Published in final edited form as:
Proteins. 2013 January ; 81(1): 119–131. doi:10.1002/prot.24167.

3Drefine: Consistent Protein Structure Refinement by Optimizing Hydrogen Bonding Network and Atomic-Level Energy Minimization

Debswapna Bhattacharya[1] and **Jianlin Cheng**[1,2,3]

[1]Department of Computer Science, University of Missouri, Columbia, MO 65211, USA

[2]Informatics Institute, University of Missouri, Columbia, MO 65211, USA

[3]Bond Life Science Center, University of Missouri, Columbia, MO 65211, USA

Abstract

One of the major limitations of computational protein structure prediction is the deviation of predicted models from their experimentally derived true, native structures. The limitations often hinder the possibility of applying computational protein structure prediction methods in biochemical assignment and drug design that are very sensitive to structural details. Refinement of these low-resolution predicted models to high-resolution structures close to the native state, however, has proven to be extremely challenging. Thus, protein structure refinement remains a largely unsolved problem. Critical assessment of techniques for protein structure prediction (CASP) specifically indicated that most predictors participating in the refinement category still did not consistently improve model quality.

Here, we propose a two-step refinement protocol, called 3Drefine, to consistently bring the initial model closer to the native structure. The first step is based on optimization of hydrogen bonding (HB) network and the second step applies atomic-level energy minimization on the optimized model using a composite physics and knowledge-based force fields. The approach has been evaluated on the CASP benchmark data and it exhibits consistent improvement over the initial structure in both global and local structural quality measures. 3Drefine method is also computationally inexpensive, consuming only few minutes of CPU time to refine a protein of typical length (300 residues).

Keywords

protein structure refinement; protein structure prediction; statistical potential; protein energy; hydrogen bonding network; energy minimization

INTRODUCTION

The goal of protein tertiary structure prediction is to accurately estimate the three-dimensional position of each atom in a protein. Comparative modeling (or homology modeling) is the most widely used technique in the field of protein structure prediction. In

Corresponding author: Jianlin Cheng, 109 EBW, University of Missouri, Columbia, MO 65211, USA, Phone: 573-882-7306, Fax: 573-882-8318, chengji@missouri.edu.

SUPPLIMENTARY INFORMATION

Supplementary Information includes two tables (Table S1 and Table S2) and a figure (Fig. S1) is available at: http://sysbio.rnet.missouri.edu/3Drefine/download.html

the traditional comparative modeling methods, an experimental protein structure (or template) that has significant sequence similarity to the target protein of interest is first identified, and then a sequence alignment between the target and the template is generated in order to use the structural information of the aligned regions of the template to construct a structural model for the target protein[1]. But, even with the best possible template and target sequence alignment, predicted models often deviate from the true native structures in terms of their atomic coordinates. Significant progress toward improving the accuracy of comparative modeling has been made during recent years by building the target structure combining the fragments from multiple templates [2-4]. The introduction of multiple templates has certainly enhanced the accuracy of structure prediction by bringing the model closer to the native structure than using a single template. Despite having largely correct backbone conformations, these models sometimes still have poor structural qualities, including irregular hydrogen bonding network, steric clashes, unphysical bond length, unrealistic bond angles, torsion angles and side-chain χ angles. Thus, direct refinement of the predicted models from their coordinate information alone with the goal of detection and correcting the errors is an essential part of computational protein structure prediction process.

The earlier studies in the field of protein structure refinement can be broadly classified into two categories: (1) methods that perform significant conformational changes in terms of backbone positioning and structural information [5-7] and (2) protocols that make small structural changes at the local level by modifying the side chain conformation[8,9] or removing the gaps and steric clashes [10,11]. The first kind of methods is more adventurous having potential of substantial improvement in the structural qualities [12]. But these techniques are computationally expensive and often inconsistent. The second type of protocols, although consistent, generally fails to bring model considerably closer to the native state. Consistent and simultaneous improvement in both global and local structural qualities of the initial models in a computationally efficient manner is, therefore, a nontrivial problem.

Protein structure refinement has received noteworthy attention in the recent critical assessment of techniques for protein structure prediction (CASP8 and CASP9). According to the results in CASP9 refinement experiment, very few methods exist which can consistently bring the initial models closer to the native structure. The majority of the improvement has been observed at the local level by modification of the physicality of the models or alteration of the side chain positions and not at the global topologies [13,14]. Encouragingly, some promising progress has been made in the recent past by development of methods that have the potential to improve both the global topologies and the local structural qualities of the predicted models by using optimized physics-based all-atom force field [15], applying knowledge-based potential [16,17] or performing Fragment-Guided Molecular Dynamics Conformation Sampling [18]. These protocols usher the way to solve the protein structure refinement problem.

In this work we present an efficient refinement protocol, called 3Drefine that is based on two steps of refinement process. We extensively test this method on CASP benchmarks having high diversity in the difficulty of the prediction targets. 3Drefine demonstrates significant potential in atomic-level protein structure refinements in terms of both global and local measures of structural qualities. Thus, we expect the protocol to be a useful addition to current state-of the-art refinement tools. We also hope that this method can be adopted as a final step in the existing protein structure prediction pipeline.

MATERIALS AND METHODS

The 3Drefine protocol refines the initial protein structures in two steps: (1) Optimizing Hydrogen Bonds network and (2) atomic-level energy minimization using a combination of physics and knowledge based force fields; both carried out using Java based molecular modeling package MESHI[19]. The justification for choosing MESHI over the other modeling packages is threefold: (1) the strict object-oriented design used in MESHI enhances possibility of code reuse by means of inheritance mechanism provided in Java, thereby reducing the development time; (2) MESHI has robust garbage-collection utility to deal with failures and (3) the open source and platform independent nature of MESHI makes it more flexible.

Optimizing Hydrogen Bonds network

Hydrogen atoms are the most frequently occurring atoms in a protein structure and play a crucial role in protein folding through hydrophobic interaction or hydrogen bonding[20-22]. Previous studies in the field suggest that accurately determining the positions of the hydrogen atoms has a major influence on applying atomic-level potentials for protein structure refinement[23-25]. Unfortunately, most of the computational methods for protein structure prediction lack the ability to consistently and correctly identify the hydrogen atoms. We, therefore, decided to first optimize the hydrogen bonding network in the initial model.

The current state-of-the-art protocols for predicting the hydrogen bonds generally follow a combination of local geometry restraint and a conformational search [26-29]. We adopt a very similar approach here to optimize hydrogen bonding network. Using MESHI [19], the hydrogen bonds in the initial full-atomic model are calculated first. The position of non-polar hydrogen atoms are determined by using fixed bond lengths and bond angles parameters supplied in the MESHI library which have been derived by a collection of 1145 protein domains as part of MESHI package. For the polar hydrogen atoms, a search is performed via “Geometry” and “Molecular Elements package” of MESHI to find out the most favorable position of hydrogen atoms satisfying hydrogen bonds with the closest neighboring atoms and considering the protonation state of each amino acid. We call this an extended atomic model.

Atomic-level Energy Minimization

We use atomic-level force fields driven by MESHI [19] for performing energy minimization on the extended atomic model. Since the current release of MESHI package does not include any established force fields, we, therefore, construct a customized all-atomic total energy of a protein model by combining the energetic contributions of the bonded interactions described in ENCAD potential [30], which is an example of a traditional molecular mechanics force field, some standard energy elements using the “Energy package” included in MESHI software and a Knowledge-based atomic pairwise potential of mean force [16]. The ENCAD molecular mechanics force fields are chosen because they are freely available and have all been implemented for use with the MESHI molecular dynamics package. We include only the bonded terms of the ENCAD potential (bond stretches, bond angle bends, and torsion angle twists) since they are stronger than the nonbonded terms [16] (van der Waals interactions, and electrostatic interactions). As some of the energy terms require the secondary structure values for accurate calculations, we set the secondary structure of the extended atomic model by using DSSP [31]. Almost all energy term requires some knowledge about the distances between the atoms. To this end, we use a fast heuristics for calculating distances in the system. Given the atom list of the system, an internal matrix of distance objects is created for all the inter-atomic distances by means of the “Distance

Matrix class" in MESHI. Following the standard convention adopted in the MESHI package to calculate distances in a computationally inexpensive way, we have made two assumptions: first, the atoms that are separated by 4 bonds or less are considered bonded and second, distances must be within a cutoff of 5.5 Å. This means that any distance between non-bonded atoms (separation of more than four covalent bonds) that is higher than 5.5 Å has been assumed infinite.

