



Amortized template matching of molecular conformations from cryoelectron microscopy images using simulation-based inference

Lars Dingeldein^{ab}, David Silva-Sánchez^c, Luke Evans^d , Edoardo D'Imprima^e , Nikolaus Grigorieff^{ef} , Roberto Covino^{b,h,1} , and Pilar Cossio^{d,i,1} 

Affiliations are included on p. 9.

Edited by Andrej Sali, University of California, San Francisco, CA; received October 1, 2024; accepted April 30, 2025

Characterizing the conformational ensemble of biomolecular systems is key to understand their functions. Cryoelectron microscopy (cryo-EM) captures two-dimensional snapshots of biomolecular ensembles, giving in principle access to thermodynamics. However, these images are very noisy and show projections of the molecule in unknown orientations, making it very difficult to identify the biomolecule's conformation in each individual image. Here, we introduce cryo-EM simulation-based inference (cryoSBI) to infer the conformations of biomolecules and the uncertainties associated with the inference from individual cryo-EM images. CryoSBI builds on simulation-based inference, a merger of physics-based simulations and probabilistic deep learning, allowing us to use Bayesian inference even when likelihoods are too expensive to calculate. We begin with an ensemble of conformations, templates from experiments, and molecular modeling, serving as structural hypotheses. We train a neural network approximating the Bayesian posterior using simulated images from these templates and then use it to accurately infer the conformation of the biomolecule from each experimental image. Training is only done once on simulations, and after that, it takes just a few milliseconds to make inference on an image, making cryoSBI suitable for arbitrarily large datasets and direct analysis on micrographs. CryoSBI eliminates the need to estimate particle pose and imaging parameters, significantly enhancing the computational speed compared to explicit likelihood methods. Importantly, we obtain interpretable machine learning models by integrating physics-based approaches with deep neural networks, ensuring that our results are transparent and reliable. We illustrate and benchmark cryoSBI on synthetic data and showcase its promise on experimental single-particle cryo-EM data.

cryo-EM | biophysics | Bayesian inference | modeling | template matching

Biomolecules continuously reorganize between alternative conformations to perform essential functions in the cell. Understanding the mechanisms of these biological functions requires knowing the structure of the various conformations and the dynamics of how one conformation reorganizes into another. Characterizing in experiments the collection of all conformations—the conformational ensemble—is challenging. Most techniques provide only ensemble averages, while single-molecule methods usually lack the structural resolution to characterize different conformations precisely. While molecular simulations provide trajectories at high temporal and spatial resolution, sampling and accuracy issues limit their ability to explore conformational ensembles. Integrative methods in structural biology combine experimental and computational techniques with the promise to describe a complete picture of biomolecular structural dynamics (1).

Cryoelectron microscopy (cryo-EM) is an experimental technique that captures different conformations of a biomolecule at the single-molecule level. In cryo-EM, a transmission electron microscope records two-dimensional projected images of a thin sample containing many identical copies of the same molecule (micrographs). Particle-picking software then identifies the two-dimensional images displaying a single copy of the molecule (particle) (2). The sample is prepared by flash-freezing an aqueous solution of randomly oriented biomolecules. Since freezing is very fast (3–5), the biomolecules are trapped in different conformations. In other words, cryo-EM provides snapshots (samples) from the entire conformational probability distribution.

In practice, reconstructing sparsely populated conformations with cryo-EM remains an outstanding challenge. The frozen sample is imaged with a limited number of

Significance

Biomolecules reorganize between alternative conformations to perform their functions. Mapping out all their different conformations is critical to understanding how they work. Cryoelectron microscopy provides 2-d snapshots of copies of a biomolecule, each in a possible conformation. If we could assign each image to a specific conformation, we could identify and count important conformations. However, assigning conformations to images is challenging because the images are noisy and show projections along unknown orientations. We developed a computational method combining Bayesian inference and machine learning to assign a biomolecule's conformation in individual cryo-EM images. Our method provides not only a most probable assignment but also an associated statistical confidence. It is very fast, enabling the analysis of massively large datasets.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2025 the Author(s). Published by PNAS. This article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](https://creativecommons.org/licenses/by-nc-nd/4.0/).

¹To whom correspondence may be addressed. Email: covino@fias.uni-frankfurt.de or pcossio@flatironinstitute.org.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2420158122/-/DCSupplemental>.

Published June 4, 2025.

electrons to avoid radiation damage, resulting in picked particles that are very noisy projections of the molecular density in unknown conformations and orientations. Obtaining high-resolution structures requires averaging over many particles to reduce the noise. Consequently, 3-d classification methods can reconstruct a limited number of different conformations (6–9), which can be combined with molecular dynamics (MD) approaches to estimate a conformational ensemble (10–13). However, these methods require grouping particles into a relatively small set of classes. Each class must contain enough particles necessary for a high-resolution reconstruction. Particles belonging to scarcely populated classes are discarded. Thus, current 3D classification methods may result in an incomplete understanding of a molecule's biological function (14). Rare conformations and transition states that occur infrequently will be missed in these analyses, which are also unsuitable for studying highly flexible biomolecules; additionally, class refinement often fails to converge, with repeated analyses yielding different particle classifications depending on the random seed used (15), and these methods depend on averaging, preventing the identification of conformations at the level of individual images.

Machine learning (ML) enabled an important step toward extracting heterogeneous reconstructions from cryo-EM, from pioneering approaches using manifold embedding (16) to state-of-the-art deep generative models (17–21). The key idea is to use ML to learn the mapping between the particles in a cryo-EM dataset and the corresponding conformational volumes (22). However, the statistical inference can become computationally intractable, especially when both the conformation and pose (projection direction and location) must be inferred simultaneously. This is why most cryo-EM ML methods, as well as non-ML variability methods (23–26), rely on “consensus” maps and costly explicit-likelihood methods to calculate the particle poses, which fail for highly flexible molecules. CryoAI (27) and its implementation in cryoDRGN (28) use direct gradient-based optimization to amortize the particle poses while still requiring the direct estimation of a pose for each particle. Additionally, these methods do not provide statistical errors when assigning each particle to a specific conformational state in the latent space, and the latent space itself lacks physical interpretability (29, 30).

Template-matching-based approaches assign each single particle to a molecular structure with high fidelity. BioEM can discriminate molecular conformations in individual particles by integrating over poses and imaging parameters within a Bayesian framework (31, 32). Recently, high-resolution template matching identified biomolecular conformations in situ using a cross-correlation-based approach in 2-d (33, 34) and 3-d (35). When the poses are sampled on a fine angular grid, and the matching is repeated for templates representing different molecular conformations, this method provides a highly accurate metric for conformational identification. However, this brute-force likelihood optimization rapidly becomes computationally prohibitive. Currently, there is no fast and tractable method for identifying molecular structures in individual cryo-EM images that also provides assignment uncertainties.

Here, we leveraged recent advances in simulation-based inference (36, 37) to develop cryo-EM simulation-based inference (cryoSBI), a computational framework that uses Bayesian inference to assign a molecular conformation to an individual cryo-EM image avoiding expensive pose searches or costly likelihood optimizations. Our framework provides amortized inference and accurate statistical confidence. Standard cryo-EM reconstruction techniques combined with molecular modeling such as advanced MD schemes (38) and AI-based structure-prediction

tools (39, 40) provide a set of structural templates that serve as structural hypotheses. We then simulate cryo-EM experiments to produce particles that answer the question: “What would an experimental particle look like if it depicted this molecular conformation?” The synthetic particles arise from random poses and imaging parameters with adequate noise levels. Neural network density estimation allows us to learn from these particles the Bayesian posterior. We illustrate the cryoSBI algorithm, benchmark it on synthetic data, and showcase its use on experimental data.

CryoSBI has several key advantages: *i*) it provides not only the most probable assignment but also a statistical confidence; *ii*) the inference is amortized, making it exceptionally fast and enabling analysis of an arbitrarily large number of images; *iii*) training ML models on simulated particles—on which we have full control—guarantees that the ML embedding is robust and interpretable. CryoSBI solves the problem of identifying molecular conformations in individual cryo-EM images, paving the way for learning the entire conformational ensemble, including rare and transient conformations, directly from experiments.

Results

Simulation-Based Inference of Single-Particle Cryo-EM. We formulate the task of inferring molecular conformations from a single-particle cryo-EM image as a Bayesian inference problem. In essence, we want to quantify the probability that a given image I depicts a molecular conformation X . Let us consider a set of molecular conformations, that is, a set of structures, and a vector θ that parameterizes them. In general, θ is simply the index of each structure in the set. In specific situations, it is possible instead to map the structures to a low-dimensional parameterization describing a conformational change. For instance, we could describe the movement between two protein domains with an angle or a distance. For clarity, we will illustrate cryoSBI using structural ensembles that represent a conformational change that can be described by a one-dimensional θ . However, the cryoSBI algorithm does not require such a low-dimensional parameterization to be formulated or to be effective.

