1$), namely, $R_t^{\alpha\beta}=\infty$ and $R_t^{\alpha\beta}=0$. The first corresponds to two replicas, $\alpha$ and $\beta$, which are independent and do not overlap at all ($Q_{\alpha\beta}=0$), while the second corresponds to replicas that coincide at the microscopic level [$Q_{\alpha\beta}=\rho\delta(\mathbf{R}_1-\mathbf{R}_2)$]. Thus, from these scaling arguments in $R_t^{\alpha\beta}$, we conclude that $Q_{\alpha\beta}$ is of the form

$$\begin{aligned} Q_{\alpha\beta}(\mathbf{R}_1,\mathbf{R}_2)&=\rho q_{\alpha\beta}\delta(\mathbf{R}_1-\mathbf{R}_2)\\ &\Rightarrow \hat{Q}^{((n+1)\infty)}\\ &=\rho\hat{q}^{(n+1)}\otimes\hat{I}^{(\infty)}, \end{aligned} \tag{3.33}$$

where off-diagonal elements of the new matrix $q_{\alpha\beta}$ are either 0 or 1, while $q_{\alpha\alpha}=1$.

The situation is completely different in $d \leqslant 2$, where an optimum in Eq. (3.32) is reached at some nontrivial intermediate scale. Note that the simple arguments of Sec. III.A.2 also indicate complications at and below $d=2$. It can be seen in both cases that the main problem in low dimensions, as opposed to $d>2$, is that most of the neighbors of each monomer in space are also neighbors along the polymer. One aspect of the complications is that replica symmetry breaking is of a more delicate form, as discussed by Shakhnovich and Gutin (1989b). Moreover, the use of the mean-field approximation is problematic for this case. We shall not further explore this low-dimensional case here and simply remark that this problem is poorly understood at present.

Coming back to $d=3$, we are now left with optimization with respect to a much simpler object, i.e., an **R**-independent matrix $q_{\alpha\beta}$. To maximize the $n$-replica free energy over $q_{\alpha\beta}$ means in fact finding the optimal grouping of replicas. Indeed, there is the following obvious transitivity rule: if, say, $R_t^{\alpha\beta}=0$ and $R_t^{\beta\gamma}=0$, meaning that conformations of replicas $\alpha$, $\beta$, and $\gamma$ are all the same, then $R_t^{\alpha\gamma}=0$ as well. In other words, if $q_{\alpha\beta}=1$ and $q_{\beta\gamma}=1$, then $q_{\alpha\gamma}=1$ as well. Using matrix row and column operations, we can organize any such matrix into block-diagonal form. This means gathering replicas that overlap in the groups and placing replicas of the same group into the same diagonal block in the matrix. One of the blocks is composed of $z=y+1$ replicas that do overlap (i.e., practically coincide) with the "target" replica 0 (i.e., conformation $\star$). The other $(n+1)-(y+1)=n-y$ replicas belong to $n/x$ groups, with $x$ replicas in each:




This is, of course, just a one-step Parisi matrix (Mezard *et al.*, 1987).

One can say that $y$ replicas here are "adsorbed" on the target conformation, which plays the role of an external field. A similar situation exists in neural networks (Amit *et al.*, 1987), where the memorized image plays a role analogous to that of the target conformation. On the other hand, the grouping of the replicas in the $x \times x$ size groups is due to spontaneous replica permutation symmetry breaking.

#### *d. Uncorrelated energy landscapes, or the random-energy model*

It is important to stress that the character of replica symmetry breaking implies that the statistical properties of a large incompressible heteropolymer globule are actually those of the random-energy model (Derrida, 1980). Indeed, the Shakhnovich and Gutin argument (3.32), with its consequence (3.33), indicates that replicas either coincide on a microscopic scale or are totally independent. In other words, they represent either the same conformation (microstate), or two unrelated conformations. This is precisely the defining property of the random-energy model, in which different microstates are assumed to be statistically independent.[12]

#### *e. Replica free energy*

Returning to our calculations, we note that our order parameter has simplified tremendously: from a complex object $Q_{\alpha\beta}(\mathbf{R}_1,\mathbf{R}_2)$, to a much simpler $q_{\alpha\beta}$, and finally to just two numbers, $x$ and $y$, which describe the grouping of replicas.

Of course, grouping of replicas costs entropy. To estimate this entropy, one has to bear in mind that conformations of the replicas coincide down to the microscopic scale if the two replicas belong to the same group, as is manifested in $Q_{\alpha\beta}(\mathbf{R}_1,\mathbf{R}_2) \sim \delta(\mathbf{R}_1-\mathbf{R}_2)$ [see Eq. (3.33)]. Consider, for instance, a group of $w$ replicas. The conformation of one of them can be chosen arbitrarily, while the $w-1$ other conformations must coincide with the chosen one. For any one replica, the entropic price associated with fixed conformation is obviously proportional to the number of monomers[13]; we write it as $-Ns$. For $(n-y)/x$ groups of $x$ replicas each and one target group of $y+1$ replicas, the entropy is therefore

[12]We remind the reader, that the estimates discussed in Sec. III.A.2 suggest the use of the random energy model. Thus it is reasonable to study heteropolymer freezing phenomenologically, i.e., assuming the validity of the random-energy model from the very beginning. Historically this approach was first implemented in the seminal work of Bryngelson and Wolynes (1987). Recently we have formulated a simple yet rigorous approach allowing one to avoid the replica trick, while systematically controlling and understanding the relevant approximations employed (Pande *et al.*, 1997c); in particular, using this approach (discussed in detail in Sec. IV.A.3), one can start with a microscopic Hamiltonian as in replica approaches and derive aspects of the freezing transition, such as the effect of interactions on the freezing temperature, without resorting to the replica trick. The price one pays for this simplicity is that one must, as in the early works of Bryngelson and Wolynes (1987), assume the validity of the random-energy model rather than deriving it. The applicability of the model will be examined in a separate section (Sec. V).

[13]As we have already noted, it would be more accurate to count contacts between monomers rather than monomers themselves, and $N$ should be understood throughout this section as the number of contacts in the maximally compact globule.

$$S=-Ns\left[\frac{n-y}{x}(x-1)+y\right]. \tag{3.34}$$

The sign is minus, because this is entropy *reduction*; $s$ is a microscopic parameter characterizing polymer chain chemical bonds (polymer rigidity). We discuss an estimate of $s$ for real proteins later in Sec. VII.B.2.a.

It is conceptually simple, although rather cumbersome, to simplify the expression for the energy (3.25) by removing both real and replica spaces and expressing the effective Hamiltonian in terms of $x$ and $y$. As there is either complete or no overlap, replicas of different groups do not interact. This is why the resulting expression for the energy has a structure similar to that of the entropy (3.34): the number of nontarget groups $(n-y)/x$ times the energy of one nontarget group, plus the energy of the target group. This is detailed in Appendix C.

The free energy of the replica system finally reads

$$\begin{aligned}\frac{F(x,y)}{NT}=\frac{n-y}{2x}\Bigg\{&\ln\det\left[\hat{I}+x\frac{\hat{\Delta}\hat{B}}{T}\right]\\&+\left\langle\vec{p}\left|x\frac{\hat{B}}{T}\left[\hat{I}+x\frac{\hat{\Delta}\hat{B}}{T}\right]^{-1}\right|\vec{p}\right\rangle\Bigg\}\\&+\frac{1}{2}\ln\det\left[\hat{I}+\frac{\hat{\Delta}\hat{B}^{\mathrm{des}}}{T^{\mathrm{des}}}+y\frac{\hat{\Delta}\hat{B}}{T}\right]\\&+\left\langle\vec{p}\left|\left[\frac{\hat{B}^{\mathrm{des}}}{2T^{\mathrm{des}}}+y\frac{\hat{B}}{2T}\right]\left[\hat{I}+\frac{\hat{\Delta}\hat{B}^{\mathrm{des}}}{T^{\mathrm{des}}}+y\frac{\hat{\Delta}\hat{B}}{T}\right]^{-1}\right|\vec{p}\right\rangle\\&+s\left[y+(n-y)\frac{(x-1)}{x}\right],\end{aligned} \tag{3.35}$$

where $\hat{B}^{\mathrm{des}}$ is the preparation matrix and $T^{\mathrm{des}}$ is the design temperature. From now on, we shall drop the superscript indications of dimensionalities, as all vectors and matrices are now in species space only, which is $q$ dimensional.

#### f. Physical meaning of the $\hat{\Delta}$ operator

Before further analysis, let us examine the physical meaning of the operator $\hat{\Delta}$ and the term $\hat{\Delta}\hat{B}$ which appear throughout our formulae. From the definition of $\hat{\Delta}$ in Eq. (3.22), we have

$$(\hat{\Delta}\hat{B})_{ik}=\sum_j (p_i\delta_{ij}-p_ip_j)B_{jk}=p_iB_{ik}-\sum_j p_ip_jB_{ij}. \tag{3.36}$$

We can always write $B_{ij}$ in terms of a sum of a homopolymeric attraction $(B_0)$ and heteropolymeric deviations: $b_{ij}=B_{ij}-\langle B\rangle$. From Eq. (3.36), we see that $\hat{\Delta}$ removes the mean interaction of species $k$ from all matrix elements $B_{kj}$. In other words, $\hat{\Delta}$ removes all homopolymeric effects. Not surprisingly, if we formally take $\hat{\Delta}\hat{B}=0$, we recover the familiar homopolymer results.

#### g. Reduction theorems

There are several cases in which the same physical system can be depicted in terms of formally different interaction matrices $\hat{B}$, $\hat{B}^{\mathrm{des}}$ and/or composition vectors $\vec{p}$. Clearly, the expressions for the freezing and folding temperatures, as well as for any other real physical quantity, must not depend on any arbitrary choice.

For example, there might be a monomer species that is formally included in the list and in the interaction matrix but that is not physically present in the chain, i.e., $p_q=0$. It is easy to check that in this case Eq. (3.35) is reduced to a smaller list of $q-1$ monomer species with a $(q-1)\times(q-1)$ interaction matrix. Another example is the case of duplicate species: consider species labeled $q$ and $q-1$ which are physically identical, i.e., they interact in identical ways to all other species. Physically, we would expect this problem to be identical to the $q-1$ species case, except with the new composition $p'_{q-1}=p_{q-1}+p_q$, and this was confirmed by Pande *et al.* (1995a, 1995b), Eq. (3.35).

These two statements, which we call "reduction theorems," are both checks of the consistency of our result (3.35) and are often useful for other applications, as well (see, for example, Du *et al.*, 1998).

### 4. Phase transitions in the large-globule model

Equation (3.35) depends on two order parameters, $x$ and $y$, implying that there are at least three phases in the corresponding phase diagram: random, glassy (or frozen), and folded (or target).[14] To see the structure of the phase diagram, which we discuss in the next section, we first look at the allowed variations of the order parameters $x$ and $y$.

For simplicity, we consider here only the small-$s$ regime. In this case, freezing transitions, which are our chief concern here, occur when $B$, or, more precisely, $\hat{\Delta}\hat{B}$, is also small. Indeed, freezing phase transitions result physically from the competition between energetic and entropic parts of the free energy (3.35), where the energetic part favors the gathering of replicas into groups while the entropic part favors the diversity of replicas. For the energy to be competitive with entropy when $s$ is small, $B$ must be small as well. This allows us to simplify Eq. (3.35), truncating it to quadratic order in $B$. Moreover, the small-$s$ and small-$\hat{\Delta}\hat{B}$ regime is the most reasonable regime to consider after we have truncated the source term to $\mathcal{O}(\phi^2)$ in Eq. (3.17). In principle, Eq. (3.35), obtained by the truncation of $\mathcal{O}(\phi^2)$, may be valid outside this regime (i.e., not-so-small values of $s$); for this case, phase transitions are discussed by Pande *et al.* (1995a, 1995b). However, if $s$ is small, then everything is simpler and all approximations are consistent with each other.

[14] We remind the reader that we consider here the collapsed maximally compact globule, and thus the globule-to-coil phase transition falls beyond the scope of the present section. This will be covered in Sec. VI.B.

As $y$ is the number of replicas whose conformation coincides with the target conformation, this value must lie between 0 and $n$. What is relevant in the replica approach is the $n\to 0$ limit; only terms that are linear in $n$ are to be considered, because higher-order terms disappear in the main equation $\langle \ln Z\rangle = \lim_{n\to 0}(\langle Z^n\rangle - 1)/n$. Accordingly, since $0\leqslant y\leqslant n$, we must linearize the free energy in $y$ as well (Pande, Grosberg, and Tanaka, 1994c, 1995b). This leads to a further simplification of Eq. (3.35):

$$\begin{aligned}\frac{F}{NT} = \mathrm{Tr}\Bigg[ (n-y)\left\{\frac{\hat\Delta\hat B}{2T} - x\frac{\hat\Delta\hat B\hat\Delta\hat B}{4T^2} + \frac{\hat P\hat B}{2T} - x\frac{\hat P\hat B\hat\Delta\hat B}{2T^2}\right\} \\ + \frac{\hat\Delta\hat B^{\mathrm{des}}}{2T^{\mathrm{des}}} + y\frac{\hat\Delta\hat B}{2T} - y\frac{\hat\Delta\hat B^{\mathrm{des}}\hat\Delta\hat B}{2TT^{\mathrm{des}}} - \frac{\hat\Delta\hat B^{\mathrm{des}}\hat\Delta\hat B^{d}}{4T^{\mathrm{des}^2}} \\ + \frac{\hat P\hat B^{\mathrm{des}}}{2T^{\mathrm{des}}} + y\frac{\hat P\hat B}{2T} - \frac{\hat P\hat B^{\mathrm{des}}\hat\Delta\hat B^{\mathrm{des}}}{2T^{\mathrm{des}^2}} - y\frac{\hat P\hat B\hat\Delta\hat B^{\mathrm{des}}}{2TT^{\mathrm{des}}} \\ - y\frac{\hat P\hat B^{\mathrm{des}}\hat\Delta\hat B}{2TT^{\mathrm{des}}}\Bigg] + s\left[y + (n-y)\frac{(x-1)}{x}\right], \end{aligned} \tag{3.37}$$

where $P_{ij}\equiv p_i p_j$.

While $y$ describes the breaking of the symmetry between $n$ replicas due to their attraction to the target replica labeled 0, $x$ describes spontaneous symmetry breaking. When we have an integer number of replicas $n$, clearly $1\leqslant x\leqslant n$: $x$ cannot be smaller than unity, because it is the number of replicas in the group. When $n\to 0$, this constraint on the number of replicas in the group is no longer applicable, but it is natural to think that formal inequalities for $x$ simply flip signs: $n\leqslant x\leqslant 1$. With this in mind, we optimize the free energy (3.37) with respect to $x$, yielding the equation that determines $x$:

$$\begin{aligned} s &= \frac{x^2}{4T^2}\mathrm{Tr}[\hat\Delta\hat B\hat\Delta\hat B + 2\hat P\hat B\hat\Delta\hat B] \\ &= \frac{x^2}{4T^2}\left[\sum_{i,j} p_i B_{ij} B_{ij} p_j - \sum_{i,j} p_i B_{ij} p_j \sum_{k,l} p_k B_{kl} p_l\right]. \end{aligned} \tag{3.38}$$

The freezing transition occurs at the temperature at which $x=1$, when replicas start to group. This is a transition to a unique ground state which is not necessarily (and most likely not) the target conformation.

To examine freezing to the target conformation, we must look at the conditions at which $y>0$. Since only the free-energy terms linear in $y$ have physical meaning in the $n\to 0$ limit, the free-energy optimum corresponds to either $y=0$ (nontarget phase) or $y=n$ (target phase). To find the corresponding critical temperature, we must examine the slope of the free energy in $y$ to determine whether $y=0$ or $y=n$ is the stable solution (Pande *et al.*, 1995b). The condition ''slope''$=0$ yields the relationship

$$\begin{aligned} s &= \mathrm{Tr}\Bigg[\frac{x}{4}\left(\frac{\hat\Delta\hat B}{T}\frac{\hat\Delta\hat B^{\mathrm{des}}}{T^{\mathrm{des}}} + 2\frac{\hat P\hat B}{T}\frac{\hat\Delta\hat B^{\mathrm{des}}}{T^{\mathrm{des}}} + \frac{\hat\Delta\hat B^{\mathrm{des}}}{T^{\mathrm{des}}}\frac{\hat\Delta\hat B}{T} \right. \\ &\quad \left. + 2\frac{\hat P\hat B^{\mathrm{des}}}{T^{\mathrm{des}}}\frac{\hat\Delta\hat B}{T}\right) - \frac{x^2}{4T^2}(\hat\Delta\hat B\hat\Delta\hat B + 2\hat P\hat B\hat\Delta\hat B)\Bigg] \\ &= \frac{x}{2TT^{\mathrm{des}}}\Bigg[\sum_{i,j} p_i B_{ij} B^{\mathrm{des}}_{ij} p_j \\ &\quad - \sum_{i,j} p_i B_{ij} p_j \sum_{k,l} p_k B^{\mathrm{des}}_{kl} p_l\Bigg] - \frac{x^2}{4T^2}\Bigg[\sum_{i,j} p_i B_{ij} B_{ij} p_j \\ &\quad - \sum_{i,j} p_i B_{ij} p_j \sum_{k,l} p_k B_{kl} p_l\Bigg]. \end{aligned} \tag{3.39}$$

By examining the structure of our results (3.38) and (3.39), we may see quite clearly their natural geometric parametrization.

### C. Geometric parametrization of the interaction matrices

While Eq. (3.39) describes the folding transition for all interaction matrices, this general approach yields formulas that are difficult to understand physically. Part of this difficulty is a lack of means to understand differences between interaction matrices. We have seen that the mean and variance of the matrix are two properties that are important in folding, but clearly matrices differ in other ways. In this section, we present a way of describing matrices using correlators that are relevant for folding and design.

To describe a given interaction matrix, we formally define an abstract vector space in which $q\times q$ interaction matrices $\hat B$ are now considered to be vectors. The rules necessary to define a vector space trivially follow: $\hat B^1 + \hat B^2 \to B^1_{ij} + B^2_{ij}$, $c\hat B \to cB_{ij}, \ldots$. As interaction matrices are symmetric ($B_{ij}=B_{ji}$), our space is $q(q+1)/2$ dimensional. Let us define the average interaction

$$\overline{B} = \sum_{i,j} p_i B_{ij} p_j, \tag{3.40}$$

which is a particular component of our vector $\hat B$, and also

$$\widehat{\delta B}_{ij} = B_{ij} - \overline{B}, \tag{3.41}$$

which represents the vector $\widehat{\delta B}$ with zero component along the direction of the mean.

Most important, for any two vectors in this space, say $\widehat{\delta B}^\dagger$ and $\widehat{\delta B}'$, we can define the scalar product as

$$\widehat{\delta B}^\dagger \cdot \widehat{\delta B}' = \sum_{ij} p_i\, \delta B^\dagger_{ij}\, \delta B'_{ij} p_j. \tag{3.42}$$

Note that this has nothing to do with the matrix product of the corresponding matrices. Accordingly, the norm, or length, of the vector $\widehat{\delta B}$ is defined via

$$\delta B^2 = \sum_{i,j} p_i\, \delta B_{ij}\, \delta B_{ij} p_j = \widehat{\delta B}\cdot\widehat{\delta B}. \tag{3.43}$$

In other words, $\delta B$ is the variance of the interaction matrix and details the diversity of the interaction energies $B_{ij}$.

Finally, we define an angle $\phi_{\dagger,}$ between vectors $\widehat{\delta B}^{\dagger}$ and $\widehat{\delta B}'$, or its cosine $g_{\dagger,}=\cos\phi_{\dagger,}$, as

$$g \equiv \frac{\widehat{\delta B}^{\dagger}\cdot\widehat{\delta B}'}{\sqrt{(\widehat{\delta B}^{\dagger}\cdot\widehat{\delta B}^{\dagger})(\widehat{\delta B}'\cdot\widehat{\delta B}')}}. \tag{3.44}$$

Obviously, $g$ can also be viewed as a connected correlator, or "cumulant," of two matrices. "Parallel" (completely correlated) interactions yield $g=1$ ($\phi=0$) and "orthogonal" (completely uncorrelated) interactions yield $g=0$ ($\phi=\pi/2$).

Thus an arbitrary interaction matrix can be written in the form

$$\hat{B}=\overline{B}+\delta B\,\hat{b}, \tag{3.45}$$

where $\hat{b}$ is a matrix of zero mean and unit variance and serves as the unit vector in our vector space. It is important to note that the angle between two vectors $\hat{B}^1$ and $\hat{B}^2$ is the same as that between the corresponding $\hat{b}^1$ and $\hat{b}^2$. Obviously, unit vectors $\hat{b}$ should be parametrized in terms of the corresponding spherical angles.

For instance, in the two-species, even composition case ($q=2, p_1=p_2=1/2$), there is only one polar angle $\theta$ to parametrize such a matrix, and the matrix can be presented in the form[15]

$$b_{ij}(\theta)=\sigma_i\sigma_j\sin\theta+\frac{1}{\sqrt{2}}(\sigma_i+\sigma_j)\cos\theta, \tag{3.46}$$

where $\sigma_i=\pm1$ are "hidden Ising spins" assigned to each monomer. What is nice about this parametrization is

$$b(\theta_1)\cdot b(\theta_2)=\cos(\theta_1-\theta_2). \tag{3.47}$$

This approach, which, is known as the generalized black-and-white model, is described in greater detail by Pande, Grosberg, and Tanaka (1997c). At particular angles the matrix described by Eqs. (3.45) and (3.46) reproduces various particular cases studied in the literature (Pande, Grosberg, and Tanaka, 1997c).

For example, one of the most widely used models is the hydrophobic-polar model, which makes the simplification that hydrophobic interactions provide the most important driving force for protein collapse. The energy of a protein globule is taken to be proportional to the number of contacts between hydrophobic monomers. In the spirit of Eq. (3.45), the matrix $\hat{B}$ should be written as

$$\begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}=-\frac{1}{4}+\frac{\sqrt{3}}{4}\begin{pmatrix} -1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{3} & 1/\sqrt{3} \end{pmatrix} \tag{3.48}$$

which, by comparison to Eq. (3.46), means that the angle $\theta$ for the hydrophobic-polar model is $\theta_{\mathrm{HP}}=\arccos[\sqrt{2/3}]\approx35°$. Curiously, this angle corresponds to the maximal skewness of the matrix (3.46) $\eta(\theta)=(1/4)\Sigma_{i,j}b_{ij}^3=-3\cos^2\theta\sin\theta$.

In the next section, we shall examine the consequences of our results (3.38), (3.39), and discuss the corresponding phase diagram.

## IV. PROPERTIES OF HETEROPOLYMERS IN THE RANDOM-ENERGY MODEL

In Sec. III, we found that a random-energy type freezing phase transition occurs in a heteropolymer globule under the appropriate circumstances. In this section, we describe the heteropolymeric properties that result from a random-energy-model description.

### A. Phase diagram

We begin with the phase diagram of the large incompressible globule model (Sec. III.B). All of the phases here are globular (i.e., compact). Since our system is described by two order parameters, $x$ and $y$, we expect three globular phases, which we call random, frozen (or glassy), and folded (or target). We have already arrived at the condition (3.38), which determines the equilibrium grouping of nontarget replicas, and the condition (3.39), which determines the conditions of the "switch" between $y=0$ (none of the replicas in the target group) and $y=n$ (all of the replicas in the target group).

In terms of our geometric definitions, Eqs. (3.38) and (3.39) read

$$s=\frac{x^2}{4T^2}\,\delta B^2; \tag{4.1}$$

$$s=\frac{x}{2TT^{\mathrm{des}}}\,\delta B\,\delta B^{\mathrm{des}}\,g-\frac{x^2}{4T^2}\,\delta B^2, \tag{4.2}$$

where the $g$ factor has been defined in Eq. (3.44), $T^{\mathrm{des}}$ is the design temperature, and $B^{\mathrm{des}}$ is the preparation matrix. We now examine the consequences of these results, starting from the simplest case, random sequences.

#### 1. Random sequences

Random sequences are obtained in the limit $T^{\mathrm{des}}\to\infty$. Freezing occurs when replicas start to group, or $x$ starts to deviate from 1, thus spontaneously breaking the permutation symmetry. Therefore the freezing temperature is given by the relation

$$T_{\mathrm{freeze}}^2=\delta B^2/4s. \tag{4.3}$$

In other words, the freezing temperature is given by the variance of the renaturation interaction matrix (Pande *et al.*, 1995a). Note that this is a transition to a unique ground state, which is not necessarily (and most likely not) the target conformation. We call this phase the *glassy* (or frozen) phase and we call the high-temperature disordered phase in which there is no form of freezing (i.e., many conformations dominate equilibrium) the *random* phase. In this sense, the random phase is actually reminiscent of a large homopolymer globule.

[15] A similar but more cumbersome expression for arbitrary $p_1$ and $p_2$ ($p_1+p_2=1$) is given by Pande, Grosberg, and Tanaka (1997c).

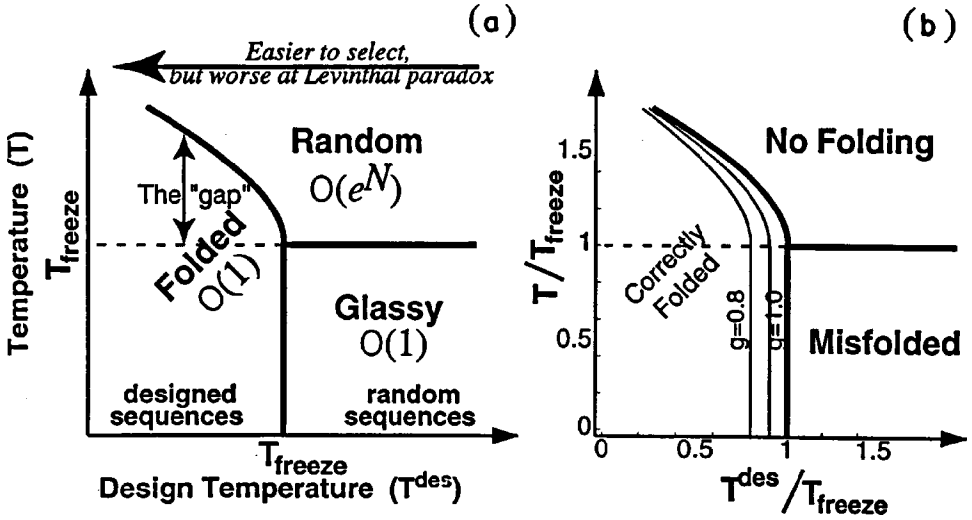



FIG. 1. Phase diagram of globular heteropolymers, based on the random-energy model. There are three phases: (1) *Random*: an exponential number of globular conformations dominate equilibrium (similar to a homopolymer globular state); (2) *Glassy*: for sequences that are not well optimized (sufficiently high design temperature $T^{\rm des}$), only one (or a few) conformations dominate below the glass temperature $T_{\rm freeze}$, but these conformations are not the target conformation; one can consider random-sequence ground states to be "poorly optimized" at $T^{\rm des}=T_{\rm freeze}$; (3) *Folded*: the target conformation $\star$ dominates equilibrium; for $T^{\rm des}<T_{\rm freeze}$, $\star$ is better optimized than the ground state of random sequences. In (a) these three phases are shown for a heteropolymer whose sequence has been designed using the same interactions that govern folding. The gap is shown in which folding is the most reliable, as the target state is thermodynamically stable, while freezing into any other conformation is not even metastable. The gap is broader for the sequences prepared at low $T^{\rm des}$. This is why lower $T^{\rm des}$ leads to better folders, while these sequences are harder to select (there are far fewer of them). Part (b) shows a similar diagram for folding and design with interactions that are somewhat different. In this case, the phases are labeled a little differently: when the system is pushed into the random globule state, no folding is observed, i.e., exponentially many conformations are found in equilibrium. If the ground state is destabilized, but the acting temperature is below the freezing temperature, the system freezes like a random sequence and the protein misfolds.

