PROTEINS: Structure, Function, and Genetics 51:21–40 (2003)

Ab Initio Construction of Polypeptide Fragments: Accuracy of Loop Decoy Discrimination by an All-Atom Statistical Potential and the AMBER Force Field With the Generalized Born Solvation Model

Paul I. W. de Bakker,* Mark A. DePristo, David F. Burke, and Tom L. Blundell
Department of Biochemistry, University of Cambridge, Cambridge, United Kingdom

ABSTRACT **The accuracy of model selection from decoy ensembles of protein loop conformations was explored by comparing the performance of the Samudrala–Moult all-atom statistical potential (RAPDF) and the AMBER molecular mechanics force field, including the Generalized Born/surface area solvation model. Large ensembles of consistent loop conformations, represented at atomic detail with idealized geometry, were generated for a large test set of protein loops of 2 to 12 residues long by a novel ab initio method called RAPPER that relies on fine-grained residue-specific phi/psi propensity tables for conformational sampling. Ranking the conformers on the basis of RAPDF scores resulted in selected conformers that had an average global, non-superimposed RMSD for all heavy mainchain atoms ranging from 1.2 Å for 4-mers to 2.9 Å for 8-mers to 6.2 Å for 12-mers. After filtering on the basis of anchor geometry and RAPDF scores, ranking by energy minimization of the AMBER/GBSA potential energy function selected conformers that had global RMSD values of 0.5 Å for 4-mers, 2.3 Å for 8-mers, and 5.0 Å for 12-mers. Minimized fragments had, on average, consistently lower RMSD values (by 0.1 Å) than their initial conformations. The importance of the Generalized Born solvation energy term is reflected by the observation that the average RMSD accuracy for all loop lengths was worse when this term is omitted. There are, however, still many cases where the AMBER gas-phase minimization selected conformers of lower RMSD than the AMBER/GBSA minimization. The AMBER/GBSA energy function had better correlation with RMSD to native than the RAPDF. When the ensembles were supplemented with conformations extracted from experimental structures, a dramatic improvement in selection accuracy was observed at longer lengths (average RMSD of 1.3 Å for 8-mers) when scoring with the AMBER/GBSA force field. This work provides the basis for a promising hybrid approach of ab initio and knowledge-based methods for loop modeling. Proteins 2003;51:21–40.**

Key words: discrimination; decoy ensembles; generalized born continuum solvation; AMBER force field; statistical potential; loop modeling

INTRODUCTION

Any method that aims to predict protein structure requires a scoring function that is able to distinguish between native and non-native conformations.[1,2] Decoy ensembles containing conformers that are different from the native structure as well as near-native conformers are a prerequisite in testing the discriminatory power of a scoring function. The primary requirement is that the native state corresponds to the lowest well on the energy surface with other non-native conformations having higher energies.

In molecular mechanics force fields such as AMBER[3] and CHARMM,[4] the potential energy usually is described as a summation of bonded terms (bond stretching, angle bending, torsion angle) and nonbonded terms (electrostatics and van der Waals interactions). Primarily resulting from the non-negligible desolvation cost of (partial) burial of polar and charged residues in the folded state with respect to the unfolded state, proper treatment of solvation is a critical element to obtain accurate effective or conformational free energies in which the solvent degrees of freedom have been integrated out.[2] This is highlighted by a number of early studies demonstrating that the intramolecular terms of the potential energy function alone are insufficient to distinguish correctly folded from incorrectly folded protein structures.[5,6] Although most of the effects of the aqueous environment of protein conformation can in principle be simulated at all-atom detail, such approaches are too time-consuming for sampling conformational space for entire polypeptide fragments because adequate conver-

Abbreviations: GBSA, Generalized Born Surface Area; PID, percent sequence identities; RMSD, root–mean–square deviation.

Grant sponsor: Cambridge European Trust; Grant sponsor: Isaac Newton Trust; Grant sponsor: NUFFIC Talentenprogramma; Grant sponsor: BBSRC; Grant sponsor: Marshall Aid Commemoration Commission.

*Correspondence to: Paul de Bakker, Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge, CB2 16A, United Kingdom. E-mail: paul@cryst.bioc.cam.ac.uk

Received 9 April 2002; Accepted 12 July 2002

gence may not be achieved within the timescales of current simulations. An additional complication is the proper treatment of long-range (electrostatic) interactions as both cutoff schemes[7] and artificially imposed periodic boundary conditions have been reported to introduce artifacts.[8]

Implicit treatment of solvation by macroscopic continuum models has grown enormously in its popularity.[9–11] Smith and Honig[12] noted that the addition of a solvation term based on Poisson–Boltzmann electrostatics to a gas-phase potential function allows the selection of conformations on the basis of calculated energies that are quite close to the crystal structure for two of three medium-sized loops in ribonuclease H. More recently, Rapp and Friesner[13] demonstrated that by using the AMBER force field augmented by the Generalized Born solvation model,[14,15] loop conformations with low root–mean–square deviation (RMSD) values to the native conformation could be identified by an extensive minimization and simulated annealing protocol for a single 8-mer and 12-mer loop in the structure of ribonuclease A.

Many potential functions have been tested on a number of de facto decoy sets such as the Holm–Sander misfolded set for whole proteins.[16] In this set, the sequences of a pair of proteins of equal length have been projected onto another backbone (with optimized placement of side chains), effectively swapping folds of two unrelated proteins. Other decoy sets have been made available at the Decoys 'R' Us database (http://dd.stanford.edu).[17] Vorobjev et al.[18] have shown that a combination of an explicit and an implicit solvent model coupled to a molecular mechanics force field can accurately produce conformational free energies to distinguish correct from incorrect conformations in the Holm–Sander decoy set. Another model that does not depend on continuum electrostatics calculations is the Gaussian solvent exclusion model, which was integrated with the CHARMM polar hydrogen potential energy function.[19] This so-called effective energy function 1 (EEF1) function also achieved good discrimination in the Holm–Sander misfolded set.[20] In another study, the combination of the CHARMM force field with the Generalized Born solvation model was effective at discrimination in the Holm–Sander set and, interestingly, also pointed to differential behavior in dynamics between native and misfolded structures.[21]

Another class of scoring functions are the so-called knowledge-based or statistical potentials that are based on statistical analyses of pair contacts or distances within protein structures. Because of their statistical nature they can, in principle, include all known (solvation) and even hitherto unrecognized physico-chemical effects implicitly. The major advantage of statistical potentials is that they are not as sensitive to minor local displacements in the structure. Indeed, molecular mechanics force fields are known to result in rough energy surfaces that do not become smoother in the vicinity of the native state.[2] Although the validity of their derivation is not entirely clear,[22,23] statistical potentials have been applied successfully to protein fold recognition,[24,25] ab initio modeling,[26,27] and protein–protein docking.[28]

Residue-based statistical potentials represent amino acid residues by single interaction sites, and their use is more appropriate in cases in which microscopic structural detail can be ignored. In fold recognition or threading, for example, it is more important to be able to recognize long-range relationships between residues far apart in sequence and to get a global measure of compatibility between the sequence of interest and a template structure. Local structural features must be taken into account, however, when assessing the quality of all-atom protein structures or modeling fragments in a protein framework. In such cases, it is essential to incorporate atomic detail as models may only differ in subtle ways. To this end, all-atom statistical potentials have been developed with their main differences in representation of atom types and their dependence on sequence separation between interacting atoms.[29–31]

The probability discrimination function (RAPDF) developed by Samudrala and Moult[30] follows the definition of the potential of mean force by Sippl[32] but is cast into a Bayesian formulation. All heavy atoms of all residue types are defined independently, resulting in a total of 167 atom types, and interaction scores are derived for all possible pairs of atom types from a non-redundant protein structure set and separated into 18 discrete distance bins. No distinction is made between local and non-local (tertiary) interactions, but this reportedly does not diminish decoy discrimination power. The performance of the RAPDF function was benchmarked in the original work against a number of decoy sets, one of which is a loop decoy set comprising conformers for 11 loops in total (five 4-residue loops and six 5-residue loops) from four different proteins. Good discrimination in the Holm–Sander decoy set has also been achieved with this RAPDF function and the all-atom statistical potential developed by Melo and Feytmans.[29] All-atom statistical potentials have also been shown to discriminate succesfully between current and older Protein Data Bank structures marked as obsolete, indicating the sensitivity to the fine structural detail of proteins.[31,33]

In this study, we focused on the selection of native and near-native loop conformations, built in the context of their native framework, from large decoy ensembles that are generated with a novel method called RAPPER. Owing to their remarkable structural diversity, loops represent a stringent test case both for conformational sampling and in addressing the problem of selecting conformers that are close to native. We demonstrated in an accompanying article that a naive approach to conformational sampling by respecting van der Waals interactions of a fixed protein framework alone is successful in generating representative ensembles for protein loop conformations, provided the use of fine-grained ϕ/ψ tables.[34] These ensembles contain both many near-native and non-native conformers and it was observed that selecting conformers from ensembles on the basis of a naive geometrical criterion (anchor geometry) gives RMSD values to the native structure that are not far from prediction accuracies reported elsewhere, illustrating the need for an objective benchmark of scoring functions.

A comprehensive analysis of the discriminatory power is presented here of two variant forms of the Samudrala-Moult RAPDF statistical potential, as well as of the AMBER molecular mechanics force field, including the Generalized Born/surface area solvation model. Additionally, we assessed the importance of the solvation model by running gas-phase calculations in parallel. We show that AMBER/GBSA gives, on average, superior performance in terms of ranking native and selecting low-RMSD conformers from the decoy sets.

We first discuss the computational details of the RAPPER program in brief and how this method is applied to generate ensembles of conformations for the entire loop test set. Two simple selection criteria and two more elaborate scoring functions (RAPDF and AMBER/Generalized Born/Surface Area [GBSA]) are subsequently assessed for their discriminatory power. We show that near-native fragments extracted from real structures, often from non-homologous proteins, can be selected purely on the basis of calculated conformational free energies. Results are also presented for conformers produced by an existing loop prediction program that makes use of an 8-state ϕ/ψ set. This study is concluded with a discussion on required accuracy of structural representation and on the implications for protein structure prediction, in particular loop modeling.

METHODS

Computational Procedure

The overall procedure can be summarized as follows: 1) Ab initio generation of ensembles of model conformations for the loop target set using RAPPER.[34] 2) Modelling of the side chains using SCWRL[35] within the context of the rest of the native protein structure. 3) The conformers of the decoy ensemble are ranked for each loop target in the test set (see Methods section) on the basis of the following scoring functions: i) Anchor geometry of the generated fragment at the C-terminal anchor ($RMSD_{anchor}$); ii) SCWRL residual clash energy; iii) Samudrala–Moult all-atom probability discriminatory function[30]; iv) energy minimization using the AMBER molecular mechanics force field,[3] including the GBSA continuum solvation model.[14,15] Thus, the generation of the decoy ensembles is separated from the assessment of scoring functions for their ability to recognize native from non-native conformers.

Ab initio modeling algorithm: RAPPER

Because the conformational sampling algorithm is described in more detail elsewhere,[34] we describe the essential steps here only briefly. The program RAPPER generates an ensemble of self-consistent conformations of a specified segment of a protein structure within the context of a fixed (native) framework. Conformational sampling is performed using a round-robin scheduling queue, where fragments are iteratively extended in the N-to-C direction with efficient checking for van der Waals overlaps using a gridbased cache. All heavy atoms (N, Cα, C, O) of the backbone and the Cβ are modelled assuming idealized stereochemistry defined by Engh and Huber.[36] Backbone ϕ/ψ dihedral angles are sampled from residue-specific propensity tables with a resolution of 5° while the ω dihedral angle is kept fixed at 180°. Atomic van der Waals radii were taken from the program PROBE[37] and scaled down by 20%. Fragments containing clashing atoms are immediately discarded. The search is further guided by a distance check between the Cα of the C-anchor residue and the outermost Cα coordinate of the growing fragment. If the fragment cannot reach this Cα atom by assuming a maximum distance of 3.7 Å per residue separation (corresponding to the average *trans*. $C\alpha_i - C\alpha_{i+1}$ distance), it is also discarded.

To ensure conformational diversity of the generated ensemble, no two fragments are allowed to have a global RMSD (without superposition of the fragments) of less than 0.2 Å. All collected fragments that bridge the loop are subjected to a final optimization step by a perturbation algorithm, in which randomly chosen dihedral angles within the fragment are continuously altered so as to improve the overlap of the N and Cα atoms of a dummy C-terminal anchor residue and the C-terminal anchor residue of the framework structure (referred to as the $RMSD_{anchor}$). A measure for the local geometry of the C-terminal anchor residue, the $RMSD_{anchor}$ is one of the simple scoring functions assessed in this study for its selection accuracy.

In the present study, ensembles of at most 1000 distinct conformers are collected for all loop targets of all lengths, except 2-residue loops where 50 conformers are sufficient.