Given a cryo-EM image I , we aim to infer the conformation θ of the molecule observed in the image. Therefore, we want to compute the Bayesian posterior $p(\theta|I)$, quantifying how compatible θ (that is associated with conformation X) is with the observed image I . The posterior can be computed by Bayes' theorem

$$p(\theta|I) \propto p(I|\theta)p(\theta), \quad [1]$$

where $p(I|\theta)$ is the likelihood of generating an image I given a molecular conformation θ and the prior $p(\theta)$ encodes all available knowledge before making the inference on how the conformations are distributed. Modeling the image formation process described by the likelihood requires taking into account the details of the experiment and the function of the electron microscope. For instance, it is often assumed that the molecule is randomly rotated, that the images are noisy, and that the microscope will introduce aberrations. The full likelihood of the image formation process is, therefore, $p(I|\theta, \phi)$, containing additional parameters ϕ required to model the details of the experiments and microscope. Usually, we are not interested in making inference of ϕ , treated as nuisance parameters. In other words, the likelihood is a marginalization

$$p(I|\theta) = \int p(I|\theta, \phi)p(\phi)d\phi, \quad [2]$$

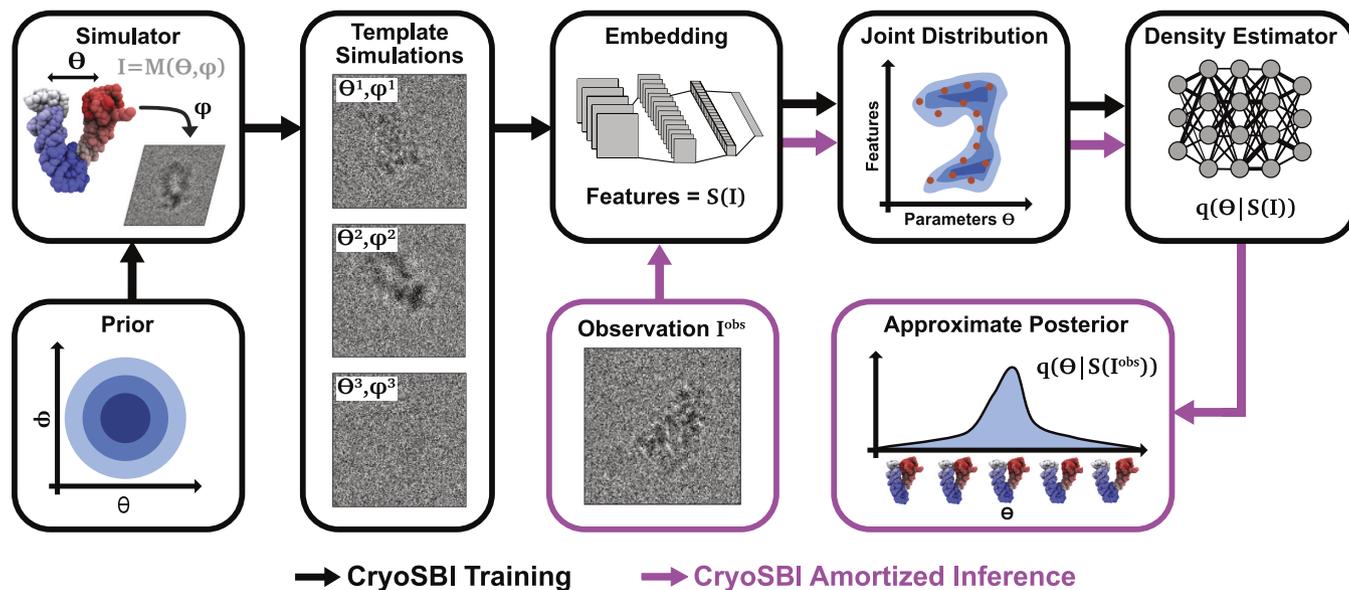


Fig. 1. Schematic representation of cryoSBI: Simulation-based inference for cryo-EM. Simulated particle images generated from template conformations θ and imaging parameters ϕ , such as rotation, defocus, and translation, are used to train the cryoSBI embedding network and posterior estimator (black boxes). θ and ϕ are sampled from a prior distribution, and fed to a forward model M to generate a particle image I . An embedding network that featuresizes the images $S(I)$ and a density estimator of the joint distribution of features and parameters are trained simultaneously using millions of simulated particle images. CryoSBI learns a computationally efficient approximation of the posterior $p(\theta|S(I))$, which takes an image as input to predict the posterior distribution over θ , enabling amortized template matching of experimental images (magenta boxes). For each experimental particle, cryoSBI provides the full posterior, indicating both the most probable conformation that generated the image and the statistical confidence of the inference.

with the prior of the imaging parameters $p(\phi)$. Maximizing the marginal likelihood $p(I|\theta)$ is computationally very expensive (32, 33) because each evaluation requires integrating over all imaging parameters. The high computational cost of approaches that require explicit likelihood optimization restricts their applicability to small sets of images.

Simulation-based inference (SBI) is an alternative approach to do Bayesian inference with intractable or computationally expensive likelihoods (36, 37, 41–44). The main idea is to replace a likelihood evaluation with a forward model to simulate synthetic data and then learn an approximation of the posterior on them. Here, we develop an SBI framework for amortized template matching of conformations from single-particle cryo-EM images (Fig. 1). While the inverse problem of inferring a conformation from a cryo-EM image I is challenging, the forward problem is much simpler. We can easily encode the image formation process described by the likelihood $p(I|\theta, \phi)$ in a cryo-EM simulator, and repeatedly sample it by running forward simulations to produce synthetic images, i.e., $I_i \sim p(I|\theta_i, \phi_i)$ with $\theta_i \sim p(\theta)$, $\phi_i \sim p(\phi)$. In this way, we accumulate a dataset of simulated images and parameters describing their generation, $\mathcal{D} = \{\theta_i, \phi_i, I_i\}$. Here, we used Neural Posterior Estimation (45) with a normalizing flow to directly approximate the Bayesian posterior from \mathcal{D} . We used an embedding network $S(I)$ to extract features and map each image into a medium-dimensional representation. We then used another neural network q as a conditional density estimator to build a surrogate model of the posterior, i.e., a statistical model that approximates the posterior, $q(\theta|S(I)) \approx p(\theta|I)$. We then trained S and q jointly on \mathcal{D} using standard supervised deep learning methods. In this way, we learned an approximation of the desired Bayesian posterior, bypassing any explicit likelihood evaluation and marginalization.

After training on the simulated images, the neural density estimator q estimates the posterior for any new experimental image. The inference is computationally efficient for two reasons:

First, it does not require any marginalization over the nuisance parameters ϕ ; and second, the inference is amortized. The computationally expensive part due to the sampling, must be paid only once upfront by repeatedly running the simulator to build the dataset \mathcal{D} . Once the conditional estimator q is trained, any new inference requires only an evaluation of the neural network underlying q .

In summary, cryoSBI involves the following steps: *i)* Prepare a set of experimental cryo-EM particles or micrographs. *ii)* Obtain a set of molecular conformations that serve as structural hypotheses—template structures—by using molecular simulations or ML methods. *iii)* Simulate many synthetic particles sampling all possible template structures and nuisance parameters. *iv)* Obtain a surrogate of the Bayesian posterior by training the embedding and conditional density estimator simultaneously on the set of simulated template particles. *v)* Perform inference on the experimental particles with the trained surrogate posterior.

Validation and Benchmark with Synthetic Data. How precisely is it possible to identify a structure in a single cryo-EM image? We answered this question by validating and benchmarking cryoSBI using synthetic data obtained from hsp90, an established benchmark model in the field (46, 47). Hsp90 comprises two chains that perform a large conformational change corresponding to their opening and closing. We selected twenty structures spanning the opening of the one of the chains, measured by the rmsd with respect to the closed structure (Fig. 2A).

