Technical Advance

Cell PRESS

Flexible Fitting of Atomic Structures into Electron Microscopy Maps Using Molecular Dynamics

Leonardo G. Trabuco,[1,2,6] Elizabeth Villa,[1,2,6] Kakoli Mitra,[3,7] Joachim Frank,[3,4,8] and Klaus Schulten[1,5,*]

[1]Beckman Institute
[2]Center for Biophysics and Computational Biology
University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[3]Howard Hughes Medical Institute, Health Research Incorporated, Wadsworth Center, Empire State Plaza, Albany, NY 12201, USA
[4]Department of Biomedical Sciences, State University of New York, Albany, NY 12222, USA
[5]Department of Physics, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
[6]These authors contributed equally to this work.
[7]Present address: Skirball Institute, New York University, New York, NY 10016, USA.
[8]Present address: Howard Hughes Medical Institute, Department of Biochemistry and Molecular Biophysics and Department of Biology, Columbia University, New York, NY 10032, USA.
*Correspondence: kschulte@ks.uiuc.edu

SUMMARY

A novel method to flexibly fit atomic structures into electron microscopy (EM) maps using molecular dynamics simulations is presented. The simulations incorporate the EM data as an external potential added to the molecular dynamics force field, allowing all internal features present in the EM map to be used in the fitting process, while the model remains fully flexible and stereochemically correct. The molecular dynamics flexible fitting (MDFF) method is validated for available crystal structures of protein and RNA in different conformations; measures to assess and monitor the fitting process are introduced. The MDFF method is then used to obtain high-resolution structures of the *E. coli* ribosome in different functional states imaged by cryo-EM.

INTRODUCTION

A key to understanding how biological systems work is to look at their structures captured in their various functional states. Experimental techniques reveal different levels of macromolecular structure: high-resolution techniques such as X-ray crystallography furnish atomic detail, but structures obtained are often in functionally undefined states; techniques such as cryo-electron microscopy (cryo-EM) image systems captured in different functional states, albeit at lower resolution (Saibil, 2000). Computational techniques can help bridge the resolution gap by adapting high-resolution crystallographic structures to EM maps, thus providing atomic detail of the system in different functional states. These techniques can also be used to analyze the physical and dynamical properties of the resulting structures, revealing astonishing views of cellular processes.

Until a few years ago, typical resolutions for EM maps of biomolecules were 15–30 Å, and high-resolution crystal structures were often available only for domains of a biomolecular complex

(Frank, 2002; Rossmann, 2000). This led to the development of so-called rigid-body docking techniques that fit atomic structures into density maps keeping the high-resolution structure rigid, usually by performing an exhaustive search over all rotational and translational degrees of freedom in real or reciprocal space, guided by some choice of similarity measure. Rigid-body docking has reached maturity, permitting today the independent docking of parts of the assembly into a map, thus producing a jigsaw structure of the macromolecular complex as a whole. The main limitation of this class of methods is that neither the interaction between the docked subunits nor the difference in conformation between the structures of the domains in crystal form and in the complex can be determined. A comprehensive review of the available rigid-body docking methods can be found in Wriggers and Chacón (2001).

Recent improvements in the resolution of cryo-EM structures motivated the development of methods to flexibly fit atomic structures into density maps. In these methods, many degrees of freedom are considered in the fit, allowing the atomic structures to undergo conformational changes that improve their correspondence to the EM map. Various approaches to flexible fitting have recently been employed and have provided insight into the structural mechanisms of a number of biomolecular processes. One of the first attempts to introduce flexibility in the fitting process, still in use today, consists in dividing a macromolecule into "rigid bodies" and fitting them independently (e.g., Volkmann et al., 2000). Refinement techniques can then be applied to the resulting structure, such as the real-space refinement originally developed for X-ray crystallography (Chapman, 1995; Chen et al., 2001). Other approaches include the use of so-called vector quantization, where a reduced representation is derived from both the atomic structure and the EM map and used as constraints in a molecular mechanics refinement (Wriggers et al., 2000; Wriggers and Birmanns, 2001); the use of a linear combination of low-frequency normal modes to deform the atomic structure, applying minimization techniques to maximize the correlation coefficient between atomic structure and EM map (Tama et al., 2004; Suhre et al., 2006); the combination of comparative modeling, based on alternative alignments and

loop conformations, and structure refinement (Topf et al., 2006); the generation of deformed structures based on the variability exhibited by the protein domains of a superfamily and subsequent selection of the best fit based on the cross-correlation coefficient (Velazquez-Muriel et al., 2006); the use of a combination of restraints imposed by the EM map and a deformable elastic network (Schröder et al., 2007); the use of Monte Carlo simulations with constraints derived from a rigidity analysis (Jolley et al., 2008); and a hierarchical approach consisting of docking the structure as one or more rigid bodies with a Monte Carlo search, followed by two refinement steps based on minimization of a scoring function and simulated annealing (Topf et al., 2008). Li and Frank (2007) recently correlated an ensemble of conformations from equilibrium molecular dynamics (MD) simulations with cryo-EM data, suggesting that a cryo-EM map can be interpreted as a conformational average and that using MD to flexibly fit structures into EM maps should yield structures representative of the conformational ensemble represented by the EM map, particularly when combined with an enhanced sampling technique.

We suggest here a novel method to perform MD simulations to flexibly fit atomic structures into EM maps, with the MD simulation incorporating EM data through an external potential. Forces proportional to the density gradient of the EM map are added in the MD simulation of the atomic structure, effectively biasing the system toward the region of conformational space of interest, i.e., one that is consistent with the density distribution of the EM map. Since large forces are applied and simulations presently are performed in vacuo, harmonic restraints must be applied to keep secondary structure elements intact, thus preventing structural distortions and "overfitting." The molecular dynamics flexible fitting (MDFF) method presented brings together several of the most desirable features of existing methods. First, MDFF takes into account all information contained in the map by avoiding the use of reduced representations or global measures of similarity to drive the fitting. Since the external potential is applied locally, it is possible to fit some components of a macromolecular assembly even when the structure of the remaining components is not available. Moreover, the degree of success of the MDFF method is expected to be independent of system size, which is an advantage over the use of optimization or Monte Carlo-based approaches that employ global-optimization criteria, where an increase in system size decreases the likelihood of successful transitions.

The two methods that MDFF is based on, namely, conventional MD simulation and 3D cryo-EM single-particle reconstruction, are introduced in the online Supplemental Data (see Molecular Dynamics Situations and 3D EM Reconstructions) that accompanies this article for the nonexpert reader. In the following, we describe how the MDFF method incorporates EM data into MD simulations and how it applies restraints to preserve the integrity of structural elements. We demonstrate MDFF by fittings into noise-free, simulated maps created from atomic structures. Finally, as an example application of MDFF, we fit atomic structures into cryo-EM maps of the *E. coli* ribosome at different resolutions.

MD Simulation with an EM-Derived External Potential

In the MDFF method, an external potential derived from the EM map is introduced into an MD simulation to steer the atoms into high-density regions. The stereochemical quality of the structure is preserved by the MD force field and also through harmonic restraints applied to enforce the integrity of secondary structure. The method, therefore, adds two extra terms to the potential energy function of an MD simulation

$$U_{total} = U_{MD} + U_{EM} + U_{SS}, \tag{1}$$

where U_{MD} is the conventional MD potential energy function (Supplemental Data, see Molecular Dynamics Situations), U_{EM} corresponds to a potential derived from the EM data, and U_{SS} is a potential that aims to preserve the secondary structure of protein and nucleic acids. We now describe the latter two terms.

