PROTEINS
STRUCTURE ■ FUNCTION ■ BIOINFORMATICS

Refinement of protein structure homology models via long, all-atom molecular dynamics simulations

Alpan Raval,[1] Stefano Piana,[1]* Michael P. Eastwood,[1] Ron O. Dror,[1] and David E. Shaw[1,2]*

[1] D. E. Shaw Research, New York, New York 10036

[2] Center for Computational Biology and Bioinformatics, Columbia University, New York, New York 10032

ABSTRACT

Accurate computational prediction of protein structure represents a longstanding challenge in molecular biology and structure-based drug design. Although homology modeling techniques are widely used to produce low-resolution models, refining these models to high resolution has proven difficult. With long enough simulations and sufficiently accurate force fields, molecular dynamics (MD) simulations should in principle allow such refinement, but efforts to refine homology models using MD have for the most part yielded disappointing results. It has thus far been unclear whether MD-based refinement is limited primarily by accessible simulation timescales, force field accuracy, or both. Here, we examine MD as a technique for homology model refinement using all-atom simulations, each at least 100 μs long—more than 100 times longer than previous refinement simulations—and a physics-based force field that was recently shown to successfully fold a structurally diverse set of fast-folding proteins. In MD simulations of 24 proteins chosen from the refinement category of recent Critical Assessment of Structure Prediction (CASP) experiments, we find that in most cases, simulations initiated from homology models drift away from the native structure. Comparison with simulations initiated from the native structure suggests that force field accuracy is the primary factor limiting MD-based refinement. This problem can be mitigated to some extent by restricting sampling to the neighborhood of the initial model, leading to structural improvement that, while limited, is roughly comparable to the leading alternative methods.

Proteins 2012; 80:2071–2079.

Key words: protein structure prediction; force fields; CHARMM22*; sampling; protein folding.

INTRODUCTION

The determination of high-resolution protein structure represents a major thrust of modern molecular biology. High-resolution structures are often required, for example, to infer the biological function of a protein[1–3] or for structure-based drug design.[2,4–7] Although the computational prediction of protein structure has been greatly facilitated by the frequent availability of crystal structures for homologous proteins, the resulting homology models are typically of low resolution, with errors of 3 Å in the atomic coordinates. The development of methods for the refinement of low-resolution homology models into more useful, high-resolution structural models is thus crucial for further progress in structure prediction.[2,8–14]

The refinement category of the biennial Critical Assessment of Structure Prediction (CASP) experiment, known as CASPR, represents a de facto benchmark for testing refinement methodologies. The organizers choose a number of target proteins, along with a homology model for each. These homology models are selected from among the best models submitted to other categories of the CASP experiment and thus represent state-of-the-art attempts at predicting a protein structure. They are particularly challenging starting points for refinement because they have generally already undergone some sort of refinement during the homology modeling process.[12] Submissions to CASPR are evaluated principally by a metric known as the global distance test total score

Additional Supporting Information may be found in the online version of the article.

*Correspondence to: Stefano Piana or David E. Shaw, D. E. Shaw Research, New York, NY 10036. E-mail: Stefano.Piana-Agostinetti@DEShawResearch.com or David.Shaw@DEShawResearch.com

Received 7 February 2012; Revised 3 April 2012; Accepted 11 April 2012
Published online 19 April 2012 in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/prot.24098

(GDT-TS), defined as the average, over four different cutoffs, of the percentage of residues in the model that are within the cutoff distance of the corresponding residues in the native structure.[15]

Commonly used approaches to structure refinement have typically relied upon some combination of physics-based[16–22] and knowledge-based[9,16–18,23–25] energy functions, and have performed local energy minimization,[17,18,24] Monte Carlo simulation,[9,16,17,23,26,27] or molecular dynamics (MD) simulation[19–23,25,28–30] to sample regions of conformational space. Many of these approaches then select the sampled conformations they predict to be most "native-like" by applying scoring functions that are unrelated to the energy function used for sampling. The application of these refinement methodologies to CASPR targets has yielded results that are largely disappointing. Examination of the top-choice model submitted by each group in the CASP8 refinement category, for example, reveals that only two of 25 groups achieved an average increase in GDT-TS over that of the initial homology model,[12] and even in these two cases the improvement was small (<0.3%). Similar results were obtained in the CASP9 experiment.[13] Although MD simulations with a physics-based force field should in principle produce structures that cluster around the native state, providing a straightforward and powerful method for refinement, this approach has, in practice, been even less successful than other homology model refinement methods. The extent to which this poor performance is attributable to limited sampling, rather than the accuracy of the energy functions, has been unclear.

Here, we address this issue in the context of MD-based refinement, by applying long, all-atom MD simulations—two to three orders of magnitude longer than previous attempts at all-atom MD-based refinement—to structure refinement of CASPR targets. We used a physics-based force field that was recently shown to successfully fold a structurally diverse set of fast-folding proteins[31] and whose ability to reproduce a number of experimental observables was as good as or better than any other force field considered in a recent assessment of force field accuracy.[32] This force field, called CHARMM22*,[33,34] was not optimized based on either previous CASPR targets or on decoys. To assess the performance of MD in an unbiased manner, we performed simulations of nearly all of the refinement targets included in the previous two CASP experiments (CASP8 and CASP9). We performed unrestrained simulations as well as simulations in which secondary structural elements were weakly restrained to the initial homology model structure. We did not use the information provided by the CASP organizers on regions within a homology model structure that deviated most from the native state in any aspect of the design of our simulations. As a test of force field accuracy, we also assessed the stability of each protein in a simulation beginning from its known native state. In total, these all-atom simulations entailed nearly 6 ms of simulated time. Collectively, our simulations reflect the state of the art in MD-based refinement of homology models, allowing us to determine the relative importance of sampling and force field quality and paving the way for future methodological improvements.