CryoSBI accurately infers molecular configurations from single images. We trained cryoSBI on synthetic cryo-EM particle images generated under realistic conditions, with random orientations, a wide range of defocus values, center translations, and signal-to-noise ratios (SNRs). We learned a surrogate model of the posterior, that we used to make inference on the synthetic particles. For each inference, we obtained an estimate of the Bayesian posterior that we could compare with the structure

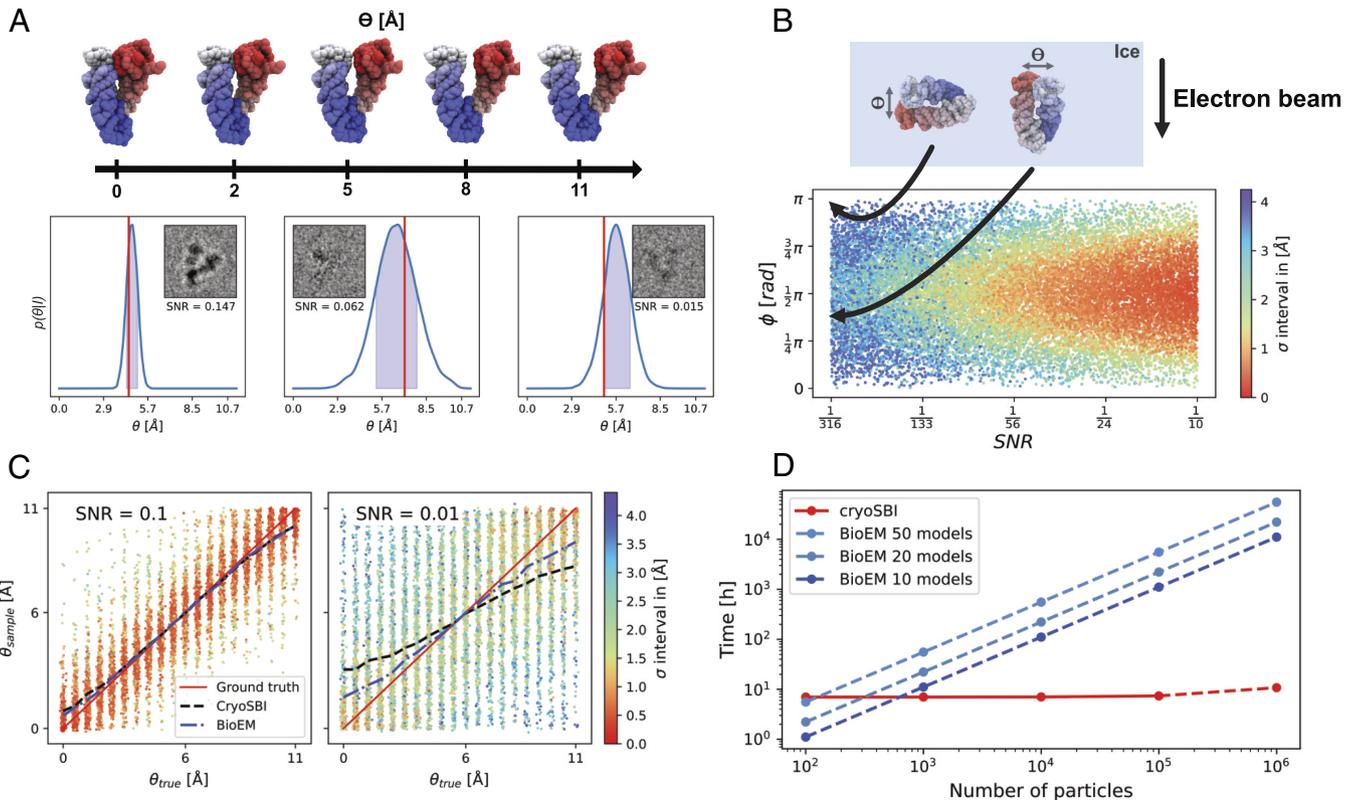


Fig. 2. Validation and benchmark with synthetic data. (A) The opening of the two arms of hsp90 defines a conformational change, quantified by θ , the rmsd w.r.t the closed conformation. We selected 20 configurations equally spaced along θ . Examples of cryoSBI inference for three particles. Each inference on synthetic images (gray-scale images in the *Insets*) resulted in a posterior (blue curve), quantifying which structural model was the most compatible with the image. Red lines represent the true configurations. The shaded region is the 2σ interval. (B) Inference precision for 10,000 images as a function of SNR and projection direction. Each point represents the σ interval (color bar, from dark blue—low confidence, to deep red—high confidence) of the posterior obtained from a synthetic image produced at a given SNR and projection angle. On *Top*, we show the schematic definition of ϕ , the angle formed between the direction of the electron beam and the direction of movement of the two arms of hsp90. (C) Inference accuracy for 10,000 images for sets with different SNR. The two scatter plots show the correlation between the estimated opening obtained as a sample of the posterior, and the true opening, using SNR = 0.1 and SNR = 0.01, respectively. We colored each sample according to σ (colors as for panel B). The black dashed line describes the average of the posterior means, while the blue line corresponds to the mean of the maximum-likelihood estimates using BioEM (32). (D) Scaling of the computational cost w.r.t. the number of images of cryoSBI and maximum likelihood methods (BioEM). Wall-times were obtained using an NVIDIA RTX A6000 GPU for cryoSBI, and an AMD EPYC 7742 CPU for BioEM.

that we actually used to produce the specific image (red line in Fig. 2A). While we focused here for clarity on a one-dimensional posterior, cryoSBI is general and can be used to inspect higher-dimensional posterior distributions (*SI Appendix, Fig. S1*). The posterior provides information both on the accuracy of the inference—whether the bulk of the distribution contains the ground truth—and the precision—the spread of the distribution. The precision estimate is accurate, as shown by the convergence of training loss and simulation-based calibration checks (*SI Appendix, Fig. S2*).

The SNR and projection direction are the main experimental factors determining how precisely we can infer a molecular configuration from a single image. As the SNR of an image decreases, the inference is still accurate but the posterior gradually broadens, corresponding to an increasing uncertainty (Fig. 2A and B). The precision also decreases for projection directions that occlude the conformational change of interest. In the case of hsp90, this occurs by projecting along a direction where one arm covers the other one (Fig. 2B), where the projection direction is parallel to the relevant conformational motion. For very low SNR and bad projection directions, the inference returns an approximately flat posterior. In other words, cryoSBI correctly tells us that these images cannot be reliably assigned to specific conformations.

A more systematic evaluation confirms that cryoSBI is accurate and precise for SNRs within experimental range (46). We compared inferred configurations to ground truths for 10,000 images, assuming high and low SNR (Fig. 2C). For SNR = 0.1, 68% of inferred configurations were accurate within 1 Å with an average uncertainty of 0.75 Å measured by the posterior σ interval (color bar). For SNR = 0.01, the accuracy of the prediction declined slightly to 68% of the inferred structures being within 2.7 Å of the true structure. The average uncertainties rose to 2.1 Å. In comparison to an explicit likelihood method (dashed lines), cryoSBI's predictions are slightly worse for low SNR. This is expected, as SBI methods tend to lose some accuracy due to the approximations of the ML models that make amortization possible, a phenomenon also observed in 3D reconstruction (48). We note that for flat posteriors, the mean is biased toward the center due to cryoSBI being trained on a finite domain of θ , therefore, we used samples of the posterior that do not exhibit this issue.

CryoSBI is fast and enables inference of very large sets of images and molecular structures. We compared the computational cost of performing inference on the synthetic dataset with cryoSBI and methods that optimize an explicit likelihood model (31, 32). These require to evaluate the likelihood and marginalize over all model parameters for each image, leading to a linear scaling

of the computational cost with the number of images and the number of conformations (Fig. 2D). The cost quickly becomes prohibitive, particularly for high-resolution template matching that may involve evaluating millions of orientations (33). CryoSBI's inference is instead amortized, that is, the largest computational cost occurs upfront to produce simulations and train the model (approximately a few hours), after which the inference is effectively free. Each inference only requires a forward pass of the trained neural network that serves as a surrogate model of the posterior. Amortization opens the door to comparing thousands of structures to datasets containing an arbitrarily large number of images.

Validating with Experimental Data. Having validated cryoSBI on synthetic data, we sought to demonstrate that it correctly maps individual experimental particles of apoferritin to their corresponding reference structure. So far, the training templates and the synthetic data have been generated with the same forward model simulator and parameter distributions. This will not be the case with experimental data, for which we must generally assume model misspecification. In other words, the data we simulate to train the model and the experimental particles that we want to make inferences on will always differ. However, we cannot access a ground truth with experimental particles and therefore first validate our method using a standard experimental benchmark system: apoferritin.

Apoferritin is a 474 kDa large cytosolic globular protein complex composed of 24 subunits forming a hollow nanocage (Fig. 3A). It is highly symmetric and rigid, making it a standard benchmark in the cryo-EM field. We used a published dataset containing 483 particles of apoferritin. Due to the absence of conformational flexibility, it is likely that the particles do not include alternative conformations. We selected the PDB structure built from the cryo-EM map reconstructed from the same dataset (49) as our ground truth (Fig. 3A). We generated a hypothetical conformational ensemble by varying our ground truth structure along two normal modes (Fig. 3A). The order parameter is

$\theta = \gamma \text{rmsd}$ to ground truth, where $\gamma = -1$ for normal mode 1 and $\gamma = 1$ for normal mode 10, which quantifies the distance of the resulting structures from the ground truth reference that sits at $\theta = 0$ by construction. We selected normal modes 1 and 10 to ensure two distinct conformational changes, thereby avoiding degeneracies caused by symmetry.