The data provided by cryo-EM reconstructions represent the Coulomb potential of the macromolecule (Supplemental Data, see 3D EM Reconstructions); the dependence of this potential on the atomic number of the composing atoms makes it roughly proportional to the mass density of the macromolecule. It is then sensible to define a potential so that when the atomic structure is placed in it, the atoms are driven through application of forces into high-density areas and away from low-density areas. This potential can be defined on a grid, thus preserving all the information contained in the EM density map. The potential energy function corresponding to this map is

$$U_{EM}(\mathbf{R}) = \sum_j w_j V_{EM}(\mathbf{r}_j), \tag{2}$$

where $\mathbf{R}$ collects all atom coordinates and

$$V_{EM}(\mathbf{r}) = \begin{cases} \xi \left[1 - \frac{\Phi(\mathbf{r}) - \Phi_{thr}}{\Phi_{max} - \Phi_{thr}} \right] & \text{if } \Phi(\mathbf{r}) \geq \Phi_{thr}, \\ \xi & \text{if } \Phi(\mathbf{r}) < \Phi_{thr.} \end{cases} \tag{3}$$

Here $\Phi(\mathbf{r})$ is the Coulomb potential revealed by cryo-EM, Φ_{max} is the maximum value of all voxels in the EM map, ξ is an arbitrary scaling factor ($\xi > 0$) discussed below, Φ_{thr} is a threshold value used to disregard data not corresponding to the biomolecule, i.e., solvent density, w_i is a weight that can be varied according to the type of atom placed in this potential, typically set to the atomic mass, and $\mathbf{r}_j$ is the position of atom j. The global minimum of U_{EM} alone corresponds to all atoms collapsed on the density maximum; however, the other two terms in Equation 1 counterbalance this effect, preserving physically sound structures.

The threshold value Φ_{thr} is selected in accordance with the density histogram of the EM map, which reveals two peaks of different density (Frank, 2006); an example is shown in Figure 1A. The first and higher peak corresponds to solvent density, whereas the second, broader peak corresponds to protein and nucleic acid density. The density histogram thus suggests a natural threshold value at the minimum between the two density peaks; clamping all values below this threshold removes the solvent contribution and yields a flat potential in the solvent regions. For cases where the solvent and biomolecule peaks are not well resolved, the average, which generally lies at the solvent peak, is chosen as the threshold, thus removing much of the undesired density while conservatively avoiding loss of macromolecular density information, as portrayed in Figure 1C.

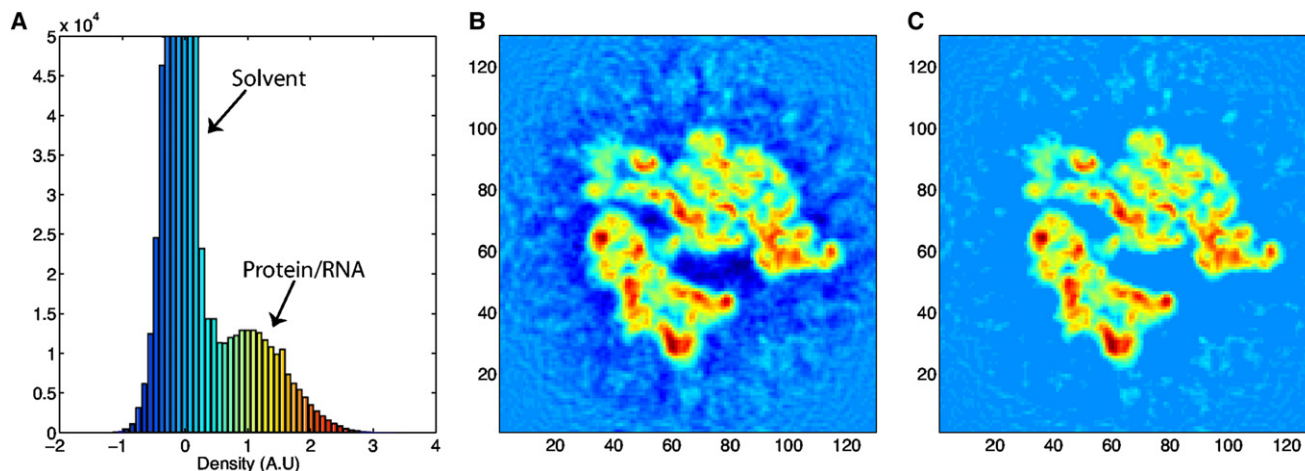The force applied to an atom inside the potential defined by Equation 3 is

Figure 1. Reconstruction of the *E. coli* Ribosome from Cryo-EM Data at ~12.8 Å Resolution
(A) A density histogram shows two distinct peaks pertaining to the solvent and macromolecule; (B) 2D slice of the density; (C) 2D slice of the density after clamping values below the average, thus homogenizing the density corresponding to the solvent surrounding the macromolecule and the bulk solvent. Unpublished data from K.M., L.G.T., E.V., A. Zavialov, M. Ehrenberg, K.S., and J.F.

$$\mathbf{f}_i^{EM} = -\frac{\partial}{\partial \mathbf{r}_i} U_{EM}(\mathbf{R}) = -w_i \frac{\partial}{\partial \mathbf{r}_i} V_{EM}(\mathbf{r}_i) \qquad (4)$$

where U_{EM} is the potential energy function introduced in Equations 2 and 3. $\mathbf{f}_i^{EM}$ can thus be tuned via the scaling factor ξ, which is the same for all atoms, and the weight w_i, which can be defined on a per-atom basis. The external forces $\mathbf{f}_i^{EM}$ are applied in MDFF using grid-steered molecular dynamics, an extension of the SMD method (Isralewitz et al., 2001; Sotomayor and Schulten, 2007) that allows an arbitrary steering potential to be defined on a grid (Wells et al., 2007). The force applied to each atom *i* depends only on the gradient of the potential derived from the EM density map at position $\mathbf{r}_i$; thus, the fitting in MDFF is performed locally.

In order to preserve the stereochemical quality of the structure and prevent overfitting in MDFF, harmonic restraints are applied to a set of internal coordinates relevant to the secondary structure of the macromolecule in its initial conformation—in many cases, the crystal structure. For protein structures, harmonic restraints are applied to ϕ and ψ dihedral angles of amino-acid residues in helices and β strands. For nucleic acids, the software package RNAView (Yang et al., 2003) is used to identify and classify base pairs. The following base pair types, which were observed to preserve the secondary structure of RNA in simulation when restrained, are selected: W/W, W/H, W/S, H/H, H/S, and stacked (W , Watson-Crick edge; H, Hoogsteen edge; S, sugar edge; see Figure 2A). For these base pairs, harmonic restraints are applied to the seven dihedral angles (α, β, γ, δ, ε, ζ, and χ) and two interatomic distances (d_1 and d_2) depicted in Figure 2B, the latter to preserve the planarity of the base pair. Thus, an extra term introducing secondary structure harmonic restraints is added to the potential energy function, namely

$$U_{SS} = \sum_{\mu} k_{\mu} (X_{\mu} - X_{\mu}^0)^2, \qquad (5)$$

where the restraints X_{μ} stand for protein dihedral angles ϕ and ψ, RNA dihedral angles α, β, γ, δ, ε, ζ, and χ, and RNA distances d_1 and d_2. The equilibrium values X_{μ}^0 are generally taken from the

initial atomic structure but can also be set to ideal values. Additional restraints can be added, such as codon-anticodon interactions in the ribosome, as discussed below.

The MDFF simulations are performed using NAMD (Phillips et al., 2005), with the CHARMM27 force field (MacKerell et al., 1998; Foloppe and MacKerrell, 2000) in vacuo, using a dielectric constant of 80. A multiple time-stepping integration scheme is used, calculating bonded interactions every 1 fs and nonbonded interactions every 2 fs; a cutoff distance of 10 Å is used for the nonbonded interactions. Temperature is maintained at 300 K using a Langevin thermostat (Brünger et al., 1984) coupled to all heavy atoms with a damping coefficient of 5 ps^{-1}. Rigid-body docking performed as the initial step of a fitting protocol in all examples presented was executed using Situs (Wriggers et al., 1999).

