

Trajectories of the ribosome as a Brownian nanomachine

Ali Dashti^{a,1}, Peter Schwander^{a,1}, Robert Langlois^b, Russell Fung^a, Wen Li^b, Ahmad Hosseinizadeh^a, Hsiao Y. Liao^b, Jesper Pallesen^{c,2}, Gyanesh Sharma^{b,3}, Vera A. Stupina^d, Anne E. Simon^d, Jonathan D. Dinman^d, Joachim Frank^{b,c,4}, and Abbas Ourmazd^{a,1,4}

^aDepartment of Physics, University of Wisconsin, Milwaukee, WI 53211; ^bDepartment of Biochemistry and Molecular Biophysics, and ^cHoward Hughes Medical Institute, Columbia University, New York, NY 10032; and ^dDepartment of Cell Biology and Molecular Genetics, University of Maryland, College Park, MD 20742

Contributed by Joachim Frank, October 8, 2014 (sent for review September 10, 2014)

A Brownian machine, a tiny device buffeted by the random motions of molecules in the environment, is capable of exploiting these thermal motions for many of the conformational changes in its work cycle. Such machines are now thought to be ubiquitous, with the ribosome, a molecular machine responsible for protein synthesis, increasingly regarded as prototypical. Here we present a new analytical approach capable of determining the free-energy landscape and the continuous trajectories of molecular machines from a large number of snapshots obtained by cryogenic electron microscopy. We demonstrate this approach in the context of experimental cryogenic electron microscope images of a large ensemble of nontranslating ribosomes purified from yeast cells. The free-energy landscape is seen to contain a closed path of low energy, along which the ribosome exhibits conformational changes known to be associated with the elongation cycle. Our approach allows model-free quantitative analysis of the degrees of freedom and the energy landscape underlying continuous conformational changes in nanomachines, including those important for biological function.

cryo-electron microscopy | elongation cycle | manifold embedding | nanomachines | translation

Ideally, one would like to “see” the conformational changes of a Brownian machine as it traverses its work cycle trajectory over the energy landscape. This information is particularly relevant for a biologically important molecular machine such as the ribosome, which is responsible for protein synthesis in all living cells. During the so-called elongation process, the ribosome repeatedly links an amino acid carried in by transfer RNA (tRNA) to the nascent polypeptide chain, with the choice of amino acid in each cycle dictated by the genetic message on the mRNA. In the eukaryotic ribosome, this process is facilitated by elongation factors eEF1A and eEF2, both GTPases.

It is believed that many intermediate conformational states must be involved in the elongation cycle of the ribosome (1), but the evidence is inferred, albeit from an impressive array of experimental techniques. Both cryogenic electron microscopy (cryo-EM) (2) and X-ray crystallographic approaches (3) have been used to determine the structures of several biochemically “trapped” states along the conformational trajectory. However, it has been pointed out that these likely represent only a fraction of the relevant conformational states, that each biochemically trapped state may correspond to more than one conformational state, and that the observed intermediate structures may be affected by the trapping process itself (1). Powerful algorithms (4, 5) have been used to sort cryo-EM snapshots into a small number of discrete classes, each presumed to represent an intermediate state (6). In some cases, however, snapshots of major ribosomal regions with large conformational flexibility have defied classification into discrete states altogether, even by the most advanced analytical methods (7). Single-molecule FRET experiments have yielded evidence for discrete conformational changes in single,

freely equilibrating pretranslocational ribosomes, and provided ensemble averages for such changes, but have been unable to provide data for short-lived intermediates (8, 9).

In a groundbreaking study, Fischer and coworkers (10) used cryo-EM to determine the structure and occupancy of different ribosomal conformational states as a function of time, obtaining the free-energy landscape through the Boltzmann factor (11). The specific process studied was that of back-translocation, a slow process (on the order of 30 min), in which the elongation cycle is partially reversed through interaction with another GTPase: LepA. Cryo-EM snapshots were classified in a hierarchical series of supervised (reference-based) steps to yield multiple conformations, differing mainly in specific preselected features: degree of intersubunit rotation, tRNA positions, and degree of head swivel of the small subunit—all changes known to be associated with the elongation work cycle of the ribosome.