Finally, the customized total energy of the extended atomic model, which is used to guide the minimization, is calculated by MESHI "Total Energy class" and consists of the following terms:

$$E_{total}=E_{bondlength}+E_{bondangle}+E_{torsion}+E_{hydrogenbonds}+E_{tether}+E_{kbpairwise} \quad (1)$$

In the following sections we will describe each of the energy terms mentioned in Eq. (1):

Bond Length energy term

When two atoms are connected by a chemical bond they tend to maintain a fixed distance (called equilibrium distance) depending on the type of the atoms participating in the bond formation. Any change in this equilibrium distance adds potential energy to the protein. As per the ENCAD potential [30], this is usually represented as:

$$E_{bondlength}=\sum \frac{1}{2}K_b(b-b_0)^2 \quad (2)$$

where b is the distance between the two bonded atoms, b_0 is their equilibrium distance and K_b is a bond stretching force-constant subject to the atom types. The parameters b_0 and K_b depend on the type of the bonds and their values can be found at the published work for ENCAD bonded energy terms [30]. This is denoted in Eq. (1) as $E_{bondlength}$ term.

Bond Angle energy term

Similar to bond length, when three atoms are connected with two chemical bonds they incline to maintain a fixed angle (called equilibrium angle) subject to the atom types. Any variation in this equilibrium angle adds potential energy to the protein. Following the standard ENCAD potential [30], this can be defined as:

$$E_{bondangle}=\sum \frac{1}{2}K_\theta(\theta-\theta_0)^2 \quad (3)$$

where θ is the distance between the two bonded atoms, θ_0 is their equilibrium distance and K_θ is a bond angle force constant. Like bond length energy term, the parameters θ_0 and K_θ related to bond angle energy term depend on the type of the bonds the atoms are involved in and their values can be found at the publication of the ENCAD bonded energy terms [30]. We symbolize this in Eq. (1) as $E_{bondangle}$ term.

Dihedral or torsions angle twist energy term

The third term in Eq. (1) has been denoted as $E_{torsion}$ and it represented in ENCAD potential [30] as:

$$E_{torsion}=\sum \frac{1}{2}K_\phi[1-\cos n(\phi_i-\phi_0)] \quad (4)$$

where n is the periodicity, φ_0 is the equilibrium value and K_φ is the half of the rotation barrier height. Values of these parameters have been described in ENCAD potential [30].

The torsion energy term has the ability to represent true dihedral angles and unrealistic or out-of-plane torsion (or dihedral) angles.

We implement all the above mentioned energy terms of the bonded interactions using the MESHI framework [19,32].

Hydrogen Bonds energy term

Hydrogen Bonds energy term calculates the energy over all the backbone Hydrogen bonds in the protein and is denoted by $E_{hydrogenbonds}$ in Eq. (1). We use a combination of *"Energy"* and "*Geometry packages*" of MESHI [19] framework to calculate the energy of the hydrogen bonds for the extended atomic model. Following the explicit hydrogen bonding potential defined in FG-MD [18] refinement method, we consider only the short range hydrogen-bonding potential with cutoff distance between the Hydrogen and the Oxygen ≤ 3 Å. This is defined as:

$$E_{HB}(d_{ij},\alpha,\beta)=\begin{cases} k_1(d_{ij}-d_0)^2+k_2(\alpha-\alpha_0)^2+k_3(\beta-\beta_0)^2 & d_{ij}\le 3.0 \\ 0 & d_{ij}>3.0 \end{cases} \quad (5)$$

where d_{ij} is the distance between hydrogen of the donor and oxygen of the acceptor, α is the N-H-O angle and β is the angle of C-O-H. Values of these parameters have been adopted from the published work of FG-MD[18] as $d_0 = 1.95 \pm 0.17$ Å, $\alpha_0 = 160.0 \pm 12.2°$, $\beta_0 = 150.0 \pm 17.5°$ and the values of the force constants are $k_1 = 2.0$, $k_2 = 0.5$ and $k_3 = 0.5$.

Tethering energy term

Protein models sometimes have unfavorable atomic interaction and these disordered atomic positions can cause large initial forces that result in artificial movement away from the original structure while performing energy minimization. One solution to avoid these large deviations is to relax the protein models gradually. But a more profound approach would be to assign tethering forces to all heavy atoms during the minimization process. The tethering constant is a force applied to fix atomic coordinates on predefined positions and the strength of tethering force affects the extent of movement of the atoms from the initial coordinates. While tethering the well-defined main chain atoms, the side chains are allowed to move and adjust their position in order to minimize the total potential energy. Tethering of protein is known to have significant impact on the rates and mechanisms of protein folding [33]. Tethering energy term, symbolized by E_{tether} in Eq. (1) is a tethering term of the C_α and C_β atoms of the model to their initial positions. We implemented the tethering energy term by means of the "*Molecular Elements package*" in the MESHI software. Tethering spring constants have been set to 1 Energy Unit/Å.

Knowledge-based atomic pairwise potential of mean force energy term

The final term in Eq. (1), $E_{kbpairwise}$ is an implementation of the knowledge-based potential of mean force [16]. The original work is based on the interaction statistics of 167 atom types derived by counting of pairwise atomic contact frequencies of proteins from a selection of 500 files from the Protein Data Bank (PDB) having high resolution (1.8 Å or better), low homology, and high quality. In the original study, the weight (w) of KB01 [16] potential has been set to w = 1.0, which is near-optimally weighted [17]. We use the same weight for our refinement protocol. The knowledge-based potential of mean force has been implemented via "*Energy package*" of MESHI [19].

Minimization Protocol

The 3Drefine minimization involves 200,000 steps of energy minimization using limited memory Broyden-Fletcher-Goldfarb-Shannon (L-BFGS) algorithm [34] or until convergence to machine precision, which is carried out by the "*Optimizers package*" in MESHI [19] framework. The backbone structure is refined primarily by the bonded terms of the ENCAD potential [30] and the knowledge-based potential of mean force energy term [16] while the Tethering energy term plays crucial role in optimizing the side chains. The hydrogen bonding network is updated during the minimization by using the explicit Hydrogen bonding energy term described in Eq. (5). The energy minimized model is the final refined model.

Data Set Preparation

To assess the performance of 3Drefine approach, we collected the refinement targets on recent critical assessment of techniques for protein structure prediction (CASP) [35]. To further test the protocol on a large benchmark of 107 CASP9 targets, we used the initial models generated by our structure prediction method, MULTICOM-REFINE [36], participated during the CASP9 experiment. The structure refinement category has been introduced since CASP8 [35]. In these experiments, the predictors were given a starting model for refinement in a blind mode. These starting models had been generated by the CASP structure prediction servers and the organizers evaluated it to be among the best prediction for each target. Although 3Drefine run has been performed after the CASP8 and CASP9 refinement experiments, we ensured same modeling conditions as the CASP blind predictors so that the performance of 3Drefine can be directly compared with the other state-of-the-art refinement methods participating in CASP8 and CASP9.

Metrics used for evaluation

We determine the quality of the structural refinement by observing the changes in global topologies of the models before and after refinement with respect to their native structures. We also determine the local structural qualities of the initial and refined model in order to measure the physical reasonableness of the structure. We have focused on GDT-HA score [37], TM-score [38] and C_α RMSD [39] which are measures of the global positioning of C_α atoms. To evaluate the local qualities of the models, we use MolProbity score [40].