Flexible-Fitting Protocol
Obtaining an optimal fit in MDFF relies on a balance among the three terms in Equation 1. The first term, U_{MD}, is given by the choice of standard force field and is here not subject to alteration; the other two terms, U_{EM} and U_{SS}, contain parameters that can be tuned and represent a trade-off: higher forces derived from the EM map (U_{EM}) and lower restraints to secondary structure (U_{SS}) can yield a better fit but can also lead to distortions of the macromolecular structure, a natural concern in flexible fitting methods generally referred to as "overfitting" (Tama et al., 2004). For U_{EM}, parameters that can be varied are the scaling factor ξ and atomic weights w_i (Equations 2 and 3). The latter are typically set to the atomic masses while ξ determines the range in magnitude of forces applied to the atoms and is usually set to values around 0.3 kcal/mol, resulting in forces typically on the order of 10–15 pN (piconewton) per atom (for a carbon atom). The gradient of the potential derived from the EM density map can be calculated before the fitting to pick a value of ξ: high values can result in high accelerations of atoms, rendering the simulation unstable. The term U_{SS} requires the selection of
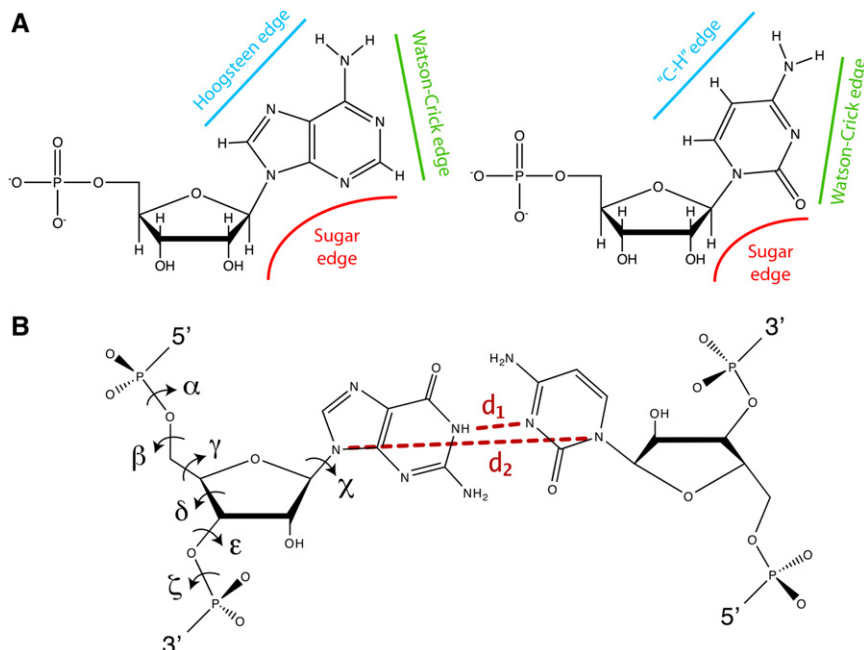
Figure 2. Harmonic Restraints Applied to Base-Paired RNA Residues
(A) RNA interaction edges for both purines (adenine is shown on the left) and pyrimidines (cytosine is shown on the right), according to Leontis and Westhof (1998); (B) dihedral angles, and the two interatomic distances (dashed lines) to which harmonic restraints are applied.

parts of the structure, which is useful, for example, for fitting protein-RNA complexes, as discussed below—the deletion is done by assuming the current fit is perfect, creating a simulated map from the fitted structure as described in the next section, and using it as a mask to selectively delete a portion of the original EM map. The new map Φ' defining U_{EM} as in Equation 3 is given by

$$\Phi' = \left(1 - \frac{\Phi_{sim}}{\max(\Phi_{sim})}\right)\Phi \qquad (6)$$

where Φ is the original EM map and Φ_{sim} is the simulated map, with a maximum voxel value of $\max(\Phi_{sim})$. Equation 6 results in potentials with smoother gradients than those obtained from difference maps.

An important advantage of multistep protocols in MDFF is evident when fitting protein-RNA complexes into EM maps. The atomic number of the atoms composing RNA is on average higher than those composing protein, which may lead to incorrect fitting of proteins due to lower-energy minima attracting them into RNA density. To avoid this problem, the RNA may be fitted first, followed by deleting the RNA contribution from the EM map once it is well fitted and by then proceeding with the fitting of the proteins. Alternatively, one may positionally restrain the RNA structure to prevent the protein from being drawn into the density the RNA occupies.

spring constants k_μ in Equation 5. Typical values used in MDFF are $k_\mu = 200$ kcal mol^{-1} rad^{-2}, or 200 kcal mol^{-1} Å^{-2} in the case of RNA distances (300 kcal mol^{-1} Å^{-2} is a typical spring constant value for a carbon-carbon bond in the CHARMM27 force field). The choice of these values renders the restrained secondary structure elements relatively stiff, thus preserving their structure throughout the simulation.

MDFF can also be performed in multiple steps, varying parameters at every step, such that the atomic structure gradually converges into the EM map. Generally, a step is considered "finished" when the MD simulation has converged as evaluated by a goodness measure, e.g., by stabilization of the root-mean-square deviation (rmsd, given by $\langle(\mathbf{r}_i - \langle\mathbf{r}_i\rangle)^2\rangle^{1/2}$, where $\mathbf{r}_i$ is the position of each atom i); one can also track the correlation between the trajectory and the EM map, as described later. Naturally, in a last step of MDFF one can switch U_{EM} and U_{SS} off and equilibrate the system subject only to the intrinsic potential U_{MD}.

In a multistep protocol of MDFF, several parameters can be varied besides those mentioned above: (1) *Temperature* can be adjusted to allow the system to overcome energy barriers, exploring a larger portion of the conformational space in less time, often done in simulated annealing simulations (Frenkel and Smit, 2002); (2) *low-pass filtering* the map in a first step can be used to avoid getting trapped in local minima, followed by use of the original map—this approach first induces overall domain motions and refines the structure subsequently on the local scale using the high-resolution information in the original map; (3) along the same line, one can initially apply strong harmonic restraints to the secondary structure and in a subsequent step *weaken the restraints* to refine the structure; (4) at any given time, any portion of the structure may be *fixed* or *restrained* to its current position (positional restraints)—this is useful, e.g., when factors or low-occupancy ligands are introduced for fitting, and the rest of the previously fitted structure can be fixed or restrained while the ligands are fitted; (5) one can also *delete the density* corresponding to

Fitting Atomic Structures into Simulated Maps
In the following sections, we demonstrate application of MDFF to noise-free, simulated EM maps as a means of validation of the method. We briefly describe how simulated maps are generated from atomic structures and how we calculate cross-correlation coefficients between a fitted structure and the EM map. Using examples where X-ray crystal structures are available in two different conformations of the same molecule, we apply the method to fit one conformation into simulated maps generated from the other.

Simulated EM Maps from Atomic Structures
Noise-free simulated EM maps can be created from atomic structures using the approach described in Stewart et al. (1993), which assumes that the EM map represents the electrostatic potential of the nuclei. The atomic number of each atom is interpolated onto a grid, and the resulting 3D density map is subsequently low-pass filtered to the desired resolution. Simulated maps can be generated from a single structure or from a set of structures obtained from an MD simulation by averaging maps created from each frame of the trajectory (see Li and Frank,

Table 1. Effect of Resolution and Grid Spacing on Flexible Fitting into Noise-Free Simulated Maps using MDFF

System (PDB Codes)	Resolution (Å)	Initial Rmsd (Å)	Initial (Local) Correlation	Grid Spacing (Å)	Final Mean Rmsd (Å)	Final Mean (Local) Correlation
ACS/CODH (1OAO)	5.0	14.82	0.677 (0.282)	1.0	0.75	0.988 (0.934)
				2.0	1.04	0.986 (0.917)
	10.0	14.33	0.778 (0.478)	1.0	1.25	0.996 (0.977)
				2.0	1.25	0.996 (0.976)
	15.0	13.92	0.817 (0.582)	1.0	2.01	0.995 (0.980)
				2.0	2.07	0.995 (0.980)
16S rRNA (2AW7, 2AVY)	5.0	3.60	0.873 (0.611)	1.0	0.65	0.992 (0.961)
				2.0	0.84	0.988 (0.939)
	10.0	3.57	0.947 (0.812)	1.0	1.04	0.996 (0.984)
				2.0	1.21	0.995 (0.976)
	15.0	3.56	0.973 (0.901)	1.0	2.06	0.992 (0.972)
				2.0	2.21	0.991 (0.965)

For each system, maps were created computationally from a given crystal structure of conformation I and a crystal structure available in an alternative conformation, II, was fitted into the computed map. The initial backbone rmsd and cross-correlation coefficients correspond to rigid-body docked structures into computed maps with grid spacing of 2.0 Å using Situs with default options. The final mean backbone rmsd values were calculated from the last 200 ps of 500 ps trajectories. The final mean cross-correlation coefficients were calculated using computed maps generated from the average of the trajectories. Corresponding local cross-correlation coefficients that consider only the molecular envelope are also shown in parenthesis (a threshold of 0.2 σ was used in these examples).

