Structure, Vol. 13, 1311–1319, September, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.str.2005.06.008

Crystallographic Refinement by Knowledge-Based Exploration of Complex Energy Landscapes

Mark A. DePristo,[1,3,*] Paul I.W. de Bakker,[1,4] Russell J.K. Johnson,[2] and Tom L. Blundell[1]
[1]Department of Biochemistry
University of Cambridge
80 Tennis Court Road
Cambridge, CB2 1GA
United Kingdom
[2]Department of Chemistry
University of Cambridge
Lensfield Road
Cambridge, CB2 1EW
United Kingdom

Summary

Although X-ray crystallography remains the most versatile method to determine the three-dimensional atomic structure of proteins and much progress has been made in model building and refinement techniques, it remains a challenge to elucidate accurately the structure of proteins in medium-resolution crystals. This is largely due to the difficulty of exploring an immense conformational space to identify the set of conformers that collectively best fits the experimental diffraction pattern. We show here that combining knowledge-based conformational sampling in RAPPER with molecular dynamics/simulated annealing (MD/SA) vastly improves the quality and power of refinement compared to MD/SA alone. The utility of this approach is highlighted by the automated determination of a lysozyme mutant from a molecular replacement solution that is in congruence with a model prepared independently by crystallographers. Finally, we discuss the implications of this work on structure determination in particular and conformational sampling and energy minimization in general.

Introduction

X-ray crystallography of macromolecules is a powerful yet challenging technology. Once the major experimental hurdles of expression, purification, and crystallization have been overcome, diffraction data can be collected, and the process of constructing a three-dimensional model that explains these data can begin (Blundell and Johnson, 1976; Drenth, 1999). This is an iterative process of model building and refinement. The former involves building or extending a putative, often incomplete, model of the contents of the crystal based upon interpretation of features in an imperfect electron density map. The latter involves adjusting the parameters of this model to improve consistency with the diffraction data (Tronrud, 2004). If everything goes well, the model becomes more complete while explaining better the diffraction data at each stage.

One of the first approaches applied to optimize crystallographic refinement of proteins was the method of least-squares with gradient descent (Watenpaugh et al., 1973) although this was most effective at high resolution when used with restraints on good initial models (Hendrickson, 1985). The introduction of molecular dynamics/simulated annealing extended refinement to lower resolutions and poorer initial models (Brunger et al., 1987). Simultaneously significant advances were made in methods for model building based on interactive computer graphics and sophisticated real-space fitting techniques (Jones and Thirup, 1986; Jones et al., 1991; Kleywegt and Jones, 1996). The integration of real-space model building and reciprocal-space refinement in an automated procedure (ARP/wARP) has further improved tolerance to incomplete and inaccurate models, although high-resolution (better than 2.3 Å) data are still required (Perrakis et al., 1999). Despite these advances, model building and refinement at lower resolutions remain outstanding problems in X-ray crystallography. Advances in optimization strategies occurred concomitantly with improvements in crystallographic target functions, from least-squares to maximum likelihood targets with bulk solvent corrections (Tronrud, 2004).

Macromolecular crystals typically diffract to resolutions poorer than ~2 Å, where the number of parameters of the model approaches or even exceeds the number of experimental observations (Tronrud, 2004). Observations about the geometry and conformation of the protein are usually included as restraints to improve the observation-to-parameter ratio. Electron density maps are often unclear in many regions, making the identification of atomic positions and molecular conformations difficult. Worse, after the estimation of phases from experimental data, these maps are based on phases calculated from incomplete and inaccurate models, and they exhibit strong bias toward the preliminary model. Finally, the energy landscape projected by the model phases and diffraction data is complex, rough, and speckled with many shallow, near isoenergetic minima.

One solution to these problems is to incorporate detailed knowledge of protein structure and conformational preferences directly into the sampling and minimization procedure at the earliest stages of modeling and throughout iterative refinement. Our approach to conformational sampling has been to reduce the problem to a minimum number of degrees of freedom, not only to improve the efficiency of conformational sampling, but also to restrict our exploration to fruitful valleys in the potential energy landscape. We use conformational constraints from the idealized geometry derived from analyses of small molecules (Engh and Huber, 1991), from ϕ/ψ propensities (Lovell et al., 2003),

*Correspondence: mark_depristo@harvard.edu
[3]Present address: Department of Organismal and Evolutionary Biology, Harvard University, 16 Divinity Street, Cambridge, Massachusetts 02138.
[4]Present address: Department of Molecular Biology, Massachusetts General Hospital, 185 Cambridge Street, Boston, Massachusetts 02114.

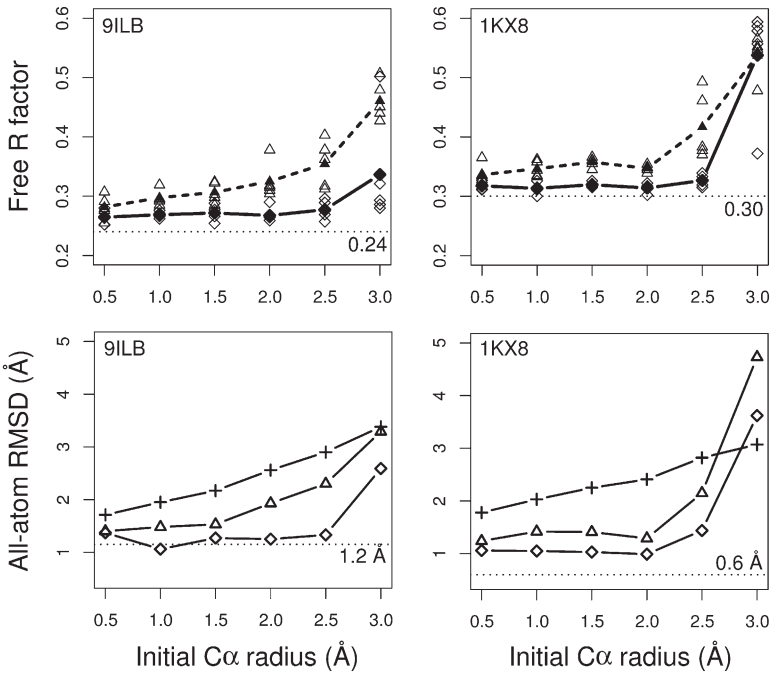

Figure 1. Refinement Quality of 9ILB and 1KX8 as a Function of Cα-Trace Sampling Radius Used to Generate Initial Models

