Advanced Review

Optimization of protein models

Dominik Gront*, Sebastian Kmiecik, Maciej Blaszczyk, Dariusz Ekonomiuk and Andrzej Koliński

Protein structure predictions, and experimentally derived protein structures, very often require certain structure improvement (refinement), which means bringing it closer to real, usually *in vivo* working conformations. In respect to the variety of protein models to be refined, computational optimization procedures could be divided into localized (applied to a small part of a structure) and global (whole structure). Generally speaking, the first problem is usually tractable, and the latter remains to be extremely challenging for systems larger then peptides or small proteins: optimization complexity and difficulty dramatically increase with the size of structures to be optimized.

How to cite this article:
WIREs Comput Mol Sci 2012, 2: 479–493 doi: 10.1002/wcms.1090

INTRODUCTION

Since the past few decades, structural refinement has been considered the last milestone in the protein structure prediction problem. The refinement (optimization) of predicted models might sound similar to other refinement protocols, such as those for X-ray and nuclear magnetic resonance (NMR) protein structure determination. The refinement tailored to predictions is, however, significantly different from its counterparts related to structure determination with experimental techniques. Crystallographers start their refinement from an all-atom initial conformation and attempt to refine phases and local conformational details, which lead to an improved model. This continues until the correlation between the diffraction data and the model is maximized. Similarly, NMR refinement starts from an atomic model and experimental data, seeking for a set of coordinates that better explains the measured observables. In both cases, an all-atom structure is refined based on true experimental data, usually fraught with experimental errors. To the contrary, in protein structure prediction refinement starts usually with a large number of models, often in very simplistic representation, being aimed at the following three objectives simultaneously: (1) to rebuild atomic details from a reduced representation, (2) to improve the shortcomings of a model, and (3) to select the final structure out of a (usually) very large pool of proposed models. Because of the broad character of these objectives, the refinement of protein models is a very difficult task, sometimes being close to the impossible.

The need for the refinement of protein predictions comes from the demand for three-dimensional (3D) protein structures. Knowledge of protein structures provides invaluable insights into the molecular basis of the machinery of life. Although spectacular progress has been made in structure determination using experimental methods, such as already mentioned X-ray crystallography and NMR spectroscopy, such methods are still very expensive, time consuming, and require highly qualified personnel. Since the very first protein structure of myoglobin was elucidated in 1960,[1] combined experimental efforts resulted in nearly 80,000 protein structures deposited in the Protein Data Bank[2] (as of June 2011). At the same time, the size of the known protein sequence universe counts in millions, with more than two million available in the release of 2011_05 UniProt[3] in May 2011. More importantly, the growth of the known sequence space is much faster than for the structure space, and the gap between the two is continuously widening. In the near future, computational methods are the only hope to bridge the gap by providing theoretical models for at least a part of the sequence data. Better optimization methods for computed models toward high-resolution structures seem to be the only way to make the computational approaches reliable

*Correspondence to: dgront@chem.uw.edu.pl
Laboratory of Theory of Biopolymers, Faculty of Chemistry, University of Warsaw, Warsaw, Poland

DOI: 10.1002/wcms.1090

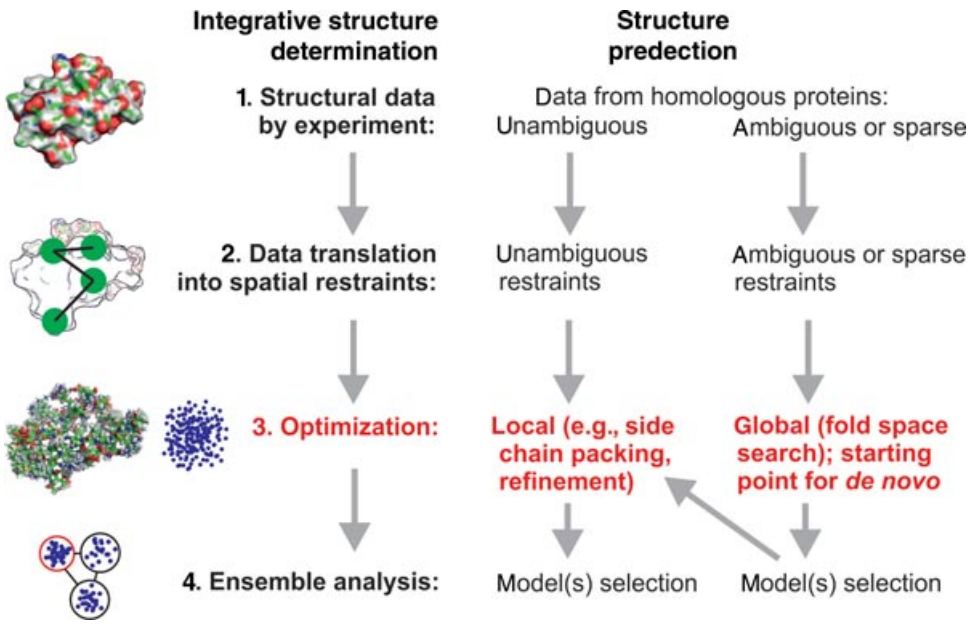

FIGURE 1 | Stages of integrative structure determination and analogous structure prediction. Structure determination by integrating varied data from experiments and modeling can be divided into four steps[12]: (1) generation of structural data by experiment, (2) data translation into spatial restraints, (3) optimization, and (4) ensemble analysis. These four steps are also characteristic for structure prediction, shown here for easy modeling cases (unambiguous data derived from homologous proteins) and difficult ones (ambiguous or sparse data) embedding fold recognition methods, together with *de novo* modeling.

and accurate. Model optimization, being the last stage of the theoretical protocols, is the final opportunity to improve model's quality and, hence, its value for a biologist. Quality of models and their potential applications (such as functional site detection[4–6] or molecular replacement[7]) as well as the state of the art of structure modeling, in general, were subject of excellent reviews, for example, by Schwede et al.,[8] Sanchez et al.,[9] Tramontano,[10] and Baker and Sali.[11]

Another important application of protein structure optimization methods is integrative structure determination, combining experimental information from varied sources. Recently, such approaches enabled the generation of atomic models of previously intractable large protein assemblies.[12] Some of these achievements have been possible due to the recently emerging experimental techniques, such as mass spectrometry of complexes and single-particle cryo-electron microscopy, and due to the availability of high-resolution structural data for individual subunits.[12,13] In general, integrative structure determination is an analogous task to protein structure prediction and similarly requires intervention of structure optimization methods as outlined in Figure 1. Apart from the possible role in integrating various experimental data, computational structure optimization has been recently used in experimental-based protocols, such as the most widespread X-ray crystallography. As demonstrated recently, X-ray structure determination can be improved, or made possible, by combining crystallographic map interpretation tools with structure prediction methodology.[14]

STRUCTURE PREDICTION METHODS AND STRUCTURAL ERROR THEY INTRODUCE

The computational methods for protein structure modeling may be divided into two very broad classes. Comparative modeling approaches (usually preceded by fold recognition procedures) use a template structure (or structures) of an already experimentally annotated protein (or proteins) as a scaffold (scaffolds) for model building. These methods can only be employed when a detectable template of a known structure is available. When such a template or templates cannot be found, only a *de novo* method can be applied. It should, however, be noted that almost all of the current prediction methods, and template-free approaches in particular, utilize known protein structures to train machine learning classifiers or to derive knowledge-based scoring functions. The commonly accepted distinction between the template-based and

