This paper was presented at a colloquium entitled "Symmetries Throughout the Sciences," organized by Ernest M. Henley, held May 11–12, 1996, at the National Academy of Sciences in Irvine, CA.[†]

Symmetry and the energy landscapes of biomolecules

PETER G. WOLYNES

School of Chemical Sciences, University of Illinois, Urbana–Champaign, Urbana, IL 61801

ABSTRACT   The role of symmetry in the folding of proteins is discussed using energy landscape theory. An analytical argument shows it is much easier to find sequences with funneled energy landscape capable of fast folding if the structure is symmetric. The analogy with phase transitions of small clusters with magic numbers is discussed.

The role of symmetry in the physics of atoms, nuclei, and elementary particles is different from its role in the biological world. In microphysics, it has been fruitful to postulate symmetry as being fundamental, thereby severely constraining the form of the underlying laws (1). Furthermore, the consequences of symmetry for dynamics and stability are profound in the subatomic world. General treatment of these symmetry effects, based on group theory, allows a nearly complete classification of states and transitions (2, 3). Indeed, even the deviations from exact symmetry have their own quantifiable consequences. In the biological world, symmetry exists too, but it often appears to be an accident (4). While accidents still can have significant consequences, rationalizable on a case-by-case basis, there has been no general unifying theory of symmetry in biology. Exact symmetry in biology would even seem to be antithetical to the notions of complexity, variety, and metamorphosis that are central to the idea of life as we know it. Nevertheless, as in microphysics, life requires stability and sameness as well as change. The apparent conflict of these notions was captured in Schrödinger's metaphor of life as an "aperiodic crystal" (5). Since Schrödinger's time, numerous studies of biological macromolecules have revealed their structure and, thereby, the aptness of his aphorism. X-ray diffraction and nuclear magnetic resonance have led to a view of proteins and nucleic acids as complex three-dimensional structures with some precision (6, 7). At the same time, dynamical studies show that the situation is much more complex than the static models alone suggest (8) and that the proteins fluctuate through a variety of conformational substates whose average is represented by the beautiful pictures now in biology textbooks. This complexity is reflected in the notion of the energy landscape (9).

In this paper, we discuss the role that symmetry plays in the structure and dynamics of biomolecules. Some of this discussion is speculative, but we shall see that the origin and importance of symmetry for biomolecules can be made much clearer in the light of a joint consideration of evolution and biomolecular energy landscapes. We will argue that, while the relationship between symmetry, stability, and dynamics is different for biological physics than for the subatomic world, there are deep similarities.

Symmetry and Evolution

Two extreme points of view can be taken in thinking about biomolecular symmetry from an evolutionary standpoint.

Symmetry of its biomolecules may provide some essential advantage to an organism and thus may reflect some important constraint on function. Alternatively, biomolecular symmetry may be a "frozen accident" of natural history. An example of thinking along the first lines was provided by Jordan (10, 11) long before a structure of any biological macromolecule was known. Jordan argued that heredity required the duplication of a molecular structure and that this duplication could be specifically carried out through the formation of a symmetric dimer in which like parts of a molecule would be attracted to like parts of its dimeric counterpart through a quantum mechanical resonance interaction. The attraction, according to Jordan, arises through the combined thermal and quantum fluctuations of the electronic structure of the molecule. These attractions are stronger if the two molecules have the same excitation spectrum. Despite the beauty of the idea, the analysis was mathematically flawed. Pauling and Delbrück (12) soundly trounced it by showing the weakness of the difference in force and argued instead that heredity and biological specificity in general arose through precise shape complementarity, so that symmetry was not necessary for these functions. The history of structural molecular biology has many times over confirmed the fruitfulness of the Pauling and Delbrück point of view.

Biomolecular symmetry may be advantageous and therefore adaptive on other grounds, however. Duplication of parts from the same instructions means less expenditure of energy and material in carrying forward genetic information. On these grounds, Crick and Watson (13) predicted that the coats of viruses would be made of repeating subunits and thus exhibit the symmetry of closed polyhedra. This beautiful symmetry of viruses was later revealed when their structures were solved by electron microscopy and x-ray diffraction (14). Larger organisms are not so constrained by the cost of passing on genetic information. Indeed, there is a tremendous amount of "noncoding DNA" in eukaryotes (15), so this cannot be the entire explanation for symmetry in more complex living things.