The Cα-trace sampling radius is shown on the x axis. The upper panels show the best R_{free} (open symbols) obtained for each of the five models for RAPPER (diamonds, continuous line) and CNS (triangles, dashed line). The average R_{free} over all five models is shown with filled symbols connected by lines. The lower panels show the average all-atom rmsd to the PDB structure of the Cα-trace (pluses), the RAPPER (diamonds), and the CNS (triangles) models connected by lines. The R_{free} factor (above) and the rmsd of the PDB structures refined by using the CNS protocol (below) are shown as horizontal dashed lines.

and from fine-grained side chain rotamers (Lovell et al., 2000; Shetty et al., 2003) calculated from high-resolution protein structures. We have previously shown that our RAPPER sampling engine can generate accurate models of irregular surface loops (de Bakker et al., 2003; DePristo et al., 2003a) and whole proteins given only coarse structural information (DePristo et al., 2003b), and fit models into electron density maps at a variety of resolutions (DePristo et al., 2004).

Our previous experience with fitting conformers into electron density maps suggested that refinement was often unable to improve the worst fit regions of a model. These regions appeared trapped in deep wells in the X-ray potential energy surface and are separated from fairly close correct conformations by large potential barriers. Such problems are common throughout protein structure modeling due to the lack of effective methods for exploring complex energy landscapes. This holds true regardless of whether the potential landscape is derived from experimental data, from molecular mechanics force fields, or from mining the database of protein structures.

In this paper, we discuss the application of the RAPPER sampling engine to the exploration of complex energy landscapes. Specifically, we demonstrate that using a measure of compatibility with an electron density map allows us to build atomic models into noisy, medium-resolution electron density maps with high specificity. Coupling this building procedure to standard molecular dynamics/simulated annealing (MD/SA) refinement (Brunger et al., 1987, 1998) greatly improves the quality and power of refinement, and, most importantly, leads to final models that fit the data better than MD/SA alone, as measured by the cross-validated R_{free} factor (Brunger, 1992). The power of this approach is illustrated by the automated determination of a novel, to our knowledge, structure of a human lysozyme mutant, by using a molecular replacement model abandoned as intractable by expert human crystallographers (R.J.K.J., unpublished data).

Results

RAPPER Refinement of Interleukin-1β and Protein A6

We applied our automated rebuilding and refinement protocol to rerefine the PDB structures of human interleukin-1β (9ILB) and the chemosensory protein A6 (1KX8). Five random conformations were generated by using Cα-trace modeling, in which models are generated under the constraint that their Cα atoms lie within 0.5, 1.0, 1.5, 2.0, 2.5, or 3.0 Å of their corresponding atoms in the PDB structure (Figure 1; Table S1 [see the Supplemental Data available with this article online]; see Experimental Procedures for details). Models generated under restraints with larger Cα radii are further from the PDB structure, and they are consequently more difficult starting points for refinement due to greater phase error and worse electron density maps. Each model was subjected to refinement with the RAPPER protocol and the CNS protocol in parallel. In short, the CNS protocol consists of 15 cycles of 3 rounds of MD/SA refinement. The RAPPER protocol differs only by replacing regions of the model with poor fit to the electron density map with ones generated with RAPPER that correlate better with the map and are consistent with protein conformational preferences.

First, both RAPPER and CNS refinement converge to reasonable final models for initial Cα restraint radii $\leq$ 2.5 Å for 9ILB and $\leq$ 2 Å for 1KX8 (Figure 1 and Table S1). Neither protocol is powerful enough to refine initial models with Cα restraints > 3 Å (data not shown). For

