Efficient Rotamer Elimination Applied to Protein Side-Chains and Related Spin Glasses

Robert F. Goldstein
Computer Center and Department of Chemistry, University of Illinois at Chicago, Chicago, Illinois 60612-7352 USA

ABSTRACT Folded proteins and spin glasses share various properties, such as seemingly random interactions between residues (spins), and one might presume that some generic behaviors of spin glasses would also be exhibited in a general way by proteins. But a comparison here shows that the side-chain conformation systems of apo-myoglobin and lysozyme are qualitatively different from specific closely related spin glass systems. This difference is manifest in the number of rotamers that can be identified as definitely not contributing to the global energy minimum. This identification is effected by using a significantly enhanced version of the Dead End Elimination theorem (Desmet, J., M. De Maeyer, B. Hazes, and I. Lasters. 1992. The dead-end elimination theorem and its use in protein side-chain positioning. *Nature.* 356:539–542), which is much more effective and efficient in eliminating rotamers. In several cases (for proteins, although not for spin glasses) this improved Dead End Elimination theorem succeeded in identifying the absolute global minimum of rotamer conformations, with no statistical uncertainty. The difference between protein and spin glass is due to correlations between the interactions of one residue pair with another pair, and probably will play an important role in the thermodynamic behavior of the protein system.

INTRODUCTION

A challenging subproblem of protein structure prediction is the determination of amino acid side-chain positions for a fixed backbone configuration. One greatly simplifying aspect was the observation (McGregor et al., 1987; Janin et al., 1978; James and Sielecki, 1983; Ponder and Richards, 1987) that each side-chain tends to occupy one of a small number of discrete conformations, called rotamers. A more recent simplification was the demonstration by Desmet et al. (1992) that one could eliminate many possible rotamers by applying a "dead-end elimination" (DEE) theorem. In practice, however, the effectiveness depends on specific circumstances, such as protein size and crystal packing. Here, I derive extensions to significantly increase the extent and efficiency of rotamer elimination. In the three cases attempted, this improved DEE theorem was able to identify the absolute global energy minimum with no statistical uncertainty.

Proteins, of course, are not the only systems with a large number of pairwise interactions between subsystems (a prerequisite for application of the DEE theorem). Spin glasses (Fischer and Hertz, 1991) are one example of such a system, and they are particularly interesting because they share some characteristics with proteins, such as many local minima on wildly different energy scales. Often one is interested in the average properties of an ensemble of spin glasses, but in this case it would be interesting to know if some properties of a specific protein side-chain system can be found in a specific, related spin glass.

Received for publication 7 September 1993 and in final form 7 February 1994.

Address reprint requests to Robert F. Goldstein, Computer Center (M/C 135) Room 124, University of Illinois at Chicago, 1940 West Taylor Street, Chicago, IL 60612-7352. Tel.: 312-996-6664; Fax: 312-996-6834; E-mail: u09872@uic.edu.

0006-3495/94/05/1335/06 $2.00

It will turn out that the dead-end elimination theorem is not so effective when applied to a spin glass system whose statistical distributions are identical to those of the side-chains. The interaction energies between a given pair of residues are usually correlated with the interaction energies between a different pair of residues, in contrast to a spin glass. This correlation will have a large effect on the number and nature of the local energy minima, and suggests that a spin glass is qualitatively different from a protein side-chain system.

DEAD END ELIMINATION

Basic algorithm

The energy of a protein with fixed backbone can be written

$$E_f = E_{\text{back}} + \sum_i E(f_i) + \sum_{i<j} E(f_i, f_j), \quad (1)$$

where a set of side-chain configurations is denoted f_i, in which residue i adopts the conformation of rotamer f_i. E_{back} is the energy of the backbone, $E(f_i)$ is the self-energy of side-chain i plus its interaction with the backbone, and $E(f_i, f_j)$ is the interaction energy of side-chains i and j. (If the energy functions were random instead of derived from molecular mechanics, Eq. 1 would describe the potential energy of a spin glass (Fischer and Hertz, 1991), in which each residue with c rotamers is isomorphic to a "spin" with multiplicity c.)