GDT-HA score

GDT-HA [37] score measure the fraction of Cα atoms that are positioned correctly. It counts the average percentage of residues with C_α distance from the native structure residues below 0.5, 1, 2, and 4 Å, respectively, after optimal structure superposition. GDT-HA is related to GDT-TS, which uses cutoffs of 1, 2, 4, and 8 Å. Therefore, GDT-HA is more sensitive to small structural errors. Because of the strong mutual correlation between GDT-HA and GDT-TS score, we chose to only use GDT-HA in our analysis. GDT-HA score has been a widely used scoring function to measure the global positing of C_α atoms in CASP experiments [41-43]. It ranges from [0, 1] with higher value indicating better accuracy.

Template Modeling score (TM-score)

TM-score [38] is a variation of the Levitt–Gerstein (LG) [44] score. It is a global measure of similarity of structural topologies. TM-score is defined as follows:

$$\text{TM}-\text{score}=max\left[\frac{1}{L_N}\sum_{i=1}^{L_T}\frac{1}{1+(d_i/d_0)^2}\right] \quad (6)$$

where L_N is the length of the native structure, L_T is the length of the aligned residues to the template structure, d_i is the distance between the i^{th} Cα pair of aligned residues, and d_0 is a scale to normalize the match difference. "*max*" represents that the maximum value is considered after optimal spatial superposition. The value of the constant d_0 is expressed in the original work [38] as:

$$d_0=1.24\sqrt[3]{L_N-15}-1.8 \quad (7)$$

Like GDT-HA score, TM-score also lie in [0, 1] with higher TM-score suggests enhanced accuracy. However, rather than using specific distance cutoffs and focusing only on the fractions of structures as described for GDT-HA score, all the residues of the modeled proteins are evaluated in the TM-score. Furthermore, TM-score does not depend on the protein length. A TM-score > 0.5 indicates that the proteins share the same fold [45].

RMSD

We evaluated the Root Mean Square Deviation (RMSD) [39] of the C_α positions of the atoms in order to determine the average distance between the C_α atoms after superposition. Similar to GDT-HA and TM-score, RMSD is a global measure of the correct positioning of the C_α atoms. However, since RMSD is based on a single superposition, lacking any kind of distance cutoffs, there is a week correlation between GDT-HA and RMSD score. Also, unlike GDT-HA or TM-score; a lower RMSD value indicates that the protein model is close to its native state. Even if the coordinates of only a few atoms undergo large atomic changes, RMDS becomes high; making RMDS sensitive to small structural errors.

MolProbity

MolProbity [40] is an all-atom measure of the physical correctness of a structure based on statistical analysis of high-resolution protein structures. It is basically a log-weighted combination of the clashscore, percentage Ramachandran not favored and percentage bad side-chain rotamers, giving one number that reflects the crystallographic resolution at which those values would be expected. The MolProbity score is calculated as:

$$\begin{aligned}\text{MolProbityscore} = {} & 0.42574 * \ln(1+clashscore) \\ & +0.32996 * \ln(1+\max(0, rota_{out}1)) \\ & +0.24979 * \ln(1+\max(0, 100\, rama_{iffy}-2))+0.5\end{aligned} \quad (8)$$

where *clashscore* is the number of unfavorable steric overlaps ≥ 0.4 Å, including Hydrogen atoms, and $rota_{out}$ and $rama_{iffy}$ are the percentages of the outliers of the side-chain rotamers and the backbone torsion angles, respectively. Thus MolProbity is sensitive to steric clashes, rotamer outliers, and Ramachandran outliers. The weighting factors were computed from a log-linear fit to crystallographic resolution on a filtered set of PDB structures, so that a protein's MolProbity score is the resolution at which it's MolProbity score would be the expected value. Thus, lower MolProbity scores indicate more physically realistic models. Unlike the other scoring measures we use, MolProbity is not native-dependent that is; the native structure is not required to calculate it. This difference makes MolProbity score significantly distinct from the other scoring function used in this work.

RESULTS AND DISCUSSION

We first present the analysis on the relative importance of the various energy terms used in 3Drefine approach. Then the overall results obtained by using 3Drefine refinement protocol have been evaluated on recent critical assessment of techniques for protein structure prediction (CASP) [35] in the refinement category along with a comparative analysis of 3Drefine against all the groups participated in CASP8 and CASP9 refinement experiments [13,14] together with a recently published refinement method called FG-MD [18]. We also examine the local qualities of the CASP8 and CASP9 refinement targets in detail before and after refinement and compared that with the qualities of the native structures. Finally we assess the performance of 3Drefine on 107 CASP9 targets using the initial models generated by our structure prediction method, MULTICOM-REFINE [36] during the CASP9 experiment.

Effects of various energy terms

To examine the detailed effects of various energy terms; we split 3Drefine into six different runs:

1. $E_{bondlength} + E_{bondangle} + E_{torsion:}$ Minimization using only the bonded terms of ENCAD potential [30]. These are the first three terms of Eq. (1).
2. Ebondlength + Ebondangle + Etorsion + Etether: Tethering energy term has been added to the bonded terms of ENCAD potential.
3. $E_{bondlength} + E_{bondangle} + E_{torsion} + E_{tether} + E_{hydrogenbonds}$: Explicit hydrogen bonding potential has been added to the ENCAD bonded terms and Tethering energy term.
4. $E_{bondlength} + E_{bondangle} + E_{torsion} + E_{hydrogenbonds} + E_{kbpairwise}$: ENCAD bonded terms together with the hydrogen bonding potential and knowledge-based potential of mean force [16]. This is basically the total energy defined in Eq. (1) without the Tethering energy term.
5. $E_{bondlength} + E_{bondangle} + E_{torsion} + E_{tether} + E_{kbpairwise}$: ENCAD bonded terms together with the Tethering energy term and knowledge-based potential of mean force [16]. The hydrogen bonding potential is omitted here from the total energy described in Eq. (1).
6. $E_{bondlength} + E_{bondangle} + E_{torsion} + E_{hydrogenbonds} + E_{tether} + E_{kbpairwise}$: This is the total energy used in the 3Drefine refinement as presented in Eq. (1).

Table I summarizes the average results on CASP8 and CASP9 refinement targets. First, the combination of only the bonded terms of ENCAD potential [30] exhibits slight degrade in the quality of global topology as measured by cumulative GDT-HA score, cumulative TM-score, average RMSD score (i.e. cumulative GDT-HA from 6.898 to 6.843, cumulative TM-score from 9.316 to 9.309 and average RMSD from 3.004 Å to 3.003 Å for CASP8 refinement targets and cumulative GDT-HA from 7.319 to 7.298, cumulative TM-score from 10.368 to 10.355 and average RMSD from 4.344 Å to 4.353 Å for CASP9 refinement targets).

After adding the Tethering energy term to the bonded terms of ENCAD potential [30], an apparent improvement is observed over the ENCAD bonded terms only with cumulative GDT-HA score, cumulative TM-score, average RMSD score of 6.890, 9.318 and 3.003 Å respectively for the CASP8 targets and 7.358, 10.372, 4.344 Å respectively for the CASP9 targets. Although, the quality of the models on an average are not improved as compared to the starting models for CASP8 targets except for a slight improvement in the cumulative

TM-score, Tethering energy term proves to be a beneficial addition to the 3Drefine total energy potential. For CASP9 targets, only the addition of Tethering energy term demonstrates improvement over the starting models in terms of cumulative GDT-HA score (from 7.319 to 7.358) and cumulative TM-score (from 10.368 to 10.372), while the average RMSD remains unaltered.

Addition of the hydrogen-bonding potential to the bonded terms of ENCAD potential and Tethering energy term seems not to affect the quality of the models compared to bonded terms of ENCAD potential and Tethering potential in terms of cumulative GDT-HA score and cumulative TM-score. However, the average RMSD score has been reduced when compared to the starting models (i.e. average RMSD from 3.004 Å to 2.978 Å for CASP8 refinement targets and from 4.344 Å to 4.339 Å for CASP9 refinement targets).