Significance

Many functions in the cell are performed by Brownian machines, macromolecular assemblies that use energy from the thermal environment for many of the conformational changes involved in their work cycles. Here we present a new approach capable of mapping the continuous motions of such nanomachines along their trajectories in the free-energy landscape and demonstrate this capability in the context of experimental cryogenic electron microscope snapshots of the ribosome, the nanomachine responsible for protein synthesis in all living organisms. We believe our approach constitutes a universal platform for the analysis of free-energy landscapes and conformational motions of molecular nanomachines and their dependencies on temperature, buffer conditions, and regulatory factors.

Author contributions: A.D., P.S., and A.O. designed research; R.F., A.H., V.A.S., A.E.S., and J.D.D. contributed new reagents/analytic tools; A.D., P.S., R.L., W.L., H.Y.L., J.P., J.F., and A.O. analyzed data; J.F. and A.O. wrote the paper; A.D. and A.O. implemented algorithms; P.S. implemented geometry transformations and algorithms; R.L. helped prepare movies and geometry transformations; R.F. and A.H. provided NLSA and parameter-fitting codes and advice; W.L. interpreted movies; H.Y.L. and J.P. preprocessed ribosome data; G.S. performed EM experiments and particle verification; V.A.S. purified yeast ribosomes; A.E.S. and J.D.D. directed ribosome purification; and J.F. directed EM experiments and interpreted movies.

The authors declare no conflict of interest.

Freely available online through the PNAS open access option.

Data deposition: The 50 3D maps generated along the free-energy trajectory (clockwise in Fig. 3B) have been deposited in the Electron Microscopy Data Bank (accession no. EMD-6044).

¹A.D., P.S., and A.O. contributed equally to this work.

²Present address: 35 W. 127th St., Apt. 1, New York, NY 10027.

³Present address: Sir Mortimer B. Davis Jewish General Hospital, Montreal, QC, Canada H3T 1E2.

⁴To whom correspondence may be addressed. Email: jf2192@columbia.edu or ourmazd@uwm.edu.

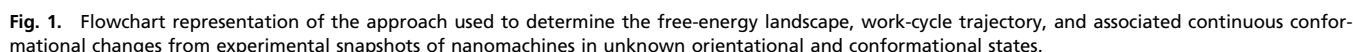
This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1419276111/-DCSupplemental.

Here, we present a new analytical approach capable of mapping continuous conformational changes of nanomachines along any trajectory in the energy landscape, without timing information, supervision, or templates. Our unbiased approach determines the number of degrees of freedom exercised during the observations, the energy landscape explored by the nanomachine, the energy trajectory traversed during the work cycle, and the continuous conformational changes associated with the work cycle. These novel capabilities constitute a powerful platform for quantitative study of the conformational and energy trajectories of nanomachines, including those engaged in a wide range of important biological processes. As reported below, the first application of this approach to an experimental cryo-EM dataset of ribosomes purified from yeast not engaged in translation reveals a closed trajectory of minimum energy along which the ribosomes appear to idle with conformational changes reminiscent of those observed during the protein elongation cycle.

A flow diagram is shown in Fig. 1, with a more technical description following the conceptual outline. Details are presented in the *SI Text*. As used, for example, in early approaches to image classification in electron microscopy (14), a snapshot can be represented as a vector in high-dimensional space by regarding the intensity value at each pixel of the snapshot as a component of the vector. Distance is a measure of similarity between the snapshots in this space. A collection of snapshots produces one or more data clouds, with discrete conformations producing separate clouds. We have previously shown that this approach is able to determine 3D structure from a collection of ultra-low signal 2D snapshots of unknown orientation (15–18), and, in the presence of orientational heterogeneity and defocus variations, distinguish

Here, we are concerned with continuous conformational changes, which produce correlations among the points representing the snapshots. These correlations define a hypersurface—a manifold—a concept found useful also in a previous approach to heterogeneity in cryo-EM (19). Using advanced machine-learning methods, we have demonstrated the ability to identify such manifolds in the presence of overwhelming noise and defocus variations (17).