Side-chain modeling

RAPPER builds all heavy atoms of the polypeptide backbone and the Cβ atom. Side chains are therefore remodeled for all generated conformers using the program SCWRL (URL: http://www.fccc.edu/research/labs/dunbrack/scwrl/) (version 2.8)[35] while taking into account the rest of the protein structure (framework) to avoid improper rotamer selection leading to steric clashes. SCWRL relies on a backbone-dependent rotamer library to select the most favorable rotamers for each loop residue and systematically resolves steric clashes between modeled side chains and the rest of the structure.

In some cases, SCWRL is unable to place side chains free of steric overlap, which is reflected by the calculated residual clash energy. This clash energy is tested for its discriminatory potential. The rationale for testing this term as a scoring function lies in the fact that RAPPER fragments are built without knowledge of potential clashes between side-chain atoms and the protein framework. Conformers with non-native backbones and therefore poorly assigned side chains may be identified on the basis of the SCWRL clash energy.

Loop target set

We used the test set for loop targets of 2 residues to 12 residues collected by Fiser et al.[38] The following structures (PDB codes given) were found to be obsolete and were replaced with the current structure (in parentheses): 2cyr

(3cyr), 4fxn (2fox), 3b5c (1cyo), and 4ptp (5ptp). Nonstandard amino acids in protein structures were stripped. Although we attempted to fix the structure when there are missing side chain atoms (reason 2, see below) with SCWRL, sometimes the side chain is close to the loop target or there are simply too many missing atoms to unambiguously determine atomic coordinates. In many of these cases the incompleteness of a structure presents a problem when performing energy minimizations. Because these shortcomings cannot be fixed easily, we decided to discard such structures altogether. Two structures (2sns and 1ivd) have such poor ϕ/ψ and bond length/angle characteristics that these, in our opinion, cannot justifiably be included in the test set. Other structures were excluded when we found forbidden or many poor backbone dihedral angles in the target loop. The following targets were removed and reasons in parentheses are explained below: 2-mers: 1bam-73 (1,2,3), 1bgc-56 (2,3), 1ede-148 (1), 1pda-41 (2,3), 1rsy-211 (2), 3cla-108 (2); 3-mers: 1bam-72 (1,2,3), 1bgc-55 (2,3), 1mpp-264 (2), 1pda-40 (2,3), 1rsy-211 (2), 3cla-107 (2); 4-mers: 1bam-92 (2,3), 1bgc-40 (2,3), 1pda-139 (2,3), 2cyp-127 (2), 3cla-27 (1,2); 5-mers: 1bam-132 (2,3), livd-449 (4), 2cyp-139 (2), 2sns-36 (4), 3cla-49 (2); 6-mers: 1mrk-241 (1, loop is right next to C-terminus), 1pda-149 (2,3), 2cyp-159 (2), 3cla-194 (2); 7-mers: 1clc-82 (2), 2sns-134 (4); 8-mers: 1clc-313 (2), 1gof-606 (problem with minimization), 1hbq-31 (1,2), livd-413 (4), 1lst-101 (1,2), 1mpp-74 (2), 2sns-17 (4), 3cox-109 (2); 9-mers: livd-244 (4), 1pda-108 (2,3), 2cyp-145 (2); 10-mers: 1ezm-237 (2), 2sns-25 (4), 3cla-96 (2); 11-mers: 1bam-149 (2,3), livd-280 (4), 1mp-p195 (2), 1pda-85 (2,3), 1rsy-230 (2), 1thg-322 (problem with minimization), 8acn-323 (1); 12-mers: 1ivd-365 (4), 2cyp-191 (2), 2sns-111 (4), 3cla-176 (2), 3cox-478 (2).

Here are the reasons for discarding targets: 1 = poor ϕ/ψ angles or B-factors in loop residues (according to WHAT_CHECK[39] and RAMPAGE[40]); 2 = missing side chain atoms; 3 = missing regions (gaps); 4 = overall poor quality.

Definition of RMSD

The root-mean-square deviation is defined as follows:

$$d_{\text{r.m.s.}} = \left(\frac{1}{N_{\text{atoms}}} \sum_{i=1}^{N_{\text{atoms}}} |\vec{r}_i(A) - \vec{r}_i(B)|^2 \right)^{1/2} \quad (1)$$

where $\vec{r}_i$(A) is the coordinate of atom i in the structure A and N_{atoms} is the number of atoms included in the calculation. In most cases, the RMSD of a structure (A) is calculated with respect to the native experimental structure (B). When reporting RMSD values it is critical to report 1) if the conformations are optimally superimposed; 2) if so, which atoms have been used to obtain the optimal superposition; and 3) which atoms are included in the calculation of the RMSD (for example, Cα-only or all heavy backbone atoms). Prior superposition of equivalent structures yields a local RMSD, reflecting similarity between polypeptide fragments disconnected from the structure they belong to, whereas no superposition gives a global RMSD between fragments as they actually appear in the context of the protein framework. In light of selecting conformations in the context of a fixed protein framework, we only use global RMSD values, calculated over all heavy backbone atoms (N, Cα, C, O), unless explicitly stated otherwise.

Scoring by the Samudrala–Moult statistical potentials

The RAMP program suite (URL: http://compbio.washington.edu) (version 0.3) developed by Ram Samadrala was downloaded and the stand-alone program potential-rapdf was used to calculate the 'probability score' of a model. Equivalent models with identical sequences (hence atom composition) can thus be ordered according to the RAPDF score. The RAMP suite comes with a range of different scoring functions, with the major differences arising from the data sets used to compile them (Ram Samudrala, personal communication). In the present study, we use the functions scop-e4-allatoms-xray-scores (RAPDF-1) and scop-150-unique-xray-scores (RAPDF-2).

Scoring by molecular mechanics minimization

Generated fragments by RAPPER exhibit good geometrical properties, as reflected by the ϕ/ψ dihedral angles, idealized bond lengths and angles, and absense of clashes with the framework structure. However, because of the rugged energy surface of a generic molecular mechanics potential energy function, conformations need to be refined first by minimization for their calculated potential energies to be meaningful for ranking. The ensemble is prefiltered on the basis of the RAPDF score and the goodness-of-fit at the C-terminal anchor residue such that no more than 50 models with good anchor geometry in total are retained to be minimized. Of all conformers with $\text{RMSD}_{\text{anchor}} \leq 0.1$ Å (or at least 50), the best 50 conformers are selected with the most favorable RAPDF scores. This ensemble is referred to as the filtered ensemble. These structures are minimized until the convergence criterion (root–mean–square gradient < 0.1 kcal/(mol.Å)) is reached or when the number of minimization steps exceeds 100. The AMBER force field[3] (URL: http://www.amber.ucsf.edu/amber/) together with the GBSA solvation model[14,15] was used in the TINKER (version 3.8) implementation (URL: http://dasher.wustl.edu/tinker/). The minimize program in the TINKER package performs a limited-memory BFGS minimization, which is known to be tolerant of poor initial structures and is recommended for minimizations of structures to a root-mean-square gradient of 1.0 to 0.1 kcal/mol.Å. Only atoms belonging to the loop and both anchor residues were allowed to move during the minimization while the native framework was frozen. Minimized fragments are subsequently ranked according to the final effective energy (conformational free energy) $W = H_{\text{intra}} + \Delta G_{\text{solv}}$ (AMBER/GBSA) and the gas-phase potential energy H_{intra}.

Ranking of Native conformers

Both initial and filtered ensembles are sorted according to RAPDF scores and AMBER/GBSA energies, respec-

tively, and the native rank is recorded per target. The native-rank curves represent the cumulative target fraction (for a particular loop length) of which the native conformer shows up in the top-1 slice. The curves in Figures 5 and 6 are plotted of top-1, 2, 5, 10, 25 ranks (of 50 conformers in the filtered ensemble) for AMBER/GBSA scoring and of top-1, 5, 10, 50, 100 (of 1000 conformers in the initial ensemble) for RAPDF scoring.

Ranking of Selected Conformers

A simple measure of the discriminatory power of a scoring function is the number of conformers with lower RMSDs than the RMSD of the conformer that is selected on the basis of it having the most favorable score. Cumulative target fractions are plotted in Figure 7 as a function of the absolute number of better conformers in the ensemble per scoring function for all targets of a given length. Comparison between curves in the same graph is possible as all functions score the same set and number of conformers (the filtered ensemble contains $n = 50$ conformers).

Predictions by the PETRA server

All loop targets of lengths 3 to 8 were submitted to the PETRA[41] web server (URL: http://www-cryst.bioc.cam.ac.uk/cgi-bin/coda/pet.cgi). The accuracy of these predictions was compared to accuracy of selected conformers by RAPDF (the same scoring function used in PETRA) in our study. We note that in the accompanying work,[34] an overall increase was observed in the lowest RMSD attainable when adopting the Deane 8-state set. Here, we focus merely on the accuracy of the actual predictions made. PETRA uses the anchor geometry and RAPDF as selection criteria, enabling a comparison to be made with results presented here. For some loop targets, no solutions were returned.

Predictions by the FREAD server

We wanted to test the effect of supplementing the RAPPER decoy ensembles with fragments taken from real protein structures. The FREAD server (URL: http://www.cryst.bioc.cam.ac.uk/cgi-bin/coda/fread.cgi) was used to collect loop candidates for all 3-mer to 8-mer targets. FREAD uses environmentally constrained substitution tables to score fragments from a non-redundant protein database.[42] Such tables reflect the likelihood of a target sequence to adopt a particular backbone conformation.[43]

For most targets, fewer than 10 models were returned, and in some cases there were no solutions. For all targets, we ensured that no FREAD model was extracted from the native target protein or from close homologs by imposing an upper sequence identity limit of 80% calculated over the entire protein sequence. The average sequence identity of added fragments to the decoy ensembles is 20%. All side chains of FREAD fragments were modelled by SCWRL. Note that all FREAD fragments automatically bypass the filtering stage so that all conformers are minimized for ranking according to AMBER/GBSA energies.

TABLE I. Statistics of the Decoy Ensembles Generated by RAPPER for All Loop Targets

Conformational statistics of ensembles			
		(RMSD) (Å)	
Loop length	n	Lowest[a]	Average[b]
2	34	0.31	0.95
3	34	0.34	1.30
4	35	0.43	1.65
5	35	0.53	2.27
6	36	0.69	3.06
7	38	0.78	3.79
8	32	1.11	4.16
9	37	1.29	5.00
10	37	1.67	5.66
11	33	1.99	6.71
12	34	2.21	6.96

[a]Refers to the average of the lowest RMSD found in each ensemble for a given loop length.
[b]Refers to the average of the ensemble-averaged RMSD of all conformers over all loop target ensembles for a given length.
The global RMSD values are relative to the native structure and include all heavy backbone atoms of the loop fragment only.

RESULTS

Defining Optimal Selection Accuracy

For a large test set of loop targets for lengths of 2 to 12 residues, ensembles were constructed by RAPPER to assess the discriminatory power of a number of scoring functions. To do this objectively, we needed first to establish baselines that corresponded to random and optimal selection. The third column of Table I lists the average of the global RMSD of the best conformer (with the lowest RMSD) found in the ensembles for a given loop length, and ranges from 0.30 Å for 2-mers to 2.20 Å for 12-mers. Thus, these average numbers refer to a generated conformer that is closest to the native structure. Of course, in terms of selection from a decoy ensemble of conformers, no method can do better than to select the conformer with the lowest RMSD from the ensemble. The lowest-RMSD curve (the lower thick line in Fig. 1) represents the baseline for optimal selection.

Defining the Accuracy of Random Selection

The fourth column of Table I lists the average of the ensemble-averaged global RMSD for all conformers within the ensembles of a given loop length. Hence, these values correspond to the expected baseline of selection accuracy for random selection. For example, for the 8-mer targets, a random selection strategy will produce loop conformations that are, on average, 4.2 Å away from their native structures. Any proper selection method aiming to rank conformers with respect to native structural similarity should in principle outperform this upper baseline (the upper thick line in Fig. 1).