To further test the effects of adding Tethering energy term, we execute the minimization by omitting the Tethering potential from the 3Drefine total energy term described in Eq. (1). The cumulative GDT-HA score has been reduced for both CASP8 (from 6.898 to 6.896) and CASP9 targets (from 7.319 to 7.262). An increase in average RMSD can also be observed for CASP9 targets (from 4.344 Å to 4.350 Å). For CASP8 targets, although the average RMSD is less than the initial models (2.998 Å), this is worse than the combination of bonded terms of ENCAD potential with Tethering energy term and hydrogen bonding potential.

To justify the use of explicit hydrogen bonding potential, we run 3Drefine minimization without the hydrogen bonding energy term in Eq. (1). The results shows further increase in the RMSD scores (2.999 Å for CASP8 refinement targets and 4.344 Å for CASP9 refinement targets).

Finally, we perform the minimization using all the energy terms as presented in Eq. (1). This approach has achieved the highest cumulative GDT-HA score, cumulative TM-score and lowest RMSD score for both CASP8 and CASP9 refinement targets.

Assessment of 3Drefine on CASP8 Refinement Experiment

We evaluate the performance of 3Drefine refinement protocol on all the 12 targets in CASP8 refinement experiment and compared it with all other groups participating in CASP8 refinement category along with FG-MD [18] which is a recent work and did not participate in CASP8. For the assessment of the results, we gather the performance of all the participating groups in CASP8 in terms of cumulative GDT-HA score, cumulative TM-score, average RMSD score and average MolProbity from the previously published works [13,18] on the assessment of CASP8 refinement experiment.

The groups have been ordered based on the cumulative GDT-HA score of refined models for all the 12 targets. Upper part of Table II summarizes the overall result of 3Drefine with FG-MD [18] and top five groups participating in CASP8. A complete list of the CASP8 groups is listed in Table S1 in the supplementary information document at http://sysbio.rnet.missouri.edu/3Drefine/download.html.

The “Null” group basically represents the starting model provided by the CASP organizers for refinement. Groups that perform worse than Null group have on average degraded the model rather than improving it. The results demonstrate other than FG-MD; 3Drefine is the method that could consistently drive the initial model closer to the experimental structure in terms of cumulative GDT-HA, TM-scores and average RMSD.

Overall, the cumulative GDT-HA and TM-score are 1.04% and 1.4% higher and RMSD is 0.123 Å lower than the second best LEE group participating in CASP8. The recent work, FG-MD, however, outperforms 3Drefine in these measures; although the performance is comparable. In terms of average MolProbity score, 3Drefine models performed better than FG-MD, suggesting improved local qualities of the structures after refinement.

When compared with the starting models, 3Drefine exhibits consistent improvement in both global and local topologies of the initial structures. Out of the 12 CASP8 initial models, 3Drefine improves GDT-HA score for 9, TM-score for 12, RMSD for 11 and MolProbity score for 10 targets. In Figures 1A and 1B, we present the score changes (i.e. score after refinement – score before refinement) of the models refined by 3Drefine in terms of GDT-HA, TM-score, RMSD and MolProbity score against the TM-score of the starting model before refinement, which show that the qualities of the models refined by 3Drefine had been improved in most of the cases.

Two representative examples of CASP8 refinement have been presented in Fig. 2. For the target TR464, 3Drefine refinement resulted in a 1.3% increase in GDT-HA score, a 0.1 % increase in TM-score and a 56 % decrease in MolProbity score. For the target TR432, the GDT-HA and TM-score improvement are 0.8 % and 0.4 % respectively while MolProbity improvement is 6 %.

Assessment of 3Drefine on CASP9 Refinement Experiment

There were 14 targets available for refinement in CASP9 Refinement Experiment with length from 69 to 159 residues [14]. Along with the initial models and global distance test total score (GDT-TS), predictors were provided with hints about the focus regions, that is, groups of residues that need refinement. To ensure strict blind prediction, we do not use the hints or the starting GDT-TS score for 3Drefine run.

A summary of 3Drefine with top five CASP9 predictors ordered based on the cumulative GDT-HA score of the first model for all 14 targets have been listed in the lower part of Table II. The results for the other groups have been adopted from the published works on CASP9 refinement assessment [14,18] and a complete list of all the CASP9 groups has been presented in Table S2 in the supplementary information file.

The results show that 3Drefine, FG-MD, ZHANG and SEOK were able to consistently refine the staring model on the basis of GDT-HA, TM-score and RMSD score. However, the both ZHANG and SEOK models have MolProbity score higher than the initial model indicating degradation in the local qualities of the structures. The MolProbity improvement for 3Drefine was 30.9% more than the best ZHANG group participating in CASP9. When compared to FG-MD, 3Drefine improves the MolProbity score by 3.8 %. Overall, 3Drefine protocol demonstrates consisted refinement of the initial model in terms of cumulative GDT-HA, cumulative TM-score, average RMSD and average MolProbity score; outperforming all the CASP9 predictors including the recent work FG-MD in terms of cumulative GDT-HA score.

In Fig. 3A-3D we present the changes in GDT-HA, TM-score, RMSD and MolProbity score before and after refinement against the TM-score of the starting models for all the 14 CASP9 targets by 3Drefine and FG-MD. There are 9,13,11 and 12 cases when 3Drefine can improve the GDT-HA, TM-score, RMSD and MolProbity scores respectively; while FG-MD do so for 11, 9, 9 and 13 CASP9 targets.

Overall, the performance of 3Drefine is comparable to FG-MD in terms of its ability to enhance global qualities of the initial structure, that is, improvement on GDT-HA, TM-score

and RMSD scores; although 3Drefine performs slightly better than FG-MD on CASP9 targets on these measures. However, with respect to the improvement of local qualities of structures, 3Drefine clearly outperforms FG-MD. The average MolProbity score of 3Drefine is 3.8 % better than that of FG-MD with an overall improvement of 16.6 % in MolProbity score over the starting models. Fig. 3E and 3F show two typical examples of refinement on CASP9 targets. For the target TR606, 3Drefine refinement resulted in a 3.1% increase in GDT-HA score, a 0.5 % increase in TM-score and a 21 % decrease in MolProbity score. There are 2.1 % and 0.6 % improvement of the GDT-HA and TM-score respectively and 8.1 % improvement in MolProbity score for the target TR622.

A closer look at the local qualities of CASP8 and CASP9 Refinement Targets

Although the global structural measures like GDH-HA score, TM-score or the C_α RMSD scores provide the accuracy predicted protein models, they are primarily focusing on the correctness of the backbone conformation of proteins and often fail to delve into finer atomic details of the predicted models. For instance, staggered χ angles are crucial details in estimating the qualities of protein structure {Lovell, 2000 #122}; but they are not considered in the global quality measures. Also the unfavorable steric clashes are strongly correlated with quality of a protein structure, with clashes reduced nearly to zero in the well-ordered parts of very high-resolution crystal structures {Arendall, 2005 #121}. In order to investigate these minute but vital aspects of models, we decide to perform a detailed analysis on the local structural qualities of the CASP8 and CASP9 refinement targets using MolProbity score - a single composite metric for local model quality.

All-atom contact, rotamers, and Ramachandran analysis are fundamental to the MolProbity structure-validation approach {Davis, 2007 #119}, which is widely accepted standard in macromolecular crystallography. CASP8 marks the first use of the MolProbity score for evaluation of non-experimental protein models. It is a very sensitive and demanding measure; attracting lot of attentions in serious works to assess the protein model qualities beyond C_α accuracy metrics {Keedy, 2009 #120}.

Table III summarizes the MolProbity score for all CASP8 and CASP9 refinement targets. For each targets the score for initial model, score after 3Drefine refinement has been presented along with the MolProbity score for the native structures. It can be clearly seen that apart from a few targets (two targets in CASP8 and two targets in CASP9); the MolProbity score is always lower in the native structures when compared with the staring models. On an average, the MolProbity score for the initial structures in CASP8 and CASP9 are 2.46 and 2.42 respectively while the native structures have an average score of 1.68 and 1.44 respectively. The difference in the MolProbity score for the initial models and the native structures undoubtedly demonstrate the need for the refinement of the local qualities of the starting structures in any refinement protocol.