2007). In the present work, simulated maps are used for three different purposes: (1) calculate cross-correlation coefficients between EM maps and atomic structures; (2) delete selected regions from the experimental EM map based on a fitted atomic structure; and (3) provide noise-free maps for validation purposes.

Cross-Correlation Coefficients
To quantify the goodness of the fit, a simulated map can be generated from the fitted atomic structure with the same target resolution as the EM map. The Pearson's correlation coefficient (usually referred to in the EM literature as the cross-correlation coefficient) between these two data sets, i.e., the simulated (S) and the experimental (E) 3D maps, can be used as a measure of similarity between them, and is given by

$$\rho_{SE} = \frac{\langle (S - \langle S \rangle)(E - \langle E \rangle) \rangle}{\sigma_S \sigma_E}, \tag{7}$$

where $\langle S \rangle$ and $\langle E \rangle$ correspond to the average voxel values of the simulated and experimental maps, respectively, and σ_S and σ_E correspond to their standard deviation (Frank, 2006). Note that the cross-correlation coefficient is normalized, i.e., $\rho_{SE} \in [-1, 1]$. All cross-correlation coefficients reported were computed considering either all voxel values in the density maps ("global correlation") or only voxels inside the molecular envelope of the simulated map ("local correlation") (Roseman, 2000), using a threshold that we report as the number of standard deviations (σ) above the mean of the simulated map. The global correlation, commonly quoted in the literature, depends sensitively on the box size arbitrarily selected by the electron microscopist. Larger boxes result in artificially higher correlation values, leading to overestimation of the quality of the fit; thus, local correlations should be preferred.

Validation Using Atomic Structures in Two Conformations
In order to validate the method and estimate the accuracy of the fitted structures, we use X-ray structures of molecules available

in two conformations as test cases. One of the structures is fitted into noise-free simulated maps of the other at 5, 10, and 15 Å resolutions, after an initial phase of rigid-body docking. Here we present two examples: acetyl-coenzyme A synthase/carbon monoxide dehydrogenase (ACS/CODH) (Darnault et al., 2003) and a 16S rRNA (Schuwirth et al., 2005). Several other examples are presented in Supplemental Data (see Fitting Atomic Structures into Simulated Maps).

All simulations used ξ = 0.3 kcal/mol and k_μ = 200 kcal mol^{-1} rad^{-2} (or 200 kcal mol^{-1} Å^{-2}) for the harmonic restraints (see Equation 5 and Figure 2B), except for 16S rRNA dihedral restraints that were enforced using k_μ = 50 kcal mol^{-1} rad^{-2}. The simulations converged in less than 200 ps of simulation but they were run for 500 ps to improve statistics. Table 1 lists the correlation coefficients and mean backbone rmsd between the fitted and target structures, calculated from the last 200 ps of trajectories (a plot of the rmsd through the entire trajectories for ACS/CODH can be found in the Supplemental Data [see Fitting Atomic Structures into Simulated Maps]). One can see that the rmsd decreases with higher resolution. It is important to note that the simulations result in a representative set of structures that fit the map. Since fitting atomic structures into low-resolution data is an indeterminate problem, a representative set of conformers should be considered when interpreting the data. The structures presented in this paper serve as an illustration of one such representative structure for each of the problems considered. Figure 3 shows the initial and target conformation for each of the systems, as well as representative structures from MDFF into 10 Å resolution simulated maps.

The final structures closely match the target structures used to generate the simulated maps; as expected, the match is less precise for lower resolutions or larger grid spacings. As an example, at 10 Å resolution, which represents a typical resolution for EM maps today, the final mean backbone rmsd is 1.25 Å for ACS/CODH and 1.04 Å for the 16S rRNA. The additional test cases presented in the Supplemental Data (see Fitting Atomic Structures
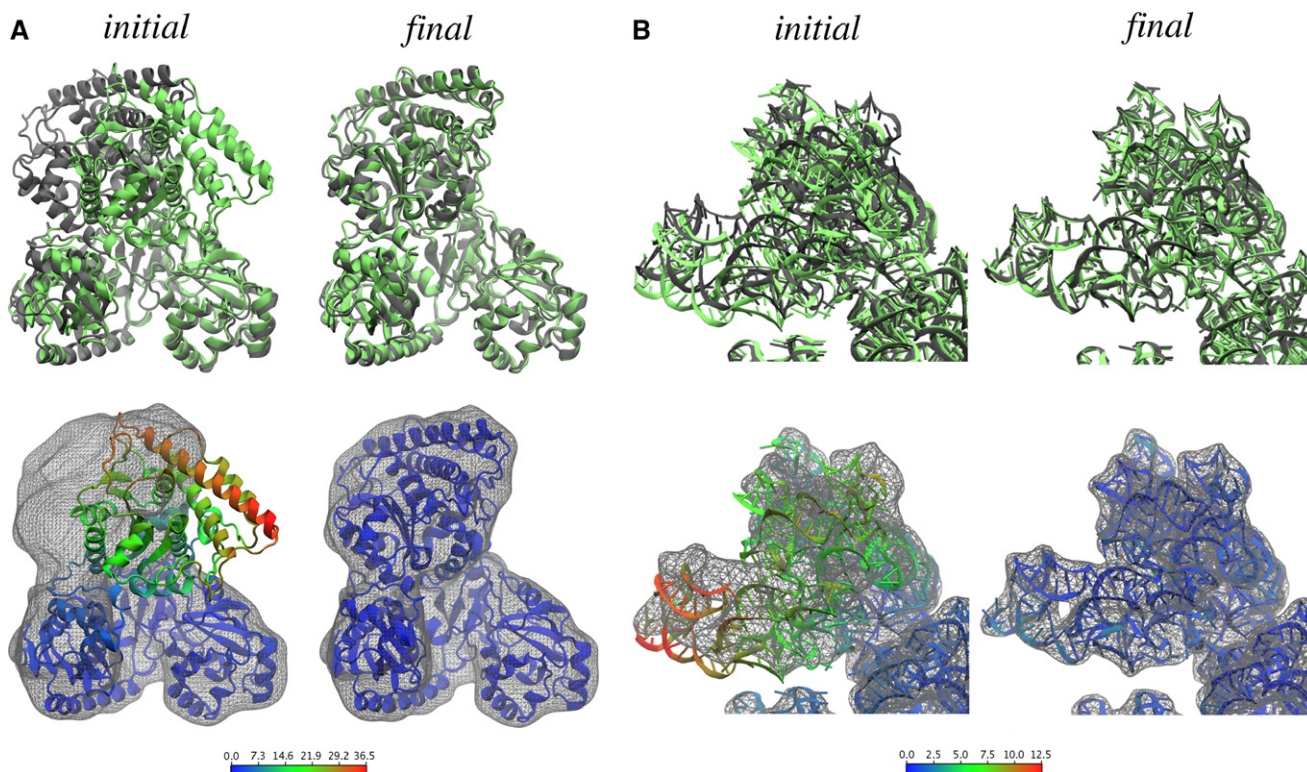
Figure 3. Validation of the MDFF Method Using X-Ray Structures in Two Conformations

