Free energy approximations in simple lattice proteins

Dirk Reith[a)]
Max-Planck-Institut für Polymerforschung, D-55128 Mainz, Germany and Research School of Chemistry, Australian National University, Canberra ACT 0200, Australia

Thomas Huber[b)]
ANU Supercomputer Facility, Australian National University, Canberra ACT 0200, Australia

Florian Müller-Plathe[c)]
Max-Planck-Institut für Polymerforschung, D-55128 Mainz, Germany

Andrew E. Torda[d)]
Research School of Chemistry, Australian National University, Canberra ACT 0200, Australia

(Received 6 September 2000; accepted 2 January 2001)

This work addresses the question of whether it is possible to define simple pairwise interaction terms to approximate free energies of proteins or polymers. Rather than ask how reliable a potential of mean force is, one can ask how reliable it could possibly be. In a two-dimensional, infinite lattice model system one can calculate exact free energies by exhaustive enumeration. A series of approximations were fitted to exact results to assess the feasibility and utility of pairwise free energy terms. Approximating the true free energy with pairwise interactions gives a poor fit with little transferability between systems of different size. Adding extra artificial terms to the approximation yields better fits, but does not improve the ability to generalize from one system size to another. Furthermore, one cannot distinguish folding from nonfolding sequences via the approximated free energies. Most usefully, the methodology shows how one can assess the utility of various terms in lattice protein/polymer models. © *2001 American Institute of Physics.* [DOI: 10.1063/1.1350575]

I. INTRODUCTION

Nature reflects many-body interactions and ensemble averages, but practical calculations usually rely on pairwise interactions and cheap approximations. For example, one is often interested in properties such as protein or polymer stability or phase behavior. These are related to free energy, but this is barely accessible, even with much patience and long simulations. This has led to the use of potentials of mean force and debates as to how they can be formulated. In this work, we consider a different problem. Given a model system, how good can a potential of mean force possibly be? Since one never knows if a better approximation can be found, can one construct machinery to test the approximation to free energy?

This subject has become particularly topical in protein modeling where there is an abundance of potentials of mean force calculated from archived protein structures.[1,2] On one side, these have been interpreted as providing realistic Helmholtz free energies.[3] On the other side, it has been stated that the resulting quantities have no bearing on properties such as stability.[4] It has also been noted that accurate free energies cannot be correctly extracted from collections of protein structures.[5]

That debate is centered on the question of whether a disparate collection of protein structures is a statistical mechanical ensemble or, at least an approximation to one. This work addresses a fundamentally different question. If we are not limited to a construction based on statistical mechanics, can any arbitrary pairwise function be fit to reproduce free energies? In other words, what are the limitations of pairwise approximations, without entering the debate over the applicability of Boltzmann-based force fields?

This can be addressed by taking a system whose Hamiltonian consists of pairwise interactions and seeing whether the true free energy (a property of the ensemble) can be meaningfully approximated by a single structure (the ground state) using an appropriately reparametrized Hamiltonian. The reparametrization can be done by fitting, rather than assuming Boltzmann statistics. For simple functions, regression methods guarantee a best fit and, by construction, answer the question of just how good a potential of mean force can be.

This strategy requires a system where one knows the free energy exactly, but this can be found for simple models such as a two-dimensional lattice polymer or protein. There, one can define the Hamiltonian (by stating the potential energy) and calculate the free energy by exhaustive enumeration (visiting every possible conformation). Doing this for every sequence of a given length provides all the information necessary for the fitting operation.

Two-dimensional lattice macromolecules may not be of great practical use, but the machinery can be used to measure certain properties and test models. For example, potentials of mean force sometimes show an artifactual size dependence. When systems are small enough to permit exhaustive enu-

[a)]Electronic mail: reith@mpip-mainz.mpg.de
[b)]Electronic mail: Thomas.Huber@anu.edu.au
[c)]Electronic mail: mplathe@mpip-mainz.mpg.de
[d)]Electronic mail: Andrew.Torda@anu.edu.au

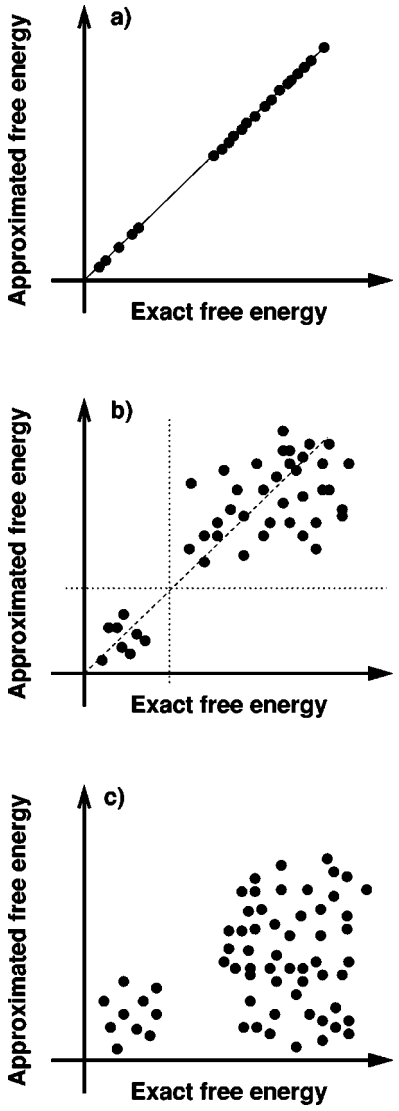

FIG. 1. Possible scenarios for the quality of free energy approximation functions.