CryoSBI correctly maps individual experimental particles of apoferritin to their corresponding reference structure. Following our pipeline, we trained a surrogate posterior on synthetic templates generated starting from the structural ensemble shown in Fig. 3A. Posteriors are peaked around $\theta = 0$, both for individual particles (Fig. 3B) and for the entire experimental dataset (Fig. 3C), indicating that we could identify the 3D structure corresponding to the individual particle accurately.

The posterior width shows that the mapping is also quite precise, with uncertainty in the order of a few Angstroms. Since we are analyzing single-molecule data (snapshots of a single protein in a specific conformation) and not averaged observables (e.g., 3D maps), we should not expect that every posterior conditioned on each image is sharply peaked precisely at the reference structure. Some particles will be more informative than others. Building a histogram by resampling each posterior conditioned on each image provides a statistical view of how informative the particles in the dataset are. Interestingly, the histogram is shaped like a funnel. Particles whose posterior samples are closest to the cryo-EM structure (i.e., centered around $\theta = 0$) are those that we can map with the highest confidence (Fig. 3D).

A Challenging Experimental Dataset. Next, we challenged cryoSBI using an experimental dataset containing particles of hemagglutinin, a homotrimeric protein complex found on the surface of influenza viruses (50). Compared to apoferritin, the hemagglutinin dataset presents several additional challenges. Hemagglutinin is more dynamic, and the particles capture a much more heterogeneous structural ensemble. In fact, only around 47% of the particles led to the reconstruction of the

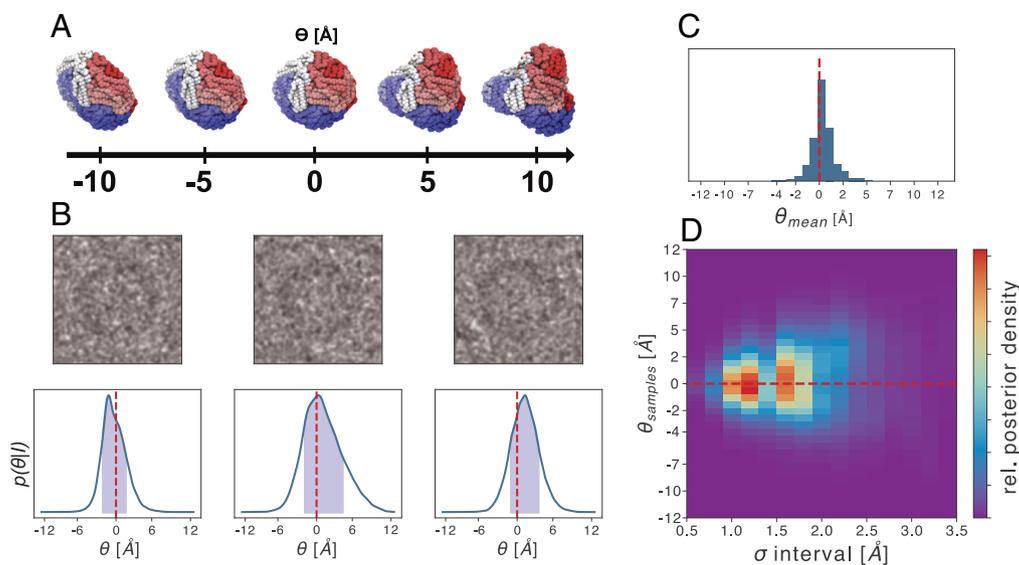


Fig. 3. CryoSBI applied to experimental particles of apoferritin. (A) Apoferritin conformational change along θ , generated by varying the ground truth reference cryo-EM structure (PDB ID 4v1w) along two normal mode directions. The conformational change is quantified by the rmsd in Å to the reference cryo-EM structure. We multiplied the rmsd by -1 for variations along mode 1 (Left half of the axis). By construction $\theta = 0$ Å for the reference structure. (B) Example of cryoSBI posteriors for experimental apoferritin particle images. The particles are low pass filtered for better visibility. The red dashed line indicates the position of the reference structure along θ . (C) Histogram of the inferred conformation from the posterior mean θ_{mean} for the particles in the dataset. (D) Two-dimensional histogram reporting 1,000 posterior samples from the posterior of each apoferritin.

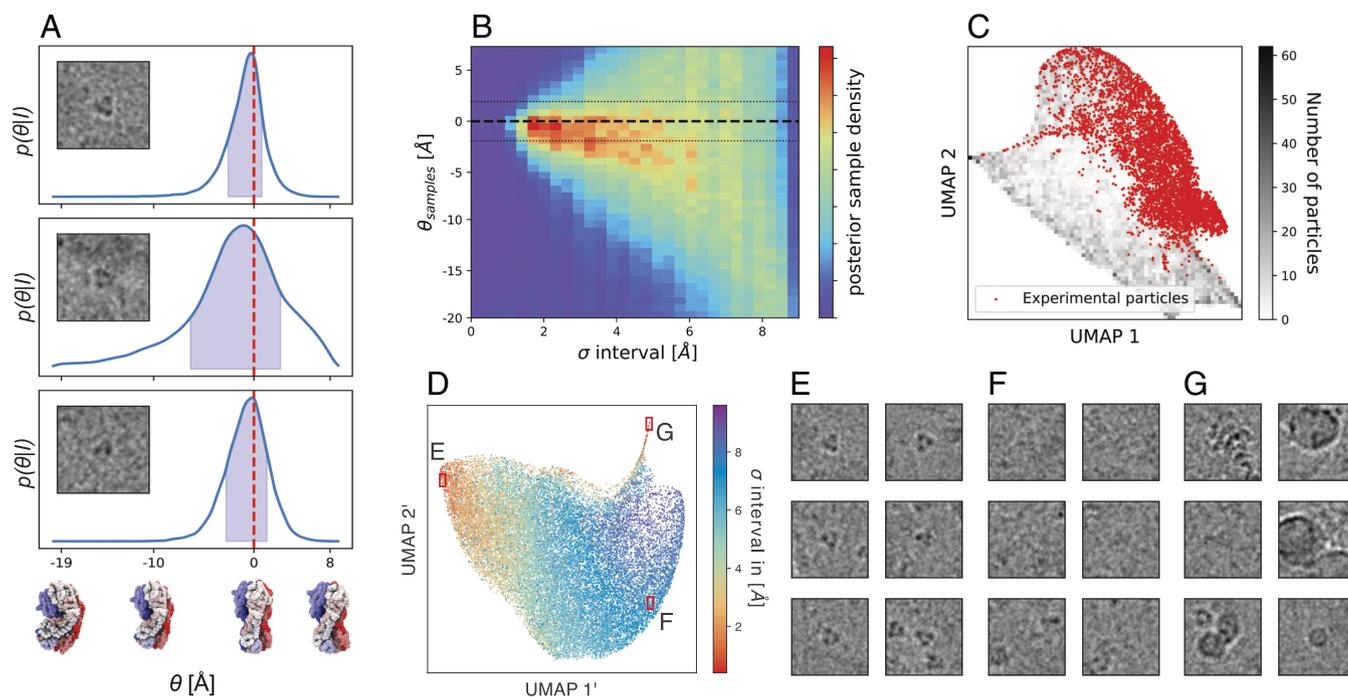


Fig. 4. CryoSBI applied to experimental particles of hemagglutinin. (A) Example of cryoSBI posteriors for experimental hemagglutinin particle images (low pass filtered for better visibility). The conformational change along θ was modeled normal mode analysis, $\theta = \gamma \text{rmsd}$ from reference cryo-EM structure (red dashed line) where $\gamma = -1$ or 1 for normal mode 1 and 2, respectively. (B) Two-dimensional histogram showing 1,000 posterior samples for each of the 271,558 posteriors. The black dashed line indicates the cryo-EM reference structure. The dotted lines show the size of the expected fluctuation around the reference structure estimated with MD simulations. (C) UMAP 2D projection of the latent space of the simulated particles used for training (gray) and a random subset of 10% of experimental particles (red). (D) UMAP 2D projection of the latent representation of a random subset of 10% of experimental particles colored according to their posterior σ interval. (E, F, and G) Experimental particles, selected from different positions in the UMAP in (D) (low pass filtered for better visibility).

published high-resolution structure. Additionally, the protein adopted a preferred orientation in the experimental sample, leading to a particle distribution that does not cover uniformly the entire space of possible orientations. This can be problematic for mapping two-dimensional projections into a three-dimensional structure.