(A) Acethyl-CoA synthase/carbon monoxide dehydrogenase; (B) 16S rRNA (only the head is shown for clarity, since it is the only region where the two conformations differ significantly). The target structures and simulated maps are shown in gray, whereas the initial and final fitted structures are shown in green (top) and colored by backbone rmsd per residue with respect to the target structures (bottom; color scales in Å). The final structures correspond to fittings into 10 Å simulated maps generated from the target structures. Movies of the fittings are included in the Supplemental Data (see Movies S1 and S2).

into Simulated Maps) also show that the accuracy obtained by applying MDFF is comparable to other methods (Jolley et al., 2008; Topf et al., 2008) and that MDFF can describe a large range of conformational changes, such as movements around a hinge and domain shearing. The fact that smaller grid spacings yield better results suggests that higher-order interpolation schemes in the grid-based force calculation can improve the fitting; indeed, the use of cubic interpolation in the current implementation of grid-steered molecular dynamics in NAMD (used in the calculations presented) yielded slightly better results when compared to linear interpolation (data not shown). The simulated maps presented in this section differ significantly from EM maps in that they do not contain noise emerging from the imaging process and numerical errors due to image processing and reconstruction, and in that they represent a single structure instead of an ensemble average as captured by cryo-EM. An attempt to address the latter is presented in the Supplemental Data (see Simulated Maps Generated from an Ensemble of Structures), where a protein structure is fitted into maps created from an ensemble of structures obtained from equilibrium MD simulations. For these tests, the fluctuation of atomic positions observed in MDFF reproduce reasonably well the fluctuations on the target map, especially for resolutions in the range of 10–15 Å.

Example Application: The E. coli Ribosome

The ribosome, a complex macromolecular machine responsible for protein synthesis in all cells, is one of the biological

systems for which cryo-EM has provided much insight to date (Frank, 2003). The ribosome undergoes several conformational changes and binds a number of cofactors throughout the process of protein synthesis. Different functional states have been extensively imaged by cryo-EM at ever-increasing resolution (Frank and Spahn, 2006). The ribosome has driven the development of methods to obtain quasi-atomic structures by combining X-ray crystallographic structures with cryo-EM maps (e.g., Wriggers et al., 2000; Gao et al., 2003; Tama et al., 2004; Mitra et al., 2005, 2006; for a review, see Mitra and Frank, 2006).

We present the ribosome as an application of MDFF, using EM maps corresponding to two functional states, namely, the *E. coli* ribosome in complex with the ternary complex EF-Tu·aminoacyl-tRNA·GDP stalled by the antibiotic kirromycin (70S-fMet-tRNA^{fMet}-Phe-tRNA^{Phe}·EF-Tu·GDP·kir) at resolutions of 9 Å (Valle et al., 2003) and 6.7 Å (J. LeBaron, R.A. Grassucci, T. Shaikh, W. Baxter, J. Sengupta, and J.F., unpublished data), and a ribosome with an accommodated A-site tRNA at 9 Å resolution (Valle et al., 2003). The first functional state corresponds to the initial selection of the ternary complex (TC) in the elongation cycle, after which EF-Tu leaves the ribosome and, subsequently, the peptide bond is formed, resulting in the second functional state, with a deacylated tRNA occupying the P site and an accommodated fMet-Phe-tRNA^{Phe} in the A site (70S-tRNA^{Phe}-MF-tRNA^{Phe}). The atomic model of the ribosome

used in this section is described in the Supplemental Data (see Atomic Model of the *E. coli* Ribosome).

The flexible fitting was performed following a multistep protocol: (1) After rigid-body docking, the 30S and 50S subunits were flexibly fitted, using $\xi = 0.3$ kcal/mol, $k_{\mu,RNA} = 200$ kcal mol^{-1} rad^{-2} or 200 kcal mol^{-1} Å^{-2} for angles and distances, respectively, and $k_{\mu,protein} = 400$ kcal mol^{-1} rad^{-2}. In addition to the secondary structure restraints, the ϕ and ψ angles of the rest of the amino acid residues were restrained with $k_{\mu,protein} = 200$ kcal mol^{-1} rad^{-2}. (2) The potential corresponding to the 5S, 16S, and 23S rRNAs was deleted, and the atoms in these chains were restrained to their current positions (positional restraints) with a relatively stiff force constant of 10 kcal mol^{-1} Å^{-2}, effectively fixing them. This allowed the fitting of ribosomal proteins to improve, since the potential energy minima corresponding to the ribosomal RNA were removed from the simulation. (3) The harmonic restraints applied to the ribosomal proteins were relaxed to 200 kcal mol^{-1} rad^{-2}, restraining only residues in helices and β sheets. (4) The remaining ligands, namely EF-Tu, tRNAs, and mRNA, were introduced using rigid-body docking. For this step, the original EM potential (before RNA deletion) was used, all ribosomal proteins were positionally restrained, and the positional restraints on the rRNAs were lifted. The same secondary structure restraints to proteins and RNAs used in the previous step were preserved and also applied to the recently introduced ligands. In addition, equivalent RNA restraints were imposed to enforce codon-anticodon interactions between the A-site mRNA codon and the tRNA. During this last step, force scaling was increased to $\xi = 1.0$ kcal/mol. Convergence times and cross-correlation coefficients for each step are presented for all maps in the Supplemental Data (see Comparison between Maps at Different Resolutions). A movie of the fitting into the 6.7 Å map illustrating the multistep protocol can be found in the Supplemental Data (see Movie S3).

The MDFF method applied to different ribosome cryo-EM maps yielded quasi-atomic models that closely fit the EM densities; indeed, the cross-correlation coefficients between the maps and the fitted structures are significantly higher than those obtained from rigid-body docking (local correlations with a threshold of 0.2 σ are given in parenthesis): 0.913 (0.764) versus 0.858 (0.632) (TC-bound ribosome at 6.7 Å), 0.919 (0.739) versus 0.835 (0.503) (TC-bound ribosome at 9.0 Å), and 0.878 (0.735) versus 0.756 (0.513) (ribosome with accommodated A-site tRNA at 9.0 Å). Selected regions for which conformational changes have been previously characterized are presented in Figure 4. It can be seen that distinct structural elements such as RNA and protein helices fit well into their corresponding densities, despite internal restraints imposed to avoid structural distortions. Local cross-correlation coefficients of different elements are presented in the Supplemental Data (see Local Cross-Correlation Coefficient Map).

An overview of the fitting into the 6.7 Å map of the TC-bound ribosome is depicted in Figure 4A, along with a detailed view of the decoding center. Figure 4B shows the X-ray structure of tRNA (crystallized in complex with EF-Tu), a fitted structure, and the same crystal structure fitted manually in a previous study (Valle et al., 2003). It can be seen that the conformation of the anticodon loop (ACL) of the A/T-site tRNA is significantly different from the one adopted in the free TC. Interestingly, our fitted

structure is very similar to the previously proposed one; however, the previous work assumed that the ACL of the A/T-site tRNA adopts the same conformation as that of the free ternary complex, and thus a model of the structure was built by interpolating coordinates between two manual fittings of the same crystal structure (Valle et al., 2003). With the MDFF method, we obtained the conformational change that the ACL undergoes when the TC binds to the ribosome without any assumptions, the fitting being driven only by the EM data. The structure of the ACL of the A/T-site tRNA obtained from our fitting has the same conformation as the ACL of the A-site tRNA observed by Selmer et al. (2006), as shown in Figure 4C. A comparison between the fittings into cryo-EM maps of the TC-bound ribosome at different resolutions (6.7 Å and 9.0 Å) is presented in the Supplemental Data (see Comparison between Maps at Different Resolutions).