Promisingly, 3Drefine exhibits improvement in the local qualities of the predicted protein models as measured by MolProbity score in CASP8 and CASP9 refinement targets. Apart from two targets in CASP8 (TR389 and TR432) and two targets in CASP9 (TR530 and TR569), the MolProbity score is always reduced compared to the starting models. Overall, the average MolProbity score for the refined models are 2.34 and 2.10 for CASP8 and CASP9 refinement targets respectively. Although the refined models are still far from achieving the average MolProbity score as the native structures, 3Drefine has certainly enhanced the local model qualities with respect to the initial structures.

Performance of 3Drefine on 107 CASP9 Targets

To further assess the performance of 3Drefine on a large set of target models, we tested the refinement protocol on 107 CASP9 targets generated by our tertiary structure predictor MULTICOM-REFINE [36] that participated in the CASP9 experiment. We selected the first predicted model generated by MULTICOM-REFINE as the starting model for 3Drefine run for each of the 107 CASP9 targets. Similar to our testing strategy of 3Drefine for CASP8 and CASP9 refinement experiments, we performed the refinement in a blind mode, that is, without the knowledge of the native structure.

We observe a consistent improvement in the global qualities of the starting models after the refinement as measured by the GDT-HA, TM-score, and RMSD score. There were 59, 89 and 87 cases when 3Drefine brings the starting model closer to the native ones with respect to GDT-HA, TM-score and RMSD score respectively.

Overall, there was a 0.4 % increase in cumulative GDT-HA score and 0.1 % increase in cumulative TM-score for the refined models over the initial structures predicted by MULTICOM-REFINE for all the 107 CASP9 targets. The average RMSD of the refined models was 0.007 Å lower than the starting models.

The changes of the RMSD score after refinement over the TM-score of the starting models has been shown as a scatter plot in Fig. 4. In Fig 5A and 5B, we present the change of GDT-HA and TM-score before and after refinement against the initial TM-score for the 107 CASP9 targets. We controlled the initial TM-score Detailed histograms of changes in the score have been shown in Fig. S1A and S1B and is available the supplementary information document. These results demonstrate the ability of 3Drefine protocol for consistent refinement of the CASP9 predicted models to bring it closer to its native structure in terms of global qualities of the structures. Most significant improvements have been observed when the TM-score of the starting model is > 0.5, that is, when the predicted models share the same fold with the native structure [45].

3Drefine refinement results for the CASP8 and CASP9 refinement experiments along with the refinement of 107 CASP9 models are freely available at: http://sysbio.rnet.missouri.edu/3Drefine/download.html

CONCLUSION

Despite attracting constant attention by the researches, protein structure refinement problem remains a largely unsolved problem [46]. Because of the strong mutual association between the back-bone positioning and side-chain conformation of a protein model [47] simultaneous refinement of both the global topologies and local structural qualities of a protein structure is intended. Unfortunately, apart from a few promising works in the recent years [15-18] majority of the structure refinement protocols fail to achieve this goal. Addressing this problem successfully would have major implications in resolving the bottleneck to apply computational protein structure prediction methods in structure-based drug design [48], protein docking [49] and prediction of biological functions based on protein structure [50]. As per the results of most recent critical assessment of techniques for protein structure prediction (CASP) refinement experiments, CASP8 and CASP9, there may be substantial room for improvement in the refinement category [13,14].

In this work, we present a computationally efficient and reliable protocol for protein structure refinement, called 3Drefine. This method is a combination of two steps of minimization: Optimizing Hydrogen Bonding Network and Energy Minimization using a composite physics and knowledge based force fields, which is implemented within the

MESHI [19] molecular modeling framework. It takes only few minutes (usually less than 5 minutes) of CPU time to refine protein models of usual length using 3Drefine.

3Drefine was tested in blind mode for CASP8 and CASP9 refinement targets in a completely automated manner, without using the knowledge about the information provided by CASP organizers to focus on certain parts of the proteins for refinement. We observe that 3Drefine has the ability to consistently bring the model closer to the native structure. The models refined by 3Drefine have shown improvement of the global topologies of the starting models as measured by GDT-HA, TM-score and C_α RMSD to native structures as well as the local structure qualities as measured by the MolProbity score. The overall results of 3Drefine were better than or comparable to the other state-of the-art methods participating in CASP refinement category.

We also tested the performance on a large benchmark of 107 CASP9 targets by using MULTICOM-REFINE [36] as a structure prediction method to generate the initial structures. 3Drefine demonstrates consistent improvement in qualities of the initial models.

Although promising, the improvement in qualities of the starting models after 3Drefine refinement is often modest. This is the case with almost all other existing state-of-the-art refinement protocols primarily due to the limited accuracy of physics based empirical force fields used predominantly in the refinement methods. Broader samplings around the initial conformation of the protein using these force fields impose the risk of degrading the model quality instead of improving it. As the result, the refinement algorithms often rely on a more conservative strategy to sample locally around the starting structures producing improvement only in general physicality of the models rather than substantially improving the backbone positioning. Also, with the progress in the structure prediction pipelines, the qualities of the starting models are getting improved. Therefore, adopting more adventurous global search techniques at the cost of consistency that can improve the overall fold in the starting models are becoming less common amongst the refinement pipelines. Even with the unadventurous strategies, the existing refinement protocols are often inconsistent as indicated by the results of CASP8 and CASP9 refinement experiment with majority of groups degrading the model qualities on an average. The unique nature of 3Drefine protocol is consistency. Around 80% of the times 3Drefine has improved the global qualities in the starting structures in CASP8 and CASP9 refinement targets. Also, the ability of 3Drefine to simultaneously improve the backbone positioning and local model qualities is encouraging. Future directions would be to use 3Drefine method in conjunction with some global search technique that can substantially improve the overall fold in the starting models together with the improvement in general physicality and local qualities of the models.

We conclude that 3Drefine can become a reliable and efficient method in protein structure refinement. The success of the protocol in improving accuracy of the initial models in a computationally inexpensive way for CASP refinement targets, where the initial model has been generated by a variety of structure prediction techniques, suggests that 3Drefine can be adopted as a final step in computational structure prediction pipeline.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank the anonymous reviewers for a very careful review. Work partially supported by a NIH grant (R01GM093123) to JC.