Despite these challenges, cryoSBI could correctly identify hemagglutinin configurations in single particles. We generated a hypothetical structural ensemble by perturbing the PDB structure along two normal modes (Fig. 4A, Bottom). We used these structures to produce synthetic templates and train a posterior model to evaluate experimental particles. The posterior could identify the reconstructed cryo-EM structure accurately and precisely (Fig. 4A). Evaluating the posterior on the entire dataset, we found a large concentration of particles ($\sim 50\%$) that we could map to the high-resolution hemagglutinin structure with high confidence, consistent with the 47% used to generate the cryo-EM reference structure (Fig. 4B). These particles are distributed in a region spanning approximately 4 Å around the high-resolution structure, consistent with thermal fluctuation at room temperature, as shown by atomistic MD simulations in explicit solvent (SI Appendix, Fig. S3). As expected, no other conformation is identified in the dataset with high confidence. For many particles, the posterior is approximately uniform and therefore uninformative.

Validating Inference with Latent Space Analysis. The embedding provides a powerful tool for analyzing the particles. The embedding is done by a neural network that encodes all particles from 128^2 pixels into a 256-dimensional representation. Conventional dimensionality reduction techniques can further

reduce the dimensionality of the representation to generate plots that allow us to visually inspect the entire dataset. Each red point in Fig. 4C corresponds to a single experimental particle, whereas synthetic particles are represented as a grayscale heat map (see SI Appendix for details on the UMAPs calculation). The two coordinates, UMAP1 and UMAP2, are nonlinear functions of the pixels defining the original images, which should quantify some “essential” features. Indeed, we can correlate different values of these coordinates to different modeling and imaging parameters, like the SNR, or conformations (SI Appendix, Fig. S4).

When examining the embedding space, an initial question is whether the synthetic hemagglutinin particles used to train our posterior are consistent with the experimental particles. Fig. 4C shows that synthetic particles are distributed in a region that contains the experimental ones. In other words, the simulator generates templates very similar to the experimental ones. Synthetic particles populate a larger region than the experimental ones. This is expected, and means that not all configurations, or imaging parameters, we considered as template hypotheses correspond to particles that constitute the ensemble captured by the cryo-EM experiment. A more quantitative statistical analysis based on the maximum mean discrepancy metric confirms that synthetic and experimental particles are very similar (Materials and Methods and SI Appendix, Fig. S5). This analysis is essential to detect model misspecification, which occurs when the posterior is trained with synthetic data that do not accurately mimic the experimental data, leading to incorrect inference. This issue is demonstrated with nonwhitened particles, where the distributions of simulated and experimental particles do not overlap (SI Appendix, Fig. S5).

We then applied a similar analysis on the experimental particles only, obtaining an insightful low-dimensional view of the entire experimental dataset. Each point in Fig. 4D corresponds to a single experimental image, colored according to the confidence with which the posterior maps it to a specific configuration. The confidence clearly describes a gradient correlating with the first reduced variable defining the plane (UMAP1'). The confidence of the inference is high for particles on the left of the plot and gradually decreases going to the right. Particles on the left of the plot (low values of UMAP1') all contain a clearly visible copy of hemagglutinin in their center (Fig. 4E). On the contrary, particles on the right (large values of UMAP1') do not contain any protein (Fig. 4F). This observation shows that UMAP1' sorts particles according to how well hemagglutinin is visible.

A slender appendix in the 2D UMAP' plot containing high-confidence points detaches from the distribution of points on the top of the plot. These particles clearly contain contaminants, e.g., unfolded proteins, structured ice, or other types of contamination (Fig. 4G). Then, why does our inference lead to a high confidence? The posterior evaluated on these particles peaks at both extreme values of the prior range (SI Appendix, Fig. S6). In other words, the posterior tells us that these particles are highly atypical and incompatible with all structures in our hypothesis ensemble, while they have low-resolution contrast that match some low-resolution template features. This observation is consistent with the appendix-like morphology of the region containing these particles. Typical particles are points with many neighbors. On the contrary, these particles are all on the border, making them very atypical. The analysis of the cryoSBI embedding is a powerful way of validating the inference and analyzing the experimental dataset.

Amortized Template Matching in a Micrograph. In single-particle cryo-EM, particle picking algorithms identify candidate particles within micrographs but often suffer from high false positive rates. Misclassification of noise or junk as target molecules requires further filtering through time-consuming steps like 3D classification. The cryoSBI posterior can match templates directly on a micrograph (Fig. 5). Fig. 5A presents a hemagglutinin micrograph ($3,824 \times 3,824$ pixels) from the same dataset as discussed so far. For this micrograph, we employed a sliding window approach, using the trained cryoSBI posterior with a box size of 256×256 pixels. Evaluating many windows

in parallel proved to be computationally efficient, achieving posterior evaluations for the entire micrograph in a couple of minutes. This allowed us to extract the posterior mean and width, which we then associated with the center of each box. Fig. 5B illustrates the effectiveness of our method in identifying the cryo-EM reference structure from particles within the micrograph using the posterior mean. Examples of boxes with centers exhibiting a mean value close to the reference are shown in Fig. 5C–E together with the posterior width. We note that because a convolutional neural network (which is translation equivariant) is used for the embedding, the exact particle position is not precisely determined. This results in a relatively wide range of pixels where the posterior mean closely matches the reference structure (red in Fig. 5C–E). Analysis of the latent space combined with the shape of the cryoSBI posterior, could also facilitate the identification of outliers versus hemagglutinin particles in the micrograph, as shown above for the picked particles.

Micrographs are likely to contain more junk and out-of-distribution images compared to a stack of previously picked particles. Since cryoSBI has not been trained on junk particles, applying it to an entire micrograph may lead to greater model misspecification than when used on preprocessed particles. Analyzing the latent space alongside the shape of the cryoSBI posterior could help distinguish outliers from hemagglutinin particles within the micrograph, as demonstrated earlier for the selected particles (Fig. 4D). However, training the model with junk particles and ice-mimicking simulations might ultimately prove more effective.

Discussion

Accurately and precisely identifying rare conformations of dynamic biomolecules from individual images is an outstanding challenge in biophysics. CryoSBI aims to overcome this problem by training a neural network with simulations that enables Bayesian inference of biomolecular conformations from single-particle cryo-EM images. Given an initial structural ensemble hypothesis—a set of 3D templates—cryoSBI builds on simulation-based inference, a merger of physics-based simulations and probabilistic deep learning, to perform very fast amortized template matching in cryo-EM. Here, we show how this approach can produce high fidelity inferences from noisy experimental cryo-EM particles and micrographs.

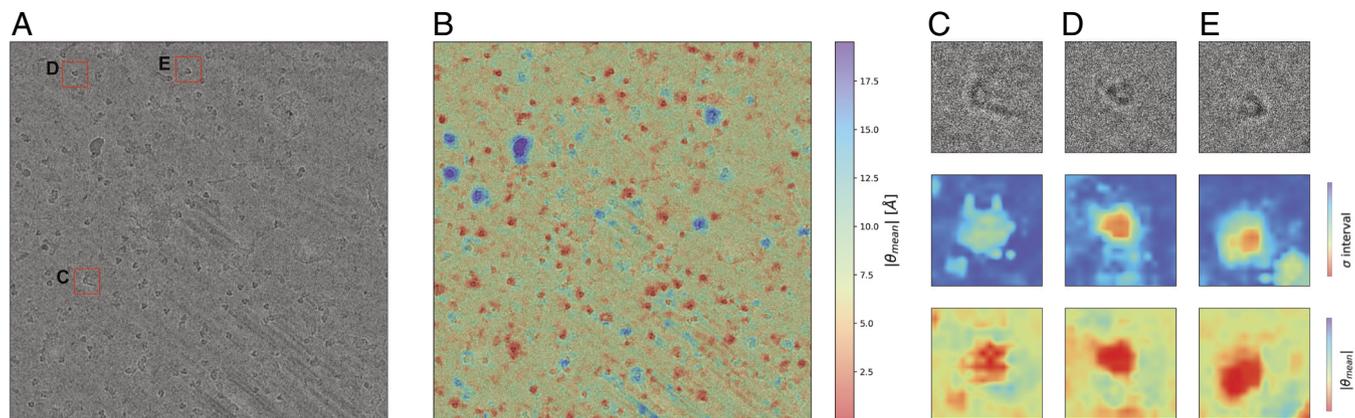


Fig. 5. Evaluating the posterior learned by cryoSBI on a micrograph directly. (A) Cropped micrograph ($3,824 \times 3,824$ pixels) from the hemagglutinin EMPIAR dataset 10026. (B) Absolute Distance in θ between the posterior mean and the cryo-EM reference structure $\theta = 0$ as a function of window position on the micrograph. (C–E) Example windows (Top) from the micrograph with a posterior mean close to the cryo-EM reference (Bottom) and a small posterior width (Middle).

