bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Effects of cryo-EM cooling on structural ensembles

Lars V. Bock*[1] and Helmut Grubmüller[1]

[1]Theoretical and Computational Biophysics Department, Max Planck Institute for Biophysical Chemistry, Göttingen, Germany

Abstract

Structure determination by cryo electron microscopy (cryo-EM) provides information on structural heterogeneity and ensembles at atomic resolution. To obtain cryo-EM images of macromolecules, the samples are first rapidly cooled down to cryogenic temperatures. To what extent the structural ensemble is perturbed by the cooling is currently unknown. Here, to quantify the effects of cooling, we combined continuum model calculations of the temperature drop, molecular dynamics simulations of a ribosome complex before and during cooling with kinetic models. Our results suggest that three effects markedly contribute to the narrowing of the structural ensembles: thermal contraction, reduced thermal motion within local potential wells, and the equilibration into lower free-energy conformations by overcoming separating free-energy barriers. During cooling, barrier heights below 10 kJ/mol were found to be overcome resulting in reduction of B-factors in the ensemble imaged by cryo-EM. Our approach now enables the quantification of the heterogeneity of room-temperature ensembles from cryo-EM structures.

Introduction

Single-particle cryo-electron microscopy (cryo-EM) is a method to resolve 3-d structures of biomolecules. Over the past years, new electron detectors and improved image processing enabled cryo-EM maps at sufficient resolution to infer the atomic structure of a large range of macromolecules[1;2;3]. Today, atomic resolution can be achieved for single proteins, visualizing the densities of individual atoms[4;5].

In general, biomolecules perform their functions in solution, however, the direct study of specimens in liquid solutions using EM is impeded because EM requires high vacuum. To preserve the biomolecules in the hydrated state, the sample is first spread into a thin film on a cryo-EM support grid. Subsequently, the grid is rapidly cooled, embedding the biomolecules in ice[6;7;8;9;10]. Because the formation of ice crystals would damage the sample, the cooling has to be rapid enough to result in vitreous (amorphous) ice[11;12]. Dubochet et. al provided an order-of-magnitude estimate for minimum cooling rate required for vitrification (10^6 K/s)[13], which would translate into cooling times shorter than 200 μs. Fast cooling rates are achieved by plunging the sample into the cryogenic liquid, mostly liquid ethane kept close to its melting temperature of ∼90 K. Alternatively, an ethane:propane mixture is used which has a lower melting temperatures and remains liquid at 77 K presumably resulting in more rapid cooling[14]. An additional benefit of studying the sample at cryogenic temperatures is the reduction of beam-induced radiation damage[15]. Radiation damage stems from the energy deposited by the electrons which can result in the ionization of the sample and the breakage of bonds which is reduced at lower temperatures. The vitrified sample is then transferred to a transmission electron microscope and the 2-d EM images of individual randomly orientated specimen are used to reconstruct 3D cryo-EM maps[16].

The rapid cooling during plunge-freezing preserves part of the structural heterogeneity generated by room temperature structural fluctuations and, hence, contains information on functional motions[17;18;19;20;21]. However, cooling of the sample is expected to perturb the structural ensemble of biomolecules. At room temperature, biomolecules can thermodynamically access more conformations and the rates for switching between conformations are faster than at low temperatures[22]. If the cooling is so rapid that essentially no barriers can be overcome in the process, the molecules are kinetically trapped in very nearby local minima and, hence, the room-temperature ensemble is fully preserved. In contrast, if the cooling rate is very slow, the