The equilibrium value of the order parameter $x$ is thus equal to

$$x=\begin{cases}1 & \text{when } T>T_{\rm freeze}\\ T/T_{\rm freeze} & \text{when } T<T_{\rm freeze}.\end{cases} \tag{4.4}$$

The linearity of $x$ in $T/T_{\rm freeze}$ below the freezing point indicates a first-order phase transition (Mezard *et al.*, 1987). We also remind the reader that according to the general interpretation (Mezard *et al.*, 1987), $x$ is equal to

$$x(T)\equiv 1-\sum_{\alpha} P_{\alpha}^2, \tag{4.5}$$

where the sum is over all maximally compact conformations $\alpha$ and $P_\alpha$ is the probability that conformation $\alpha$ is found in equilibrium i.e., $P_\alpha=W_\alpha/Z$, where $W_\alpha=\exp[-\mathcal{H}(\text{conf}=\alpha,\text{seq})/T]$ is the Boltzmann weight of conformation $\alpha$ and $Z=\Sigma_\beta W_\beta$ is the partition function. Clearly, $x$ is related to the number of thermodynamically relevant (micro)states, $\mathcal{M}$: assuming $P_\alpha\sim 1/\mathcal{M}$, we get $\mathcal{M}=1/(1-x)$. Thus this number decreases sharply below $T_{\rm freeze}$.

Note that Eq. (4.2) does not contain the design temperature $T^{\rm des}$ at all, which means that the freezing temperature, at which the glassy phase becomes thermodynamically favored over the random phase, is independent of $T^{\rm des}$. This is clearly the result of the statistical independence of states: while design at finite $T^{\rm des}$ does affect the energy of the chain in the target conformation $\star$, it does not affect the energies of any other conformation and thus does not affect $T_{\rm freeze}$. This is reflected in Fig. 1(a), in which the phase diagram is shown in terms of the acting temperature $T$ (at which a polymer "lives" in solution) and design temperature $T^{\rm des}$ (which controls the type of the sequence). We see that at sufficiently high $T^{\rm des}$ there is a line of freezing transitions that is horizontal (i.e., independent of $T^{\rm des}$). Heteropolymer freezing for random-sequence heteropolymers was first described by Shakhnovich and Gutin (1989a).

### 2. Designed sequences

Equation (4.2) describes freezing to the target conformation $\star$. We call this the *folding transition*, and the corresponding phase is the *folded* (or target) *phase*. In this phase, the system freezes to the target conformation. The folding transition is marked by the conditions at which replicas start to group with the target replica, i.e., $y$ switches from 0 to $n$. To write the equation that defines the folding phase boundary, it is convenient to define the boundary temperature $T_{\rm freeze}^{\rm des}$ according to the equation

$$(T_{\rm freeze}^{\rm des})^2=\delta B_{\rm des}^2/4s. \tag{4.6}$$

Equation (4.6) is similar to Eq. (4.3), except it includes a preparation matrix $\hat{B}^{\rm des}$ instead of $\hat{B}$; physically, $T_{\rm freeze}^{\rm des}$

is the temperature at which a random heteropolymer would undergo a freezing transition provided its conformations are governed by the $\hat{B}^{\rm des}$ interaction matrix. $T^{\rm des}_{\rm freeze}$ sets a natural scale for $T^{\rm des}$. Solving for the condition in which $y$ switches from 0 to $n$, we find the boundary of the folded phase, which is easier to write in terms of the critical value of the design temperature:

$$\frac{T^{\rm des}_{\rm cr}}{T^{\rm des}_{\rm freeze}} = \begin{cases} g\,\dfrac{2T/T_{\rm freeze}}{1+T^2/T^2_{\rm freeze}} & \text{for } T \geqslant T_{\rm freeze} \\ g & \text{for } T \leqslant T_{\rm freeze}, \end{cases} \tag{4.7}$$

where $T^{\rm des}_{\rm cr}$ is the critical value of $T^{\rm des}$ at which the design of sequences becomes effective enough to make the target conformation the ground state, and $g$ is defined in Eq. (3.44).

Equation (4.7) implies that the boundary between the folded and glassy phases is always vertical on the phase diagram. Equation (4.7) has a clear physical meaning: as both phases are characterized with negligibly small entropies, the transition between the two cannot be controlled by a temperature change. As $T^{\rm des}$ is a quenched property (due to the quenched sequence), for $g=1$, a given sample may exhibit either a regular freezing transition to the glassy phase or a folding phase transition to the target phase, but not both. As long as we consider the vicinity of the triple point on the phase diagram (Fig. 1), Eq. (4.7) can also be rewritten in terms of $T_{\rm fold}$, the temperature at which the random-to-folded phase transition occurs:

$$\frac{T_{\rm fold}}{T_{\rm freeze}} = 1+\sqrt{2}\left[1-\frac{T^{\rm des}}{gT^{\rm des}_{\rm freeze}}\right]^{1/2}. \tag{4.8}$$

#### *a. Design and folding governed by the same set of interactions*

We now consider the simplest case: $\hat{B}=\hat{B}^{\rm des}$ or $g=1$, for which $T^{\rm des}_{\rm freeze}=T^{\rm des}$. The phase diagram shown in Fig. 1(a) represents the central point of this review. Its applicability to specific systems such as proteins will be discussed in Sec. VII. The main feature of the diagram is that for well-designed sequences, the folding phase transition occurs at a higher temperature than the freezing transition of a random sequence.

To gain insight into the physical nature of this result, let us consider a model with a discrete set of conformations, such as a lattice model, and discuss the spectrum of energies of all conformations. For a random sequence, this spectrum typically looks like those shown in Fig. 2: there is a very high density of states at medium energy and a few, sparse low-energy levels. This is a typical random-energy-type spectrum (see property 2, Sec. II.A.2). At low $T^{\rm des}$, sequence design selects sequences whose energy in the target state is low: design "pulls down" the target-state energy. Decreasing $T^{\rm des}$ lowers the target-state energy of the corresponding sequences. As long as the assumption of the random-energy model is valid, the energies of all other conformations remain statistically unaffected by design, and thus design does not pull the energies of nontarget conformations, as is illustrated in Fig. 2. It is therefore not a surprise that, at low $T^{\rm des}$, freezing to the target state occurs at a higher temperature than that for freezing to a randomly chosen conformation of the glassy phase.

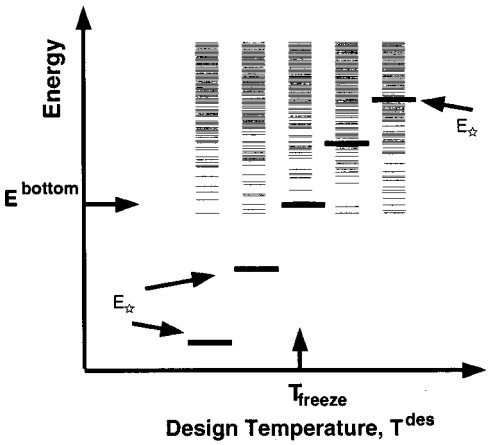



FIG. 2. Typical random-energy-model spectra. Design affects one particular energy level. When the design temperature $T^{\rm des}$ is high, this level is lost in the sea of others. When $T^{\rm des}$ decreases, it is pulled down out of the sea of random-energy states.

Furthermore, if we consider a polymer whose sequence has been prepared at $T^{\rm des}<T_{\rm freeze}$, we then find a *gap* between the folding temperature $T_{\rm fold}$ and the glass temperature $T_{\rm freeze}$. This gap in temperature reflects the gap in the energy spectrum between the low "pulled" target-state energy and the other higher levels.

The question of the gap has been extensively discussed in the literature,[16] both in terms of the difference between temperatures $T_{\rm fold}$ and $T_{\rm freeze}$ and in terms of the gap between energy levels. Based on phenomenological arguments, it has been argued that in order for a protein to fold successfully, there must be two distinct temperature scales, with a gap between them (Shakhnovich, 1997). It seems reasonable to identify these two characteristic temperatures with $T_{\rm fold}$ and $T_{\rm freeze}$, respectively, which are well-defined quantities. There was also a debate in the literature (Klimov and Thirumalai, 1996; Shakhnovich, 1997) as to the existence, meaning, and significance of the gap in the energy spectrum. As we see, in the random-energy theory there is a one-to-one correspondence between the two gaps ($T_{\rm fold}-T_{\rm freeze}$ and $E-E_\star$.) We shall show, however, that a more accurate theory, which incorporates corrections to the random-energy model and deviations from it, predicts that a small number of states have energies pulled down together with that of the native state, so that the idea of a gap in the energy spectrum appears valid but requires further clarification (see Sec. VII below). More-

[16]See, for example, Goldstein *et al.*, 1992; Shakhnovich and Gutin, 1993; Pande *et al.*, 1995b; Socci and Onuchic, 1995; Klimov and Thirumalai, 1996; Shakhnovich, 1997.

over, these states appear to be decisively important for the kinetics of folding.

From the discussion above, we can draw one more conclusion regarding random sequences in our model. The limit $T^{\rm des}\to\infty$ means that all sequences are weighted equally. Thus the target conformation $\star$ has no meaning. However, it is often convenient to consider folding properties of random sequences, where $\star$ is now the ground state of a given random sequence. In this sense, one can view random sequences as sequences designed at the temperature that is equal to that of freezing: $T^{\rm des}=T_{\rm freeze}$. Indeed, when $T^{\rm des}=T_{\rm freeze}$, design has not yet done anything to the overall appearance of the spectrum of energy levels, since $E_\star$ lies where one normally expects to find the lower energy levels. Due to design, we only know which particular state is the ground state, while in all other respects the properties of the sequence are just that of any random sequence. In other words, for $T^{\rm des}=T_{\rm freeze}$, we select ''typical'' (i.e., not optimized) sequences that have the desired ground state $\star$. Similarly, $T^{\rm des}=T_{\rm freeze}+\epsilon$, where $\epsilon$ is small and positive, has the physical meaning of selecting sequences that have $\star$ as a low-energy state, but not necessarily the lowest.

#### *b. Design and folding governed by two different sets of interactions*

When design and folding are governed by two different sets of interactions, the interactions can be different in a number of ways. Most simply, one can imagine that the polymer folds under different solvent conditions compared to the environment in which it was designed. As another example, consider a computer simulation of protein folding or design. One must approximate the nature of the interaction potentials involved. This leads directly to a problem: we believe that protein sequences have been designed by Nature, and this natural design was obviously governed by natural interactions; we are now trying to fold this same polymer using simulated potentials, which can be considered as a distorted or ''noisy'' copy of the ''true'' interactions. Equivalently, consider the computer design of sequences using approximated potentials and then folding experimentally *in vitro*.

The difference between interactions affects the phase diagram. According to Eq. (4.7), the value $T_{\rm cr}^{\rm des}$ is proportional to $g$. The appearance of a positive but incomplete correlation ($0<g<1$) leads to an affine deformation of the boundary of the target-phase region, as illustrated in Fig. 1(b). At $g=0$, the target region disappears, and it does not exist at $g<0$. This is clear because when matrices $\hat{B}$ and $\hat{B}^{\rm des}$ are anticorrelated, design does not help, but rather inhibits folding into the desired conformation, by raising the energy of $\star$. This consideration also makes clear that the very existence of a correctly folded phase is due to design: design optimizes the ground-state energy to such an extent that moderate mistakes do not harm its stability. We also conclude that the ''angle'' between matrices, given by the factor $g$ (3.44), is an adequate measure of their proximity.

One can use the difference between interactions to simulate mutations. Experimentally, it is extremely informative and useful to design new mutant proteins (by means of *protein engineering*) and examine the changes in folding in response to the mutations. In light of our discussion, we can assume that mutations change the degree of optimization of the native-state conformation (which presumably remains close enough to the original, unmutated native-state conformation). As long as that is the case, a simple way to simulate mutants would be to change the degree of optimization by designing the sequence under somewhat distorted interactions, in other words, using interactions that are ''turned'' by a certain mutation angle $\delta\theta$ compared to folding ones (see Sec. III.C). This approach has been fruitfully used by Pande *et al.* (1997b) and Vendruscolo *et al.* (1997).

### 3. ''No-replica'' derivation of the phase diagram

The theory derived above in Sec. III.B is a truly microscopic theory, in the sense that its starting point is a microscopic, albeit simplified, Hamiltonian (2.5). The fundamental approximation involved in this theory can be called one-step replica symmetry breaking with no overlap between groups. This is equivalent to the approximation of uncorrelated energies of the (micro)states. We have provided proof of the validity of this approximation for a large incompressible globule in 3D, both through a simple physical argument (see Sec. III.A.3) and by more formal considerations (Sec. III.B.3.c), as was first done by Shakhnovich and Gutin (1989a). However, if one is interested only in getting to specific results, one may choose to skip the question of whether the random-energy model is valid and explicitly take this model as an assumption, or (educated) guess, from the very beginning, as was done originally by Bryngelson and Wolynes (1987). Then it is possible to derive the entire phase diagram (Fig. 1) based on this assumption. In this section, we briefly describe this derivation (for details, see Pande *et al.*, 1997c).

The approach is based on Eq. (2.4), which states that for a random-energy model above the freezing temperature, the quenched average free energy $\langle\ln Z\rangle$ coincides with the annealed average $\ln\langle Z\rangle$. Thus one need compute only the annealed average free energy, which is considerably simpler than calculating the quenched free energy. To compute $\ln\langle Z\rangle$, we resort to the first nontrivial order of the high-temperature expansion, which for the Hamiltonian (2.5) yields

$$F_{\rm ann\ av}(T)\simeq\mathcal{Q}\left[\overline{B}-\frac{\delta B^2}{4T}\right]-T\mathcal{Q}s, \tag{4.9}$$

where $\mathcal{Q}\sim N$ is the total number of contacts in the globule, and $e^{\mathcal{Q}s}$ is the total number of (maximally compact) conformations. Here $s$ is the same as in Eq. (3.34) above. Given that above $T_{\rm freeze}$, $F(T)=F_{\rm ann\ av}$, we examine the

condition that the entropy vanishes,

$$S(T_{\rm freeze}) = -\left.\frac{\partial F}{\partial T}\right|_{T=T_{\rm freeze}} = 0, \tag{4.10}$$

which signals the freezing transition, and recover precisely Eq. (4.3). Furthermore, we use the probability distribution (2.14) and consider the design average native-state energy

$$\langle E_\star({\rm seq})\rangle = \sum_{\rm seq} P^\star_{\rm seq} E_\star({\rm seq}) = \frac{\Sigma_{\rm seq}\Pi_{I=1}^N p_{s_I} \exp\left[-\dfrac{\mathcal{H}^{\rm des}({\rm seq},\star)}{T^{\rm des}}\right]\cdot \mathcal{H}({\rm seq},\star)}{\Sigma_{\rm seq}\Pi_{I=1}^N p_{s_I} \exp\left[-\dfrac{\mathcal{H}^{\rm des}({\rm seq},\star)}{T^{\rm des}}\right]}, \tag{4.11}$$

with Hamiltonians $\mathcal{H}^{\rm des}$ and $\mathcal{H}$ defined according to Eq. (2.5) with matrices $\hat{B}^{\rm des}$ and $\hat{B}$, respectively. Simple algebra allows us to rewrite this in the form

$$\frac{\langle E_\star({\rm seq})\rangle}{T} = -\frac{1}{W^\star}\left.\frac{\partial W^\star}{\partial n}\right|_{n\to 0} = -\left.\frac{\partial \ln W^\star}{\partial n}\right|_{n\to 0}, \tag{4.12}$$

where $W^\star$ is the annealed average partition function

$$W^\star = \left\langle \exp\left[-\frac{\mathcal{H}_{\rm eff}({\rm seq},\star)}{T}\right]\right\rangle = \sum_{\rm seq} P^{(0)}_{\rm seq} \exp[-\mathcal{H}_{\rm eff}({\rm seq},\star)/T], \tag{4.13}$$

and the effective Hamiltonian $\mathcal{H}_{\rm eff}$ is given by

$$\frac{\mathcal{H}_{\rm eff}({\rm seq},\star)}{T} = \frac{\mathcal{H}^{\rm des}({\rm seq},\star)}{T^{\rm des}} + n\frac{\mathcal{H}({\rm seq},\star)}{T}. \tag{4.14}$$

This represents our usual Hamiltonian (2.5), except with a new interaction matrix,

$$\frac{\hat{\mathcal{B}}_{\rm eff}}{T} = \frac{\hat{B}^{\rm des}}{T^{\rm des}} + n\frac{\hat{B}}{T}. \tag{4.15}$$

To lowest order in $1/T$ and $1/T^{\rm des}$, the annealed average free energy is independent of $\star$ (because all compact conformations have the same number of monomer contacts $\mathcal{Q}$), and it is given by

$$-\frac{\ln W^\star}{T} = \mathcal{Q}\left[\overline{\mathcal{B}_{\rm eff}} - \frac{\delta\mathcal{B}^2_{\rm eff}}{4T}\right], \tag{4.16}$$

in complete analogy with Eq. (4.9). We compute $\delta\mathcal{B}^2_{\rm eff}$ using the scalar product rule (3.42) and thus obtain

$$\langle E_\star\rangle = \mathcal{Q}\left[\overline{B} - 2\frac{\delta B^{\rm des}\,\delta B g}{2T^{\rm des}}\right]. \tag{4.17}$$

Finally, comparing this with the free energy of the random phase, which above $T_{\rm freeze}$ is given by Eq. (4.9), we recover the result (4.7) for the folding phase transition.

The above derivation is interesting in two respects. First, the $n\to 0$ limit in Eq. (4.12) appears independently of replicas. As all of our operations mathematically parallel those one would employ with the replica trick, this derivation sheds light on the replica trick itself. Indeed, here $n$ has the physical meaning of a source field that tests the response of the designed sequence to folding interactions.

Second, the nature of our theory is connected with the truncated high-temperature expansion for an annealed system. Note that the single most important approximation of the replica analysis is the truncation to $\phi^2$ terms in Eq. (3.18), which is equivalent to a high-temperature expansion. Physically, the high-temperature expansion is surprisingly applicable in the heteropolymer freezing problem. At high temperatures, the system is dominated by a single energy scale of the interaction matrix: the variance $\delta B^2$. For matrices with uncorrelated elements, i.e., Gaussian random interactions as in the independent-interaction model, $\delta B^2$ is the only energy scale and thus there are no further terms of the high-temperature expansion. Correlations exist in more realistic matrices, and thus the consideration of higher-order terms in the high-temperature expansion is necessary to capture these details. Such effects describe new and interesting physics, including microphase segregation in the globule (Chan and Dill, 1991; Sfatos *et al.*, 1993) at low temperature. We shall not discuss these effects here.

### B. How many designed sequences are there?

#### 1. Sequence-space entropy

We mentioned in the Introduction that the evolution paradox results from the need to select relatively few "good" sequences out of a very large sequence space. Moreover, it is qualitatively clear that better optimized sequences are more difficult to select. However, we can make this statement quantitative and ask, What is the fraction of all sequences with the given degree of optimization, or with the given native-state energy $E_\star$? It is easier to work with canonical design and ask how many sequences there are in the ensemble that result from the design at a given $T^{\rm des}$. Indeed, as long as we bypass the kinetics of sequence selection and consider the canonical design model with an equilibrium distribution in the sequence space (2.14), then the question of the number of sequences is immediately answered in terms of the corresponding entropy in the sequence space:

$$S_\star(T^{\rm des}) = -\frac{\partial(-T^{\rm des}\ln Z_\star)}{\partial T^{\rm des}}, \tag{4.18}$$

where

$$Z_\star = \sum_{\text{seq}} \exp[\mathcal{H}(\text{seq},\star)/T^{\text{des}}]. \tag{4.19}$$

(We assume for simplicity here that the design is performed with the Hamiltonian $\mathcal{H}=\mathcal{H}^{\text{des}}$.)

To begin with, it is useful to keep in mind the expression for the total number of sequences, which follows, for instance, from Eq. (4.18) in the $T^{\text{des}}=\infty$ limit:

$$\mathcal{N} = \sum_{\text{seq}} 1 = q_{\text{eff}}^N;$$

$$\ln q_{\text{eff}} = -\sum_{i=1}^{q} p_i \ln p_1 \leqslant \ln q, \tag{4.20}$$

where the effective number of monomer species $q_{\text{eff}}$ reaches its maximal value of $q$ if and only if all species are equally abundant ($p_i = 1/q$).

In principle, the entropy $S_\star(T^{\text{des}})$ depends on both $T^{\text{des}}$ and target conformation $\star$, as indicated by a subscript. However, in the volume approximation adopted here this $\star$ dependence is negligible. Indeed, if we resort to a high $T^{\text{des}}$ expansion, then to the order $\mathcal{O}(1/T^{\text{des}^2})$ we get

$$S_\star = \ln \mathcal{N} - \langle \delta\mathcal{H}^2 \rangle / 2T^{\text{des}^2}; \tag{4.21}$$

$\delta\mathcal{H}$ is the part of the energy due to $\widehat{\delta B}$, and $\langle\cdots\rangle$ means an average over sequences. Similar to what we have in Eq. (4.9), we find that $\langle \delta\mathcal{H}^2 \rangle = \mathcal{Q}\delta B^2/2$ is independent of $\star$, because it depends only on the number of bonds $\mathcal{Q}$, which is the same for all maximally compact conformations including $\star$. It is convenient to express $\delta B$ in terms of $T_{\text{freeze}}$ using Eq. (4.3); we then get in terms of the number of sequences designed at the temperature $T^{\text{des}}$ for the conformation $\star$

$$\mathcal{N}_\star(T^{\text{des}}) = \mathcal{N} \exp[-\mathcal{Q}s(T_{\text{freeze}}/T^{\text{des}})^2]. \tag{4.22}$$

First let us look at what our answer yields for $T^{\text{des}} = T_{\text{freeze}}$. We know that the ensemble of sequences prepared at $T^{\text{des}} = T_{\text{freeze}}$ includes all—let us call it $\mathcal{N}_\star^{(0)}$—sequences for which $\star$ is the ground state. As each sequence has its ground state on one of the compact conformations, and as we neglect the dependence on the particular compact conformation $\star$, it is not surprising that what we get for $\mathcal{N}_\star^{(0)}$ is the number of all sequences $\mathcal{N}$ divided by the number of compact conformations, $\exp(\mathcal{Q}s)$.

We can now rewrite Eq. (4.22) in terms of $\mathcal{N}_\star^{(0)}$:

$$\mathcal{N}_\star(T^{\text{des}}) = \mathcal{N}_\star^{(0)} \exp[\mathcal{Q}s(1-(T_{\text{freeze}}/T^{\text{des}})^2)]. \tag{4.23}$$

It is plausible, although we do not have any proof, that the random-energy model and its consequence, Eq. (4.23), remain valid to a reasonable approximation even when the volume approximation fails and $\star$ dependence becomes significant due to the sublinear in $N$ (or $\mathcal{Q}$) terms in entropy. If that is the case, then $\mathcal{N}_\star^{(0)}$ should be interpreted as the *designability* of the conformation $\star$ with respect to the given set of interactions $\hat{B}$, and Eq. (4.23) describes how the number of designed sequences decreases with respect to pure designability when one wants to select really well-optimized sequences.

#### 2. What is the best possible design?

Another interesting aspect of formula (4.23) is that it predicts a limit to improvements: since there should be at least one sequence, $\mathcal{N}_\star(T^{\text{des}}) \geqslant 1$, there is a minimal design temperature. Noting that $\mathcal{Q} \sim N$ and, introducing the conformational entropy per monomer $\omega$ such that $\mathcal{Q}s = N\omega$, we get

$$T^{\text{des}}/T_{\text{freeze}} \geqslant \omega/\ln q_{\text{eff}}. \tag{4.24}$$

This shows the value of $T^{\text{des}}$ at which the design procedure yields the best possible sequence, with the fewest frustrations.

Thus when laying the groundwork for optimal design, it is good to have more different species (larger $q_{\text{eff}}$) and less conformational freedom (smaller $\omega$, or $s$). More species allows better selectivity, while less freedom makes it easier to freeze. The latter shows, in particular, the possible role of secondary structure in proteins. Secondary structure greatly reduces conformational freedom and thus presumably creates room for better sequence design; this is precisely what is exploited in the experimental *de novo* design schemes (Hecht *et al.*, 1990; Dahiyat and Mayo, 1997). On the other hand, Eq. (4.24) also implies that design may be difficult if only few monomer species are in play. For instance, black-and-white models should not be considered representative for design, and indeed computational tests of the hydrophobic-polar model are not always successful (Yue *et al.*, 1995). Clearly, if $q_{\text{eff}} < e^\omega$, then design is entirely impossible. For instance, commonly employed computational models on a cubic lattice, such as 18-, 27-, 36-, and 48-mers, have $\omega$ values of, respectively, 0.42, 0.43, 0.51, and 0.53 (Pande, Joerg, *et al.*, 1994); accordingly, $e^\omega \approx 1.7$ barely leaves room for design for a hydrophobic-polar model with two species even if $q_{\text{eff}} = q$.

### C. Other random-energy-type systems: Screened polyampholytes

Up to now, we have considered only short-range interactions. Although it is commonly held that hydrophobic interactions provide the dominant force (Chan and Dill, 1993), some amino acids are charged at physiological pH and thus Coulomb interactions are potentially relevant (Creighton, 1992). In this section, we address a simple random-energy-model estimate of freezing assuming that it is driven purely by long-range interactions. That is, we consider a neutral polyampholyte chain (a chain with an equal number of positively and negatively charged monomers) and we assume that the chain is in a dense globular state. Actually, the globular state is very natural for polyampholytes with no net charge (Dobrynin and Rubinshtein, 1995; Kantor and Kardar, 1995a), but we shall not consider the density dependence on the conditions and instead assume that

maximal compactness is supported, say, by an external box of size $R$. Even though there are no mobile counterions in the globule, we also assume that conformational flexibility of the chain is sufficient that Coulomb interactions are well screened (Wittmer, 1993).

For this case, the freezing transition is expected to be in conformity with the random-energy model; the freezing temperature can then be estimated via the condition $\mathcal{M}n(T)\sim 1$, where $\mathcal{M}=e^{\omega N}$ ($\omega N=s\mathcal{Q}$) is the number of (compact) conformations and $n(T)$ is the density of states for the plasma of disconnected monomers at an energy in equilibrium at the temperature $T$; this latter value[17] is given by

$$\frac{n}{n_0}\simeq\begin{cases} e^{-c(E/E_0)^2}=e^{-c(E_0/T)^2} & T>E_0 \\ e^{-c(E/E_0)^3}=e^{-c(E_0/T)^3} & T<E_0, \end{cases} \tag{4.25}$$

where $E_0=Ne^2/R$ and $c$ is a constant of order unity.

Assuming $R=aN^{1/3}$, one obtains (Pande, Grosberg, Kardar, *et al.*, 1996)

$$T_{\text{freeze}}\sim e^2/a\omega^{2/3}. \tag{4.26}$$

Most importantly, this is independent of $N$, suggesting that the freezing transition is a true phase transition. We shall see, however, that the long range of interactions leads to a departure from the random-energy model (see Sec. V.E).

### D. Verifications of heteropolymer freezing and design

#### 1. Experiments

We must admit that little is known experimentally about the freezing transition of designed heteropolymers. In terms of experiments on proteins, Sauer (1996) examined a library of random amino acid sequences. A considerable fraction exhibited proteinlike folding properties. Much remains to be done in both theory and experiment to make a quantitative comparison of these data with theory.

Also of interest are experiments on the freezing and design of synthetic polymers. Normally, in a solution of heteropolymers, there is a very broad distribution of sequences. It is overwhelmingly likely that each sequence will be found once, in any experimentally realizable volume of the solution. In Nature, the protein synthesis machinery of the living cell circumvents this problem. However, one can turn this problem upside down and examine a polymeric gel, which is in a sense a single, macroscopic polymer. While our theory is not directly applicable to gels (see Panyukov and Rabin, 1996), and in particular we do not know if the random-energy model is applicable to gels, experiments on heteropolymer gels are being actively conducted by one of us (Tanaka *et al.*, 1998).