MATERIALS AND METHODS

Homology model and native structures were downloaded from the CASP website (http://predictioncenter.org) and used as starting structures for our simulations. We carried out unrestrained, explicit-solvent MD simulations of 25 CASPR targets culled from the previous two CASP experiments (TR389 to TR488 from CASP8, the remainder from CASP9), representing all CASPR targets from these experiments except TR461 and TR476. We did not simulate TR461 because the corresponding native state structure (PDB entry 3DH1) contains a zinc ion in its core; this information was not provided by CASP and thus could not be incorporated in the refinement protocol. We did not simulate TR476 because its side chains were missing in the homology model. The simulated targets include two domains of TR429 simulated separately (these are labeled TR429A, containing residues 1–100, and TR429B, containing residues 101–176).

For each refinement target, we carried out an all-atom simulation at least 100 μs in length starting from the initial homology model; these simulations were conducted at 300 K under neutral pH conditions. In addition, we applied weak harmonic restraints with a spring constant of 0.07 kcal/mol/Å^2 on all C_α atoms in secondary structural elements and carried out a restrained simulation from the homology model for each refinement target that was also at least 100 μs long. The restraint level of 0.07 kcal/mol/Å^2 was found by optimizing GDT-TS values realized in the simulation of a single target (TR389); this target was chosen because its native state is highly unstable in simulation, making it one of the most difficult targets to refine. Finally, we carried out unrestrained simulations starting from the native state for each refinement target, except target TR435, which was excluded because its PDB structure has 11 missing residues (residues 62–72). All native-state simulations were at least 40 μs long.

All simulations were performed on Anton, a special-purpose machine for MD simulation.[35,36] We used the CHARMM22* force field,[33,34] which has proven to be among the most accurate current force fields[32] and to be adequate for folding a number of fast-folding proteins of different structural classes.[31] All production simulations were carried out without distant electrostatics, using a shifted Coulomb potential with cutoffs ranging from 9.0 to 11.3 Å. In an earlier study,[31] it was found that this

approach is accurate enough to enable *ab initio* folding in simulation of several fast-folding proteins.[32] Furthermore, we performed several control simulations with an Ewald-based treatment of distant electrostatics, and did not find large differences in GDT-TS between these control simulations and those performed with the shifted Coulomb potential (Supporting Information Fig. S1). From a practical point of view, the use of the shifted Coulomb potential with a cut-off is motivated by the fact that it leads to a speed-up of a factor of 1.5–2 in our simulations, allowing simulation rates of 7–20 μs/day on a 128-node Anton machine or 20–30 μs/day on a 512-node Anton machine. Finally, we also performed control simulations of CASP8 homology models and native states using variants of the ff99SB force field,[37] namely ff99SB-ILDN[38] and ff99SB*-ILDN[38,39] with TIP4P-Ew water.[40] These were chosen as they also performed well in recent tests of force field quality.[32,41] Although there are occasional differences between the ff99SB and CHARMM22* simulations in individual targets, overall the ff99SB simulations did not result in significant differences in native state stabilities and quality of homology model refinement.

GDT-TS calculations were carried out using the Local-Global Alignment (LGA) program.[15] Molecular graphics were produced using Visual Molecular Dynamics (VMD).[42] Further simulation and analysis details are provided in the Supporting Information.

RESULTS AND DISCUSSION

The 25 CASPR targets studied range from 63 to 192 amino acids in length, with a median length of 121 amino acids (Supporting Information Table S1). The initial GDT-TS values for these targets range from 53.4 to 91.5, with a mean GDT-TS of 74.2 (Supporting Information Table S1). The initial RMSD of their C_α atoms from their respective native structures ranges from 1.3 Å to 7.5 Å, with a mean initial RMSD of 3.4 Å (Supporting Information Table S1). The targets are diverse with respect to their secondary structural content, and include all-alpha proteins, all-beta proteins, as well as proteins with long loop regions.

In all simulations starting from the homology model, we found structural fluctuations with characteristic timescales up to several microseconds in length, as indicated by both GDT-TS as well as RMSD from the native state (autocorrelation functions of these two metrics are shown in Supporting Information Fig. S2 and Fig. 1). In conjunction with evidence of structural improvement over several microseconds in restrained simulations (discussed below), these findings suggest that simulations on timescales shorter than a microsecond may not generate sufficient sampling of the neighborhood of the initial state.

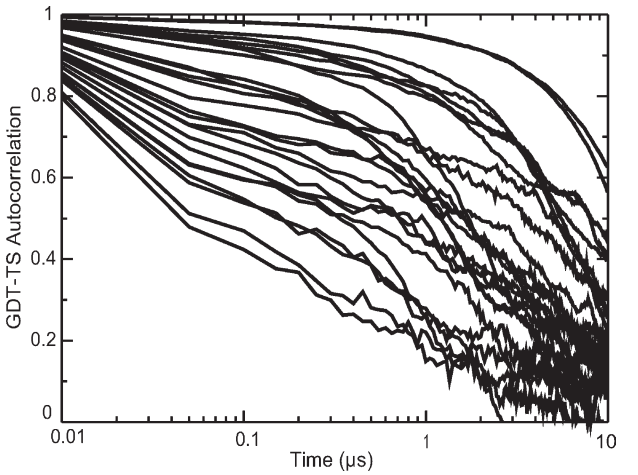

Figure 1

Fluctuations at the microsecond scale in unrestrained simulations of all 25 refinement targets, as indicated by the autocorrelation function (ACF) of the GDT-TS time series. Averaging was performed over the entire trajectory to compute the ACF at each time lag. For every target, correlations persist over time intervals of at least 1 μs.