*To whom correspondence should be addressed. Email: lbock@mpibpc.mpg.de

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

molecules can overcome most barriers and equilibrate into the conformations thermodynamically accessible at the low temperature, i.e., the lowest free energy minima. For intermediate cooling rates, one would expect that conformational changes with rates above a certain threshold (low barriers) would equilibrate into local free-energy minima during the cooling, while for conformational changes with slower rates (high barriers), the high-temperature ensemble would be largely preserved. For the slow conformational changes where the high-temperature ensemble is preserved, the ensemble after cooling is expected to depend on the temperature before cooling[23]. An example for this scenario is given by cryo-EM reconstructions of ribosomes which were kept at different temperatures prior to cooling (37 °C, 18 °C, 4 °C)[17]. The 30S body rotation angle of ribosomes cooled down from a temperature of 37 °C showed a broad distribution. With lower temperature prior to cooling, the distribution narrowed, showing that information of the angle distribution is preserved during cooling. These differences of the distributions indicate that the 30S body rotation is too slow to equilibrate during the cooling process. Further, in a recent study, Chen et al. found that the captured conformation a ketol-acid reductoisomerase differs dramatically for different temperatures prior to cooling ranging from 4 °C to 70 °C[24].

The effect that cooling down from an ensemble at lower temperature leads to a more homogeneous ensemble is used to obtain high-resolution structures by keeping the sample at temperatures below the physiological temperature before plunging, often at 4°C (277.15 K)[25;4;5]. Further, the observation that some conformational changes are slower than the cooling rate is used in time-resolved cryo-EM which allows to obtain structural information for different time points of a system started out of equilibrium[26;27]. To that aim, e.g., after starting a reaction by mixing the reactants, the sample is frozen after different time intervals. Kinetic information can then be obtained from counting the specimens in the different states at the different time points. The fastest accessible rates are determined by the shortest time interval from mixing to the frozen sample, currently reaching 6 ms[28].

The effects of low temperatures on the dynamics of proteins have been studied extensively using Mössbaurer spectroscopy[29;30;31;32], x-ray crystallography[22;33;34;35;36], neutron scattering[37;38], IR spectroscopy[39], NMR[40], and molecular dynamics (MD) simulations[41;42;43;44;45;46;47;48;49]. The kinetic effects of cooling on biomolecules in crystals are expected to be very different from those embedded in thin solvent layers for single partice cryo-EM. In contrast to the thin solvent layers, the crystal diameters are typically larger than 50 μm leading to cooling rates between 50 K/s and 700 K/s (ref.[50]), which is slower than the rates needed to vitrify pure water[51;36]. However, vitrification of water is achieved by added cryoprotectants[52] and by the highly concentrated proteins acting as cryoprotectants themselves[53;54;55]. Interestingly, the temperature dependence of the B-factors (Debye-Waller Factors, temperature factors) shows a glass transition between 180 K and 220 K from a smaller to a larger slope[29;31;37;41;42;43;35;44;45;46;38;56;48;57], e.g. from 1.2 Å^2 per 100 K to 6.4 Å^2 per 100 K (ref.[35]). In contrast to temperatures above the glass-transition temperature, crystallised ribonuclease-A cannot bind or unbind an inhibitor below the glass transition temperature indicating that the conformational change necessary for binding and unbinding is not accessible[34]. A dependency of the cooled ensemble on the cooling rate was hinted at by MD simulations of the protein carboxymyoglobin started from a single structure with very rapid cooling rates of 2 K/ps and 0.2 K/ps where a lower potential energy conformation was reached after the slower cooling, indicating that kinetics affect the cooled ensemble[45].

Here, we investigate the effects of plunge-freezing on structural ensembles by combining molecular dynamics (MD) simulations with kinetic models of the cooling process. As a first step, for different water-layer thicknesses, we used a continuum model to estimate the temperature drop rates after plunging a water layer into liquid ethane. To probe how the cooling rate affects structural ensembles, we used explicit-solvent all-atom MD simulations of a ribosome·EF-Tu complex with decreasing temperatures at different rates (2 K/ps to 1.6 K/ns) to cool down an ensemble of 41 ribosome conformations. We chose the ribosome·EF-Tu complex because it includes very rigid parts, e.g., the surrounding of the peptidyl-transferase center, and large flexible parts which undergo conformational changes on 100-ns timescales, e.g., the L1 stalk[58]. Next, we trained and cross-validated kinetic models of the cooling process using the results from the MD simulations to estimate the effects of the cooling during plunge-freezing. Finally, we combined the kinetic model with the temperature drops expected for water-layer thicknesses used in cryo-EM experiments. This combination allowed us to predict the width of the structural ensemble, i.e., the B-factors before cooling from the B-factors observed in cryo-EM structures, thereby improving the interpretation of cryo-EM structures and their comparison to results from other experiments.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