Unfortunately, it is not possible to determine the conformational changes from such a description, because the local rates of change in multidimensional manifolds obtained from machine-learning techniques are, in general, unknown (21, 22) and cannot be easily related to the underlying changes in the system under observation (15, 23). To overcome this well-known difficulty, we introduce an additional step, in which the cloud of points is mapped to another coordinate system, where the local rates of change can be determined exactly, and related to the underlying conformational changes. This step leads to a representation of conformational change in terms of a universal parameter



(metric). The density of points in this space can now be related to the energy landscape sampled by the system through the Boltzmann factor $e^{-\Delta G/k_B T}$ (10), with ΔG denoting the change in the Gibbs free energy, k_B the Boltzmann constant, and T the temperature. The locus of minimum energy in this landscape represents the trajectory traversed by the machine during its work cycle. In each viewing direction, 2D movies can be compiled to reveal the conformational changes along the path of minimum energy or indeed any chosen trajectory. 3D movies can be compiled by stepping along such a trajectory and, in each step, performing a 3D reconstruction by integrating the information from many different viewing directions.

Analytical Procedure

The analysis is illustrated below in more technical terms with reference to a set of 849,914 experimental cryo-EM snapshots of 80S ribosomes from yeast, obtained in the course of a study of translational initiation by a plant virus (*SI Text*). The procedure of purification rendered the ribosome free of mRNA and most of the tRNAs.

Manifolds defined by data clouds are, in general, nonlinear: the data lie on intrinsically curved rather than flat hypersurfaces. Such manifolds can be identified and described (embedded) by graph-theoretic machine-learning techniques often used for dimensionality reduction. The so-called diffusion map embedding algorithm used in our work yields a description of a curved manifold in terms of the orthogonal eigenfunctions (more precisely eigenvectors) of known operators, specifically the Laplace–Beltrami operator (21, 22, 24) (*SI Text*). We use a specially developed kernel to deal with the substantial defocus variations encountered in cryo-EM data (17) (*SI Text*). In Fig. 2, we show representative cryo-EM snapshots (Fig. 2A) and a typical data manifold (Fig. 2B) described in terms of the first two eigenvectors of the Laplace–Beltrami operator. [In general, there is no simple correspondence between the eigenvectors obtained from this nonlinear analysis, and those obtained with linear approaches, such as principal component analysis (PCA) and singular value decomposition (SVD).] The manifold produced by the experimental snapshots is, in fact, nonlinear, with an intrinsic dimensionality of five (*SI Text*).

Fig. 2 describes the data in terms of the eigenfunctions of the Laplace–Beltrami operator with respect to an unknown metric (22). The absence of information on the metric precludes a consistent description of the conformational changes in terms of a known universal parameter. We solve this problem by mapping the manifold to another space, in which the eigenfunctions are known exactly. For this, we use an approach used in nonlinear Laplacian spectral analysis (NLSA) (25), a technique capable of performing SVD on nonlinear manifolds (nonlinear SVD for short) (*SI Text*). Briefly, one considers a collection of supervectors formed by concatenating snapshots falling within a window moving over the data vectors. The snapshots within each supervector are ordered according to the projections of the points representing them on a line through the origin, making an angle θ with, say, the horizontal axis (Fig. 2 and *SI Text*). By virtue of this projection, the arrangement picks out the conformational evolution along the selected line characterized by θ , with the random ordering of conformational changes along other directions assuming the character of noise. Nonlinear SVD (25) is then used to extract characteristic images (topos) and their evolutions (chronos) from these supervectors (*SI Text*). Each topo/chrono pair constitutes an element of a biorthogonal decomposition of the conformational changes along the selected line. Noise-reduced snapshots can be reconstructed from the topo/chrono pairs with significant (above-noise) singular values and embedded to obtain the manifold characteristic of the conformational changes along the selected line (*SI Text*). By construction, this manifold is one-dimensional, described in terms of known eigenfunctions, viz. $\cos(k\pi\tau)$, $k = 1, 2, 3, \dots$, and governed by the single parameter τ (*SI Text*).

Given a sufficiently dense collection of radial lines, each making an angle θ with the horizontal axis, the conformational changes can be described in any direction in the multidimensional space of conformations by the parameter $\tau(\theta)$. The mapping from a space of unknown metric and hence eigenfunctions to one characterized by known eigenfunctions governed by a single parameter $\tau(\theta)$ allows a consistent description of the conformational changes (*SI Text* and Fig. S1). This description of conformational changes in a given projection direction can be related to descriptions in other projection directions by assuming that the same conformational spectrum is viewed in all projection directions. In other words, the histograms of occupancy vs.