As the study of DNA sequence confirms in the higher organisms, gene duplication is quite common and provides a route to symmetry. A fair fraction of enzymes and other functional molecules are oligomers (4). The most famous example is hemoglobin (see Fig. 1). The tetrameric hemoglobin molecule has a two-fold symmetry, as revealed by x-ray diffraction. The oligomerization allows the possibility of cooperative binding with all of the concomitant functional advantages. The natural historians tell us that this function evolved ≈600 million years ago, and the missing link between the monomeric and oligomeric form can be seen fixed in the genes of the sea lamprey whose hemoglobin subunits are monomeric but associate upon binding oxygen (16). The

[†]The present paper was started as a collaboration between P.G.W. and Hans Frauenfelder. As the paper progressed, it became clear that the crucial ideas were entirely those of P.G.W. Hans Frauenfelder therefore removed his name, but nevertheless presented the paper at the conference that P.G.W. was unable to attend.

FIG. 1.   The hemoglobin molecule is a tetramer $\alpha_2\beta_2$. If the subunits were identical it would have 222 symmetry.

similarity between the α and β subunits is so close that we can be quite confident that this gene duplication occurred as a historical fact. In this scenario, the actual symmetry of the oligomer would seem to be a kind of accident, and the differences between the α and β chains, which prevent a full 222 symmetry from being exact, is a fossilized mark of this history.

It is the inexact symmetries of biological molecules that are most striking. At the tertiary fold level, a single nonoligomeric biomolecule can show symmetry in its overall structure, even though the sequences of the approximate symmetrically related portions of the chains are not the same. An example of this type is shown in Fig. 2. The overall chain traces of the residues from 1 to 88 and from 89 to 174 of the γ crystallin clearly have the same shape, but their sequences are not identical. The symmetry is therefore not exact; furthermore, the similarity in sequence is sufficiently small that we might not realize the two sequences are related by evolution at first. Nevertheless, each section of the molecule is so large that, again, we can be fairly confident after seeing the structure that, in fact, this approximate symmetry arose through gene duplication and fusion followed by genetic drift for a very long period of time.

The most striking examples of approximate biomolecular symmetry arise, however, for smaller proteins that lack any obvious subunit structure or symmetry in sequence. In Fig. 3, we show a four-helix bundle protein, cytochrome b562. There is no obvious exact symmetry in sequence, but the tertiary fold is easily described using the language of symmetry. The molecule is made of four helices of nearly the same length. (One has a poorer hydrogen bonding, so it actually appears to be shorter than it really is.) There is also a twist ultimately due to chirality, which is important, but, ignoring this, it is a symmetric bundle of four cylinders. It is at this crude level of approximation that a symmetry is apparent.

Protein folds generally can be described as belonging to families (48). Many, but not all, of these families exhibit simple approximate symmetries when we idealize them, ignoring the

fine details of the chirality of the helices—e.g., approximate mirror symmetries, like the four-helix bundle, or approximate polyhedral symmetries, like the globins (see Fig. 3). The hemoglobin subunit, when the helices are rendered as cylinders, looks like a polyhedron. Symmetry is not universal, since some families do not exhibit even this approximate symmetry. Why are these approximately symmetric structures so prevalent? Again it is possible that the approximate symmetry is an accident of natural history tracing back to primordial proteins. It has been hypothesized that the original proteins were much shorter than most present day proteins and that the larger proteins were later assembled from these units (17, 18). This attractive hypothesis was made soon after the discovery of split genes or exons in eukaryotes, which could have encoded these folding units. The hypothesis remains a bit controversial, but, if correct, even the symmetry of these smaller proteins would be an accidental result of gene duplication and drift. An alternate view, we pursue below is that this kind of approximate biomolecular symmetry in smaller proteins arises not from duplication and drift but from some kind of convergence in evolution. In this view, the symmetry conferred an advantage even for organisms possessing the primordial proteins and/or symmetry continues to be reinvented during evolutionary history. Under this view, in any event, we must explain why such approximate symmetry of tertiary structure should persist and be prevalent in the absence of obvious symmetry in sequence. An explanation lies in the energy landscape of foldable proteins.

The Energy Landscape of Foldable Proteins and the Minimal Frustration Principle

To function efficiently, most proteins must fold. Much recent theoretical work suggests that achieving a folded structure spontaneously is not trivial (ref. 19; ref. 20 and references therein; refs. 21–23). Foldability requires a specially shaped energy landscape resembling a funnel (20, 21). Not all protein sequences possess an obviously "funneled" landscape and certainly not all sequences have surfaces funneled to the same degree. A heteropolymer with a completely random sequence generically possesses a rugged energy landscape without a funnel, because bringing together two segments of a random sequence will usually give rise to conflicting interactions between the connected side chains, and, therefore, many alternative local minima of the energy can be found. We therefore say a random heteropolymer typically is "frustrated" [a term from solid state physics (24, 25)], because these conflicts make the search through the landscape very difficult. At temperatures low enough to stabilize the ground state conformational dynamics is sluggish or "glassy." States that are geometrically globally different can be nearly ground states and can act as traps slowing the folding process.