Instability of homology model structures and native states is consistent with force field errors

In the majority of unrestrained simulations, the simulated structure drifts away from the native structure on timescales ranging from less than a microsecond to several microseconds (Fig. 2, light purple curves, and Supporting Information Fig. S3). This drift can correspond to complete unfolding over tens of microseconds, as in TR569, and can be interrupted by forays towards the native state, as in TR389, TR464, and TR557. To examine the possibility that the decrease in GDT-TS in unrestrained simulations could be an artifact of limited sampling, we ran simulations starting from the native state (Fig. 2, orange curves). In the majority of targets, the structure is not stable in the native state and drifts towards GDT-TS values comparable to those realized by the homology model simulations. Furthermore, for six targets (TR429A, TR453, TR462, TR469, TR530, and TR624), there is direct evidence that the homology model and native state simulations eventually sample overlapping regions of conformational space (see Supporting Information text and Supporting Information Fig. S4). Although these simulations have different starting conformations, these findings suggest that for these simulation timescales, sampling is not a major factor preventing MD-based refinement simulations from reaching the native state. Instead, these findings are consistent with the interpretation that the global free-energy minimum of the force field used here is different from the X-ray or NMR structure for most CASP targets, even though the force field has been shown to be able to

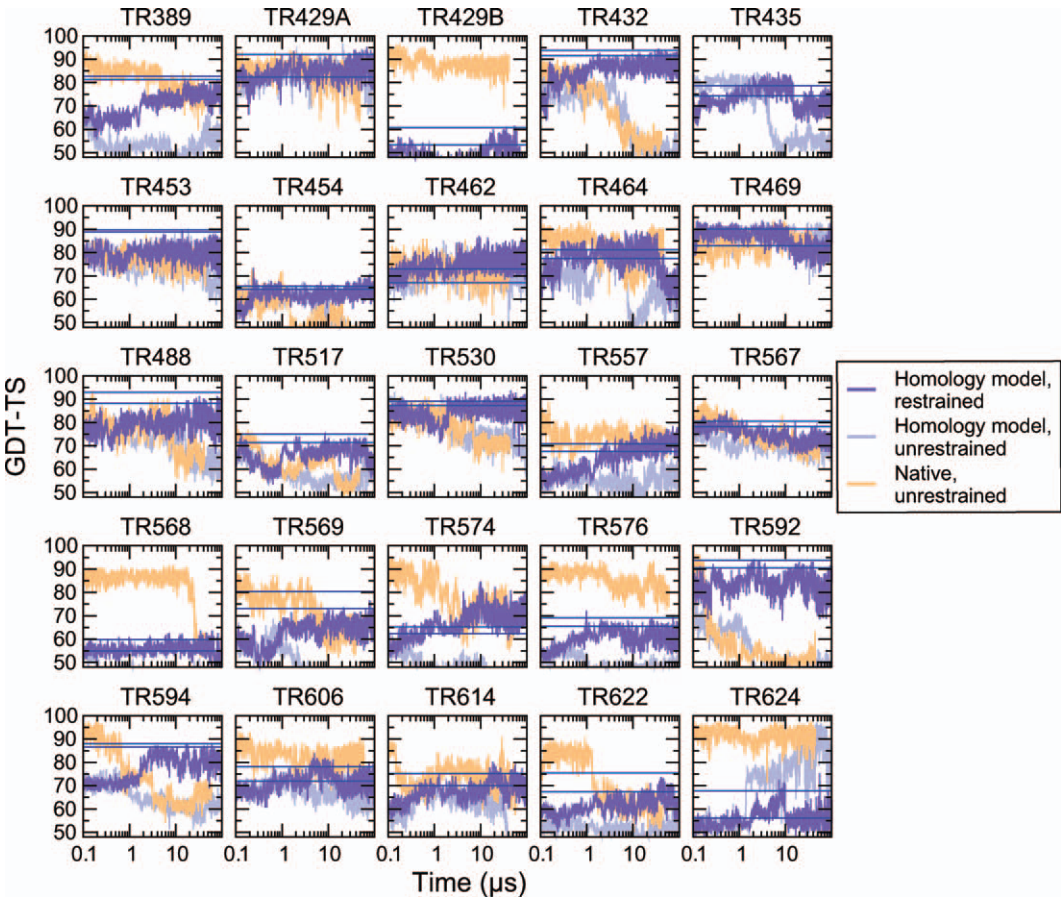

Figure 2

GDT-TS as a function of time for unrestrained (light purple) and restrained (purple) simulations starting from the homology model, and for unrestrained (orange) simulations starting from the native state. The horizontal blue lines in each case refer to GDT-TS values of the homology model (lower horizontal line) and of the best refined structure (upper horizontal line) reported by CASPR participants.

obtain the correct minimum for several fast-folding proteins.[31]

Restrained simulations show more consistent improvement in model quality over tens of microseconds

Although it would be desirable to address the force field issue directly, we explored the extent to which an imperfect energy function could be alleviated by introducing position restraints that prevent the structure from drifting away considerably from the starting homology model. This strategy has been used frequently in previous MD-based refinement attempts; it relies on the assumption that the starting model is reasonably accurate, as turns out to be the case for the majority of CASPR targets. We ran 100-μs simulations with weak restraints to prevent large drift from the starting homology model and we observed, on the whole, improvement over the homology model (Fig. 2, purple curves). For all targets, we note that several frames improve upon the original homology model. (We defer the practical question of how to select frames to the following section).