Binding of the TC to the ribosome induces a conformational change in the GTPase-associated center (GAC) (Valle et al., 2003). Figure 4D presents a comparison of the GAC conformation in the TC-bound ribosome and the ribosome in complex with an accommodated A-site tRNA. In the first conformation (closed), the GAC approaches the 50S to interact with the TC; when EF-Tu leaves the ribosome, the second conformation (open) arises in which the GAC lobe moves back to its original position. One can see that rigid-body docking of the crystal structure to the ribosomal cryo-EM map with an accommodated A-site tRNA shows a good fit for the GAC, revealing that the crystal structure captures this conformation. Flexible fitting obtains a closer match to the EM map for this state and reveals the closed conformation of the GAC and TC in the TC-bound ribosome. The details of the atomic structures obtained from the 6.7 Å map will be discussed elsewhere (J. Sengupta, E.V., L.G.T., J. LeBaron, W.T. Baxter, T. Shaikh, R.A. Grassucci, P. Nissen, M. Ehrenberg, K.S., and J.F., unpublished data).

Even at subnanometer resolution that permits the identification of secondary structure elements, some regions are not well defined in an EM map. For example, the switch regions of EF-Tu in the 6.7 Å TC-bound ribosomal EM map are not resolved, presumably due to their high flexibility. Though information about their structure is not directly available in the EM map, MD simulations can be performed in the presence of the EM potential to assess the feasibility of different conformations of the switches: an unfeasible conformation will either change during simulation, or alter the quality of the fit of neighboring domains. The conformational dynamics of the interaction of EF-Tu with the ribosome can be deduced from the computational studies, even though a crystal structure of a ribosome-bound EF-Tu is not yet available, and density for some key elements governing the interaction is missing in the EM map (J. Sengupta, E.V., L.G.T., J. LeBaron, W.T. Baxter, T. Shaikh, R.A. Grassucci, P. Nissen, M. Ehrenberg, K.S., and J.F., unpublished data).

DISCUSSION

We have developed a novel method, MDFF, for combining atomic structures and EM maps to reveal atomic details of macromolecular complexes in functional states. High-resolution structures of complexes imaged by cryo-EM permit a better interpretation of the data, e.g., by characterizing the flexibility of
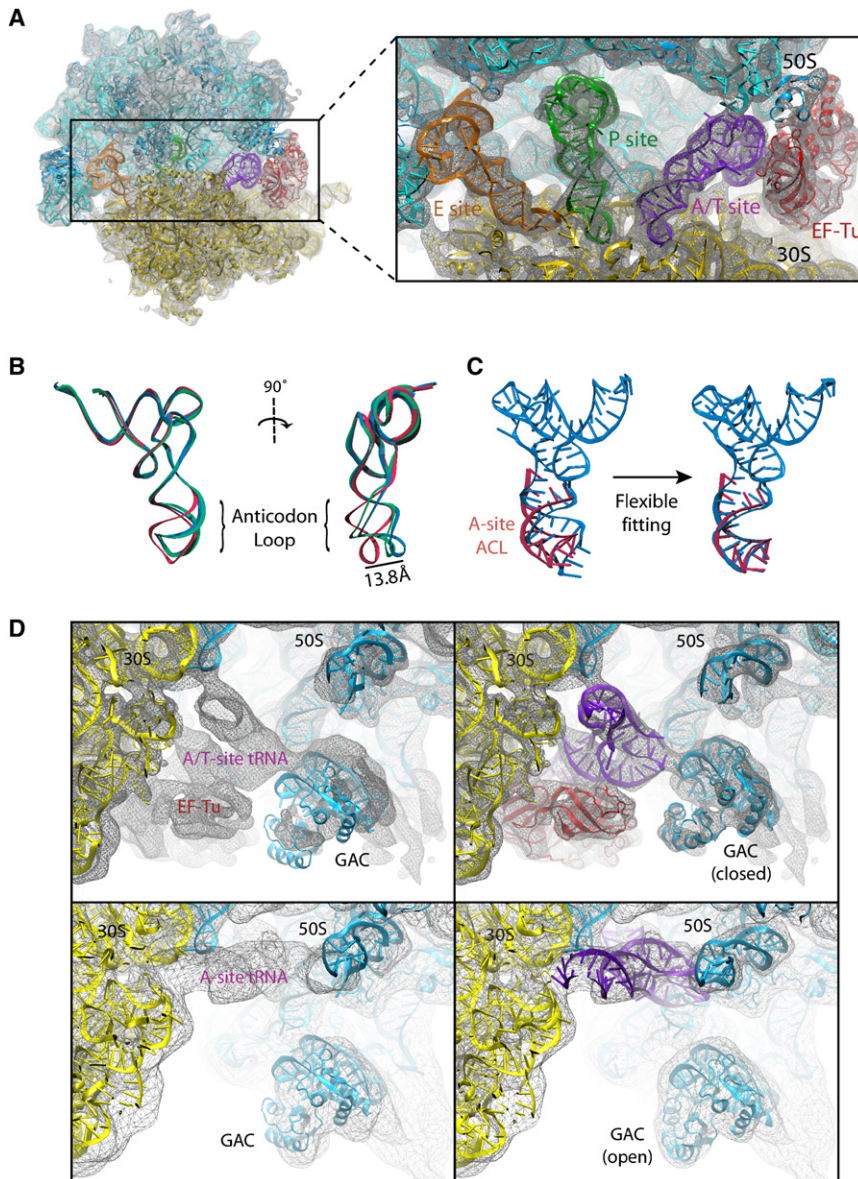
Figure 4. Fitting into the TC-Bound Ribosome Cryo-EM Map at 6.7 Å Resolution by Means of MDFF

(A) Overview of the all-atom ribosome structure fitted into the 6.7 Å map, with a close view into the decoding center (inset).

(B) Conformation of tRNA in the A/T site. The crystal structure from the free TC used as a starting point for the fitting (PDB code: 1OB2; unpublished data) is shown in red; the A/T tRNA model obtained by applying the MDFF method to the 6.7 Å map is shown in blue; the A/T tRNA model previously obtained using a 9.0 Å map constructed by interpolating two manual fittings of tRNA (PDB code: 1OB2) is shown in green (Valle et al., 2003).

(C) Conformation of tRNA in the A/T site (blue) compared to a partial crystal structure of the A-site tRNA (Selmer et al., 2006) (red). The crystal structure from the free TC used as a starting point for the fitting (PDB code: 1OB2; unpublished data) is shown on the left; the A/T tRNA model obtained by applying the MDFF method to the 6.7 Å map is shown on the right.

(D) Conformational dynamics of the GTPase-associated center. Shown are differences in the conformation of the GTPase-associated center between the TC-bound ribosome (EM map at 6.7-Å resolution, top), and the accommodated ribosome (EM map at 9 Å resolution, bottom). Rigid-body docked structures into the corresponding maps, used as initial coordinates for flexible fitting, are shown on the left; flexibly fitted structures are shown on the right.

the complex, identifying key interactions between its composing elements, or locating bound cofactors. The MDFF method takes advantage of the impressive advances both in X-ray crystallography and cryo-EM, by using MD simulations that incorporate EM data through a potential driving an atomic structure into a conformation corresponding to an EM map. The method has several characteristics that bring together the best features of previously developed methods: (1) MDFF avoids the use of *reduced representations* which necessarily discard some of the information contained in the crystal structure or the EM map. (2) MDFF considers information contained in the map through a potential, such that the fitting is performed *locally*, i.e., independently of other regions of a molecule. (3) MDFF uses *global measures* of the fit only to assess convergence and not to drive the fit. Methods that rely on global measures to obtain a good fit become less efficient as the *size of the system* increases. (4) MDFF can be used to fit both *protein and nucleic acids*, as well

as systems composed of both. (5) MDFF can fit parts of the complex independently when complete atomic structures are not available. (6) MDFF does not require *user input* to divide a molecule into pieces to flexibly fit them into the map; rather, the flexibility is indigenous to the molecular structure, with additional restraints dictated by the secondary structure. (7) MDFF follows *multistep protocols* that permit adjusting the method to face various challenges; e.g., in the case of systems composed of protein and nucleic acids, fitting can be performed on nucleic acids first and the protein component second. (8) MDFF ensures *stereochemical correctness* during the fitting process, obviating the need for post-fitting refinement, which often results in deviations from the EM map. In fact, the structures obtained by MDFF can be used as initial coordinates for further simulation studies, in particular, for an equilibration testing the stability of the model arrived at. (9) MDFF uses restraints to preserve *secondary structure* elements and other structural features in order to prevent overfitting; in this respect it bears some similarity to real-space refinement (Chapman, 1995), but MDFF is more automated and more adaptive. (10) MDFF can represent the structural variability present in the experimental map by providing a representative set of fitted structures instead of a single one. (11) MDFF can be extended to take advantage of other enhanced-sampling techniques

such as the use of elastic network models (Tama et al., 2004; Schröder et al., 2007).