The cryoSBI framework is based on Bayesian inference, which allows us to include prior knowledge and accurately assess uncertainties. CryoSBI can start from standard reconstruction, or AI-based structure prediction, to obtain one or a few reasonable conformations of the biomolecule of interest. Then, MD simulations and AI-methods can generate a hypothetical conformational ensemble from the initial few structures. This ensemble, which does not need to be entirely accurate but should provide a list of structural templates, becomes the prior for our inference. The outcome of the inference is the posterior, a distribution describing the probability that a given particle contains a specific structural template. The cryoSBI posterior not only gives us information about the most probable conformation but also, crucially, provides an accurate statistical CI given by the posterior width. The ability to go beyond point estimates is crucial to distinguish particles that provide useful structural information—characterized by a peaked posterior—from those for which the inference leads to very broad posterior distributions and are, therefore, not informative. In other words, some particles will be too noisy, or originate from a specific pose, such that a precise inference is not feasible, and the cryoSBI posterior will indicate this.

CryoSBI provides amortized inference, enabling the analysis of massive datasets. Given initial structural templates, cryoSBI trains an embedding and neural posterior density using simulated particles. Simulated template particles range over different conformations, poses, and other values of all the nuisance parameters associated with the image formation process. All these simulations are performed once upfront, with which we can train an embedding and a neural density estimator for the inference. The inference is a function only of the templates, marginalized over all other parameters, including the pose. The inference is amortized, i.e., we do not have to perform any optimization to solve the inference problem from scratch for each particle. We have only to evaluate the trained posterior, a forward pass evaluation of a neural network that takes milliseconds. Therefore, in contrast to most traditional and ML reconstruction frameworks, cryoSBI bypasses the pose and defocus estimate, resulting in an extremely fast inference and enabling its efficient application to micrographs and datasets containing millions of particles. Moreover, training with simulations enables connecting the underlying physics of the experiment to the embedding features of the neural network.

We still face significant challenges. In this work, we have demonstrated that cryoSBI can accurately infer large conformational changes in synthetic data and correctly identify conformations in experimental data. Our current focus is on expanding this framework to detect small conformational changes in large proteins and identify multiple alternative states. However, the most critical challenge is identifying and overcoming model misspecification. Any parametric inference is only as good as the model we assume to describe the underlying physical process. In our case, this is the structural ensemble that we use as a starting hypothesis and the simulator encoding the image formation process. Here, we have shown that problems occur when our ensemble is missing structures or when the simulator is missing features depicted in the experimental particles (e.g., Fig. 4G). This can result in “hallucinations”—high-confidence inferences that are completely incorrect (note that this differs from “model-bias” in cryo-EM, where noise is aligned during reconstruction). These problems may become more pronounced when analyzing micrographs directly. However, they could be mitigated by incorporating junk particles into the training data, ensuring the ratio of target molecules to noise/junk matches

the expected distribution across the micrograph, and explicitly training a network for both detection and inference to better define the task.

An additional limitation of the current implementation is that new models must be trained for each new biomolecule of interest. The embedding and neural density estimation are specific to the conformational ensemble used during training, making the model nontransferable across different biomolecules. Future development could leverage recent advancements in generative-AI protein structure generation, which connect protein sequence with their structure or structural ensemble (39, 40, 51). These approaches allow us to represent conformational ensembles (52), creating a natural integration point with cryoSBI. Additionally, pretraining the posterior or embedding networks on diverse proteins and conformations could enable efficient fine-tuning for specific biomolecules, significantly reducing computational costs for new applications.

Importantly, we have shown that the analysis of the latent space provides a powerful framework to diagnose model-misspecification. The accuracy of cryoSBI's inference is only guaranteed if the distribution of points corresponding to the simulated particles largely contains the points corresponding to the experimental ones. This is equivalent to saying that the templates capture the underlying physical features in the cryo-EM data. A small overlap in the latent space would instead immediately reveal inadequate template simulations (e.g., *SI Appendix, Fig. S5*). To this end, we have initiated a thorough quantitative study of the cryoSBI latent space (53), which outlines a validation and physical interpretation of the latent space through manifold learning and dimension reduction techniques. There are many exciting opportunities for this direction, such as reducing model-misspecification by learning from data, which is a very active field of research. Moreover, utilizing equivariant embedding representations may enhance the efficiency of the training process and reduce the number of required simulations.

CryoSBI can improve current cryo-EM reconstruction pipelines. The posterior confidence can be used to classify and sort particles with sharp posteriors (54) and weight the particle contribution to the reconstruction. It could also be used to improve imaging conditions, such as the electron dose per frame, by monitoring the sharpness of posterior widths. An advantage of cryoSBI is that it provides a per-particle measurement plus an error, not suffering from orientational bias. Moreover, it can explain the relation between the conformational motion of interest and the projection direction (Fig. 2B). Therefore, it could be combined with the ML heterogeneous reconstruction methods to sieve particles that do not provide information along the relevant conformational motion due to their projection direction. Even though we focused on inferring conformations in this work, cryoSBI can also be directly used to infer all other parameters involved in the image formation process, such as pose and defocus, setting priors, and ranges for cryo-EM reconstruction.

Cryo-EM has heavily relied on averaging particles. Even the state-of-the-art ML heterogeneous reconstruction methods rely on starting from a consensus volume and struggle with highly dynamic systems. CryoSBI provides a single-particle inference that can contribute to overcoming several problems, such as identifying rare conformations, structural intermediates (transition states), or studying highly flexible biomolecules by leveraging structural hypotheses from molecular simulations. The amortized inference could significantly speed up the recovery of free-energy landscapes from cryo-EM (22, 47, 55, 56) by quickly comparing millions of particles to thousands of structures. Data-driven

techniques applied to the latent space of experimental particles can lead to the discovery of new metastable states and the learning of the overall organization of the conformational landscape (53). Moreover, cryoSBI could speed-up the identification of biomolecules in situ, for studying their cellular contexts and environment-dependent properties by comparing the particles to the templates within the embedding space.

In summary, cryoSBI not only provides accurate structural inferences but also quantifies uncertainties through Bayesian posterior distributions. CryoSBI is a modular and flexible framework. The simulator, the embedding network, and the density estimator can readily integrate more sophisticated algorithms to overcome challenges emerging from complex datasets. Future work will focus on addressing model misspecification and enhance the simulator for more realistic scenarios, such as in situ scenarios, and expanding its capabilities to detect subtle conformational changes. These improvements will be vital for studying molecular conformational ensembles of flexible biomolecules in biologically relevant environments.

Materials and Methods

CryoSBI Simulations. We begin with a conformational ensemble parameterized with a 1D degree of freedom θ , where $f: \mathbb{R}^{3N_{\text{atom}}} \rightarrow \mathbb{R}$ maps a structure X to a real number $\theta = f(X)$, which can be discretized, θ_i , to select a specific 3D template. The goal of cryoSBI is to learn the posterior probability $p(\theta_i|I)$. To train cryoSBI, we run millions of forward model simulations of cryo-EM template images by using the forward model described in *SI Appendix*. We sample from the prior parameter distributions $p(\theta_i, \phi) = p(\theta_i)p(\phi)$, where ϕ contains the nuisance and imaging parameters, and generate a corresponding synthetic image (details in *SI Appendix*). The embedding network for the image featurization and a neural density estimator, a function that approximates the joint distribution of model parameters and features, are trained simultaneously as described below.

Embedding Network. We used a modified ResNet-18 architecture (57) as embedding network $S_\psi(I)$ with parameters ψ to learn a compressed representation of the images I . We adapted the ResNet-18 to grayscale images and 256-dimensional output.

Learning the Posterior. We used the Neural Posterior Estimation (NPE) algorithm to approximate the posterior distribution from synthetic particles (45). NPE uses a neural network density estimator q_φ of parameters φ to approximate the posterior, $p(\theta|I) \approx q_\varphi(\theta|S_\psi(I))$. For each system, we created a large dataset of N synthetic particles $I_n \sim p(I|\theta_n, \phi_n)$ by drawing from the prior conformations, $\theta_n \sim p(\theta)$, nuisance imaging parameters, $\phi_n \sim p(\phi)$, and then running a total of N forward model cryo-EM template simulations, with $n = 1, \dots, N$. We then trained jointly the embedding network and the density estimator by maximizing the average log-likelihood of the posterior probability under the training samples,

$$\mathcal{L}(\varphi, \psi) = \frac{1}{N} \sum_{n=1}^N \log q_\varphi(\theta_n|S_\psi(I_n)). \quad [3]$$

We used a Neural Spline Flow (58) as density estimator, containing five transformation stages. The neural network in each transformation stage contained 12 layers. The first ten layers had 256 hidden nodes, while the last two had 128 and 64 hidden nodes, respectively. We trained the network using an AdamW optimizer (59), with a learning rate of 0.0003, gradient clipping with a maximum norm of 5, and a batch size of 256. Because the forward model simulation is inexpensive, we generated synthetic images during training on demand and did not store the training set. Therefore, each batch of images was newly generated, allowing training until convergence and preventing overfitting. In *SI Appendix*, we provide details about how we evaluate the trained posterior.