Consider two rotameric states of residue α, namely g_α and h_α, and let $\{f'\}$ denote the conformations of all the side-chains *except* for side-chain α. If every protein state with g_α is higher in energy than the corresponding state with h_α, for all possible configurations of the non-α residues, then

$$E(g_\alpha) + \sum_j E(g_\alpha, f'_j) > E(h_\alpha) + \sum_j E(h_\alpha, f'_j) \quad \forall\{f'\}, \quad (2)$$

and the rotamer g_α *cannot* belong to a first-order (or higher)

local minimum. (In an nth order local minimum, the change of any n or fewer residues must result in higher energy. Thus, the global minimum is an Nth-order local minimum, where N is the total number of residues.)

Equation 2 is not easy to check computationally, but one can compute the weaker condition as a strict bound

$$E(g_\alpha) + \sum_j \min_{f'} E(g_\alpha, f'_j) > E(h_\alpha) + \sum_j \max_{f'} E(h_\alpha, f'_j), \quad (3)$$

which is precisely the DEE theorem (equation 2 from Desmet et al., 1992). If Eq. 3 is true for a specific g_α and h_α, then g_α cannot contribute to *any* local minimum. This does not resolve any difficulties with multiple minima, although it does reduce the combinatorial explosion of rotameric states.

Improved algorithm

Equation 3 can be made more effective by subtracting the right-hand side from the left in Eq. 2 before applying the min operator

$$E(g_\alpha) - E(h_\alpha) + \sum_j \min_{f'}[E(g_\alpha, f'_j) - E(h_\alpha, f'_j)] > 0. \quad (4)$$

In words, Eq. 3 says that a given rotamer of a particular residue (g_α) cannot contribute to a local energy minimum if the smallest possible energy of a conformer that contains g_α is larger than the largest possible energy of a conformer with a different rotamer in the same residue (h_α). This is certainly true, but is overly restrictive. Equation 4, in contrast, says that g_α will not contribute to a local minimum if the energy of a conformation with g_α can always be lowered by just changing g_α to h_α, keeping the other non-α residues frozen. (Strictly speaking, the above sentences should reference the strict energy bounds produced by the max and min operators, rather than the energies themselves, but the effect is conceptually similar.) This large difference in effectiveness between Eqs. 3 and 4 is illustrated in Fig. 1.

Equation 4 would appear to be less efficient because one must compute the min function for each pair of rotamers in a given residue, whereas in Eq. 3 one must compute the max and min functions only for each single rotamer in a residue. But the increased effectiveness of Eq. 4 means that in practice, there are many fewer combinations to consider in later iterations of this process, and this leads to overall greater efficiency.

Going further, Eq. 2 can be generalized to

$$E(g_\alpha) + \sum_j E(g_\alpha, f'_j) > \sum_{r=1}^{R} C_r \left[E(h_\alpha^{(r)}) + \sum_j E(h_\alpha^{(r)}, f'_j) \right] \quad \forall\{f'\}, \quad (5)$$

where the coefficients C_r satisfy $\sum_{r=1}^{R} C_r = 1$ and $C_r \geq 0$, but are otherwise arbitrary. If Eq. 5 is true, the energy can be lowered by changing from g_α to at least one of the $h_\alpha^{(r)}$ conformations, but which one may depend on the conformation of the rest of the protein. (See Fig. 1). A computationally tractable form, similar to Eq. 4, is

$$E(g_\alpha) - \sum_{r=1}^{R} C_r E(h_\alpha^{(r)}) + \sum_j \min_{f'} \left[E(g_\alpha, f'_j) - \sum_{r=1}^{R} C_r E(h_\alpha^{(r)}, f'_j) \right] > 0 \quad (6)$$