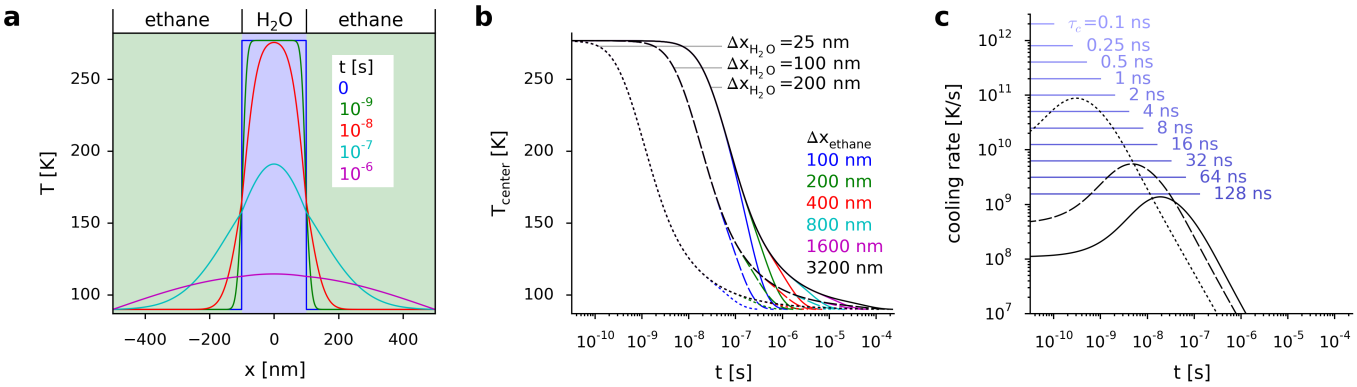

Figure 1: **Plunge-freezing of water layer embedded in ethane** (**a**) Solution of the heat equation for a layer of water (thickness $\Delta x_{\mathrm{H_2O}}$=200 nm) surrounded by two layers of ethane (thickness $\Delta x_{\mathrm{ethane}}$=400 nm, each). The temperature profile T is shown for different times t. The temperatures at the left and right borders are kept at 90 K. (**b**) The temperature T_{center} at $x = 0$ nm is shown as function of time t for different thicknesses of the water and ethane layers. (**c**) The cooling rate at $x = 0$ nm is shown as function of t for different water-layer thicknesses and for $\Delta x_{\mathrm{ethane}}$=3200 nm. Cooling rates for linear temperature decreases with different cooling time spans τ_c are shown as blue lines.

Results

Temperature drop during plunge-freezing

The most common procedure to cool cryo-EM samples is to plunge the support grid with thin films of the sample into liquid ethane kept close to its melting temperature of 90 K. To estimate the rate and shape of the temperature drop within the sample, we considered a system consisting of three layers (Fig. 1a), one layer of water with a thickness of $\Delta x_{\mathrm{H_20}}$ surrounded by two layers of ethane with a thickness of $\Delta x_{\mathrm{ethane}}$. The initial temperature profile at the time of plunging was set to 90 K for the ethane layers and 277.15 K for the water layer (Fig. 1a, blue line). To calculate the temperature profile evolution over time, the heat equation was solved numerically, with the temperature at the outer boundaries kept at 90 K. In experiments, different ice thicknesses between 15 nm and 200 nm depending on preparation and specimen were observed[59;8;60;61;5]. To estimate the effect of the water-layer thickness on the temperature drop, we used three values for $\Delta x_{\mathrm{H_20}}$, a minimal value of 25 nm slightly larger than the diameter of the ribosome and two values of 100 nm and 200 nm which capture the range of observed thicknesses. In the experiments, the layer of ethane is much larger than the water layer. To assess the effect of $\Delta x_{\mathrm{ethane}}$ on the temperature drop within the water layer, we used $\Delta x_{\mathrm{ethane}}$ values ranging from 100 nm to 3200 nm.