Latent-Space Analysis. To detect model misspecification, we compared the latent representations from the embedding network of both experimental and simulated particles using two approaches. The first is a qualitative visual inspection relying on the dimensionality reduction technique UMAP (60) (*SI Appendix*). In a more quantitative way, we compared the distributions of the latent representations using the maximum mean discrepancy (MMD) metric. MMD is a standard metric to statistically test whether two independent sets of samples come from two different distributions (61) and is often used in the SBI community to test model misspecification (62). Let $S_1 = \{s_1\}$ and $S_2 = \{s_2\}$ be two datasets containing the 256-dimensional latent representations. For the MMD, we used a Gaussian kernel with Euclidean distances in the embedding space, so that for a pair of s_1, s_2 , the kernel is $k_\epsilon(s_1, s_2) = \exp(-\|s_1 - s_2\|^2/\epsilon)$, where $\|\cdot\|^2$ is the L_2 -norm. We choose the bandwidth ϵ as the median of pairwise squared distances between the datasets S_1 and S_2 .

Biomolecular Systems. To validate and showcase cryoSBI, we analyzed three biomolecular systems: hsp90, apoferritin, and hemagglutinin (for details, see *SI Appendix*). For hsp90, we generated synthetic data representing a conformational change of the opening of one of its chains, quantified by the rmsd relative to the closed structure across 20 templates. Simulations involved varying structural B-factors, orientations, defocus, and SNRs. For apoferritin, experimental data from EMPIAR 10026 (49) were analyzed. The templates were generated via normal mode analysis, and cryo-EM simulations were conducted with similar settings. Hemagglutinin images and micrographs from EMPIAR 10532 (50) were used to infer the cryo-EM structure from individual particles. Normal modes were used as structural hypotheses. Experimental particles were downsampled and whitened to align with simulated data. For details about the cryoSBI parameter, ranges, training, hemagglutinin micrograph analysis, and MD simulations, see *SI Appendix*. Across all cases, cryoSBI enabled detailed inference of the molecular structures from the individual cryo-EM images.

Data, Materials, and Software Availability. The code is available at GitHub (<https://github.com/flatironinstitute/cryoSBI>) and is based on LAMPE (63), a PyTorch implementation for simulation-based inference. Data and scripts necessary to reproduce all the results in this paper are freely accessible at the Zenodo repository (<https://zenodo.org/records/14593608>) (64).

ACKNOWLEDGMENTS. The Flatiron Institute is a division of the Simons Foundation. E.D. acknowledges Istituto di Ricovero e Cura a Carattere Scientifico Humanitas Research Hospital for financial support. L.D. and R.C. acknowledge the support of Goethe University Frankfurt, the Frankfurt Institute of Advanced Studies, the Landes-Offensive zur Entwicklung Wissenschaftlich-ökonomischer Exzellenz (LOEWE) Center for Multiscale Modelling in Life Sciences of the state of Hesse, the Collaborative Research Center (CRC) of the German Science Foundation 1507: Membrane-associated Protein Assemblies, Machineries, and Supercomplexes (P09), and the International Max Planck Research School on Cellular Biophysics, as well as computational resources and support from the Center for Scientific Computing of the Goethe University and the Jülich Supercomputing Centre. L.D., D.S.-S., and R.C. thank the Flatiron Institute for hospitality while a portion of this research was carried out.

Author affiliations: ^aInstitute of Physics, Faculty of Physics, Goethe University Frankfurt, Frankfurt am Main 60438, Germany; ^bFrankfurt Institute for Advanced Studies, Frankfurt am Main 60438, Germany; ^cDepartment of Applied Mathematics, Yale University, New Haven, CT 06520; ^dCenter for Computational Mathematics, Flatiron Institute, New York, NY 10010; ^eIstituto di Ricovero e Cura a Carattere Scientifico Humanitas Research Hospital, Unit of Correlative Light Emission Microscopy Core, Rozzano, Milan 20089, Italy; ^fRNA Therapeutics Institute, University of Massachusetts Chan Medical School, Worcester, MA 01605; ^gHHMI, University of Massachusetts Chan Medical School, Worcester, MA 01605; ^hInstitute of Computer Science, Faculty of Computer Science and Mathematics, Goethe University Frankfurt, Frankfurt am Main 60325, Germany; and ⁱCenter for Computational Biology, Flatiron Institute, New York, NY 10010

Author contributions: R.C. and P.C. designed research; L.D., D.S.-S., L.E., R.C., and P.C. performed research; E.D. and N.G. contributed new reagents/analytic tools; L.D., D.S.-S., L.E., R.C., and P.C. analyzed data; and L.D., R.C., and P.C. wrote the paper.