We measured the improvement of simulated structures over the initial homology model in several ways. First, we considered sliding time windows of length 1 μs and computed the maximum and mean GDT-TS of 20 frames within the window as two measures of structural improvement as a function of time (see Supporting Information for details). We found that the start times of windows containing the highest GDT-TS frames varied from a few microseconds to tens of microseconds (Supporting Information Fig. S5). The average structural improvement (across all targets), as assessed by either of the two measures, is thus uniformly high between ~1 μs and ~50 μs, with a peak at 10 μs (Fig. 3, black and gray curves).

To assess the best structure in a simulation trajectory as a function of the duration of the trajectory, we also measured the highest GDT-TS encountered up to a certain time t, for approximately logarithmically spaced values of t: 100 ns, 500 ns, 1 μs, 5 μs, 10 μs, 50 μs, and

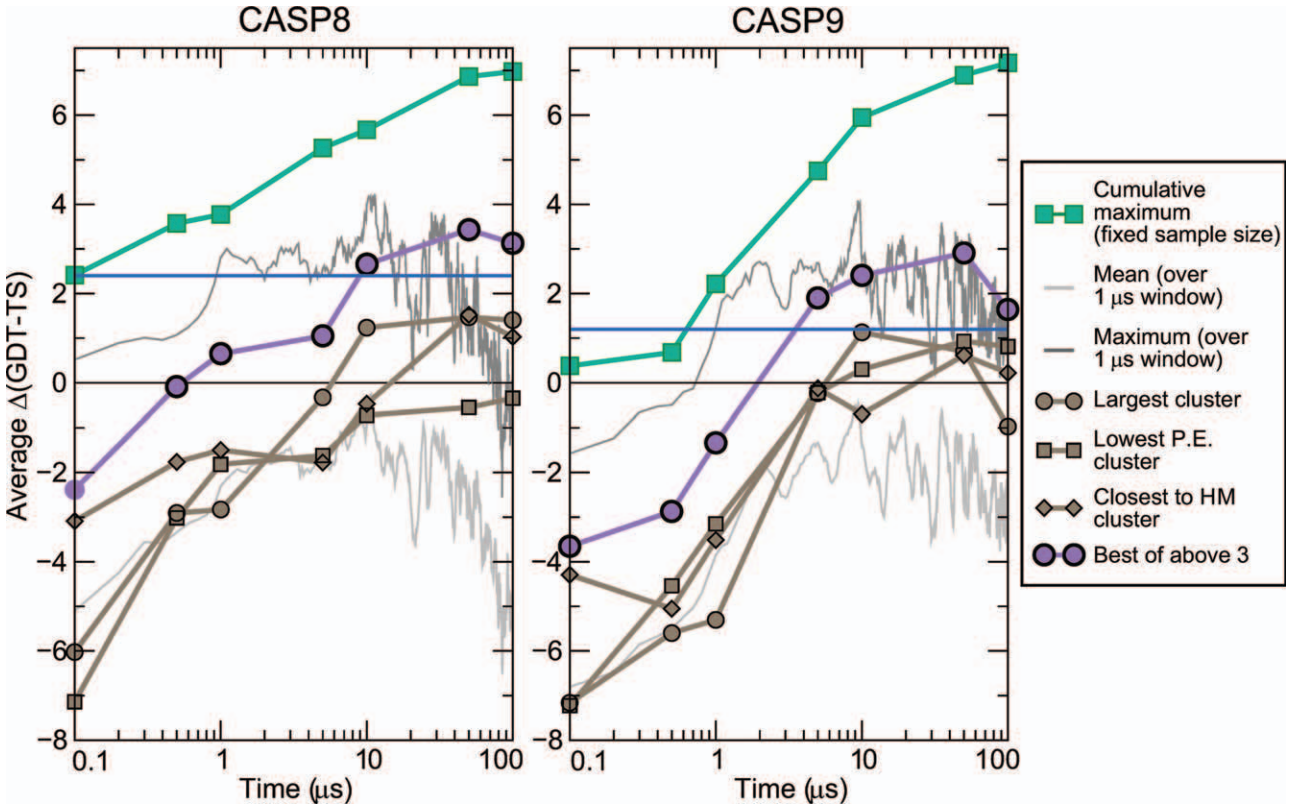

Figure 3

Average structural improvement and average prediction performance, assessed separately for restrained simulations of CASP8 (left panel) and CASP9 (right panel) targets. All curves represent plots of the average difference between a calculated GDT-TS value and the GDT-TS value of the initial homology model, averaged separately over the 11 CASP8 targets (left panel) and the 14 CASP9 targets (right panel). This average difference is labeled as "Average Δ(GDT-TS)" on the *y*-axis. The continuous curves represent maximum (dark gray) and mean (light gray) GDT-TS in a moving, 1-μs time window. The abscissa of each point on these curves represents the start time of the window. The green curve represents the highest average Δ(GDT-TS) obtained up to the time represented by the abscissa, with a fixed sample size of 500 frames used for each time point. The remaining curves represent performance of three structure prediction protocols. In each case, the predicted structure consisted of a representative frame from the largest cluster (brown, filled circles), the cluster with lowest mean potential energy (brown, filled squares), and the cluster in which the representative frame was closest to the initial homology model (brown, filled diamonds). The violet curves represent average improvements of the best of these three prediction protocols. In each panel, the horizontal blue line represents the performance of the group with the highest average GDT-TS improvement among CASPR submissions where, for each target, the highest GDT-TS model was picked from the five submitted models.