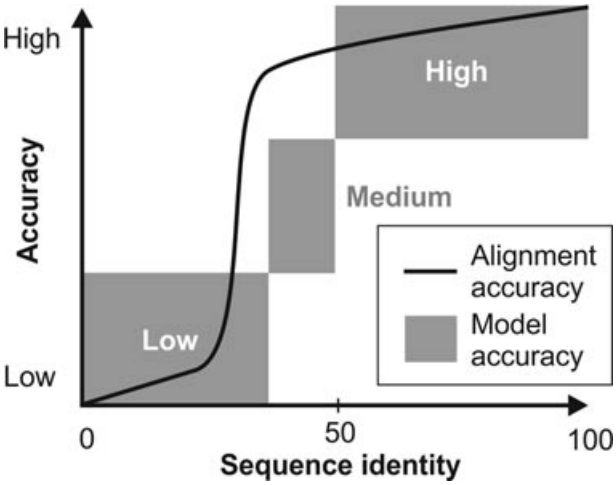

FIGURE 2 | Approximate dependence of sequence alignment accuracy and expected models accuracy on the percentage of sequence identity. The classification of model accuracy (see Table 1) practically agrees with alignment accuracy, which falls into one of the following three zones of sequence similarity (defined by Rost[15]): safe, twilight, and midnight zone. The twilight zone denotes a huge drop in alignment accuracy (roughly in the range of 25%–30% of sequence identity) from the safe zone (high level of sequence similarity—proteins also have similar structures and functions) to the midnight zone (low level of sequence similarity—protein similarity cannot be detected from sequence comparisons alone).

the template-free methods is based on the use of a particular structure (or a few of them) for gathering information about the overall fold, assignment of secondary structure elements, their mutual orientation in space, and so on. The modeling process is usually guided by alignment, which is a correspondence between the two sequences (a target and a template), and consequently defines the expected similarity of their structures. The accuracy of the final model critically depends on alignment accuracy, whereas alignment accuracy strongly depends on the sequence similarity between the target and the template (see Figure 2). Typically, a number of homology search tools, such as PSI-BLAST,[16] profile-to-profile alignments,[17,18] or Hidden Markov Models,[19] could be used for the identification of proteins of known structure which are evolutionary related to the query sequence. Depending on the sequence similarity between the query and the templates, as well as the alignment coverage, the modeling process branches either into comparative modeling or into *de novo* prediction (see Figure 1). The alignment may be considered a recipe that defines which parts of the template structure may be directly used to construct some parts of the target. Remaining fragments of the target (usually loops) have to be modeled explicitly. The amino acid sequences of the structural pieces taken from a template frequently differ from the respective sequence of a target; therefore, residue side chains have to be reconstructed and repacked. The optimization of side-chain conformations is a very important and common application of the structure refinement methods.

In contrast to comparative modeling, *de novo* methods do not assume any knowledge about the architecture of a target. Therefore, it is expected that these methods will sample a substantial number of different protein topologies (global optimization, see Figure 1) and that the proper one will be selected based on energy criteria. In practice, these two goals are, to a large extent, mutually exclusive. To effectively sample the immense conformational space of a polypeptide chain, coarse-grained representations are commonly used to reduce the number of degrees of freedom treated in an explicit way and to flatten the energy landscape. Usually, such coarse-grained force fields are not accurate enough to distinguish between alternative protein topologies and will most likely fail in the selection of the correct one. In such cases, the common approach is to convert reduced space models into the all-atom representation, followed by their refinement and rescoring.[20] At the reconstruction step, where the missing atoms are added, the number of degrees of freedom necessary to represent the modeled system may increase several times. Resulting energetic end entropic barriers impose a considerable challenge for sampling methods and require large computational efforts.

In principle, errors can occur at each step of the modeling process, even in very easy modeling cases when a high-resolution model is expected of a similar accuracy to experimentally derived structures (for a short outline of error sources, see Table 1). Template-based methods rely upon a template structure which occasionally may be wrongly assigned. Errors of this kind are almost impossible to fix, especially if the selected template is qualitatively different from the correct one. Even when the template has been identified correctly, but it does not share sufficient sequence similarity to the target sequence, alignment between the two sequences is a nontrivial task. In such cases, model structure refinement should be applied already at this stage. Small errors in alignment, such as a shift of a few residues, may be tolerable for some methods, depending on how structural information is incorporated in the modeling process. The most straightforward approach—just copying the appropriate structural fragments from a template into the model—is the most error prone. A more error-tolerant

TABLE 1 | Application of Protein Models According to Their Accuracy[8,11,21] (See also Figure 2 for the Classification of Model Accuracy According to Sequence Identity)

Model Accuracy (Sources of Errors)	Model Applications
High: Target-template sequence alignment higher than 50% sequence identity (about 1 RMSD for the main chain: mistakes in side-chain packing, small shifts of the core main chain regions, incorrect loop reconstruction). Medium: Target-template sequence alignment between 30% and 50% sequence identity (about 90% of the main chain modeled with 1.5 Å RMSD error, alignment: small errors in nonconserved segments, incorrect loop reconstruction). Low: Target-template sequence alignment lower than 30% sequence identity (substantial alignment errors, suboptimal template selection)	Studying protein interactions with small molecules: enzyme mechanisms, structure-based drug design, ligand docking. Molecular replacement, protein design (stable, crystallizable variants, structural support for mutagenesis), protein–protein docking (prediction of protein partners). Integrative modeling (e.g., NMR structure refinement, modeling into low-resolution density maps), defining antibody epitopes. Functional relationships from structure: finding functional sites by motif searching, identifying conserved (functional) surface patches

RMSD, root mean square deviation; NMR, nuclear magnetic resonance.

approach is to follow a *de novo* modeling protocol, with, however, strong restraints acquired from the template structure. A floating restraint function that allows small shifts in the definition of restraints (say, ±1 or 2 residues) may be used in such cases to accommodate alignment ambiguities.[22] Another advantage of such approaches is that incorporating the information from several structural templates is straightforward. It should be, however, pointed out that none of these methods can fix severe alignment errors. In a recently proposed method, template-based modeling can be performed without any alignment prior to the modeling,[23] which successfully alleviates alignment-induced errors. The structural modeling is done directly onto a multifeatured scaffold created from a template structure and projected on a 3D grid in the Cartesian space. The method can be described as alignment refinement coupled with simultaneous structure refinement. Alternatively, alignment-related errors may be handled by generating many alternative alignments (from thousands to millions). Each alignment is used independently in a structure modeling process and the final model is selected from the multitude of models based on alternative alignments.[24]

Due to the shortcomings of the fold search (global optimization) and/or subsequent model selection methods (see Figure 1) *de novo* and fold recognition approaches usually introduce much more serious errors than comparative modeling. In the scenarios where template information is not available or it is very sparse and limited, some of the native secondary structure elements may be missing, mispredicted, or incorrectly oriented within the model structure. It is also very likely that the overall topology of the model is incorrect. Subsequent structure optimization is usually performed on a local scale (see Figure 1) and includes small changes in the orientation of secondary structure elements, loop remodeling, and side-chain repacking. Typically, a very large number of low-resolution trial conformations are proposed that sample very broadly the conformational space available for a given polypeptide sequence. All-atom refinement and scoring is frequently combined with structural clustering as a method of selection of a handful of final models out of millions of trial conformations.[25,26]