meration, one can ascertain the extent to which this is a methodological problem and to what extent it is a fundamental limitation on the potential of mean force. Perhaps more interesting is the ability to examine different lattice models with different levels of sophistication. In the simplest case one could define the potential energy of a two-dimensional lattice protein or polymer with a classical hydrophobic-polar (HP) model.[6] This has only three interaction parameters (ϵ_{HH}, ϵ_{HP}, and ϵ_{PP}) which might be set at $(-1, 0, 0)$, respectively.[6a] The first step would be to see if some new set of (ϵ^*_{HH}, ϵ^*_{HP}, and ϵ^*_{PP}) could be found so as to reproduce the exact free energies. Next, one could try additional terms. For example, monomers could interact with empty sites on the lattice, reminiscent of lattice-solvent models.[7] This would introduce additional solvent interaction parameters (κ^*_H and κ^*_P) to be fit simultaneously. There is no rigorous basis for the additional terms since they do not exist in the original potential energy, but they are typical of the kinds of elaboration added to simple models. They may even be no more than pragmatic tools for the approximation of free energies. One can continue in this vein and add other terms to try to capture the entropic contributions. For example, one could base a term on second-nearest neighbors on the lattice. Results from such calculations are given in the next section.

The calculations below are simplified by considering a subset of the fitting problem particularly relevant to proteins. Most protein sequences do not fold to a specific structure and are not seen in nature. The set of viable ''folding'' protein sequences is actually tiny compared to the set of possible protein sequences.[8] As a consequence, a real potential of mean force need not even work for all possible sequences. In the fitting calculations, we do not even attempt to fit to all possible protein sequences, but instead look for an approximation which works for ''folding'' sequences. In lattice models it is particularly easy to identify the folding sequences. These are normally taken to be those with a single lowest energy conformation[9] since degeneracy of ground states would correspond to a protein which moved between equally attractive conformations rather than a single stable structure.

At the risk of overinterpreting some very simple models, one can pursue the concept of folding sequences and ask whether a free energy approximation based on folding sequences can be used to recognize the difference between folding and nonfolding sequences. This is an important issue in protein and sometimes polymer design where one does not merely want a low energy sequence-structure pair, but rather one wants a sequence which folds to a single stable conformation. Viewed schematically, we would like to know if a free energy approximation works as shown in Fig. 1(a) where there is excellent agreement between predicted and observed free energies. With a less predictive fit, one may see the situation of Fig. 1(b). The approximation, being based on the folding sequences only, is not good at predicting free energies for nonfolding sequences with their degenerate ground states, but it is perfectly adequate at distinguishing folding from nonfolding sequences. In the last case, one may see the situation of Fig. 1(c). Even though there is a correlation between predicted and observed free energies, there is such overlap between predicted values for folding and nonfolding sequences that the function would not be useful for protein design.

In the following section, there are a series of calculations demonstrating issues which can be addressed. The calculations begin with an HP-like model[6] of a two-dimensional square lattice protein/heteropolymer where it is possible to visit every conformation and every possible sequence up to length 16. Since this model suffers from frequent degeneracy of ground states,[8] similar calculations were also performed using a potential energy matrix designed to alleviate degeneracy while preserving the property that similar monomers tend to aggregate.[9] Because of the artificiality of the model, comparisons were also done with a relaxed definition of degeneracy. Finally, examples of artificial free energy approximations (solvent interaction and second-nearest neighbor terms) were tested.

TABLE I. Interaction matrices and values of ϵ_{ij} interaction parameters.

2 monomer types					
	HP			HA	
	H	P		H	P
H	−1	0	H	−2.3	−1
P	0	0	P	−1	0

II. MODEL

Chains consisted of a set $\{\sigma_i\}$ of connected monomers of type σ at position i. The potential energy, E, depended on the corresponding set of coordinates $\{\mathbf{r}_i\}$ as well as sequence and was given by

$$E(\{\sigma_i\},\{\mathbf{r}_i\})=\sum_{i<j}\ \epsilon_{\sigma_i\sigma_j}\Delta(\mathbf{r}_i,\mathbf{r}_j), \tag{1}$$

where the summation runs over all i, j, pairs and $\Delta(\mathbf{r}_i,\mathbf{r}_j)$ is a switching function dependent on the coordinates of the interacting particles. It is equal to 0 for most pairs or 1 if the coordinates are adjacent in space ($|\mathbf{r}_i-\mathbf{r}_j|=1$), but not adjacent in the sequence ($|i-j|>1$). $\epsilon_{\sigma_i\sigma_j}$ gave the strength of the interaction type. σ adopted one of two types as given by an interaction matrix from Table I. The two-monomer-type matrices were the classic HP model[6] and one labeled HA,[9] loosely based on real protein statistics.[2] The labels are only a notational convenience and may not be the same as those of the original authors.