References

1. Sali A, Blundell T. Comparative protein modelling by satisfaction of spatial restraints. Protein Structure by Distance Analysis. 1994; 64:C86.
2. Fischer D. 3D-SHOTGUN: A novel, cooperative, fold-recognition meta-predictor. Proteins: Structure Function and Bioinformatics. 2003; 51:434–441.
3. Cheng J. A multi-template combination algorithm for protein comparative modeling. BMC Structural Biology. 2008; 8:18. [PubMed: 18366648]
4. Wu S, Zhang Y. LOMETS: a local meta-threading-server for protein structure prediction. Nucleic acids research. 2007; 35:3375–3382. [PubMed: 17478507]
5. Zhang Y. Template-based modeling and free modeling by I-TASSER in CASP7. Proteins: Structure Function and Bioinformatics. 2007; 69:108–117.
6. Joo K, Lee J, Lee S, Seo JH, Lee SJ. High accuracy template based modeling by global optimization. Proteins: Structure Function and Bioinformatics. 2007; 69:83–89.
7. Misura KMS, Chivian D, Rohl CA, Kim DE, Baker D. Physically realistic homology models built with ROSETTA can be more accurate than their templates. Proceedings of the National Academy of Sciences. 2006; 103:5361–5366.
8. Wang Q, Canutescu AA, Dunbrack RL. SCWRL and MolIDE: computer programs for side-chain conformation prediction and homology modeling. Nature protocols. 2008; 3:1832–1847.
9. Krivov GG, Shapovalov MV, Dunbrack RL Jr. Improved prediction of protein side-chain conformations with SCWRL4. Proteins: Structure Function and Bioinformatics. 2009; 77:778–795.
10. Levitt M. Accurate modeling of protein conformation by automatic segment matching. Journal of molecular biology. 1992; 226:507–533. [PubMed: 1640463]
11. Eswar N, Webb B, Marti-Renom MA, Madhusudhan M, Eramian D, et al. Comparative protein structure modeling using Modeller. Curr Protoc Protein Sci. 2007; 2:15–32.
12. Qian B, Raman S, Das R, Bradley P, McCoy AJ, et al. High-resolution structure prediction and the crystallographic phase problem. Nature. 2007; 450:259–264. [PubMed: 17934447]
13. MacCallum JL, Hua L, Schnieders MJ, Pande VS, Jacobson MP, et al. Assessment of the protein-structure refinement category in CASP8. Proteins: Structure Function and Bioinformatics. 2009; 77:66–80.
14. MacCallum JL, Pérez A, Schnieders MJ, Hua L, Jacobson MP, et al. Assessment of protein structure refinement in CASP9. Proteins: Structure Function and Bioinformatics. 2011
15. Jagielska A, Wroblewska L, Skolnick J. Protein model refinement using an optimized physics-based all-atom force field. Proceedings of the National Academy of Sciences. 2008; 105:8268.
16. Summa CM, Levitt M. Near-native structure refinement using in vacuo energy minimization. Proceedings of the National Academy of Sciences. 2007; 104:3177.
17. Chopra G, Kalisman N, Levitt M. Consistent refinement of submitted models at CASP using a knowledge-based potential. Proteins: Structure Function and Bioinformatics. 2010; 78:2668–2678.
18. Zhang J, Liang Y, Zhang Y. Atomic-Level Protein Structure Refinement Using Fragment-Guided Molecular Dynamics Conformation Sampling. Structure. 2011; 19:1784–1795. [PubMed: 22153501]
19. Kalisman N, Levi A, Maximova T, Reshef D, Zafriri-Lynn S, et al. MESHI: a new library of Java classes for molecular modeling. Bioinformatics. 2005; 21:3931–3932. [PubMed: 16105898]
20. Rose GD, Fleming PJ, Banavar JR, Maritan A. A backbone-based theory of protein folding. Proceedings of the National Academy of Sciences. 2006; 103:16623–16633.
21. Dill KA, Bromberg S, Yue K, Fiebig KM, Yee DP, et al. Principles of protein folding--a perspective from simple exact models. Protein science: a publication of the Protein Society. 1995; 4:561. [PubMed: 7613459]
22. Wallin S, Shakhnovich EI. Understanding ensemble protein folding at atomic detail. Journal of Physics: Condensed Matter. 2008; 20:283101.
23. Engler N, Ostermann A, Niimura N, Parak FG. Hydrogen atoms in proteins: Positions and dynamics. Proceedings of the National Academy of Sciences. 2003; 100:10243.

24. McDonald IK, Thornton JM. Satisfying hydrogen bonding potential in proteins. Journal of molecular biology. 1994; 238:777–793. [PubMed: 8182748]
25. Chen J, Brooks CL III. Can molecular dynamics simulations provide high-resolution refinement of protein structure? Proteins: Structure Function and Bioinformatics. 2007; 67:922–930.
26. Vriend G. WHAT IF: a molecular modeling and drug design program. Journal of molecular graphics. 1990; 8:52. [PubMed: 2268628]
27. Brünger AT, Karplus M. Polar hydrogen positions in proteins: empirical energy placement and neutron diffraction comparison. Proteins: Structure Function and Bioinformatics. 1988; 4:148–156.
28. Word JM, Lovell SC, Richardson JS, Richardson DC. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation1. Journal of molecular biology. 1999; 285:1735–1747. [PubMed: 9917408]
29. Li Y, Roy A, Zhang Y. HAAD: a quick algorithm for accurate prediction of hydrogen atoms in protein structures. PloS one. 2009; 4:e6701. [PubMed: 19693270]
30. 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. Computer Physics Communications. 1995; 91:215–231.
31. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers. 1983; 22:2577–2637. [PubMed: 6667333]
32. Amir EAD, Kalisman N, Keasar C. Differentiable, multi-dimensional, knowledge-based energy terms for torsion angle probabilities and propensities. Proteins: Structure Function and Bioinformatics. 2008; 72:62–73.
33. Friedel M, Baumketner A, Shea JE. Effects of surface tethering on protein folding mechanisms. Proceedings of the National Academy of Sciences. 2006; 103:8396–8401.
34. Liu DC, Nocedal J. On the limited memory BFGS method for large scale optimization. Mathematical programming. 1989; 45:503–528.
35. Moult J, Hubbard T, Fidelis K, Pedersen JT. Critical assessment of methods of protein structure prediction (CASP): round III. Proteins: Structure Function and Bioinformatics. 1999; 37:2–6.
36. Wang Z, Eickholt J, Cheng J. MULTICOM: a multi-level combination approach to protein structure prediction and its assessments in CASP8. Bioinformatics. 2010; 26:882–888. [PubMed: 20150411]
37. Zemla A. LGA: a method for finding 3D similarities in protein structures. Nucleic acids research. 2003; 31:3370–3374. [PubMed: 12824330]
38. Zhang Y, Skolnick J. Scoring function for automated assessment of protein structure template quality. Proteins: Structure Function and Bioinformatics. 2004; 57:702–710.
39. Kabsch W. A solution for the best rotation to relate two sets of vectors. Acta Crystallographica Section A: Crystal Physics Diffraction Theoretical and General Crystallography. 1976; 32:922–923.
40. Chen VB, Arendall WB, Headd JJ, Keedy DA, Immormino RM, et al. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallographica Section D: Biological Crystallography. 2009; 66:12–21.
41. Kopp J, Bordoli L, Battey JND, Kiefer F, Schwede T. Assessment of CASP7 predictions for template-based modeling targets. Proteins: Structure Function and Bioinformatics. 2007; 69:38–56.
42. Cozzetto D, Kryshtafovych A, Fidelis K, Moult J, Rost B, et al. Evaluation of template-based models in CASP8 with standard measures. Proteins: Structure Function and Bioinformatics. 2009; 77:18–28.
43. Keedy DA, Williams CJ, Headd JJ, Arendall WB III, Chen VB, et al. The other 90% of the protein: Assessment beyond the Cαs for CASP8 template-based and high-accuracy models. Proteins: Structure Function and Bioinformatics. 2009; 77:29–49.
44. Levitt M, Gerstein M. A unified statistical framework for sequence comparison and structure comparison. Proceedings of the National Academy of Sciences. 1998; 95:5913.
45. Xu J, Zhang Y. How significant is a protein structure similarity with TM-score= 0.5? Bioinformatics. 2010; 26:889–895. [PubMed: 20164152]

46. Kryshtafovych A, Venclovas Č, Fidelis K, Moult J. Progress over the first decade of CASP experiments. Proteins: Structure Function and Bioinformatics. 2005; 61:225–236.

47. Keskin O, Bahar I. Packing of sidechains in low-resolution models for proteins. Folding and Design. 1998; 3:469–479. [PubMed: 9889162]

48. Wieman H, Kristin T, Anderssen E, Drablos F. Homology-based modelling of targets for rational drug design. Mini reviews in medicinal chemistry. 2004; 4:793–804. [PubMed: 15379646]

49. Méndez R, Leplae R, Lensink MF, Wodak SJ. Assessment of CAPRI predictions in rounds 3–5 shows progress in docking procedures. Proteins: Structure Function and Bioinformatics. 2005; 60:150–169.