To illustrate the idea behind those experiments, let us imagine the synthesis of a polyampholyte gel with balanced amounts of positively and negatively charged monomers. Let us first assume that there is no added salt in the pre-gel solution. Then the positively and negatively charged ''prospective'' monomers (i.e., those molecules that will be polymerized) can cause screening. Assuming that the conversion of the gelation reaction is 100%, we obtain a gel in which the screened pattern of charges of the pre-gel solution is imprinted as a three-dimensional analog of the heteropolymer sequence. This sequence is highly nonrandom: on the scale just above the pre-gel Debye length, predominantly positive and predominantly negative regions should be found to be alternating in space, reminiscent of an ionic crystal, such as NaCl. This gel should readily collapse upon submersion into a poor solvent. Consider now another gel, prepared with an abundance of salt present in the pre-gel solution. In this case, monomers are screened from each other by the added salt and thus are likely to form a completely random pattern of charges imprinted on the gel, with either positive or negative predominance of order $\sqrt{V}$ for any volume $V$ at every scale. Obviously the salt cannot be washed away from this gel: positive and negative counterions remain captured in the respective regions of the gel. Although there are many complications in these gel experiments, which will be reported elsewhere, this brief discussion illustrates that design can be implemented in a physical experiment that does not utilize proteins or any form of biotechnology.

#### 2. Computer simulations

To examine the validity of our model, we now compare our results to computer simulations of lattice heteropolymers. First, we examine the behavior of the $x(T)$ order parameter [see Eq. (4.5)] for both random and designed sequences for maximally compact 27-mers (Fig. 3). We then demonstrate that the overlap parameter ($Q$) has a very well-defined bimodal distribution (Fig. 4), thus confirming the first-order nature of freezing. Furthermore, we examine the density of compact (globular) states for a well-designed 36-mer sequence (Fig. 5). Pande, Joerg, *et al.* (1994) calculated the energy of all 84 731 192 compact conformations of the 36-mer to

[17]Although Eq. (4.25) represents a fundamental property of a regular plasma, we were surprised not to find it in standard textbooks. It is derived in the following way. In general, the density of states (DOS) $n(E)$ is defined via the relation $Z(T)$-$\int n(E)e^{-E/T}dE$ with the partition function $Z(T)$. Under normal thermodynamic conditions, this integral is of the saddle-point type, yielding $n(E)=e^{S(E)}=\exp[-\partial F/\partial T]$.. For a plasma, there is the Debye length scale $R_D\sim(TR^3/e^2N)^{1/2}=R\sqrt{T/E_0}$, which necessitates two regimes: (i) for $T>E_0$, the size of the system is smaller than Debye length ($R<R_D$, which is of no interest for regular plasma applications, but may be the case for an globule), the system is not screened, and one can compute free energy and entropy using a high-temperature expansion, yielding the Gaussian DOS; (ii) for $T<E_0$ the system is screened ($R>R_D$), and free energy is about $T$ per Debye volume $F\sim-(R^3/R_D^3)T$, which upon computing entropy yields Eq. (4.25). It is not surprising that screening strongly suppresses the DOS as particles must be finely arranged to screen each other, and there are relatively few such states.

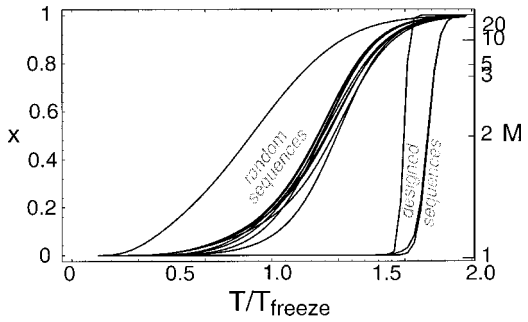



FIG. 3. Order parameter $x$ as a function of temperature $T$ calculated from the enumeration of all maximally compact 27-mer conformations. We see that random sequences have a less cooperative transition at a lower temperature (at $T \approx T_{\text{freeze}}$) than do designed sequences (which have a transition at $T \approx 1.5 T_{\text{freeze}}$). However, in all the cases shown here (27-mer sequences, independent-interaction-model interactions), the ground state is unique, so there is a freezing transition.

create this histogram. We note that the gap between the ground state and the peak of the distribution is rather pronounced.

We confirmed the phase diagram calculated analytically (Fig. 6) using computer simulations of compact 27-mers. The computer simulations generated chains using Monte Carlo annealing at a given design temperature $T^{\text{des}}$. Next, the partition function for maximally compact conformations was exactly calculated, i.e., the energy of all 103 346 maximally compact states was calculated. The folding temperature was determined by the temperature at which the order parameter [Eq. (4.5)] was $x(T) = 0.9$. The results are shown in Fig. 6. This demonstration employed the 7-Potts model (Pande, Grosberg, and Tanaka, 1994b) as it had been shown to be "most heteropolymeric" for 27-mers. Other models show similar behavior. Due to finite-system effects, we see that the folding temperature becomes constant for $T^{\text{des}} < 0.5$, as we have reached the maximum degree of optimization for 27-mers. Also, near $T^{\text{des}}/T_{\text{freeze}} \sim 1$, there are large

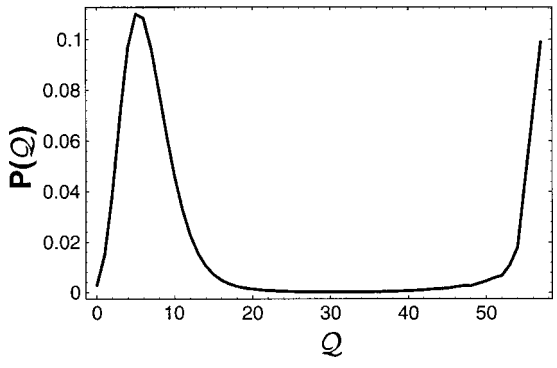



FIG. 4. Probability $P$ as a function of $Q$, the number of contacts, for a well-designed 48-mer sequence near the coexistence temperature. We see two distinct peaks at $Q = 8$ and $Q = 57$ contacts (48-mers have at most 57 contacts). This bimodal distribution clearly demonstrates that freezing is a first-order transition.

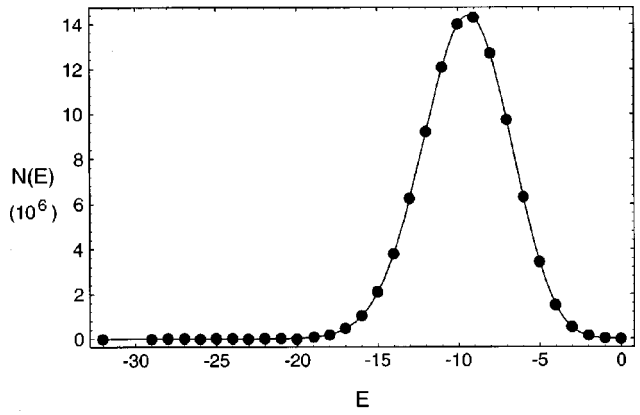



FIG. 5. Density of compact states for a 36-mer calculated by full enumeration. We see a Gaussian distribution centered at $E = -9$ and a ground state at $E = -32$.

fluctuations due to small size effects, which lead to small quantitative deviations from our theory.

To join the computer simulations with the analytic predictions, we measured the freezing temperature $T_{\text{freeze}}$ for random sequences directly from the simulation data at high $T^{\text{des}}$. With the measured value of $T_{\text{freeze}}$ and the calculated value of the variance of the interaction matrix, we plotted the theoretical curves using only the lowest-order term $\mathcal{O}(1/T^2)$. In our comparison between theory and simulation, there are no free parameters. The agreement is quite striking. However, we should also stress that the freezing calculation described here is that of globules (as described by the large-globule theory presented in the previous section). For phase transitions from a coil to the native state, the agreement with the large-globule theory is of course not as good. This is fundamentally important for proteins, as we are generally interested in the coil-to-native-state transition, not globule-globule transitions. A major part of the disagreement, however, involves the failure of the random-energy model for this system, which will be discussed in the next section.

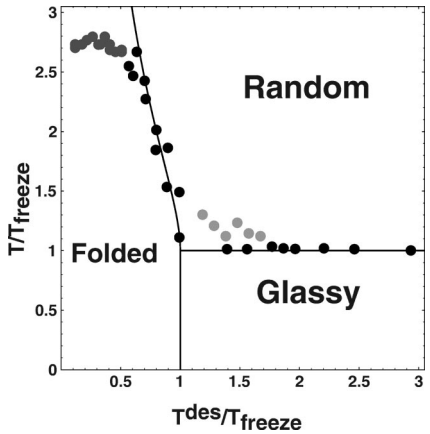



FIG. 6. Quantitative agreement of the phase diagram for the analytic model and simulation data for designed 7-Potts-model heteropolymers: compact 27-mers on the $3 \times 3 \times 3$ cubic lattice and analytic prediction with no free parameters.

## V. A DEEPER LOOK INTO APPLICABILITY OF THE RANDOM-ENERGY MODEL

### A. Important deviations from the random-energy model

As should be clear from the previous discussion, the applicability of the random-energy model is not just an academic problem of the validity of a particular approximation. We have demonstrated that this model is an extremely useful paradigm yielding insightful results regarding equilibrium thermodynamic properties of the heteropolymer freezing transition. Moreover, it offers a simple method of sequence design, i.e., manipulating one target energy level without affecting all other levels.

On the other hand, to fully understand protein folding, we need to go further and to examine why and how proteins deviate from the model. Indeed, as we have already mentioned, a pure random-energy model would leave the Levinthal paradox unresolvable, and, therefore, folding kinetically impossible on any reasonable time scale: an uncorrelated energy landscape cannot (by definition) provide any bias toward the native state, and an unbiased random walk search for a single point (the native state) in multidimensional conformational space will last practically forever.

While the random-energy model is rarely mentioned in experimental studies on proteins, the assumption of statistical independence of states is implicit in the motivation of several experiments; for example, *de novo* protein design (Regan and DeGrado, 1988; Hecht *et al.*, 1990; Dahiyat and Mayo, 1996; DeGrado, 1997; Lazar *et al.*, 1997) relies on the assumption that the selection of sequences which lower the energy of a desired conformation will not also lower the energies of other competing conformations. An opposite intuition is also prevalent in many works, such as computational generation for a given sequence of a low, but not the lowest, energy conformation (Holden, 1995), if the random-energy model were valid, then a low but not lowest energy conformation would tell us nothing about the ground state.

To understand the random-energy model better, and also to prepare for the discussion of other effects in Sec. VI, let us consider the applicability of the model following Pande, Grosberg, Joerg, and Tanaka (1996).

### B. Energy correlator as a measure of the validity of the random-energy model

Local rearrangement of conformations is the simplest means of going beyond the random-energy model, as illustrated in Fig. 7. To make this clear, imagine two conformations, of which one conformation, $\alpha$, represents a small local rearrangement of the other conformation, $\beta$ (see Fig. 7). As the energies $E_\alpha$ and $E_\beta$ are given as sums over all pairs of monomers that are in contact (we are speaking now about short-range interactions), they are dominated by identical contributions and differ only due to the small region of difference between $\alpha$ and $\beta$. Clearly these two energies are strongly dependent, which violates the defining rule of the random-energy model, i.e., the statistical independence of energy levels (2.2). Moreover, this consideration has an immediate bearing on our understanding of design: as a design, in a sense, pulls down the energy of the desired target conformation relative to typical conformations, it should also affect similar conformations, albeit to a smaller degree.

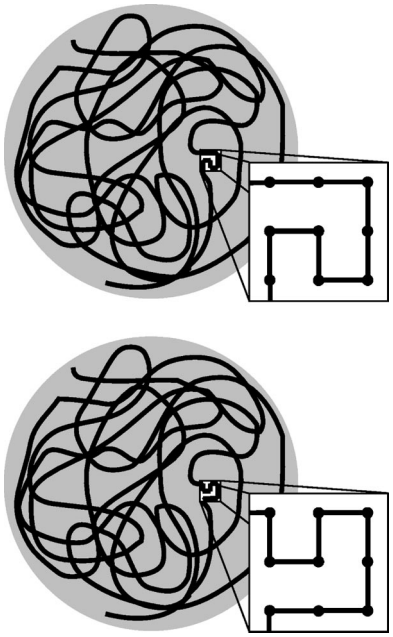


FIG. 7. Local conformation fluctuation. Two different though closely related conformations are shown; everywhere within the shaded area they are identical, while the local difference is seen in the window.

To make the above argument quantitative, one can directly take the Hamiltonian (2.5) and calculate the degree of statistical dependence of conformations $\alpha$ and $\beta$ by the connected energy-energy correlation function,

$$\langle E_\alpha E_\beta\rangle - \langle E_\alpha\rangle\langle E_\beta\rangle = \frac{1}{2}\delta B^2\left[\mathcal{Q}_{\alpha\beta} + \kappa\mathcal{K}_{\alpha\beta} - \frac{c}{N}\right], \tag{5.1}$$

where $\kappa = \Sigma_{i,j,k\neq i} p_i b_{ij} p_j b_{jk} p_k$, $\delta B^2$ and $b_{ij}$ are defined via Eq. (3.45), and the last term is due to the constraint of a fixed overall composition ($c$ is a constant of order unity). When interactions are short range, $\mathcal{Q}_{\alpha\beta}$ is our standard overlap parameter between conformations $\alpha$ and $\beta$ (the number of contacts these conformations have in common; see Sec. II.B.2.b), and

$$\mathcal{K}_{\alpha\beta} \equiv \sum_{I\neq J\neq K} \Delta(\mathbf{r}_I^\alpha - \mathbf{r}_J^\alpha)\Delta(\mathbf{r}_J^\beta - \mathbf{r}_K^\beta) \tag{5.2}$$

describes the covariance of the coordination number of monomers in conformations $\alpha$ and $\beta$.

It is natural that the overall scale of the energy correlator be set by the standard deviation of the interaction matrix $\delta B$. The appearance of the $\mathcal{Q}_{\alpha\beta}$ term in the correlator is also natural and corresponds precisely to our qualitative argument regarding locally rearranged conformations: obviously the conformations shown in Fig. 7 have very high overlap $\mathcal{Q}$. For the special interac-

tions for which the coefficient $\kappa$ vanishes, $\mathcal{Q}_{\alpha\beta}=0$ means statistical independence between states $\alpha$ and $\beta$. Physically this reflects the fact that, as these states have no contacts in common, their energies will have no terms in common and should be statistically independent; analogously, for the case $\mathcal{Q}_{\alpha\beta}=\mathcal{Q}_{\max}=\mathcal{Q}$, the two conformations are identical and thus are trivially statistically dependent.

The appearance of $\mathcal{K}_{\alpha\beta}$ is more unexpected. It describes the correlation in coordination number of monomers in conformations $\alpha$ and $\beta$. The coefficient of $\mathcal{K}_{\alpha\beta}$, $\kappa=\Sigma_i p_i[\Sigma_j p_j b_{ij}]^2\geqslant 0$, represents an intricate property of the interaction matrix. $\kappa$ can be interpreted as a measure of the average "virtual interaction" between monomer species $i$ and $k$ mediated through $j$; it can also be interpreted as a correlator of interaction matrix elements. Thus there are qualitative differences between systems with vanishing and nonvanishing $\kappa$.

Energy correlations can be caused by conformational overlaps, and this is the only source of correlations for a system whose interactions correspond to $\kappa=0$. In this case, the validity of the random-energy model may rest on the statistical rarity of overlapping conformations. For other systems, with $\kappa\neq 0$, there is a residual statistical dependence between states—even nonoverlapping ones, with $\mathcal{Q}_{\alpha\beta}\approx 0$—and the validity of the model is even more problematic. Although for a finite system with a fixed composition there is the possibility of compensation due to the last term in Eq. (5.1), a random-energy model for this case would remain questionable because of higher-order correlations. In the next section, we discuss the two possibilities $\kappa=0$ and $\kappa\neq 0$ in more detail and give some computational examples. Then, in Sec. V. D, we employ sequence design as a tool for systematically testing violations of the random-energy model.

### C. Sources of energy correlations

#### 1. Conformations

The first term of the energy correlator equation (5.1) is proportional to the number of contacts that conformations $\alpha$ and $\beta$ have in common, $\mathcal{Q}_{\alpha\beta}$; clearly, identical contacts give identical contributions to energies of both conformations, thus yielding a dependence. However, if large $\mathcal{Q}_{\alpha\beta}$ are rare, then the random-energy model remains a good approximation. Let us examine this for simple computational models typically employed.

Specifically, we perform the following calculation. We first choose one conformation $\alpha$ (a "center" in conformational space) and count (by enumeration) how many other conformations $\beta$ are within a given overlap "distance" $\mathcal{Q}$ from $\alpha$; we then average this over $\alpha$ (over different centers): $P(\mathcal{Q})=(1/\mathcal{M})\Sigma_{\alpha\beta}\delta(\mathcal{Q}_{\alpha\beta}-\mathcal{Q})$, where $\mathcal{M}$ is the number of conformations in the given conformational space. Figure 8 presents the results of such a study for a variety of cubic lattice conformational spaces; the data are presented in logarithmic scale, i.e., in terms of entropy $\Delta S(\mathcal{Q})=\ln P(\mathcal{Q})$.

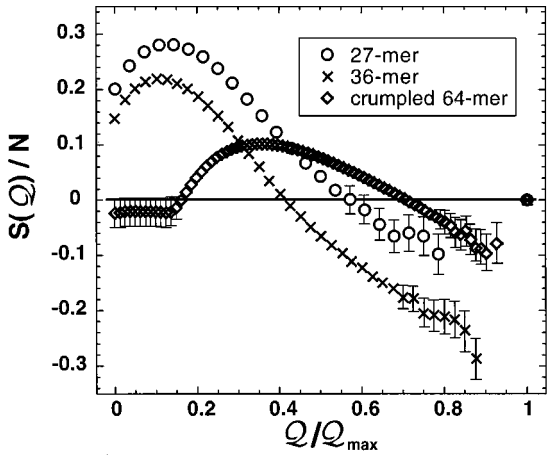



FIG. 8. Entropy $S$ as a function of $\mathcal{Q}$, the number of contacts, where $S(\mathcal{Q})\equiv\ln[P(\mathcal{Q})]$, for compact 27-mers, compact 36-mers, and compact and crumpled 64-mers. Although all distributions are peaked at small $\mathcal{Q}$, the region with virtually no overlapping conformations at large $\mathcal{Q}$ is narrower and less pronounced for crumpled conformations. On the other hand, it gets broader and deeper as chain length gets longer, as is indicated by the comparison of compact 27- and 36-mers.

One type of conformational space, which is typically used in freezing simulations, includes all maximally compact conformations. Pande, Grosberg, Joerg, and Tanaka (1996), examined this space for 27-mers and 36-mers, filling, respectively, $3\times3\times3$ and $3\times3\times4$ regions of the cubic lattice. The first and most notable feature seen in Fig. 8 is that $P(\mathcal{Q})$ is peaked at low overlap and has a deep minimum, where $P(\mathcal{Q})<1$ ($\Delta S<0$) at large overlaps. It should be understood that $P(\mathcal{Q})<1$ means that, for most of the central conformations, there are no other conformations with the given value of overlap $\mathcal{Q}$; only a few centers have one or more conformations with which they have $\mathcal{Q}$ contacts in common. This is precisely the situation that favors the random-energy model as small rearrangements are not possible and $P(\mathcal{Q})$ has virtually no contributions in the large-$\mathcal{Q}$ range. We also note that these features are more pronounced in 36-mers than in 27-mers. It is natural to conjecture that there exists an overall tendency toward sharpening of the small-$\mathcal{Q}$ peak with increasing chain length $N$. This agrees well with analytic arguments given above. Another simulation study (Pande, Grosberg, Kardar, *et al.*, 1996), performed for maximally compact conformations visiting (once and only once) all the cubic lattice sites within the sphere of a radius $R\leqslant 10$ (that is, $N$ up to about 4000), indicates that the most probable value of $\mathcal{Q}$ scales as

$$\mathcal{Q}_{\text{most prob.}}\sim N^{\gamma},\quad \gamma\approx 0.75. \tag{5.3}$$

As $\mathcal{Q}_{\max}$ is proportional to $N$, the ratio $\mathcal{Q}/\mathcal{Q}_{\max}$ vanishes in the thermodynamic limit (see also Franz *et al.*, 1999).

Following this logic, one would expect a sharper peak at low $\mathcal{Q}$ and a deeper minimum at strong overlaps for 64-mers than we found for 36-mers. Figure 8 shows that this is *not* the case for another type of conformational space, the *crumpled conformation*, exemplified by the

data for crumpled 64-mers (confined within $4\times4\times4$). Here we follow the terminology of Grosberg *et al.* (1988) and Pande, Grosberg, Joerg, and Tanaka (1996) and call a conformation crumpled if it fills each subvolume completely before entering the next one. Technically, the $4\times4\times4$ cube can be broken down into eight $2\times2\times2$ subcubes. Crumpled conformations fill every site in a given subcube before entering a new one. Thus, in crumpled conformations, neighbors along the chain are likely to be neighbors in space. In this sense, crumpled conformations perhaps crudely model the effect of secondary structure that localizes contacts. By construction, crumpled conformations can be relatively easily rearranged locally, and this is why there are much closer relations among them. The validity of the random-energy model is very problematic for crumpled conformations.

The most spectacular departure from the model for crumpled 64-mers is seen in the distribution of overlaps; this is similar to Fig. 8, except not for all states—only for those that are ground states for a given sequence and interactions. Multiple degeneracy of states is typical for lattice models with discrete values of interaction energies, giving rise to the discrete random-energy model, or DREM (see Gutin and Shakhnovich, 1993 and Sec. II.A.3 above). For either the basic model or the discrete random-energy model to be valid, these ground states should all be unrelated, and this is indeed the case for regular maximally compact 27- or 36-mers with Potts interactions, whose ground states yield a $P(\mathcal{Q})$ indistinguishable from that of conformations taken at random (Shakhnovich and Gutin, 1990; Gutin and Shakhnovich, 1993). The situation is dramatically different for crumpled 64-mers: upon enumerating the energies of all conformations for 1000 sequences with Ising interactions and comparing the ground states, Pande, Joerg, *et al.* (1994) have shown that the model fails for this conformational space (see also Pande, Grosberg, Joerg, and Tanaka, 1996). This does not merely demonstrate that the random-energy model breaks down for crumpled 64-mers in particular, but also reveals that we cannot assume its validity *a priori* for all conformation spaces in general, even in three dimensions.

### 2. Interactions

Apart from the variance of the elements of the interaction matrix, which determines the overall scale of the energy correlator, Eq. (5.1), interactions also play a more dramatic role in the second term in Eq. (5.1), as $\kappa$ vanishes for only certain models. Note that $\mathcal{K}_{\alpha\beta}$ in Eq. (5.2), for maximally compact conformations on the lattice, is only very weakly conformation dependent; it depends exclusively on the position of the chain ends in conformations $\alpha$ and $\beta$. The appearance of a conformation-independent term signals a departure from the random-energy model, as even states with vanishing $\mathcal{Q}$ are statistically dependent. What forms of interactions have this residual statistical dependence?

The simplest case to examine is a system with just two monomer species, with even composition ($p_1=p_2=1/2$). In this case, for the matrix (3.46), it is easy to compute $\kappa=(1/4)\cos^2\theta$. Thus $\kappa$ vanishes for an Ising model ($\theta=90°$), but it is essential for, say, hydrophobic-polar interactions ($\theta\approx35°$).

From the point of view of energy correlations [Eq. (5.1)], the Miyazawa and Jernigan $20\times20$ matrix (Miyazawa and Jernigan, 1985), a set of amino acid interaction potentials derived from protein statistics, seems like the hydrophobic-polar model plus some noise; indeed, $\kappa\neq0$ for the Miyazawa-Jernigan matrix, as well. Another example is the case of one monomer species that interacts particularly strongly with all others; in this case, there is a correlated contribution to the conformation energies even if they have no bond in common. Finally, even a perfectly ordered interaction matrix, such as the $q$-Potts model (in which values are either $-1$ along the diagonal or 1 off diagonal), can yield a nonzero $\kappa$ if the polymer composition is not even ($p_i\neq1/q$). We have to remind the reader, however, that although none of these models favor the random-energy model because of the residual energy correlation for even unrelated conformations, this effect can be due to the negative fixed-composition term in Eq. (5.1); thus, in some cases (e.g., certain values of $N$), the model may appear to be a good approximation due to a complicated cancellation of factors.

The opposite examples, where $\kappa=0$, include models with an even composition and a symmetric contribution from monomer species, such as the Potts model or the independent-interaction model, in which $B_{ij}$ are taken from a Gaussian distribution (Shakhnovich and Gutin, 1989a). For these models, the conformation-independent term vanishes and the only statistical dependence comes from conformational overlap.

## D. Sequence design as a tool for examining the applicability of the random-energy model

As previously discussed, the validity of the random-energy model has important implications for sequence design. If the model were rigorously valid, the design would affect a target conformation only, by pulling down its energy level and not affecting other levels. However, in reality, other closely related conformations should also be involved in the design. To examine this property, let us consider the following computer "experiment": we perform sequence design for a particular target conformation $\star$ and we see how (to which degree, in which direction, etc.) other states are affected. In this sense, sequence selection acts as a field in which we test the response of the energies of other conformations to the manipulation of the energy of $\star$.

In Fig. 9, we see from the density of states, $P(E,\mathcal{Q})\equiv\Sigma_\alpha\delta(\mathcal{Q}-\mathcal{Q}_{\alpha\star})\,\delta(E-\mathcal{H}_\alpha)$, for two differently designed sequences with Ising interactions, that sequence selection acts on all states $\alpha$ to a degree that is roughly linear in $\mathcal{Q}_{\alpha\star}$. Therefore the energy of a particular conformation $\alpha$ is largely determined by $\mathcal{Q}_{\alpha\star}$.

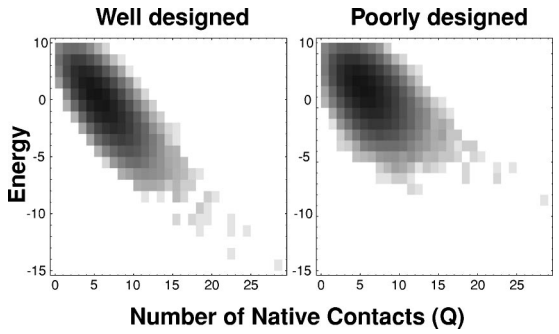



FIG. 9. Probability density $P(E,\mathcal{Q})$ for maximally compact 27-mer conformations ($\mathcal{Q}_{\max}=28$), Ising interactions ($B_{ij}=\delta_{ij}$) for a well-designed sequence ($T^{\rm des}=0.2T_{\rm freeze}$) and a poorly designed sequence ($T^{\rm des}=0.8T_{\rm freeze}$). (We remind the reader that a design with $T^{\rm des}\approx T_{\rm freeze}$ yields sequences that are practically random, except that $\star$ is among their low-energy states.) Darker regions correspond to more states. We see that design pulls down the energy of the ground state, but also pulls down the energy of related states to a degree proportional to the similarity with the native state; low-energy excitations from the native state are therefore very similar to the native state. However, for poorly designed sequences, this is not the case and low-energy excitations typically have little overlap with the native state. The zero-point energy is chosen to be the value of the energy of a typical sequence in a typical conformation, i.e., $\langle\overline{\mathcal{H}}\rangle$, and is therefore independent of $T^{\rm des}$.

Thus we see that sequence selection does not just pull down the energy of a desired conformation but rather affects the whole density of states. This is a dramatic deviation from the random-energy model. Note, on the other hand, that this is consistent with the assumption of the Go model (see Sec. II.A.1), as energy appears to be roughly linear in $\mathcal{Q}$; we shall see later (Sec. VI) that this is precisely what one should expect for a well-designed sequence.