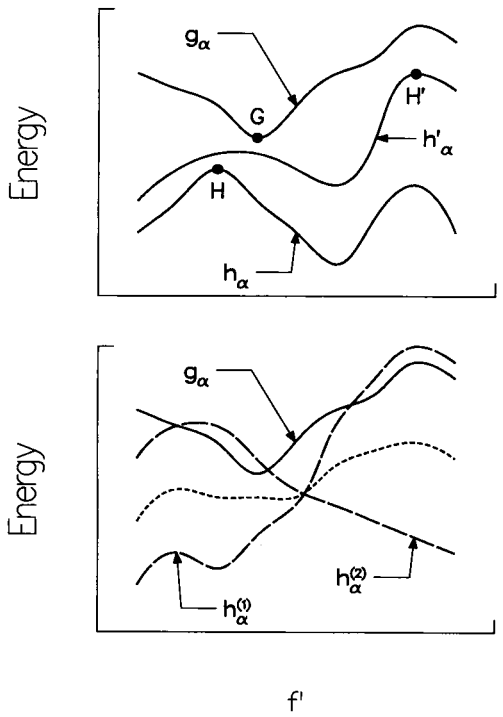

FIGURE 1 Prototypical energies as a function of the protein conformation $\{f'\}$ while the conformation of residue α is held fixed at g_α or some other conformation, as marked. (*Upper panel*) Point G, the lowest energy when residue α is fixed at rotamer g_α, is higher than point H, the highest energy when residue α is fixed at rotamer h_α. Thus g_α is identified as dead-ending relative to h_α by virtue of Eq. 3. Equation 3 would not identify G_α as dead-ending relative to h'_α because H' is higher in energy than G. Equation 4, however, would recognize that the g_α curve is higher everywhere in energy than the corresponding h'_α curve for fixed $\{f'\}$, and would therefore mark g_α as dead-ending. (*Lower panel*) Although there are regions of $\{f'\}$ where g_α is lower in energy than $h_\alpha^{(1)}$ and other regions where g_α is lower than $h_\alpha^{(2)}$ there is no region where g_α is lower in energy than both alternatives. Indeed, g_α is higher everywhere in energy than the average of $h_\alpha^{(1)}$ and $h_\alpha^{(2)}$ (*dotted line*). Thus, g_α would be marked as dead-ending by Eq. 6 but not by Eqs. 3 or 4.

Equation 6 (and Eq. 4 as a special case with $R = 1$) is much more capable of detecting dead-ending rotamers than Eq. 3 (referred to as $R = 0$), although the advantage of $R > 1$ over $R = 1$ is sometimes small in practice.

Desmet et al. (1992) noted that Eq. 3 can be generalized to apply to pairs, or higher order combinations, of residues. These eliminated pairs can then be used to restrict the set of conformations, f', over which the max and min functions are evaluated in Eq. 3, and thereby lead to the elimination of still more rotamers. The DEE theorem with $R \neq 0$ generalizes in the same way. Note, however, that dead-end elimination of pairs may actually eliminate pairs that form part of first order

local minima and, therefore, *can* alter the local minimum structure of the overall potential. In general, one applies the DEE theorem iteratively to all residues until no further rotamers can be eliminated.

A search to eliminate pairs (or higher order combinations) is expensive, and can be sped up by not searching every pair of residues. The residue pairs with large interactions are precisely those that are likely to have rotamer pairs that can be eliminated and will lead to the later elimination of single rotamers. So for efficiency, an interaction threshold can be used to limit the search for dead pairs and triples. Further, the value of R can be reduced for those residues with a large number of rotamers.

The improved dead-end elimination theorem was applied to apo-myoglobin (Mattevi et al., 1991) and lysozyme (Bell et al., 1991) in vacuo, as shown in Fig. 2, and the efficiency of various approximations is shown in Fig. 3. (A comparison of results with crystallographic coordinates is not presented, because the effectiveness of the DEE theorem is in question here, not the accuracy of the underlying potential surface. Also, of course, the DEE theorem is not generally powerful enough to identify uniquely a global minimum by itself.) Clearly the $R \neq 0$ calculation can be more effective than the $R = 0$ case by tens of orders of magnitude. The resulting smaller set of rotamers will then be much more susceptible to attack by other techniques, such as the genetic algorithm (Tuffery et al., 1991).