100 μs (Supporting Information Fig. S6). On average, the improvement over the initial homology model, Δ(GDT-TS), of the best encountered structure in restrained simulation increases with the length of the simulation trajectory. This monotonic growth in Δ(GDT-TS) of the best encountered structure is, however, expected to occur simply because of the accumulation of samples as the simulation progresses in time. Thus, in order to carry out an unbiased assessment of long simulations, we calculated the maximum Δ(GDT-TS) encountered up to a certain time, while ensuring that the maximum was always computed over the same number (500) of randomly selected samples (simulated frames). (See Supporting Information Fig. S6 and Supporting Information text for details.). In this modified computation, average Δ(GDT-TS) values increase roughly linearly with log(*t*) over three orders of magnitude (Fig. 3, green curves), which demonstrates the benefit of running long-timescale, restrained MD simulations for structural refinement.

Assessment of blind structure predictions derived from restrained simulation trajectories

We applied three simple protocols to extract predicted structures from simulation trajectories. In each case, we clustered the restrained simulation trajectories using *k*-means clustering, with different values of *k*. The predicted structure under each protocol was identified as the representative frame (i.e., the frame with lowest average C_α RMSD from all other frames in the cluster) from a consensus cluster, with the consensus taken across the

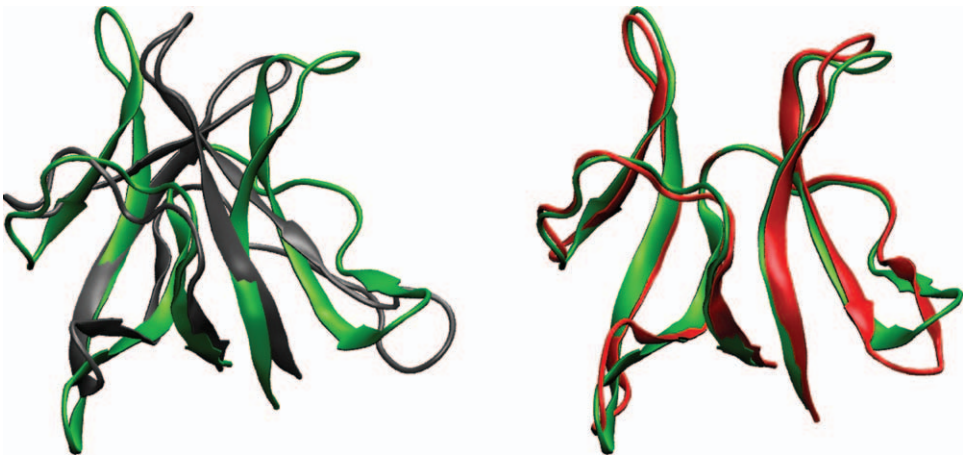

Figure 4

Refinement to a native-like conformation in an unrestrained simulation of TR624. The initial homology model (gray) has a C_α RMSD from the native structure (green) of over 5 Å. The simulation trajectory was clustered using *k*-means clustering with $k = 5, 10, \ldots, 95$. For each value of *k*, the representative frame from the most populated cluster was identified. The most frequently occurring representative frame (red), reached at a simulation time of 19.1 μs with a GDT-TS value of 94.2, is shown superimposed on the native structure. The simulated structure shows excellent agreement of the backbone with that of the native structure, except in a short loop region. Refinement to the native-like conformation in this case required the reorientation of four β strands.

different values of *k*. The consensus cluster was identified in the following ways (see Supporting Information for details): (a) as the largest cluster, (b) as the lowest average potential energy cluster, and (c) as the cluster with the lowest average C_α RMSD from the initial homology model. The rationale behind method (c) is that, because we are addressing a refinement problem, it is likely that the native structure is close to the initial homology model.

We carried out a series of structure predictions based on the three protocols identified above. In each case, we clustered restrained simulation trajectory data up to a time *t* to evaluate prediction performance as a function of simulated time (Supporting Information Fig. S6). The three protocols performed comparably on average (Fig. 3, brown curves). For each target, we also computed the best of three GDT-TS (i.e., the maximum GDT-TS among the three predictions, for each value of *t*). The average best of three Δ(GDT-TS) over all targets increased up to about 50 μs, followed by a slight dip at 100 μs (Fig. 3, violet curves). This dip is consistent with the observation that structural improvement in time windows beyond 50 μs is lower than those at prior times (Fig. 3, gray curves), because force field limitations appear to become relevant at timescales on the order of tens of microseconds (see the discussion on instability of native states below). Although the prediction at 100 μs is based on clustering of all frames up to 100 μs, the large number of low GDT-TS frames in the 50–100 μs time range appears to confound the detection of high GDT-TS frames in a sea of low GDT-TS "noise." Average prediction performance is highest at around 50 μs, when the best of three Δ(GDT-TS) values is 3.4 for CASP8 targets and 2.9 for CASP9 targets. While these results represent only modest overall improvements, they are nevertheless comparable to the values reported by the best CASPR submissions.[12,13] Similar trends are observed for two other GDT-based scores, GDT-HA and GDC-SC, as well as the average C_α RMSD from the native state (Supporting Information Fig. S7). While greater improvement in GDT-TS (as well as GDT-HA and RMSD) is possible by averaging structures over a suitably chosen time window, we refrain from doing so since that would result in unphysical structures with potentially bad contacts. Each of our predicted structures corresponds to a single snapshot of the simulation and thus represents a physically plausible structure.