Fast-folding proteins can access their thermodynamically stable organized (folded) structure at temperatures where glassy dynamics and traps can be avoided. To do this, the interactions in the folded structure must act in concert more effectively than expected in the most random case. This idea is called the "principle of minimal frustration" (19) and is a generalization of many older notions, such as the idea of Gō (26) that local secondary and nonlocal tertiary interactions must be consistent with each other.

The principle of minimal frustration can be made quantitative using energy landscape theory. Using the simplest approximations of the statistical mechanics of disordered systems, it can be shown to be equivalent to the assertion that the ground state energy of the protein is significantly (i.e., many standard deviations of the energies in the rugged part of the landscape) lower than the bulk of the collapsed states of the protein (27, 28). Such an energy landscape allows folding into the correct state at temperatures where glassy traps are

Fig. 2. γ crystallin is a monomeric protein, but clearly its structure arises from gene duplication and drift. The two modules are similar but not identical in sequence.

unimportant. When the ground state energy is very low, structures similar to the ground state will also be low in energy, thus leading to a funnel structure of the energy landscape as shown in Fig. 4. Many analytical and simulation studies of model heteropolymer problems confirm the outlines of the energy landscape description quickly summarized above. The quantitative form of these ideas can also be used to find and evaluate energy functions for protein structure prediction (27, 28) and to design foldable protein sequences (29) if one is given exact energy functions. The most rapidly foldable proteins have the most pronounced funnel structure to their landscape and are quite stable.

Again we must emphasize that minimal frustration does not mean no frustration. The symmetry and apparent uniqueness of the folded configuration is only approximate, leading to a large number of distinct conformational substates with different reaction characteristics in the overall folded state.

Symmetry and Minimally Frustrated Energy Landscapes

It is natural to suppose that symmetry, minimal frustration, and the funneled nature of the landscape are somehow con-

nected. The consistency of geometry in symmetric structures suggests some preselection and consistency of interaction— i.e., an argument for the validity of the minimal frustration principle (30). This connection is easily understandable when one deals with systems with precisely symmetric interactions between identical particles, such as those in nuclei or atomic clusters. In fact, the notion of funneled landscapes has been taken over to the world of the thermodynamics and kinetics of simple (nonbiological) clusters of atoms (31). For atomic clusters, an especially stable ground state, removed in energy from the bulk of the states allowing a crisp and rapid melting transition, can be achieved only when highly symmetric arrangements of the atoms are possible (32). This property is related to the phenomenon of the "magic numbers" for the stability of van der Waals clusters. Ar_{13}, which is capable of forming a close packed icosahedron, has a funneled landscape easily accessible upon cooling. On the other hand, Ar_8 has a much rougher landscape and the ground state is not easily accessed kinetically. We must unfortunately admit the current lack of a deep understanding or mathematical proof of why symmetric arrangements are always so stable for such systems
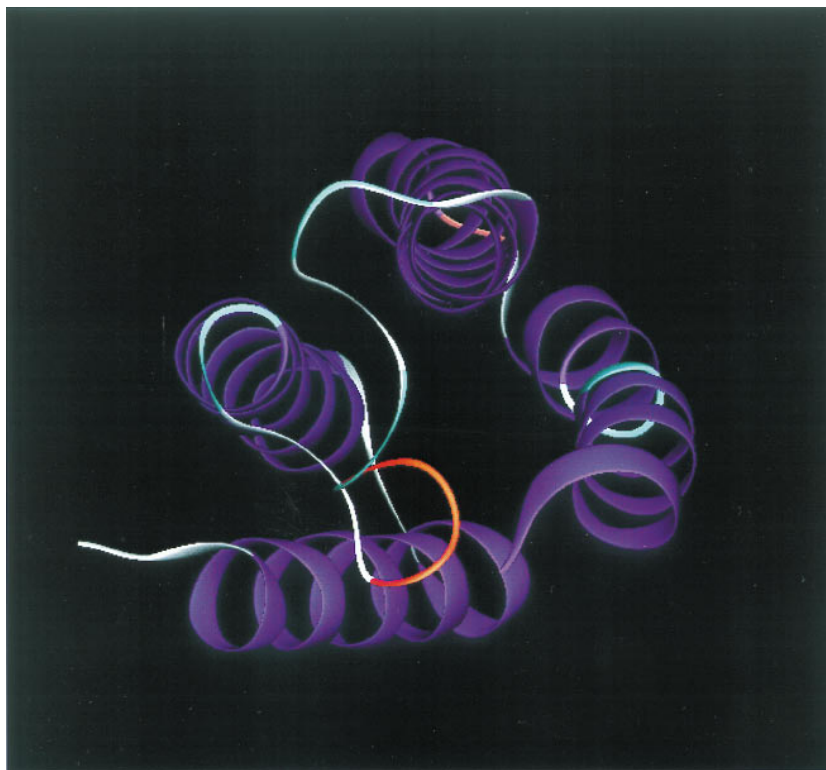
FIG. 3.   A four-helix bundle molecule, cytochrome b562. The ribbon diagram actually exaggerates the dissymmetry because of the demanding algorithm for secondary structure assignment.