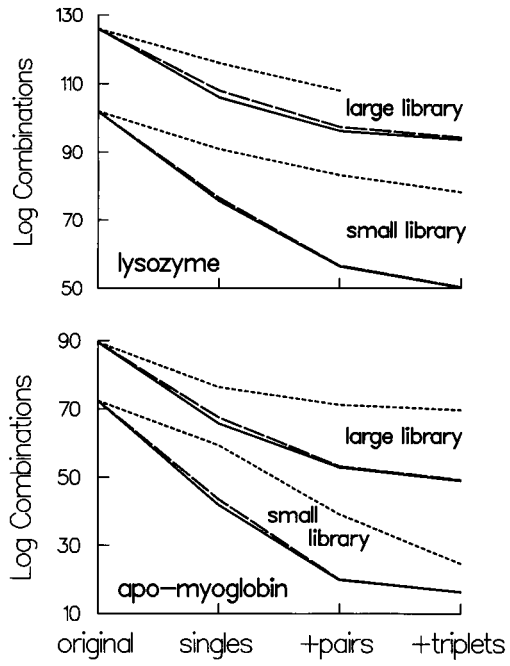

FIGURE 2 Reduction in total rotamer combinations as the DEE theorems are applied to single rotamers ("singles"), to singles plus pairs of rotamers ("+pairs"), and to singles plus pairs plus triplets of rotamers ("+triplets"). The dotted lines represent $R = 0$ calculations (original theorem from Desmet et al., 1992) for all three phases. Dashed ($R = 1$ for singles, $R = 1$ for pairs whose product of rotamers were less than 50, $R = 0$ for other pairs, and $R = 0$ for triplets) and solid (same as dashed, but $R = 2$ for singles) represent improvements presented here. (Raising the $R = 1$ cutoff for pairs, i.e., including more pairs in the $R = 1$ calculation, did not materially affect the effectiveness of rotamer reduction here. It matters greatly, however, for renormalized residues.) Calculations for the upper three curves in each panel used a large rotamer library (Desmet et al., 1992) plus the proline rotamers from Ponder and Richards (1987); lower three curves in each panel used a smaller rotamer library (Tuffery et al., 1991) plus the above-mentioned proline rotamers. Serines were not constrained by any existing di-sulfide bridges. Energy threshold for pairs was 0.1 kcal/mol, and for triplets was 0.3 kcal/mol (see Fig. 3). Backbone coordinates were taken from Brookhaven Protein Data Bank (Bernstein et al., 1977; Abola et al., 1987) entries 7LZM and 5 MBA. The protein interactions were computed with the potential surface from AMBER (Weiner et al., 1984) and OPLS (Jorgensen and Tirado-Rives, 1988) using the MOIL package (Elber et al., 1993), with a cutoff distance of 9 A. Run times (RS/6000 model 320) for computation of the rotamer-rotamer interactions ranged from 3 min for the apo-myoglobin calculation with the small rotamer library to 31 min for the lysozyme calculation with the large rotamer library.

Renormalized residues

A further gain in effectiveness can be obtained by combining two or more residues into one "super-residue." This renormalized residue has the disadvantage of having many more rotameric states (and, therefore, the energy matrix takes up more memory) without changing the overall energetics. Obviously, if each residue has c rotamers, the renormalized residue will have c^2 rotamers. But it has an advantage in that what used to be dead-ending pairs can now be completely eliminated as dead-ending "singles." That is, the c^2 rotamers might be substantially reduced due to dead-ending pairs between the original residues that were known at the time the residues were combined.