For each sequence $\{\sigma_i\}$, the potential energy of every conformation $\{\mathbf{r}_i\}_m$ was calculated [Eq. (1)] and the conformational integral given by

$$Z_{\{\sigma_i\}}=\sum_m\ \exp\left(\frac{-E(\{\sigma_i\},\{\mathbf{r}_i\}_m)}{kT}\right), \tag{2}$$

where the summation runs over every conformation m and the temperature T was set to 0.1 unless otherwise stated. The exact free energy F was given by

$$F_{\{\sigma_i\}}=-kT\ln(Z_{\{\sigma_i\}}). \tag{3}$$

In all cases, we work in reduced units with Boltzmann's constant $k=1$ and the potential energies were offset so the lowest energy state for a sequence was zero. Note that the free energy is labeled as a property of the sequence $\{\sigma_i\}$.

For the parameter fitting, an approximate free energy F^* was defined,

$$F^*(\{\mathbf{r}_i\}^0,\{\epsilon^*\})=\sum_{i<j}\ \epsilon^*_{\sigma_i\sigma_j}\Delta(\mathbf{r}_i,\mathbf{r}_j). \tag{4}$$

Because this is an approximation, we write it explicitly as a function of the set of fit parameters, $\{\epsilon^*\}$ and coordinates of the ground state, $\{\mathbf{r}_i\}^0$. Since $F^*(\{\mathbf{r}_i\}^0,\{\epsilon^*\})$ is linear in the interaction parameters, fitting was done with general least-squares regression[10] to minimize an error given by

$$\text{error}=\sum_p\ (F_{\{\sigma_i\}_p}-F^*(\{\mathbf{r}_i\}^0,\{\epsilon^*\}))^2, \tag{5}$$

TABLE II. Statistics of problem size. N_{seq} is the number of sequences possible for two or three monomer types, N_{conf} is the number of possible conformations, and N_{fold} the number of ''folding'' conformations. Values are given for the interaction matrices as named in the text and with a definition of folding as the number of ground states (N_{ground}) of one or two.

			N_{fold}			
			$N_{\text{ground}}=1$		$N_{\text{ground}}=2$	
Length	N_{seq}	N_{conf}	HP	HA	HP	HA
6	64	36	7	22	11	52
8	256	272	7	35	45	147
10	1024	3034	6	104	114	480
12	4096	15 037	87	782	417	2362
14	16 384	110 188	386	2770	1598	8924
16	65 536	802 075	1539	12 252	7255	37 980

where the summation runs all over p (folding) sequences of a specific length. More generally, the approximate free energy $F^*(\{\mathbf{r}_i\}^0,\{\epsilon^*\})$ could be redefined to include extra arbitrary terms as described below.

For the fits including a ''solvation'' term, a free energy approximation was defined by

$$F^*(\{\mathbf{r}_i\}^0,\{\epsilon^*\},\{\kappa^*\})=\sum_{i<j}\ \epsilon^*_{\sigma_i\sigma_j}\Delta(\mathbf{r}_i,\mathbf{r}_j)+\sum_i\ \kappa^*_{\sigma_i}s(\mathbf{r}_i), \tag{6}$$

where $s(\mathbf{r}_i)$ is the number of empty lattice sites adjacent to the position $\mathbf{r}_i$ and $\kappa^*_{\sigma_i}$ is an adjustable parameter indexed by the type of monomer i.

For the approximations using second nearest pairs, we define

$$F^*(\{\mathbf{r}_i\}^0,\{\epsilon^*\},\{\rho^*\})$$

$$=\sum_{i<j}\ \epsilon^*_{\sigma_i\sigma_j}\Delta(\mathbf{r}_i,\mathbf{r}_j)+\sum_{i<j}\ \rho^*_{\sigma_i\sigma_j}u(\mathbf{r}_i,\mathbf{r}_j), \tag{7}$$

where $u(\mathbf{r}_i,\mathbf{r}_j)$ is usually 0, but set to 1 if the particles are diagonal neighbors (distance $|\mathbf{r}_i-\mathbf{r}_j|=\sqrt{2}$) and separated by more than one residue in the chain. Finally, one can add all the extra terms and consider an approximation based on both first- and second-nearest neighbor terms and solvation. One will then end up with a $(3+3+2)=8$-parameter space in the case of any two-monomer model:

$$F^*(\{\mathbf{r}_i\}^0,\{\epsilon^*\},\{\kappa^*\})=\sum_{i<j}\ \epsilon^*_{\sigma_i\sigma_j}\Delta(\mathbf{r}_i,\mathbf{r}_j)+\sum_{i<j}\ \rho^*_{\sigma_i\sigma_j}u(\mathbf{r}_i,\mathbf{r}_j)$$

$$+\sum_i\ \kappa^*_{\sigma_i}s(\mathbf{r}_i). \tag{8}$$

III. RESULTS

In all cases, parameters were fit considering all folding sequences. For checking if the fits generalized to nonfolding sequences, statistics were collected from a sampling of the order of 10^3 sequences chosen randomly from the 10^5 to 10^7 possibilities.