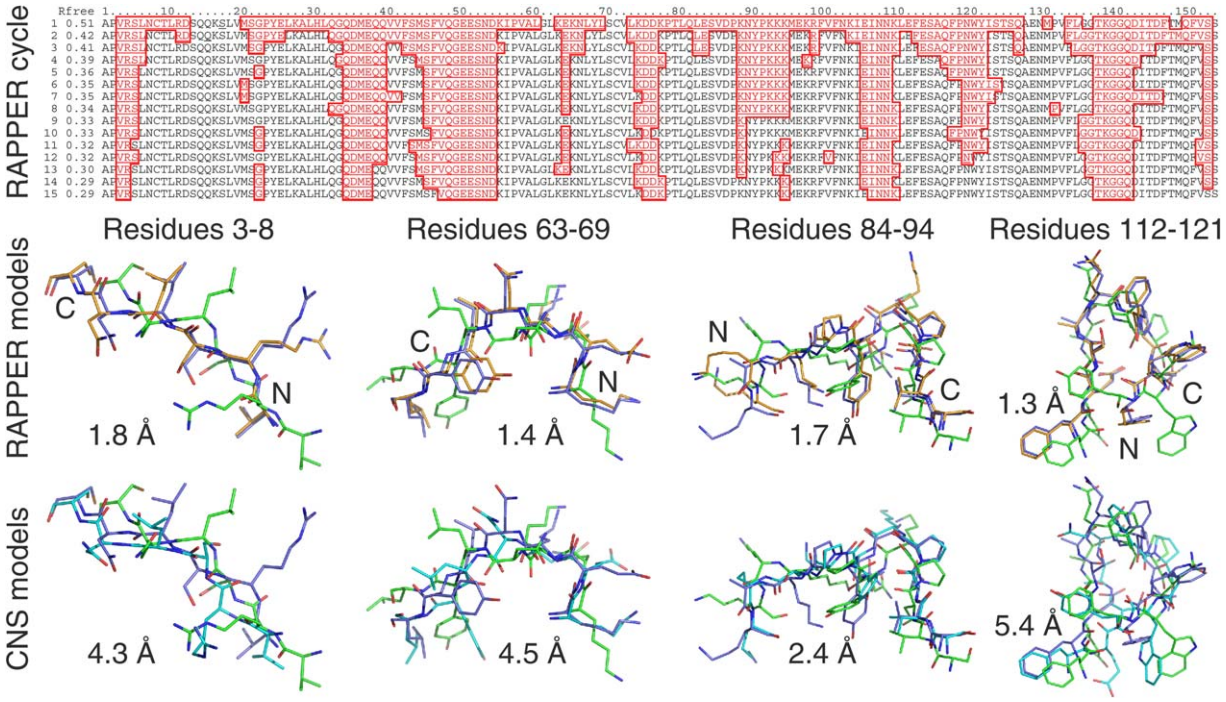

Figure 2. Refinement Trajectory of a 3 Å Cα-Trace Model of 9ILB, with RAPPER Generating a Model with an R_{free} of 0.29 and CNS Generating One of Only 0.45

The upper panel shows the amino acid sequence, the R_{free} factor, and residues remodeled by RAPPER in red for each of the 15 refinement cycles. The lower two panels show stick models of the PDB (blue), initial Cα-trace (green), final RAPPER model (gold), and final CNS model (cyan). Models are annotated by their global all-atom rmsd to the PDB structure.

9ILB, refinement of the PDB structure leads to an R_{free} of 0.27 with the RAPPER protocol and 0.24 with CNS; for 1KX8, the equivalent values are 0.31 and 0.30 (note that the larger R_{free} following RAPPER refinement of the PDB structure reflects the fact that the original R_{free} set was unavailable. CNS remains close to the PDB structure and hence was an artificially low R_{free} factor). Against this baseline, RAPPER completely refines initial models with Cα restraints of up to 2.5 Å radius, while CNS never approaches its baseline but diverges with increasing Cα restraint radius. For 1KX8, again RAPPER refines to near baseline for $\leq$ 2.5 Å restraint radius, while CNS remains 4%–6% above baseline until 2.5 Å restraint radius, beyond which refinement breaks down. As the RAPPER and CNS protocols both produce models with equivalent but slightly tighter B factor distributions than the PDB structure, B factors do not account for the differences in R_{free} factors among RAPPER and CNS models.

In terms of three-dimensional atomic coordinates, both RAPPER and CNS move the initial models significantly closer to the PDB structure, as assessed by the all-atom rmsd between the models and the PDB structure. Given that equivalent models for medium-resolution crystal structures differ significantly from each other (DePristo et al., 2004), it is necessary to relate the model rmsds to a baseline obtained by refining the PDB structure by using the CNS protocol (Figure 1 and Table S1). From these calculations we expect rmsds no lower than 1.2 Å from the PDB for 9ILB, and 0.6 Å for 1KX8. The proximity of the refined models to the PDB coordinates is especially striking when compared to these baselines. For 9ILB, RAPPER remains within 0.3 Å rmsd of baseline, while CNS diverges at 2 Å and is no longer able to move the initial models close to the PDB structure. For 1KX8, RAPPER consistently finds models within 0.4 Å rmsd of the baseline, until a total breakdown at 3 Å, while CNS performs slightly worse up to 2 Å, then diverges rapidly.

Individual Model Refinement

Detailed analysis of the refinement trajectory of individual models is needed to move beyond discussion of overall trends between RAPPER and CNS (Figure 2 and Figure S1). RAPPER reconstruction clearly contributes most during the first rounds of reconstruction, dropping rapidly to near-final levels within three or four cycles. This is consistent with the approximate exponential decay in the number of reconstructed amino acids during the first few cycles, as conformations that are incorrectly fitted are rapidly replaced by better ones. The few stragglers remaining at later stages are residues trapped in poorly fit conformations in regions of poor electron density. Interestingly, poor initial models require many passes of reconstruction in problematic regions. During refinement, incremental improvements in model phases lead to better electron density maps that allow for the identification of misfit regions and their subsequent reconstruction.

The initial models are poor, with side chain conforma-

tions grossly different from the PDB structure, sharing only ~60% of χ_1 < 40°. Electron density maps derived from such models are consequently terrible. Even starting from such maps, RAPPER builds conformations that are in excellent agreement with the PDB structure. CNS, in contrast, is often unable to move the main chain and side chains away from their initial positions and toward correct conformations. The PDB, RAPPER, and CNS models differ most in the positioning of side chains, especially large, bulky ones like arginine, tryptophan, and tyrosine, where the CNS models appear trapped near their initial conformations in the Cα-trace models. In regions in which the electron density is poor, both RAPPER and CNS fail to produce well-fit conformations. In the RAPPER protocol, this results in continuous regeneration of the same regions (see Figure 2, which illustrates that after 15 cycles, several regions are still candidates for remodeling).

RAPPER refinement produces models with good structural quality metrics, such as deviation from ideal bond geometry, observed ϕ/ψ distributions, expected side chain rotamers, and atomic packing (Table S2). These models have only slightly poorer ϕ/ψ, rotamer, and clash scores than the PDB structure. The systematically worse values for the CNS models presumably reflect worse refinement (i.e., higher R factors) and not an intrinsic difference in the quality of models produced by CNS. Support for this view comes from the correlation between R_{free} factors and structural quality metrics, independent of model origin (R^2 = 0.66 for a linear regression of R_{free} and clash score among all models in Table S2). Indeed, this relationship confounds comparison of structural quality for models that differ substantially in their R factors. Nevertheless, the models produced by RAPPER have good validation statistics, comparable to those of the PDB structures, which is encouraging given the complete automation of RAPPER and the low resolution of the diffraction data.

Blind Test Structure Determination of Lysozyme Mutant

To test further the performance of RAPPER and assess its practical value, we undertook a blind structure determination experiment. An initial molecular replacement model and diffraction data were obtained for a human lysozyme mutant by expert crystallographers (B.F. Luisi and R. Johnson) and were provided to us. As a completely independent operation, model building and refinement proceeded in parallel by using standard techniques used by the experts and fully automated refinement with both protocols.