Contrary to experimentally derived structures, where error level and their distribution along the structure may be deduced from the discrepancy between the model and the experimental measurements, theoretical models do not provide any clues on their accuracy. The error estimation, in this case, imposes a great challenge on its own. Model quality assessment (MQA) problem, that is, the possibility of predicting the accuracy of structural models has become a well-defined field of active research, reviewed recently by Kryshtafovych and Fidelis.[27]

PHYSICS-BASED OPTIMIZATION

Molecular dynamics (MD) simulations enable us to follow atomic motions of molecular systems. As such, they provide invaluable information about the dynamic aspects of proteins, on top of the structural information, which is available from crystallographic studies.[28] Given computational costs, studying folding with explicit solvent fully atomic MD has been limited to only small systems, such as peptides and small proteins.[29] A tempting approach is to use all-atom MD for protein model refinement, but despite considerable efforts, all-atom refinement has shown limited success.[30] Different aspects and issues have

been discussed in the literature and are briefly summarized below.[31–35]

In the study of 12 small, single-domain proteins with different topologies, Lee et al.[31] noticed that although explicit solvent MD and implicit solvent-free energy calculations are very successful in ranking native structures and filtering models generated by the Rosetta method; any refinement of the models was beyond the scope of the suggested methodology. It has been commented that perhaps longer simulation times are needed for more systematic structure refinement.

The problem of simulation time in fully atomic refinement was further elucidated by Fan et al.[32] They concluded that very long simulations are needed to overcome kinetic barriers. To observe major structural changes, simulations on at least a microsecond time scale will be needed. Increasing the temperature to improve sampling was quite effective when the initial model was close to the target structure but did not work (in fact, often resulting in a major loss of structure accuracy) when the initial model was far from the target and not in a local potential energy well.

Moreover, current force fields are 'not perfect' and in some cases can generate nonnative conformations even when the simulations are carried for (apparently) sufficiently long times.

Wroblewska and Skolnick[36] employed a benchmark set of 150 nonhomologous proteins in the experiment aimed at recognition of the native structure from decoys. They demonstrated that the MM-GBSA (Molecular Mechanics with Generalized Born Surface Area) energy failed the test when all the structures were sufficiently minimized and concluded that some of the earlier successes in recognizing native structures by physics-based all-atom force fields were artifacts of decoy preparation procedures.

In another study aimed at the structural refinement of a set of single-domain, nonhomologous proteins with different folds, a carefully optimized all-atom Amber ff03 potential was used.[34] Structural improvements were observed for 70% of the models on average and 10% of decoys were refined to near experimental accuracy, below 2.5 Å.[30] However, this result was obtained for a compact decoy set, spanning the range of 0–8 Å Cα crmsd (coordinate root mean square deviation) from the native structure. For more distant structures, a reasonable correlation of MD energy with similarity to the native structure is rather improbable.

Feig et al.[35,37] proposed an iterative structure refinement protocol with an idealized scoring function. Selection of crmsd as the idealized scoring function enabled clear separation of the sampling problem from the scoring problem. This way the focus was on an assessment of the performance of sampling methods at different resolution levels. It has been found that CHARMM19[38] and SICHO[39] models lead to initially rapid refinement of a small set of proteins, but they are eventually outperformed by an all-atom model with a CHARMM22[40] force field with distance-dependent dielectric implicit solvent approximation. Also, the PRIMO[41] model performed well in generating near-native models with reduced representation of conformational space, and thus with reduced computational costs.

Finally, it is worth mentioning that apart from the refinement approaches that purely depend on physics-based methods, there are also a number of methods that refine structures combining physics-based methods with the use of experimental restraints[42,43]

KNOWLEDGE-BASED OPTIMIZATION

The protein structure is a densely packed system of atoms with a 3D network of connections imposed by covalent and hydrogen bonds. This makes its energy landscape extremely rugged. A refinement process can be defined as a walk on this surface in the quest for energy minimum better than the starting one, and it is a very difficult task. Knowledge-based force fields open a possibility to alleviate some of the minima and to make the landscape smoother. On the contrary to MM force fields, statistical potentials are often multidimensional and context dependent. For example, the statistical description of hydrogen bonding interaction recently proposed by Grishaev and Bax[44] is based on six-dimensional statistics derived from a subset of known protein structures. In another example,[45] backbone-dependent potential for rotamer assessment is a function of Phi, Psi, and all-Chi angles, which results in a function of six variables for Arginine and Lysine. Probably the most far-reaching case of such a type of approach is the CABS force field,[46] where interactions between side groups and model hydrogen bonds are explicitly connected to a complex and multi-featured molecular context. Such potentials, based on multidimensional statistics, enable capturing subtle effects observed in proteins that cannot be captured by MM interactions.

Hierarchical Approach to Optimization

In the case of physics-based optimization, the walk on the energy surface is guided by forces. Conversely, knowledge-based approaches usually combine Monte

Carlo sampling with various structural micromodifications (moves) to drive sampling toward the desired solution. The overall refinement problem may be split into a multilevel hierarchy, roughly corresponding to the levels of structure optimization. In some of the approaches, various independent assessment methods are used to indicate the regions of a model structure that have to be refined. Sampling efforts addressed to different parts of a refined model may actually directly depend on the local structural quality assigned to these areas.

Side-Chain Optimization

Side-chain refinement may be combined with backbone moves or performed as a separate task. Indeed, many methods have been proposed in the literature,[47] devoted solely to side-chain conformation optimization. In general, the methods attempt to optimize a fitness function (energy, score, etc.) according to different optimization methods: dead-end elimination (DEE)[48,49] theorem, gradient-based minimization, genetic algorithms,[50] and Monte Carlo sampling. These approaches have various execution times, but a reduction in the run time usually leads to corresponding accuracy tradeoff.[51]

Loop Modeling

Loop modeling is another important concept in comparative protein structure prediction that may be detached from other protocols and considered a separate task. During loop modeling, it is most often assumed that the structural core of the target has been already modeled with good accuracy and only the loops have to be reconstructed. In the cases when the template protein is closely related to the target, this assumption is generally fulfilled. Technically a loop region is subjected to conformational sampling, whereas the rest of the structure contributes only to the energy of a loop conformation. In principle, the protein core much better preserves its structure in the course of molecular evolution, whereas loops are usually variable in both sequence and structure even in the same protein family. However, it is often the case that the loop cannot be refined without simultaneously refining adjacent portions because of structural inaccuracies surrounding the loop region. As shown by Jacobson and colleagues,[52] extending the refinement area to include only the loop and the adjacent regions may improve prediction accuracy.

The loop structure and sequence variability makes them the most difficult regions to model by comparative modeling approaches (loop modeling is often termed a 'small *de novo* prediction' problem), and in general, the accuracy of homology models is the lowest in loop regions. However, loops are important for protein functions and play critical roles in protein recognition. In general, loop modeling methods[53,54] can be divided into two classes: database search and *de novo*. Methods from the former group attempt to find in structural databases main chain segments that fit the anchor regions of a loop. In the latter, conformational space of a loop is sampled to minimize its energy. Loop conformations may be generated by Monte Carlo search procedure,[55] kinematic loop closure,[56,57] Cyclic Coordinate Descent method,[58] or *de novo* hierarchical procedures.[52]

Recent studies on proteins structure and function provided another type of target for molecular simulation, similar to the loop modeling but even less tractable: intrinsically disordered regions. The regions, typically longer than loops, exist as ensembles of fluctuating structures lacking in stable structure and many of them undergo coupled folding and binding processes. Even though molecular simulation methods have only recently been applied to the study of protein disorder, they have already provided a new structural insight into disordered state ensembles.[59]

Examples of Knowledge-Based Optimization Methods

A number of different approaches have been proposed over the recent years for solving the refinement problem. The last edition of the critical assessment of techniques for protein structure prediction (CASP) competition alone had 34 groups participating in this category. Components of these methods and their key ideas are discussed in the relevant sections of this review. Here, we provide a more detailed description of two leading methods. Both of them combine knowledge- and physics-based elements.