50. Arakaki AK, Zhang Y, Skolnick J. Large-scale assessment of the utility of low-resolution protein structures for biochemical function assignment. Bioinformatics. 2004; 20:1087–1096. [PubMed: 14764543]

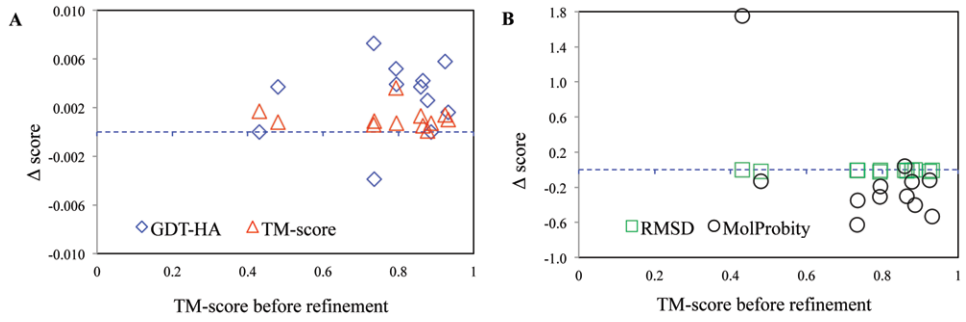

Fig. 1.

Changes in global and local structural qualities using 3Drefine on CASP8 Refinement Targets

(A) Scatter plot of changes in GDT-HA and TM-score. A positive change indicates the quality of the model of a target has been improved by refinement.

(B) Scatter plot of changes in RMSD and MolProbity-score. A negative change indicates the quality of the model of a target has been improved by refinement.

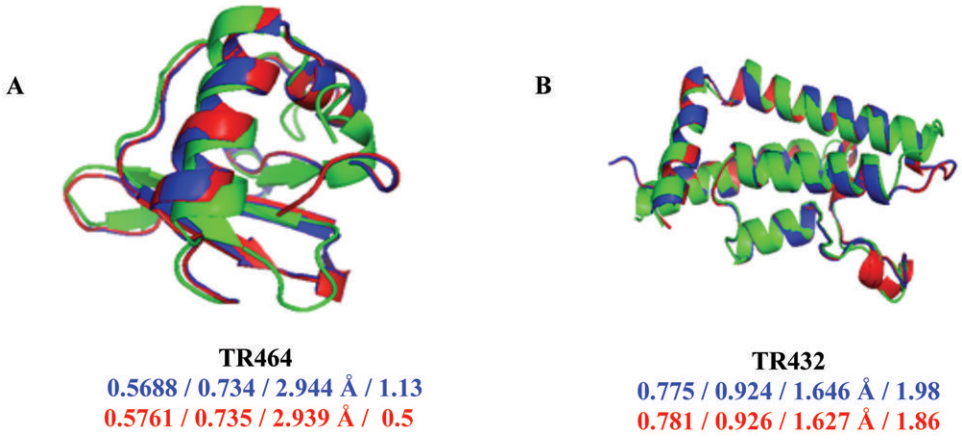

Fig. 2.

Structural superposition of initial model (blue) and refined model using 3Drefine (red) on the native structure (green) for two CASP8 Targets. The values under each model indicate GDT-HA, TM-score, RMSD and MolProbity score respectively before (blue) and after (red) refinement.

(A) Structural superposition for Target TR464.

(B) Structural superposition for Target TR432.

Figures were prepared in PyMOL (The PyMOL Molecular Graphics System, Version 1.4.1, Schrödinger, LLC.).

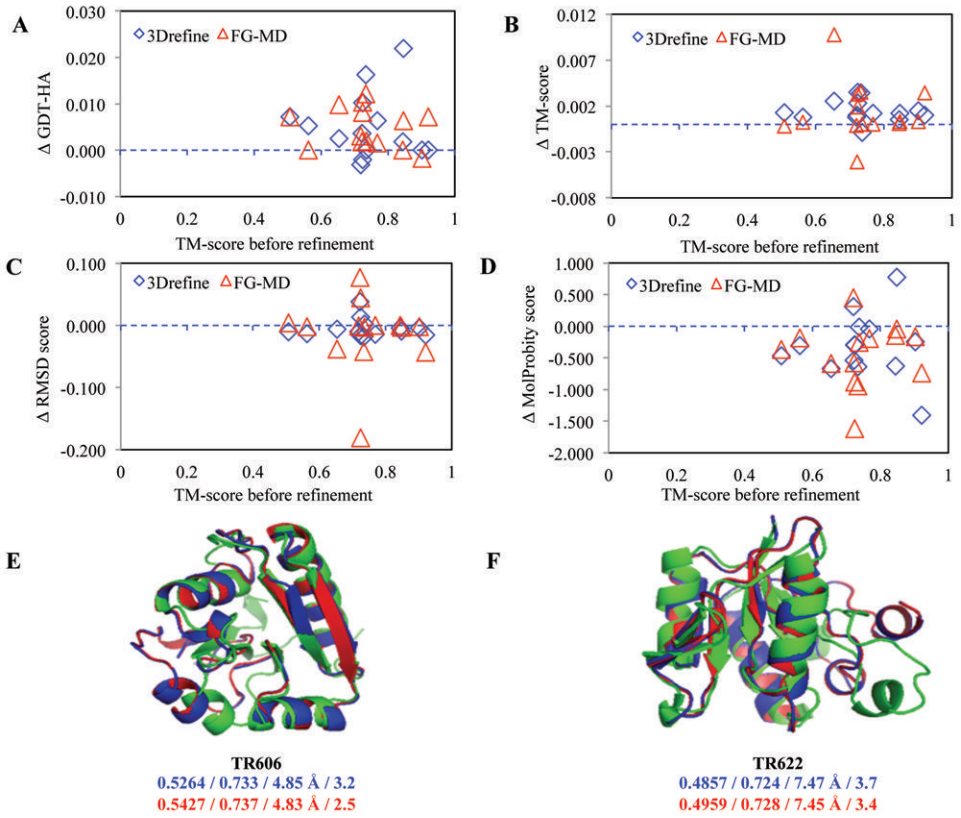

Fig. 3.

Changes in local and global qualities of CASP9 Refinement Targets

(A) Scatter plot for GDT-HA score changes for 3Drefine and FG-MD

(B) Scatter plot for TM-score changes for 3Drefine and FG-MD

(C) Scatter plot for RMSD-score changes for 3Drefine and FG-MD

(D) Scatter plot for MolProbity-score changes for 3Drefine and FG-MD

(E) Structural superposition of initial model (blue) and refined model using 3Drefine (red) on the native structure (green) for CASP9 target TR606.

(F) Structural superposition of initial model (blue) and refined model using 3Drefine (red) on the native structure (green) for CASP9 target TR624.

The values under each model indicate GDT-HA, TM-score, RMSD and MolProbity score respectively before (blue) and after (red) refinement.

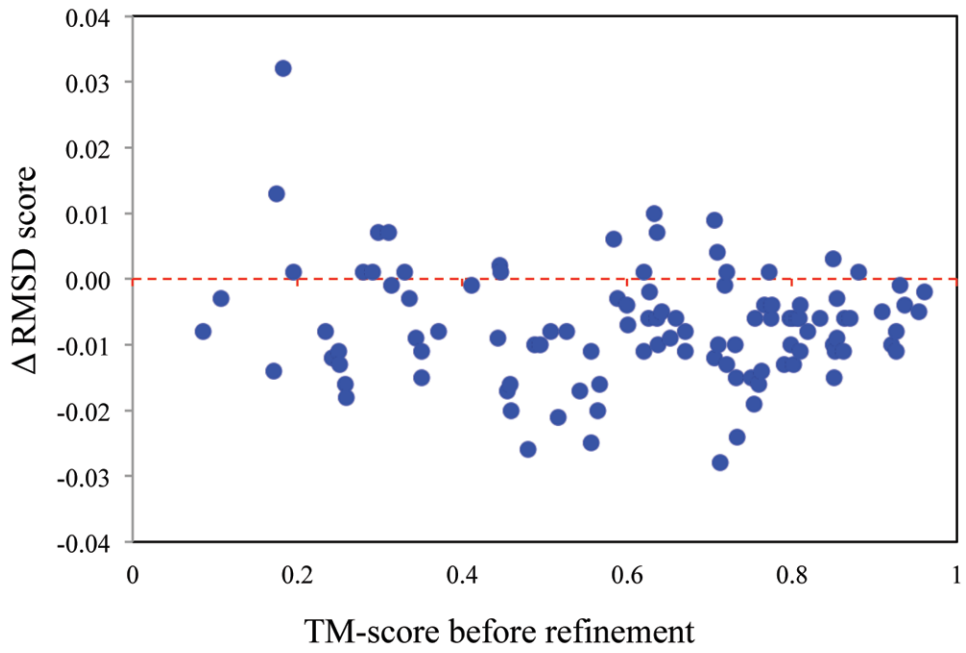

Fig. 4.
Scatter plot of RMSD changes for 107 CASP9 Targets
(Initial models generated using MULTICOM-REFINE).