First, and foremost, the initial molecular replacement model provided to us was abandoned as intractable by the crystallographers, due to poor refinement quality and a large difference in R and R_{free} factors. A superior MR solution was obtained and solved against the full 2.5 Å diffraction data. The structure determination process, nevertheless, required significant intervention to correct for conformational changes in surface-exposed loops and side chain rearrangements. This structure refined to an R_{free} of 0.27 by using REFMAC with the Babinet bulk solvent correction (Murshudov et al., 1997). Refining this structure against the truncated 2.8 Å diffraction data results in an R factor of 0.21 and an R_{free} of 0.28. Thus, we expect to do no better than an R_{free} of 0.28 by using the automated refinement protocols.

RAPPER refined automatically the original MR model against the 2.8 Å reflection set to an R_{free} of 0.31 (Figure 3 and Table 1). The final model is close to the PDB structure (rmsd = 0.9 Å). The CNS model, on the other hand, languishes with an R_{free} of 0.37. The contrast between RAPPER and CNS is starker across all Cα-trace models, with RAPPER consistently refining these structures to R_{free} values of 0.30–0.37, while the equivalent range for CNS is 0.37–0.48. In other words, the RAPPER models are between 0.05 and 0.16 lower in R_{free} than those produced by using CNS alone. The model with the lowest R_{free} produced by CNS started from the MR solution and has an R_{free} factor of 0.37. The best RAPPER model, with an R_{free} of 0.30, started from a Cα-trace sample, is better than the refined MR model, and is almost as good as the 1W08 refinement against the original data.

There are significant real-space structural differences among the MR, PDB, Cα-trace, RAPPER, and CNS models (Table 1 and Table S3). First, the MR solution is far from the final PDB structure in main chain and all-atom rmsd, but has a large number of side chains in the correct rotameric states. Both RAPPER and CNS refine initial CNS models far away from their initial conformation. The 5% higher χ_1 conservation of CNS compared to the initial models indicates that CNS is more constrained by the initial side chain conformation than RAPPER. The refined RAPPER models are highly similar to each other, with little difference between the best and worst models. The CNS models, in contrast, are much less similar and are only slightly closer to the PDB structure than the initial models, reflecting the inability of CNS to improve these initial models. The RAPPER and CNS models have equivalent main chain rmsds but different side chain rmsds and χ_1% compared to the PDB structure, indicating that poor side chain positioning is severely limiting CNS refinement.

Interestingly, the best model produced by RAPPER is only slightly closer to the PDB structure than the CNS model, despite having an R_{free} factor that is 7% lower. The RAPPER and CNS models differ most significantly at bulky side chains, where the CNS conformations are often grossly misfit relative to the PDB structure (Figure 3). Other errors in the CNS models are at sites of multiple conformations, where CNS incorrectly places a single conformation between the two PDB conformations, while the RAPPER conformation overlaps completely with one of the PDB conformations. The fine details of the RAPPER conformation appear better, almost completely overlapping with the PDB structure, while the CNS model differs slightly at almost every site. Finally, model quality near the T70N mutation is excellent for both RAPPER and CNS, despite the large conformational change in residues 69–73 to accommodate the substitution.

Discussion

Our previous experiences with knowledge-based protein modeling (Blundell et al., 1987, 1988; Sali and Blun-

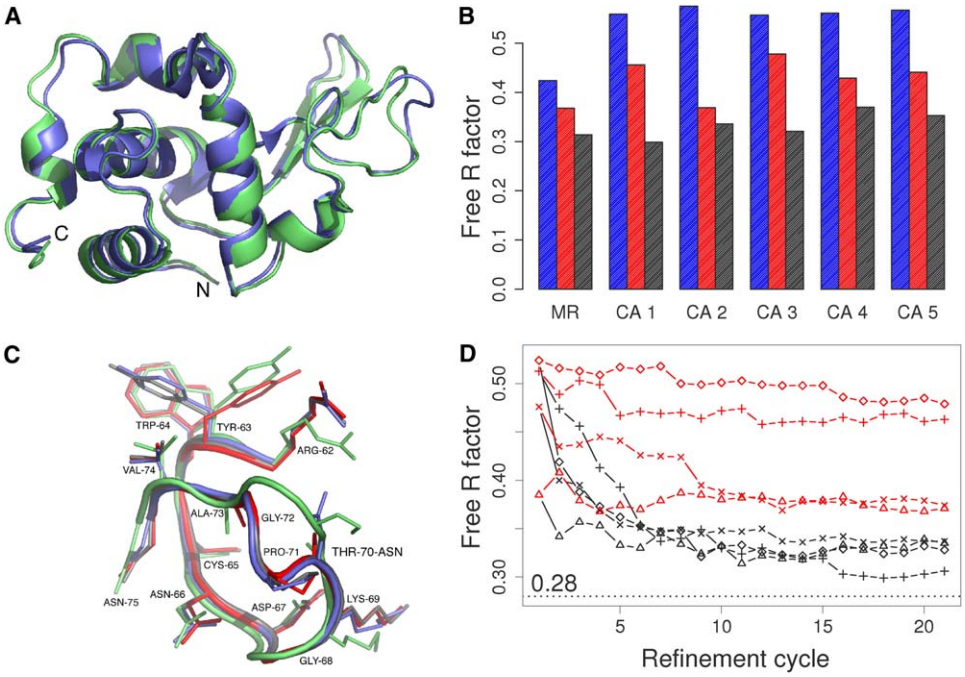

Figure 3. Refinement of the Lysozyme Mutant

(A) Cartoon diagram of MR (green) and PDB structure (blue), superimposed on Cα atoms.

(B) Bar graph of R_{free} factors for initial MR and Cα-trace models (blue), and final CNS (red) and RAPPER (black) model.

(C) Residues 62–75 focused on the mutation T70N in the MR model (green), PDB structure (blue), CNS model (red), and RAPPER model (gray) with side chain conformations as sticks.