Rosetta

In the Rosetta method,[60] short structural fragments are used to construct polypeptide chains. The fragments (most typically three and nine residue long) are extracted from the known protein structures[61] and represented by main chain and single united atoms for side groups. During the sampling process, model chains undergo a long series of deletions/insertions of such building blocks. After short simulated annealing in the reduced space, the best scoring model is selected, reconstructed to the all-atom representation (including all hydrogen atoms), and subjected to Monte Carlo refinement with deep

energy minimization. All-atom Rosetta energy comprises 21 knowledge-based terms, such as a multidimensional hydrogen bond,[62] solvation energy, or a score based on Ramachandran (i.e., Phi/Psi) statistics.

There are several refinement protocols in Rosetta, designed for different refinement goals. All of them comprise small backbone perturbations, rotamer repacking, and gradient-based energy minimization. These three components could be to a great extent customized, depending on a particular problem. Backbone perturbations are performed either by a small random change of Phi/Psi/Omega angles or by a backrub[63] move. Low-energy rotamer selection[64] employs a rotamer library in Monte Carlo simulated annealing. The resulting low energy conformation is further optimized according to rotamer trials with the side-chains minimization (RTMIN) protocol.[65] In a single trial, all rotameric conformations for a certain amino acid side chain are attempted. Each rotamer energy is minimized in the dihedral space with the rest of the protein held fixed. A rotamer that introduces the largest drop in the total energy is introduced. The algorithm proceeds by randomly selecting another amino acid in the protein chain and repeating the RTMIN step. Such a procedure enables us to go beyond the limitations of a discrete rotamer library and samples a continuous spectrum of side-chain conformations. At the end of the refinement calculations, all-atom energy function may be subjected to Davidon–Fletcher–Powell minimization either in the dihedral (Phi, Psi, Omega, and Chi) or in the Cartesian space.

During the whole-structure refinement protocol applied to every model obtained from Rosetta-reduced space simulation, Phi/Psi angles are given very little freedom to move. The main purpose of this protocol is to repack side chains and to find the lowest value of all-atom energy for a given conformation of a protein backbone. It has been shown[66] that Rosetta all-atom energy may be used to identify the correct (i.e., native-like) model providing that the input structure subjected to the refinement protocol is already close (about 2.5 Å or better) to the true answer.

During loop refinement, Rosetta holds all the secondary structure elements fixed. Trial loop conformation may be generated according to one of a few different algorithms implemented in Rosetta, for example, CCD,[58] from a loop library or inverse kinematics.[67] A new loop conformation is then subjected to side-chain repacking and refinement as described above. The loop modeling protocol can be easily applied to refine any arbitrary part of a protein structure.

KnowMIN

Levitt et al.[68] derived an all-atom knowledge-based scoring function from a nonredundant database of 500 high-quality experimental protein structures. All atoms were classified into 167 atom types as done previously in Samudrala's RAPDF KB potential.[69] A distance histogram was built separately for each pair of atom types and transformed into a potential of mean force according to Lu and Skolnick[70] formalism. In the next step, the potentials were approximated with a spline polynomial. Finally, the knowledge-based potential was combined with classical molecular mechanics force fields: GROMOS96, OPLS-AA, AMBER99, and ENCAD just by replacing the relevant nonbonded terms. The MESHI package[71] (see below) was used to control the refinement process and conduct the gradient-based minimization of the energy function.

MESHI optimization routines can be also accessed using the Beautify program, devised to refine Cα models resulting from various fold recognition and comparative modeling techniques. Starting from alpha carbon coordinates, the program reconstructs all the necessary main chain atoms and missing fragments (e.g., loops) and refines the model in its backbone representation. Finally, the side chains are added and the structure is subjected to the second refinement step. Both the refinement steps are done by direct gradient-based energy minimization according to the MESHI knowledge-based force field.

CRITICAL ASSESSMENT OF THE PERFORMANCE OF REFINEMENT METHODS

A few papers attempted to compare physics-based refinement methodology with its knowledge-based counterpart. Summa and Levitt[72] performed *in vacuo* energy minimization to compare the performance of various physics-based force fields: (OPLS)-AA,[73] AMBER99,[74,75] GROMOS96,[76] and ENCAD[77] with a statistically derived KB potential. The authors found that the knowledge-based potential performed best and refined almost all proteins toward the native structure. AMBER99 was found to be the second best. In general, physics-based potentials moved models away from the native state, thus making the refinement results worse than the initial conformations. These findings can be explained at least to some extent by the fact that the *in vacuo* environment is artificial for proteins and a native structure may no longer be in its global free energy minimum. *In vacuo* conditions, therefore, favor knowledge-based force fields that

implicitly include solvent effects, for example, encoded in distance distributions observed in crystal structures. In a subsequent study, Levitt and coworkers[78] tested the role of a solvent in protein structure refinement. Implicit solvent GBSA combined with an all-atom OPLS-AA force field has been compared with a knowledge-based approach. The energy of 729 near-native models collected for 75 different proteins was optimized according to a limited-memory variant of a BFGS minimization algorithm. Overall, GBSA performed better than KB for 30 of the 75 proteins. However, GBSA moved 29 proteins away from the native state, whereas KB worsened just 3.

The refinement of a native structure is an important test for a given methodology that enables decoupling conformational sampling efficiency from energy function performance. Misura and Baker[79] tested the Rosetta refinement protocol on a set of 10 small proteins and concluded that the sampling protocol appears to be the main limitation. Although their backbone and rotamer search protocols were reasonably successful in recovering rotameric states in the native backbone or in refining a perturbed native structure, the performance in refining *de novo* models was considerably worse. The refined *de novo* models feature poorer packing and consistently worse attractive components of Lennard–Jones energy. Some of these barriers, however, may result from the fact that the sampling is conducted in internal coordinates (Phi,Psi,Omega, and Chi angles) while keeping bond lengths and planar angles fixed to idealized values. Nevertheless, the authors showed improvement in crmsd distance for all the 10 test cases during the refinement process.

FIGURE 3 | Results of TR592 refinement from the CASP9 refinement competition. Top: Three superimposed structures are shown—native in blue, starting model in gray [coordinate root mean square deviation (crmsd): 1.26], and the best model in magenta (crmsd: 0.96). Significant improvement from the starting model can be observed in the loop region marked with a circle. It is noteworthy that this region was pointed out by the competition assessors to the participating groups as one of the main areas for refinement. Bottom: In the same size and orientation as mentioned above; all the predicted models (designated by participating groups as the best) are shown in magenta, together with the native in blue (a few heavily mismatched models were removed for clarity).

LESSONS FROM THE TWO LATEST CASPs—IS THERE ANY PROGRESS?

Refinement Category in CASP

In the last two editions of CASP, one of the categories was protein structure refinement. In this category, organizers suggested starting models for a small set of targets. Moreover, in most cases, predictors were informed which fragments of models needed significant improvements (for an example, see Figure 3).