Selection by Anchor Goodness-of-Fit, SCWRL Clash Energy, and Samudrala–Moult Statistical Potentials

The performance in selection from the decoy ensembles for two simple scoring functions ($RMSD_{anchor}$ and

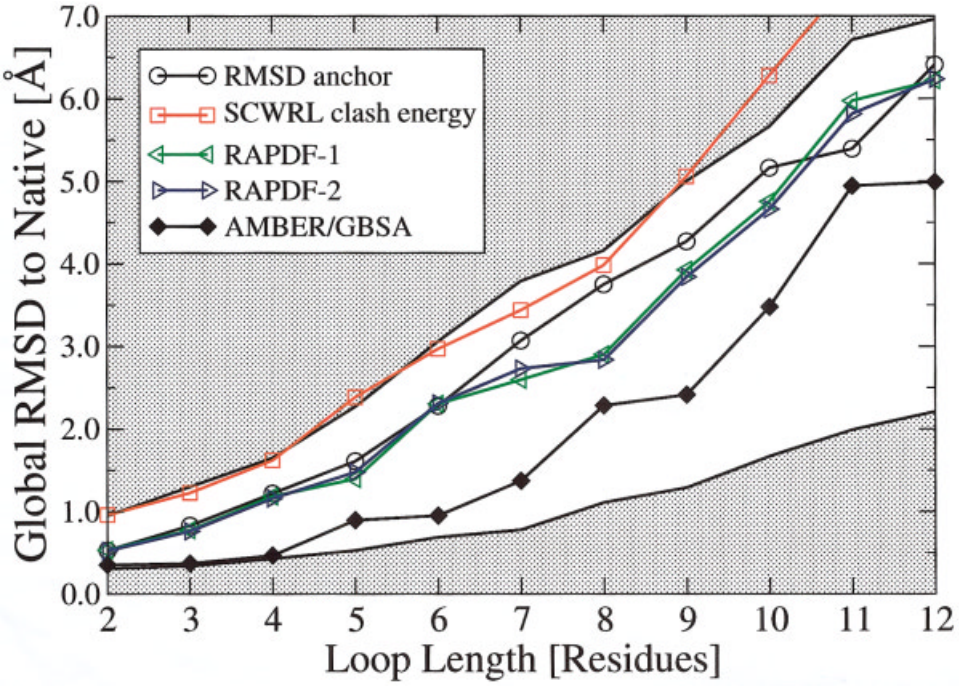

Fig. 1. Plot of the average global RMSD (Å) to native of selected conformers using different scoring functions as a function of loop length. (**A**) Anchor goodness-of-fit ($RMSD_{anchor}$); (**B**) SCWRL residual clash energy; (**C**) RAPDF score of the scop-e4-allatoms-xray-scores function; (**D**) RAPDF score of the scop-150-unique-xray-scores function; and (**E**) AMBER/GBSA minimization. For selections using (**A**), (**B**), (**C**), and (**D**), all conformers in the generated RAPPER ensemble are scored directly ($n = 1000$), whereas selections by (**E**) the decoy ensemble are filtered down to $n = 50$ conformers according to a combination of the $RMSD_{anchor}$ and RAPDF scores (see Methods section). The baselines for random selection (as average RMSD of all conformers in the RAPPER ensembles) and best-possible selection (average of the best RMSD conformer in RAPPER ensembles) are plotted as the thick lines as a function of loop length: the lower the curve, the better the discriminatory power of the function. RMSD values of selected conformers are calculated over all heavy backbone atoms and are averaged over all targets of a given length. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

TABLE II. Selection Accuracy Given as the Average Global (RMSD) (Å) of Selected Conformers for All Loop Targets of a Given Length Using Four Scoring Functions: Anchor Goodness-of-Fit ($RMSD_{anchor}$), $SCWRL_{Residual\ Clash\ Energy}$, and Samudrala–Moult RAPDF Statistical Potentials

Selection accuracies					
		(RMSD) (Å)			
Loop length	n	$RMSD_{anchor}$	SCWRL clash energy	RAPDF-1	RAPDF-2
2	34	0.53	0.96	0.53	0.53
3	34	0.84	1.23	0.78	0.76
4	35	1.22	1.62	1.19	1.16
5	35	1.60	2.38	1.40	1.48
6	36	2.29	2.97	2.30	2.31
7	38	3.09	3.44	2.60	2.73
8	32	3.75	3.98	2.90	2.84
9	37	4.27	5.06	3.92	3.84
10	37	5.16	6.28	4.75	4.66
11	33	5.47	7.47	5.97	5.81
12	34	6.41	7.67	6.22	6.24

SCWRL residual clash energy) and two more advanced functions (RAPDF-1 and -2) is given in Table II. For all these scoring functions, the conformer with the most favorable score was defined as the selected conformer from the decoy ensemble. The RMSD accuracies achieved by these scoring functions are plotted in Figure 1 together with the baselines for random and optimal selection (lower and upper thick lines).

Anchor RMSD

The goodness-of-fit at the anchor residue was calculated as the global RMSD of the N and Cα atoms of the C-terminal anchor residue of the generated conformer and the existing framework anchor residue (see Methods section). Using the $RMSD_{anchor}$ as the scoring function, the global RMSD increases steadily from 0.5 Å for the 2-mer targets to 1.2 Å for 4-mers, to 3.7 Å for 8-mers to 6.4 Å for 12-mers. The RMSD curve for the $RMSD_{anchor}$ (circles in Fig. 1) lies roughly 0.5 Å below the baseline of random selection for all loop lengths. The relative accuracy of this scoring function deteriorates at longer lengths, as will be shown by calculating the correlation between $RMSD_{anchor}$ and the RMSD of all conformers in each ensemble.[34]

SCWRL residual clash energy

The SCWRL residual clash energy had hardly any discriminatory power as a scoring function, reflected by the poor RMSD values for selected conformers in Table II. The RMSD curve corresponding to selection on the basis of this score (Fig. 1) essentially revolves around the baseline for random selection. For loop lengths longer than nine residues, it actually performed worse than random, suggesting that the SCWRL energy is anti-correlated with RMSD.

Samudrala–Moult RAPDF Statistical Potential

Not surprisingly, the Samudrala–Moult RAPDFs both perform much better than the SCWRL clash energy to identify low-RMSD conformers. The level of selection accuracy of the RAPDFs is, however, about the same as that of the $RMSD_{anchor}$ function for short loop lengths of up to six residues. Only for loops of lengths seven to 10 residues can a slight improvement be observed compared to the $RMSD_{anchor}$ function. The average global RMSD ranges from 0.5 Å for 2-mers to 1.2 Å for 4-mers, 2.8 Å for 8-mers and 6.2 Å for 12-mers.

It is clear from the RMSD curve in Figure 1 that the RAPDF-1 and RAPDF-2 are entirely interchangeable as the differences in selection accuracy are practically negligible. Indeed, this is confirmed by the calculated correlation of $r = 0.997$ between RAPDF-1 and RAPDF-2 for all RAPPER conformers of all targets in the data set.

Selection by Molecular Mechanics Force Field

Because of the computational cost involved in the calculation of the AMBER potential energy, especially when including the GBSA solvation term, the ensembles are subjected to a filter (see Methods section) so as to retain only 50 conformers with favorable RAPDF scores and good anchor geometry. The average RMSD of the conformers in these filtered ensembles is 1.0 Å for all 4-mer targets, 3.1 Å for 8-mers and 6.3 Å for 12-mers. These conformers are then energy minimized with the AMBER force field. Figure 1 shows the RMSD curve for selection on the basis of the AMBER potential energies. The level of selection accuracy by minimization of the AMBER/GBSA potential energy was substantially better than by ranking according to either of the RAPDF functions. The average minimized RMSD of the selected conformer (with the lowest potential energy) started at 0.4 Å for 2-mer targets and increased to 0.5 Å for 4-mers, 2.3 Å for 8-mers up to 5.0 Å for 12-mers. For loops of 2 to 4 residues long, the average RMSD of selection was at the level of the optimal baseline. Although for some targets at longer lengths (8- and 12-mers) the conformers in the filtered ensemble had a relatively high RMSD, ranking by the AMBER/GBSA function, on average, allowed more accurate selection than on the basis of RAPDF alone.

Effect of the GBSA Solvation Term on Selection

To investigate the added value of the GBSA solvation model for the sake of selection, all conformers were also minimized in vacuo using the AMBER force field. In Figure 2, the RMSD of selection is plotted as a function of loop length for conformers minimized with and without the GBSA solvation term in the potential energy function, both before and after minimization. The average RMSD selection accuracy when using gasphase calculations is slightly decreased with respect to the AMBER/GBSA selection accuracies. At most lengths, the RMSD difference in selection is at most 0.2 Å with the GBSA-based calculations showing greater predictive power. The difference is most pronounced for the 9-mers and 10-mers targets, where the GB/SA minimizations have an average RMSD of selection that is 0.9 Å and 0.7 Å better than the corresponding gas-phase calculations, respectively.

In Figure 3, the RMSD of every conformer selected by gas-phase minimization is plotted for all loop targets against the RMSD of the conformer selected by GBSA minimization from the corresponding ensemble. Obviously, there are many instances where the AMBER gas-phase calculations give rise to lower RMSDs than continuum solvation minimizations and vice versa. Yet on average, for all targets of a given length, the AMBER/GBSA minimization protocol gave a better selection of conformers that are slightly closer to the native conformation.

Effect of minimization

Especially for the shorter loops of less than nine residues long, minimizing the conformers resulted invariably in an average RMSD improvement of ≈0.1 Å unless the initial conformations were far away from the native structure. This is evidence that the force field is capable of moving the structures toward the native conformation. The decrease in RMSD of the selected conformer upon minimization in the gas phase was, on average, of the same magnitude as the improvement seen with AMBER/GBSA minimizations. Although these improvements are modest on an absolute scale, certainly for short loops this is impressive on an absolute scale as the conformations are already so close to the native structure. We emphasize, however, that the purpose of minimizing the conformers is to be able to rank them in a meaningful way, not to perform more conformational sampling or to improve their conformations per se.

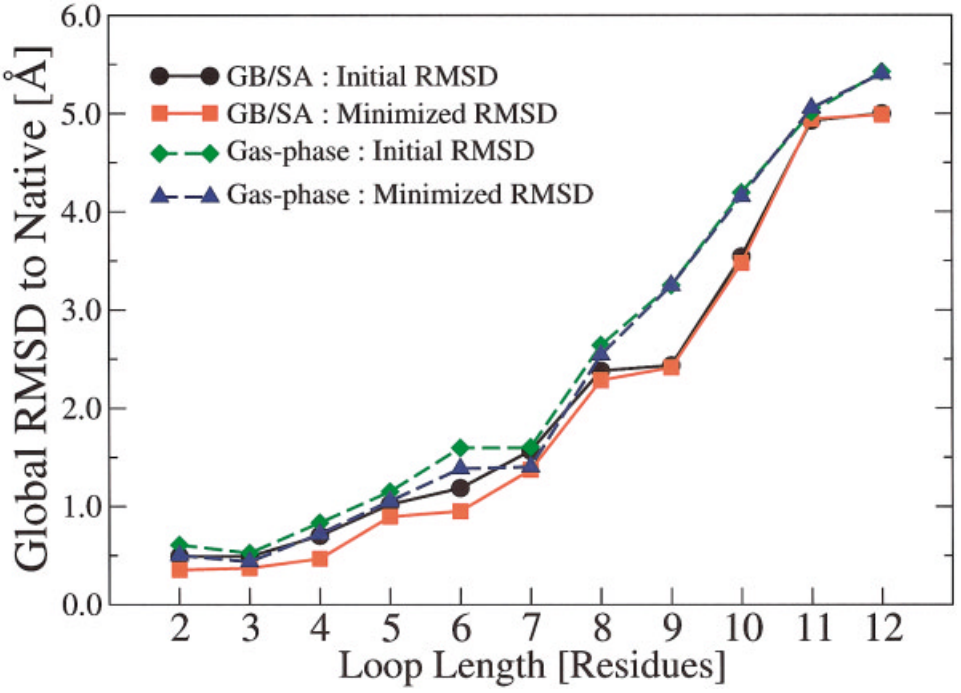

Fig. 2. Plot of the average global RMSD (Å) to native of selected conformers by energy minimization using the AMBER force field with (GB/SA: solid line/squares) and with-out (gas-phase: dashed line/triangles) the GBSA solvation model. The RMSD of the initial conformers before minimization was plotted as well (GBSA: solid line/circles and gas-phase: dashed line/diamonds). The initial sets of the 50 candidate models after filtering (see Methods section) were identical for both AMBER runs. The RMSD curve of the minimized GBSA conformers was consistently lower than the other curves. Minimization systematically moved conformers toward a lower RMSD for both GBSA and gas-phase calculations. Inclusion of the GBSA solvation energy term improved the RMSD of selected conformers over all loop lengths with respect to gas-phase calculations. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Correlations Between Scoring Functions and RMSD

The ideal scoring function has perfect correlation with the RMSD to the native conformation, allowing straightforward detection of the best available candidate by virtue of flawless ranking of the conformers present in the ensemble. In the absense of such a function, it is useful to characterize the level of correlation between the RMSD and tested scoring functions. The higher the correlation with RMSD to native, the more suitable the scoring function for selecting near-native conformers.

Figure 4 shows the correlations between RMSD and $\text{RMSD}_{\text{anchor}}$, SCWRL clash energy, and RAPDF and AMBER/GBSA potential energy for all individual 4-mer ($n = 35$), 8-mer ($n = 32$), and 12-mer ($n = 34$) target ensembles, respectively. The correlations of $\text{RMSD}_{\text{anchor}}$, SCWRL clash energy, and RAPDF [Fig. 4 (a–c)] are calculated for the entire RAPPER ensembles (1000 conformers), whereas the correlation of AMBER/GBSA potential energy [Fig. 4 (d)] and AMBER (gas-phase) potential energy [Fig. 4 (e)] is calculated only for the filtered ensembles (50 conformers). To allow a direct comparison between RAPDF and AMBER/GBSA potential energy, the correlation of RAPDF with RMSD is recalculated for the conformers in the filtered ensemble only [dashed line in Fig. 4 (d)].

The correlation of the $\text{RMSD}_{\text{anchor}}$ [Fig. 4 (a)] is relatively high ($\langle r \rangle = 0.6$) and consistent for all 4-mer target ensembles whereas the correlation diminished at 8-mer ($\langle r \rangle = 0.3$) and 12-mer ($\langle r \rangle = 0.2$) lengths. At longer lengths, the variability of the correlation of the $\text{RMSD}_{\text{anchor}}$ is striking, with some target ensembles showing high correlation and others showing no correlation at all ($r \approx 0$).