A. Simple approximations

Table II shows the sizes and properties of the systems examined. Extensive calculations were carried out on the HP

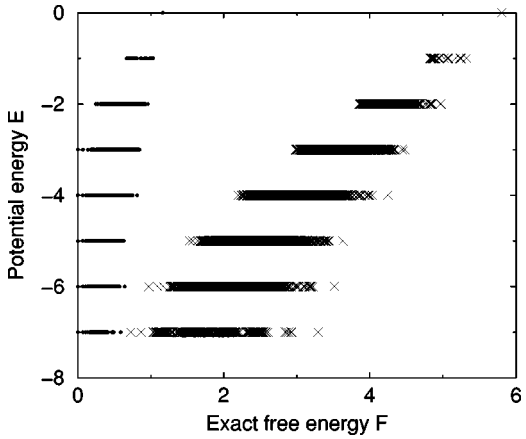

FIG. 2. Comparison of free and potential energies for the HP model, $N=16$ at temperatures $T=0.1$ (dots) and $T=0.5$ (crosses).

and HA model for chains up to length $N=16$. Initial conditions were determined by examining the HP model. Before considering any approximations, the simplest system can be used to show the effect of temperature and the role of entropy. Figure 2 shows a comparison of free and ground state potential energies for the HP model, $N=16$ at two different temperatures. Obviously, the free energy is correlated with the potential energy, but the fit to a line becomes worse as the temperature increases and the entropic contribution to F increases. The temperature $T=0.1$ was used in all subsequent calculations since it includes some entropic contribution and has been shown to be below the temperature of phase transition for this kind of system. The figure shows another limitation of this model. Aside from the well-known, high ground state degeneracy of the HP model,[8] the plot shows that this simple lattice system has very few energy levels. This means that potential energies are restricted to a relatively small number of discrete values, but that free energies will span something closer to a continuum. Any attempt to approximate or predict free energies can also be seen as an attempt to account for this entropic spread of free energies.

One can now apply the fitting procedure to the original functional form and show how the simple functional form with the original number of parameters can be adjusted to reproduce free energies (bring the points closer to a straight line). As an example, Table III shows the fit quality and resulting ϵ^* parameters for length $N=12-16$. The smaller chains ($N<12$) do not have enough energy levels to make a fit for three parameters well defined. While one could argue that the fits are better than random, the correlation coefficients are always less than 0.5 and the approximation is not usefully predictive. The quality (or lack thereof) of the fit is easily seen in Fig. 3. The top plot shows the inability of the simple approximation to fit even the folding sequences alone. For completeness, the bottom plot shows all the sequences. Clearly this simple fit is so bad that even given the ground state of a sequence, one cannot predict whether or not it will be a folding sequence.

TABLE III. Fits of approximated free energies for the HP model for lengths 12–16 without additional terms. ϵ^*_{HH}, ϵ^*_{HP}, and ϵ^*_{PP} are the approximating parameters given in Eq. (4). The parameter r has the meaning of correlation coefficient.

N	Model	ϵ^*_{HH} (10^{-5})	ϵ^*_{HP} (10^{-5})	ϵ^*_{PP} (10^{-5})	χ^2	r	χ^2	r
					Folding		Nonfolding	
12	HP	1.78	5.95	3.19	6×10^{-8}	0.34	0.01	−0.21
14	HP	0.85	23.11	10.73	5×10^{-6}	0.45	0.02	0.08
16	HP	1.94	14.85	9.46	1×10^{-6}	0.35	0.05	0.04

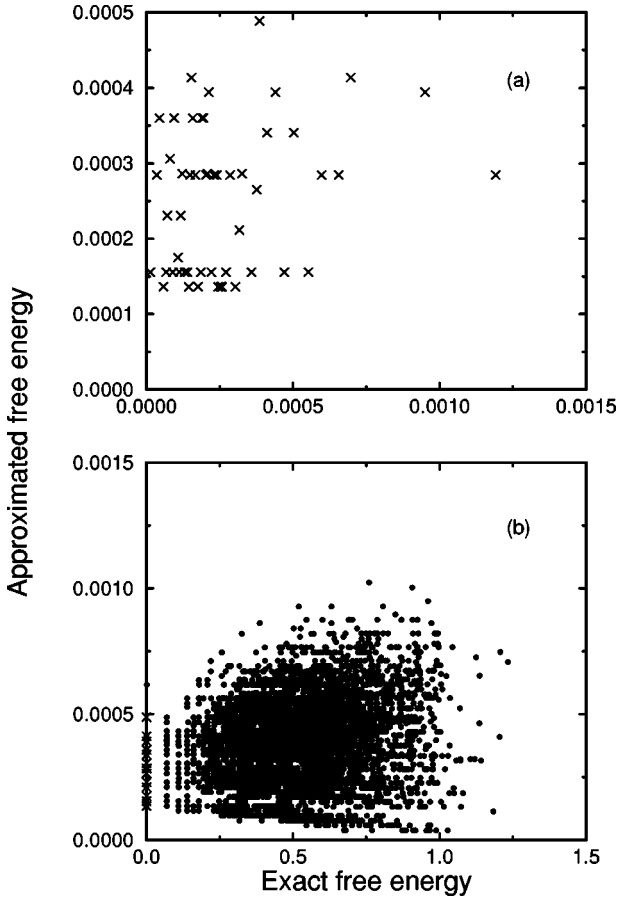

FIG. 3. Fit using simple approximation for HP data for 16-mers. Crosses denote folding sequences and dots nonfolding sequences. (a) Folding sequences only (no degeneracy) and (b) all sequences. More than 10^5 points were used in the calculation, but for clarity, a sampling of 4726 points is shown.