A natural concern about flexible fitting methods is overfitting. We have shown that the harmonic restraints proposed are enough to prevent overfitting into noise-free maps. However, the degree of overfitting into experimental EM maps prevented by applying the restraints presented here cannot be evaluated at this time, and thus users are discouraged from softening the tight restraints suggested in this paper. The choice of restraints in MDFF might be improved further, e.g., through incorporation of restraints based on rigidity analysis (Jacobs et al., 2001; Jolley et al., 2008). A limitation intrinsic to our method is that rotation of structural elements are difficult to capture; this may be addressed, e.g., by enhancing the lowest frequency normal modes in the simulation (Zhang et al., 2003) or by refining with MDFF models generated with a flexible fitting method based on normal modes (Tama et al., 2004; Suhre et al., 2006). Naturally, with the current restraints it is not possible to capture conformational changes that involve unfolding of secondary structure elements. Even though the fitting in MDFF is performed locally, proteins that interact with nucleic acids should not be fitted in the absence of the latter, since they might be attracted to the higher nucleic acid density. An interesting alternative involves the use of local, but not global, cross-correlation maps to define U_{EM}.

Naturally, MDFF results cannot be better than the EM data used permit. Any artifact in the molecular density reported in the EM map can be propagated to the fitted structure. The quantitative use of the density map requires high-quality data collection and efficient 3D reconstruction algorithms, as extensively discussed in Frank (2006). The fitting and the interpretation of the resulting structures must be done taking into account factors that prevent the 3D reconstruction from accurately representing the molecular density. An additional concern that the MDFF technique might raise is computational cost. Indeed, the necessary calculations are extensive, akin to other available flexible-fitting methods (Schröder et al., 2007; Wriggers and Birmanns, 2001). However, with the prospect of automated cryo-EM data eventually being able to match the speed of throughput achieved by other structural methods (Zhu et al., 2001), the advantages of not requiring ad hoc user input and of incorporating as much of information from the original data as possible, together with expected increases in computer power, will make MDFF a feasible and attractive tool for obtaining atomic structures from the wealth of EM data. Moreover, by using NAMD, the computational cost of the MDFF method scales linearly with system size, permitting the application of MDFF to large macromolecular complexes.

The success of the MDFF method is evident from the quality of the high-resolution structures obtained for cryo-EM maps of ribosomal complexes. The ribosome represents one of the greatest challenges to fitting methods due to its sheer size, its lack of symmetry, and its mixed composition of protein and RNA. We have recently fitted atomic structures to several ribosomal cryo-EM maps to study the conformational changes that take place during the decoding process. (K.M., L.G.T., E.V., A. Zavialov, M. Ehrenberg, K.S., and J.F., unpublished data; J. Sengupta, E.V., L.G.T., J. LeBaron, W.T. Baxter, T. Shaikh, R.A. Grassucci, P. Nissen, M. Ehrenberg, K.S., and J.F., unpublished data). Typical global cross-correlation coefficients obtained improve from

~0.8 after rigid-body docking, to ~0.9 after completion of the flexible fitting.

In recent years, it has become evident that cellular functions are carried out by assemblies of interacting macromolecules (Alberts, 1998), many of them existing only transiently in the cell. In order to provide a comprehensive description of such complexes, spanning the atomic and the system level, data obtained from various structural techniques must be combined into high-resolution structures with the aid of theoretical approaches (Sali et al., 2003; Alber et al., 2007). MD is a method of choice, being increasingly used to refine macromolecular structures, and is an established tool for studying structural dynamics of large biomolecules. The conceptual grounds of the methodology presented here can easily be extended to other sources of structural data by including them in MD simulations of atomic or coarse-grained structures (Shih et al., 2007; Arkhipov et al., 2006) as external potentials that merge several levels of description into high-resolution structures.

As structural methods become more prolific, automated software capable of interpreting intermediate-resolution structures at the atomic level is becoming crucial; indeed, it seems likely that in years to come this type of approach will yield the main, and perhaps only, source of atomic structures for large macromolecular complexes in functional conformations. The software should be widely available and should not require very detailed technical expertise or ad hoc input, but rather should be easy to use and general. These requirements are met by the MDFF method. The method is currently implemented in the molecular dynamics software package NAMD (Phillips et al., 2005), and the automated setup and analysis of the simulations is done through the molecular visualization software package VMD (Humphrey et al., 1996).

SUPPLEMENTAL DATA

Supplemental Data include an overview of MD simulations and 3D EM reconstructions; several additional validation cases for MDFF, including density maps generated from an MD simulation; a description of the atomic model of the *E. coli* ribosome used in the current work; an analysis of the effect of resolution on the fitting into experimental cryo-EM maps; a local cross-correlation coefficient map for a fitted ribosome structure; and three movies illustrating the MDFF method. Supplemental Data can be found with this article online at http://www.structure.org/cgi/content/full/16/5/673/DC1/.

REFERENCES

Alber, F., Dokudovskaya, S., Veenhoff, L.M., Zhang, W., Kipper, J., Devos, D., Suprapto, A., Karni-Schmidt, O., Williams, R., Chait, B.T., et al. (2007). Determining the architectures of macromolecular assemblies. Nature 450, 683–694.

Alberts, B. (1998). The cell as a collection of protein machines – preparing the next generation of biologists. Cell *92*, 291–294.

Arkhipov, A., Freddolino, P.L., and Schulten, K. (2006). Stability and dynamics of virus capsids described by coarse-grained modeling. Structure *14*, 1767–1777.

Brünger, A.T., Brooks, C.L., III, and Karplus, M. (1984). Stochastic boundary conditions for molecular dynamics simulations of ST2 water. Chem. Phys. Lett. *105*, 495–498.

Chapman, M.S. (1995). Restrained real-space macromolecular atomic refinement using a new resolution-dependent electron-density function. Acta Crystallogr. A *51*, 69–80.

Chen, L.F., Blanc, E., Chapman, M.S., and Taylor, K.A. (2001). Real space refinement of acto-myosin structures from sectioned muscle. J. Struct. Biol. *133*, 221–232.

Darnault, C., Volbeda, A., Kim, E.J., Legrand, P., Vernède, X., Lindahl, P.A., and Fontecilla-Camps, J.C. (2003). Ni-Zn-[Fe-S] and Ni-Ni-[Fe-S] clusters in closed and open α subunits of acetyl-CoA synthase/carbon monoxide dehydrogenase. Nat. Struct. Biol. *10*, 271–279.

Foloppe, N., and MacKerrell, A.D., Jr. (2000). All-atom empirical force field for nucleic acids: I. Parameter optimization based on small molecule and condensed phase macromolecular target data. J. Comput. Chem. *21*, 86–104.

Frank, J. (2002). Single-particle imaging of macromolecules by cryo-electron microscopy. Annu. Rev. Biophys. Biomol. Struct. *31*, 309–319.

Frank, J. (2003). Electron microscopy of functional ribosome complexes. Biopolymers *68*, 223–233.

Frank, J. (2006). Three-Dimensional Electron Microscopy of Macromolecular Assemblies (New York: Oxford University Press).

Frank, J., and Spahn, C.M.T. (2006). The ribosome and the mechanism of protein synthesis. Rep. Prog. Phys. *69*, 1383–1417.