The correlation of the SCWRL clash energy [Fig. 4(b)] was substantially weaker than the $\text{RMSD}_{\text{anchor}}$, with an average correlation coefficient $\langle r \rangle = 0.3$ for 4-mer, $\langle r \rangle = 0.1$ for 8-mer and $\langle r \rangle = 0.0$ for 12-mer target ensembles.

The correlation of RAPDF with RMSD [Fig. 4(c)] was only slightly less ($\langle r \rangle = 0.5$) than the correlation of $\text{RMSD}_{\text{anchor}}$ for the 4-mer target ensembles. This is in part because of the anti-correlation ($r < 0$) observed for a single target whereas for $\text{RMSD}_{\text{anchor}}$ positive correlation was observed across all targets. The correlation of RAPDF also exhibited a somewhat larger variability than that of the $\text{RMSD}_{\text{anchor}}$. Of the 8-mer target ensembles, there are 5 cases where the correlation of RAPDF was negative, whereas in many other cases the correlation is still relatively high ($\langle r \rangle = 0.4$). The overall correlation of RAPDF in the 12-mer target ensembles is reduced to a weak $\langle r \rangle = 0.1$, again in some instances showing good correlation and in others poor correlation.

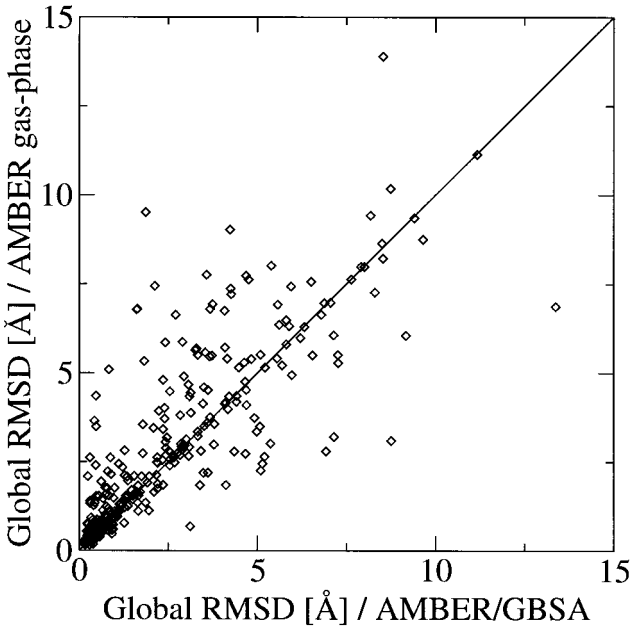

Fig. 3. Scatter plot of the global RMSD (Å) of selected RAPPER conformers using the AMBER force field with the GBSA solvation model (*x*-axis) vs. without solvation (i.e. gas-phase) (*y*-axis) for all loop targets of all lengths (n = 385). The RMSD includes all heavy backbone atoms. Minimized conformers below the diagonal line $y = x$ have lower RMSD values (closer to native) when minimized in vacuum. Minimized conformers above the line have lower RMSD values when minimized, including the GBSA solvation model. The Pearson correlation coefficient is r = 0.85.

The correlation between the AMBER/GBSA potential energy and the RMSD of the minimized conformers [Fig. 4(d)] is calculated for all filtered ensembles and averages to $\langle r \rangle = 0.6$ for 4-mers, $\langle r \rangle = 0.3$ for 8-mers, and $\langle r \rangle = 0.1$ for 12-mers. For 4-mer target ensembles, in only eight cases did we observe low or negative correlations, and in many other cases the correlation was much better. Interestingly, the correlation of RAPDF in this filtered ensemble ($\langle r \rangle$ = 0.0) is, in all but one case, substantially lower than the correlation of the AMBER/GBSA potential energy. For 8-mer target ensembles, the average correlation drops to $\langle r \rangle = 0.3$ but is still greater than the RAPDF correlation of $\langle r \rangle$ = 0.2. For seven targets, the RAPDF was better correlated with RMSD than the AMBER/GBSA potential energy of the conformers in these ensembles. For the 12-mer target ensembles, the correlation of the AMBER/GBSA potential energy and RAPDF dropped to $\langle r \rangle = 0.1$ with less than half of the target ensembles demonstrating better RMSD correlation with the RAPDF function than with the AMBER/GBSA potential energy.

The correlation of the AMBER gas-phase potential energy with RMSD is also plotted [Fig. 4(e)]. For only five targets of the 35 4-mers were the AMBER gas-phase energies better correlated with RMSD than the AMBER/GBSA energies. For as many as 14 of the 32 8-mer targets and 12 of the 34 8-mer targets, the vacuum potential energies were better correlated with RMSD.

Clearly, variability in correlation was seen at all target lengths and for all tested scoring criteria. The lack of consistency in correlation is not surprising because no two loop targets are the same, nor are the conformational ensembles for them generated by RAPPER. The exception to this is the consistent positive correlation of the $\mathrm{RMSD}_{\mathrm{anchor}}$ observed for the 4-mer target ensembles. Having the lowest average correlation with RMSD to native, the SCWRL clash energy cannot reliably be used for ranking of conformers, which is underlined by the high RMSD of selection shown in Figure 1. The average correlation of the RAPDF is about as good as the correlation of the $\mathrm{RMSD}_{\mathrm{anchor}}$ for target ensembles of all lengths, in basic agreement with the observed average RMSD of selection.

Analysis of the correlation of the AMBER/GBSA energies is complicated by the fact that the ensembles are prefiltered by $\mathrm{RMSD}_{\mathrm{anchor}}$ and RAPDF. Perhaps in some targets good conformers with low RMSDs are filtered out because of bad RAPDF scores, leaving only bad conformers for ranking in the filtered ensemble. A quick way to test this is to see whether the native conformation would have been filtered out. The big red circles in the correlation plots [Fig. 4(d)] indicate targets for which the native conformer does not satisfy the RAPDF/$\mathrm{RMSD}_{\mathrm{anchor}}$ criteria to end up in the filtered ensemble. This is the case in six of the 4-mer targets, three of the 8-mer targets, and only one in the 12-mer targets. In all of these cases, exclusion of the native conformation from the filtered ensemble does not seem to result in an apparent decrease in the correlation between AMBER/GBSA energies and RMSD of the conformers. Inclusion of the native conformer, however, does not guarantee that the other conformers in the filtered ensemble are good near-native conformers, as demonstrated by the relatively high average RMSD of all conformers present in the filtered ensembles, especially for the 8-mer and 12-mer targets. Thus, no simple relationship can be detected between the mean RMSD of the conformers in the filtered ensemble and the level of correlation between the AMBER/GBSA energies and RMSD.

Ranking of Native Conformers

Here we evaluated the ranking power of the scoring functions against the background of the decoy ensembles by following the rank of the native conformations. Figures 5 and 6 show native ranking curves for the RAPDF and AMBER/GBSA functions as a function of loop length (see Methods section). These curves represent the fraction of loop targets in which the native conformation is ranked within the top slice of all conformers, depending on the exact size used for the slice (top 1, 5, 10, 50, or 500 for RAPDF and top 1, 2, 5, 10, and 25 for AMBER/GBSA).

RAPDF

The lowest curve (top-1) in Figure 5 for the RAPDF function shows the fraction of targets of a given loop length for which the native conformer has the most favorable score. The observation that this curve is not unity implies that the RAPDF in many cases attributes a more favorable score to non-native conformers (yet perhaps near-native ones). This also reflects the quality of the decoy ensemble: often a RAPPER conformer is ranked by the RAPDF as

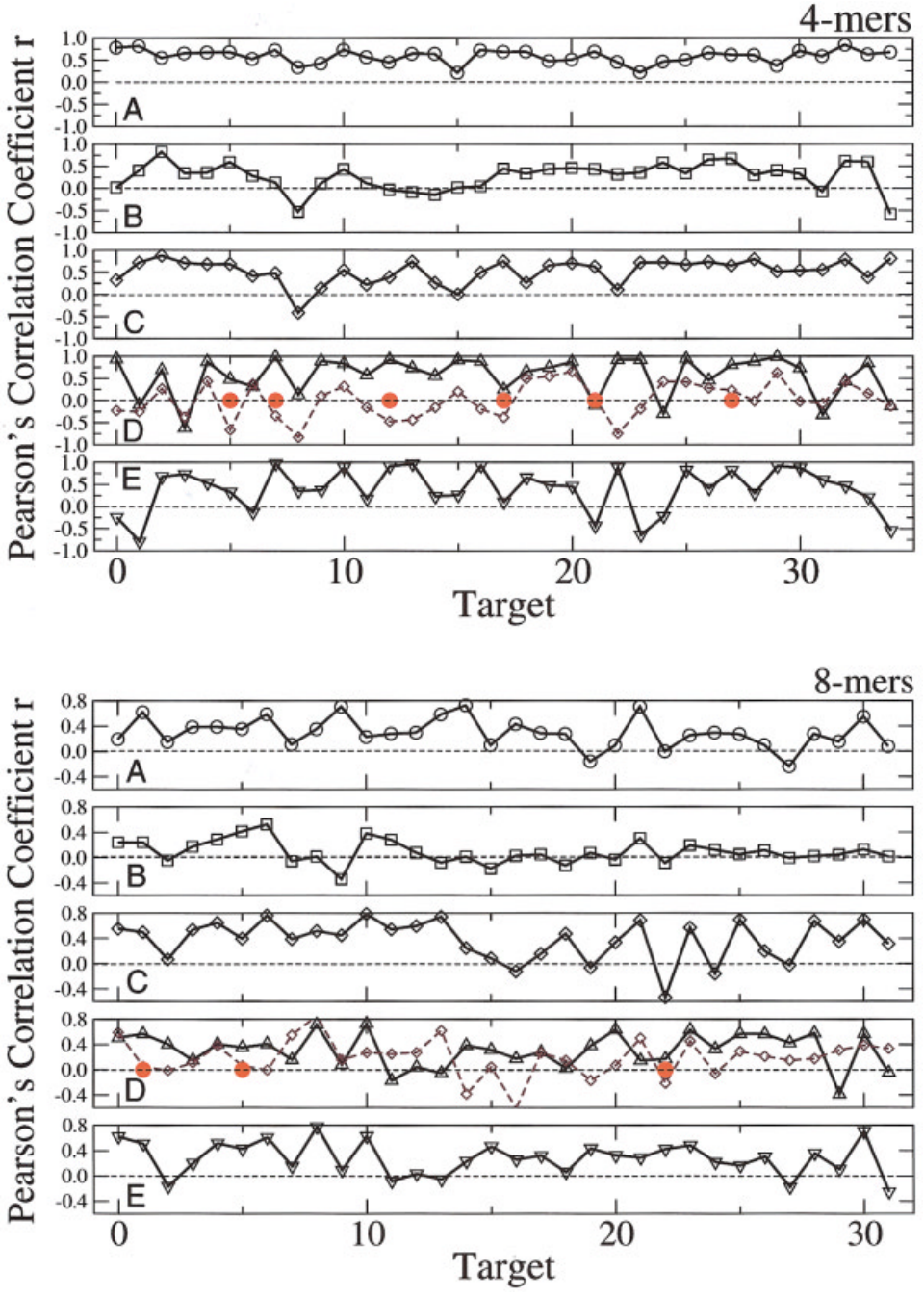

Fig. 4. Pearson's correlation coefficient *r* is plotted per individual (i) 4-mer; (ii) 8-mer; and (iii) 12-mer loop target for the relationship between the global RMSD (Å) of all generated RAPPER conformers and the following scoring functions: (**A**) Anchor goodness-of-fit ($RMSD_{anchor}$; circles); (**B**) SCWRL residual clash energy (squares); (**C**) RAPDF score of the scop-e4-allatoms-xray-scores function (diamonds); (**D**) energy minimization using the AMBER/GBSA force field (triangle up). (**E**) energy minimization using the AMBER force field (gas-phase; triangle down). For (**A**), (**B**), and (**C**), the correlation coefficient is calculated over all RAPPER generated decoy conformations ($n = 1000$), whereas for (**D**) and (**E**) only $n = 50$ conformers are included due to filtering (see Methods section). To allow a comparison between the AMBER/GBSA and RAPDF results, the RMSD correlation was also plotted for the RAPDF function for all conformers in the filtered ensemble (dashed line; diamonds). The big red circles (in **D**) indicate that the native conformation of that particular target does not satisfy the criteria to end up in the filtered ensemble. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