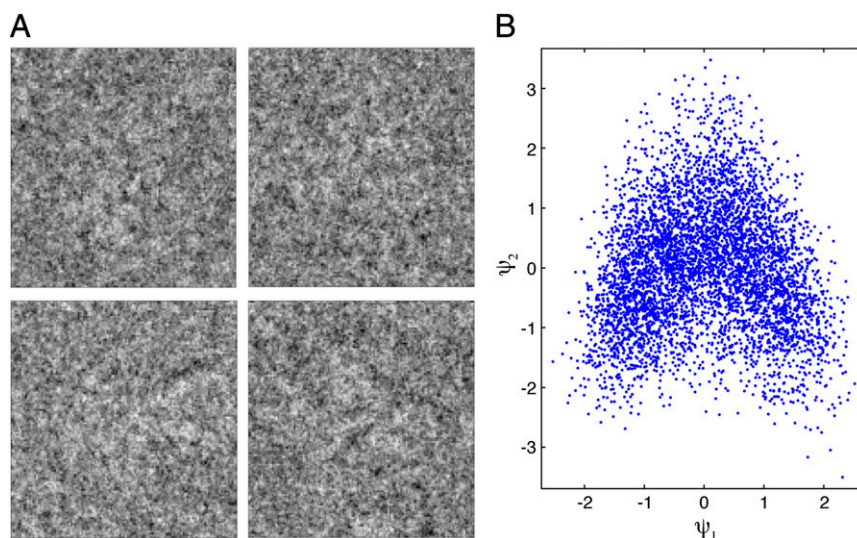


Fig. 2. (A) Representative cryo-EM snapshots, and (B) 2D view of a typical conformational manifold. The manifold is derived from a measure of similarity among $\sim 1,500$ cryo-EM snapshots of ribosome particles viewed within a tight orientational aperture. The axes ψ_1, ψ_2 represent the first two eigenvectors obtained by the diffusion map algorithm with a kernel able to deal with defocus variations.

conformational parameter $\tau(\theta)$ for different projection directions all represent the same spectrum of conformations and can thus be equalized. This equalization allows a universal description of the conformational spectrum across all projection directions. For any selected projected direction, the energy landscape along any line can be determined from the density of points along that line via the Boltzmann factor. Energy landscapes of higher dimension (with the dimensionality given by the number of eigenvectors) can be reconstructed by a tomographic extension of this approach (Fig. 3 and *SI Text*). Here, we concentrate on the first two eigenvectors, which contain some of the most interesting biological information (see below).

With this information in hand, the structural evolution of the system, including 2D movies of the conformational changes and the system's thermodynamic properties, can be quantitatively investigated in any viewing direction (*SI Text* and *Movie S1*). For any point on the energy landscape, a 3D structure map can be compiled by integrating the 2D information from many viewing directions into a 3D representation (Fig. 4) (*SI Text* and *Movies S2–S4*). Thus, 3D movies can be compiled by stepping along any selected trajectory in the energy landscape and reconstructing a 3D structure map in each step. The conformational changes along the closed minimum-energy trajectory of Fig. 3*B* are summarized in Fig. 3 and exemplified in Fig. 4. The associated movies (*Movies S2–S4*) present the ribosome as it evolves, with each movie corresponding to a different viewing direction selected for optimal visibility of the biologically relevant domain

motions (Fig. 3*A*). We note that, by construction, these movies are based on similarity rather than time. Accordingly, identical trajectories in opposite directions cannot be distinguished.

Results

The trough of minimum free energy is seen to form a closed, roughly triangular path (Fig. 3*B*), with variations of <0.9 kcal/mol ($1.5k_B T$ at room temperature) in the energy positions of the deepest points. The energy difference between the lowest and highest points of the entire landscape, corresponding to states with the highest and lowest nonzero occupancy in the experiment, amount to 3.8 ± 0.65 kcal/mol ($6.5 \pm 1.0k_B T$ at room temperature), with the uncertainty stemming chiefly from the low occupancy of high-energy states. The energy range covered nonetheless indicates that so-called transition states with energies several times the thermal energy have been probed in our experiment.