However, this effect does not necessarily detrimentally affect folding to $\star$. For models with a degree of statistical antidependence (e.g., fixed, even hydrophobic-polar composition), selection may act to push *up* the energy levels of conformations other than $\star$. Moreover, as this statistical antidependence varies between sequences, it could be considered as a sequence selection criterion (in addition to optimization of the ground state) and therefore is a possible means to "design out" unwanted conformations (Yue *et al.*, 1995).

### E. Energy correlations for long-range interactions

While it is commonly held that short-range interactions (and, in particular, hydrophobic collapse) predominate, some amino acids are charged and thus Coulomb interactions are potentially relevant. In this section, we address how long-range interactions between monomers lead to a departure from the random-energy model. Therefore we do not include the effects of screening by counterions, which may render Coulomb interactions short range for proteins in physiological conditions.

Long-range Coulomb interactions are described by the Hamiltonian (2.5) with $f(r)=1/r^{(d-2)}$ (in $d$ dimensions) and $B_{ij}=Bs_is_j$, where $s_i\in\{\pm1\}$ is the charge of monomer $i$. When one computes the energy correlator (5.1) for this case, it turns out that $\kappa=0$. Moreover, the relevant value for judging statistical dependence is not the standard overlap $\mathcal{Q}_{\alpha\beta}$ (2.8), but

$$\mathcal{Q}^{\rm LR}_{\alpha\beta}=\sum_{I\neq J} f(\mathbf{r}^\alpha_I-\mathbf{r}^\alpha_J)f(\mathbf{r}^\beta_I-\mathbf{r}^\beta_J). \tag{5.4}$$

This quantity is different indeed from the short-range case $\mathcal{Q}_{\alpha\beta}$. The main physical difference between the short-range and long-range cases is the fact that polymeric bonds always keep monomers within long-range interaction distances: one can "hide" monomers in the short-range case to bring down $\mathcal{Q}^{\rm SR}$, by rearrangements of the conformation. In the long-range case, the presence of polymeric bonds creates a correlation that cannot be significantly changed by changes in conformation, as such changes keep monomers within the long-range interaction scale. Thus, for two conformations chosen at random, one would expect that $\mathcal{Q}^{\rm SR}$ would be negligible, but perhaps $\mathcal{Q}^{\rm LR}$ may not be.

A simple scaling argument demonstrates this. First, we consider the value of maximum overlap, i.e., overlap with the same conformation. For each of the $N$ monomers, there is a contribution from $\mathcal{O}(r^{d-1})$ monomers at a distance $r$ away. Thus $\mathcal{Q}_{\max}\sim\int dr\,Nr^{d-1}f^{-2}(r)$. For short-range interactions, this integral is dominated by contributions at a microscopic (a single monomer) length scale and we get $\mathcal{Q}^{\rm SR}_{\max}\sim N$. For long-range interactions, we find that while contributions from monomers far away are smaller, there are more of them. For Coulomb interactions $f(r)=1/r^{d-2}$, $\mathcal{Q}^{\rm LR}_{\max}$ is dominated by a maximal length scale. Thus, for a polymer packed in a volume of size $R$, we get $\mathcal{Q}^{\rm LR}_{\max}\sim NR^d/R^{2(d-2)}\sim NR^{4-d}$.

Next we consider the long-range overlap between two conformations chosen at random. In this case, $\mathcal{Q}^{\rm LR}_{\rm rand}$ is still infrared divergent and dominated by the largest length scale $R$: we have about $N^2$ of monomer pairs, each contributing about $R^{-2(d-2)}$, yielding $\mathcal{Q}^{\rm LR}_{\rm rand}\sim N^2R^{-2(d-2)}$.

Interestingly, for densely packed (maximally compact) polymers, when $R\sim N^{1/d}$, both $\mathcal{Q}^{\rm LR}_{\max}$ and $\mathcal{Q}^{\rm LR}_{\rm rand}$ appear to have the same scaling:

$$\mathcal{Q}^{\rm LR}_{\max}\sim\mathcal{Q}^{\rm LR}_{\rm rand}\sim N^{4/d}. \tag{5.5}$$

In particular, the ratio $\Gamma\equiv\mathcal{Q}^{\rm LR}_{\rm rand}/\mathcal{Q}^{\rm LR}_{\max}$ does not vanish, even in the thermodynamic limit. Moreover, this residual overlap does not appear to be conformation dependent (for typical pairs of conformations) and thus we expect that $P(\mathcal{Q}^{\rm LR})\sim\delta(\mathcal{Q}^{\rm LR}-\mathcal{Q}^{\rm LR}_{\rm rand})$.

This result is fundamentally different from the short-range case. The existence of a residual overlap in the long-range case means that the random-energy model is not valid, as there is always a residual statistical dependence. Note that $\Gamma$ does depend on the fractal dimensionality of the chains. While $\Gamma$ is a finite constant when conformations are globular, it vanishes in the thermody-

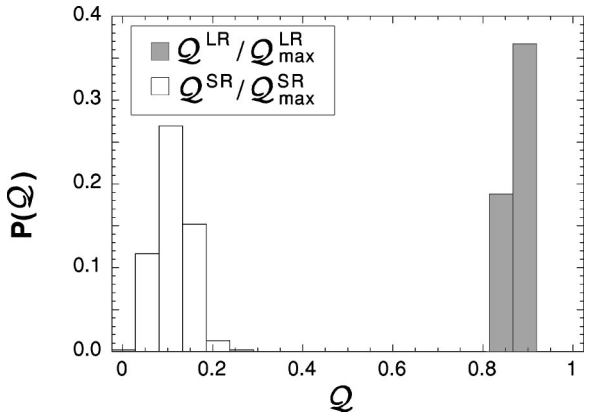



FIG. 10. Short-range (SR) and long-range (LR) probability distributions $P(\mathcal{Q}^{LR})$ and $P(\mathcal{Q}^{SR})$, obtained from 64-mers on a cubic lattice. Due to finite size effects, there is some residual overlap in the short-range case (here peaked at 0.1). However, we expect the short-range residual overlap to vanish in the thermodynamic limit, while the long-range overlap does not.

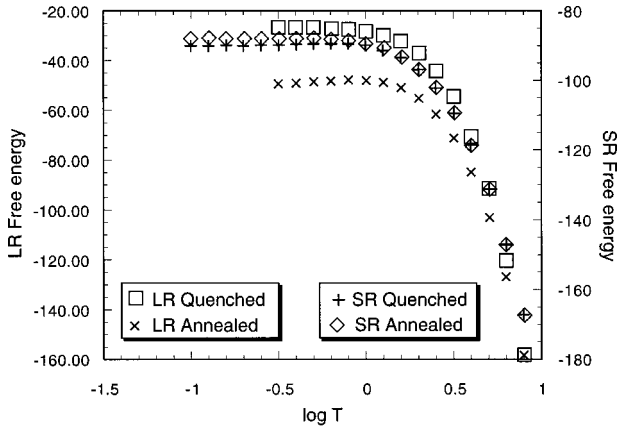



FIG. 11. Comparison of the average free energy calculated by the exact enumeration of all maximally compact 36-mer states for long-range (LR) Coulomb and short-range (SR) antiferromagnetic Ising model interactions. The average is performed over 80 random sequences and both the quenched and annealed free energies, $\langle \ln Z \rangle$ and $\ln\langle Z \rangle$, respectively, are shown. The quenched and annealed short-range cases agree remarkably well, reflecting the validity of the random-energy model, whereas the long-range cases agree only at high temperature.

namic limit when at least one of the conformations under comparison is less compact than the other, when the conformations differ in their scaling $R \sim N^{1/d_f}$, where $d_f$ is the fractal dimension of the chain. Thus, for an ensemble of conformations that consists of compact and noncompact chains, we must expect a more complicated distribution $P(\mathcal{Q}^{LR})$.

Numerical tests reported by Pande, Grosberg, Kardar, *et al.* (1996) fully support the above scaling arguments and the conclusion (5.5). According to these authors, the scaling exponents $\gamma$ given by $\mathcal{Q}^{LR} \sim N^{\gamma}$ appear to be the same within error for random pairs and maximum overlap for compact conformations. Furthermore, the fits agree well with the predictions $\gamma^{LR}_{max} = \gamma^{LR}_{rand} = 4/3$ [to be compared with the trivially expected $\gamma^{SR}_{max} = 1$ and the numerical result $\gamma^{SR}_{rand} \approx 0.75$, Eq. (5.3)]. Another test, reported in Fig. 10, is free of these problems. In this figure, we show both short-range and long-range overlap distributions averaged over 1000 pairs of maximally compact conformations of 64-mers. The difference is absolutely clear: while $\mathcal{Q}^{SR}$ is almost always small, $\mathcal{Q}^{LR}$ is not.

Yet another demonstration of nonrandom behavior for a heteropolymer with long-range interactions is shown in Fig. 11. This figure reports calculations of Coulomb energy for all 84 731 192 compact conformations of 36-mers on the $3 \times 3 \times 4$ lattice. In Fig. 11, we plot both the quenched $\langle \ln Z \rangle$ and annealed $\ln\langle Z \rangle$ averaged free energies vs temperature; the average was performed over 80 random sequences of an equal number of plus and minus charges. For comparison, the same computation is reported in the figure for the short-range (antiferromagnetic Ising) interactions. The quenched and annealed averages are the same within experimental error for short-range interactions above $T_{freeze}$, as it should be in the random-energy model according to Eq. (2.4). On the other hand, these quantities are very different for long-range interactions (except for the extreme of infinite temperature, where the energy, and therefore the nature of interactions, is irrelevant). Note that the sequences used in Fig. 11 are the same; the only difference is what quantity is averaged ($\langle Z \rangle$ or $\langle F \rangle$) and what type of interaction is used, long-range or short-range.

### F. Is the random-energy model valid?

To summarize the discussion of this section, the random-energy model is a good approximation only as long as we are interested in maximally compact three-dimensional conformations, symmetric short-range interactions, even compositions, and purely thermodynamic properties. In these cases, correlated conformations are statistically rare and in the course of freezing the system "jumps" between complete overlap and no overlap. However, as soon as we violate any of the conditions (noncompact conformations; restricted set of compact conformations, e.g., crumpled; dimension below three; long-range interactions; uneven composition; or asymmetric short-range interactions, e.g., hydrophobic-polar or Miyazawa-Jernigan), the model breaks down. Furthermore, even when it holds, there are conformations close to the native one that are affected by the design; although they are of no significance in thermodynamics, it is natural to expect that they are exceptionally important for kinetics.

## VI. PROPERTIES OF HETEROPOLYMERS DUE TO CORRELATIONS IN ENERGY LANDSCAPE

### A. Beyond the random-energy approximation

#### 1. The random-energy model as a mean-field approximation

In Sec. IV we described the properties of heteropolymers using the approximation of uncorrelated energies. There are two reasons for exploring an alternate ap-

proach. On the one hand, one may want to go beyond the mean-field approximation, considering fluctuations around both folded and unfolded states. On the other hand, it is desirable to relax the (interdependent) conditions of space uniformity and incompressibility of the globule, allowing for local density fluctuations. Physically, both these alternate approaches have to do with fluctuations, which are immensely important for understanding related problems, such as folding kinetics. Although this goes beyond the scope of this review, it is worth mentioning that in order to address kinetics one has to understand not only small fluctuations and the neighborhoods (in conformation space) close to native and unfolded states, but also all the states along the pathway between native and unfolded phases; in other words, kinetics involves large fluctuations. In its most general form, the problem of fluctuations in heteropolymers and proteins remains to be solved.

#### 2. Non-random-energy theory: A work in progress

In the current literature, there are works directed towards a new theory that will deal with energy correlations among different conformations. Some of these take the phenomenological route, following Bryngelson and Wolynes (1987). In this case, attempts are made to formulate some generalization or replacement for the random-energy model. Other works follow the microscopic route closer to the ideas of Shakhnovich and Gutin (1989a), with the goal of deriving something more general than the random-energy model. Obviously, there are intermediate approaches, mixing ideas of different schools and groups. Whatever the route, it is the general understanding that correlation between energies of conformations is the key aspect of the emerging theory. Indeed, we mentioned in Sec. V that energy correlations are inevitable for a variety of reasons. Thus in this section we discuss how energy correlations may enter the theory and what consequences they may lead to. As our subject here is not a mature theory, we shall restrict ourselves to the simplest possible arguments and considerations.

### B. Interpolation expression for non-random free energy

#### 1. Contributions to the free energy

We shall concentrate primarily on energy correlations due to the lack of compactness of the polymer, because this factor seems to be most obvious from both theoretical and experimental viewpoints. Experimentally we can easily reduce the overall hydrophobic compression of a protein globule by adding some denaturant, such as urea, to the solution. Theoretically we have to address the coil phase and coil-globule transition in a heteropolymer phase diagram. Thus we continue to consider a model based on the Hamiltonian (2.5) with short-range pairwise interactions. However, we relax the condition of maximal compactness and consider all conformations, subject only to chain connectivity and excluded-volume constraints. Such an approach is attractive because it incorporates density as another order parameter along with the native overlap $q$.

We specify the (macro)state of the polymer in terms of two order parameters, namely, $\phi_\alpha$, which is the (averaged) volume fraction occupied by monomer links within the polymer,[18] and $q_\alpha \equiv \mathcal{Q}_{\alpha\star}/\mathcal{Q}_{\max}$, which is the (fractional) overlap of the current conformation $\alpha$ with the native one $\star$, which in turn is set up by the sequence design. We note that $\phi_\alpha$ is proportional to the number density of the polymer; it is also easy to understand that to a mean-field approximation $\phi_\alpha \simeq q_{\alpha\alpha}$. Indeed, $\phi$ is the fraction of occupied sites averaged over the volume of the polymer, while $q_{\alpha\alpha}$ is the fraction of occupied sites among those sites neighboring to monomers. We shall not attempt to specify the coordinate dependence of the density $\phi$, and this should not be understood as consideration of a large spatially uniform globule; quite the opposite, we resort to a Flory-style argument, saying that the chain is described by a single scale expressed as either an average density $\phi$ or the overall size $R \sim a(N/\phi)^{1/3}$ (e.g., rms gyration radius).

We saw in Sec. IV that the random-energy-model results of the previous sections could be derived from a high-temperature expansion. Now, as we do not assume externally supported compactness any longer, we have to clarify that this high-temperature expansion is justified insofar as temperature is high compared to the scale set by the heterogeneity of interactions $\delta B$, not the mean interactions $\overline{B}$ (see Sec. III.C for definitions of $\delta B$ and $\overline{B}$). Thus our plan is to write the Hamiltonian (2.5) as $\mathcal{H} = \overline{\mathcal{H}} + \delta\mathcal{H}$, where the $\overline{\mathcal{H}}$ and $\delta\mathcal{H}$ parts are due to $\overline{B}$ and $\delta B$, respectively, and then perform an expansion in $\delta\mathcal{H}$. We first of all note that

$$\overline{\mathcal{H}}_\alpha \simeq E_{\text{homo}}(\phi) \tag{6.1}$$

is just the homopolymer energy, which to a mean-field approximation is readily written as a function of polymer density $\phi$ only, meaning that it is independent of $q$ as well as of other properties of the conformation $\alpha$. This allows us to write (using here $h = \mathcal{H}/T$ for brevity)

$$\sum_\alpha^{(\phi q)} e^{-\overline{h}_\alpha - \delta h_\alpha} \equiv \sum_\alpha^{(\phi)} e^{-\overline{h}_\alpha} \frac{\Sigma_\alpha^{(\phi q)} e^{-\overline{h}_\alpha - \delta h_\alpha}}{\Sigma_\alpha^{(\phi)} e^{-\overline{h}_\alpha}}$$

$$\simeq \sum_\alpha^{(\phi)} e^{-\overline{h}_\alpha} \frac{\Sigma_\alpha^{(\phi q)} e^{-\delta h_\alpha}}{\Sigma_\alpha^{(\phi q)} 1} \frac{\Sigma_\alpha^{(\phi q)} 1}{\Sigma_\alpha^{(\phi)} 1}, \tag{6.2}$$

where we use the following notations: $\Sigma_\alpha^{(\phi q)}$ means summation restricted to conformations of a given density and native overlap [i.e., $\Sigma_\alpha^{(\phi q)} \cdots = \Sigma_\alpha \cdots \delta(q_\alpha - q)\,\delta(\phi_\alpha - \phi)$, and similarly $\Sigma_\alpha^{(\phi)} \cdots = \Sigma_\alpha \cdots \delta(\phi_\alpha - \phi)$]. Thus the free energy of a given (quenched) sequence heteropolymer can be written in the form

[18] It would be more accurate to define $\phi_\alpha$ as the "fraction of connected bonds"—the ratio of the number of neighbors of all monomers to the maximal possible number of neighbors—see Eq. (2.7)..

$$F_{\rm seq}(\phi,q) = -T\ln\left[\sum_{\alpha}^{(\phi q)} \exp\left[-\frac{\overline{\mathcal{H}}_\alpha + \delta\mathcal{H}_\alpha}{T}\right]\right]$$
$$= F_{\rm homo}(\phi) + \delta F_{\rm seq}(\phi,q) - TS(q,\phi), \quad (6.3)$$

where

$$F_{\rm homo}(\phi) = -T\ln\left[\sum_{\alpha}^{(\phi)} \exp(-\overline{\mathcal{H}}_\alpha/T)\right] \quad (6.4)$$

is the free energy of a homopolymer with averaged interactions and a given degree of compactness $\phi$;

$$S(q,\phi) = \ln\left[\sum_{\alpha}^{(\phi q)} 1 \Big/ \sum_{\alpha}^{(\phi)} 1\right] \quad (6.5)$$

describes the entropy associated with the selection of conformations with the given native overlap $q$ out of all conformations with the given degree of compactness $\phi$; and

$$\delta F_{\rm seq}(\phi,q) = -T\ln\overline{[e^{-\delta\mathcal{H}/T}]}_{\phi q} \quad (6.6)$$

is the purely heteropolymeric part of the free energy. The $\phi q$ average is defined as

$$\overline{[\mathcal{A}]}_{\phi q} = \sum_{\alpha}^{(\phi q)} \mathcal{A}_\alpha \Big/ \sum_{\alpha}^{(\phi q)} 1 \quad (6.7)$$

for every $\mathcal{A}$. Let us discuss contributions to the free energy (6.3) one by one.

### 2. Homopolymer collapse contribution

Historically, homopolymer coil-globule collapse was the first naive model for protein folding (Ptitsyn and Eizner, 1965). Indeed, the homopolymer free energy appears as the first term in the free energy (6.3). Qualitatively, this term describes the entropy price associated with overall chain compaction, the entropy price associated with approaching high packing density, and the energy gain due to the overall predominant attraction of monomers—something typically associated with burst hydrophobic collapse in the biophysical literature. For the $F_{\rm homo}(\phi)$, one can use either a simple Flory-type expression [see, for example, Eqs. (13.2)–(13.4) in Grosberg and Khokhlov, 1994] or a more sophisticated Lifshitz theory. Since we want to consider a very dense globule, the usual approximation neglecting all but leading virial terms is not accurate enough here. Perhaps the simplest expression providing a reasonable interpolation between low-density (virial) and high-density (almost incompressible) regions is

$$F_{\rm homo}(\phi) \simeq NT\left[1 - \chi\phi + \frac{1-\phi}{\phi}\ln(1-\phi)\right], \quad (6.8)$$

where the Flory-Huggins constant is roughly $\chi \approx -(z/2)\overline{B}/T$ and $z = \mathcal{Q}_{\max}/N$ is the coordination number. We omit both the polymer elongation entropy term (which is relevant for chain swelling, not collapse) and the polymer compression entropy term [which is always outweighed by the excluded-volume repulsion hidden in the logarithmic term; see Witelski *et al.* (1998) for further details].

One may ask what happens to our Eq. (6.3) if we consider a homopolymer. It is clear from the definition (6.6) that $\delta F_{\rm seq}(\phi,q) = 0$ vanishes for a homopolymer, and since $q$ has no meaning for a homopolymer, one has to average over $q$ (by summing $\exp[-F_{\rm seq}(\phi,q)/T]$ over $q$). According to Eq. (6.5), this kills the entropy reduction term and leaves us with only the homopolymer contribution, as it must.

### 3. Heteropolymer interaction part of free energy

The truly heteropolymeric dependence on the quenched sequence is in the $\delta F_{\rm seq}(\phi,q)$ term. Following our plan, we resort to the high-temperature expansion yielding

$$\delta F_{\rm seq}(\phi,q) \simeq +\overline{[\delta\mathcal{H}]}_{\phi q} - \frac{1}{2T}\overline{[\delta\mathcal{H}^2]}_{\phi q} + \frac{1}{2T}\overline{[\delta\mathcal{H}]}^2_{\phi q}. \quad (6.9)$$

Now we have to average this free energy over sequences. We use the design probability distribution $P^\star_{\rm seq}$ (2.14). We note that the homopolymeric part $\overline{\mathcal{H}}$ does not show up in $P^\star_{\rm seq}$ [formally, it cancels out in Eq. (2.14) because $\overline{\mathcal{H}}$ is sequence independent], and thus

$$P^\star_{\rm seq} \simeq \left[1 - \frac{\delta\mathcal{H}^{\rm des}(\star)}{T^{\rm des}}\right]\prod_{I=1}^{N} p_{s_I}. \quad (6.10)$$

Performing the average over sequences $\delta F(\phi,q) = \Sigma_{\rm seq} P^\star_{\rm seq}\,\delta F_{\rm seq}(\phi,q)$, we immediately encounter energy correlators similar to what we examined in the previous section; see Eq. (5.1). We shall restrict ourselves here to the case in which $\Sigma_i p_i\,\delta B_{ij} = 0$; this corresponds to the symmetric Ising model $\theta = \pi/2$ for the model with just two monomer species, Eq. (3.46). We also shall not fix the overall composition of the polymer, so that the two last terms vanish in the energy correlator Eq. (5.1). For this case, we arrive at

$$\delta F(\phi,q) \simeq -\mathcal{Q}_{\max}\left[q\,\frac{\delta B^{\rm des}\,\delta B g}{2T^{\rm des}} + (\phi - \overline{[q_{\alpha\beta}]}_{\phi q})\,\frac{\delta B^2}{4T}\right], \quad (6.11)$$

where the fractional overlap $q_{\alpha\beta} = \mathcal{Q}_{\alpha\beta}/\mathcal{Q}_{\max}$ is averaged similarly to Eq. (6.7) with both conformations $\alpha$ and $\beta$ bound to have the same numbers of contacts ($\phi$) and of native contacts $q$.

To proceed, we must estimate the typical overlap between conformations $\overline{[q_{\alpha\beta}]}_{\phi q}$. For large $q$, there are several possible ways of choosing $q$ native contacts, and thus we would expect that

$$\overline{[q_{\alpha\beta}]}_{\phi q} \approx q. \quad (6.12)$$

Although this is a poor approximation for conformations with low $q$, in the region where $q$ is small we do not expect the error in the approximation above to have

significant effects. Furthermore, we stress that the role of this term is subordinate to that of the polymer entropy $S(q,\phi)$, which we estimate below. Thus we can write

$$\delta F(\phi,q) \simeq -\mathcal{Q}_{\max}\left[q\,\frac{\delta B^{\rm des}\,\delta B g}{2T^{\rm des}}+(\phi-q)\,\frac{\delta B^2}{4T}\right]. \quad (6.13)$$

Our result (6.13) has a simple physical interpretation. There are $q\mathcal{Q}_{\max}$ native contacts, and each of them contributes an energy $\epsilon=-\delta B^{\rm des}\,\delta B g/2T^{\rm des}$, which is governed by how the sequence is designed, or how the native state is optimized; $\epsilon$ gets smaller for lower preparation temperature $T^{\rm des}$ and for better alignment $g$ [see Eq. (3.44)] between preparation and folding interaction matrices. On the other hand, all other non-native contacts contribute energies that are governed solely by $-\delta B^2/T$ and are independent of the design. This is also a natural result of our formulae, as $\phi-q$ may be interpreted as the number of non-native contacts. Thus Eq. (6.13) represents just a reformulation of the Go model.

#### 4. Conformational entropy

We now turn to the last term in Eq. (6.3), which is the conformational entropy $S(q,\phi)$. We gave a qualitative discussion of this quantity in Sec. III.A.2. Computing $S(q,\phi)$ is the major unsolved part of the problem, and we shall discuss the expected properties of the $S(q,\phi)$ entropy and deduce the implications for the phases and properties of a heteropolymer.

To begin with, let us note that $\exp[S(q,\phi)]$, considered as a function of $q$ at a fixed $\phi$, satisfies an obvious normalization condition,

$$\int_0^{\phi}\exp[S(q,\phi)]dq=1, \quad (6.14)$$

and represents the probability distribution of all conformations with the given degree of compactness $\phi$ over their respective numbers of native contacts $q$. For maximally compact conformations, when $\phi=1$, we discussed the $\exp[S(q,\phi)]$ distribution and gave computational arguments suggesting that it has two distinct peaks. One of the peaks is at $q=1$, which is the native state itself, with zero entropy. The other peak is at small $q$ (see Fig. 8) and is expected to approach a $\delta$ function at $q=0$ when $N\to\infty$ [see also Eq. (5.3) as to where this peak is located for finite $N$]. The entropy of this state is $s\mathcal{Q}$. When $\phi$ gets smaller, the native state itself, which is assumed to be maximally compact, does not contribute, but still there should be a peak at very large $q$ because the somewhat softened packing constraint allows for some local wigglings of the chain around its native conformation (see also Fig. 7). To determine how much the $q$ peak broadens and grows requires understanding the geometry of relatively dense (though not maximally dense) nativelike conformations. Clearly this should be sensitive to the local geometry of both polymeric connections between monomers and monomers themselves, including their side groups (if any; see also the very recent work of Klimov and Thirumalai, 1998). In experimental terms, these conformations comprise the molten globule state (Kuwajima, 1989; Haynie and Freire, 1993; Ptitsyn and Uversky, 1994; Ptitsyn, 1995; see below for some further discussion). Apart from these conformations, the problem has not been solved so far even for the simplest chain on a cubic lattice.

To gain better insight, it is advantageous to rewrite the definition (6.5) in the form

$$\exp[S(q,\phi)]=\sum_{\text{Choices of } q}\left[\frac{\Sigma_\alpha^{(\phi,\ \text{particular } q)}1}{\Sigma_\alpha^{\phi}1}\right]. \quad (6.15)$$

Here the internal sum runs over conformations with a fixed particular set of $q$ native contacts, while the external sum runs over all possible ways of choosing those $q$. Thus the internal sum corresponds to what can be called loop and bond entropy, as the former corresponds to the number of ways to place the chain loops between fixed native contacts, while the latter corresponds to the number of ways to position the native contacting pairs themselves. As the latter quantity is proportional to $q$, we can write

$$\exp[S(q,\phi)] = \sum_{\text{Choices of } q}\exp[S_{\rm loop}(\phi,\text{particular } q)+Nsq]. \quad (6.16)$$

Replacing the remaining sum by the number of ways to choose a set of $Nq$ native contacts times the properly averaged value, we obtain

$$S(q,\phi)=S_{\rm mix}(q,\phi)+\langle S_{\rm loop}(\phi,\text{particular } q)\rangle_{\rm ann}+Nsq, \quad (6.17)$$

where, as follows from Eq. (6.16), $\langle\cdots\rangle_{\rm ann}$ means an annealed average (in the sense that the number of states, not the entropy, has to be averaged). Thus we recover Eq. (3.5), derived earlier (see Sec. III.A.2) from simplistic considerations. While Eq. (6.17) is exact, we see now that the simple expressions (3.2) and (3.4) are not always valid. Indeed, there are a couple of complications:

- First of all, how much freedom is there to choose native contacts given the numbers of native ($q$) and non-native contacts? If the density is high, and the number of native contacts to be made is also high, then the choice is likely to be very much restricted. Indeed, if the globule is very compact ($\phi$ is close to unity), and almost all of the native contacts are present ($q$ also close to unity), then there are few possibilities for local rearrangements of the conformation. On the other hand, if the globule is very loose ($\phi$ is small), the few native contacts it can have ($q\leqslant\phi$) can be chosen arbitrarily. Thus we expect mixing entropy to be suppressed compared to the naive combinatoric expression $-(q/\phi)\ln(q/\phi)-[1-(q/\phi)]\ln[1-(q/\phi)]$.
- The fact that loop entropy must be an annealed average means that the main contribution may not be due to configurations with all loops of comparable lengths $1/q$, as we assumed in Eq. (3.2). For in-

stance, if we consider an artificial model in which all monomers forming native contacts are allowed (and required) to be at one point in space (in violation of the excluded-volume restriction), we can compute the annealed average loop entropy. It turns out that when $q$ is large (more specifically, close to $\phi$), then the main contribution comes from configurations in which the few missing contacts cluster together along the polymer, and this yields $(d/2)\ln 2(q/\phi)\ln(q/\phi)$ instead of our naive result $(d/2)(q/\phi)\ln(q/\phi)$. Note that $2/\ln 2 \approx 2.89 \approx 3$. Thus already at $d=3$ we come close to what happens only in $d=2$ in the simplified model (Sec. III.A.2), namely, $q \ln q$ terms from mixing and loop entropies may cancel each other, thus washing out the entropic barrier between states with low $q$ and those with high $q$.