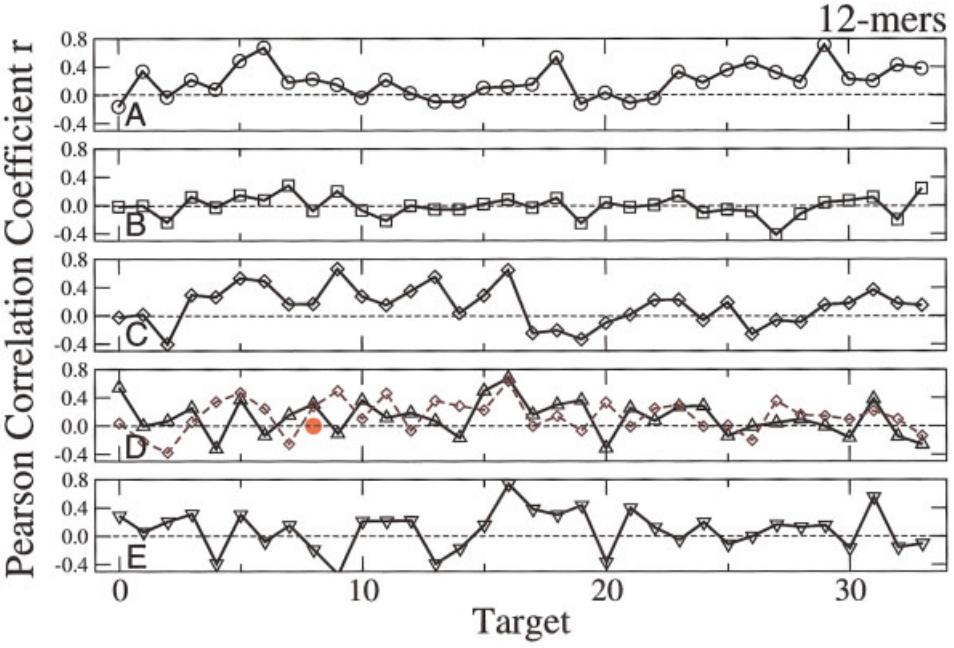

Figure 4. (Continued.)

having a better score than the native conformer. At shorter loop lengths (three to seven residue loops), a relatively high error rate is observed in ranking the native conformation, indicating low specificity of the RAPDF or a disproportional improvement of the decoy ensemble (in terms of producing more native-like conformers). These results suggest that the decoy ensembles of shorter lengths pose a real challenge to the RAPDF to recognize the native as the most favorable conformer. The results also show that the sampling at longer lengths is becoming less effective, as the ensemble size of 1000 is not enough to cover the accessible space adequately. At longer loop lengths, in 70% of all targets, the native conformation is located in the top-10 of all conformers of the ensemble. Looking at the uppermost top-100 curve, the native conformation is always ranked in the top 10% of the ensemble in at least two-thirds of all targets, for all loop lengths.

The overall trend is that the RAPDF is definitely better than random for selecting the native conformation against the decoy background. For the shorter lengths, the decoy ensembles contain many loop fragments that are ranked with better RAPDF scores. The dip of the ranking curve at four residues may reflect optimal sampling for this length.

AMBER/GBSA

Turning to the same curves for the calculated AMBER/GBSA energies after minimization of the conformers in the filtered ensembles ($n = 50$), we observed a steady improvement in the ranking ability of the force field with length (Fig. 6). As noted before, for some targets the native conformer is excluded due to the filtering procedure. In about half of the 4-mer targets, the native conformation had the most favorable AMBER/GBSA potential energy whereas for 8-mers 65% of the targets placed the corresponding native conformer at the top. This improved for the 12-mers to more than 90% of the targets, which reflect the steady increase of the average RMSD (less native-like) of all conformers at longer lengths (see Fig. 1).

Interestingly, the native ranking for the 3-mer and 4-mer targets was relatively weak because in many targets the native is not even placed within the top-25 (out of 50) because of RAPPER conformers that are ranked with better energies than the native one. The plateau of the top-1 curve at longer loop lengths (9-mers and longer) suggests that the decoy ensembles generated for these targets do not pose a real challenge for the potential function of the AMBER/GBSA force field. This was not the case for the shorter loop targets: in many cases RAPPER conformers had more favorable energies than their native competitors.

The native ranking (top-1) on the basis of the AMBER gas-phase potential function is also shown in Figure 6 (dashed line), which is significantly lower than the corresponding AMBER/GBSA line, a clear indication of lower discriminatory power.

Ranking of selected conformers

Although the RMSD of selected conformers to the native conformation is a useful measure for predictive quality of a scoring function, it does not gauge the performance with respect to the presence of other conformers in the ensemble that have lower RMSD values (and should have been selected instead). Here we analyzed how many conformers in the ensemble (excluding the native conformation, of course) had better RMSDs than the actual selected one (see Methods section). These error-rates are plotted in Figure 7 for the 4-mer, 8-mer, and 12-mer targets. Each curve represents the cumulative fraction of the loop tar-

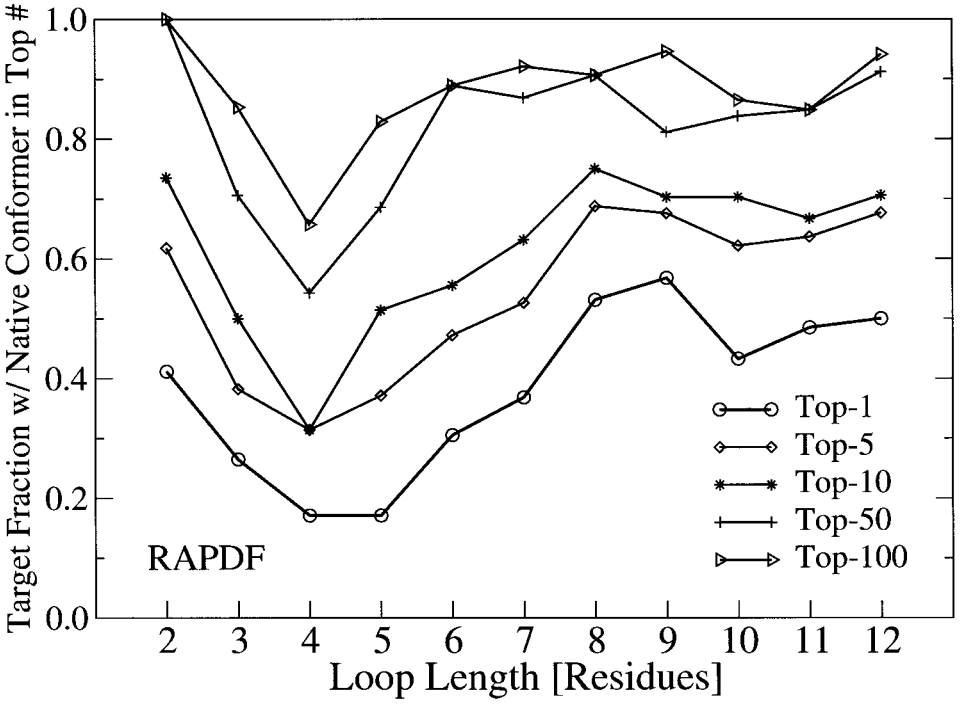

Fig. 5. Ranking of the native conformer with the RAPDF function for all loop lengths displayed as the fraction of loop targets with native ranked in the top 1/5/10/50/100 of the RAPPER decoy ensembles (n = 1000). The lowest curve corresponds to native being ranked with the most favorable score according to the RAPDF function.

gets as a function of the number of conformers with lower RMSDs than the selected one. Given the size of the filtered ensemble (50), an error rate of 25 conformers with better RMSDs is as good as random selection (indicated by the dotted vertical lines). The lower the number of better conformers, the better the overall discriminatory power of the scoring function.

For 4-mer, 8-mer, and 12-mer targets, the leftmost curve corresponds to the AMBER/GBSA potential energy function with a significantly lower error rate than the other scoring functions. Although the SCWRL clash energy, $RMSD_{anchor}$, and RAPDF are roughly equivalent in terms of error rate for the 4-mer targets, the AMBER/GBSA energy was clearly selecting conformers with lower RMSD. The error rate curve is also plotted for the AMBER gas-phase function and shows a lower discriminatory power compared to the AMBER/GBSA function.

For the 8-mer targets, the AMBER/GBSA function was superior to the AMBER gas-phase function and RAPDF, which was in turn better than the $RMSD_{anchor}$ and SCWRL functions. The same conclusion can be made for the 12-mer targets, with the exception that the performance of the RAPDF improved only marginally, considering that for 70% of the targets the AMBER/GBSA function is to be preferred to the other functions.

Note that although these results give valuable insight into the ranking power of the scoring functions, they do not reflect the absolute accuracy of the selected conformer in terms of RMSD to native. Furthermore, one should be aware that these are cumulative plots and that there are quite a few targets where the AMBER/GBSA function performs worse than other functions (see Fig. 3). Nevertheless, the overall performance of the AMBER/GBSA energy function is superior to the other functions, including the AMBER gas-phase function.

Effect of Including Fragments From Experimental Structures

Although the AMBER/GBSA energy function is successful at recognizing the native loop conformation for many targets and also tends to select near-native conformers with low RMSD from the decoy sets, for some of the targets, conformers are still selected with poor RMSDs from the RAPPER ensembles. This severely affects the reported average RMSD of selections. This is due in part to inadequate filtering by the RAPDF and $RMSD_{anchor}$ criteria and likely due to structural approximations from discrete ϕ/ψ states, fixed bond lengths, and angles and a fixed ω dihedral angle. Conformational sampling is also likely to be limited, especially at longer lengths, when ensembles of 1000 conformers are too small to cover phase space adequately and do not contain enough near-native conformers. The aim here was to obtain protein fragments from a knowledge-based source and to add these to the decoy ensembles.

For targets of 3 to 8 residues long, the RAPPER ensembles were supplemented with at most 10 fragments extracted from the protein database on the basis of the target sequence using the program FREAD[42] (see Methods section). We ensured that the FREAD fragments were not

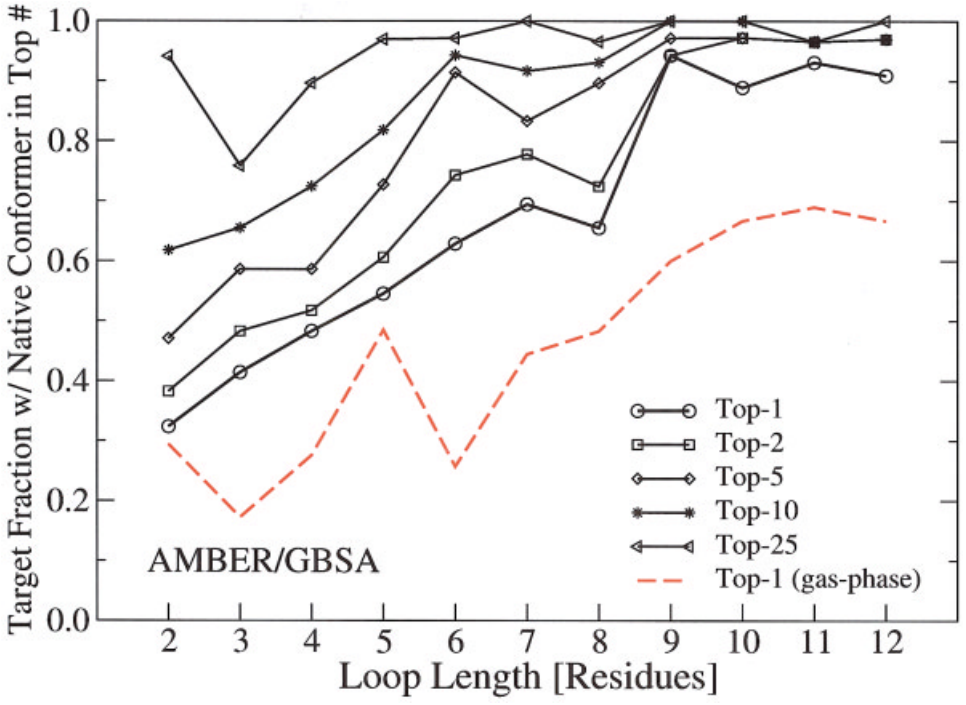

Fig. 6. Ranking of the native conformer with the AMBER/GBSA energy function for all loop lengths displayed as the fraction of loop targets with native ranked in the top 1/2/5/10/25 of the filtered ensembles (n = 50). Also included is the curve corresponding to native ranked as the most favorable conformation (top-1) according to AMBER gas-phase energies (dashed red line). The higher the curve, the better the discriminatory power with respect to native ranking. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

derived from the original native structures nor from close homologs by imposing a maximal global sequence identity cutoff at 80% to the native sequence. It must be emphasized that the sole purpose of this exercise was to ascertain whether the AMBER/GBSA potential energy function was capable of recognizing potential native-like fragments that have been added to the filtered ensemble, regardless of their origin.

The results are summarized in Table IV. The improved performance is directly evident from a quick comparison between the second and third columns which list the average RMSD for AMBER/GBSA selection before (same as in Table III) and after the addition of FREAD fragments. FREAD is in some cases unable to return solutions; the fifth column (No. of Failures) lists the number of targets for which it could not find any fragments. The number of targets for which a FREAD fragment is selected by virtue of having the lowest AMBER/GBSA potential energy of all conformers, and how many times this leads to an actual improvement (reduction) of the RMSD with respect to the native loop, is listed in the "Selected" and "Improved" columns, respectively. The average RMSD change per replaced fragment (ΔRMSD) demonstrates that in the relatively few cases that the FREAD fragment replaced an originally selected RAPPER fragment, the reduction in RMSD was impressive at longer lengths. For example, for the 8-mer targets, the average decrease in global RMSD was 2.3 Å for each FREAD fragment that successfully replaced a poorly selected RAPPER fragment in the original ensemble. At this length, the decrease in average selected RMSD is considerable: the (RMSD) of selected conformers drops to 1.3 Å for all 8-mer targets. In contrast, we note that the average RMSD of selection does not significantly change with respect to the original (RMSD) (Table II) when scoring using the RAPDF function (data not shown).