In the extensive assessment of CASP8 results,[80] the authors suggest that refinement methods should *primum non nocere*—first do not harm. In other words, methods can be judged as successful if the quality of the model returned by the method is higher than the quality of the starting structure. Results from last two CASPs show that this criterion is really challenging. The main reason is very high accuracy of the starting models. As shown in Figure 4, the majority of models have crmsd to the native structure below 2 Å. Starting models given by organizers are chosen among the best submitted for the regular structure prediction CASP category. Obviously, the groups, which provide the best models, likely perform model optimization before submitting their results. Therefore, the real goal for the participants in the refinement

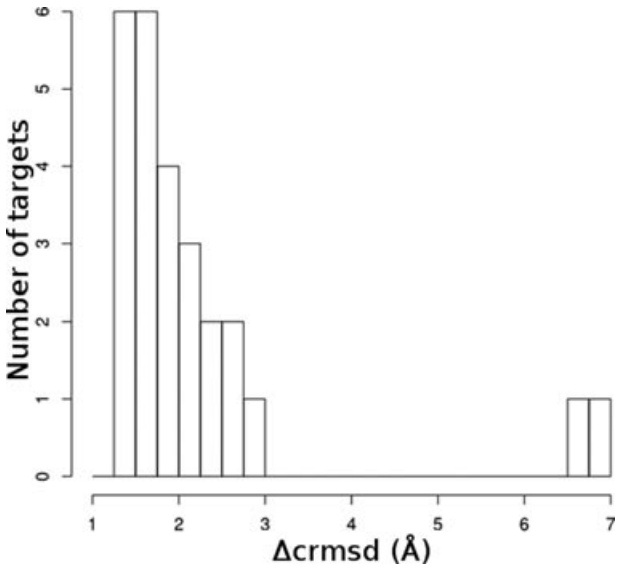

FIGURE 4 | Distribution of coordinate root mean square deviation (crmsd) of the starting structures for the refinement in CASP8 and CASP9. The majority of models are close to the native structure.

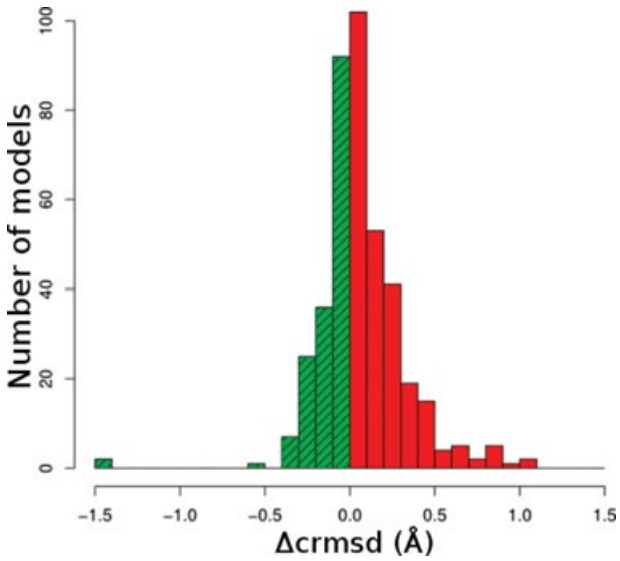

FIGURE 5 | Distribution of coordinate root mean square deviation (crmsd) differences between structures before and after refinement, obtained during the CASP9 experiment. Green color (negative Δcrmsd) means improvements and red (positive Δcrmsd) means worsening of the original models.

category was to improve structures which have been already refined.

CASP Results

MacCallum et al.[80] used many metrics (such as GDT_TS, GDT_HA, GDT_SC, MolProbity, MCRS, etc.) to measure changes in the accuracy of resulting structures in the CASP8 refinement category. For most of the metrics, there were more failures than improvements. Obviously, this is a general statement for the entire set of the structures submitted and does not show differences between the groups. In fact, some of them performed much better than others, but only a few groups were able to improve the average scores of the structures.

Here, we present results from CASP9 based on the scores provided by CASP organizers.[81] Figure 5 shows the distribution of crmsd differences between structures before and after refinement (Δcrmsd). The majority of structures obtained after refinement are worse than the starting ones (positive Δcrmsd). However, in most cases the absolute Δcrmsd value is very small (below 0.1 Å). Specific information for each group is presented in Figure 6.

Successful Methods

Table 2 shows the best groups from the last CASP according to the combination of scores used by CASP assessors. Only two groups (Baker, Levitt/KnowMIN)

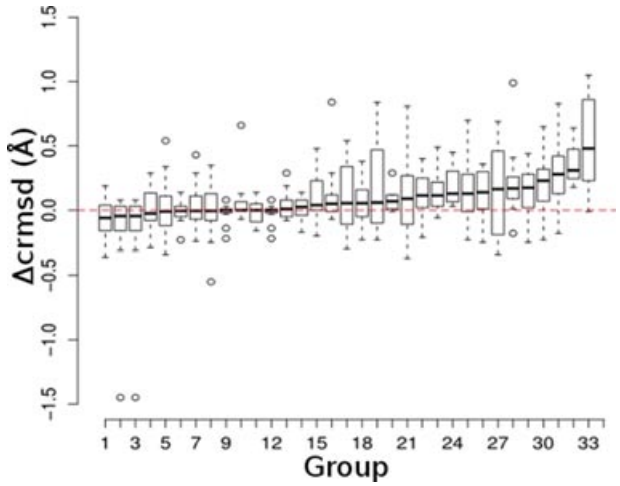

FIGURE 6 | Distribution of Δcrmsd (coordinate root mean square deviation) for each method. Successful results are below the red dashed line. Only for eight of 34 methods have the sets of the resulting structures mean Δcrmsd lower than zero.

were able to repeat their success from the CASP8 refinement experiment and were ranked among the top four. The two best groups (Baker, Foldit)[82,83] used the Rosetta modeling approach, summarized above. Other successful methods employ MD (SchroderLab) or direct energy minimization of the knowledge-based potential of mean force (Levitt).[84] In CASP9, the KnowMin group (previously, Levitt) extended their

TABLE 2 | Top Ranked Groups from CASP9 Refinement Category

Rank	Group Name	Group Members	Methods	CASP8 Rank
1	Baker	J. Thompson, T.J. Brunette, D.E. Kim, F. Khatib, D. Gront, F. DiMaio, R. Wang, R. Vernon, B. Kim, J. Pei, S. Cooper, M. Tyka, D. Baker	Monte Carlo minimization of Rosetta full-atom energy	1
2	Foldit	F. Khatib, S. Cooper, J. Thompson, I. Makedon, J. Barbero, Z. Popović, D. Baker, Foldit players	Minimization of Rosetta full-atom energy; online player modifications	New group
3	Knowmin	G. Chopra, M. Levitt	Energy minimization (knowledge- and physics-based potentials)	4 (previously LevittGroup)
4	SchroderLab	A. Wojtyczka, G.F. Schröder	Molecular dynamics (physics-based potentials)	New group

procedure using a combination of knowledge- and physics-based potentials.

It should be noted that differences between the top-ranked groups are not statistically significant and strongly depend on the metrics used for evaluation. This is a consequence of the character of the task, which aims at subtle changes of structures and the arbitrary character of all metrics which focus on different features. Moreover, the limited number of targets (CASP8—12, CASP9—14) does not allow for statistically significant analysis of the results and for changing them between the last two CASP editions. However, we agree with Ken Dill, assessor of the refinement category, who said during the last CASP conference that if there was any progress it was little. See also Table 3 for the list of example software (online servers and stand-alone software) for protein structure optimization.