One could write an interpolation expression for the entropy $S(q,\phi)$ accounting for the qualitative discussion above. We shall not do this because such an expression would include a number of rather restrictive and crude approximations. Nevertheless, the qualitative picture that arises is robust. What is important about our discussion here is the steps beyond the random-energy approximation: states with partial overlaps to the native state are explicitly included. A somewhat different version of a theory that violates the random-energy model has been developed by Plotkin *et al.* for nondesigned (1996) and designed (Plotkin *et al.*, 1997) sequences. Although their formulae are somewhat different from what we have just described, they incorporate the same physical factors, such as a relaxed compactness constraint, and quenched and annealed constraints on configurations of loops. Therefore the next step is to examine the physical consequences of these factors.

### C. Phase diagram of a nonrandom heteropolymer

To describe the phase diagram of this system, we have to optimize the free energy, Eq. (6.3), with its contributions (6.8), (6.13), and (6.17) with respect to both the fraction of native contacts $q$ and the overall compactness $\phi$, subject to the constraint $q \leqslant \phi$. We first note that the coefficient of the term linear in $q$ in Eq. (6.13) may flip its sign (depending, for instance, on $T^{\mathrm{des}}$); this is where the folding phase transition occurs in a random-energy theory [see Eq. (3.39)]. However, as we now have entropy terms with a nontrivial $q$ dependence, the folding transition point gets somewhat shifted from its location in random-energy theory; most important, this transition is generally not from $q=0$ to $q=1$ but from some small yet positive $q$ to some $q$ less than unity and close to $q=\phi$.

On the other hand, for the $q=0$ case, we recover the homopolymer free energy, yielding a coil-globule transition between phases with $\phi \approx 0$ (coil) and some nonzero positive $\phi$ (globule). If $q=0$, the transition point is given by $\overline{B} - \delta B^2/T + 1/z = 0$. To the accuracy of the present theory, this transition is second order (as $\phi$ itself is continuous), and it goes from the coil ($q \approx 0$, $\phi \approx 0$) to the random globule ($q \approx 0$, $\phi \neq 0$). However, if, for instance, $T^{\mathrm{des}}$ is low enough, then the phase that competes with the coil is not the random globule, but the folded globule ($q \approx \phi \neq 0$). The transition from the coil to the folded globule is first order.

Thus we identify the following states:

Coil: $q \approx 0$, $\phi \approx 0$. An extreme denaturant, such as the $6M$ GuCl that Tanford first used in 1968 (Tanford, 1968), corresponds to a very good solvent for all monomers in a protein. In fact, the $T/2$ factor of the repulsive homopolymeric term is not only stronger than $\overline{B}$, but is even stronger than the effective attraction due to heteropolymeric effects $\overline{B} - \delta B^2/T$. In this case, the protein is effectively a homopolymer and conformational entropy dominates: the chain swells and looks roughly like a self-avoiding random walk. The number of native contacts is negligible. This coil regime is also what one would expect from a random-energy theory.

Unfolded state: $q \ll 1$, $\phi \ll 1$. However, as the overall solvent quality becomes just a little worse, some native contacts appear, although the particular choice of native contacts fluctuates greatly (which is more favorable entropically), i.e., loops are of the annealed type. The random-energy model predicts that the unfolded state will have no native contacts and will be either a coil or a random globule (depending on whether monomers on average attract each other). An unfolded state with native contacts results from the crossover from annealed to quenched loops; without this effect, there would simply be a first-order transition from coil to molten globule.

Folded globule: $q \simeq \phi \simeq 1$. Essentially all native contacts are present and there are only minor conformational fluctuations. There is virtually no conformational entropy and instead all of the contact energy is present. However, unlike the pure random-energy model, which formally predicts exactly zero entropy, nonrandom theory yields that even the folded phase consists of several states structurally very close to the native one (Frauenfelder and Wolynes, 1994). This corresponds to the protein native state.

Molten globule: $q$ is large, but $q<1$. Unlike the unfolded regime, there are *many* native contacts and there is a nonvanishing overlap between conformations found in equilibrium. Thus the molten globule has a much lower energy (many more native contacts) and entropy than the unfolded state; loop entropy is negligible and the mixing entropy is responsible for stabilizing the molten globule. Although this is consistent with the molten globule seen experimentally in many proteins (see Ptitsyn, 1995 and Uversky and Ptitsyn, 1996), we stress that in the analytic model outlined here the molten globule does not represent a separate phase, while experimentally it does. This can be directly explained by the inclusion of the non-random-energy theory presented above, while taking account of the small-scale fluctuations of nonpointlike monomers (Shakhnovich and Finkelstein, 1982). The effect of side chains and the existence of the molten globule phase will be discussed in greater detail in the next section.

All these results are summarized on the schematic phase diagram shown in Fig. 12. In our model, there is only a single phase transition, from coil to folded globule; however, as we distinguish unfolded and molten globule states, we can say that the transition occurs between the unfolded and the molten globule states. This transition is a true phase transition, driven by loop entropy: one cannot continuously collapse a heteropolymer loop into the nucleus of the globule. This transition has been seen experimentally in proteins but was not previously physically understood (Ptitsyn, 1995).

### D. Simulations

The regimes discussed in the previous section—unfolded (UF), molten globule (MG), and native state (NS)—have been found as distinct thermodynamic phases in recent lattice-model simulations (Pande and Rokhsar, 1997). The lattice model discussed here is a 36-mer on a cubic lattice with independent-interaction model interactions. In the past, the ground state of the lattice model was often associated with the molten globule, as one assumed that the lattice model represented a coarse-grained model of proteins and thus could not distinguish between molten globule and the native state, which differ in the nature of side-chain packing; thus one might think that the native state and molten globule could both be represented by the ground-state conformation in the lattice model. However, while the lattice model does not explicitly have side chains, it does have a sense of packing. Off-lattice models (see, for example, Pande and Rokhsar, 1997) without side chains can continuously "breathe," i.e., gradually swell, when placed in any compact conformation. Side-chain packing prevents this type of uniform motion. In the lattice model, the packing of the monomers performs the same function. Thus packing in the lattice model reduces the number of almost-native states, just as side chains do.

As the lattice model has a packing character, one would expect from the Shakhnovich-Finkelstein theory that there would be a first-order molten-globule/native-state transition in addition to a first-order unfolded/molten-globule transition, as predicted in the previous section. This is precisely what was found. The free energy shown in Fig. 13 clearly has three free-energy minima, indicating the unique phases. The number of native contacts is a good order parameter and delineates the unfolded (small $\mathcal{Q}$), molten globule (intermediate $\mathcal{Q}$), and native-state (maximal $\mathcal{Q}$) phases. Note that the unfolded minima are not at $\mathcal{Q}=0$, but rather have many native contacts.

Experimentally, one must add denaturant to stabilize the molten globule phase. This destabilizes the native state by making native contacts less favored, but not to the extent that the unfolded state dominates. In simulations, the molten globule phase was stabilized in a similar manner by adding noise to the interaction matrix, i.e., $B^{\text{fold}}=\sqrt{1-g^2}B^{\text{des}}+g^2B^{\text{noise}}$, where $B^{\text{noise}}$ is a Gaussian random, symmetric matrix with unit variance and vanishing mean and $g$ is the degree of destabilization we wish to perform. Thus $g$ is related to the denaturant concentration. The phase diagram for this system as a function of temperature $T$ and denaturant $g$ is shown in Fig. 14.

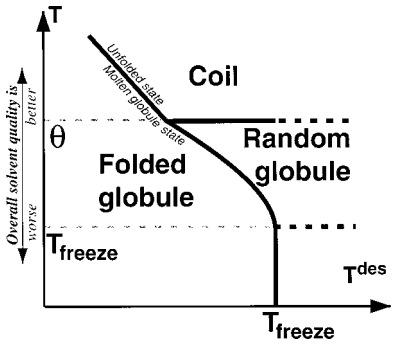



FIG. 12. Phase diagram based on a nonrandom theory. This phase diagram uses the same control variables in Fig. 1(a), but incorporates several features resulting from energy correlations that could not be present in a random-energy-model phase diagram [Fig. 1(a)]. Most notable is the coil phase and the transitions to coil from both folded and random globule states. The coil-random globule transition occurs close to the averaged $\theta$ point, is relatively smooth, and is generally similar to the coil-globule transition in an "averaged" homopolymer, i.e., a homopolymer with the effective attraction $\overline{B}-\delta B^2/4T$. Accordingly, the random globule-frozen globule phase transition occurs generally when solvent is poor. By contrast, the coil-folded globule transition is a purely heteropolymeric effect: it is first order and occurs in a moderately good solvent; strong attractive contacts between monomers designed to be coupled in the native state outweigh the modest overall repulsion in a good solvent. Furthermore, for well-designed (low-$T^{\text{des}}$) sequences, there are some native contacts even in the coil phase close to the transition, giving rise to the experimentally observed unfolded state (even though it is not a separate phase). Similarly, the folded globule close to the transition has some conformational mobility and represents a mixture of states structurally similar to the native one. The appearance of these states similar to the native state is precisely due to energy correlations. While molten, this globule state is not a separate phase in our model; it is expected to be a phase in a more sophisticated model that allows for small-scale rotations for, say, side atomic groups of the monomers (i.e., side-chain packing; Shakhnovich and Finkelstein, 1982). There is evidence that the molten globule state is a separate phase for some proteins (Ptitsyn, 1995), although this is not universally accepted. Our simple nonrandom theory is not applicable for poorly designed sequences, with $T^{\text{des}}>T_{\text{freeze}}$, and this is why it does not yield a glassy phase; phase transitions in that region are shown by dotted lines.

Finally, the loop and mixing entropy arguments employed in previous sections can be directly tested computationally. In particular, we have argued that entropy causes the first-order unfolded/molten-globule and unfolded/native-state transitions. Moreover, this entropy is the result of a delicate cancellation of the loop and mixing entropy terms. To test this computationally, we have calculated the mixing entropy by recording the probability $p_i$ that a particular set of contacts $i$ are found in our Monte Carlo simulation. The mixing entropy is

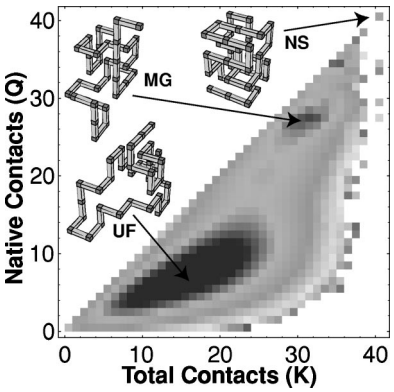



FIG. 13. Free energy vs the number of native contacts ($\mathcal{Q}$) and the total number of contacts ($K$) of a 36-mer lattice model designed heteropolymer near the three-phase coexistence temperature and denaturant value. Note that the number of contacts $K$ is directly related to the averaged volume fraction $\phi$ in the theory.

therefore $S_{\rm mix}(\mathcal{Q})=\Sigma_i\delta(\mathcal{Q}_i-\mathcal{Q})p_i\ln p_i$. One can also calculate the total entropy $S_{\rm total}(\mathcal{Q})$ by calculating the probability that the system will reach a conformation with $\mathcal{Q}$ contacts (in comparison to a *particular* set of $\mathcal{Q}$ contacts $i$ used in the mixing entropy calculation). As $S_{\rm bond}(\mathcal{Q})$ is the linear term in $S_{\rm total}(\mathcal{Q})$, i.e., $S_{\rm bond}(\mathcal{Q})=-\mathcal{Q}(S_{\rm total}(\mathcal{Q}_{\rm max})-S_{\rm total}(0))/\mathcal{Q}_{\rm max}$, we can calculate the loop entropy: $S_{\rm loop}(\mathcal{Q})=S_{\rm total}(\mathcal{Q})-S_{\rm bond}(\mathcal{Q})-S_{\rm mix}(\mathcal{Q})$.

The loop, mixing, and sum of the loop and mixing entropies are shown in Fig. 15. We see that the sum of the mixing and loop entropies leads to an entropy maximum (free-energy minimum) at small $\mathcal{Q}$ and an entropic minimum (free-energy maximum) at intermediate $\mathcal{Q}$. As the energy is linear in $\mathcal{Q}$, energetics cannot lead to a barrier in $F(\mathcal{Q})$ and thus this entropic cancellation, shown here computationally and discussed and derived analytically in previous sections, is the cause of the first-order phase transition in this system.

### E. Nonrandom properties of compact polyampholytes

In this section we consider another example of effects due to correlations of energies between states. We have

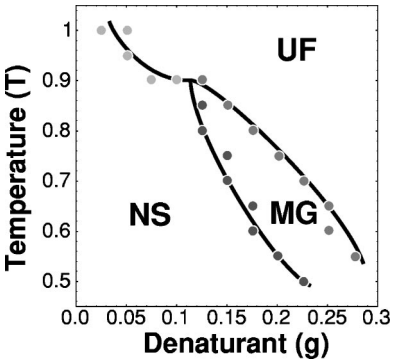



FIG. 14. Phase diagram for a 36-mer under denaturation. We see three phases: unfolded, molten globule, and native state. The boundaries of these phases were calculated by the examination of the free energy as a function of the number of native contacts and the total number of contacts (Fig. 13).

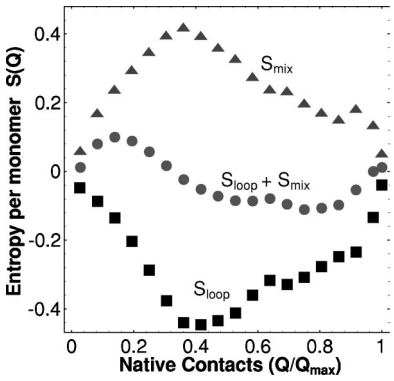



FIG. 15. Polymer entropies $S_{\rm mix}$, $S_{\rm loop}$, and $S_{\rm mix}+S_{\rm loop}$ for a 36-mer. There is an entropic minimum (which corresponds to a free-energy maximum) at intermediate $\mathcal{Q}$, due to the cancellation of loop and mixing entropies. It is this free-energy maximum that makes the transition first order.

already seen that such correlations exist for the case in which interactions are long range (see Sec. V.E). These correlations lead to the non-self-averaging behavior of even the mean of energy spectra for polyampholytes. To show this, we compare the widths of the energy spectra for both quenched and annealed polyampholytes (we assume that the polymer, whether quenched or annealed, is externally maintained in a maximally compact state). These two quantities are given by

$$\sigma^2_{\rm ann}\equiv\langle\overline{(E^2)}\rangle-\langle\overline{E}\rangle\langle\overline{E}\rangle$$

$$\sigma^2_{\rm quen}\equiv\langle\overline{(E^2)}\rangle-\langle(\overline{E})^2\rangle, \tag{6.18}$$

where $\overline{\cdots}$ and $\langle\cdots\rangle$ denote averaging over conformations and sequences, respectively. The difference between these two quantities,

$$\Gamma\equiv\sigma^2_{\rm ann}-\sigma^2_{\rm quen}=\langle(\overline{E})^2\rangle-\langle\overline{E}\rangle\langle\overline{E}\rangle \tag{6.19}$$

characterizes the variation of the energies from one sequence to the other. In the random-energy model, $\Gamma=0$; in fact, $\Gamma=0$ can serve as another definition of the random-energy model.

In the annealed case, the energy variance is given by $\sigma^2_{\rm ann}=B^2\mathcal{Q}_{\rm max}$; this result is directly seen from Eqs. (5.1) and (5.4), or by the physical argument that the annealed case can access all possible states (both conformations and sequences), and thus the width of the energy states must be maximal. On the other hand, as $\Gamma$ can be interpreted as a correlator between typical states, we get $\Gamma=B^2\mathcal{Q}_{\rm rand}$. Thus $\sigma^2_{\rm quen}=B^2(\mathcal{Q}_{\rm max}-\mathcal{Q}_{\rm rand})$. This makes sense physically as correlation in the energies should narrow the width of the energy spectra, while anticorrelation should broaden it. We also see that when there is no correlation ($\Gamma=0$), $\sigma_{\rm ann}=\sigma_{\rm quen}$, as in the random-energy model for $T>T_{\rm freeze}$. We summarize these results (in units of $B^2$) as follows:

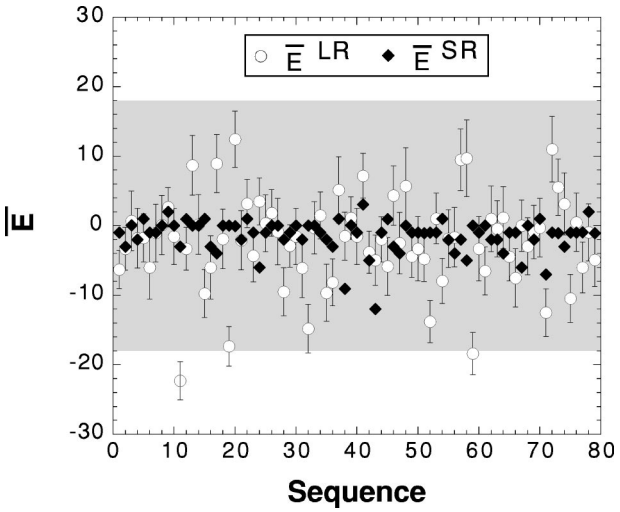



FIG. 16. Mean and width of the energy spectra for 80 sequences of 36-mers, determined by full enumeration over all maximally compact conformations (see text for details).

$$\sigma^2_{\rm ann} = \mathcal{Q}_{\rm max},$$
$$\sigma^2_{\rm quen} = \mathcal{Q}_{\rm max} - \mathcal{Q}_{\rm rand},$$
$$\Gamma = \mathcal{Q}_{\rm rand}. \tag{6.20}$$

Therefore we can put all of these ideas together to form the following picture. For the short-range case above the freezing temperature, $\mathcal{Q}^{\rm SR}_{\rm rand}=0$. Thus $\Gamma=0$ and we expect the mean of the energy spectra for a given sequence not to vary between sequences. Also, the width of the energy spectra for a given sequence should be large (as large as the annealed case, which is the maximum possible). As the variation of the mean of the energy spectra between sequences $\Gamma$ is much smaller than the typical width of the energy spectra $\sigma^2_{\rm quen}$, we see that disorder is not important. It is only for $T \leqslant T_{\rm freeze}$ that this argument breaks down as the system becomes correlated, i.e., $\Gamma$ approaches $\mathcal{Q}^{\rm SR}_{\rm max}$ due to freezing and we see that disorder is important.

For the long-range case, $\mathcal{Q}^{\rm LR}_{\rm rand}$ does not vanish and is significant. We expect the width of the energy spectra to be small and the means to vary widely from sequence to sequence. Thus in the long-range case disorder is always relevant.

In Fig. 16, this is tested computationally using the results from our exact enumeration of all globular states of cubic lattice 36-mers in $d=3$. We see that for short-range interactions, the mean of the spectra is indeed well defined and the width of the spectra (gray region) is large. For long-range interactions, the mean is poorly defined and indeed the variance of the mean of the spectra between sequences is greater than the width of the spectra (error bars). Quantitatively, the simulation yields $\Gamma$ as measured by the variance of the mean to be $\Gamma_{\rm mean}=41$ and the averaged spectra width to be $\sigma^2_{\rm quen}=14$. Thus we computationally calculate that for the maximally compact 36-mers $\mathcal{Q}^{\rm LR}_{\rm rand}/\mathcal{Q}^{\rm LR}_{\rm max}\approx 0.8$. The same value as found by directly calculating the overlaps $\mathcal{Q}$ is 0.62. Thus quantitative predictions with the above formulae and physical picture are also reasonably accurate.

However, the behavior in the thermodynamic limit may be different. For freezing involving a Gaussian density of states, the important scale to consider is the width of the Gaussian $\sigma^2_{\rm quen}/B^2 = \mathcal{Q}_{\rm max} - \mathcal{Q}_{\rm rand}$. For short-range interactions, $\mathcal{Q}_{\rm rand}=0$ and $\mathcal{Q}_{\rm max}\sim N$; thus the width scales as $\sigma_{\rm quen}\sim N^{1/2}$. While this width yields the scaling from the ground state to the peak of the distribution, it also gives the scaling of other energy-spectrum-related quantities such as the typical energy gap. For long-range interactions, $\mathcal{Q}_{\rm rand}$ does not vanish, but note that both $\mathcal{Q}_{\rm rand}$ and $\mathcal{Q}_{\rm max}$ scale like $N^{2-2/d}$. We expect that $\sigma_{\rm quen} \sim N^{1-1/d}$ and the energy spectra for a given sequence should be wider (in the thermodynamic limit) for the long-range case than for the short-range case. However, we see from the computational results for short chains that since $(\mathcal{Q}^{\rm LR}_{\rm max} - \mathcal{Q}^{\rm LR}_{\rm rand})/\mathcal{Q}^{\rm LR}_{\rm max}$ is small compared to unity, the width is smaller. Thus, for some $N$, there should be crossover between these two regimes.

## VII. APPLYING HETEROPOLYMER THEORY TO PROTEINS

The subject matter of the present review—the statistical theory of a single heteropolymer molecule—can potentially have many different applications, ranging from synthetic copolymers (especially those made to simulate proteins) to ''nontraditional'' polymers, such as flux lines. However, since the primary application is proteins, we devote this section to a discussion of insights into proteins gained from the theory of heteropolymers.

### A. Is heteropolymer theory applicable for proteins?

Our models capture the physics of designed heteropolymers, with quenched sequences that fold to a globular native state. One may question whether the incorporation of these features is sufficient or whether other protein-specific details are important for folding. We think that quenched designed heterosequence and globular state do indeed capture the most fundamental aspects of folding. This notwithstanding, there are some other aspects which are certainly missing in the models presented in this review. We summarize some of these below.

#### 1. Secondary structure

Potentially the most important drawback of the models discussed in this review is the omission of protein secondary structure, such as $\alpha$ helices and $\beta$ sheets, which fit poorly in the current theoretical scheme. In particular, if we consider a protein as a large polymer globule, then we expect the Flory theorem to apply. This stipulates that pieces of the chain inside the globule should obey Gaussian statistics. Gaussian statistics of polymer loops play an important role in the theories discussed in the previous sections. On the other hand, there is substantial experimental evidence that elements of secondary structure play an important role in many aspects of folding. While one can assert that the mono-

mers in our models are not the bare amino acids, but some renormalized entities such as elements of the secondary structure, these renormalizations have not yet been rigorously performed. Some important steps in this direction have been reported by Doniach *et al.* (1996; see also the review article of Garel *et al.*, 1996), who examined geometrical competition between secondary structure (simplified to chain preference to go straight) and overall compactness. When both factors are present, there exist two distinct globular phases, of which one lacks secondary structure and is liquid-like, while the other is highly ordered and crystallike. (Similar albeit less sophisticated models have also been considered by Grosberg, 1984; Zhou *et al.*, 1996; and Sear, 1997.)

### 2. Designability of conformations

On the topic of conformational restrictions manifested in secondary structure, we mention the following computer experiment: using 27-mers with a certain $2\times2$ interaction matrix [which corresponds to $\theta=0.09$; see Eq. (3.46)], Li *et al.* (1996) were able to perform a complete enumeration and energy calculation of all 103 346 maximally compact conformations for all $2^{27}$ possible sequences. The authors then defined the designability of each conformation $\star$, $\mathcal{N}_\star$, as the number of sequences for which this conformation serves as the ground state. The designability varied widely for different conformations. In general, the idea that some of the conformations may better fit the needs of proteins than others was suggested and studied earlier (Finkelstein *et al.*, 1993). While it is tempting to naively associate features of 27-mer conformations (or even some two-dimensional conformations) directly with some properties of secondary structure, such an approach is at best speculative and appears to have no physical or biological foundations. However, we admit that the very fact of a broad distribution of designabilities remains largely unexplained. We stress that this effect has its origin in the finite size of the globule and in the role of its shape and surface. Indeed, as we have seen several times in this review, all compact native conformations $\star$ have the same ground-state energy (4.17), stability gap or folding temperature (4.8), and designability (4.22) as long as volume approximation (or linear in $N$ terms in energy and entropy) is considered. Of course, given the modest size of real proteins, the inclusion of their surface (and, in general, sublinear in $N$) properties is an important unsolved problem.

### 3. Monomers are not pointlike

In a more realistic model of proteins, monomers (taken either as bare amino acid residues or pieces of secondary structure) would not be pointlike: their orientations affect both the polymer connections of each monomer to its neighbors along the chain and contacts and interactions with other monomers in space. Furthermore, these monomers have internal degrees of freedom, allowing for oscillations, fluctuations, and even conformational isomerizations.

None of these factors is currently considered in the simplified models discussed here. All-atom models contain these properties, but they are often computationally too intractable to simulate folding trajectories. Ideally one could add these properties from simpler models to examine these effects in a tractable model. Some important steps in this direction have been reported by Micheletti, Seno, *et al.* (1998), who simulated the polymer with different monomers having different excluded volumes, and by Klimov and Thirumalai (1998), who suggested a lattice model for the chain with side groups, yielding a remarkable increase in the cooperativity of folding.

These additional degrees of freedom are relevant in order to address the question of the entropy of proteins at zero temperature (Frauenfelder and Wolynes, 1994; Goldanskii, 1989). Does it vanish, as required by the Nernst theorem for each equilibrium system, or is there some residual positive entropy, as is the case for disordered systems such as glasses? Or are proteins a crystallike or glass-like system? Apart from terminology, we remark that in the framework of the simple models considered above, such as lattice models, the answer is strict: the entropy vanishes, not only at zero temperature, but everywhere below freezing temperature $T_{\text{freeze}}$. Although nobody has been able to show the nature of the low-temperature entropy experimentally, there is little doubt that a more sophisticated model, which would incorporate internal degrees of freedom, allowing for small-scale rotations and/or oscillations of the monomers, would exhibit a freezing transition similar to that examined above (with a residual entropy present at $T<T_{\text{freeze}}$ and related to the aforementioned small-scale fluctuations).

Theoretically one could imagine two distinct kinds of such models: those with vanishing zero-temperature entropy, in which small-scale degrees of freedom equilibrate at each temperature, or one with a nonzero entropy and a glasslike behavior of small-scale fluctuations. Judging from the experimental data, as far as we know, it looks more plausible that the latter scenario is more applicable to proteins (Frauenfelder and Wolynes, 1994). However, whichever of these alternatives is closer to the truth has little bearing on the problem of folding. Folding occurs at a vastly larger scale and, from that point of view, it seems almost certain that the main features of folding, such as evolutionary requirements for foldable sequences, are correctly captured by some coarse-grained models and at best weakly depend on the small-scale features of monomers important in low-temperature experiments.