The average percent sequence identities (PID) are also listed (last two columns) for all FREAD fragments added and for those selected by the AMBER/GBSA function as the best conformer. Clearly, the average PID of about 20% indicates that the added fragments from FREAD are not biased toward fragments derived from homologous proteins. The average PID of 35% for the selected FREAD fragments confirms that the selected fragments do not come from close homologs. We emphasize that the selection procedure is insensitive to the origins (RAPPER or FREAD) of the scored conformers and the real task is to identify these fragments purely on the basis of their energies.

Example 1: AMBER/GBSA Ranking Better Than AMBER and RAPDF

This example is of an 8-mer loop target (sequence: QFSGNYYG) corresponding to residues 72–79 of guanylate kinase from yeast (PDB code 1gky). This loop contains a type I′ β-turn. The conformation selected by the AMBER/GBSA energy function was of excellent quality (global RMSD of the backbone of 0.8 Å) whereas the ones selected

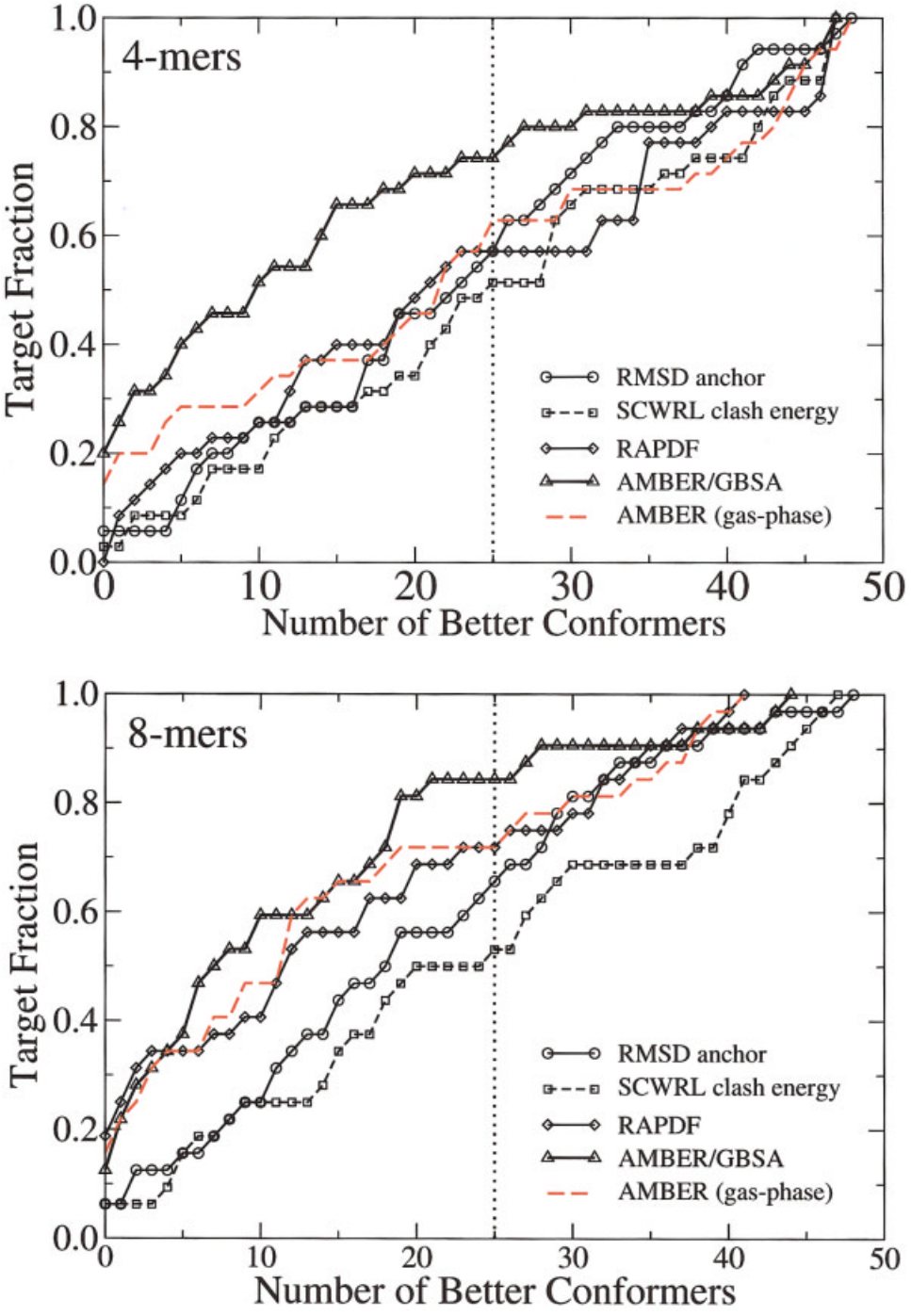

Figure 7. Discrimination curves for different scoring functions for all (i) 4-mer; (ii) 8-mer; and (iii) 12-mer loop targets. Each curve is the cumulative fraction of loop targets as a function of the number of conformers with lower (better) RMSDs than the selected conformer by the scoring function in the filtered ensemble ($n = 50$) to allow a quantitative comparison with the AMBER/GBSA energy function. The optimal curve coincides with the vertical x-axis (no other conformers with lower RMSDs) whereas random selection (no discrimination) corresponds to the dashed vertical line at $x = 25$. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

by the AMBER function in the gas-phase and the RAPDF had much higher RMSD values (5.0 and 3.8 Å, respectively). The β-hairpin conformation in the native (in blue) and selected (in orange) loops are shown in Figure 8. Inclusion of the solvation term appears critical for selection, probably due to many favorable interactions with solvent, as the gas-phase potential energy tends to favor a conformer with increased packing against the framework

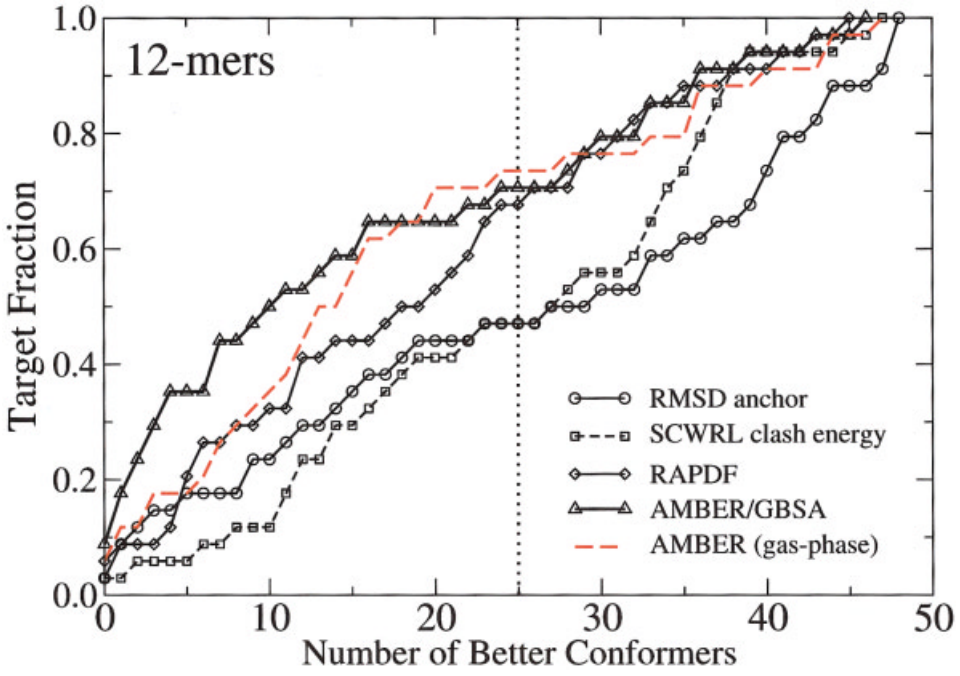

Figure 7. (Continued.)

structure (not shown), which is not observed in the native structure.

When the filtered ensemble was supplemented by fragments collected by the FREAD server, the conformer with the most favorable conformational free energy (after AMBER/GBSA minimization) was a FREAD fragment from aspartic proteinase from *Penicillum janthenellum* (PDB code 1bxo) and had an RMSD of 0.9 Å with respect to the native loop (not shown). Remarkably, the average sequence identity between the target structure and the original structures of the added FREAD fragments was very low (≈12%). Furthermore, the aspartic proteinase was not homologous to the guanylate kinase, which confirmed that selection was entirely the result of superior ranking by the AMBER/GBSA energy function rather than the fortuitous inclusion of a fragment from a close homolog.

The average RMSD of the conformers in the filtered ensemble for this target was 4.3 Å whereas there was only a single conformer with a lower RMSD (0.8 Å) than the selected one. The difference in the calculated conformational free energies of these two low-RMSD conformers was minor (−276.4 vs −282.1 kcal/mol). This highlights the discriminatory power of the AMBER/GBSA energy function: from the accessible conformational space it was able to recognize virtually the best conformation available.

Example 2: Effect of Adding FREAD Fragments

Here we give an example of the 8-mer loop target in thymidylate synthase of *Escherichia coli* (PDB code 1tys), covering residues 83–90 (sequence: WADENGDL) between helix-5 and helix-6 (magenta backbone ribbon in Fig. 9). The selected RAPPER fragment by AMBER/GBSA (yellow ribbon) had a poor RMSD of 5.1 Å whereas the selected RAPPER fragment by AMBER in the gas-phase (green ribbon) had a better but still not very good RMSD of 3.1 Å. From the same ensemble supplemented with nine fragments collected by the FREAD server, we were able to select a conformer of 1.5 Å (blue ribbon) that came from residues 97–104 (sequence: WKQEDGTI) of the thermostable thymidylate synthase homologue of *Bacillus subtilis* (PDB code 1bkp). The two thymidylate synthase structures shared 37% sequence identity whereas the loop sequences had three amino acid residues in common. Ranking was performed solely on the basis of AMBER/GBSA conformational free energy and without explicit knowledge that this FREAD fragment was extracted from a homologue. Moreover, the eight other FREAD fragments included in the filtered ensemble had less favorable conformational free energies and global RMSD values ranging from 4.0 Å to 6.1 Å.

The lowest observed RMSD in the original RAPPER ensemble was 1.34 Å and the average RMSD of all generated fragments was 4.6 Å, which ruled out a fundamental sampling problem in terms of generating near-native backbone conformations. The problem of not selecting such near-native conformations using AMBER/GBSA can be attributed to aberrant filtering, as only three conformers of the filtered ensemble have a global RMSD of below 2 Å.

Accuracy of PETRA Predictions

To compare our results to an existing loop prediction program, we submitted the entire test set to the PETRA web server. The results are summarized in Table V. For all tested loops of three to eight residues long, the global

RMSD of a PETRA prediction was, on average, 1.0 Å worse than the expected RMSD of random selection from the RAPPER generated ensemble (see Table I). The predictions made on the basis of the fit to the anchors (Best-Fit Prediction) were not significantly different from those made on the basis of the knowledge-based potential (Best-Energy Prediction).

DISCUSSION

In this study, we presented a method called RAPPER for efficient conformational searching to generate ensembles of conformations for protein fragments. All generated conformers followed three fundamental rules to guarantee local structural health: Engh and Huber geometry, residue-specific ϕ/ψ propensities, and excluded volume. These rules can be regarded as negative filters that specify which conformers are forbidden. The generated ensembles effectively map out the conformational freedom available to the loop in the context of a fixed protein structure, and contain both near-native and non-native fragments.

We stress that by deconvoluting the conformational search from the selection process, the conformers in the ensemble are guaranteed to be unbiased with respect to the scoring functions. Because the decoy ensembles are identical for all scoring functions, resulting differences in discrimination performance are thus solely attributable to inadequacies of the functions.