We examine the 3D movies showing the conformational changes along the closed triangular trajectory, analyzing in detail seven structure maps along the way (*SI Text* and *Movie S5*). Using rigid-body fitting of domains while making use of a published cryo-EM map of the yeast ribosome (26) for reference, we observe combinations of four motions previously described in the literature and associated with the elongation cycle, as follows: (i) ratchet-like intersubunit rotation, a counterclockwise rotation of the small subunit about an axis normal to the subunit interface (Fig. 3*A*, solvent view) (27, 28); (ii) rotation (closing movement) of

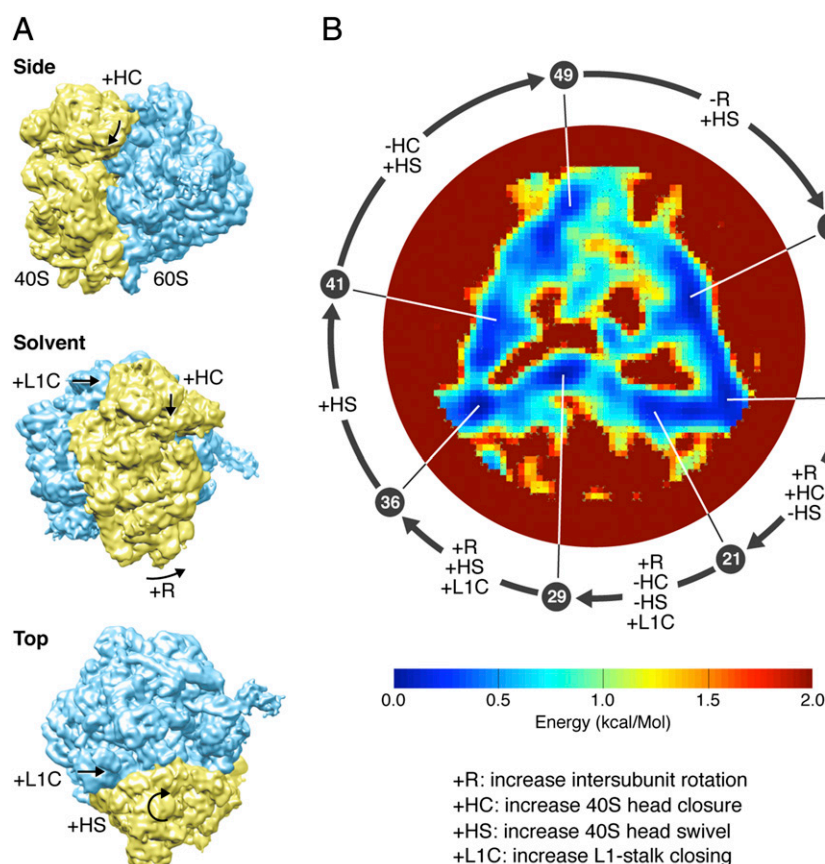


Fig. 3. (A) Three views of a cryo-EM map of the 80S ribosome from yeast (32), with arrows indicating four key conformational changes associated with the elongation work cycle of the ribosome. (B) The energy landscape traversed by the ribosome. The color bar shows the energy scale. The energy range has been truncated at 2 kcal/mol to show details of the triangular trough. The error in energy determination along the closed triangle is 0.05 kcal/mol. The roughly triangular minimum free-energy trajectory is divided into 50 states. The arrows indicate the structural changes between 7 selected states, each identified by its place in the sequence of 50 states.

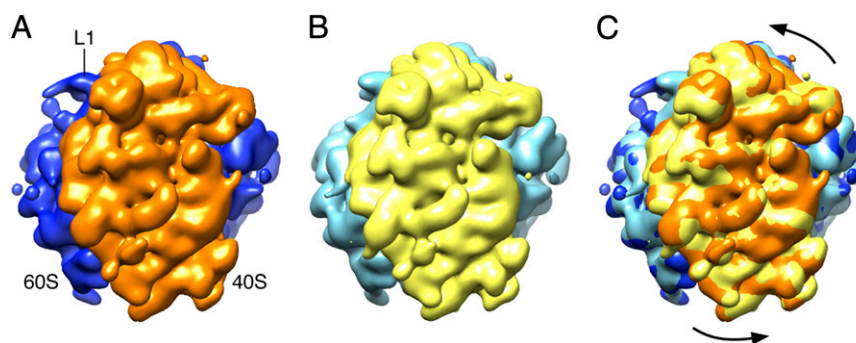


Fig. 4. Example of conformational changes along the trajectory: ratchet-like motion. (A) Unrotated ribosome, map 14 in Fig. 3*B*. (B) Maximally rotated ribosome, map 36. (C) Superposition of the two maps. The full set of frames showing continuous conformational changes is shown in three common viewing directions in Movies S2–S4.

the L1 stalk toward the intersubunit space (Fig. 3*A*, solvent and top views) (29); (iii) small subunit head swivel, a rotation of the subunit head about its long axis (Fig. 3*A*, top view) (30); and (iv) small subunit head closure, a “nodding” motion of the head (Fig. 3*A*, side and solvent views) (31). The first three motions are correlated and known to be associated with the process of mRNA–tRNA translocation (30, 32), whereas the fourth is known to accompany the selection of cognate aminoacyl tRNA during the process of decoding (31).