with symmetric (and therefore minimally frustrated) interactions. But the observation of extra stability through symmetry is nevertheless not only very plausible for systems of equal particles but also a fact. Admitting this fact, there still remains a significant puzzle for heteropolymers, since the interactions within proteins are far from symmetric. Indeed, an explanation based on the symmetries of a homopolymer seems out of the question on several thermodynamic grounds. Soluble homopolymers of amino acids are either in a random coil state or helical fibers. Simulations of homopolymers on lattices usually give nearly degenerate ground states and insufficient driving force toward the native structure when it is unique. It can also be demonstrated that even for highly complex nonsymmetric structures, there are sequences capable of giving a protein with a funneled landscape that leads to an asymmetric structure but where the frustration has nevertheless been minimized (29).

In answering this last puzzle, the symmetry of heteropolymers, it is important to recall that the observation requiring explanation is not that symmetrical protein folds are universal but that they are prevalent. Several studies suggest all shapes are not equally easy to design by a random process like evolution (33–36). In the context of RNA secondary structure folding where the secondary structural folding problem can be explicitly solved, Schuster *et al.* (34) have shown that some shapes have many sequences that can code for them while others have few (34). In the more difficult computational problem of protein tertiary folding, Yue and Dill (35), using lattice models of proteins with two-letter codes, both hydrophobic and hydrophilic, have shown that symmetric patterns giving nondegenerate ground states are easy to design and have argued that this is the origin of the symmetry prevalence. Govindarajan and Goldstein (36) have not discussed the symmetry issue explicitly but have tried to quantify the statistical distribution of designable proteins using landscape theory and recently lattice models. Using some ideas from statistical mechanics of energy landscapes, we now can make the connection between symmetry and designability of funnel landscapes more precise and explicit.

For concreteness, consider an ensemble of heteropolymers of fixed length N and average composition. For purposes of illustration, the energy function for this heteropolymer can be taken to have the pairwise additive form:

$$E(\{r_i\}) \{a_i\} = \sum v_{ij}(a_i, a_j, (\mathbf{r}_i - \mathbf{r}_j)) + E_{\text{backbone}}(\{r_i\}), \quad [1]$$

where $\{r_i\}$ is the set of polymer subunit locations and $\{a_i\}$ represents the sequence information. The pair potential v_{ij} depends on the distance between and identity of the residues. The potential $E_{\text{backbone}}(\{r_i\})$ maintains chain connectivity and those backbone effects inherent to poly-amino acids that allow α-helices and β-sheets to form. We now consider the designability of a specific structure $\{r_i\}$. If the landscape is to be funneled, this structure must have an energy considerably below that for the typical random structure. The latter quantity, according to the simplest theories, should be a function only of the composition but not the "sequence order." Thus, to ensure a funneled landscape, we must ask that the energy of that configuration with locations of residues at $\{R_i\}$, $E(\{\mathbf{R}_i\})$ $\equiv E_F(\{a_i\})$ must reach some desired low value. The folded energy E_F is a function now only of the sequence $\{a_i\}$. The actual value of the energy difference needed depends on the biological constraint for the time scale of folding. This involves the physics of the stability folding time correlation. Finding the most designable structures by a random process corresponds with finding the structures such that the number of sequences corresponding to them is the largest. This can be thought of equivalently as the entropy of a statistical mechanics problem (29), where the sequence is the variable rather than the locations, $\{\mathbf{R}_i\}$.