This result may be a property of the simple model used, but another trend is clear from the results. The fit parameters, ϵ^*_{HH}, ϵ^*_{HP}, and ϵ^*_{PP} are clearly dependent on system size. This is a severe problem with this kind of free energy approximation and has been seen in other contexts.[5] This is discussed below.

B. Additional terms

Perhaps one cannot predict free energies using only Eq. (4) (simple HP model). This does not mean that some other, more elaborate, function will not work. This can be tested by adding terms which are not in the functional form of the original potential energy. We investigated cases which could be seen as a solvation term [Eq. (6)], a second-nearest neigh-

TABLE IV. Parameter values for fits for chains of length $N=12-16$ for the HP model. Solv and 2nd nbor indicate if ''solvent'' and second nearest neighbor terms were used. ϵ^*_{ab} are the fit nearest neighbor interaction parameters as given by Eq. (4). κ^*_H and κ^*_P are the ''solvation'' parameters for hydrophobic and polar groups corresponding to Eq. (6) and the ρ^*_{ab} correspond to the second-nearest neighbor interaction parameters of Eq. (7). r_{fold} and r_{all} refer to the correlation coefficient for the folding sequences and all sequences, respectively.

N	Solv	2nd nbor	ϵ^*_{HH}	ϵ^*_{HP}	ϵ^*_{PP}	κ^*_H	κ^*_P	ρ^*_{HH}	ρ^*_{HP}	ρ^*_{PP}	r_{fold}	r_{all}
12			1.78	5.95	3.19						0.34	−0.21
14			0.85	23.11	10.73						0.45	0.08
16			1.94	14.85	9.46						0.35	0.04
12	X		−6.41	4.56	2.42	5.34	1.43				0.68	0.60
14	X		−17.67	13.96	3.71	13.94	4.06				0.63	0.86
16	X		−11.20	16.95	9.32	12.88	1.98				0.67	0.78
12		X	−6.87	4.64	1.36			5.88	1.19	3.32	0.41	0.76
14		X	−43.07	−2.4	−2.65			25.38	12.86	7.17	0.23	0.76
16		X	−26.87	5.47	7.62			17.08	6.84	3.91	0.37	0.75
12	X	X	−5.29	5.77	0.47	3.48	−1.44	2.13	2.43	5.53	0.72	0.10
14	X	X	−36.54	2.56	0.23	7.93	3.99	14.97	6.27	1.16	0.66	0.81
16	X	X	−14.52	15.81	10.61	12.24	3.57	1.95	−7.47	−2.17	0.68	0.83

bor term [Eq. (7)], and both artificial terms [Eq. (8)]. Labels such as ''solvation'' imply a physical basis, but one could regard the terms simply as numerical tools. Table IV gives the quality of the fit, as well as the final parameters for various system sizes. Since the largest number of unknown parameters is eight, there is never a problem with underdetermined parameter sets.

Adding a second-nearest neighbor term increases the quality of the fit for folding sequences, but has the most dramatic effect for nonfolding sequences. Figure 4 shows the fit quality for the 16-mer. One may consider a correlation coefficient of 0.75 (for the nonfolding sequences) to be successful and informative, but there is another way to view the data. A substantial number of points have a true free energy around 0.5 units, but are predicted as having very low free energies. Such an approximation would lead to incorrectly classifying sequences as folding. Figure 5 shows a better fit achieved with solvation terms.

Simultaneously fitting to solvation and second-nearest neighbor terms adds no significant improvement in the correlation coefficient. For example, in the 16-mers $r=0.67$ with solvent alone and moves to $r=0.68$ with both kinds of terms. The quality of the data does not justify the extra parameters. Furthermore, there is no sign of a useful threshold for distinguishing folding from nonfolding sequences.

C. Additional models

The fits described so far have been based on a model which produces very few folding sequences (from 87 to 1539) and has few possible potential energy levels. Furthermore, the definition of folding is extremely narrow. Perhaps the poor fits are merely a reflection of the model's limitations. One way to test this is to relax the restrictions and see if the trends change in any way.

The first approach is to use a more complex interaction matrix. The HA model still has only two monomer types, but suffers less from degenerate ground states[9] as shown in Table II. The quality of fits is summarized in Table V and the results for the HA model are generally the same as for the HP model. The simple fit is appalling. Second-nearest neighbor terms give a marked improvement and a solvation term is more effective. The best fit is given by the formulation with the most terms (second-nearest neighbor+solvation), but the improvement is too small to justify the additional