Proceeding clockwise along the minimum free-energy pathway, these motions occur in the sequence encountered during the elongation cycle: starting at the apex of the triangle and moving to the right side of the base (Fig. 3*B*), we see the small subunit head closure as a prominent motion, corresponding to the tRNA selection step. Next, proceeding along the base of the triangle, three coordinated movements of small subunit body, head swivel, and L1 stalk occur—all expected during translocation—while the small subunit head remains in the closed position. Finally, moving up from the left side of the base to the apex, we observe the reversal of head closing (the conformational change associated with the disengagement of the mRNA from the decoding center), as well as the reversal of the intersubunit motion.

Discussion

The energy landscape constructed by the analysis of the ribosome ensemble shows multiple interconnected paths of relatively low energy. Of these, a path of roughly triangular shape clearly stands out, encompassing roughly one third of the molecules. Comparison of 3D reconstructions from data at selected points along this path reveals conformational changes previously observed in the elongation cycle of fully programmed, translating ribosomes. Because the ribosomes captured in our snapshots were not engaged in protein synthesis, lacking mRNA and aminoacylated tRNAs, we must assume that their idle motions in the thermal environment sample the conformational space permitted by the degrees of freedom, thus exhibiting the conformational changes that would be productive in the presence of the ligands of the translational machinery (mRNA, aa-tRNA, eEF2, and eEF1A). Indeed, idling of the pretranslocational ribosome in the absence of elongation factors along the direction of the most prominent conformational changes (intersubunit motion and opening/closing of L1 stalk) has been previously observed by single-molecule FRET (8, 9) and cryo-EM (33). Idling of the empty ribosome along the same path has also been inferred from a series of X-ray structures (34). It remains to be established by the approach outlined in this paper whether fully programmed ribosomes follow the same path identified here. As to the other paths traversing the free-energy landscape (Fig. 3*B*), it is tempting to speculate that

they may represent alternative routes of the molecular machine under different buffer and temperature conditions.

We now discuss the salient features of our approach. First, the usefulness of movies of Brownian machines has been rightly questioned, because each trajectory of a single machine is strongly influenced by stochastic factors and thus is unique (1). The movies presented here, however, integrate information from a large ensemble of snapshots, each stemming from an object viewed only once. By using manifolds to capture the properties of the entire dataset, and nonlinear singular value analysis to suppress noise, our approach offers an efficient means for extracting the ensemble kinetics, i.e., the information common to a collection of objects, each viewed in an initially unknown orientational and conformational state.

Most successful experimental studies of the conformational spectra of biological machines have been hitherto restricted to sorting snapshots into a small number of classes (10, 11), using templates in some form, relying on timing information, or a combination of these tools (10). In contrast, our approach naturally yields detailed conformational trajectories and energy landscapes without a priori information or assumptions and at moderate computational expense (*SI Text*).

Of course, the maximum number of detectable conformational states is limited by various factors. These factors include the frequency (via the Shannon–Nyquist theorem) with which the conformational spectrum has been sampled, the information content of the individual snapshots, and, ultimately, the atomic nature of the object itself. In our analysis, the conformational signal stems primarily from the most heavily sampled projection directions, in each of which the conformational spectrum is randomly sampled by up to 5,000 snapshots. The number of meaningfully distinct conformational states is governed by the signal-to-noise ratio, which determines the statistical confidence with which neighboring states can be distinguished. In this study, the requirement for neighboring conformational states to be separated by 3 SDs (3σ) means that ~50 conformational states can be distinguished (*SI Text*). This number is about an order of magnitude larger than previously achieved without timing information or templates. In combination with the recent availability of large cryo-EM datasets with near-atomic resolution (7, 35), our approach promises the possibility to extract conformational information with unprecedented detail.

In recent years, increasing computing power has fueled efforts to simulate the conformational trajectories of molecular machines, particularly the ribosome (36), by molecular dynamics. Experimental results obtained with the new approach presented here offer the promise to guide these efforts and provide the means for verifying important modeling assumptions.