(D) Individual refinement trajectories for MR (triangles) and Cα-trace (other symbols) models for RAPPER (black) and CNS (red). The R_{free} factor of the 1W08 structure refined against the 2.8 Å diffraction data is shown as a dashed horizontal line.

dell, 1993) and conformational sampling for loop modeling (de Bakker et al., 2003; DePristo et al., 2003a), comparative modeling (de Bakker, 2003), and X-ray structure determination (DePristo et al., 2004) motivated us to extend RAPPER with a sensitive scoring scheme for modeling building into electron density maps. Our primary interest was to explore how much traditional crystallographic refinement techniques (Tronrud, 2004)—such as least-squares (Hendrickson, 1985), conjugate gradient (Murshudov et al., 1997), or MD/SA (Brunger et al., 1987)—could be improved when supplemented with all-atom, knowledge-based conforma-

Table 1. Structural Comparisons among Lysozyme Mutant Models

	Pairwise RMSD (Å)		
	Main Chain	All Atoms	$\chi_1 < 40°$ (%)
MR model versus PDB 1W08	0.8	1.3	89.4
Initial Cα-Trace Models Based on MR Solution			
Versus PDB 1W08[a]	0.8 (0.04)	1.6 (0.04)	57.9 (3.3)
Versus final RAPPER models[a]	1.1 (0.1)	1.9 (0.3)	59.1 (3.7)
Versus final CNS models[a]	1.1 (0.1)	1.7 (0.2)	64.6 (12.7)
Final RAPPER Models			
Versus final CNS models[a]	0.6 (0.2)	1.4 (0.2)	75.0 (6.9)
Versus final RAPPER models[b]	0.3 (0.3)	0.7 (0.5)	89.2 (7.6)
Versus final PDB models[a]	0.4 (0.2)	1.0 (0.2)	85.7 (3.3)
Final CNS Models			
Versus final CNS models[b]	0.5 (0.3)	1.1 (0.7)	76.9 (15.5)
Versus final PDB models[a]	0.5 (0.1)	1.4 (0.2)	74.2 (8.4)
Best Model Versus PDB 1W08			
RAPPER (model 0)	0.4	0.9	88.5
CNS (model MR)	0.4	1.0	87.5

Comparison of MR, PDB, RAPPER, and CNS models using pairwise main chain and all atom rmsd and the percent of side chains in the similar rotameric state ($\chi_1 < 40°$). The table is divided into four sections, each comparing one set of models against all others. Values in parentheses are standard deviations for comparisons involving multiple structures.

[a] Averaged over five equivalent models (model 1 versus model 1, 2 versus 2, et cetera).

[b] Averaged over ten pairwise models (model 1 versus models 2–5, 2 versus 3–5, et cetera).

tional sampling. We chose to compare only with MD/SA refinement, given its well-deserved reputation for power and large radius of convergence; it is likely that comparisons with gradient descent methods would be even more favorable to RAPPER. It is important to recognize that this work does not constitute a criticism of CNS refinement, as CNS is an essential part of the RAPPER refinement protocol. Nevertheless, it does enable us to identify where the MD/SA approach to exploring complex energy landscapes can be improved and to suggest how this can be achieved.

We suspected that the largest benefits would be at medium (2–3 Å) or low (>3 Å) resolution, where traditional methods have difficulty. In this study, we focused exclusively on the medium-resolution range, given its importance in macromolecular crystallography (Kleywegt, 2002). The use of two previously deposited structures allowed us to evaluate RAPPER under controlled circumstances, while the blind structure determination provided a more realistic challenge.

One important conclusion from the comparison of RAPPER and CNS is that MD/SA refinement is unable to escape from the deep wells in the X-ray potential surface projected by incorrect conformations, especially those involving bulky side chains. The nature of MD/SA refinement sees very large potential energy barriers between such conformations. The heating/cooling cycle of simulated annealing is intended to alleviate this problem, but even the high-temperature phase of simulated annealing clearly has difficulties overcoming these barriers. It is doubtful that baking the model at even higher temperatures will radically improve the quality and power of MD/SA refinement. RAPPER circumvents the barrier-crossing problem by employing a discrete sampling scheme. When coupled with a detailed side chain conformer library, RAPPER can efficiently explore conformational space, a property particularly advantageous for bulky side chains.

Beyond leaping energy barriers, a further advantage of sampling only main chains with good ϕ/ψ characteristics and rotameric side chains is avoiding unproductive and dangerous parts of conformational space. The issue is that conformations with poor ϕ/ψ or nonrotameric side chains often fit well into low-resolution electron density maps and thus lead like the Siren's call away from correct conformations and toward disaster. Restricting sampling to the important degrees of freedom and enforcing prior constraints help overcome the limited specificity, model bias, and phase error of medium-resolution maps.

The average resolution of crystal structures deposited in the PDB is 2.2 Å, shows no downward trend (Kleywegt, 2002), and may indeed increase as focus shifts toward macromolecule complexes (Harrison, 2004). Even now, over half of new crystal structures are solved within the medium-resolution range (2–3 Å), where RAPPER refinement has proven most valuable. The importance of tools like RAPPER for automated structure determination will likely increase given the exponential growth of crystallographic studies and the rise of structural genomics (Gerstein et al., 2003).

In all fairness, RAPPER is no panacea for crystallographic model building and refinement. First, conformational sampling under strict crystallographic constraints is computationally demanding, especially in the early stages when the model is poor. The inflexible use of the real-space correlation function leads to endless reconstruction of regions that never achieve a well-fit conformation. Perhaps the most general critique is that, unlike the excellent program ARP/wARP (Perrakis et al., 2001), the method presented here is limited to refinement of an existing model and cannot be used to automatically construct parts of the structure de novo. On the other hand, RAPPER is free of the atomicity assumption that restricts ARP/wARP to higher-resolution data sets. One final concern is that RAPPER refinement depends critically on the register of the provided model. Detecting such situations is trivially obvious, however, as register errors make it impossible to build through particular regions of the electron density map. Indeed, this behavior forms the basis for current work on de novo model building at resolutions up to 5 Å (N. Furnham, A.S. Dore, D.Y. Chirgadze, P.I.W.d.B., M.A.D., T.L.B., unpublished data).