### 4. Molten globule state as a distinct phase

The most obvious drawback of a model that disregards the local freedom and chemical specifics of amino acids is the fact that the molten globule state does not appear as a distinct phase in the phase diagram (Fig. 12). For proteins, the molten globule state is defined (Ptitsyn, 1995) as a state with a nativelike backbone con-

formation, but a high degree of fluctuation (hence the name "molten"). In terms of more familiar quantities, such as the fraction of native contacts $q$, the molten globule state has $q$ close to, but less than, unity. Molten globule-stateshave been seen experimentally in many proteins (Ptitsyn, 1995), and the phase transitions between molten globule, unfolded, and native states have been observed (Uversky and Ptitsyn, 1996). Recently the molten globule state has been observed to be a distinct phase in proteinlike lattice heteropolymer simulations (Pande and Rokhsar, 1997). Recent simulations of off-lattice Monte Carlo and molecular dynamics simulations of a small protein (fragment $B$ of protein $A$) have also shown a molten-globule-like state (Boczko and Brooks, 1995; Olszewski *et al.*, 1996; Pande and Rokhsar, 1997). The molten globule state is believed to be stabilized by the entropy of local fluctuations of the monomers (Shakhnovich and Finkelstein, 1982) and the mixing entropy of choosing native contacts (see Sec. VI.B).

The unfolded state is also poorly understood. Although there is growing experimental evidence (Shortle, 1996) that the denatured phase of proteins corresponds to the unfolded regime described above, and an unfolded state with a nonvanishing fraction of native contacts has been seen in recent simulations of both proteinlike lattice-model heteropolymers and an off-lattice simulation of fragment $B$ of protein $A$, a small three-helix bundle (Pande and Rokhsar, 1997), experimental understanding of these states is still incomplete.

Our models do not consider protein-specific chemical aspects; they exclude the effect of side-chain packing and thus cannot produce a first-order molten-globule-to-folded-phase transition. We do, however, consider the forces that stabilize the molten globule—competition between mixing entropy and native contact energy—to be a general property of heteropolymers, and therefore we expect that analogous regimes may be found in synthetic heteropolymers without side chains (Chakraborty *et al.*, 1997).

Finally, interactions in proteins are neither pairwise nor purely short range. Although we discussed briefly the role of long-range interactions, we did not consider the mixed case in which both long- and short-range interactions act simultaneously, in a potential conflict.

### B. What does the theory of designed heteropolymers tell us about proteins?

With all the aforementioned complications considered, the heteropolymer theory described here undoubtedly captures many of the key properties of proteins, including the fact that they are polymers, that they are in the compact globular form, and that they have quenched sequences of different monomers. Potentially the best argument to support these claims is that proteins exhibit correlations in their sequences that are fully consistent with the ideas of design and energy optimization (see Sec. I.C.3.b above). It is therefore perfectly justified to look at proteins in light of what we have learned about model heteropolymers.

#### 1. Where in the heteropolymer phase diagram are proteins? The stability gap

If we forget about secondary structure and other complications and consider proteins to be designed heteropolymers, then from our discussions of the previous sections we argue that proteins should function (and should therefore typically be found *in vivo*) at $T > T_{\rm freeze}$. Indeed, for $T < T_{\rm freeze}$, polymer kinetics will be drastically slowed down due to trapping into low-energy conformations unrelated to the native-state conformation. Obviously, proteins must also be in the folded globule state, and thus it is only the gap region $T_{\rm fold} > T > T_{\rm freeze}$ that is appropriate for proteins.

We have also mentioned many times that sequence selection at lower $T^{\rm des}$ provides numerous advantages due to better optimization and elimination of frustration in the ground state. As the gap width $T_{\rm fold} - T_{\rm freeze}$ increases monotonically when $T^{\rm des}$ gets smaller [see Eq. (4.8)], we can equivalently say that this wide $T_{\rm fold} - T_{\rm freeze}$ gap is advantageous in terms of minimal frustration or optimal ground state. This seemingly trivial reformulation is useful. Indeed, we should keep in mind that real proteins have not been prepared following the design prescription based on Eq. (2.14); on the contrary, the design approach, assuming it is correct, pretends only to mimic the results of a complex kinetic evolution process. If we are given a protein, or even a set of several different proteins, there is no obvious way to attach a certain value of $T^{\rm des}$ to them. On the other hand, both $T_{\rm fold}$ and $T_{\rm freeze}$ can be measured—the former directly by observing the folding transition, and the latter by examining randomized sequences of the same composition. Thus the dimensionless gap width

$$\Delta = (T_{\rm fold} - T_{\rm freeze})/T_{\rm freeze} \tag{7.1}$$

is believed to be the important characteristic of proteins.

While our derivation proceeds from the idea of proteins as designed heteropolymers, similar conclusions have been made using physical arguments considering the nature of the random-energy model [see, in particular, Goldstein *et al.* (1992) and Wolynes (1994), and references therein].

Let us now summarize which properties of proteins can be explained by treating them as large-$\Delta$ heteropolymers:

- A broad gap provides stability against folding mistakes of various kinds.
- A broad gap provides stability against mutations. To see this, we can reinterpret Eq. (4.7), resolving it with respect to $g$ and saying that it yields the minimal value of $g$ or maximal value of mutation angle $\theta$ ($\cos\theta = g$) for which the system remains in the folded phase. Replacing $T^{\rm des}$ with $\Delta$ using that same Eq. (4.7), but with $g = 1$ and $T = T_{\rm fold}$, and assuming also for simplicity that $\theta$ is small, we arrive at the

conclusion that a protein remains correctly folded until mutations are weaker than

$$\theta < \theta^* \simeq \Delta. \tag{7.2}$$

With greater optimization, a greater number of mutations will be necessary to destabilize the native state to the point at which renaturation is not possible.

- A broad gap provides for a highly cooperative folding transition, as observed in proteins. Poorly designed narrow-gap heteropolymers (best exemplified by heteropolymers with random sequences) have noncooperative, almost second-order transitions, whereas well-designed sequences have extremely cooperative, first-order folding transitions (Nymeyer and Onuchic, 1997; Pande and Rokhsar, 1997). In poorly designed chains, the energy of nonnative contacts is essentially the same as that of native contacts. Thus there is no strong energetic preference toward native contacts and the system samples many non-native contacts, which leads to broadening and smoothing of the transition. For well-designed sequences, typically only native contacts are found, which leads to a sharp first-order transition.
- It has been argued in the literature that the same logic applies to kinetic properties: sequences with a broader gap are better solvers of the Levinthal paradox (see, in particular, Klimov and Thirumalai, 1996 and Shakhnovich, 1997).

The idea of the gap as the necessary condition for good folding has been suggested by Goldstein *et al.* (1992) and Shakhnovich and Gutin (1993) and has recently been reviewed by Shakhnovich (1997). There have also been numerous works in which authors have naively interpreted the gap mechanically, as the gap between the lowest and second-lowest energy states (conformations). This interpretation would indeed be correct in a pure, narrowly defined random-energy model in which design affects only the native state $\star$ and even the second-lowest energy level is completely unaffected. However, as we know, nonrandom effects are inevitable, and thus in reality—*in vivo, in vitro*, and *in silico*—there are always energy levels not very far from the lowest one; they correspond to states that are structurally closely related to the native, and thus low-energy excitations correspond to structurally modest fluctuations. This is why a real understanding of the stability gap requires a statistical rather than a mechanical approach, i.e., consideration of the gap between temperatures $T_{\rm fold}$ and $T_{\rm freeze}$. Not surprisingly, the mechanical approach leads to confusion, because the energy gap between the two lowest levels is in general neither necessary nor sufficient for anything useful with respect to folding.

### 2. Determining parameters for proteins

Although enormous effort is being devoted to experimental studies of proteins, there is a fundamental difficulty in determining $\delta B$, $s$, and other heteropolymer parameters based on experimental information. The problem is that monomers of the theory are not to be identified with amino acid residues, but with some renormalized entities; as we do not know how to perform the renormalization, the only way out is to treat heteropolymer parameters phenomenologically.

#### a. Conformational entropy s

We remind the reader that the value of $s$ was introduced in Eq. (3.34), as the entropy reduction per bond due to the placing of one replica exactly in the conformation of the other. Equivalently, this parameter also appeared in Eq. (4.9), as the entropy per contact for a dense globular conformation, where $\exp \mathcal{Q}s$ is the number of such conformations. In reality, conformations of replicas coincide down to a microscopic but finite scale—the $\delta$ function in $Q_{\alpha\beta}(\mathbf{R}_1,\mathbf{R}_2) \sim \delta(\mathbf{R}_1-\mathbf{R}_2)$ (3.33) is concentrated on neighboring points $\mathbf{R}_1$ and $\mathbf{R}_2$. Equivalently, the conformation itself is defined only up to fluctuations on a microscopic scale. If this scale corresponds to a monomer volume $v$, and if the chain connection between monomers is associated with a length scale $a$ then we can estimate $s \simeq \ln(a^3/v)$. The $v/a^3$ parameter is well known in polymer theory (Grosberg and Khokhlov, 1994), and it also involves chain flexibility. For a moderately flexible polypeptide chain of a protein, it is believed that $v/a^3 \approx 0.2$ (Creighton, 1992). Thus

$$s \approx 1.6 \text{ for a flexible polymer.} \tag{7.3}$$

This estimate completely neglects the role of secondary structure.

Another estimate can be obtained using lattice models. As maximally compact conformations have been enumerated for 18-, 27-, 36-, and 48-mers (Pande, Joerg *et al.*, 1994), we can easily calculate $s$ for those models, and it turns out to be 0.39, 0.41, 0.45, and 0.46, respectively. It has been argued (Leopold *et al.*, 1992; Luthey-Schulten *et al.*, 1995) that the conformational space of the cubic lattice 27-mer is similar to that of a protein with 60 to 80 residues, as it has a similar number of contacts in the compact state, around 30, and the number of conformations is also believed to be similar because it is suppressed for proteins due to secondary structure. With this reasoning, we can estimate

$$s \approx 0.4 \text{ with secondary structure.} \tag{7.4}$$

#### b. Protein interactions

The determination of the $20\times 20$ $\hat{B}$ matrix of interaction energies for amino acids is a difficult task and lies beyond the scope of the present review. We mention only that the most advanced methods are "knowledge based," that is, they employ the statistics of the protein structures (conformations) taken from the data bank (Abola *et al.*, 1987), rather than quantum chemistry or other microscopic theoretical computations or experimental measurements. We mention four studies devoted

to the determination of $\hat{B}$ using various knowledge-based approaches: (Miyazawa and Jernigan, 1985; Kolinski *et al.*, 1993; Hinds and Levitt, 1994; Mirny and Shakhnovich, 1996). For example, Miyazawa and Jernigan (1985) counted all the contacts between monomer species $i$ and $j$ throughout all proteins from the data bank and then related $B_{ij}$ to the logarithm of that number. This makes the assumption that the contacts in different proteins are governed by a Gibbs distribution, an assumption that may seem unfounded, as the set of proteins in the database has really nothing to do with a thermodynamic Gibbs ensemble. However, as long as the random-energy model is valid, these two ensembles have the same distributions, as we have mentioned previously in this review. Moreover, as long as we believe that proteins are designed heteropolymers, we can argue that the ensemble of proteins is actually a Gibbs ensemble corresponding to the temperature $T^{\rm des}$. Thus what has been determined by Miyazawa and Jernigan (1985) approximates the quantity $B_{ij}/T^{\rm des}$ (see also Finkelstein *et al.*, 1993 and the Appendix in Pande, Grosberg, and Tanaka, 1995c). Unfortunately, the Miyazawa and Jernigan (1985) approach to potential determination relies heavily on the validity of the random-energy model and also assumes a direct relation between contacts (pair correlation function) and energies ($B_{ij}$), which is problematic for a condensed system such as a globule.

The different estimates of interaction matrices have recently been compared by Du *et al.* (1998) using the idea that correlation between interaction matrices should be defined according to Eq. (3.44). This analysis yielded the conclusion that all known matrices agree poorly with each other. The very existence of different poorly correlated estimates raises the question, How accurate must potentials be for simulations of protein folding? This question is addressed in the next section.

### c. Potentials for simulations of protein folding

Consider that proteins have been indeed designed according to our theoretical model. This means that they were prepared with interactions $\hat{B}^{\rm des}$ at some temperature $T^{\rm des}<T^{\rm des}_{\rm freeze}$, and now they "work" at some other temperature $T$, such that $T^{\rm des}_{\rm freeze}<T<T_{\rm fold}$. Their work is governed by their natural interactions, that is, by the same matrix $\hat{B}^{\rm des}$ as was supposedly used for design. We now take some other artificial matrix $\hat{B}$ and try (for example, by means of computer simulation) to recover the correct renaturation. In terms of our phase diagram, correct renaturation occurs when and only when the system remains in the folded phase. As a greater degree of error in the potentials pushes the phase boundary to the left, determined by the value of the factor $g$ (3.44), we conclude that correct renaturation at temperature $T$ is possible for a chain designed at $T^{\rm des}$ if and only if the degree of correlation is sufficient (or the angle is small enough) between $\hat{B}^{\rm des}$ and $\hat{B}$, namely, when

$$g>g^*=\frac{T^{\rm des}}{T^{\rm des}_{\rm freeze}}\,\frac{1+(T/T_{\rm freeze})^2}{2T/T_{\rm freeze}}, \tag{7.5}$$

where $g^*$ is defined from the condition that the phase boundary $T_{\rm fold}$, given by Eq. (4.7), goes through the given point $(T^{\rm des},T)$. The purely theoretical value of $T^{\rm des}$ can be replaced with the measurable value of the gap $\Delta$ (7.1), as we did before. This yields the result, which we write for simplicity in terms of the angle $\theta$ defined as $\cos\theta=g$ and assuming also that $\theta$ is small,

$$\theta<\theta^*=\left[\Delta^2-\left(\frac{T-T_{\rm freeze}}{T_{\rm freeze}}\right)^2\right]^{1/2}\leqslant\Delta. \tag{7.6}$$

We conclude from this equation that minimal "correctness" of the interaction matrix $g^*$ or $\theta^*$ is defined by the degree of optimization of the ground state that we would like to renature, and this is quantified in terms of the gap $\Delta$, as well as several other properties discussed above. A similar but more general estimate can also be obtained from the non-random-energy-model free energy, by the condition that the system remain inside the folded phase, whether below the transition to the random globule or to the coil (or unfolded state).

We stress that this error limit is independent of the length of the polymer. This is to be compared with the error limit of about $1/\sqrt{N}$ found by Bryngelson (1994) for the possibility of reconstructing a randomly chosen conformation of the glassy (frozen) globule phase. In the ensemble of random sequences, sequences with very stable ground states are very rare (exponentially so); thus these ground states are typically very unstable with respect to error-based renaturation, especially for long chains. In contrast, our treatment is a comparison between the types of freezing (to the native or some random conformation). This treatment is therefore independent of the length of the polymer chain and is essentially of a different nature than the pure random-energy-model effect examined by Bryngelson (1994).

Pande, Grosberg, and Tanaka (1995c) have addressed the question of which types and values of errors in the determination of interactions lead to a particular value of $g$ factor. None of the linear ($B_{ij}=\beta B^p_{ij}+B_0$) systematic errors contribute at all, yielding $g=1$ (because these kinds of errors contribute to homopolymer terms only and do not affect the selectivity of monomer interactions to one another). To get an idea of random mistakes, we examine the case in which each matrix element of the renaturation matrix is that of the design matrix plus some noise, which is normally (and independently for different entries) distributed with a variance $\sigma$. This gives $\sigma\approx10\%$ for $g^*\approx0.95$. Thus even conservative estimates leave room for about 10% error in potentials.

### d. Composition

There are 20 different types of amino acids used as monomer species in proteins. Their occurrence $p_i$ is not even, differing from 0.05 and ranging from 0.015 to 0.087. Thus the effective number of species (4.20) is $q_{\rm eff}\approx18$.

### C. How optimal are present-day proteins?

#### 1. Quantitative aspects of evolutionary optimization

While our discussion directly considers the temperatures $T$, $T_{\rm freeze}$, $T_{\rm fold}$, and $T^{\rm des}$, our theory should be understood in a broader way. The strength of interactions can be controlled by temperature and also by a variety of other factors, such as solvent composition and mutations from a wild-type sequence. This is particularly important for the interpretation of $T^{\rm des}$. Indeed, proteins are believed to have been "designed" by evolution, at both the prebiological and biological stages. Our $T^{\rm des}$ based description envelops many different factors. From that point of view, although our theory is microscopic in nature, it operates with parameters that should be understood as purely phenomenological; in particular, $T^{\rm des}$ is just a phenomenological descriptor of the degree of optimization of the native conformation. Moreover, as the natural scale for $T^{\rm des}$ is $T_{\rm freeze}$, we conclude that a quantitative measure of the quality of sequence design is the ratio $T^{\rm des}/T_{\rm freeze}$: the smaller $T^{\rm des}/T_{\rm freeze}$ yields a broader gap region and thus better folders. Alternatively and even more phenomenologically one can judge the quality of design based on the value $\Delta$ of the gap itself.

One can ask what the numerical values of $T^{\rm des}/T_{\rm freeze}$ or of $\Delta$ are for proteins found in modern living organisms. The information available at present is insufficient to make any solid statement on that matter. To get some rough idea, we can use the Miyazawa and Jernigan (1985) interaction matrix for amino acid residues. As we explained earlier, it represents the quantity $B_{ij}^{\rm MJ} = B_{ij}/T^{\rm des}$. On the other hand, according to Eq. (4.3) we have $\delta B = 2s^{1/2}T_{\rm freeze}$. Thus the variance of the Miyazawa-Jernigan matrix yields $\delta B^{\rm MJ} = 2s^{1/2}T_{\rm freeze}/T^{\rm des}$. This variance is easy to compute directly ($\delta B^{\rm MJ} \approx 1.5$). If we adopt the estimate (7.4) for $s$, we arrive at

$$T^{\rm des}/T_{\rm freeze} \approx 0.85 \tag{7.7}$$

Given the uncertainties in determination of both $s$ and $\delta B$, this value should not be considered solidly determined, but rather as an illustrative example. However, according to Eq. (4.7), it translates into the "gap"

$$\Delta = (T_{\rm fold} - T_{\rm freeze})/T_{\rm freeze} \approx 0.6, \tag{7.8}$$

which is consistent with the value independently calculated by Onuchic *et al.* (1995).

#### 2. How far has evolution progressed?

What we want to stress is not so much the numerical values of either $T^{\rm des}/T_{\rm freeze}$ or $\Delta$, but the very possibility of quantifying the evolutionary optimization of proteins to perfect their functions, and of quantifying it in terms of measurable quantities.

This should be viewed in the broader context of evolution theory. People used to think that everything in biology is optimized to absolute perfection (and it is hard to avoid thinking this way when looking at the amazing performance of living organisms). In sharp and dramatic contradiction, the selection of the "best" out of the sea of possibilities is prohibitively consuming both of time and of materials (Monod, 1971; Volkenstein, 1994). In the last decades, the so-called neutral theory of evolution (Kimura, 1983) has been proven broadly applicable. This theory stipulates that the lion's share of all mutations are neutral, carrying neither functional nor structural advantages/disadvantages. In line with this way of thinking, proteins are no longer viewed as absolutely optimal (Ptitsyn and Volkenstein, 1986), implying that the degree of optimization may have been improving over the course of evolution and may even keep improving slowly at the present stage of evolution. It is exciting to think about evolution in terms of $T^{\rm des}/T_{\rm freeze}$ or $\Delta$ considered as a time-dependent (on an evolutionary scale) dynamic variable.

From that point of view, one may ask how far the optimization can go. One answer to this lies in Eq. (4.24), which says that the sequences are not yet the best as long as $T^{\rm des}/T_{\rm freeze} \geqslant s/\ln q_{\rm eff} \approx 0.2$, or $\Delta < 1.2$. Of course, progressive optimization poses increasingly difficult selection problems, and it seems safe to assume that evolution can never approach anywhere near the limit set by Eq. (4.24). It is also undoubtedly true that there are other factors affecting the "quality" of proteins besides $T^{\rm des}/T_{\rm freeze}$ or $\Delta$; in particular, reliability and the rate of folding kinetics can be improved dramatically, even with constrained $T^{\rm des}/T_{\rm freeze}$ or $\Delta$, provided that ground-state energy optimization, described by $T^{\rm des}/T_{\rm freeze}$ or $\Delta$, is already good enough; see Mirny *et al.* (1998). Nevertheless, our estimates indicate that there is still plenty of room left for further evolutionary improvements of proteins. Indeed, this has been seen in experiments, as *de novo* designed proteins are typically unfolded at higher temperatures, reflecting greater energetic stabilization and thus lower $T^{\rm des}$ (Dahiyat and Mayo, 1996).

## VIII. CONCLUSIONS

We have been pursuing a statistical approach to proteins, which means that we did not examine any specific example of a protein, but rather general properties of some broad ensembles of them. Conceptually, this approach can be viewed as the examination of what could have happened during evolution, rather than what has happened in the only example that we know, i.e., proteins. Technically, this approach implies the consideration of certain averages over the ensemble of random or designed sequences.

We stress that, although averaging over the realizations of disorder might seem to be the only sensible approach for a physicist, it is actually not so in biology. Indeed, the synthetic apparatus of the living cell (when considered as a piece of equipment for physical experiments) is able to produce *macroscopic amounts of identical copies* of a disordered system (protein molecules). This is to be compared with the unimaginable method of preparing two samples of, for example, a spin glass with

microscopically identical structures. Until recently, only the opposite approach was possible in biology, where one routinely examines a particular protein. These two approaches are complementary [see also the famous talk by Bohr (1935)]: by examining a particular protein, one can learn little regarding ensemble properties (including those members of the ensemble that never existed, but presumably could have appeared if evolution had chosen another path); on the other hand, by examining an ensemble, one can learn almost nothing about the specifics of any particular protein. Recently, due to biotechnological experimental methods, the examination of ''general properties'' of proteins has been gathering momentum: there are works on random-sequence polypeptides, with further selection performed based on their folding properties (Ptitsyn, 1995; Sauer, 1996; Dahiyat and Mayo, 1997); completely artificial sequence polypeptides, constructed according to certain *a priori* considerations (Hecht *et al.*, 1990); and artificially created mutations of natural protein sequences (Fersht, 1995). In this review we have been considering theoretical aspects of such an approach. What were the results?

First, under certain circumstances, there is a freezing phase transition for a heteropolymer with a random sequence. In the simplest case—when (i) the polymer is externally supported in a dense globular state (e.g., in a box or poor solvent overall), but conformations are not restricted otherwise, and (ii) interactions between monomers are short range—the freezing temperature $T_{\text{freeze}}$ is governed by the diversity of heteropolymeric interactions (e.g., variance of the interaction matrix). A freezing transition occurs from a state with typically no overlap with the ground state to a state with complete overlap, that is, to the ground state itself. This discontinuity is largely due to the nature of conformational space, as relatively few local rearrangements of the conformations exist in a dense globule because of the excluded-volume constraint. Thus small correlations exist between most relevant states (giving a dominant contribution to the partition function). This in turn gives rise to an approximation in which one assumes no correlations between any states; this approximation is equivalent to the random-energy model.

Thus even a heteropolymer with a random sequence can have a unique ground-state conformation. Nevertheless, the ground state of a random sequence is not proteinlike. Indeed, while formally there is one ground-state conformation, random-energy-model low-energy states are only $\sim\sqrt{N}$ from each other. These states are typically structurally unrelated to the native state. This leads to nonproteinlike behavior, including a complete alteration of the ground-state conformation due to destabilization of the original ground state caused by, for example, only a minute change of solvent composition.

However, the random-energy view of an uncorrelated energy landscape invokes a simple way to select, or design, sequences whose energy minimum is well pronounced. Sequences with these properties are very atypical (exponentially rare) in the ensemble of all (random) sequences. However, as soon as sequence selection is weighted with the Boltzmann factor $\exp[-E_\star/T^{\text{des}}]$, the energy $E_\star$ of the desired native conformation $\star$ is about $N$ below most of the other states, which remain statistically unaffected as long as the random-energy-model is valid and there are no correlations between states. A lower design (or selection) temperature $T^{\text{des}}$ leads to stronger pulling down of the ground-state energy $E_\star$ or, in other words, better energy optimization of the native state or less frustration in the native state.

Sequence design appears to be a highly successful paradigm for understanding many aspects of proteins. First, it is consistent with the fact that protein sequences have indeed been affected by evolutionary selection. Second, design predicts nonrandomness of protein sequences; moreover, the character of correlations along the chain, based on the interactions of monomers in other directions, is at least qualitatively consistent with design. Third, design explains why and how the native state could be thermodynamically stable while all unrelated states are not stable and thus do not serve as traps. This suggests that design must be viewed as performed somewhat below $T_{\text{freeze}}$ while proteins have to work somewhat above $T_{\text{freeze}}$ ($T^{\text{des}}<T_{\text{freeze}}$, but $T>T_{\text{freeze}}$; see phase diagram, Fig. 2). Fourth, design explains the remarkable stability of the native state to a variety of perturbations, including solvent composition, mutations, etc. (which is clearly demonstrated by the examination of folding and design governed by somewhat different interactions). Finally, the idea of sequence design is a strong stimulus for experiments, both with polypeptide chains and with synthetic polymers and gels.

This also suggests that one should view the design temperature (or rather the ratio $T^{\text{des}}/T_{\text{freeze}}$) as a phenomenological parameter defining the degree of selection. Lower values of $T^{\text{des}}/T_{\text{freeze}}$ mean better energy optimization and thus better resolution of the Levinthal paradox. However, there are far fewer sequences with a smaller value of $T^{\text{des}}/T_{\text{freeze}}$, and thus it is more difficult to select them—they are worse at resolving the evolution paradox. One might expect that the current state of evolution represents a compromise between these two opposing factors.

Although the random-energy view of an uncorrelated energy landscape, when complemented with the idea of design, explains several aspects of proteins, it is not successful in other respects. First and foremost, a pure random-energy model would leave the Levinthal paradox unresolvable. Indeed, in a pure random-energy model, the system, even when it is conformationally very close to the native state, still does not have any hint of the location of the ''goal'' and keeps walking randomly in highly dimensional conformation space; in other words, a pure random-energy model does not provide any bias towards the native state (or anywhere else). Thus to understand the kinetics of folding and to finally resolve the Levinthal paradox one must consider deviations from the random-energy model, with a correlated energy landscape. To this end, the terminology of a ''protein folding funnel'' has been introduced (Leopold

*et al.*, 1992). The original use (Leopold *et al.*, 1992) painted a physical picture in which correlations between energies of different states provide the bias for the system to fold to its native state, although it has come to have a broader application (Bryngelson *et al.*, 1995; Dill and Chan, 1997).

However, to understand how this physical picture is connected with proteins, the first step is to understand the nature of the preferred kinetic path (or reaction coordinate) for protein folding kinetics. The elucidation of the correct reaction coordinate remains unsolved. Moreover, to build a theory that would be able to exploit the attractive phenomenological idea of a correlated landscape biasing kinetics, one has to determine the physical and geometric meaning of the reaction coordinate, including why the system travels along this path, what happens to the directions perpendicular to the path, why states are correlated along the path, and why this correlation drives the system toward the native state along this path. These questions lie at the heart of a nonrandom-energy theory of heteropolymers.

We must admit that this picture of heteropolymers is much less developed, and the problems are poorly understood at present. We know several factors that drive the system away from the random-energy model and provide energy correlations, including a long-range component in the interactions, some conformational restrictions (e.g., excess of local crumples), overall swelling of the globule, density fluctuations, the surface of the globule, and local fluctuations due to small-scale internal degrees of freedom of monomers. We understand that these factors bring about new phases, or distinct (macro)states, of the polymer, such as the unfolded state (which is not simply a self-avoiding random walk) and the molten globule state. We also understand that these correlations have an important effect on sequence design, since they lower the energies of several conformations even though selection is aimed at lowering specifically a target state. These ''accompanying'' conformations are structurally related to the native one; geometrically, they form a kind of braid of conformations intertwined around the native conformation. Furthermore, the closer the conformation in this braid is to the native conformation, the more strongly it is affected by the design. Thus it is natural to expect that such conformations can indeed provide an energetically biased path toward the native state.