Frenkel, D., and Smit, B. (2002). Understanding Molecular Simulation From Algorithms to Applications (California: Academic Press).

Gao, H., Sengupta, J., Valle, M., Korostelev, A., Eswar, N., Stagg, S.M., Van Roey, P., Agrawal, R.K., Harvey, S.C., Sali, A., et al. (2003). Study of the structural dynamics of the *E. coli* 70S ribosome using real space refinement. Cell *113*, 789–801.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD – Visual Molecular Dynamics. J. Mol. Graph. *14*, 33–38.

Isralewitz, B., Gao, M., and Schulten, K. (2001). Steered molecular dynamics and mechanical functions of proteins. Curr. Opin. Struct. Biol. *11*, 224–230.

Jacobs, D.J., Rader, A.J., Kuhn, L.A., and Thorpe, M.F. (2001). Protein flexibility predictions using graph theory. Proteins *44*, 150–165.

Jolley, C.C., Wells, S.A., Fromme, P., and Thorpe, M.F. (2008). Fitting low-resolution cryo-EM maps of proteins using constrained geometric simulations. Biophys. J. *94*, 1613–1621.

Leontis, N.B., and Westhof, E. (1998). Conserved geometrical base-pairing patters in RNA. Q. Rev. Biophys. *31*, 399–455.

Li, W., and Frank, J. (2007). Transfer RNA in the hybrid P/E state: Correlating molecular dynamics simulations with cryo-EM data. Proc. Natl. Acad. Sci. USA *104*, 16540–16545.

MacKerell, A., Jr., Bashford, D., Bellott, M., Dunbrack, R.L., Jr., Evanseck, J., Field, M.J., Fischer, S., Gao, J., Guo, H., Ha, S., et al. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins. J. Phys. Chem. B *102*, 3586–3616.

Mitra, K., and Frank, J. (2006). Ribosome dynamics: insights from atomic structure modeling into cryo-electron microscopy maps. Annu. Rev. Biophys. Biomol. Struct. *35*, 299–317.

Mitra, K., Schaffitzel, C., Shaikh, T., Tama, F., Jenni, S., Brooks, C.L., Ban, N., and Frank, J. (2005). Structure of the *E. coli* protein-conducting channel bound to a translating ribosome. Nature *438*, 318–324.

Mitra, K., Schaffitzel, C., Fabiola, F., Chapman, M., Ban, N., and Frank, J. (2006). Elongation arrest by SecM via a cascade of ribosomal RNA rearrangements. Mol. Cell *22*, 533–543.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. J. Comput. Chem. *26*, 1781–1802.

Roseman, A.M. (2000). Docking structures of domains into maps from cryo-electron microscopy using local correlation. Acta Crystallogr. D Biol. Crystallogr. *56*, 1332–1340.

Rossmann, M.G. (2000). Fitting atomic models into electron-microscopy maps. Acta Crystallogr. D Biol. Crystallogr. *56*, 1341–1349.

Saibil, H.R. (2000). Macromolecular structure determination by cryo-electron microscopy. Acta Crystallogr. D Biol. Crystallogr. *56*, 1215–1222.

Sali, A., Glaeser, R., Earnest, T., and Baumeister, W. (2003). From words to literature in structural proteomics. Nature *422*, 216–225.

Schröder, G.F., Brunger, A.T., and Levitt, M. (2007). Combining efficient conformational sampling with a deformable elastic network model facilitates structure refinement at low resolution. Structure *15*, 1630–1641.

Schuwirth, B.S., Borovinskaya, M.A., Hau, C.W., Zhang, W., Vila-Sanjurjo, A., Holton, J.M., and Cate, J.H.D. (2005). Structures of the bacterial ribosome at 3.5 Å resolution. Science *310*, 827–834.

Selmer, M., Dunham, C.M., Murphy, F.V., IV, Weixlbaumer, A., Petry, S., Kelley, A.C., Weir, J.R., and Ramakrishnan, V. (2006). Structure of the 70S ribosome complexed with mRNA and tRNA. Science *313*, 1935–1942.

Shih, A.Y., Freddolino, P.L., Arkhipov, A., and Schulten, K. (2007). Assembly of lipoprotein particles revealed by coarse-grained molecular dynamics simulations. J. Struct. Biol. *157*, 579–592.

Sotomayor, M., and Schulten, K. (2007). Single-molecule experiments in vitro and in silico. Science *316*, 1144–1148.

Stewart, P.L., Fuller, S.D., and Burnett, R.M. (1993). Difference imaging of adenovirus: Bridging the resolution gap between X-ray crystallography and electron microscopy. EMBO J. *12*, 2589–2599.

Suhre, K., Navaza, J., and Sanejouand, Y.H. (2006). NORMA: a tool for flexible fitting of high-resolution protein structures into low-resolution electron-microscopy-derived density maps. Acta Crystallogr. D Biol. Crystallogr. *62*, 1098–1100.

Tama, F., Miyashita, O., and Brooks, C.L., III. (2004). Normal mode based flexible fitting of high-resolution structure into low-resolution experimental data from cryo-EM. J. Struct. Biol. *147*, 315–326.

Topf, M., Baker, M.L., Marc-Renom, M.A., Chiu, W., and Sali, A. (2006). Refinement of protein structures by iterative comparative modeling and cryoEM density fitting. J. Mol. Biol. *357*, 1655–1668.

Topf, M., Lasker, K., Webb, B., Wolfson, H., Chiu, W., and Sali, A. (2008). Protein structure fitting and refinement guided by cryo-EM density. Structure *16*, 295–307.

Valle, M., Zavialov, A., Li, W., Stagg, S.M., Sengupta, J., Nielsen, R.C., Nissen, P., Harvey, S.C., Ehrenberg, M., and Frank, J. (2003). Incorporation of aminoacyl-tRNA into the ribosome as seen by cryo-electron microscopy. Nat. Struct. Biol. *10*, 899–906.

Velazquez-Muriel, J.A., Valle, M., Santamaría-Pang, A., Kakadiaris, I.A., and Carazo, J.M. (2006). Flexible fitting in 3D-EM guided by the structural variability of protein superfamilies. Structure *14*, 1115–1126.

Volkmann, N., Hanein, D., Ouyang, G., Trybus, K.M., DeRosier, D.J., and Lowey, S. (2000). Evidence for cleft closure in actomyosin upon ADP release. Nat. Struct. Biol. *7*, 1147–1155.

Wells, D.B., Abramkina, V., and Aksimentiev, A. (2007). Exploring transmembrane transport through α-hemolysin with grid-steered molecular dynamics. J. Chem. Phys. *127*, 125101–125110.

Wriggers, W., and Birmanns, S. (2001). Using Situs for flexible and rigid-body fitting of multiresolution single-molecule data. J. Struct. Biol. *133*, 193–202.

Wriggers, W., and Chacón, P. (2001). Modeling tricks and fitting techniques for multiresolution structures. Structure *9*, 779–788.

Wriggers, W., Milligan, R.A., and McCammon, J.A. (1999). Situs: A package for docking crystal structures into low-resolution maps from electron microscopy. J. Struct. Biol. *125*, 185–195.

Wriggers, W., Agrawal, R.K., Drew, D.L., McCammon, A., and Frank, J. (2000). Domain motions of EF-G bound to the 70S ribosome: Insights from a handshaking between multi-resolution structures. Biophys. J. *79*, 1670–1678.

Yang, H., Jossinet, F., Leontis, N., Chen, L., Westbrook, J., Berman, H., and Westhof, E. (2003). Tools for the automatic identification and classification of RNA base pairs. Nucleic Acids Res. *31*, 3450–3460.

Zhang, Z., Shi, Y., and Liu, H. (2003). Molecular dynamics simulations of peptides and proteins with amplified collective motions. Biophys. J. *84*, 3583–3593.

Zhu, Y., Carrangher, B., Kriegman, D.J., Milligan, R.A., and Potter, C.S. (2001). Automated identification of filaments in cryoelectron microscopy images. J. Struct. Biol. *135*, 302–312.