Dramatic improvement in model quality over tens of microseconds for TR624

As discussed earlier, while most unrestrained simulations lead to structures drifting away from the native state because of force field errors, there is some evidence that regions of the protein that are markedly different between homology model and native state structures (and thus in need of substantial refinement) are correctly identified in unrestrained simulations (Supporting Information Fig. S8). Also, there are a few targets where the structure drifts towards the native state in unrestrained simulations over several microseconds (TR429A, TR462, TR469, and TR624). In the specific case of TR624, dramatic improvement in structural quality takes place over roughly 20 μs, followed by a long (~50 μs) stable phase,

followed by fluctuations away from the native state. Over the course of the first 20 μs, the GDT-TS value increases from about 56 to over 90, and the C_α RMSD to the native structure decreases from 5.2 Å to about 1 Å (Fig. 2 and Supporting Information Fig. S9), representing refinement from a very low-resolution structure all the way to a native-like structure. The refinement in this case is characterized by two sharp transitions (Supporting Information Fig. S9). The first transition at ~1 μs is characterized by the first and fourth beta hairpins moving apart, whereas the second transition at ~20 μs is characterized by the stabilization of the long loop (residues 26–35) between the second and third beta hairpin. This simulation highlights the potential of unrestrained, all-atom MD to refine structures to their native states. A simple clustering protocol (see Supporting Information for details) applied to the entire trajectory robustly identifies the native-like state among the most populated clusters (Supporting Information Table S2 and Fig. 4), without the need for sophisticated scoring schemes.

To assess the robustness of the TR624 simulation results, we ran three additional simulations of TR624 with different initial velocities; within a few microseconds, each simulation transitioned to a conformation whose C_α RMSD to the native structure was less than 2 Å. We also carried out additional simulations starting from different homology models of progressively worse quality. We find that the most accurate of these three homology models, with initial GDT-TS of 52.9, reaches a native-like conformation in less than a microsecond. On the other hand, simulations of the homology models further from the native state (GDT-TS of 25.0 and 46.0) do not adopt native-like conformations even after 100 μs of simulation (see Supporting Information and Supporting Information Fig. S10). We conclude that, even in cases where the force field correctly describes the native state, refinement within a timescale of 100 μs may not be achievable for low-quality models that deviate substantially from the native state.

CONCLUSIONS

To our knowledge, the simulations reported here represent the longest refinement simulations to date by at least two orders of magnitude. In most cases, our unrestrained simulations appear to be sufficiently sampled: for all but six targets (TR389, TR429B, TR557, TR569, TR574, and TR576), GDT-TS values of unrestrained simulations from the homology model and native state eventually overlap (Fig. 2). Of the 18 targets for which convergence of GDT-TS occurs within 100 μs, simulations of 12 (TR432, TR454, TR462, TR488, TR517, TR530, TR567, TR568, TR592, TR594, TR614, and TR622) drift towards GDT-TS values below 80. For six targets (including four targets that drift towards high GDT-TS values in both sets of unrestrained simulations), there is evidence that simulations initiated in the homology model and in the native state sample overlapping regions of conformational space. These findings suggest that, in most cases, limited sampling is not the critical obstacle in our refinement attempts. The fact that both native state and homology model simulations realize low GDT-TS values is consistent with the interpretation that the structure that realizes the global free-energy minimum for the force field employed is not the X-ray or NMR structure, even though there is evidence that the force field employed is among the best available[32] and is sufficiently accurate to simulate the *ab initio* folding of a number of small, fast-folding proteins.[31] We suspect that the underlying reason for this disparity is that the free energy surface of fast-folding proteins is mostly characterized by a single, well-defined minimum corresponding to the folded state with few, if any, competing minima having similar free energy. The energy landscape of a fast-folding protein might thus be expected to be robust in the face of small inaccuracies in the force field representation. In contrast, refinement targets typically display long and flexible regions that are likely to have multiple minima of comparable free energy, and residual errors in the force field may drive the system toward substantially different conformations. A large number of control simulations were also performed using variants of the ff99SB force field, which were also found to be among the best performing in recent studies of force field accuracy.[32,41] In these simulations, we did not observe any significant improvement in the quality of the refinement. This indicates that the results presented here are not just the result of an unfortunate choice of force field, but are likely representative of the state of the art in MD-based homology model refinement.

For most of the weakly restrained 100-μs simulations, we observe a small improvement with respect to the starting model. We find that the frames with the highest GDT-TS are reached within the first 50 μs, suggesting a refinement protocol in which weakly restrained simulations are run for tens of microseconds. We used relatively simple methods based on clustering to predict native protein structure from these trajectories. This approach yielded a level of refinement comparable to that of the best submissions to CASPR. Taken together, the data from our restrained and unrestrained simulations suggest that the use of restraints can mitigate to some extent the problem of having an imperfect energy function, but not eliminate it.