Although focused on the interesting and practical problem of crystallographic refinement, the approach described here is more general than simply refining

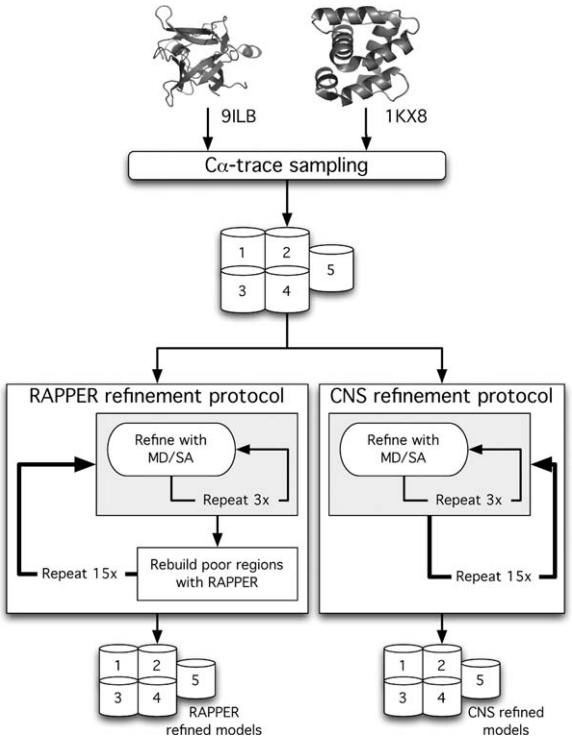

Figure 4. Flowchart of the Current Study

Five initial models were generated by using Cα-trace sampling. Each of these five models was subjected to refinement according to the RAPPER and CNS protocols, and two sets of five refined models were produced. The CNS protocol (right) consists of three iterations of MD/SA refinement (see Experimental Procedures for details), producing a single refinement model. Each such model is fed back into the inner MD/SA cycle 15 times. The RAPPER protocol differs from the CNS protocol only in that RAPPER is used to identify and reconstruct poorly fit regions of the model after each cycle of MD/SA refinement.

crystal structures. Almost all computational problems in structural biology—such as structure prediction, rational protein design, protein-ligand docking—can be recast as a search for low-energy conformations on a complex energy landscape. The improvements over molecular dynamics/simulated annealing presented here suggest that similar improvements can be made by supplementing traditional search and minimization procedures with knowledge-based sampling schemes like RAPPER.

Experimental Procedures

RAPPER Modeling

The discrete conformational sampling program RAPPER was used to generate both the initial conformational samples and X-ray-restrained models (Figure 4). Briefly, RAPPER employs a discrete conformational sampling algorithm to construct an ensemble of all-atom models that each satisfies a set of user-defined restraints (DePristo et al., 2003b). The algorithm proceeds by iteratively extending a polypeptide chain in the N-terminal to C-terminal direction. B factors were reset to 20 Å^2 for main chain and 30 Å^2 for side chain atoms in all models.

Initial Models

Models were generated with RAPPER, by enforcing Cα restraints (with 0.5–3.0 Å radii) and side chain centroid restraints (with 3 Å radii) centered on the atoms in the Protein Data Bank (PDB) structure (DePristo et al., 2003b). To facilitate R and R_{free} comparisons with PDB structures, waters were copied from the PDB into the initial models but were freely refined.

Goodness-of-Fit of a Protein Model into an Electron Density Map

The correlation coefficient between the SA omit map, ρ_o, and the map calculated from the atomic coordinates and scattering factors of the entire putative model, ρ_c:

$$\text{fit}(\rho_o, \rho_c) = \frac{\Sigma\,(\hat{\rho}_o{}^{*}\,\hat{\rho}_c)}{\sqrt{\Sigma\,\hat{\rho}_o^2{}^{*}\Sigma\,\hat{\rho}_c^2}}$$

summed over all indices *x*,*y*,*z* in the map, where $\hat{\rho}_c = \rho_c - \bar{\rho}_c$, ρ_c is the electron density at point *x*,*y*,*z* in the map, $\bar{\rho}_c$ is the average electron density of the map, and σ_c is the number of rms deviations above $\bar{\rho}_c$. The sum is evaluated only at points in the map of significant density, i.e., those where either $\sigma_c \geq 0.5$ or $\sigma_o \geq 0.5$. The fit function varies from 1.0, a perfect correlation between the maps, to zero for unrelated maps; negative correlations are scored as zero. The fit score of a set of atoms, *A*, is simply the fit over all *x*,*y*,*z* within 3 Å of any atom in *A*. This correlation function was inspired by the real-space fit of Jones et al. (1991) and Kleywegt and Jones (1996).

Automated Reconstruction of Protein Models in an Electron Density Map

We automatically identify and reconstruct poorly fit regions of a protein model based on their fit score to a $2F_o–F_c$ simulated annealing omit electron density map (Hodel et al., 1992). First, the fit of each residue of the input model is scored, and those with fit scores < 0.8 are flagged for reconstruction. Next, each contiguous region of poorly fit residues is rebuilt with RAPPER as follows. An ensemble of 100 conformations is extended by a single residue, while all main chain atoms are required to lie in positive electron density and the model Cα atoms are required to lie within 2 Å of their positions in the poorly fit conformation. Next, the side chain conformation with the best goodness-of-fit that satisfies side chain centroid restraints (see RAPPER modeling) from the finely detailed (0.5 Å) side chain conformer library is selected (Shetty et al., 2003). This library includes thousands of individual side chain conformations for each amino acid derived from high-quality residues in high-resolution protein structures (Shetty et al., 2003). The all-atom conformers in the ensemble are then ranked according to their goodness-of-fit, and the lowest 90% are ejected from the ensemble. This is repeated six times, such that the population of conformations is iteratively enriched in well-fit conformations. Finally, the best-fitting conformation is selected when 100 complete polypeptide chains have been generated for the reconstructed region.