We introduced expectation baselines for random selection (corresponding to the average RMSD of all decoys) and for optimal selection (corresponding to the best conformer with the lowest RMSD among the decoys), yielding a window of opportunity within which the scoring function has to present its discriminatory power. The RAPDF statistical potentials performed as well as or for 7- to 10-residue loops, only slightly better than the goodness-of-fit at the C-terminal anchor residue as a function for selection. The anchor goodness-of-fit performed consistently better than a random selection strategy. The existence of many conformers in the ensemble that are actually lower in RMSD to native than the selected fragment and the fact that for so many targets the native conformer did not have the most favorable score, points to the rather limited discriminatory power of the RAPDF function. These results are in keeping with the observed "lack of universality" of statistical potentials for recognition of native folds[44] and a more recent claim that it is impossible to find a pair potential (with a 9 Å distance cutoff) by linear optimization that recognizes all native conformations from a large set of decoys.[45] Context-dependent features (like

TABLE III. Selection Accuracy Given as the Average Global (RMSD) (Å) of Selected Conformers for All Loop Targets of a Given Length on the Basis of Energy Minimization with the AMBER force field (A) including the GB/SA Solvation model and (B) in Vacuum

Selection accuracies by AMBER force field					
		(RMSD) (Å)			
		AMBER/GBSA		AMBER gas-phase	
Loop length	n	Initial	Minimized	Initial	Minimized
2	34	0.49	0.35	0.61	0.50
3	34	0.49	0.37	0.53	0.44
4	35	0.70	0.47	0.83	0.72
5	35	1.02	0.90	1.15	1.05
6	36	1.19	0.95	1.60	1.39
7	38	1.57	1.37	1.60	1.40
8	32	2.38	2.28	2.64	2.54
9	37	2.44	2.41	3.25	3.25
10	37	3.54	3.48	4.19	4.16
11	33	4.93	4.94	5.01	5.05
12	34	5.00	4.99	5.42	5.41

The initial RMSD values refer to the conformers before minimization.

TABLE IV. Selection Accuracy for FREAD-Supplemented Ensembles Given as the Average Global RMSD (Å) of Selected Conformers for Loop Targets of Three to Eight Residues on the Basis of Energy Minimization with the AMBER Force Field, Including the GB/SA Solvation Model

Effect of supplementing ensembles with real protein fragments									
	(RMSD)							(PID) (%)	
Loop length	Original ensemble	Supplemented by FREAD	N[a]	#Failures[b]	#Selected[c]	#Improved[d]	Δ(RMSD)	Selected[e]	Supplemented[f]
3	0.37	0.35	34	3	13	8	−0.1	32	19
4	0.47	0.47	35	3	7	4	−0.0	28	23
5	0.90	0.81	35	3	8	6	−0.4	37	19
6	0.95	0.74	36	4	10	8	−0.8	44	20
7	1.37	0.82	38	3	14	14	−1.5	34	22
8	2.28	1.26	32	2	14	12	−2.3	41	21

RMSD values are given for minimized conformers.
[a]Refers to the number of loop targets.
[b]Number of targets for which FREAD does not find any solutions.
[c]Number of targets for which a FREAD fragment is selected by AMBER/GBSA.
[d]Number of targets when this involves a reduction of the RMSD (i.e. improved selection). The Δ(RMSD) is the average change in RMSD per selected FREAD fragment to effect the observed reduction in (RMSD). PID refers to percentage sequence identity of the original structure of the FREAD fragment to the loop target structure.
[e]Average PID of the selected FREAD fragments.
[f]All added FREAD fragments.

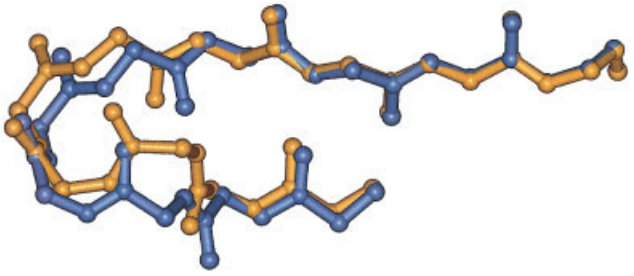

Fig. 8. Ball-and-stick representation of the 8-mer loop target 1gky:72–79, with the native conformation displayed in blue and the selected RAPPER conformer by AMBER/GBSA in orange. The global RMSD of the RAPPER conformer is 0.8 Å for all heavy backbone atoms.

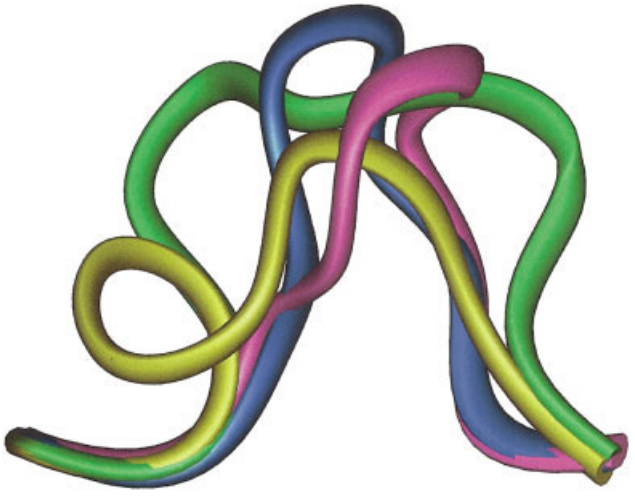

Fig. 9. Solid ribbon representation of the 8-mer loop target 1tys:83–90, with the native conformation displayed in blue and the selected RAPPER conformer by AMBER/GBSA in yellow, the selected RAPPER conformer by AMBER (gas-phase) in green, and the selected FREAD conformer by AMBER/GBSA in magenta. The FREAD conformer has a global RMSD of 0.9 Å for all heavy backbone atoms.

the local environment of the loops dictated by protein and solvent atoms) are perhaps too important to be treated in an implicit, average manner inherent to the formulation of statistical potentials in general.

Interestingly, no significant difference was observed in selection accuracy between the two variants of the RAPDF tested here, namely scop-e4-allatoms-xray-scores and scop-150-unique-xray-scores. The former is derived from a larger set of 1953 SCOP[46] domains, whereas the latter is derived from only a subset of these (488 domains in total) with the addition of four protein domains not included in the longer list (Ram Samudrala, personal communication). This illustrates the robust nature of knowledge-based interaction potentials, given differing sets of protein structures for their derivation. Given the rapid increase of high-quality structures in the PDB over the last years due to structural genomics projects, it will be interesting to see whether more recent versions of the RAPDF can give improved performance.

Filtering conformers in the ensembles by applying a combination of the RAPDF function and the anchor goodness-of-fit reduces the number of conformers in the ensemble that can be scored using the AMBER/GBSA force field. The average global RMSD values of all heavy backbone atoms are 0.5 Å for 4-mers, 2.3 Å for 8-mers, and 5.0 Å for 12-mers. When fragments derived from experimental structures are added to the ensemble, these values improve markedly at longer lengths: The average global RMSD is 1.3 Å for 8-mers. Although the minimization protocol is primarily used to arrive at meaningful potential energies so that conformers can be ranked, it is noteworthy that conformers, on average, move consistently toward the native conformation with a concomitant decrease in their RMSD values.

The above findings support the widely held belief that statistical potentials are less sensitive to local structural perturbations (RAPDF uses 1 Å bins) and less accurate than molecular mechanics force fields.[20] Because single-point energy calculations are cheap to compute, this study supports a rational basis for a two-step scoring scheme consisting of the RAPDF statistical potential as a filter and the AMBER/GBSA force field for accurate decoy discrimination. Nevertheless, provided sufficient computing resources, it will undoubtedly be better to bypass the RAPDF altogether and evaluate all conformers in the ensemble directly using the AMBER/GBSA force field. In a comparative modelling context, however, the success of this approach will largely depend on how models are generated (in terms of bad interactions) and on the extent and nature of such non-native interactions.

We observed that the selection accuracy (RMSD) follows by no means a normal distribution. Many selected conformers of all loop lengths had very low RMSD values whereas few others had very high RMSDs, thereby skewing the average of the entire distribution. This is illustrated by the enormous improvement in the (RMSD) when for only few targets near-native FREAD fragments are selected (see Table IV). In some cases this is due to aberrant filtering by the RAPDF scores and in others due to the inadequacy of the energy function.

Does Conformational Free Energy Distinguish Loop Conformations?

Our results with AMBER/GBSA and AMBER gas-phase calculations (see Fig. 3) are in agreement with Pellequer and Chen,[47] who concluded that, in their words, "the solvation energy term plays a confusing role, sometimes discriminating reference-like (. . .) geometry and many times allowing non-reference-like conformations to have the lowest conformational free energies." In their analysis, candidate conformations for antibody hypervariable loops were evaluated using the CHARMM potential together with the solvation free energy based on the Poisson–Boltzmann equation for the electrostatics term and a surface area term for the nonpolar term. We stress that, although the spread in Figure 3 is certainly remarkable, overall selective performance observed in the present study is worse when the solvation term is omitted. Nevertheless, it is likely that the conformation of some classes of loops is predominantly determined by intramolecular inter-

TABLE V. Performance of Loop Prediction Given as Average Global (RMSD) (Å) by PETRA for all 3-mer to 8-mer Targets in the Test Set.

Prediction accuracies by PETRA				
		(RMSD) (Å)		
Loop length	Published test set[a]	Best energy[b] prediction	Best fit[b] prediction	Average[c]
3	1.2	1.63	1.48	1.56
4	1.9	2.61	2.74	2.67
5	2.6	3.84	3.65	3.74
6	2.9	4.26	4.47	4.37
7	3.5	5.22	4.93	5.08
8	3.8	5.56	5.56	5.56

[a]Average global RMSD values for predictions made on the original benchmark set.[41]
[b]Conformations with the best RMSD fit to the anchor regions and lowest knowledge-based energy as calculated by PETRA.
[c]Average RMSD of the two predictions. The listed RMSD values (relative to the native structure) are taken over all heavy backbone atoms of the loop fragment.

actions so that the inclusion of the solvation free energy gives ambiguous results when ranking loop conformations. A more detailed look into the actual local structural environment is warranted in order to understand this dichotomy.

How Much Structural Accuracy in Representation Is Needed?

We have demonstrated that the prediction results from the PETRA server for the entire test set are on average much worse than the random selection baseline. The ab initio method called PETRA[41] was developed for the prediction of peptide fragments of three to eight amino-acid residues long. The PETRA algorithm relies on a precomputed database encoding all possible self-avoiding conformations, assuming a representative set of eight ϕ/ψ dihedral angle pairs, which are optimized for conformational coverage of protein structures. Loop conformations in PETRA are selected on the basis of the goodness-of-fit at the anchor residues and the Samudrala–Moult statistical potential. We noticed that the prediction results for the current loop test set were much worse than those for the original benchmark set, suggesting that the current Fiser et al.[38] test set represents a more stringent benchmark (and is perhaps an atypical one).

As pointed out in the original PETRA work, there is a general lack of correlation between the statistical potential score of generated conformers and the global RMSD to native below a certain level. Furthermore, it is observed that if the statistical potential is not applied as a cutoff filter, the prediction results become worse. Given our comparative success with Samudrala–Moult RAPDF selection, it is very likely that because of the compromised accuracy of representation and the absence of side-chain atoms beyond Cβ, near-native conformations may not be favored with respect to alternative conformations resulting from a negligible signal. This is consistent with the observation that the accuracy of selection using the Samudrala–Moult RVPDF (with one virtual atom per residue) is much lower than that of the all-atom RAPDF function.[30] Given that most of the specificity of the statistical potential comes from atom–atom interactions between side chains, the lack of correlation between RMSD and the RAPDF in PETRA is hardly surprising. Hence, we suggest that the reason for the poor performance of PETRA in prediction is 2-fold: 1) we have demonstrated elsewhere[34] that near-native conformers have, on average, much higher RMSD values than when using a fine-grained ϕ/ψ set like in RAPPER (in other words, there are fewer native-like conformers); 2) the lack of side chain atoms results in the inability of the scoring function to select near-native conformations.

A related issue is the accuracy of the representation of side chain atoms. The current RAPPER program does not explicitly model side chain atoms but instead relies on the SCWRL method for assigning side chain rotamers to residues of the built fragments. Although more recent sidechain rotamer libraries[48] are likely to be of value, they do not circumvent the general problem of disallowing a rotamer state due to a bad clash that may be relieved by only subtle local readjustments. Especially when the local environment is closely packed so that viable side chain orientations are sparse due to limited freedom, this is likely to become a problem. In other words, single conformations representing each rotamer state may well not be sufficient for accurate modeling.

It is clear that some loops in proteins contain features that are not straightforward to model because of deviations from idealized geometry and the presence of *cis* peptide bonds. We note that 17 targets of the test set have *cis* peptide bonds and that the selected RAPPER models for these targets (by RAPDF and AMBER/GBSA) have substantially worse RMSD values (data not shown) than the average RMSD as shown in Figure 1. For one 11-mer target containing a *cis* peptide bond (1phf:100-110), the lowest RMSD in the generated RAPPER ensemble is roughly as high as the average RMSD (around 7 Å), which indicates that none of the generated models (with non-native *trans* peptide bonds) are sufficiently close to native. This example highlights that the ability to model *cis* peptide bonds is crucial for accurate modelling and motivates us to implement this into the RAPPER program.

Limitations of Conformational Sampling in RAPPER

As demonstrated elsewhere,[34] for all loops of up to 10 residues, the vast majority of conformers in the generated ensembles ($n = 1000$) have a local RMSD of <1 Å, a cutoff defined by Fidelis et al.[49] to indicate that the conformer can be considered near-native. For longer loops, however, the conformational sampling starts to become less effective and near-native conformers are no longer guaranteed to be generated given the fixed size of the ensemble. The size of 1000 conformers is probably too limited to provide conformational coverage at longer lengths sufficiently well. We are currently looking into other discrete sampling

methods that allow us to sample longer lengths with greater efficiency and effectiveness.