COMPUTER OR HUMAN OPTIMIZATION?

The arbitrary partitioning of the refinement problem into separate concepts, implemented as distinct subroutines or protocols, attempts to utilize the experience of human experts in an automated fashion. So far, however, no one succeeded in substituting a human with a computer. The FoldIt[82] refinement method goes in the opposite direction: the computer is substituted with a human. The program itself resembles a video game. The player can move around a particular residue or a secondary structure element dragging them with a mouse. The computer's role is to quickly minimize all-atom energy according to the Rosetta scoring function. Interestingly, many players lack any biological training and they play FoldIt like any other video game. Despite this, they can achieve results comparable to or even better than human experts or specialized computer programs. In fact, during the ninth round of CASP experiment, FoldIt was the third-best performing group,[92] as ranked by an overall score averaged over the best (cherry-picked) model for each target. This shows that the currently used algorithms can still be improved.

CONCLUSION

Examples from the recent literature show that the difficulty of the structure optimization problem grows very fast with the distance between the starting conformation and the true answer. Some authors suggest that the problem is, in general, unsolvable for structures more than 2–3 Å apart from the native.[66] Reaching such a high accuracy is already a great challenge for *de novo* modeling methods and, if achieved, it should be considered a very successful case rather than a typical result. Extending the radius of convergence of the structure optimization methods is probably the most important direction for future improvements. As already discussed, optimization protocols comprise primarily two components: an energy function (i.e., scoring) and a sampling method, which define two areas of research. The two, however, strongly depend on each other and should not be considered in separation. For instance, the optimization of a protein molecule in its all-atom representation enables one to account for subtle energetic effects[93] and to accurately describe side-chain packing. At the same time, the densely packed system of atoms represented by spheres that strongly repel one another at close distances imposes a great challenge for the sampling method. The excluded volume term that is often expressed as r^{-12} renders a great number of local minima on the energy landscape. Therefore, even when

TABLE 3 | Example Software for Protein Structure Optimization

	Name	Address	Comments
Online Servers	Kobamin	http://csb.stanford.edu/kobamin/	Direct energy minimization using knowledge-based potentials[84] (see section *KnowMIN* in the text)
	FG-MD	http://zhanglab.ccmb.med.umich.edu/FG-MD/	MD-based algorithm. FG-MD identifies analogous fragments from the Protein Data Bank (using TM align[85]), which enables deriving spatial restraints.
	Protein Refinement Server	http://silvio.cs.uno.edu/proteinrefinementserver/	Near-native structure refinement using knowledge-based potentials.[72]
	FoldIt	http://fold.it/portal/	Online game,[82] which uses Rosetta full-atom energy and allows manual modifications.
Stand-Alone Software	Rosetta	http://www.rosettacommons.org/	A complex package[60] for protein structure prediction, which supports several structure refinement protocols (see the text for details).
	Modeller	http://salilab.org/modeller/	Modeller[86] is a state-of-the-art protein structure modeling tool, which performs *inter alia de novo* modeling of loops[87] and optimization of protein structure models with respect to a flexibly defined objective function.
	PLOP	http://www.jacobsonlab.org/plop_manual/plop_overview.htm	Part of the PRIME package. PLOP is a program for protein modeling using all-atom energy functions, which employs truncated Newton minimization, side-chain optimization,[88,89] loop prediction,[90] and the prediction of helix positions.[91]
	Beautify	http://www.cs.bgu.ac.il/~meshi/	Part of the MESHI package.[71] The program enables structure reconstruction (backbone, side chains) and refinement (direct energy minimization according to knowledge-based energy function).

MD, Molecular dynamics.

a particular scoring function places the native conformation in the global minimum, it is quite likely that the sampling process will get trapped before reaching it.[79] Apparently, FoldIt players manage to avoid it by manually navigating the conformational search.

Finally, it is worth noting that in all benchmarks, as well as during the CASP experiments, the results of the optimization process are compared to the experimental structure, most typically established by an X-ray diffraction experiment conducted in cryogenic conditions. To some extent, this explains the favorable results by knowledge-based methods, obtained in a process of aggressive minimization of a scoring function. The function that is most commonly derived from already solved high-resolution protein structures. The dynamic aspect of a structure is completely missing in this picture. However, even in their native state, proteins access many different conformations and importantly their motions are related to the function they perform. Characterization of the protein structural dynamics, thus linking the static structures with their function, is one of the major challenges in biology, and computer simulation methods are already a standard tool for that purpose. Although contemporary physics-based force fields are rough approximations of the real potential energy function, current estimations are that they are able to capture the meaningful aspects of protein dynamics[94] and the same is possible using a knowledge-based force field approach.[95] Hopefully the next generations of methods will optimize a protein as an ensemble rather than as a single conformation. This would provide us clues about function but also give some estimation of entropy for competing energy minima and help in the final model selection.

ACKNOWLEDGMENTS

This work was performed with support from the Polish National Science Center (Grant No. NN301071140) and from the Polish Ministry of Science and Higher Education (IP2010 011170).

REFERENCES

1. Kendrew JC, Dickerson RE, Strandberg BE, Hart RG, Davies DR, Phillips DC, Shore VC. Structure of myoglobin: a three-dimensional Fourier synthesis at 2 Å. Resolution. *Nature* 1960, 185:422–427.
2. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. The Protein Data Bank. *Nucleic Acids Res* 2000, 28:235–242.
3. Wu CH, Apweiler R, Bairoch A, Natale DA, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, et al. The Universal Protein Resource (UniProt): an expanding universe of protein information. *Nucleic Acids Res* 2006, 34(suppl 1):D187–D191.
4. DeWeese-Scott C, Moult J. Molecular modeling of protein function regions. *Proteins* 2004, 55:942–961.
5. Piedra D, Lois S, de la Cruz X. Preservation of protein clefts in comparative models. *BMC Struct Biol* 2008, 8:2. Available at: http://www.biomedcentral.com/1472-6807/8/2
6. Chakravarty S, Sanchez R. Systematic analysis of added-value in simple comparative models of protein structure. *Structure* 2004, 12:1461–1470.
7. Qian B, Raman S, Das R, Bradley P, McCoy AJ, Read RJ, Baker D. High-resolution structure prediction and the crystallographic phase problem. *Nature* 2007, 450:259–U257.
8. Schwede T, Sali A, Honig B, Levitt M, Berman HM, Jones D, Brenner SE, Burley SK, Das R, Dokholyan NV, et al. Outcome of a workshop on applications of protein models in biomedical research. *Structure* 2009, 17:151–159.
9. Sanchez R, Pieper U, Melo F, Eswar N, Marti-Renom MA, Madhusudhan MS, Mirkovic N, Sali A. Protein structure modeling for structural genomics. *Nat Struct Biol* 2000, 7:986–990.
10. Tramontano A. The role of molecular modelling in biomedical research. *FEBS Lett* 2006, 580:2928–2934.
11. Baker D, Sali A. Protein structure prediction and structural genomics. *Science* 2001, 294:93–96.
12. Robinson CV, Sali A, Baumeister W. The molecular sociology of the cell. *Nature* 2007, 450:973–982.
13. Zhang J, Baker ML, Schroder GF, Douglas NR, Reissmann S, Jakana J, Dougherty M, Fu CJ, Levitt M, Ludtke SJ, et al. Mechanism of folding chamber closure in a group II chaperonin. *Nature* 2010, 463:379–383.
14. DiMaio F, Terwilliger TC, Read RJ, Wlodawer A, Oberdorfer G, Wagner U, Valkov E, Alon A, Fass D, Axelrod HL, et al. Improved molecular replacement by density- and energy-guided protein structure optimization. *Nature* 2011, 473:540–543.
15. Rost B. Twilight zone of protein sequence alignments. *Protein Eng* 1999, 12:85–94.
16. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 1997, 25:3389–3402.
17. Gront D, Kolinski A. BioShell—a package of tools for structural biology computations. *Bioinformatics* 2006, 22:621–622.
18. Gront D, Kolinski A. Utility library for structural bioinformatics. *Bioinformatics* 2008, 24:584–585.
19. Eddy SR. Profile hidden Markov models. *Bioinformatics* 1998, 14:755–763.
20. Kmiecik S, Gront D, Kolinski A. Towards the high-resolution protein structure prediction. Fast refinement of reduced models with all-atom force field. *BMC Struct Biol* 2007, 7:43.
21. Hershkovitz E, Sapiro G, Tannenbaum A, Williams LD. Statistical analysis of RNA backbone. *IEEE/ACM Trans Comput Biol Bioinf* 2006, 3:33–46.
22. Kolinski A, Betancourt MR, Kihara D, Rotkiewicz P, Skolnick J. Generalized comparative modeling (GENECOMP): a combination of sequence comparison, threading, and lattice modeling for protein structure prediction and refinement. *Proteins* 2001, 44:133–149.
23. Kolinski A, Gront D. Comparative modeling without implicit sequence alignments. *Bioinformatics* 2007, 23:2522–2527.
24. Chivian D, Baker D. Homology modeling using parametric alignment ensemble generation with consensus and energy-based model selection. *Nucleic Acids Res* 2006, 34:e112.
25. Shortle D, Simons KT, Baker D. Clustering of low-energy conformations near the native structures of