More importantly, it means that a DEE pair comparison between the renormalized residue and a third residue is actually a DEE *triple* comparison on the un-renormalized residues; and a DEE pair comparison between two such renormalized residues actually represents an original *quartic* comparison. Thus, the renormalized residues allow the selective application of very high order DEE comparisons, and can thereby identify combinations of residues that contribute to high order minima (but not the global minimum). Another way to say that is the renormalization may turn some second-order minima into first-order minima, and make them easier to identify and eliminate.

Obviously, this technique can be taken to absurd limits. If all residues are combined into one (very large!) super-residue, the quadratic interactions between residues would vanish, and the global minimum would be trivial to find. But that would require storage of something like 10^{100} bytes.

For this renormalization technique to be practical, one must combine a few residues into renormalized residues, apply the DEE theorem through at least pair comparisons, and repeat iteratively. This technique was tried on three of the four systems shown in Fig. 2 (apo-myoglobin with both rotamer libraries and lysozyme with the small library). In all three cases, selective renormalization of residues, coupled with extensive DEE calculations at the pair level with $R = 1$, led to absolute identification of the global energy minimum.

For apo-myoglobin with the small library, the entire process took only a few cpu minutes (see Fig. 4). But for the

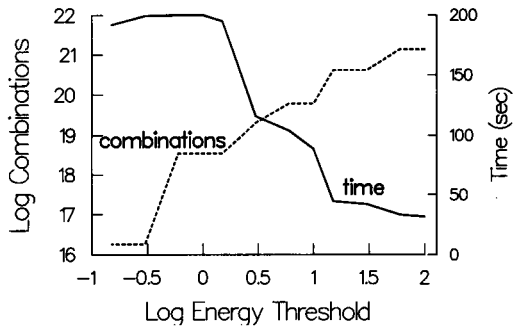

FIGURE 3 Effects of the energy threshold parameter on rotamer elimination effectiveness and run time. The full calculation was run for apo-myoglobin with the smaller rotamer library (lowest *solid line* in second panel of Fig. 2) with differing energy threshold values (the thresholds for pairs are plotted on the x-axis; the values for triples are 3 times those for pairs). That is, the DEE theorem was applied only to those pairs or triples of residues if the maximum minus minimum interaction energy exceeded the threshold. The solid line shows the increase in rotamer elimination as the threshold for checking pairs and triples is decreased. The dotted line shows the corresponding increase in run time (on an IBM RS/6000 model 320). When the threshold was set to zero (not shown on the log plot) so that all pairs and triples of residues were checked, the rotamer elimination was no more effective, but the run time increased to 500 s.

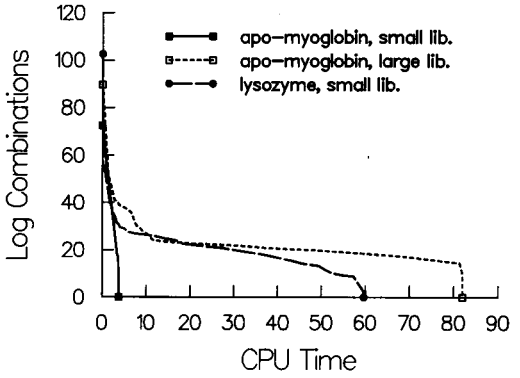

FIGURE 4 The DEE theorem was alternated with residue renormalization, and in the three cases here, the global energy minimum was found. The cpu time (on a small workstation: RS6000 model 320) is in minutes for apo-myoglobin with the small rotamer library; it is in hours for the other two cases.

other two cases, the computation took 60–80 cpu hours, not counting several tens of hours for trial and error in deciding which residues to renormalize. The inital reduction of rotamer combinations to about 10^{30} ocurred quickly, but it was followed by roughly 30 iterations of very slow improvement to about about 10^{15} combinations, and finally finished with a quick descent to the global minimum through a few fast iterations. The upshot is that the renormalization technique does work, and in favorable circumstances, can identify the global minimum out of over 10^{100} possibilities. But full convergence is not guaranteed.