The existence of a lower bound on the best attainable RMSD of a RAPPER fragment (the optimal baseline) for a given target conformation[34] provides a rationale for adding fragments from external sources like FREAD. For well-defined local structures, loops can be detected by sequence signals and this is where knowledge-based methods are likely to prove their utility despite pessimistic predictions that such methods do not provide adequate conformational coverage, and probably never will, especially at longer loop lengths.[49]

Protein loop prediction

One of the obvious applications of the RAPPER method of conformational sampling and selection is the prediction of protein loop structures or structurally variable regions. We have reported a significant advance both in conformational sampling[34] and in the selection accuracy of near-native loop conformations compared to the PETRA prediction method. Although it must be stressed that loop modeling on real comparative models is a more complex problem than modeling loops on native structures, the steps of conformation generation and decoy discrimination need to be addressed separately before we are in a position to assess the selection of loop decoy conformers on approximate protein models. Although practical loop prediction is beyond the scope of this work, we find that our results are comparable with the accuracies reported by the approach described by Fiser et al.,[38] which is integrated into the MODELLER package.[50] From 500 independent optimizations using molecular dynamics with simulated annealing, they obtain an average global RMSD of 0.9 Å for 4-mers, 1.8 Å for 8-mers, and 3.9 Å for 12-mers (using their ad hoc conversion factor of 1.5 to calculate global RMSD values from local RMSD values). Starting from initial random conformations for the atoms of the target loop, MODELLER performs a continuous optimization in Cartesian space with respect to the objective function which includes bonded energy terms from the CHARMM-22 force field and statistical potentials describing intramolecular main chain and side chain dihedral propensities and non-bonded atom–atom interactions.

The selection accuracy of the all-atom RAPDF statistical potential is not that much worse than the level of accuracy obtained by Fiser et al.,[32] perhaps somewhat surprising considering that conformers are not explicitly minimized with respect to the RAPDF function in our method. The net gain in selection accuracy by MODELLER might be due to either the explicit minimization of conformers or an intrinsic difference between their energy function and the RAPDFs tested here in terms of discriminatory power. Regardless, because all-atom statistical potentials and molecular mechanics potential energy functions like AMBER (even including the relatively expensive GBSA solvation term) both require $O(N^2)$ time, there are few grounds to perform minimization of an all-atom statistical potential, given the greater accuracy of the molecular mechanics force field. Additional algorithmic improvements, such as the frozen atom approximation in the calculation of the Generalized Born solvation free energy,[51] may alleviate some of the computational burden that currently prevents us from scoring entire ensembles of conformations.

CONCLUSIONS

We presented a statistical exploration of the discriminatory power of an all-atom statistical potential and the AMBER molecular mechanics force field, including the GBSA solvation model, over a wide range of loop structures where the native framework is used. Neither the goodness-of-fit at the anchor residue ($RMSD_{anchor}$) nor the residual clash energy calculated by SCWRL can reliably be used for quantitative ranking of conformers. The selection accuracy of the RAPDF statistical potentials is better, and thanks to the cheap nature of calculating scores for individual conformations, these potentials can be used as a first-pass filter to reduce the decoy ensemble size to a more manageable number. The AMBER molecular mechanics force field, including the Generalized Born solvation model produces excellent energetic rankings. Selected conformers with the most favorable conformational free energies have average global RMSD values of 2.3 Å for 8-residue loops. By adding conformers extracted from experimental structures to the ensembles, the selection accuracy improves dramatically, as the average global RMSD value drops to 1.3 Å for 8-residue loops.

The inclusion of the GBSA solvation term to the potential function contributes significantly to the observed discriminatory power in terms of the average RMSD of the selected conformers at all loop lengths. Nevertheless, in many cases omitting the solvation term results in the selection of conformers with lower RMSD than when the GBSA term is included. We have to remind ourselves that that GBSA itself is a useful simplification of the continuum treatment and that specific interactions between protein atoms and solvent molecules are not accounted for, when such effects might actually play a role in loop stability. Perhaps in those cases the simplification offered by GBSA breaks down, and more accurate schemes are required to arrive at more realistic estimates of solvation free energies.

In the context of the current study, overall improvement in selection accuracy can be expected if scoring functions in general are better trained to discard bad structures (negative filtering) rather than try to recognize the best ones (positive selection). Computer programs such as RAPPER that build ensembles of protein-like conformations in all-atom detail will be extremely useful in the pursuit of developing better discriminatory functions.

ACKNOWLEDGMENTS

PDB thanks Charlotte Deane and Hampapathalu Nagarajaram for their help during the initial stages of the project and Dr. Ram Samudrala for supplying the RAPDF functions and Dr. Andras Fiser for supplying the benchmark test set. The authors are grateful to the Cambridge European Trust, Isaac Newton Trust, NUFFIC Talentenprogramma and BBSRC (to PDB), and the Marshall Aid Commemoration Commission (to MDP) for financial support.

REFERENCES

1. Moult J. Comparison of database potentials and molecular mechanics force fields. Curr Opin Struct Biol 1997;7:194–199.
2. Lazaridis T, Karplus M. Effective energy functions for protein structure prediction. Curr Opin Struct Biol 2000;10:139–145.
3. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM Jr, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A second generation force field for the simulation of proteins and nucleic acids. J Am Chem Soc 1995;117:5179–5197.
4. MacKerell AD Jr, Bashford D, Bellott M, Dunbrack RL Jr., Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reither III WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 1998;102:3586–3616.
5. Novotny J, Bruccoleri R, Karplus M. An analysis of incorrectly folded protein models. Implications for structure predictions. J Mol Biol 1984;177:787–818.
6. Novotny J, Rashin AA, Bruccoleri RE. Criteria that discriminate between native proteins and incorrectly folded models. Proteins 1988;4:19–30.
7. Smith PE, Pettitt BM. Peptides in ionic solutions: A comparison of the Ewald and switching function techniques. J Chem Phys 1991;95:8430–8441.
8. Hünenberger PH, McCammon JA. Effect of artificial periodicity in simulations of biomolecules under Ewald boundary conditions: A continuum electrostatics study. Biophys Chem 1999;78:69–88.
9. Honig B, Nicholls A. Classical electrostatics in biology and chemistry. Science 1995;268:1144–1149.
10. Warshel A, Papazyan A. Electrostatic effects in macromolecules: Fundamental concepts and practical modeling. Curr Opin Struct Biol 1998;8:211–217.
11. Bashford D, Case DA. Generalized Born models of macromolecular solvation effects. Annu Rev Phys Chem 2000;51:129–152.
12. Smith KC, Honig B. Evaluation of the conformational free energies of loops in proteins. Proteins 1994;18:119–132.
13. Rapp CS, Friesner RA. Prediction of loop geometries using a Generalized Born model of solvation effects. Proteins 1999;35:173–183.
14. Still WC, Tempczyk A, Hawley RC, Hendrickson T. Semianalytical treatment of solvation for molecular mechanics and dynamics. J Am Chem Soc 1990;112:6127–6129.
15. Qiu D, Shenkin PS, Hollinger FP, Still WC. The GB/SA continuum model for solvation. A fast analytical method for the calculation of approximate Born radii. J Phys Chem A 1997;101:3005–3014.
16. Holm L, Sander C. Evaluation of protein models by atomic solvation preference. J Mol Biol 1992;225:93–105.
17. Samudrala R, Levitt M. Decoys 'R' Us: A database of incorrect conformations to improve protein structure prediction. Protein Sci 2000;9:1399–1401.
18. Vorobjev YN, Almagro JC, Hermans J. Discrimination between native and intentionally misfolded conformations of proteins: ES/IS, a new method for calculating conformational free energy that uses both dynamics simulations with an explicit solvent and an implicit solvent continuum model. Proteins 1998;32:399–413.
19. Lazaridis T, Karplus M. Effective energy function for proteins in solution. Proteins 1999;35:133–152.
20. Lazaridis T, Karplus M. Discrimination of the native from misfolded protein models with an energy function including implicit solvation. J Mol Biol 1999;288:477–487.
21. Dominy BN, Brooks CL III. Identifying native-like protein structures using physics-based potentials. J Comput Chem 2002;23:147–160.
22. Thomas PD, Dill KA. Statistical potentials extracted from protein structures: How accurate are they? J Mol Biol 1996;257:457–469.
23. Ben-Naim A. Statistical potentials extracted from protein structures: Are these meaningful potentials? J Chem Phys 1997;107:3698–3706.
24. Jones DT, Tress M, Bryson K, Hadley C. Successful recognition of protein folds using threading methods biased by sequence similarity and predicted secondary structure. Proteins 1999;S3:104–111.
25. Domingues FS, Koppensteiner WA, Jaritz M, Prlic A, Weichenberger C, Wiederstein M, Floeckner H, Lackner P, Sippl MJ. Sustained performance of knowledge-based potentials in fold recognition. Proteins 1999;S3:112–120.
26. Simons KT, Ruczinski I, Kooperberg C, Fox BA, Bystroff C, Baker D. Improved recognition of native-like protein structures using a combination of sequence-dependent and sequence-independent features of proteins. Proteins 1999;34:82–95.
27. Samudrala R, Xia Y, Huang E, Levitt M. Ab initio protein structure prediction using a combined hierarchical approach. Proteins 1999;S3:194–198.
28. Robert CH, Janin J. A soft, mean-field potential derived from crystal contacts for predicting protein-protein interactions. J Mol Biol 1998;283:1037–47.
29. Melo F, Feytmans E. Novel knowledge-based mean force potential at atomic level. J Mol Biol 1997;267:207–222.
30. Samudrala R, Moult J. An all-atom distance-dependent conditional probability discriminatory function for protein structure prediction. J Mol Biol 1998;275:895–916.
31. Rojnuckarin A, Subramaniam S. Knowledge-based interaction potentials for proteins. Proteins 1999;36:54–67.
32. Sippl MJ. Calculation of conformational ensembles from potentials of mean force. An approach to the knowledge-based prediction of local structures in globular proteins. J Mol Biol 1990;213:859–883.
33. Melo F, Feytmans E. Assessing protein structures with a non-local atomic interaction energy. J Mol Biol 1998;277:1141–1152.
34. DePristo MA, de Bakker PIW, Lovell SC, Blundell TL. Ab initio construction of polypeptide fragments: Efficient generation of accurate, representative ensembles. Proteins 2003;51:41–55.
35. Bower MJ, Cohen FE, Dunbrack RL Jr. Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: A new homology modeling tool. J Mol Biol 1997;267:1268–1282.
36. Engh RA, Huber R. Accurate bond and angle parameters for X-ray protein-structure refinement. Acta Crystallogr A 1991;47:392–400.
37. Word JM, Lovell SC, LaBean TH, Taylor HC, Zalis ME, Presley BK, Richardson JS, Richardson DC. Visualizing and quantifying molecular goodness-of-fit: small-probe contact dots with explicit hydrogen atoms. J Mol Biol 1999;285:1711–1733.
38. Fiser A, Do RKG, Šali A. Modeling of loops in protein structures. Protein Sci 2000;9:1753–1773.
39. Hooft RWW, Vriend G, Sander C, Abola EE. Errors in protein structures. Nature 1996;381:272.
40. Lovell SC, Davis IW, Arendall WB III, de Bakker PIW, Word JM, Prisant MG, Richardson JS, Richardson DC. Structure validation by Cα geometry: ϕ/ψ and Cβ deviation. Proteins 2002;50:437–450.
41. Deane CM, Blundell TL. A novel exhaustive search algorithm for predicting the conformation of polypeptide segments in proteins. Proteins 2000;40:135–144.
42. Deane CM, Blundell TL. CODA: a combined algorithm for predicting the structurally variable regions of protein models. Protein Sci 2001;10:599–612.
43. Topham CM, McLeod A, Eisenmenger F, Overington JP, Johnson MS, Blundell TL. Fragment ranking in modelling of protein structure. Conformationally constrained environmental amino acid substitution tables. J Mol Biol 1993;229:194–220.
44. Rooman M, Gilis D. Different derivations of knowledge-based potentials and analysis of their robustness and context-dependent predictive power. Eur J Biochem 1998;254:135–143.
45. Tobi D, Elber R. Distance-dependent, pair potential for protein folding: Results from linear optimization. Proteins 2000;41:40–46.
46. Murzin AG, Brenner SE, Hubbard T, Chothia C. SCOP: a structural classification of proteins database for the investigation of sequences and structures. J Mol Biol 1995;247:536–540.
47. Pellequer JL, Chen SW. Does conformational free energy distinguish loop conformations in proteins? Biophys J 1997;73:2359–2375.
48. Lovell SC, Word JM, Richardson JS, Richardson DC. The penultimate rotamer library. Proteins 2000;40:389–408.
49. Fidelis K, Stern PS, Bacon D, Moult J. Comparison of systematic search and database methods for constructing segments of protein structure. Prot Eng 1994;7:953–960.
50. Šali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993;234:779–815.
51. Guvench O, Weiser J, Shenkin P, Kolossvary I, Still WC. Application of the frozen atom approximation to the GB/SA continuum model for solvation free energy. J Comput Chem 2002;23:214–221.