small proteins. *Proc Natl Acad Sci USA* 1998, 95:11158–11162.

26. Gront D, Kolinski A. HCPM—program for hierarchical clustering of protein models. *Bioinformatics* 2005, 21:3179–3180.
27. Kryshtafovych A, Fidelis K. Protein structure prediction and model quality assessment. *Drug Discov Today* 2009, 14:386–393.
28. Scheraga HA, Khalili M, Liwo A. Protein-folding dynamics: overview of molecular simulation techniques. *Annu Rev Phys Chem* 2007, 58:57–83.
29. Duan Y, Kollman PA. Pathways to a protein folding intermediate observed in a 1-microsecond simulation in aqueous solution. *Science* 1998, 282:740–744.
30. Jagielska A, Wroblewska L, Skolnick J. Protein model refinement using an optimized physics-based all-atom force field. *Proc Natl Acad Sci USA* 2008, 105:8268–8273.
31. Lee MR, Tsai J, Baker D, Kollman PA. Molecular dynamics in the endgame of protein structure prediction. *J Mol Biol* 2001, 313:417–430.
32. Fan H, Mark AE. Refinement of homology-based protein structures by molecular dynamics simulation techniques. *Protein Sci* 2004, 13:211–220.
33. Verma A, Wenzel W. Protein structure prediction by all-atom free-energy refinement. *BMC Struct Biol* 2007, 7:12+.
34. Wroblewska L, Jagielska A, Skolnick J. Development of a physics-based force field for the scoring and refinement of protein models. *Biophys J* 2008, 94:3227–3240.
35. Feig M, Gopal S, Vadivel K, Stumpff-Kane A. Conformational sampling in structure prediction and refinement with atomistic and coarse-grained models. In: Kolinski A, ed. *Multiscale Approaches to Protein Modeling*.. New York: Springer; 2011, 85–109.
36. Wroblewska L, Skolnick J. Can a physics-based, all-atom potential find a protein's native structure among misfolded structures? I. Large scale AMBER benchmarking. *J Comput Chem* 2007, 28:2059–2066.
37. Stumpff-Kane AW, Maksimiak K, Lee MS, Feig M. Sampling of near-native protein conformations during protein structure refinement using a coarse-grained model, normal modes, and molecular dynamics simulations. *Proteins* 2008, 70:1345–1356.
38. Neria E, Fischer S, Karplus M. Simulation of activation free energies in molecular systems. *J Chem Phys* 1996, 105:1902–1921.
39. Kolinski A, Skolnick J. Assembly of protein structure from sparse experimental data: an efficient Monte Carlo model. *Proteins* 1998, 32:475–494.
40. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, et al. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J Phys Chem B* 1998, 102:3586–3616.
41. Gopal SM, Mukherjee S, Cheng YM, Feig M. PRIMO/PRIMONA: a coarse-grained model for proteins and nucleic acids that preserves near-atomistic accuracy. *Proteins* 2010, 78:1266–1281.
42. Bruenger AT, Karplus M. Molecular Dynamics Simulations with Experimental Restraints. *Acc Chem Res* 1991, 24:54–61.
43. Dolenc J, Missimer JH, Steinmetz MO, van Gunsteren WF. Methods of NMR structure refinement: molecular dynamics simulations improve the agreement with measured NMR data of a C-terminal peptide of GCN4-p1. *J Biomol NMR* 2010, 47:221–235.
44. Grishaev A, Bax A. An empirical backbone-backbone hydrogen-bonding potential in proteins and its applications to NMR structure refinement and validation. *J Am Chem Soc* 2004, 126:7281–7292.
45. Amir E-AD, Kalisman N, Keasar C. Differentiable, multi-dimensional, knowledge-based energy terms for torsion angle probabilities and propensities. *Proteins: Struct Funct Bioinf* 2008, 72:62–73.
46. Kolinski A. Protein modeling and structure prediction with a reduced representation. *Acta Biochim Pol* 2004, 51:349–371.
47. Vásquez M. Modeling side-chain conformation. *Curr Opin Struct Biol* 1996, 6:217–221.
48. Desmet J, Maeyer MD, Hazes B, Lasters I. The dead-end elimination theorem and its use in protein side-chain positioning. *Nature* 1992, 356:539–542.
49. Desmet J, Spriet J, Lasters I. Fast and accurate side-chain topology and energy refinement (FASTER) as a new method for protein structure optimization. *Proteins* 2002, 48:31–43.
50. Tuffery P, Etchebest C, Hazout S, Lavery R. A new approach to the rapid determination of protein side chain conformations. *J Biomol Struct Dyn* 1991, 8:1267–1289.
51. Voigt CA, Gordon, Mayo SL. Trading accuracy for speed: a quantitative comparison of search algorithms in protein sequence design1. *J Mol Biol* 2000, 299:789–803.
52. Sellers BD, Zhu K, Zhao S, Friesner RA, Jacobson MP. Toward better refinement of comparative models: predicting loops in inexact environments. *Proteins* 2008, 72:959–971.
53. Fiser A, Do RK, Šali A. Modeling of loops in protein structures. *Protein Sci* 2000, 9:1753–1773.
54. Soto CS, Fasnacht M, Zhu J, Forrest L, Honig B. Loop modeling: sampling, filtering, and scoring. *Proteins* 2008, 70:834–843.
55. Jamroz M, Kolinski A. Modeling of loops in proteins: a multi-method approach. *BMC Struct Biol* 2010, 10:5+.
56. Go N, Scheraga HA. Ring closure and local conformational deformations of chain molecules. *Macromolecules* 1970, 3:178–187.