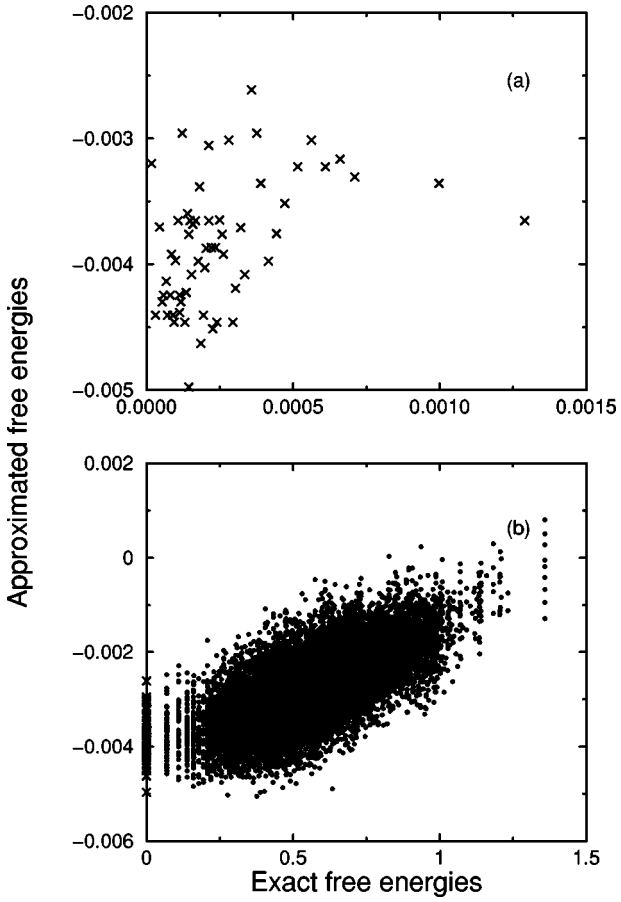

FIG. 4. Fit including second-nearest neighbor terms for the HP model, 16-mers. Crosses denote folding sequences and dots nonfolding sequences. (a) Folding sequences only and (b) all sequences. More than 10^6 points were used in the fitting, but for clarity, a sampling of 5622 points is shown.

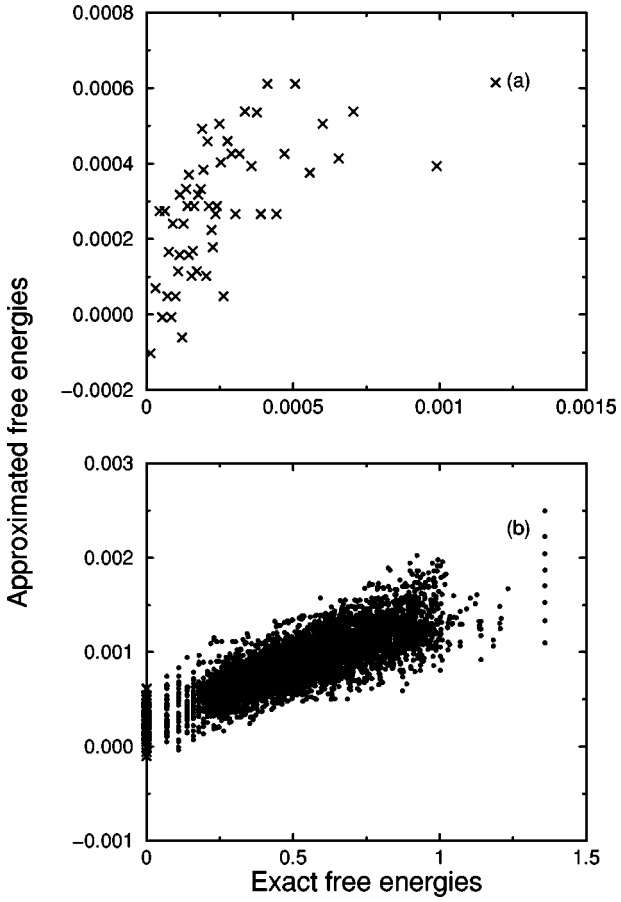

FIG. 5. Fit including solvation neighbor terms for the HP model, 16-mers. Crosses denote folding sequences and dots nonfolding sequences. (a) Folding sequences only and (b) all sequences.

adjustable parameters. The HA model results in a more intricate energy surface, but free energies cannot be reliably fit and predicted.

The definition of folding is another weakness of the approach. It has been said that a protein sequence will fold if it has a unique ground state, but in a two-dimensional lattice model, this may be an unreasonable restriction. The fits on 16-mers were repeated, but using a relaxed definition of folding where a sequence is allowed to have two conformations of lowest energy. These are labeled with ''folding def=2'' in Table V. In general, the results are the same as for the restricted definition of folding (folding def=1). Simple fits using three adjustable interaction parameters are not useful and fits using second-nearest neighbor terms are better. The best fits are seen with solvation terms where the correlation coefficient is often around $r=0.8$ which may almost be useful.

TABLE V. Fit quality with model variations. The model refers to either HP or the 2 monomer HA interaction matrix given in Table I. Solv and 2nd nbor refer to the inclusion of solvation or second-nearest neighbor terms. r_{fold} and r_{all} are the correlation coefficients with the same meaning as in Table IV. ''Folding def'' states, for given sequence, the number of conformations allowed to have ground state energies in order to classify the given sequence as a folding sequence.