This paper has reviewed the theory of heteropolymer freezing and design up to the limits of our current understanding. In the introduction to this review, we quoted the final sentences of the *Reviews of Modern Physics* article by Lifshitz *et al.* (1978) on homopolymer globules, and we mentioned that our present review begins precisely where Lifshitz *et al.* (1978) left off 20 years ago. It is hard to resist guessing what will be discussed in yet another *RMP* review in another 20 years. We feel that the concepts of heteropolymer freezing and design are solidly established. Although we have been able to discuss them for simple models only, the ideas discussed seem to be useful as a first step toward better approximations. In particular, it appears almost certain that factors such as secondary structure, rotations and deformations of monomers (as they are not points), and density fluctuations and surface of the globule should and could be taken into account, even though that would require generalizations of the model based on Hamiltonian (2.5). Finally, the elucidation of the mechanism of folding kinetics, including the nonrandom nature of design and the geometry of folding pathways, remains an exciting and challenging unsolved problem.

## APPENDIX A: GLOSSARY OF PHASES

Since the terminology involved is often vague and confusing, we present a glossary with short explanations of some terms as we have used them in the present review.

*Coil* completely open phase of a polymer, either a Gaussian or swollen (self-avoiding) random-walk state. A coil is a strongly fluctuating state in which long-range correlations enforced by the chain involve the entire polymer.

*Denaturation* abrupt transition of a protein from its native state to a denatured state.

*Denatured state of a protein* the state at which a protein function is completely lost, although chemically the protein remains completely intact. Commonly attributed to the loss of native three-dimensional fold.

*Freezing* phase transition of a heteropolymer between a ''normal'' state with exponentially many $[\mathcal{O}(e^N)]$ conformations and a frozen or folded state in which only one or very few conformations $[\mathcal{O}(1)]$ dominate equilibrium.

*Frozen state* same as the glassy state.

*Folded state* the globular state dominated by the correct native conformation or by *a priori* chosen target conformation; similar to target state.

*Folding of a protein* experimentally observed phenomenon in which a protein macromolecule under appropriate solvent conditions goes spontaneously from an open (or at least denatured) state to the native one.

*Glassy state* phase state of a heteropolymer globule that is dominated by one or very few conformations.

*Globule* or *globular phase* condensed phase of a polymer chain. As opposed to the coil, correlations are suppressed in a globule, such that density fluctuations are mild.

*Misfolded state* glassy state of a chain whose sequence has been designed under ''distorted'' interactions, different from folding ones.

*Molten globule* a phase of protein globules, which is usually close to the native globule as regards the amount of secondary structure and overall character of the chain fold; it is, however, somewhat less compact than native. Transitions from the molten globule state to coil and to native globule are accompanied by considerable latent heat and are believed to be first order.

*Random globule* globular phase of a heteropolymer that is dominated by many $[\mathcal{O}(e^N)]$ conformations and in this sense is similar to a homopolymer globule.

*Target state* globular phase of a designed heteropolymer that is dominated by the conformation $\star$ chosen by the design.

*Unfolded state* state in which there are some native contacts between monomers, i.e., nonzero overlap to the native state, but the degree of openness is closer to the coil, though denser. It remains unclear whether there is a phase transition between this state and the coil.

## APPENDIX B: PROPERTIES OF THE RANDOM-ENERGY MODEL

### 1. Density of states for the random-energy model

We first examine the energy spectrum of a typical random-energy-model realization of disorder. It is easy to generate realizations of this spectrum computationally; a few are shown as examples in Fig. 1. Here typical spectra consist of a very dense region, with many states at high and relatively modest energies, and a low-energy part of the spectrum, which is discrete and composed of only a few levels. Formally speaking, the spectrum is discrete everywhere, as there is a finite number of states (and levels). However, this number is huge, and thus the difference between continuous and discrete (large and small density of states) is indeed defined to logarithmic accuracy. By inspection of Fig. 1, one sees that the continuous part of the spectrum looks identical for all realizations of the disorder, while the discrete part is very individual and looks completely different for different realizations.

We can gain insight into these properties of the random-energy-model spectrum by examining the density of states $n(E)$: $n(E)\Delta E$ is the number of states with energies between $E$ and $E+\Delta E$. It is very easy to find the expectation value for $n(E)$:

$$\langle n(E)\rangle = \mathcal{M}P(E). \tag{B1}$$

This value is huge (due to $\mathcal{M}$) whenever $E$ is not far from the central part of the spectrum; this is why an astronomically large set of states forms an almost continuous spectrum at all energies where the probability (2.3) is not very small. When the density of states is so large, it is about the same in all particular realizations, so that $n(E)\simeq\langle n(E)\rangle$. This argument, however, is valid only as long as $\mathcal{M}P(E)>1$, or $E>E^{\text{bottom}}$, where

$$\mathcal{M}P(E^{\text{bottom}})\sim 1 \Rightarrow E^{\text{bottom}} \simeq -N\mathcal{E}\sqrt{2\omega}. \tag{B2}$$

If we go to low energies $E<E^{\text{bottom}}$ where this breaks down, then the expected number of energy levels in an interval $\Delta E$ becomes so low that there is not a single level in the interval of interest:

$$n(E)=\begin{cases}\mathcal{M}P(E) & \text{when } E>E^{\text{bottom}}\\ \text{random peaks} & \text{when } E<E^{\text{bottom}}.\end{cases} \tag{B3}$$

### 2. Typical and atypical realizations in the random-energy model

To gain deeper insight into the properties of the random-energy model, let us look at the energy differences between low-energy states. First of all, one can easily write the probability distribution for the ground-state energy. Indeed, for some state with energy $E$ to be the ground state, all other states must be of higher energy; for each state, the corresponding probability is $\int_E^{+\infty}P(E)dE$ and, most important, as all other $\mathcal{M}-1$ states are independent, we get

$$\wp_{\text{ground}}(E)=P(E)\left[\int_E^{+\infty}P(E)dE\right]^{\mathcal{M}-1}. \tag{B4}$$

Given that $\mathcal{M}$ is so astronomically large, this amounts to[19]

$$\wp_{\text{ground}}(E)\simeq\begin{cases}P(E) & \text{when } E<E^{\text{bottom}}\\ 0 & \text{when } E>E^{\text{bottom}}.\end{cases} \tag{B5}$$

Thus, for the overwhelming majority of the realizations, the ground-state energy is about $\mathcal{E}\sqrt{N}$ below the boundary of the continuous spectrum, $E^{\text{bottom}}$. Note that this boundary itself is about $\mathcal{E}N$ below the mean energy [see Eq. (B2) above]. In this sense, both $E^{\text{bottom}}$ and the ground-state energy are the same to thermodynamic accuracy, and to that accuracy they are self-averaging, independent of the disorder.

Similarly, we can easily compute the probability distribution for the energy difference between the lowest and the second-lowest states. For the two lowest states

$$\begin{aligned}\wp(E_1,E_2)&=P(E_1)P(E_2)\left[\int_E^{+\infty}P(E)dE\right]^{\mathcal{M}-2}\\ &\simeq\begin{cases}P(E_1)P(E_2) & \text{when } E<E^{\text{bottom}}\\ 0 & \text{when } E>E^{\text{bottom}},\end{cases}\end{aligned} \tag{B6}$$

where $E=\max\{E_1,E_2\}$, and finally for the gap $\Delta E=|E_1-E_2|$ we have

$$\wp(\Delta E)=\int\wp(E_1,E_2)\,\delta(\Delta E-|E_1-E_2|)dE_1\,dE_2. \tag{B7}$$

Simple analysis indicates that $\wp(\Delta E)$ is peaked at $\Delta E\sim\mathcal{E}\sqrt{N}$.

---

[19]The following is a more accurate estimate. As long as $N\gg 1$, we can use the saddle-point approximation to get $\int_{E^{\text{bottom}}}^{+\infty}P(E)dE\simeq 1-k\exp[-(E^{\text{bottom}})^2/2N\mathcal{E}^2]\simeq 1-k\exp[-N\omega]$ where $k$ is a numerical constant. This value is very close to unity, but its deviation from unity is just about $1/\mathcal{M}$. Therefore just immediately above $E^{\text{bottom}}$ the value $\int_{E^{\text{bottom}}}^{+\infty}P(E)dE$ becomes small enough that its $\mathcal{M}$th power practically vanishes.

Thus, for a typical realization of the random-energy-model system, all discrete energy levels are just a little bit below the lower threshold of the continuous spectrum and they are very close to each other. Note that the differences between them, which scale as $\sqrt{N}$, are negligible in the thermodynamic limit.

It is vitally important for protein applications that, besides the typical realizations of disorder, there are some rare atypical realizations for which the spectrum looks quite different. In particular, there are some, albeit exponentially rare, realizations for which the ground-state energy is order $N$ below the threshold $E^{\text{bottom}}$, to select them is actually the main purpose of sequence design.

### 3. Thermodynamics of the random-energy model

Consider now the thermodynamics of the random-energy model. While, in principle, one may wish to compute the partition function and the free energy for each particular realization of the disorder, this is clearly impractical for most applications, and what one does instead is to note that the averaged free energy is dominated by the typical realizations of the disorder. Thus we are first of all interested in an average of the form

$$F(T)=\langle F_{\text{seq}}(T)\rangle=-T\langle \ln Z_{\text{seq}}(T)\rangle. \tag{B8}$$

To average the logarithmic function is a tedious mathematical task, and this is precisely why disordered systems are so difficult to examine theoretically. This is the place where the famous replica trick (Mezard *et al.*, 1987) enters. The main good news about the random-energy model is that one does not need to resort to this trick.

Indeed, $Z_{\text{seq}}(T)$, the partition function for a given realization of the disorder, is just the sum over all states $i=1,2,\ldots,\mathcal{M}$, and it can always be rewritten in terms of the density of states:

$$Z_{\text{seq}}(T)=\sum_{i=1}^{\mathcal{M}} \exp[-E_i/T]=\int_{-\infty}^{\infty} n(E)e^{-E/T}\,dE. \tag{B9}$$

At high enough temperature, this sum is dominated by the states of high entropy [large $n(E)$], where the spectrum is continuous and independent of sequence. This means that complications arising from differences between individual realizations of the disorder do not occur in this temperature region and disorder is, in a way, irrelevant. Indeed, as long as the saddle point of the integral

$$Z=\int_{-\infty}^{\infty} \mathcal{M}P(E)e^{-E/T}\,dE\simeq \mathcal{M}P(E_{\text{saddle}})e^{-E_{\text{saddle}}/T}, \tag{B10}$$

belongs to the continuous-spectrum region $E>E^{\text{bottom}}$, the first line of Eq. (B3) is valid, and thus we get a *partition function* that is independent of the disorder.[20] For the Gaussian distribution (2.3), $E_{\text{saddle}}=-N\mathcal{E}^2/T$, we get that $E_{\text{saddle}}>E^{\text{bottom}}$ is valid at $T>T_{\text{freeze}}=\mathcal{E}/\sqrt{2\omega}$. Thus at $T>T_{\text{freeze}}$ we can safely average the partition function over disorder (as it does not depend on disorder!) and arrive at a free energy that is also independent of the disorder.

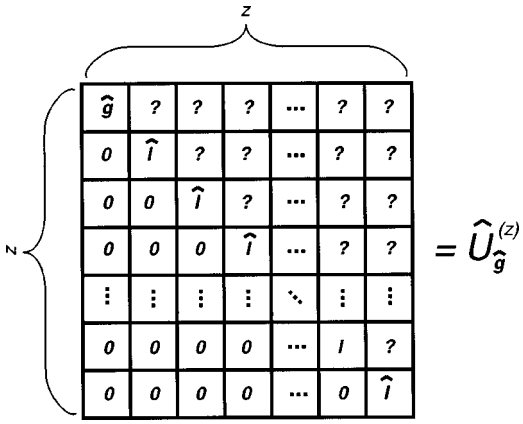
This is not valid at lower temperatures. What happens there is that just one or few low-energy states dominate the partition function. In principle, one could expect that at this low temperature, the thermodynamics of a particular sample would strongly depend on the disorder. Note, however, that typical differences between low-energy states are only about $\sqrt{N}$ and they are negligible in thermodynamic limit. As the free energy is a continuous function of the temperature, we arrive at Eq. (2.4), the powerful conclusion (see also Koukiou, 1993) that disorder is irrelevant above the freezing point for the random-energy model.

## APPENDIX C: ROTATION OF REPLICA SPACE

Our main goal here is to simplify Eq. (3.25).

### 1. Three auxiliary lemmas

Lemma 1. Consider an auxiliary problem of the matrix




[20]The fact that the partition function in the random-energy model is independent of disorder is very different from what one would expect in a typical disordered system. In extensive quantities (linear in the number of particles $N$), differences in contributions from each particle get averaged out when we sum up $N$ of them and take the thermodynamic limit (because of the central-limit theorem). Due to this property, extensive quantities are called "self-averaging" and are in this sense independent of disorder. Quantities such as the partition function (which scales exponentially in $N$) typically do not average in this form (for the partition function, terms with lower free energy dominate the average due to the exponential). Thus the fact that the random-energy-model partition function is self-averaging is very unusual and results from the assumption of the statistical independence of states.

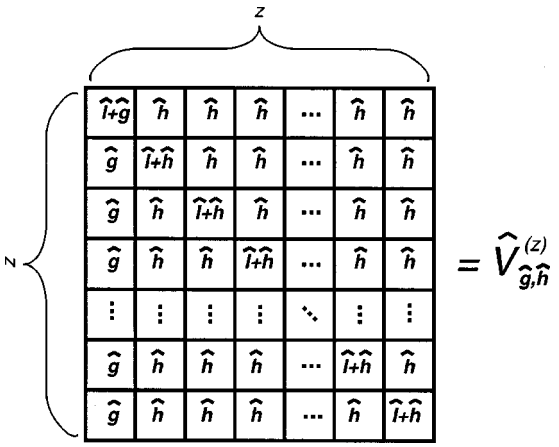
This is a block matrix, where $\hat{g}$ is a $q\times q$ matrix and $\hat{I}$ is an identity matrix of the same size $q\times q$. The question is how to find the determinant of this matrix.

It can be shown by expansion over the elements of the first column, then over the elements of the first column of the remaining minor, and by repeating this operation $q$ times, that

$$\det[\hat{U}^{(z)}_{\hat{g}}]=\det\hat{g} \tag{C1}$$

independently of the blocks placed in the upper right triangle (shown conventionally with question marks).

Lemma 2. Consider another auxiliary problem of the following block matrix:




Here $\hat{g}$ and $\hat{h}$ are matrices $q\times q$; they generally do not commute to each other. $\hat{I}$ is an identity matrix of the same size $q\times q$. The total size of the block matrix $\hat{V}^{(z)}_{\hat{g},\hat{h}}$ is therefore $zq\times zq$. The question is to find the inverse of the matrix $\hat{V}^{(z)}_{\hat{g},\hat{h}}$.

It turns out that this inverse is in fact the matrix of the same structure, namely,

$$(\hat{V}^{(z)}_{\hat{g},\hat{h}})^{-1}=\hat{V}^{(z)}_{\hat{e},\hat{f}},$$

$$\text{where}\quad \hat{e}=-(\hat{I}+(z-1)\hat{h}+\hat{g})^{-1}\hat{g}$$

$$\text{and}\quad \hat{f}=-(\hat{I}+(z-1)\hat{h}+\hat{g})^{-1}\hat{h}. \tag{C2}$$

The result can be easily proved using the block matrix multiplication rule.

Lemma 3. Consider an auxiliary problem of the scalar product

$$\langle\vec{\rho}^{(qz)}|\hat{W}^{(qz)}|\vec{\rho}^{(qz)}\rangle,$$

where $\vec{\rho}^{(qz)}=\vec{p}^{(q)}\otimes\vec{i}^{(z)}=p_i$ (does not depend on replica indices $\alpha$), and $\hat{W}^{(qz)}$ is a block matrix comprised of blocks $\hat{W}^{(q)}_{\alpha\beta}$. Obviously this scalar product is reduced to the scalar products of smaller dimensionality $q$, that is, purely in species space, summed over all the blocks of the matrix:

$$\langle\vec{\rho}^{(qz)}|\hat{W}^{(qz)}|\vec{\rho}^{(qz)}\rangle=\left\langle\vec{p}^{(q)}\middle|\left(\sum_{\alpha\beta}\hat{W}^{(q)}_{\alpha\beta}\right)\middle|\vec{p}^{(q)}\right\rangle. \tag{C3}$$

2. Some general properties of the "direct-product" operation for matrices

We repeat the definition: for two (not necessarily square) matrices $\hat{A}^{(r\times r')}\otimes\hat{B}^{(s\times s')}$ is the $rs\times r's'$ matrix, built up by substitution of the $s\times s'$ block $A_{uv}\hat{B}^{(s\times s')}$ for each matrix element of $\hat{A}^{(r\times r')}$. For example, $|\vec{\rho}^{(nq)}\rangle=\vec{\rho}^{(n)}\otimes\vec{p}^{(q)}$.

(1) By matrix row and column operations, it is easy to show that the rule is commutative, i.e.,

$$\hat{A}^{(r)}\otimes\hat{B}^{(s)}=\hat{B}^{(s)}\otimes\hat{A}^{(r)}. \tag{C4}$$

(2) Block matrix multiplication rule: It is well known that the operation of block matrix multiplication is carried out in the same scheme as normal matrix multiplication, except that the multiplication of elements is replaced by the matrix multiplication of blocks. This can be written as

$$(\hat{A}^{(r)}\otimes\hat{B}^{(s)})\cdot(\hat{A}'^{(r)}\otimes\hat{B}'^{(s)})$$
$$=(\hat{A}^{(r)}\hat{A}'^{(r)})\otimes(\hat{B}^{(s)}\hat{B}'^{(s)}). \tag{C5}$$

(3) Commutation of $\hat{A}^{(r)}\otimes\hat{B}^{(s)}$ and $\hat{A}'^{(r)}\otimes\hat{B}'^{(s)}$ depends on commutation of both pairs $A^{(r)}\&\hat{A}'^{(r)}$ and $\hat{B}^{(s)}\&\hat{B}'^{(s)}$ (this directly follows from the previous property).

(4) The determinant of a block-diagonal matrix equals the product of the determinants of the diagonal blocks. In particular,

$$\det(\hat{A}^{(r)}\otimes\hat{I}^{(s)})=(\det\hat{A}^{(r)})^s. \tag{C6}$$

(5) Matrix operation with a vector:

$$\hat{A}^{(r)}\otimes\hat{B}^{(s)}|\vec{a}^{(r)}\otimes\vec{b}^{(s)}\rangle=\hat{A}^{(r)}|\vec{a}^{(r)}\rangle\otimes\hat{B}^{(s)}|\vec{b}^{(s)}\rangle. \tag{C7}$$

(6) Scalar product of vectors:

$$\langle\vec{a}^{(r)}\otimes\vec{b}^{(s)}|\vec{a}'^{(r)}\otimes\vec{b}'^{(s)}\rangle=\langle\vec{a}^{(r)}|\vec{a}'^{(r)}\rangle\langle\vec{b}^{(s)}|\vec{b}'^{(s)}\rangle. \tag{C8}$$

The proof of all these properties is straightforward.

*a. Elimination of real-space coordinates*

We have to simplify the expression for the operator $\hat{\mathcal{B}}_{\text{eff}}$ (3.25) given that the $\hat{Q}$ matrix is of the form (3.33). The real-space integration is performed trivially due to the $\delta$ functions everywhere, thus producing the factor $N/\rho$ (which is the overall volume of the globule):

$$\mathcal{H}_{\text{eff}}=\frac{N}{\rho}\left[\langle\vec{\rho}|\hat{\mathcal{B}}_{\text{eff}}|\vec{\rho}\rangle+\frac{1}{2}\ln\det\hat{\mathcal{B}}_{\text{eff}}\right],$$

$$\hat{\mathcal{B}}_{\text{eff}}=\frac{1}{2}(\hat{\mathcal{B}}^{1/2})^{(q(n+1))}[\hat{I}^{(q(n+1))}+\rho(\hat{q}^{(n+1)}\otimes\hat{\Delta}^{(q)})$$
$$\times\hat{\mathcal{B}}^{(q(n+1))}]^{-1}(\hat{\mathcal{B}}^{1/2})^{(q(n+1))}. \tag{C9}$$

*b. Elimination of replica variables*

As to the elimination of replicas, we perform it through several steps:

(1) As the matrix $\hat{q}$ is of one-step replica symmetry-breaking shape, as explained above, with one distinct group of $y+1$ replicas and $(n-y)/x$ groups of $x$ replicas each, we can view the matrix $\hat{M}=\hat{I}+\rho\hat{q}\otimes\hat{\Delta}\hat{B}$ as an $(n+1)\times(n+1)$ block matrix in replica space, where each matrix element is a $q\times q$ matrix in species space. This block matrix is of the same structure as $\hat{q}$, with one $(y+1)\times(y+1)$ superblock and $(n-y)/x$ of $x\times x$ superblocks.

(2) The determinant in the first term in the free energy is decomposed into the product of determinants of superblocks.

(3) Vector $\vec{\rho}$ is composed of $n+1$ blocks $p_i$, thus making the second term in the free energy the sum of independent contributions from the groups of replicas. Along with the previous step, this means that different groups of replicas do not interact, and this is why they contribute independently to the free energy.

(4) The effective replica energy $E$ is now presented in the form

$$\frac{E(x,y)}{N}=\epsilon_y+\frac{n-y}{x}\epsilon_x, \tag{C10}$$

where $\epsilon_y$ and $\epsilon_x$ are the (independent) contributions from the corresponding groups of replicas. (Note that the replica entropy is also of the same form.)

(5) Both $\epsilon_y$ and $\epsilon_x$ have almost the same form as $\mathcal{H}_{\rm eff}$ (3.25), except that a simpler matrix $\hat{\tilde{q}}$, with all matrix elements 1, appears instead of $\hat{q}$:

$$\epsilon_z=\frac{1}{2}\ln\det[\hat{I}^{(zq)}+\rho\hat{\tilde{q}}^{(z)}\otimes\hat{\Delta}^{(q)}\hat{B}^{(zq)}] + \frac{1}{2\rho}\langle\vec{\rho}|\hat{B}^{(zq)}[\hat{I}^{(zq)}+\rho\hat{\tilde{q}}^{(z)}\otimes\hat{\Delta}^{(q)}\hat{B}^{(zq)}]^{-1}|\vec{\rho}\rangle, \tag{C11}$$

where $z$ is either $x$ or $y+1$, i.e., the number of replicas in the group.

(6) To simplify the first term (with determinant), we define a rotation unitary operator

$$\hat{\mathcal{R}}^{(z)}_{\alpha\beta}=\frac{1}{\sqrt{z}}\exp\left[\frac{2\pi i}{z}(\alpha-1)(\beta-1)\right],\quad 1\leqslant\alpha,\beta\leqslant z. \tag{C12}$$

It is easy to check that this operator transforms $\hat{\tilde{q}}^{(z)}$ into diagonal form, where one diagonal matrix element is 1, while all others are 0:

$$\hat{\mathcal{R}}^{(z)}\hat{\tilde{q}}^{(z)}(\hat{\mathcal{R}}^{(z)})^{-1}=\hat{\lambda}^{(z)},\quad \text{where } \hat{\lambda}_{\alpha\beta}=z\,\delta_{\alpha1}\delta_{1\beta}. \tag{C13}$$

We also define $\hat{\mathcal{R}}^{(zq)}=\hat{I}^{(q)}\otimes\hat{\mathcal{R}}^{(z)}$ and note that the determinant is not changed upon rotation. We write

$$\begin{aligned}
&\det[\hat{I}^{(zq)}+\rho\hat{\tilde{q}}^{(z)}\otimes\hat{\Delta}^{(q)}\hat{B}^{(zq)}]\\
&\quad=\det[\hat{\mathcal{R}}^{(zq)}]\det[\hat{I}^{(zq)}+\rho\hat{\tilde{q}}^{(z)}\otimes\hat{\Delta}^{(q)}\hat{B}^{(zq)}]\det[\hat{\mathcal{R}}^{(zq)}]^{-1}\\
&\quad=\det[\hat{I}^{(zq)}+\rho(\hat{\mathcal{R}}^{(zq)})\hat{\tilde{q}}^{(z)}\otimes\hat{\Delta}^{(q)}\hat{B}^{(zq)}(\hat{\mathcal{R}}^{(zq)})^{-1}]\\
&\quad=\det[\hat{I}^{(zq)}+\rho(\hat{\mathcal{R}}^{(z)})\hat{\tilde{q}}^{(z)}(\hat{\mathcal{R}}^{(z)})^{-1}\otimes\hat{\Delta}^{(q)}(\hat{\mathcal{R}}^{(zq)})\hat{B}^{(zq)}(\hat{\mathcal{R}}^{(zq)})^{-1}]\\
&\quad=\det[\hat{I}^{(zq)}+\rho\hat{\lambda}^{(z)}\otimes\hat{\Delta}^{(q)}(\hat{\mathcal{R}}^{(zq)})\hat{B}^{(zq)}(\hat{\mathcal{R}}^{(zq)})^{-1}].
\end{aligned} \tag{C14}$$

As $\hat{B}^{(zq)}$ is diagonal in replica space, $\hat{B}^{(zq)}=\hat{B}^{(q)}_{\alpha}\delta_{\alpha\beta}$, we have

$$\begin{aligned}
&((\hat{\mathcal{R}}^{(zq)})\hat{B}^{(zq)}(\hat{\mathcal{R}}^{(zq)})^{-1})_{\alpha\beta}\\
&\quad=\sum_{\gamma\delta}\hat{\mathcal{R}}_{\alpha\gamma}\hat{B}^{(q)}_{\gamma}\delta_{\gamma\delta}(\hat{\mathcal{R}}^{(zq)})^{-1}_{\delta\beta}\\
&\quad=\frac{1}{z}\sum_{\gamma}\exp\left[\frac{2\pi i}{z}(\alpha-\beta)(\gamma-1)\right]\hat{B}^{(q)}_{\gamma}.
\end{aligned} \tag{C15}$$

Taking into account the simple structure of $\hat{\lambda}$ (C13), we arrive at

$$\begin{aligned}
&\hat{\lambda}^{(z)}\otimes\hat{\Delta}^{(q)}(\hat{\mathcal{R}}^{(zq)})\hat{B}^{(zq)}(\hat{\mathcal{R}}^{(zq)})^{-1}\\
&\quad=\delta_{\alpha1}\sum_{\gamma}\exp\left[\frac{2\pi i}{z}(1-\beta)(\gamma-1)\right]\hat{\Delta}^{(q)}\hat{B}^{(q)}_{\gamma}.
\end{aligned} \tag{C16}$$

(7) First consider a nontarget group of $z=x$ replicas. In this group, all the replicas are identical, meaning that $\hat{B}^{(q)}_{\gamma}=\hat{B}^{(q)}$ does not depend on the replica index $\gamma$. This yields

$$\hat{\lambda}^{(z)}\otimes\hat{\Delta}^{(q)}(\hat{\mathcal{R}}^{(zq)})\hat{B}^{(zq)}(\hat{\mathcal{R}}^{(zq)})^{-1}=x\,\delta_{\alpha1}\delta_{1\beta}\hat{\Delta}^{(q)}\hat{B}^{(q)} \tag{C17}$$

and thus

$$\det[\hat{I}^{(xq)}+\rho\hat{\tilde{q}}^{(x)}\otimes\hat{\Delta}^{(q)}\hat{B}^{(xq)}]=\det[\hat{I}^{(q)}+\rho x\hat{\Delta}^{(q)}\hat{B}^{(q)}]. \tag{C18}$$