Refinement Protocols

MD/SA refinement and map generation were performed by using the CNS program. The RAPPER refinement protocol involves the following steps. An initial model is subjected to 15 cycles of MD/SA refinement (Brunger et al., 1998) and RAPPER reconstruction, followed by a final round of MD/SA refinement. MD/SA refinement is comprised of 3 cycles of 100 steps of geometry minimization, and then torsion-angle simulated annealing starting at 5000 K, slow cooling in 25 K steps, and finally 200 steps of gradient minimization. A maximum likelihood target plus a bulk solvent correction is used for the X-ray term. The CNS protocol is identical to the RAPPER protocol, except that the reconstruction step with RAPPER is omitted.

Protein Structures and Reflection Data

All structures and reflection data were obtained from the Protein Data Bank. The signaling molecule human interleukin-1β (Yu et al., 1999) (PDB code 9ILB) is a 153 amino acid protein solved at 2.28 Å to an R/R_{free} of 0.157/0.210 by using TNT least-squares refinement. The original R_{free} set for 9ILB was not available, so a new set comprising 10% of the reflections was generated; note that it is the same set used in DePristo et al. (2004). The chemosensory protein A6 (Lartigue et al., 2002) (PDB code 1KX8), composed of 112 amino acids, was solved at 2.80 Å to an R/R_{free} of 0.265/0.281 by using CNS with a flat bulk solvent model. The original R_{free} set was obtained from the PDB and comprises 10.4% of the 5096 reflections. Free sets were excluded from all refinement and map calculations.

Structure Solution of the T70N Human Lysozyme Mutant

The expression, purification, and crystallization of a destabilized mutant (T70N) of human lysozyme (residues 19–148) were carried out by R.J.K.J. (Johnson et al., 2005). Diffraction data were collected from a small crystal with dimensions of 100 × 50 × 50 μmi^3 on an R axis IV area detector to a resolution of 2.5 Å. The data were processed and scaled with the program DENZO/SCALEPACK (Otwinowski and Minor, 1997). The R_{sym} was 7.5% (14.6% in the 2.59–2.50 bin), and the completeness was 95.1%. The crystals belong to space group $P2_12_12_1$ with cell dimensions a = 31.117 Å, b = 56.177 Å, c = 62.489 Å. An R_{free} set comprising 5% of the 3649 reflections was set aside. The program AMORE was used to obtain molecular replacement (MR) solutions (Navaza, 1994). Manual model building was performed with Quanta (Accelrys), and refinement was performed with REFMAC (Murshudov et al., 1997).

An MR solution based on the wild-type human lysozyme structure (PDB code 1JSF) was obtained. The 2.5 Å data were truncated to 2.8 Å, and manual model building and refinement starting with the MR structure and reflection set were performed (B.F. Luisi and R. Johnson). This process converged to a generally poor model with an abnormal R/R_{free} ratio and was abandoned as a potentially false MR solution.

At the same time, the 2.8 Å truncated data and initial MR solution were provided to us for a blind assessment of refinement with RAPPER. The Cα atoms and side chain centroids of the MR solution were used to generate five conformational samples as described above (see "Initial Models"). These five models plus the MR solution were refined by following the protocols described above, except that every three cycles of refinement (and RAPPER reconstruction), CNS was used to add waters at sites of large (>3σ), unexplained density in a $F_o–F_c$ difference map using distance and hydrogen bond filters (the default parameters in the standard CNS water_pick.inp script). Between 79 and 93 waters were added to the RAPPER models, and between 49 and 74 were added to the CNS models; the PDB structure has 88 waters.

After concluding that the initial MR solution was intractable, the crystallographers obtained a superior MR solution, and they solved this structure by using the full 2.5 Å diffraction data. This new MR model was based on the same residues 19–148 of the human lyso-

zyme structure, but it excluded the loop that includes the T70N mutation. This model was placed at a different crystallographic origin, and it was oriented slightly differently once superimposed with the initial MR trial. Manual model building and refinement were again performed (B.F. Luisi and R. Johnson), progressing until R and R_{free} factors of 0.185 and 0.268, respectively, were obtained with excellent stereochemistry. In this model, large adjustments were made in the loop at the site of mutation, and counter-ions and waters were included in the model.

The final, or PDB, model is used for all comparisons with the RAPPER refinement. It must be emphasized that the entire manual modeling building and refinement process was conducted simultaneously with and independently from the RAPPER calculations. Note that, unfortunately, the MR solution used for RAPPER and CNS included the wild-type threonine residue and not the asparagine mutation, as the original threonine residue was present in the MR solution used to initiate RAPPER and CNS refinement.

Miscellaneous

All crystallographic calculations were performed with CNS, unless noted otherwise. Structural analyses, including rmsd calculations, were performed with RAPPER. Molecular figures were produced with PyMOL (DeLano, 2002). Similarity in side chain rotamer state between two structures was assessed as the percentage of first side chain dihedral angles (χ_1) within 40° of each other. All models, refinement scripts, and the RAPPER program itself are available for download from the RAPPER website at http://www-cryst.bioc.cam.ac.uk/rapper/. RAPPER currently supports reciprocal-space refinement with either CNS or REFMAC and can easily be adapted to other refinement strategies.

Supplemental Data

Supplemental Data including one figure and three tables are available at http://www.structure.org/cgi/content/full/13/9/1311/DC1/.

Acknowledgments

We acknowledge Ian Tickle for useful suggestions on electron density map scoring schemes and Nick Furnham, Jane Richardson, David Richardson, and Ian Davis for helpful discussions and encouragement. We also thank Ben Luisi and Chris Dobson for crystallographic data and for many stimulating discussions about analysis and interpretation. This work was funded by the Marshall Aid Commemoration Commission, the U.S. National Science Foundation, the Cambridge Overseas Trust (M.A.D.), the Cambridge European Trust, Isaac Newton Trust, and the Biotechnology and Biological Sciences Research Council (P.I.W.d.B.).

Received: March 17, 2005
Revised: June 3, 2005
Accepted: June 8, 2005
Published: September 13, 2005

References

Blundell, T.L., and Johnson, L.N. (1976). Protein Crystallography (New York: Academic Press).

Blundell, T.L., Sibanda, B.L., Sternberg, M.J., and Thornton, J.M. (1987). Knowledge-based prediction of protein structures and the design of novel molecules. Nature *326*, 347–352.