Elementary statistical thermodynamics (37) argues that for large enough systems, maximizing the entropy, given an energy constraint, is equivalent to minimizing the free energy for a specific value of a pseudotemperature T^* chosen in such a way that the Boltzmann average energy of the structures at that
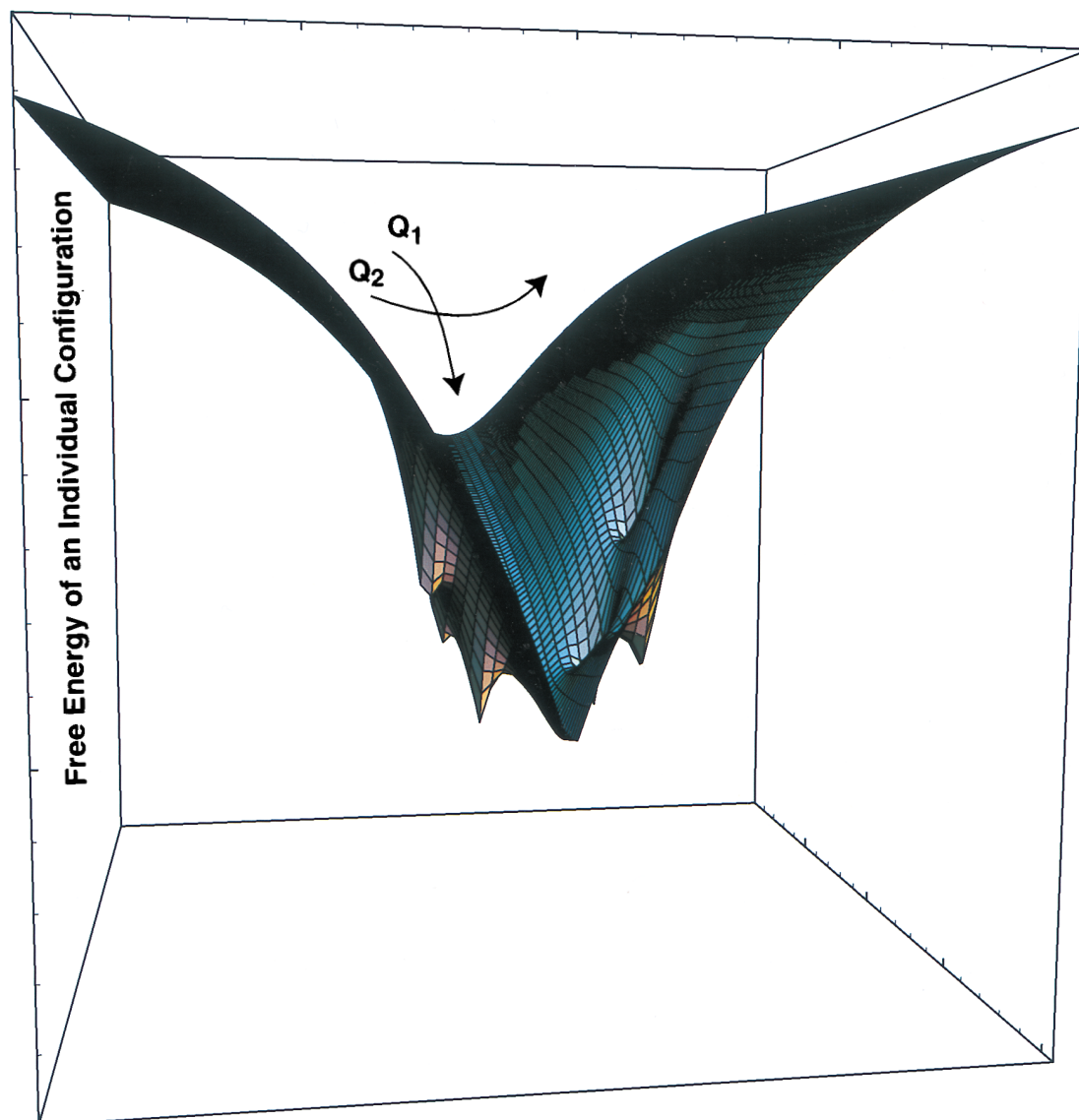
FIG. 4.   A protein folding funnel. The width of the funnel represents configurational entropy. The depth is the free energy of an individual configuration. The funnel is shown as anisotropic. The coordinate Q_1, which destroys the symmetry rapidly, changes the energy, while Q_2, preserving symmetry gives smaller energy changes.

pseudotemperature is equal to that selected energy value. In fact, just such a pseudo-Boltzmann ensemble is the basis of some formal approaches to the statistical mechanics of polymers satisfying the minimal frustration principle (38, 39). To be consistent with a low energy ground state, T^* must be lower than physiological temperature and the glass transition temperature.

For a specific structure, this free energy can be written using a partition sum:

$$Z = \sum_{\{a_i\}} e^{-E(\{R_i\}, \{a_i\})/k_{\mathrm{B}}T^*} = e^{-F(\{r_i\})/k_{\mathrm{B}}T^*}. \qquad [2]$$

This partition sum is an average over possible sequences with a given average composition. Thus we can see that the effective free energy governing the prevalence of sequences for a given structure is a Hamiltonian lacking disorder. The sequence dependence is per force averaged over! The effective free energy function contains only equivalent sites. The energy corresponds to what is known as the "annealed" average of the partition function for the originally disordered system. The connection to symmetry and the simple cluster system can be made especially transparent when the number of types of amino acids is large. In this case, the individual pairs of interactions or bonds are nearly independent and the energy of a given pair and the average over sequence used to compute pseudo-Boltzmann probability for a structure is just a product of the average Boltzmann factors for each pair. This approximation, while exact for large numbers of amino acid types, is not a bad approximation even when there are few. It is equivalent to the quasichemical or Bethe and Kirkwood approach used in the statistical mechanics of magnetic systems (37). The sequence averaged energy function, appropriate for determining the designability of this sequence, is in fact identical with the energy function of an equivalent homopolymer. The interactions of this homopolymer are attractive (and may be much stronger than those presenting any actual homopolymer of amino acids). If the pair interactions were precisely Gaussian in their distribution, they would give a free energy of the form:

$$F(\{\mathbf{r}_i\}) = \sum_{ij} -\langle v_{ij}\rangle_{a_i,a_j} - \frac{1}{2k_{\mathrm{B}}T^*} \langle \delta v_{ij}^2\rangle_{a_i,a_j} + E_{\mathrm{backbone}}(\{\mathbf{r}_i\})$$

$$[3]$$

Here $\delta v_{ij} = v_{ij} - \langle v_{ij} \rangle_{a_i a_j}$ is the fluctuating part of the interaction representing the fact that an individual protein is a heteropolymer. We see the most easily designable structures (i.e., the most numerous in a random sampling with the minimal frustration constraint) are the minimum energy structures for an effectively homoattractive polymer with some specific backbone interactions (the expression in Eq. **3**). Just as for atomic clusters, we would expect the most stable arrangements of the equivalent homopolymer to correspond with symmetrical arrangements. If the backbone interactions were entirely isotropic and very weak, the resulting highly symmetric structures would in fact be the same as those corresponding with the magic numbers of van der Waals clusters.

Two effects will change the preferred symmetries from those of atomic clusters. First, the inadequacy of the independent pair or quasichemical approximation would imply that the actual sequence-averaged free energy would not be additive as in Eq. **3**, but would have nonadditive interactions. Nevertheless, it would still be symmetrical between the particles. The favored symmetries for nonadditive forces may be different from additive ones. This is what occurs for metallic clusters in contrast to van der Waals clusters. More important, however, is that the backbones of proteins favor linear hydrogen bonded arrangements such as α-helices, so one really must consider the most symmetrical arrangements of α helices or β sheets but with homoattractive interactions. Just such a route to understanding the families of protein folds was undertaken by Murzin and Finkelstein (40) a rather long time ago when they discussed the bundle and polyhedral packings of helices, neglecting the heteropolymeric aspect (41). Similar ideas have been used for the homoattractive β-sheets by Chothia *et al.* (42). We see that the validity of those crystallographically inspired approaches can be based on the combination of the minimal frustration principle for the energy landscape providing a force for convergence in evolution with the assumption of otherwise sufficient neutrality and time for drift so that the most common adaptive sequences are those that are observed.

It is amusing that the evolutionarily averaged potential that gives rise to the symmetry, according to this view, is very much analogous to the quantum mechanical van der Waals force. The term $\langle v_{ij} \rangle_{a_i a_j}$ contributes too but is partially canceled by the fact that it is also present in the disordered molten globule, as is a fraction T^*/T of the fluctuation part. Indeed, it is the average interaction over evolutionary fluctuations in sequence as opposed to quantum mechanical or thermal fluctuations in electronic structure that enters the probability of finding an abundant structure. Despite the validity of the Pauling and Delbrück viewpoint for each individual biomolecule, in a sense, a ghost of Jordan's mechanism for biomolecular symmetry resurfaces through an analogy involving the statistical mechanics of evolution and minimally frustrated energy landscapes.

Consequences of Symmetry for the Biomolecular Energy Landscape

For a particular protein, biomolecular symmetry will not be exact unless it is the result of gene duplication. It is a quantitative issue how much the forces leading to minimal frustration, partially determined by symmetry considerations, are dominant over the aspects of the landscape arising from randomness. For protein folding itself, and even more so for function, much needs to be done to address these issues, but we will make here a few observations and speculations.

First, the convergent evolution viewpoint on biomolecular symmetry could be most directly investigated by combinatorial synthesis experiments in which random or partially random proteins are produced through genetic engineering. Two efforts along this line already give some information. Sauer's

group at the Massachusetts Institute of Technology has synthesized random amino acid chains from a small library (43). They showed that a small fraction (0.5%) showed the two-state-like thermodynamics associated with a funneled energy landscape. It is very interesting that of this fraction many of the species were oligomers and, therefore, may be taking advantage of symmetry to fold. On the other hand, Hecht and coworkers (44) has synthesized a random library of proteins whose hydrophobic pattern were forced to obey the exact symmetries of the four-helix bundle, and, in this case, the fraction of sequences with apparently funneled energy landscapes for the monomer is considerably higher. These studies in combinatorial synthesis of proteins, however, point out the complexity in applying quantitatively the statistical mechanical analysis described earlier. Clearly, specific packing forces involving side chains need to be taken into account to achieve fully protein-like behavior.