(8) Consider now a target group of $z=y+1$ replicas. In this case, $\hat{B}^{(q)}_{\gamma}=\hat{B}^{(q)}_{\rm des}$ for $\gamma=1$ and $\hat{B}^{(q)}_{\gamma}=\hat{B}^{(q)}$ otherwise. We write, therefore,

$$\begin{aligned}
&\hat{I}^{(zq)}+\hat{\lambda}^{(z)}\otimes\hat{\Delta}^{(q)}(\hat{\mathcal{R}}^{(zq)})\hat{B}^{(zq)}(\hat{\mathcal{R}}^{(zq)})^{-1}\\
&\quad=\hat{I}^{(q)}\delta_{\alpha\beta}+\delta_{\alpha1}\hat{\Delta}^{(q)}(\hat{B}^{(q)}_{\rm des}-\hat{B}^{(q)})\\
&\qquad+(y+1)\,\delta_{\alpha1}\delta_{1\beta}\hat{\Delta}^{(q)}\hat{B}^{(q)}.
\end{aligned} \tag{C19}$$

This is a block matrix of a peculiar form such that only the upper block is nonzero in the first column; for that reason, its determinant is equal to the product of determinants of diagonal blocks (see Lemma 1). Thus

$$\begin{aligned}
&\det[\hat{I}^{((y+1)q)}+\rho\hat{\tilde{q}}^{(y+1)}\otimes\hat{\Delta}^{(q)}\hat{B}^{((y+1)q)}]\\
&\quad=\det[\hat{I}^{(q)}+\rho\hat{\Delta}^{(q)}(y\hat{B}^{(q)}+\hat{B}^{(q)}_{\rm des})].
\end{aligned} \tag{C20}$$

(9) As to the second term in $\epsilon_z$ (C11), it is easily computed using Lemma 2. Indeed, $\hat{B}^{((y+1)q)}$ is a block-diagonal matrix with one block $\hat{B}^{(q)}_{\rm des}$ and $y$ others $\hat{B}^{(q)}$. On the other hand, $\hat{\tilde{q}}^{(y+1)}\otimes\hat{\Delta}^{(q)}$ is a block matrix with

every block the same $\hat{\Delta}^{(q)}$. Therefore the matrix in question, $[\hat{I}^{(zq)}+\rho\hat{\tilde{q}}^{(z)}\otimes\hat{\Delta}^{(q)}\hat{B}^{(zq)}]$, is exactly of the form $\hat{V}^{(z)}_{\hat{g},\hat{h}}$, where $\hat{g}=\hat{\Delta}^{(q)}\hat{B}^{(q)}_{\rm des}$ and $\hat{h}=\hat{\Delta}^{(q)}\hat{B}^{(q)}$. Using the block matrix multiplication rule, it is easy to compute $\hat{B}^{((y+1)q)}\hat{V}^{(y+1)}_{\hat{e},\hat{f}}$ (see Lemma 2) and then to use the result of Lemma 3. This finally gives

$$\frac{1}{2\rho}\langle\vec{\rho}|\hat{B}^{(zq)}[\hat{I}^{(zq)}+\rho\hat{\tilde{q}}^{(z)}\otimes\hat{\Delta}^{(q)}\hat{B}^{(zq)}]^{-1}|\vec{\rho}\rangle$$
$$=\frac{\rho}{2}\langle\vec{p}|(\hat{B}^{(q)}_{\rm des}+y\hat{B}^{(q)})[\hat{I}^{(q)}+\rho y\hat{\Delta}^{(q)}\hat{B}^{(q)}+\rho\hat{\Delta}^{(q)}\hat{B}^{(q)}_{\rm des}]^{-1}|\vec{p}\rangle. \quad \text{(C21)}$$

(10) A similar expression for a nontarget group of $x$ replicas can be derived from here by formally putting $\hat{B}^{(q)}_{\rm des}\to\hat{B}^{(q)}$ and $y+1\to x$; this gives

$$\frac{1}{2\rho}\langle\vec{\rho}|\hat{B}^{(zq)}[\hat{I}^{(zq)}+\rho\hat{\tilde{q}}^{(z)}\otimes\hat{\Delta}^{(q)}\hat{B}^{(zq)}]^{-1}|\vec{\rho}\rangle$$
$$=\frac{\rho}{2}\langle\vec{p}|x\hat{B}^{(q)}[\hat{I}^{(q)}+\rho x\hat{\Delta}^{(q)}\hat{B}^{(q)}]^{-1}|\vec{p}\rangle. \quad \text{(C22)}$$

Further details regarding this calculation can be found in the appendices for the papers by Pande *et al.* (1995a, 1995b).

## ACKNOWLEDGMENTS

We have benefited greatly from discussions with R. Baldwin, J. Banavar, A. Chakraborty, D. Chandler, K. Dill, C. Dobson, S. Doniach, W. Eaton, I. Ya. Erukhimovich, A. M. Gutin, C. Joerg, M. Kardar, A. R. Khokhlov, M. Levitt, S. Marqusee, V. Munoz, D. Nelson, J. N. Onuchic, K. Plaxco, Y. Rabin, D. S. Rokhsar, E. I. Shakhnovich, and P. G. Wolynes. We would also like to thank D. Butts, R. Du, C. Pethick, D. S. Rokhsar, and E. Shender for their critical reading of the manuscript. V.S.P. acknowledges support from the Miller Institute of Basic Research at the University of California at Berkeley and the Physical Biosciences Division, Lawrence Berkeley National Laboratories (LBNL). The work was supported by the National Science Foundation (NSF DMR-9616791). Computations were performed on Project SCOUT (ARPA Contract No. MDA972-92-J-1032) and on the T3E at NERSC, LBNL.

## REFERENCES

Abe, H., and N. Go, 1981, Biopolymers **20**, 1013.
Abkevich, V. I., A. M. Gutin, and E. I. Shakhnovich, 1994, Biochemistry **33**, 10026–10036.
Abkevich, V. I., A. M. Gutin, and E. I. Shakhnovich, 1995, Protein Sci. **4**, 1167.
Abkevich, V., A. Gutin, and E. Shakhnovich, 1996, Proc. Natl. Acad. Sci. USA **93**, 830.
Abola, E. E., F. C. Bernstein, S. H. Bryant, T. F. Koetzle, and J. Weng, 1987, "Protein Data Bank," in *Crystallographic Databases-Information Content, Software Systems, Scientific Applications*, edited by F. H. Allen, G. Bergerhoff, and R. Sievers (Data Commission of the International Union of Crystallography, Bonn/Cambridge/Chester), p. 107.
Amit, D. J., H. Gutfreund, and H. Sompolinsky, 1987, Ann. Phys. (N.Y.) **173**, 30.
Anfinsen, C. B., 1973, Science **181**, 223.
Binder, K. and P. Young, 1986, Rev. Mod. Phys. **68**, 801.
Binder, K., 1995, Ed., *Monte Carlo and Molecular Dynamics Simulations in Polymer Science* (Oxford University, New York).
Birshtein, T. M., and O. B. Ptitsyn, 1964, *Conformations of Macromolecules* (Nauka, Moscow) [Interscience, New York (1966)].
Boczko, E. M., and C. L. Brooks III, 1995, Science **269**, 393.
Bohr, N., 1935, Lecture given at the congress in Copenhagen.
Bryngelson, J. D., 1994, J. Chem. Phys. **100**, 6038.
Bryngelson, J. D., J. N. Onuchic, N. D. Socci, and P. G. Wolynes, 1995, Proteins: Struct., Funct., Genet. **21**, 167.
Bryngelson, J. D., and P. G. Wolynes, 1987, Proc. Natl. Acad. Sci. USA **84**, 7524.
Bryngelson, J. D., and P. G. Wolynes, 1989, J. Phys. Chem. **93**, 6902.
Caflisch, A., and M. Karplus, 1994, Proc. Natl. Acad. Sci. USA **91**, 1746.
Camacho, C. J., and D. Thirumalai, 1993, Proc. Natl. Acad. Sci. USA **90**, 6369.
Chakraborty, A., E. I. Shakhnovich, and V. S. Pande, 1997, J. Chem. Phys. **108**, 1683.
Chan, H. S., S. Bromberg, and K. A. Dill, 1995, Philos. Trans. R. Soc. London, Ser. B **348**, 6.
Chan, H. S., and K. A. Dill, 1991, Annu. Rev. Biophys. Biophys. Chem. **20**, 447.
Chan, H. S. and K. A. Dill, 1993, Phys. Today **46**(10), 24.
Copart, J., and F. Candau, 1993, Macromolecules **26**, 1333.
Creighton, T. E., 1992, *Protein Folding* (W. H. Freeman, New York).
Cule, D., and T. Hwa, 1997, preprint cond-mat/9701117.
Daggett, V., and M. Levitt, 1992, Proc. Natl. Acad. Sci. USA **89**, 5142.
Dahiyat, B., and S. Mayo, 1997, Science **278**, 82.
Dahiyat, B. I., and S. L. Mayo, 1996, Protein Sci. **5**, 895.
de Gennes, P.-G., 1979, *Scaling Concepts in Polymer Physics* (Cornell University, Ithaca).
DeGrado, W. F., 1997, Science **278**, 80-1.
Derrida, B., 1980, Phys. Rev. Lett. **45**, 79.
Derrida, B., 1985, J. Phys. (France) Lett. **46**, L401.
Derrida, B., and E. Gardner, 1986a, J. Phys. C **19**, 2253.
Derrida, B., and E. Gardner, 1986b, J. Phys. C **19**, 5783.
Deutsch, J. M., and T. Kurosky, 1996, Phys. Rev. Lett. **76**, 323.
Dill, K. A., and H. S. Chan, 1997, Nat. Struct. Biol. **4**, 10.
Dinner, A., A. Sali, M. Karplus, and E. Shakhnovich, 1994, J. Chem. Phys. **101**, 1444.
Dobrynin, A. V., and M. Rubinstein, 1995, J. Phys. II **5**, 677.
Doi, M., and S. F. Edwards, 1986, *The Theory of Polymer Dynamics* (Oxford University, Oxford).
Doniach, S., J. Bascle, T. Garel, and H. Orland, 1995, J. Mol. Biol. **254**, 960.
Doniach, S., T. Garel, and H. Orland, 1996, J. Chem. Phys. **105**, 1601.

Dressler, D., and H. Potter, 1990, *Discovering Enzymes* (Scientific American Library, New York).
Du, R., A. Yu. Grosberg, and T. Tanaka, 1998, Folding Des. **3**, 203.
Du, R., V. S. Pande, A. Yu. Grosberg, T. Tanaka, and E. I. Shakhnovich, 1997, J. Chem. Phys. **108**, 334.
Edwards, S. F., P. R. King, and P. Pincus, 1980, Ferroelectrics **30**, 3.
English, A. E., S. Mafe, J. A. Manzanares, X.-H. Yu, A. Yu. Grosberg, and T. Tanaka, 1996, J. Chem. Phys. **104**, 8713.
Fersht, A. R., 1995, Proc. Natl. Acad. Sci. USA **92**, 10869.
Finkelstein, A. V., 1997, Curr. Opin. Struct. Biol. **7**, 60.
Finkelstein, A. V., A. Ya. Badretdinov, and A. M. Gutin, 1995, Proteins: Struct., Funct., Genet. **23**, 142.
Finkelstein, A. V., A. M. Gutin, and A. Ya. Badretdinov, 1993, FEBS Lett. **325**, 23.
Flory, P., 1953, *Principles of Polymer Chemistry* (Cornell University, Ithaca).
Flory, P., 1969, *Statistical Mechanics of Chain Molecules* (Interscience, New York).
Franz, S., M. Mezard, and G. Parisi, 1994, Int. J. Neural Syst. **3**, 195.
Franz, S., T. Gard, and H. Orland, 1999, Eur. Phys. J. B **11**, 463.
Frauenfelder, H., and P. G. Wolynes, 1994, Phys. Today **47**(3), 58.
Frauenkron, H., U. Bastolla, E. Gerstner, P. Grassberger II, and W. Nadler, 1998, preprint, cond-mat/9806323.
Friedrichs, S., and P. G. Wolynes, 1989, Science **246**, 371.
Garel, T., L. Leibler, and H. Orland, 1994, J. Phys. II **4**, 2139.
Garel, T., and H. Orland, 1988a, Europhys. Lett. **6**, 597.
Garel, T., and H. Orland, 1988b, Europhys. Lett. **6**, 307.
Garel, T., H. Orland, and D. Thirumalai, 1996, in *Recent Developments in Theoretical Studies of Proteins*, edited by Ron Elber (World Scientific, Singapore), p. 197.
Go, N., 1983, Annu. Rev. Biophys. Bioeng. **12**, 183.
Godzik, A., A. Kolinski, and J. Skolnick, 1995, Protein Sci. **10**, 2107.
Goldanskii, V. I., 1989, *Tunneling Phenomena in Chemical Physics* (Gordon and Breach, New York).
Goldenfeld, N., 1992, *Lectures on Phase Transitions and the Renormalization Group* (Addison-Wesley, Reading, MA).
Goldstein, R., Z. A. Luthey-Shulten, and P. G. Wolynes, 1992, Proc. Natl. Acad. Sci. USA **89**, 4918.
Govindarajan, S., and R. A. Goldstein, 1996, Proc. Natl. Acad. Sci. USA **93**, 3341.
Govindarajan, S., and R. A. Goldstein, 1997, Biopolymers **42**, 427.
Grosberg, A., 1984, Biofizika **29**, 621.
Grosberg, A. Yu., 1986, Biofizika **31**, 1045.
Grosberg, A. Yu., and A. R. Khokhlov, 1994, *Statistical Physics of Macromolecules* (AIP, New York).
Grosberg, A. Yu., and D. V. Kuznetsov, 1992, Macromolecules **25**, 1970.
Grosberg, A. Yu., S. K. Nechaev, and E. I. Shakhnovich, 1988, J. Phys. (France) **49**, 2095.
Grosberg, A. Yu., and E. I. Shakhnovich, 1986, Sov. Phys. JETP **64**, 1284.
Guo, Z., and D. Thirumalai, 1996, J. Mol. Biol. **263**, 323.
Gutin, A., V. Abkevich, and E. Shakhnovich, 1995, Proc. Natl. Acad. Sci. USA **92**, 1282.
Gutin, A. M., and E. I. Shakhnovich, 1993, J. Chem. Phys. **98**, 8174.
Gutin, A. M., and E. I. Shakhnovich, 1994, J. Chem. Phys. **100**, 5290.
Haynie, D. T., and E. Freire, 1993, Proteins: Struct., Funct., Genet. **16**, 115.
Hecht, M. H., J. S. Richardson, and D. C. Richardson, 1990, Science **249**, 884.
Higgs, P. G., and J. F. Joanny, 1991, J. Chem. Phys. **94**, 1543.
Hinds, D., and M. Levitt, 1994, J. Mol. Biol. **243**, 668.
Holden, C., 1995, ''Random Samples,'' Science **269**, 1821.
Irback, A., C. Peterson, and F. Potthast, 1996, Proc. Natl. Acad. Sci. USA **93**, 9533.
Kantor, Y., and M. Kardar, 1995a, Phys. Rev. E **51**, 1299.
Kantor, Y., and M. Kardar, 1995b, Phys. Rev. E **52**, 835.
Kantor, Y., H. Li, and M. Kardar, 1994, Phys. Rev. E **49**, 1383.
Kauffman, S,, and A. D. Elington, 1999, Curr. Opin. Chem. Biol. **3**, 256.
Khalatur, P. G., V. A. Ivanov, N. P. Shusharina, and A. R. Khokhlov, 1998, Russ. Chem. Rev. **47**, 855.
Khokhlov, A. R., and P. G. Khalatur, 1998, Physica A **249**, 253.
Kimura, M., 1983, *The Neutral Theory of Molecular Evolution* (Cambridge University Press, New York).
Klimov, D. K., and D. Thirumalai, 1996, Phys. Rev. Lett. **76**, 4070–4073.
Klimov, D. K., and D. Thirumalai, 1998, Folding Des. **3**, 127.
Kolinski, A., A. Godzik, and J. Skolnick, 1993, J. Chem. Phys. **98**, 7420.
Kolinski, A., and J. Skolnick, 1994, Proteins: Struct., Funct., Genet. **18**, 338.
Koukiou, F., 1993, J. Phys. A **26**, L1207.
Kurosky, T., and J. M. Deutsch, 1995, J. Phys. A **28**, 1387.
Kuwajima, K., 1989, Proteins: Struct., Funct., Genet. **6**, 87.
Laidig, K. E., and V. Daggett, 1996, Folding Des. **1**, 335.
Lau, K. F., and K. A. Dill, 1989, Macromolecules **22**, 3986.
Lazar, G. A., J. R. Desjarlais, and T. M. Handel, 1997, Protein Sci. **6**, 1167.
Leopold, P. E., M. Montal, and J. N. Onuchic, 1992, Proc. Natl. Acad. Sci. USA **89**, 8721.
Levinthal, C., 1968, in *Mossbauer Spectroscopy in Biological Systems*, edited by P. Debrunner, J. C. M. Tsibris, and E. Münck (University of Illinois, Urbana).
Levitt, M., M. Gerstein, E. Huang, S. Subbiah, and J. Tsai, 1997, Annu. Rev. Biochem. **66**, 549.
Levitt, M., and A. Warshel, 1975, Nature (London) **253**, 694.
Li, A., and V. Daggett, 1994, Proc. Natl. Acad. Sci. USA **91**, 10430.
Li, H., R. Helling, C. Tang, and N. Wingreen, 1996, Science **273**, 666.
Lifshitz, I. M., 1968, Zh. Eksp. Teor. Fiz. **55**, 2408 [Sov. Phys. JETP **28**, 1280 (1969)].
Lifshitz, I. M., A. Yu. Grosberg, and A. R. Khokhlov, 1978, Rev. Mod. Phys. **50**, 683.
Lumry, R. T., R. Biltonen, and J. F. Brandts, 1966, Biopolymers **4**, 917.
Luthey-Schulten, Z., B. E. Ramirez, and P. G. Wolynes, 1995, J. Phys. Chem. **99**, 2177.
Mezard, M., G. Parisi, and M. Virasoro, 1987, *Spin Glass Theory and Beyond* (World Scientific, Singapore/Teaneck, NJ).
Micheletti, C., J. Banavar, A. Maritan, and F. Seno, 1998, Phys. Rev. Lett. **80**, 5683.
Micheletti, C., F. Seno, A. Maritan, and J. Banavar, 1998, Phys. Rev. Lett. **80**, 2237.

Mirny, L., V. Abkevich, and E. Shakhnovich, 1998, Proc. Natl. Acad. Sci. USA **95**, 4976.
Mirny, L., and E. Domany, 1996, Proteins: Struct., Funct., Genet. **26**, 391.
Mirny, L A., and E. I. Shakhnovich, 1996, J. Mol. Biol. **264**, 1164.
Mirsky, A. E., and L. Pauling, 1936, Proc. Natl. Acad. Sci. USA **22**, 439.
Miyazawa, S., and R. Jernigan, 1985, Macromolecules **18**, 534.
Monod, J., 1971, *Chance and Necessity: an Eassy on the Natural Philosophy of Modern Biology* (Knopf, New York).
Morrissey, M. P., and E. I. Shakhnovich, 1996, Folding Des. **1**, 391.
Nymeyer, H., and J. N. Onuhic, 1997, preprint.
Olszewski, K. A., A. Kolinski, and J. Skolnick, 1996, Proteins: Struct., Funct., Genet. **25**, 286.
Onuchic, J. N., N. S. Socci, Z. Luthey-Schulten, and P. G. Wolynes, 1996, Folding Des. **1**, 441.
Onuchic, J. N., P. G. Wolynes, Z. Luthey-Shulten, and N. D. Socci, 1995, Proc. Natl. Acad. Sci. USA **92**, 3626.
Panchenko, A., Z. Luthey-Schulten, and P. G. Wolynes, 1995, preprint.
Pande, V. S., A. Yu. Grosberg, C. Joerg, and T. Tanaka, 1996, Phys. Rev. Lett. **76**, 3987.
Pande, V. S., A. Yu. Grosberg, M. Kardar, C. Joerg, and T. Tanaka, 1996, Phys. Rev. Lett. **77**, 3565.
Pande, V. S., A. Yu. Grosberg, and T. Tanaka, 1994a, Proc. Natl. Acad. Sci. USA **91**, 12972.
Pande, V. S., A. Yu. Grosberg, and T. Tanaka, 1994b, Proc. Natl. Acad. Sci. USA **91**, 12976.
Pande, V. S., A. Yu. Grosberg, and T. Tanaka, 1994c, J. Phys. II **4**, 1771.
Pande, V. S., A. Yu. Grosberg, and T. Tanaka, 1995a, Phys. Rev. E **51**, 3381.
Pande, V. S., A. Yu. Grosberg, and T. Tanaka, 1995b, Macromolecules **28**, 2218.
Pande, V. S., A. Yu. Grosberg, and T. Tanaka, 1995c, J. Chem. Phys. **103**, 9482.
Pande, V. S., A. Yu. Grosberg, and T. Tanaka, 1997a, J. Chem. Phys. **107**, 5118.
Pande, V. S., A. Yu. Grosberg, and T. Tanaka, 1997b, Folding Des. **2**, 109.
Pande, V. S., A. Yu. Grosberg, and T. Tanaka, 1997c, Biophys. J. **73**, 3192.
Pande, V. S., A. Yu. Grosberg, and T. Tanaka, 1997d.
Pande, V. S., A. Yu. Grosberg, T. Tanaka, and D. S. Rokhsar, 1998, Curr. Opin. Struct. Biol. **8**, 68.
Pande, V. S., C. Joerg, A. Yu. Grosberg, and T. Tanaka, 1994, J. Phys. A **27**, 6231.
Pande, V. S., and D. S. Rokhsar, 1998, Proc. Natl. Acad. Sci. USA **95**, 1490.
Panyukov, S., and Y. Rabin, 1996, Phys. Rep. **269**, 1.
Parisi, G., 1997, in *Physics of Complex Systems: Proceedings of the Enrico Fermi School of Physics, Course CXXXIV*, edited by F. Mallamace and H. E. Stanley (Soc. It. Fisica, Bologna), p. 517.
Park, B. H., and M. Levitt, 1995, J. Mol. Biol. **249**, 493.
Plotkin, S. S., J. Wang, and P. G. Wolynes, 1996, Phys. Rev. E **53**, 6271.
Plotkin, S. S., J. Wang, and P. G. Wolynes, 1997, J. Chem. Phys. **106**, 2932.
Privalov, P. L., 1979, Adv. Protein Chem. **33**, 167.
Ptitsyn, O. B., 1995, Adv. Protein Chem. **47**, 83.
Ptitsyn, O. B., and E. Yu. Eizner, 1965, Biofizika **10**, 3 [Biophysics USSR **10**, No. 1 (1965)].
Ptitsyn, O. B., and V. N. Uversky, 1994, FEBS Lett. **341**, 15.
Ptitsyn, O. B., and M. V. Volkenstein, 1986, J. Biomol. Struct. Dyn. **4**, 137.
Ramakrishnan, R., and J. E. Pekny, 1995, ''An Ergodic Scheme for the Generation of Compact Self-Avoiding Walks on a Lattice,'' preprint.
Ramanathan, S., and E. I. Shakhnovich, 1994, Phys. Rev. E **50**, 3907.
Regan, L., and W. R. DeGrado, 1988, Science **241**, 976.
Rey, A., and J. Skolnick, 1993, Proteins: Struct., Funct., Genet. **16**, 8-28.
Sasai, M., and P. G. Wolynes, 1990, Phys. Rev. Lett. **65**, 2740.
Sauer, R. T., 1996, Folding Des. **1**, R27.
Sear, R., 1997, J. Chem. Phys. **107**, 7477.
Seno, F., M. Vendruscolo, A. Maritan, and J. Banavar, 1996, Phys. Rev. Lett. **77**, 1901.
Sfatos, C., A. Gutin, and E. Shakhnovich, 1993, Phys. Rev. E **48**, 465.
Shakhnovich, E. I., and A. V. Finkelstein, 1989, Biopolymers **28**, 1667.
Shakhnovich, E. I., 1994, Phys. Rev. Lett. **72**, 3907.
Shakhnovich, E. I., 1996, Folding Des. **1**, R50.
Shakhnovich, E. I., 1997, Curr. Opin. Struct. Biol. **7**, 29.
Shakhnovich, E. I., and A. V. Finkelstein, 1982, Dokl. Akad. Nauk SSSR **267**, 1247.
Shakhnovich, E. I., and A. V. Finkelstein, 1989, Biopolymers **28**, 1667.
Shakhnovich, E. I., and A. Gutin, 1989a, Biophys. Chem. **34**, 187.
Shakhnovich, E. I., and A. Gutin, 1989b, J. Phys. A **22**, 1647.
Shakhnovich, E. I., and A. Gutin, 1990, J. Chem. Phys. **93**, 5967.
Shakhnovich, E. I., and A. M. Gutin, 1993, Proc. Natl. Acad. Sci. USA **90**, 7195.
Shortle, D., 1996, FASEB J. **10**, 27.
Socci, N. D., W. S. Bialek, and J. N. Onuchic, 1994, Phys. Rev. E **49**, 3440.
Socci, N. D., and J. N. Onuchic, 1994, J. Chem. Phys. **101**, 1519.
Socci, N. D., and J. N. Onichic, 1995, J. Chem. Phys. **103**, 4732.
Socci, N. D., J. N. Onuchic, and P. G. Wolynes, 1996, J. Chem. Phys. **104**, 5860.
Sun, S., R. Brem, H. S. Chan, and K. A. Dill, 1995, Protein Eng. **8**, 1205.
Tanaka, T., T. Enoki, A. Yu. Grosberg, S. Masamune, T. Oya, Y. Takaoka, K. Tanaka, C. Wang, and G. Wang, 1998, Ber. Bunsenges. Phys. Chem. **102**, 1529.
Tanford, C., 1968, Adv. Protein Chem. **23**, 121.
Thirumalai, D., and Z. Guo, 1995, Biopolymers **35**, 137.
Thomas, P. D., and K. A. Dill, 1996, Proc. Natl. Acad. Sci. USA **93**, 11628.
Tirado-Rives, , J. M. Orozco, and W. L. Jorgensen, 1997, Biochemistry **36**, 7313.
Ueda, Y., H. Taketomi, and N. Go, 1975, Int. J. Pept. Protein Res. **7**, 445.
Uversky, V. N., and O. B. Ptitsyn, 1996, Folding Des. **1**, 117.
Van Gunsteren, W. F., P. H. Hunenberger, H. Kovacs, A. E. Mark, and C. A. Schiffer, 1995, Philos. Trans. R. Soc. London, Ser. B **348**, 49.

Vendruscolo, M., A. Maritan, and J. R. Banavar, 1997, Phys. Rev. Lett. **78**, 3967.
Volkenstein, M. V., 1959, *Configurational Statistics of Polymer Chains* (Moscow) [Interscience, New York, 1963].
Volkenstein, M. V., 1994, *Physical Approaches to Biological Evolution* (Springer Verlag, Berlin/New York).
Witelski, T., A. Grosberg, and T. Tanaka, 1998, J. Chem. Phys. **108**, 9144.
Wittmer, J., A. Johner, and J.-F. Joanny, 1993, Europhys. Lett. **24**, 263.
Wolynes, P. G., 1994, in *Statistical Mechanics, Protein Structure, and Protein-Substrate Interactions*, NATO Advanced Research Workshop, Cargèse, Corsica, France, 1993 (Plenum, New York), p. 26.
Wolynes, P. G., J. N. Onuchic, and D. Thirumalai, 1995, Science **267**, 1619.
Wolynes, P. G., Z. Luthey-Schulten, and J. N. Onuchic, 1996, Chem. Biol. **3**, 425.
Yue, K., and K. A. Dill, 1993, Phys. Rev. E **48**, 2267.
Yue, K., K. M. Fiebig, P. D. Thomas, H. S. Chan, K. A. Dill, and E. I. Shakhnovich, 1995, Proc. Natl. Acad. Sci. USA **92**, 325.
Zhou, Y., C. Hall, and M. Karplus, 1996, Phys. Rev. Lett. **77**, 2822.