Blundell, T.L., Carney, D., Gardner, S., Hayes, F., Howlin, B., Hubbard, T., Overington, J., Singh, D.A., Sibanda, B.L., and Sutcliffe, M. (1988). 18th Sir Hans Krebs lecture. Knowledge-based protein modelling and design. Eur. J. Biochem. *172*, 513–520.

Brunger, A.T. (1992). The free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472–474.

Brunger, A.T., Kuriyan, J., and Karplus, M. (1987). Crystallographic R factor refinement by molecular dynamics. Science *235*, 458–460.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

de Bakker, P.I.W. (2003). Ab initio Sampling of Polypeptide Conformations and the Prediction of Protein Structure (Cambridge, UK: University of Cambridge).

de Bakker, P.I.W., DePristo, M.A., Burke, D.F., and Blundell, T.L. (2003). Ab initio construction of polypeptide fragments: accuracy of loop decoy discrimination by an all-atom statistical potential and the AMBER force field with the Generalized Born solvation model. Proteins *51*, 21–40.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System (http://www.pymol.org/).

DePristo, M.A., de Bakker, P.I.W., Lovell, S.C., and Blundell, T.L. (2003a). Ab initio construction of polypeptide fragments: efficient generation of accurate, representative ensembles. Proteins *51*, 41–55.

DePristo, M.A., de Bakker, P.I.W., Shetty, R.P., and Blundell, T.L. (2003b). Discrete restraint-based protein modeling and the Cα-trace problem. Protein Sci. *12*, 2032–2046.

DePristo, M.A., de Bakker, P.I.W., and Blundell, T.L. (2004). Heterogeneity and inaccuracy in protein structures solved by x-ray crystallography. Structure *12*, 831–838.

Drenth, J. (1999). Principles of Protein X-Ray Crystallography, Second Edition (Heidelberg, Germany: Springer-Verlag).

Engh, R.A., and Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. Acta Crystallogr. A *47*, 392–400.

Gerstein, M., Edwards, A., Arrowsmith, C.H., and Montelione, G.T. (2003). Structural genomics: current progress. Science *299*, 1663.

Harrison, S.C. (2004). Whither structural biology? Nat. Struct. Mol. Biol. *11*, 12–15.

Hendrickson, W.A. (1985). Stereochemically restrained refinement of macromolecular structures. In In Methods in Enzymology, H.W. Wyckoff, C.H.W. Hirs, and S.N. Timasheff, eds. (New York: Academic Press, Inc.), pp. 252–270.

Hodel, A., Kim, S.H., and Brunger, A.T. (1992). Model bias in macromolecular crystal structures. Acta Crystallogr. A *48*, 851–858.

Johnson, R.J.K., Christodoulou, J., Dumoulin, M., Caddy, G.L., Alocer, M.J., Murtagh, G.J., Kumita, J.R., Larsson, G., Robinson, C.V., Archer, D.B., Luisi, B., and Dobson, C.M. (2005). Rationalising Lysozyme Amyloidosis: Insights from the Structure and Solution Dynamics of T70N Lysozyme. J. Mol. Biol., in press(1986). Model bias in macromolecular crystal structures. EMBO J. *5*, 819–822.

Jones, T.A., and Thirup, S. (1986). Using known substructures in protein model building and crystallography. EMBO J. *5*, 819–822.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Kleywegt, G.J. (2002). Homo crystallographicus–quo vadis? Structure *10*, 465–472.

Kleywegt, G.J., and Jones, A.T. (1996). Efficient rebuilding of protein structures. Acta Crystallogr. D Biol. Crystallogr. *52*, 829–832.

Lartigue, A., Campanacci, V., Roussel, A., Larsson, A.M., Jones, T.A., Tegoni, M., and Cambillau, C. (2002). X-ray structure and ligand binding study of a moth chemosensory protein. J. Biol. Chem. *277*, 32094–32098.

Lovell, S.C., Word, J.M., Richardson, J.S., and Richardson, D.C. (2000). The penultimate rotamer library. Proteins *40*, 389–408.

Lovell, S.C., Davis, I.W., Arendall, B., III, de Bakker, P.I.W., Word, J.M., Prisant, M.G., Richardson, J.S., and Richardson, D.C. (2003). Structure validation by Cα geometry: ϕ, ψ, and Cβ deviation. Proteins *50*, 437–450.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr. D Biol. Crystallogr. *53*, 240–255.

Navaza, J. (1994). AMoRe: an automated package for molecular replacement. Acta Crystallogr. A *50*, 157–163.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–325.

Perrakis, A., Morris, R., and Lamzin, V.S. (1999). Automated protein model building combined with iterative structure refinement. Nat. Struct. Biol. *6*, 458–463.

Perrakis, A., Harkiolaki, M., Wilson, K.S., and Lamzin, V.S. (2001). ARP/wARP and molecular replacement. Acta Crystallogr. D Biol. Crystallogr. *57*, 1445–1450.

Sali, A., and Blundell, T.L. (1993). Comparative protein modelling by satisfaction of spatial restraints. J. Mol. Biol. *234*, 779–815.

Shetty, R.P., de Bakker, P.I.W., DePristo, M.A., and Blundell, T.L. (2003). Advantages of fine-grained side chain conformer libraries for protein modelling. Protein Eng. *16*, 963–969.

Tronrud, D.E. (2004). Introduction to macromolecular refinement. Acta Crystallogr. D Biol. Crystallogr. *60*, 2156–2168.

Watenpaugh, K.D., Sieker, L.C., Herriott, J.R., and Jensen, L.H. (1973). Refinement of the model of a protein: rubredoxin at 1.5 Å resolution. Acta Crystallogr. B *29*, 943–956.

Yu, B., Blaber, M., Gronenborn, A.M., Clore, G.M., and Caspar, D.L. (1999). Disordered water within a hydrophobic protein cavity visualized by X-ray crystallography. Proc. Natl. Acad. Sci. USA *96*, 103–108.

Accession Numbers

The final model and structure factors have been deposited in the PDB with code 1W08.