Model	Folding def	Solv	2nd nbor	r_{fold}	r_{all}
HA	1			0.09	0.30
HA	1		X	0.10	0.35
HA	1	X		0.43	0.80
HA	1	X	X	0.46	0.78
HA	2			0.06	0.20
HA	2		X	0.11	0.34
HA	2	X		0.43	0.81
HA	2	X	X	0.46	0.80
HP	2			0.04	−0.49
HP	2		X	0.10	0.75
HP	2	X		0.21	0.84
HP	2	X	X	0.29	0.78

IV. DISCUSSION

Before interpreting these results it is worth seeing their place in the context of free energy approximations in different systems and models. At one extreme, one has atomistic proteins or polymers where accurate free energies are not easy to obtain experimentally or from simulation. At the next level of approximation, one may accept data which comes from different systems (archived structures) and it may be useful to assume a Boltzmann-distribution of observations.[1,2] One may note the problems in methodology[4] or use iterative methods to remedy them.[11] Continuing to simpler models, it is possible to construct examples where the sampling methodology does not produce very accurate free energy estimates.[5] Going to the simplest and smallest systems, it is possible to see what the limits of pairwise approximations to free energy ultimately are. Unfortunately, there is no continuum of work going from two-dimensional toy models to real proteins, so some speculation is warranted.

It is also important to compare results showing the limitations of other approaches. Several workers have asked whether functions can be constructed so as to favor native conformations over all others.[12] Others have shown that for some formulations this way is exceedingly difficult.[13] Here, we are not interested in recognizing the single native conformation, nor entering into the argument as to whether this is the lowest free energy state. Instead, we ask whether the free energy can be approximated for a collection of states over a range of related systems.

One should also note small differences when making comparisons to other work. The Hamiltonians used here come from rather arbitrary models for potential energy and different results would be obtained with more sophisticated examples. The simplest model used here is referred to as the HP model because of its similarity to classic work,[6] but there is a significant difference. In the original description of the HP model, the interaction matrix (Table II) was regarded as the free energy of the system. Here, we use the interaction matrix for the potential energy and calculate the resulting free energy exactly. This is valid since any plausible potential energy model will do. All that matters is that, given some potential energy, the free energy can be rigorously calculated.

Given all of these caveats, one may find trends in the results. The first three rows of Table IV show that the simple contact interaction matrix cannot be adjusted to reproduce both potential energies and entropic contributions simultaneously at the specified temperature. The next three lines of

Table IV (with a solvation term) are intuitively expected. The HH interaction term (ϵ^*_{HH}) is always the most negative and of the two solvent interaction parameters, the polar one, κ^*_{P}, is always the more favorable. The results with the second-nearest neighbor terms are less expected. Table IV shows the ϵ^*_{HH} parameter to be most favorable, but the ρ^*_{HH} is not favorable. Trying to interpret this physically would suggest that HH interactions were favorable at a distance of 1 unit, but unfavorable at a distance of $\sqrt{2}$. This is, of course, a complete artifact. HH interactions at a distance of $\sqrt{2}$ are not considered in the original potential energy. The distribution of the interaction types at this distance is somewhat correlated with the exact free energy and this is reflected in the fit ρ^*_{HH} parameter. In one sense, these results are typical of fitting in real applications. A term may have a title such as second nearest neighbor interaction, but numerically, it is another degree of freedom in the fitting and in moving to accommodate the data, it may not behave as its name suggests.

Although the systems here are small and simple, it is worth trying to compare to larger or more elaborate systems. From the first few lines of Table IV, it is clear that the $\epsilon^*_{\sigma_i\sigma_j}$ parameters are not transferable between different system sizes. This problem is seen with the more common Boltzmann/knowledge-based methodology on simple systems.[5] This reflects the parameters adjusting to accommodate the different ratios of exposed to buried sites. A more interesting question is whether the terms labeled ''solvation'' remove the size dependence. If they truly accounted for solvation and were additive with the pairwise terms, the size dependence would vanish. The numerical contribution from the solvation would scale appropriately with the surface to buried ratio as the system size was changed. From Table IV, this is certainly not happening. The reason is that the intuitive separation of contributions is not appropriate. Although a term might be labeled ''solvation,'' the parameters are adjusted to reflect the contributions from all the structures and reflect distributions of all the structures which contribute to the partition function. By simple arguments, one can see that the terms are not independent since a large number of monomer—monomer interactions is highly correlated with a smaller number of monomer–solvent (empty site) interactions. This can be confirmed by examination of the correlation matrices from the fitting procedure (data not shown). It is certainly true that better, multibody representations of solvent exist[14] and should lead to better size independence. These, however, would run counter to the spirit of simple pairwise approximations.

There is another way to view the result from the so-called solvation terms. They can be compared to the second nearest neighbor terms for efficacy. In this case the results are interesting. While neither term is obviously more realistic, the interaction with empty lattice sites is the better term for modeling free energy. This bodes well for the use of the term in lattice simulations[7] and perhaps, by analogy, the use of solvent exposure terms in atomistic simulations.

With any lattice model, there is always the question of how transferable the results are to a more realistic system or even a different simple model. From the literature, one knows that changing the representation would change the results. For example, it is known that about 2% of sequences on a square lattice have nondegenerate ground states, but the fraction is much higher on triangular lattices.[15] Similarly, using a larger number of monomer types would also give more folding sequences[16–18] (and a better dispersion of potential energy levels). Unfortunately, given the exponential growth of calculation size, no systematic calculations could be performed with more than two monomer types.