(The final protein structures are not discussed here, because the central point is the algorithm by which the global minimum was found, rather than the exact force field that governs the nature of the minimum. When statistical techniques, such as Monte Carlo, are used, it is quite reasonable to compare the best minimum found with a crystal structure. But in this case, there is no doubt that the global mimimum was found *for the particular force field used*, because the algorithm is deterministic. Of course, it would be quite surprising if this mimimum did *not* differ from the crystal structure, because the force field was applied in vacuo.)

Two aspects of the renormalization technique seem particularly important for its success. First, even when the renormalized residues become large, the $R = 1$ DEE theorem for pairs must be applied. This is computationally very expensive, but the extra effectiveness seems crucial.

Second, it is very important to pick "good" residues to renormalize. Unfortunately, it is not obvious what makes a pair of residues "good," except in hindsight. I had good success picking pairs of residues that had a large fraction of their rotamer combinations identified as dead, even after all single rotamers that could be eliminated were elminated. The idea was that these particular residue pairs represent strongly interacting residues, and that once renormalized many "single" rotamers of this new residue would be eliminated easily with the concomitant savings in computer storage.

Nevertheless, it was important not to pick too many pairs to renormalize at one time. Generally, I picked no more than four residues to renormalize into two new residues, even if more than four good candidates were apparant. Very often, the nature of the renormalization candidates changed radically after application of the DEE theorems, and the wrong choice of residues to renormalize could expand memory requirements to the point of impracticality. It was much better to renormalize slowly and to compute many interations of the DEE theorem, than to quickly reduce the total combinations by 10 orders of magnitude and then fail to converge further.

SPIN GLASS COMPARISON

The DEE theorem is not limited to proteins. Equation 1 resembles the energy of a spin glass with completely anisotropic interactions between pairs of spins (rotamer-rotamer interactions) in a random external field (rotamer-backbone interactions). Because spin glass models have been used to model generic protein folding (Byrngelson and Wolynes, 1990), one might ask if the physics of a real protein side-chain system is similar to that of a specific idealized spin glass.

Interactions between spins in a idealized spin glass are drawn from a random distribution, although the properties of that distribution may depend on the distance between spins. What that distribution should be for protein side-chains is unclear; however, one normally assumes that the distribution for one spin or pair of spins is *uncorrelated* with other distributions for other spins or pairs.

Thus, one can test the "glassiness" of the system by randomizing the interactions between each pair of residues, and between each residue and the backbone, without changing any of the underlying interaction magnitudes. That is, consider the interaction energies between two residues. If each

residue has c rotamers, then the interactions are described by a $c \times c$ matrix. This matrix can describe the interaction of two "spins" in a spin glass, with the elements drawn at random from some underlying distribution. But whatever the underlying distribution is, it is clearly unchanged by simply randomizing the *order* of elements in the $c \times c$ interaction matrix. Only the correlation between one matrix and another is destroyed, and an uncorrelated spin glass is created.

The DEE theorem was applied to many such "randomly phased proteins," shown in Fig. 5. The reduction in rotameric combinations was much larger for the native protein than for any randomly phased construct. This may mean that there are more low-order (less cooperative) local minima in a native protein than in a related spin glass. If so, then the transition from molten globule to folded protein might be accomplished much more easily than the analogous process of hopping from one local minimum to another (with the attendent possible trapping) in a spin glass. Clearly, the correlated nature of side-chain interactions will need to be accounted for in any complete thermodynamic or kinetic treatment of such systems.

A more detailed comparison with spin glasses would be fruitful, but is beyond the scope of this paper. The generalized random energy model (Derrida and Gardner, 1986a, b) can represent at least some of the correlations seen here in the side chain systems, and Derrida (1987) has indicated how one might search for phase transitions in nonsymmetric systems similar to side chains. But these analyses depend on knowledge of the underlying random distribution of spin interactions, and it is not yet clear what that distribution is, although it is clear in the present work that that distribution is strongly correlated.