While the symmetry is approximate for any given native protein, it will contribute to the overall shape of the energy landscape. We might expect there to be a tendency for the symmetry to be preserved in the most important routes to folding and for breaking of symmetry along such a route to be more costly. Again, there is some anecdotal evidence of symmetry preservation along the dominant folding routes when we examine folding of systems such as triosphosphate isomerase. This molecule has an overall three-fold symmetry, but, on the long time scales, it seems to fold in two steps. In the first a two-fold symmetry is preserved and the final third unit is added last (49). The statistical mechanical analysis of folding for symmetrical proteins may also show some correlation of the typical barrier heights with the degree of symmetry of the molecules. The landscape or funnel geometry for a specific minimally frustrated protein will have a strong correlation to the landscape of the sequence averaged effective homopolymer and the energy (see Fig. 4) should vary differently along collective coordinates preserving symmetry from those breaking it.

The individuality of proteins with a common symmetric fold arises from the fluctuations away from the ideal average landscape. It is interesting that the magnitude of these fluctuations is again dependent on the δv^2 interaction that enters the annealed average. For quenched disorder, according to spin glass theory, however, the relevant disorder averaged force is scaled by the actual temperature, not T^*, and is therefore weaker. Different heteropolymers exhibit non-self-averaging behavior when $\delta v^2/k_B T$ is sufficiently strong to cause a glass transition. In the formal statistical mechanics of disordered systems, the glass transition corresponds to "broken replica symmetry" (45). In the language used by the particle physicists, we can restate the minimal frustration principle as "the folding transition that breaks the isotropic symmetry of the globule down to the point symmetry of the folded state occurs at higher temperature than the temperature of broken replica symmetry in order for fast folding to occur."

The arguments presented here have focused on symmetry and the energy landscape involved in folding, but there may be important consequences for function as well as for folding. It is important to remember that Monod, Wyman, and Changeux proposed their two-state model for the allostery of hemoglobin on the basis that the dyadic symmetry of the molecule had to be preserved (46, 47). Since that time, especially in the kinetics, other nonsymmetric states have been invoked and found to be important in the dynamics. We can see the present arguments would, however, be consistent with the Monod, Wyman, and Changeux perspective. Clearly, if one wishes to have an allosteric protein, it must have two different structures. By the same counting arguments that make single symmetric structures more easily designable, pairs of symmetric structures should be an easier design goal if one needs two-state-like switching behavior. It is striking that Eaton has recently

Colloquium Paper: Wolynes

Proc. Natl. Acad. Sci. USA 93 (1996)   14255

described the multistate complex kinetics of hemoglobin allostery using the basic Monod, Wyman, and Changeux symmetry-derived notion of two symmetric *R* and *T* states with a simple superposition of multistate relaxation through an energy landscape superimposed on these two basic structures (E. R. Henry, J. Hofrichter, and W. Eaton, personal communication).

We see that the beautiful symmetries of biomolecules may indeed reflect basic principles about the energy landscape in biology, just as symmetry relations do in particle physics. The extra key, however, is the notion of evolution, which can reconcile the symmetry of a family and the diversity of individuals. Whether the powerful mathematical machinery describing symmetry via group theory used in field theory and condensed matter theory will also help understand the biomolecular energy landscape through the choice of collective coordinates and by describing the nature of defects in biomolecular structure is a question for future development.

Note Added in Proof. Several papers relevant to the theme of this paper have appeared recently. These should be consulted for alternate viewpoints and approaches (50–52).