When pooling together all the results obtained for the various CASP8 and CASP9 targets, we did not find a significant correlation between the level of refinement and simple properties such as length of the protein or secondary structure content, suggesting that trivial corrections to the force field may not be sufficient to improve refinement results (see Supporting Information, Supporting

Information Fig. S11, and Supporting Information Table S3). We did find a weak, negative correlation between the level of refinement achieved and the initial GDT-TS value, suggesting that it is harder to improve models that are already good. We also found a weak, negative correlation with the proportion of surface area that is involved in crystal contacts in the native state (Supporting Information Fig. S11 and Supporting Information Table S3), suggesting that at least some of the difficulties in achieving consistent refinement may be ascribed to differences between the protein structure in solution and the crystallographically determined structure.

What might we expect if we were to run long, unrestrained simulations with improved force fields that correctly describe the native state? An example is provided by TR624, for which we observe dramatic (and for CASPR targets, unprecedented) refinement occurring over tens of microseconds of simulation. With improved force fields, it appears plausible that we may find more examples of dramatic refinement on these timescales. Perhaps unsurprisingly, the identification of native-like frames in simulation trajectories only appears to be a difficult problem when the force fields are not accurate enough to enable frequent sampling of the native state on the timescales studied here. In our view, it is probably more beneficial in the long run to focus on the development of better force fields than on the development of sophisticated methodologies for scoring structures realized in simulation.

ACKNOWLEDGMENTS

The authors thank Rebecca Kastleman and Mollie Kirk for editorial assistance.

REFERENCES

1. Laskowski RA, Watson JD, Thornton JM. ProFunc: a server for predicting protein function from 3D structure. Nucleic Acids Res 2005;33:W89–W93.
2. Zhang Y. Protein structure prediction: when is it useful? Curr Opin Struct Biol 2009;19:145–155.
3. Roy A, Kucukural A, Zhang Y. I-TASSER: a unified platform for automated protein structure and function prediction. Nat Protoc 2010;5:725–738.
4. Becker OM, Dhanoa DS, Marantz Y, Chen D, Shacham S, Cheruku S, Heifetz A, Mohanty P, Fichman M, Sharadendu A, Nudelman R, Kauffman M, Noiman S. An integrated *in silico* 3D model-driven discovery of a novel, potent, and selective amidosulfonamide 5-HT1A agonist (PRX-00023) for the treatment of anxiety and depression. J Med Chem 2006;49:3116–3135.
5. Ekins S, Mestres J, Testa B. *In silico* pharmacology for drug discovery: applications to targets and beyond. Br J Pharmacol 2007;152:21–37.
6. Brylinski M, Skolnick J. A threading-based method (FINDSITE) for ligand-binding site prediction and functional annotation. Proc Natl Acad Sci USA 2008;105:129–134.
7. Haupt VJ, Schroeder M. Old friends in new guise: repositioning of known drugs with structural bioinformatics. Brief Bioinform 2011;12:312–326.
8. Moult J. A decade of CASP: progress, bottlenecks and prognosis in protein structure prediction. Curr Opin Struct Biol 2005;15:285–289.
9. Misura KM, Baker D. Progress and challenges in high-resolution refinement of protein structure models. Proteins 2005;59:15–29.
10. Giorgetti A, Raimondo D, Miele AE, Tramontano A. Evaluating the usefulness of protein structure models for molecular replacement. Bioinformatics 2005;21(Suppl 2):ii72–ii76.
11. Qian B, Raman S, Das R, Bradley P, Mc Coy AJ, Read RJ, Baker B. High-resolution structure prediction and the crystallographic phase problem. Nature 2007;450:259–264.
12. MacCallum JL, Hua L, Schnieders MJ, Pande VS, Jacobson MP, Dill KA. Assessment of the protein-structure refinement category in CASP8. Proteins 2009;77(Suppl 9):66–80.
13. MacCallum JL, Perez A, Schneiders MJ, Hua L, Jacobson MP, Dill KA. Assessment of protein structure refinement in CASP9. Proteins 2011;79(Suppl S10):74–90.
14. Kryshtafovych A, Venclovas C, Fidelis K, Moult J. Progress over the first decade of CASP experiments. Proteins Suppl 2005;7:225–236.
15. Zemla A, Venclovas C, Reinhardt A, Fidelis K, Hubbard TJ. Numerical criteria for the evaluation of *ab initio* predictions of protein structure. Proteins 1997;Suppl 1:140–150.
16. 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.
17. Lin MS, Head-Gordon T. Reliable protein structure refinement using a physical energy function. J Comput Chem 2011;32:709–717.
18. Summa CM, Levitt M. Near-native structure refinement using in vacuo energy minimization. Proc Natl Acad Sci USA 2007;104:3177–3182.
19. Fan H, Mark AE. Refinement of homology-based protein structures by molecular dynamics simulation techniques. Protein Sci 2004;13:211–220.
20. Chen J, Brooks CL, III. Can molecular dynamics simulations provide high-resolution refinement of protein structure? Proteins 2007;67:922–930.
21. Ishitani R, Terada T, Shimizu K. Refinement of comparative models of protein structure by using multicanonical molecular dynamics simulations. Mol Simul 2008;34:327–336.
22. Kannan S, Zacharias M. Application of biasing-potential replica-exchange simulations for loop modeling and refinement of proteins in explicit solvent. Proteins 2010;78:2809–2819.
23. Lu H, Skolnick J. Application of statistical potentials to protein structure refinement from low resolution *ab initio* models. Biopolymers 2003;70:575–584.
24. Chopra G, Kalisman N, Levitt M. Consistent refinement of submitted models at CASP using a knowledge-based potential. Proteins 2010;78:2668–2678.
25. Zhu J, Fan H, Periole X, Honig B, Mark AE. Refining homology models by combining replica-exchange molecular dynamics and statistical potentials. Proteins 2008;72:1171–1188.
26. Han R, Leo-Macias A, Zerbino D, Bastolla U, Contreras-Moreira B, Ortiz AR. An efficient conformational sampling method for homology modeling. Proteins 2007;71:175–188.
27. Kim DE, Blum B, Bradley P, Baker D. Sampling bottlenecks in de novo protein structure prediction. J Mol Biol 2009;393:249–260.
28. Lee MR, Tsai J, Baker D, Kollman PA. Molecular dynamics in the endgame of protein structure prediction. J Mol Biol 2001;313:417–430.
29. Chopra G, Summa CM, Levitt M. Solvent dramatically affects protein structure refinement. Proc Natl Acad Sci USA 2008;105:20239–20244.
30. Chen J, Brooks CL, III, Khandogin J. Recent advances in implicit solvent-based methods for biomolecular simulations. Curr Opin Struct Biol 2008;18:140–148.
31. Lindorff-Larsen K, Piana S, Dror RO, Shaw DE. How fast-folding proteins fold. Science 2011;334:517–520.
32. Lindorff-Larsen K, Maragakis P, Piana S, Eastwood MP, Dror RO, Shaw DE. Systematic validation of protein force fields against experimental data. PLoS ONE 2012;7:e32131.
33. MacKerrel AD, Jr., Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FTK, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R,

Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M. All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 1998;102:3586–3616.
34. Piana S, Lindorff-Larsen K, Shaw DE. How robust are protein folding simulations with respect to force field parameterization? Biophys J 2011;100:L47–L49.
35. Shaw DE, Dror RO, Salmon JK, Grossman JP, Mackenzie KM, Bank JA, Young C, Deneroff MM, Batson B, Bowers KJ, Chow E, Eastwood MP, Ierardi DJ, Klepeis JL, Kuskin JS, Larson RH, Lindorff-Larsen K, Maragakis P, Moraes MA, Piana S, Shan Y, Towles B. Millisecond-scale molecular dynamics simulations on Anton. Proceedings of the Conference on High Performance Computing, Networking, Storage and Analysis (SC09). New York: ACM; 2009.
36. Shaw DE, Maragakis P, Lindorff-Larsen K, Piana S, Dror RO, Eastwood MP, Bank JA, Jumper JM, Salmon JK, Shan Y, Wriggers W. Atomic-level characterization of the structural dynamics of proteins. Science 2010;330:341–346.
37. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins 2006;65:712–725.
38. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, Shaw DE. Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins 2010;78:1950–1958.
39. Best RB, Hummer G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. J Phys Chem B 2009;113:9004–9015.
40. Horn HW, Swope WC, Pitera JW, Madura JD, Dick TJ, Hura GL, Head-Gordon T. Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew. J Chem Phys 2004;120:9665–9678.
41. Beauchamp KA, Lin Y-S, Das R, Pande VS. Are Protein Force Fields Getting Better? A Systematic Benchmark on 524 Diverse NMR Measurements. J Chem Theory Comput 2012. DOI:10.1021/ct2007814.
42. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14:33–38.
43. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79:926–935.
44. Fennell J, Gezelter JD. Is the Ewald summation still necessary? Pairwise alternatives to the accepted standard for long-range electrostatics. J Chem Phys 2006;124:234104.
45. Shan Y, Klepeis JL, Eastwood MP, Dror RO, Shaw DE. Gaussian split Ewald: a fast Ewald mesh method for molecular simulation. J Chem Phys 2005;122:054101.
46. Tuckerman M, Berne BJ, Martyna GJ. Reversible multiple time scale molecular dynamics. J Chem Phys 1992;97:1990–2001.
47. Bowers KJ. Scalable algorithms for molecular dynamics simulations on commodity clusters. Proceedings of the ACM/IEEE Conference on Supercomputing (SC06). New York: ACM; 2006.
48. Nosé S. A unified formulation of the constant temperature molecular dynamics methods. J Chem Phys 1984;81:511.
49. Hoover WG. Canonical dynamics: equilibrium phase-space distributions. Phys Rev A 1985;31:1695–1697.
50. Martyna GJ, Klein ML, Tuckerman M. Nosé-Hoover chains: the canonical ensemble via continuous dynamics. J Chem Phys 1992;97:2635–2643.
51. Tu T, Rendleman CA, Borhani DW, Dror RO, Gullingsrud J, Jensen MØ, Klepeis JL, Maragakis P, Miller P, Stafford KA, Shaw DE. A scalable parallel framework for analyzing terascale molecular dynamics simulation trajectories. Proceedings of the ACM/IEEE Conference on Supercomputing (SC08). New York: ACM; 2008.
52. Heinig M, Frishman D. STRIDE: a Web server for secondary structure assignment from known atomic coordinates of proteins. Nucl Acids Res 2004;32:W500–502.
53. Ihaka R, Gentleman R. R: a language for data analysis and graphics. J Comput Graph Stat 1996;5:299–314.