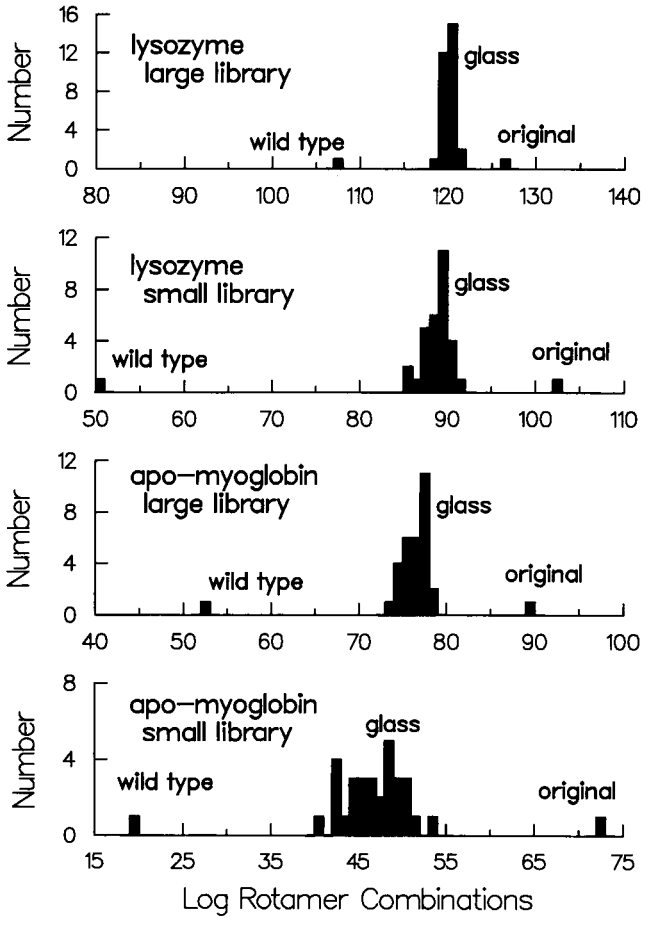

FIGURE 5 The DEE theorem was applied to 30 randomized structures for each protein/rotamer library combination, and the histograms of the resulting numbers of rotamer combinations are presented here. (The lysozyme large library calculation included only single residues at $R = 1$. The apomyoglobin large library calculation also included $R = 1$ calculations for pairs whose product of rotamers were less than 50, and $R = 0$ calculations for other pairs. And the small library calculations included, in addition, calculations of triples at $R = 0$.)

CONCLUSIONS

It is possible, by application of the dead-end elimination theorem, to identify rotamers that cannot contribute to local energy minima of a certain or higher order. If enough rotamers can be eliminated by recursive application, the global minimum can be found; in fact, this was established in three example cases.

The DEE theorem applies to force fields in which the protein side-chains interact in a pair-wise additive fashion. Although the examples considered here were in vacuo calculations, it would be easy to include solvent effects either at the one-residue level (a "wetting" term could describe a single side-chain that sticks into the solvent) or the two-residue level (one residue polarizes the solvent, and a second residue interacts with that polarization). Only true three-body terms (polarization of the solvent might depend on the conformation of a third residue, for example) could not be included without modifying the DEE equations and significantly increasing the computer requriements for successful application.

The DEE theorem does not work well for spin glasses that are closely related to protein side-chain systems. Apparently, in a native protein the interactions between residues A and B are correlated with the interactions between residues B and C in a way that is atypical of most generic spin glasses. The protein correlations may well result from the three-dimensional nature of the protein interactions. That is, residue B may well have large steric interactions with residues A and C, but there is probably no single conformation of B that overlaps strongly with both A and C. But this type of correlation is generally not present in a random spin glass.

The correlations that do exist for native proteins, at least when the protein backbone is fixed in the crystal structure coordinates, seem to make the DEE theorem very effective. They will drastically affect the ruggedness of the potential energy "landscape," and they may well affect the overall folding pathway of the protein.