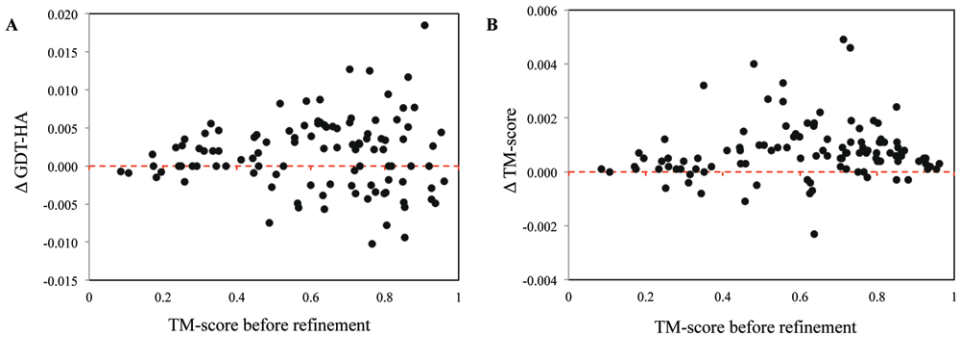

Fig. 5.

Refinement results for 107 CASP9 Targets using 3Drefine (Initial structures generated using MULTICOM-REFINE)

(A) GDT-HA score changes. A positive change indicates the quality of the model of a target has been improved by refinement

(B) TM score changes. A positive change indicates the quality of the model of a target has been improved by refinement.

Table I

Effects of various energy terms of 3Drefine refinement on CASP8 and CASP9 Refinement Targets *

	#	Energy Terms[a]	GDT-HA[b]	TM-score[c]	RMSD[d]
CASP8	1	$E_{bondlength} + E_{bondangle} + E_{torsion}$	6.843	9.309	3.003
	2	$E_{bondlength} + E_{bondangle} + E_{torsion} + E_{tether}$	6.890	9.318	3.003
	3	$E_{bondlength} + E_{bondangle} + E_{torsion} + E_{tether} + E_{hydrogenbonds}$	6.709	9.306	2.978
	4	$E_{bondlength} + E_{bondangle} + E_{torsion} + E_{hydrogenbonds} + E_{kbpairwise}$	6.896	9.318	2.998
	5	$E_{bondlength} + E_{bondangle} + E_{torsion} + E_{tether} + E_{kbpairwise}$	6.941	9.326	2.999
	6	$E_{bondlength} + E_{bondangle} + E_{torsion} + E_{hydrogenbonds} + E_{tether} + E_{kbpairwise}$	**6.979**	**9.362**	**2.953**
CASP9	1	$E_{bondlength} + E_{bondangle} + E_{torsion}$	7.298	10.355	4.353
	2	$E_{bondlength} + E_{bondangle} + E_{torsion} + E_{tether}$	7.358	10.372	4.344
	3	$E_{bondlength} + E_{bondangle} + E_{torsion} + E_{hydrogenbonds} + E_{tether}$	7.347	10.376	4.339
	4	$E_{bondlength} + E_{bondangle} + E_{torsion} + E_{hydrogenbonds} + E_{kbpairwise}$	7.262	10.344	4.350
	5	$E_{bondlength} + E_{bondangle} + E_{torsion} + E_{tether} + E_{kbpairwise}$	7.386	10.382	4.344
	6	$E_{bondlength} + E_{bondangle} + E_{torsion} + E_{hydrogenbonds} + E_{tether} + E_{kbpairwise}$	**7.388**	**10.388**	**4.339**

[a] Combination of various energy terms used in 3Drefine refinement.

[b] Cumulative GDT-HA score.

[c] Cumulative TM-score.

[d] Average RMSD in Å.

* Numbers in bold represent the best score in each category.

Table II

Comparison of 3Drefine results (with FG-MD and Top Five Groups in CASP8 and CASP9 Refinement Experiments)

	Group Name	No. of Targets[a]	GDT-HA[b]	TM-score[c]	RMSD[d]	MolProbity[e]
CASP8	FG-MD[f]	12	6.979	9.362	2.953	2.575
	3Drefine	12	6.932	9.329	2.994	2.349
	Null[g]	12	6.898	9.316	3.004	2.706
	LEE	12	6.86	9.195	3.117	2.613
	LevittGroup	12	6.701	9.16	3.047	2.875
	FAMSD	12	6.562	8.746	3.056	2.796
	SAM-T08-human	12	6.523	9.084	3.105	2.762
	YASARARefine	12	6.407	9.155	3.359	1.071
CASP9	3Drefine	14	7.388	10.388	4.339	2.101
	FG-MD[f]	14	7.387	10.386	4.331	2.183
	ZHANG	14	7.365	10.396	4.338	3.042
	SEOK	14	7.359	10.399	4.259	3.436
	Null[g]	14	7.319	10.368	4.344	2.521
	FAMSD	14	7.284	10.348	4.309	2.55
	FAMS-MULTI	14	7.284	10.348	4.44	2.55
	KNOWMIN	14	7.194	10.182	4.74	2.179

[a] Number of CASP targets in the Refinement Experiment.

[b] Cumulative GDT-HA score based on the first submitted model.

[c] Cumulative TM-score based on the first submitted model.

[d] Average RMSD based on the first submitted model with respect to the native structure in Å.

[e] Average MolProbity score based on the first submitted model.

[f] Not a participating group in CASP8 and CASP9 Refinement Experiment.

[g] The initial models for the CASP refinement experiment.

Table III

MolProbity scores for CASP8 and CASP9 Refinement Targets

	Target Name	Initial MolProbity[a]	Refined MolProbity[b]	Native MolProbity[c]
CASP8	TR389	2.68	2.72	1.05
	TR429	2.61	2.47	1.81
	TR432	2.01	3.76	1.1
	TR435	2.42	2.07	2.0
	TR453	1.13	0.5	1.48
	TR454	3.14	2.83	0.86
	TR461	2.41	1.88	1.62
	TR462	2.06	1.87	2.7
	TR464	3.15	2.75	2.55
	TR469	2.53	2.23	1.86
	TR476	1.98	1.86	2.66
	TR488	3.38	3.25	0.5
Average		**2.46**	**2.34**	**1.68**
CASP9	TR517	1.4	1.36	1.02
	TR530	0.9	1.67	0.64
	TR557	1.5	1.49	1.14
	TR567	1.4	0.77	3.6
	TR568	1.5	1.19	0.56
	TR569	0.7	1.01	2.05
	TR574	3.6	2.93	0.50
	TR576	3.7	3.16	1.46
	TR592	3.5	2.09	1.94
	TR594	2.9	2.65	1.01
	TR606	3.2	2.56	0.81
	TR614	4	3.7	1.67
	TR622	3.7	3.4	1.42
	TR624	1.9	1.43	2.33
Average		**2.42**	**2.10**	**1.44**

[a] MolProbity score for the starting models.

[b] MolProbity score after 3Drefine refinement.

[c] MolProbity score for Native Structures.