1. Frauenfelder, H. & Henley, E. (1991) *Subatomic Physics* (Prentice–Hall, Englewood Cliffs, NJ), 2nd. Ed.
2. Wigner, E. (1932) *Gruppentheorie und ihre Anwendung unf die Quantenmechanik der Atomspektren* (Vieweg, Braunschweig, Germany).
3. Gell-Mann, M. & Neeman, Y. (1964) *The Eightfold Way* (Benjamin, New York).
4. Monod, J. (1969) in *Symmetry and Function of Biological Systems at the Macromolecular Level*, eds. Engstrom, A. & Strandbert, B. (Wiley, New York), pp. 15–29.
5. Schrödinger, E. (1949) *What is Life?* (Cambridge Univ. Press, Cambridge, U.K.).
6. Perutz, M. (1992) *Protein Structure: New Approaches to Disease and Therapy* (Freeman, New York).
7. Branden, C. & Tooze, J. (1991) *An Introduction to Protein Structure* (Garland, New York).
8. Frauenfelder, H., Parak, F. & Young, R. D. (1988) *Annu. Rev. Biophys. Biophys. Chem.* **17,** 451–479.
9. Frauenfelder, H. & Wolynes, P. G. (1994) *Phys. Today* **47,** 58–64.
10. Jordan, P. (1938) *Phys. Z.* **39,** 711–714.
11. Jordan, P. (1939) *Z. Phys.* **113,** 431.
12. Pauling, L. & Delbrück, M. (1940) *Science* **92,** 77–79.
13. Crick, F. & Watson, J. D. (1956) *Nature (London)* **177,** 473–475.
14. Caspar, D. L. D. & Klug, A. (1962) *Cold Spring Harbor Symp. Quant. Biol.* **27,** 1–24.
15. Darnell, J., Lodish, H. & Baltimore, D. (1986) *Molecular Cell Biology* (Freeman, New York).
16. Dickerson, R. E. & Geis, I. (1983) *Hemoglobin: Structure, Function, Evolution and Pathology* (Benjamin, Menlo Park, CA).
17. Gilbert, W. (1978) *Nature (London)* **271,** 501.
18. Blake, C. C. F. (1978) *Nature (London)* **273,** 267.
19. Bryngelson, J. D. & Wolynes, P. G. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 7524–7528.
20. Bryngelson, J. D., Onudic, J., Socci, N. & Wolynes, P. (1995) *Proteins Struct. Funct. Genet.* **21,** 167–195.
21. Leopold, P. E., Montal, M. & Onuchic, J. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 8721–8725.
22. Dill, K. A., Bromberg, S., Yue, K., Fiebig, K. M., Yee, D. P., Thomas, P. D. & Chan, H. S. (1995) *Protein Sci.* **4,** 561–602.
23. Shakhnovich, E. & Gutin, A. M. (1993) *Proc. Natl. Acad. Sci. USA* **90,** 7195–7199.
24. Anderson, P. W. (1978) *J. Less Common Met.* **62,** 291–294.
25. Toulouse, G. (1977) *Commun. Phys. (London)* **2,** 115–126.
26. Gō, N. (1983) *Annu. Rev. Biophys. Bioeng.* **12,** 183–210.
27. Goldstein, R. A., Luthey-Schulten, Z. A. & Wolynes, P. G. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 4818–4822.
28. Goldstein, R. A., Luthey-Schulten, Z. A. & Wolynes, P. G. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 9029–9033.
29. Shakhnovich, E. & Gutin, A. (1993) *Protein Eng.* **6,** 793–800.
30. Wolynes, P. G. (1992) in *Spin Glasses and Biology*, ed. Stein, D. (World Scientific, Singapore), pp. 225–259.
31. Ball, K. D., Berry, R. S., Kunz, R. E., Li, F.-Y., Proykova, A. & Wales, D. J. (1996) *Science* **271,** 963–965.
32. Hoare, M. R. & Pal, P. (1975) *Adv. Phys.* **24,** 645–678.
33. Finkelstein, A., Gutin, A. & Badredtinov, A. (1993) *FEBS Lett.* **325,** 23–28.
34. Schuster, P., Fontana, W., Stadler, P. & Hofecker, I. (1994) *Proc. R. Soc. London Ser. B* **255,** 279–284.
35. Yue, K. & Dill, K. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 146–150.
36. Govindarajan, S. & Goldstein, R. A. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 3341–3345.
37. Hill, T. L. (1962) *Introduction to Statistical Thermodynamics,* (Addison–Wesley, Reading, PA)
38. Ranganathan, S. & Shakhnovich, E. (1994) *Phys. Rev. E Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.* **50,** 1303–1312.
39. Pande, V., Grossberg, A. & Tanaka, T. (1994) *J. Phys.* **4,** 1771–1784.
40. Murzin A. & Finkelstein, A. (1988) *J. Mol. Biol.* **204,** 749.
41. Chirgadze, Y. N. (1987) *Acta Crystallogr. A* **43,** 405.
42. Chothia, C., Levitt, M. & Richardson, D. (1977) *Proc. Natl. Acad. Sci. USA* **73,** 3793.
43. Davidson, A. R. & Sauer, R. T. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 2146.
44. Kamtekar, S., Schiffer, J. M., Xiong, H., Babik, J. M. & Hecht, M. H. (1993) *Science* **262,** 1680–1685.
45. Mezard, M., Parisi, G. & Virasoro, M. (1986) *Spin Glass Theory and Beyond* (World Scientific, Singapore).
46. Monod, J., Changeux, J. P. & Jacob, F. (1963) *J. Mol. Biol.* **6,** 306–329.
47. Monod, J., Wyman, J. & Changeaux, J. P. (1965) *J. Mol. Biol.* **12,** 88–118.
48. Richardson, J. (1981) *Adv. Protein Chem.* **34,** 167.
49. Mathews, C. R. (1993) *Annu. Rev. Biochem.* **62,** 653–683.
50. Li, H., Helling, R., Tang, C. & Wingreen, N. (1996) *Science* **273,** 666–669.
51. Lindgard, P. A. & Bohr, H. (1996) *Phys. Rev. Lett.* **77,** 779–782.
52. Kellman, M. (1996) *J. Chem. Phys.* **105,** 2500–2508.