57. Coutsias EA, Seok C, Wester MJ, Dill KA. Resultanta and loop closure. 2006, 106:176–189+.
58. Canutescu AA, Dunbrack RL. Cyclic coordinate descent: a robotics algorithm for protein loop closure. *Protein Sci: Publ Protein Soc* 2003, 12:963–972.
59. Rauscher S, Pomes R. Molecular simulations of protein disorder. *Biochem Cell Biol* 2010, 88:269–290.
60. Leaver-Fay A, Tyka M, Lewis SM, Lange OF, Thompson J, Jacak R, Kaufman K, Renfrew PD, Smith CA, Sheffler W, et al. Rosetta3 an object-oriented software suite for the simulation and design of macromolecules. *Methods Enzymol* 2011, 487:545–574.
61. Gront D, Kulp DW, Vernon RM, Strauss CEM, Baker D. Generalized fragment picking in Rosetta: design, protocols and applications. *PLoS ONE*. 2011, 6:e23294.
62. Morozov AV, Kortemme T, Tsemekhman K, Baker D. Close agreement between the orientation dependence of hydrogen bonds observed in protein structures and quantum mechanical calculations. *Proc Natl Acad Sci USA* 2004, 101:6946–6951.
63. Betancourt MR. Efficient Monte Carlo trial moves for polypeptide simulations. *J Chem Phys* 2005, 123:174905+.
64. Kuhlman B, Baker D. Native protein sequences are close to optimal for their structures. *Proc Natl Acad Sci USA* 2000, 97:10383–10388.
65. Wang C, Schueler-Furman O, Baker D. Improved side-chain modeling for protein–protein docking. *Protein Sci: Publ Protein Soc* 2005, 14:1328–1339.
66. Bradley P, Malmstrom L, Qian B, Schonbrun J, Chivian D, Kim DE, Meiler J, Misura KM, Baker D. Free modeling with Rosetta in CASP6. *Proteins* 2005, 61 Suppl 7:128–134.
67. Mandell DJ, Coutsias EA, Kortemme T. Sub-angstrom accuracy in protein loop reconstruction by robotics-inspired conformational sampling. *Natl Methods* 2009, 6:551–552.
68. Summa CM, Levitt M. Near-native structure refinement using in vacuo energy minimization. *Proc Natl Acad Sci USA* 2007, 104:3177–3182.
69. Samudrala R, Moult J. An all-atom distance-dependent conditional probability discriminatory function for protein structure prediction. *J Mol Biol* 1998, 275:895–916.
70. Lu H, Skolnick J. A distance-dependent atomic knowledge-based potential for improved protein structure selection. *Proteins* 2001, 44:223–232.
71. Kalisman N, Levi A, Maximova T, Reshef D, Zafriri-Lynn S, Gleyzer Y, Keasar C. MESHI: a new library of Java classes for molecular modeling. *Bioinformatics* 2005, 21:3931–3932.
72. Levitt M, Summa CM. Near-native structure refinement using in vacuo energy minimization. *Proc Natl Acad Sci USA* 2007, 104:3177–3182.
73. Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL. Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J Phys Chem B* 2001, 105:6474–6487.
74. Wang JM, Cieplak P, Kollman PA. How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? *J Comput Chem* 2000, 21:1049–1074.
75. Sorin EJ, Pande VS. Exploring the helix-coil transition via all-atom equilibrium ensemble simulations. *Biophys J* 2005, 88:2472–2493.
76. van Gunsteren WF BS, Eising AA, Hunenberger PH, Kruger P, Mark AE, Scott WRP, Tironi IG. *Biomolecular Simulation: The GROMOS96 Manual and User Guide*. Zürich, Switzerland: Vdf Hochschulverlag AG an der ETH Zürich; 1996.
77. Levitt M, Hirshberg M, Sharon R, Daggett V. Potential-energy function and parameters for simulations of the molecular-dynamics of proteins and nucleic-acids in solution. *Comput Phys Commun* 1995, 91:215–231.
78. Chopra G, Summa CM, Levitt M. Solvent dramatically affects protein structure refinement. *Proc Natl Acad Sci USA* 2008, 105:20239–20244.
79. Misura KM, Baker D. Progress and challenges in high-resolution refinement of protein structure models. *Proteins* 2005, 59:15–29.
80. MacCallum J, Hua L, Schnieders M, Pande V, Jacobson M, Dill K. Assessment of the protein-structure refinement category in CASP8. *Proteins* 2009, 77(suppl 9):66–80.
81. The official website of 9th Community Wide Experiment on the Critical Assessment of Techniques for Protein Structure Prediction. Available at: http://predictioncenter.org/casp9/. (Accessed August 3, 2011).
82. Cooper S, Khatib F, Treuille A, Barbero J, Lee J, Beenen M, Leaver-Fay A, Baker D, Popovic Z, Players F. Predicting protein structures with a multiplayer online game. *Nature* 2010, 466:756–760.
83. Raman S, Vernon R, Thompson J, Tyka M, Sadreyev R, Pei J, Kim D, Kellogg E, DiMaio F, Lange O, et al. Structure prediction for CASP8 with all-atom refinement using Rosetta. *Proteins* 2009, 77(suppl 9):89–99.
84. Chopra G, Kalisman N, Levitt M. Consistent refinement of submitted models at CASP using a knowledge-based potential. *Proteins* 2010, 78:2668–2678.
85. Zhang Y, Skolnick J. TM-align: a protein structure alignment algorithm based on the TM-score. *Nucleic Acids Res* 2005, 33:2302–2309.
86. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A. Comparative protein structure modeling using Modeller. *Curr Protoc Bioinf* 2006, Chapter 5:Unit 5 6.

87. Fiser A, Do RKG, Sali A. Modeling of loops in protein structures. *Protein Sci* 2000, 9:1753–1773.
88. Jacobson MP, Friesner RA, Xiang ZX, Honig B. On the role of the crystal environment in determining protein side-chain conformations. *J Mol Biol* 2002, 320:597–608.
89. Jacobson MP, Kaminski GA, Friesner RA, Rapp CS. Force field validation using protein side chain prediction. *J Phys Chem B* 2002, 106:11673–11680.
90. Jacobson MP, Pincus DL, Rapp CS, Day TJF, Honig B, Shaw DE, Friesner RA. A hierarchical approach to all-atom protein loop prediction. *Proteins Struct Funct Bioinf* 2004, 55:351–367.
91. Li X, Jacobson MP, Friesner RA. High-resolution prediction of protein helix positions and orientations. *Proteins Struct Funct Bioinf* 2004, 55:368–382.
92. MacCallum JL, Pérez A, Schnieders MJ, Hua L, Jacobson MP, Dill KA. Assessment of protein structure refinement in CASP9. *Proteins: Struct Funct Bioinf* 2011, 79:74–90.
93. Song Y, Tyka M, Leaver-Fay A, Thompson J, Baker D. Structure-guided force-field optimization. *Proteins* 2011, 79:1898–1909.
94. Orozco M, Rueda M, Ferrer-Costa C, Meyer T, Perez A, Camps J, Hospital A, Gelpi JL. A consensus view of protein dynamics. *Proc Natl Acad Sci USA* 2007, 104:796–801.
95. Kmiecik S, Kolinski A. Characterization of protein-folding pathways by reduced-space modeling. *Proc Natl Acad Sci USA* 2007, 104:12330–12335.