1. M. P. Rout, A. Sali, Principles for integrative structural biology studies. *Cell* **177**, 1384–1403 (2019).
2. Y. Cheng, Single-particle cryo-EM at crystallographic resolution. *Cell* **161**, 450–457 (2015).
3. J. Dubochet *et al.*, Cryo-electron microscopy of vitrified specimens. *Q. Rev. Biophys.* **21**, 129–228 (1988).
4. L. V. Bock, H. Grubmüller, Effects of cryo-EM cooling on structural ensembles. *Nat. Commun.* **13**, 1709 (2022).
5. N. J. Mowry, C. R. Kruger, M. Drabbel, U. J. Lorenz, Direct measurement of the critical cooling rate for the vitrification of water. *arXiv [Preprint]* (2024). <http://arxiv.org/abs/2407.01087> (Accessed 20 June 2024).
6. S. H. Scheres, Relion: Implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
7. D. Lyumkis, A. F. Brilot, D. L. Theobald, N. Grigorieff, Likelihood-based classification of cryo-EM images using frealign. *J. Struct. Biol.* **183**, 377–388 (2013).
8. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
9. P. Cossio, G. Hummer, Likelihood-based structural analysis of electron microscopy images. *Curr. Opin. Struct. Biol.* **49**, 162–168 (2018).
10. M. Bonomi, R. Pellarin, M. Vendruscolo, Simultaneous determination of protein structure and dynamics using cryo-electron microscopy. *Biophys. J.* **114**, 1604–1613 (2018).
11. S. E. Hoff, F. E. Thomsen, K. Lindorff-Larsen, M. Bonomi, Accurate model and ensemble refinement using cryo-electron microscopy maps and Bayesian inference. *PLoS Comput. Biol.* **20**, e1012180 (2024).
12. C. Blau, L. Yvonesdotter, E. Lindahl, Gentle and fast all-atom model refinement to cryo-EM densities via a maximum likelihood approach. *PLoS Comput. Biol.* **19**, e1011255 (2023).
13. T. Wlodarski *et al.*, Bayesian reweighting of biomolecular structural ensembles using heterogeneous cryo-EM maps with the cryoensemble method. *Sci. Rep.* **14**, 18149 (2024).
14. M. Beck, R. Covino, I. Hänelt, M. Müller-McNicoll, Understanding the cell: Future views of structural biology. *Cell* **187**, 545–562 (2024).
15. C. Sorzano *et al.*, On bias, variance, overfitting, gold standard and consensus in single-particle analysis by cryo-electron microscopy. *Acta Crystallogr. Sec. D Struct. Biol.* **78**, 410–423 (2022).
16. A. Dashti *et al.*, Trajectories of the ribosome as a Brownian nanomachine. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 17492–17497 (2014).
17. E. D. Zhong, T. Bepler, B. Berger, J. H. Davis, CryoDRGN: Reconstruction of heterogeneous cryo-EM structures using neural networks. *Nat. Methods* **18**, 176–185 (2021).
18. M. Chen, S. J. Ludtke, Deep learning-based mixed-dimensional Gaussian mixture model for characterizing variability in cryo-EM. *Nat. Methods* **18**, 930–936 (2021).
19. A. Punjani, D. J. Fleet, 3D flexible refinement: Structure and motion of flexible proteins from cryo-EM. *bioRxiv [Preprint]* (2021). <https://doi.org/10.1101/2021.04.22.440893> (Accessed 17 June 2024).
20. I. Hamitouche, S. Jonic, Deepemnma: Resnet-based hybrid analysis of continuous conformational heterogeneity in cryo-EM single particle images. *Front. Mol. Biosci.* **9**, 965645 (2022).
21. J. Schwab, D. Kimanius, A. Burt, T. Deendoooven, S. H. W. Scheres, DynaMight: Estimating molecular motions with improved reconstruction from cryo-EM images. *Nat. Methods* **21**, 1855–1862 (2024).
22. W. S. Tang, E. D. Zhong, S. M. Hanson, E. H. Thiede, P. Cossio, Conformational heterogeneity and probability distributions from single-particle cryo-electron microscopy. *Curr. Opin. Struct. Biol.* **81**, 102626 (2023).
23. A. Punjani, D. J. Fleet, 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM. *J. Struct. Biol.* **213**, 107702 (2021).
24. M. A. Gilles, A. Singer, Cryo-EM heterogeneity analysis using regularized covariance estimation and kernel regression. *Proc. Natl. Acad. Sci. U.S.A.* **122**, e2419140122 (2025).
25. D. Herreros *et al.*, Estimating conformational landscapes from cryo-EM particles by 3d Zernike polynomials. *Nat. Commun.* **14**, 154 (2023).
26. R. Vuillemot *et al.*, Extracting continuous conformational landscapes from cryo-EM single particle datasets using 3D-to-2D flexible fitting based on molecular dynamics simulation. *J. Mol. Biol.* **435**, 16795167951 (2023).
27. A. Levy *et al.*, "Cryoai: Amortized inference of poses for ab initio reconstruction of 3d molecular volumes from real cryo-EM images" in *European Conference on Computer Vision* (Springer, 2022), pp. 540–557.
28. A. Levy, G. Wetzstein, J. N. Martel, F. Poitevin, E. Zhong, Amortized inference for heterogeneous reconstruction in cryo-EM. *Adv. Neural Inf. Process. Syst.* **35**, 13038–13049 (2022).
29. R. R. Lederman, B. Toader, "On manifold learning in plato's cave: Remarks on manifold learning and physical phenomena" in *2023 International Conference on Sampling Theory and Applications (SampTA)* (IEEE, 2023), pp. 1–7.
30. D. A. Klindt, A. Hyvärinen, A. Levy, N. Miolane, F. Poitevin, Towards interpretable cryo-EM: Disentangling latent spaces of molecular conformations. *Front. Mol. Biosci.* **11**, 1393564 (2024).
31. P. Cossio, G. Hummer, Bayesian analysis of individual electron microscopy images: Towards structures of dynamic and heterogeneous biomolecular assemblies. *J. Struct. Biol.* **184**, 427–437 (2013).
32. P. Cossio *et al.*, Bioem: Gpu-accelerated computing of Bayesian inference of electron microscopy images. *Comput. Phys. Commun.* **210**, 163–171 (2017).
33. J. P. Rickgauer, N. Grigorieff, W. Denk, Single-protein detection in crowded molecular environments in cryo-EM images. *eLife* **6**, e25648 (2017).
34. B. A. Lucas *et al.*, Locating macromolecular assemblies in cells by 2d template matching with cistem. *eLife* **10**, e68946 (2021).
35. S. Cruz-León *et al.*, High-confidence 3D template matching for cryo-electron tomography. *Nat. Commun.* **15**, 3992 (2024).
36. K. Cranmer, J. Brehmer, G. Louppe, The frontier of simulation-based inference. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 30055–30062 (2020).
37. L. Dingeldein, P. Cossio, R. Covino, Simulation-based inference of single-molecule experiments. *Curr. Opin. Struct. Biol.* **91**, 102988 (2025).
38. H. Jung *et al.*, Machine-guided path sampling to discover mechanisms of molecular self-organization. *Nat. Comput. Sci.* **3**, 334–345 (2023).
39. J. Abramson *et al.*, Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
40. S. Lewis *et al.*, Scalable emulation of protein equilibrium ensembles with generative deep learning. *bioRxiv [Preprint]* (2024). <https://doi.org/10.1101/2024.12.05.626885> (Accessed 13 December 2024).
41. G. Papamakarios, I. Murray, Fast ϵ -free inference of simulation models with Bayesian conditional density estimation. *Adv. Neural Inf. Process. Syst.* **29**, 06376 (2016).
42. L. Dingeldein, P. Cossio, R. Covino, Simulation-based inference of single-molecule force spectroscopy. *Mach. Learn. Sci. Technol.* **4**, 025009 (2023).
43. M. Dax *et al.*, Real-time gravitational wave science with neural posterior estimation. *Phys. Rev. Lett.* **127**, 241103 (2021).
44. P. J. Gonçalves *et al.*, Training deep neural density estimators to identify mechanistic models of neural dynamics. *eLife* **9**, e56261 (2020).
45. J. M. Lueckmann, J. Boelts, D. Greenberg, P. Gonçalves, J. Macke, "Benchmarking simulation-based inference" in *International Conference on Artificial Intelligence and Statistics* (PMLR, 2021), pp. 343–351.
46. E. Seitz, F. Acosta-Reyes, P. Schwander, J. Frank, Simulation of cryo-EM ensembles from atomic models of molecules exhibiting continuous conformations. *bioRxiv [Preprint]* (2019). <https://doi.org/10.1101/864116> (Accessed 8 July 2024).
47. J. Giraldo-Barreto *et al.*, A Bayesian approach to extracting free-energy profiles from cryo-electron microscopy experiments. *Sci. Rep.* **11**, 13657 (2021).
48. C. Donnat, A. Levy, F. Poitevin, E. D. Zhong, N. Miolane, Deep generative modeling for volume reconstruction in cryo-electron microscopy. *J. Struct. Biol.* **214**, 107920 (2022).
49. C. J. Russo, L. A. Passmore, Ultrastable gold substrates for electron cryomicroscopy. *Science* **346**, 1377–1380 (2014).
50. Y. Z. Tan, J. L. Rubinstein, Through-grid wicking enables high-speed cryo-EM specimen preparation. *Acta Crystallogr. Sec. D Struct. Biol.* **76**, 1092–1103 (2020).
51. A. Levy *et al.*, Solving inverse problems in protein space using diffusion-based priors. *arXiv [Preprint]* (2024). <http://arxiv.org/abs/2406.04239> (Accessed 9 July 2024).
52. A. Maddipati *et al.*, Inverse problems with experiment-guided AlphaFold. *arXiv [Preprint]* (2025). <http://arxiv.org/abs/2502.09372> (Accessed 12 March 2025).
53. L. Evans *et al.*, Cryo-EM images are intrinsically low dimensional. *arXiv [Preprint]* (2025). <https://doi.org/10.48550/arXiv.2504.11249> (Accessed 15 May 2025).
54. J. Zhu *et al.*, A minority of final stacks yields superior amplitude in single-particle cryo-EM. *Nat. Commun.* **14**, 7822 (2023).
55. W. S. Tang *et al.*, Ensemble reweighting using cryo-EM particle images. *J. Phys. Chem. B* **127**, 5410–5421 (2023).
56. L. Evans *et al.*, Counting particles could give wrong probabilities in cryo-electron microscopy. *bioRxiv [Preprint]* (2025). <https://doi.org/10.1101/2025.03.27.644168> (Accessed 1 April 2025).
57. K. He, X. Zhang, S. Ren, J. Sun, "Deep residual learning for image recognition" in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2016), pp. 770–778.
58. C. Durkan, A. Bekasov, I. Murray, G. Papamakarios, Neural spline flows. *Adv. Neural Inf. Process. Syst.* **32**, 7511–7522 (2019).
59. I. Loshchilov, F. Hutter, Decoupled weight decay regularization. *arXiv [Preprint]* (2017). <http://arxiv.org/abs/1711.05101> (Accessed 10 July 2024).
60. L. McInnes, J. Healy, J. Melville, Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv [Preprint]* (2018). <http://arxiv.org/abs/1802.03426> (Accessed 12 July 2024).
61. A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, A. Smola, A kernel two-sample test. *J. Mach. Learn. Res.* **13**, 723–773 (2012).
62. M. Schmitt, P. C. Bürkner, U. Köthe, S. T. Radev, Detecting model misspecification in amortized Bayesian inference with neural networks. *arXiv [Preprint]* (2021). <http://arxiv.org/abs/2112.08866> (Accessed 11 July 2024).
63. F. Rozet *et al.*, Lampe: Likelihoodfree amortized posterior estimation. <https://pypi.org/project/lampe/>. Accessed 19 May 2025.
64. L. Dingeldein *et al.*, Amortized template-matching of molecular conformations from cryo-electron microscopy images using simulation-based inference: Code and data. Zenodo. <https://zenodo.org/records/14593608>. Deposited 8 February 2025.