Quantitative results may not be transferable between systems, but it may be useful to see which trends continue with more realistic systems. Some results should be clear with more computational time. If one wants to pursue simple proteins, it would be possible to see how long chains must be before parameters become less size dependent. Some results are less easy to anticipate. It has been shown that properties such as designability and foldability are strongly dependent on the interaction matrix and most importantly, the number of monomer types.[16] Unfortunately, when doing exhaustive calculations on these systems, the computational effort is factorial in the chain size and exponential in the number of monomer types. One may have to use a sampling approach to tackle these systems at the cost of losing the elegance of exhaustive enumeration.

Overall, the results suggest that it will always be difficult to reproduce free energies, when limited to two-body terms and the problem would be even harder if one wanted an approximation which worked at more than one temperature. On the other hand, the observed fit of free energies is quite encouraging. With increasing system size (more protein/polymer like), the quality of fit improves. This leaves open the question of whether one can get a good enough approximation to at least recognize folding sequences in larger proteins or polymers. If one has a limited application area and is satisfied with an approximation for some small range of systems, then the use of apparently artificial terms may be practical.

ACKNOWLEDGMENTS

We gratefully acknowledge Adrian Cootes, Burkhard Dünweg, Ralf Everaers, and Kurt Kremer for advice and fruitful discussions.

[1] H. Tanaka and H. A. Scheraga, Macromolecules **9**, 142 (1976); D. T. Jones, W. R. Taylor, and J. M. Thornton, Nature (London) **358**, 86 (1992); M. J. Sippl, J. Comput. Aid. Mol. Des. **7**, 473 (1993); D. T. Jones and J. M. Thornton, Curr. Opin. Struct. Biol. **6**, 210 (1996).
[2] S. Miyazawa and R. L. Jernigan, Macromolecules **18**, 534 (1985).
[3] M. J. Sippl, M. Ortner, M. Jaritz, P. Lackner, and H. Floeckner, Folding Des. **1**, 289 (1996); M. J. Sippl, J. Mol. Biol. **260**, 644 (1996).
[4] A. Ben-Naim, J. Chem. Phys. **107**, 3698 (1997).
[5] P. D. Thomas and K. A. Dill, J. Mol. Biol. **257**, 457 (1996).
[6] K. F. Lau and K. A. Dill, Macromolecules **22**, 3986 (1989); (a) There is the possibility of confusion in the term ''free energy.'' The interaction terms in simple models such as the HP model are thought of as to include implicitly effects of degrees of freedom neglected in the model, such as internal degrees of freedom within the residues or solvent molecules. Being effective interactions, they also incorporate entropic contributions from the neglected degrees of freedom and they are often referred to as ''free energies.'' However, they do not incorporate the conformational entropy of the protein (or polymer) chain. In the present article, we use them as potential energies to evaluate the conformational distribution of

the protein, from which we obtain pairwise interactions which now also contain the conformational entropy. These are referred to as potentials of mean force or free energies.
[7] M.-H. Hao and H. A. Scheraga, Physica A **244**, 124 (1997); S. Premilat and O. Collet, Europhys. Lett. **39**, 575 (1997).
[8] H. S. Chan and K. A. Dill, J. Chem. Phys. **95**, 3775 (1991); **100**, 9238 (1994); A. M. Gutin and E. I. Shakhnovich, *ibid*. **98**, 8174 (1993); E. I. Shakhnovich, Folding Des. **3**, R45 (1998).
[9] H. Li, R. Helling, C. Tang, and N. Wingreen, Science **273**, 666 (1996).
[10] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge University Press, Cambridge, England, 1992).
[11] P. D. Thomas and K. A. Dill, Proc. Natl. Acad. Sci. U.S.A. **93**, 11628 (1996).
[12] R. I. Dima, J. R. Banavar, and A. Maritan, Protein Sci. **9**, 812 (2000); F. Seno, A. Maritan, and J. R. Banavar, Proteins **30**, 244 (1998); J. von Mourik, C. Clementi, A. Maritan, F. Seno, and J. R. Banavar, J. Chem. Phys. **110**, 10123 (1999); C. Clementi, A. Maritan, and J. R. Banavar, Phys. Rev. Lett. **81**, 3287 (1998); V. N. Maiorov and G. M. Crippen, J. Mol. Biol. **227**, 876 (1992).
[13] M. Vendruscolo, R. Najmanovich, and E. Domany, Proteins **38**, 134 (2000); M. Vendruscolo and E. Domany, J. Chem. Phys. **109**, 11101 (1998).
[14] J. M. Sorenson and T. Head-Gordon, Folding Des. **6**, 523 (1998).
[15] A. Irbäck and E. Sandelin, J. Chem. Phys. **108**, 2245 (1998).
[16] N. E. G. Buchler and R. A. Goldstein, Proteins **34**, 113 (1999).
[17] C. Micheletti, A. Maritan, and J. R. Banavar, J. Chem. Phys. **110**, 9730 (1999).
[18] R. Backofen, S. Will, and E. Bornberg-Bauer, Bioinformatics **15**, 234 (1999).