I thank Dr. Johan Desmet for a copy of the rotamer library used in Desmet et al. (1992).

REFERENCES

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. F. H. Allen, G. Bergerhoff, and R. Sievers, editors. Data Commission of the International Union of Crystallography. 107–132.

Bell, J. A., K. Wilson, X.-J. Zhang, H. R. Faber, H. Nicholson, and B. W. Matthews. 1991. Comparison of the crystal structure of bacteriophage T4 lysozyme at low, medium, and high ionic strengths. *Proteins Struct. Funct. Genet.* 10:10–21.

Bernstein, F. C., T. F. Koetzle, G. J. B. Williams, Jr., E. F. Meyer, M. D. Brice, J. R. Rodgers, O. Kennard, T. Shimanouchi, and M. Tasumi. 1977. The protein data bank: a computer-based archival file for macromolecular structures. *J. Mol. Biol.* 112:535–542.

Byrngelson, J. D., and P. G. Wolynes. 1990. A simple statistical field theory of heteropolymer collapse with application to protein folding. *Biopolymers.* 30:177–188.

Derrida, B., and E. Gardner. 1986. Magnetic properties and the function q(x) of the generalised random-energy model. *J. Phys. C: Solid State Phys.* 19:5783–5798.

Derrida, B., and E. Gardner. 1986. Solution of the generalised random energy model. *J. Phys. C: Solid State Phys.* 19:2253–2274.

Derrida, B. 1987. Dynamical phase transition in non-symmetric spin glasses. *J. Phys. A: Math. Gen.* 20:L721–L725.

Desmet, J., M. De Maeyer, B. Hazes, and I. Lasters. 1992. The dead-end elimination theorem and its use in protein side-chain positioning. *Nature.* 356:539–542.

Elber, R., A. Roitberg, C. Simmerling, R. F. Goldstein, G. Verkhivker, H. Li, and A. Ulitsky. 1993. MOIL: a molecular dynamics program with emphasis on conformational searches and reaction path calculations. *In* Statistical Mechanics, Protein Structure and Protein Substrate Interactions. S. Doniac, editor. Plenum Press, New York.

Fischer, K. H., and J. A. Hertz. 1991. Spin Glasses. Cambridge University Press, Cambridge. 397 pp.

James, M. N. G., and A. R. Sielecki. 1983. Structure and refinement of Penicillopepsin at 1.8 Å resolution. *J. Mol. Biol.* 163:299–361.

Janin, J., S. Wodak, M. Levitt, and B. Maigret. 1978. Conformation of amino acid side-chains in proteins. *J. Mol. Biol.* 125:357–386.

Jorgensen, W. L., and J. Tirado-Rives. 1988. The OPLS potential functions for proteins. Energy minimizations for crystals of cyclic peptides and crambin. *J. Am. Chem. Soc.* 110:1657*a*. (Abstr.)

Mattevi, A., G. Gatti, A. Coda, M. Rizzi, P. Ascenzi, M. Brunori, and M. Bolognesi. 1991. Binding mode of azide to ferric Aplysia Limacina myoglobin. Crystallographic analysis at 1–9 Angstroms resolution. *J. Mol. Recog.* 4:1*a*. (Abstr.)

McGregor, M. J., S. A. Islam, and M. J. E. Sternberg. 1987. Anaylsis of the relationship between side-chain conformation and secondary structure in globular proteins. *J. Mol. Biol.* 198:295–310.

Ponder, J. W., and F. M. Richards. 1987. Tertiary templates for proteins: use of packing criteria in the enumeration of allowed sequences for different structural classes. *J. Mol. Biol.* 193:775–791.

Tuffery, P., C. Etchebest, S. Hazout, and R. Lavery. 1991. A new approach to the rapid determination of protein side chain conformations. *J. Biomol. Struct. Dyn.* 8:1267–1289.

Weiner, S. J., P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, and P. Weiner. 1984. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* 106:765–784.