Figure 1b shows the temperature at the center of the water layer T_{center} as a function of time. For the thinnest water layer (25 nm), temperatures below 150 K are reached within 4 ns, whereas for the thicker layers (100 nm, 200 nm) it takes 64 ns and 250 ns, respectively. Temperatures below 100 K are reached within 120 ns, 1.9 μs and 7.6 μs. Before reaching 150 K, for water thicknesses between 25 nm and 200 nm, cooling rates in the range between 10^8 K/s and 10^{11} K/s are observed (Fig. 1c). We therefore decided to use cooling rates between $2 \cdot 10^9$ K/s and $2 \cdot 10^{12}$ K/s (blue lines) for the subsequent molecular dynamics (MD) temperature-quench (T-quench) simulations. As can be seen, the slowest cooling rates capture water thicknesses between 100 nm and 200 nm, whereas the faster cooling rates describe systems with minimal water layer thickness. Further, these cooling rates are consistent with the estimated minimum cooling rate of 10^6 K/s required for vitrification[13].

For small values of $\Delta x_{\mathrm{ethane}}$, the temperature drops more rapidly, because temperatures larger than 90 K reach the outer boundaries earlier. With the largest $\Delta x_{\mathrm{ethane}}$ values, this deviation occurs when the temperature is close to the ethane temperature (<100 K), indicating that for $\Delta x_{\mathrm{ethane}} = 3200$ nm the effect of the boundaries on the relevant part of the cooling process is small and therefore will no be considered further.

Apart from the water-layer thickness, the temperature drop also depends on the position within the layer with the slowest drop in the center (Fig. 1a), which is relevant, because in the time between the spreading

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of the sample onto the grid and the plunging, the biomolecules tend to adsorb to the air-water interface[62].

Effects of different cooling rates on the ensemble

To quantify the effects of different cooling rates on structural ensembles of large biomolecular complexes typically studied by cryo-EM, we used T-quench simulations, i.e. all-atom explicit-solvent MD simulations of a ribosome complex with linearly decreasing temperatures at different rates. A similar approach was proposed in a recent viewpoint article[63]. To that aim, we repeated T-quench simulations 41 times for each cooling rate, starting from 41 snapshots of an ensemble at 277.15 K (4 °C), the temperature used in the experiment before plunge-freezing[25]. The latter ensemble was generated from a 3.5-μs MD simulation started from a cryo-EM structure of the ribosome·EF-Tu complex with the bound antibiotic kirromycin[25]. During the simulation, the deviation from the starting structure measured by the root mean square deviation (rmsd) approaches 5 Å (Fig. S1a), which is similar to the rmsd obtained from 2-μs simulations of the same system at 300 K presented earlier[64]. The observation that, at 300 K, similar rmsd values are reached in shorter time suggests that the dynamics on the timescales of the simulations is markedly slowed down by the lower temperature, as expected.

During cooling, one would expect that the structural heterogeneity of biomolecules decreases as the biomolecules equilibrate into local free-energy minima at the lower temperatures. As a result, the broader room-temperature ensemble should become narrower, depending on whether or not the relevant energy barriers can be overcome during the cooling process. To quantify this process, we recorded the distribution of root mean square fluctuations (rmsf) of atoms during the T-quench simulations, which corresponds to the width of distribution of atom positions and therefore serves as a measure of structural heterogeneity. As a reference, an ensemble of structures that is sufficiently converged with respect to this observable was required. To that aim, we extracted structures of the ribosome complex at intervals of 50 ns from the 3.5-μs simulation at 277.15 K (Fig. S1a). Then, we grouped the 41 structures from the time points between 0 μs and 2 μs into an ensemble and calculated rmsf values for all atoms resolved in the cryo-EM structure[25]. This rmsf calculation was repeated for intervals 0.1–2.1 μs, 0.2–2.2 μs, ..., and 1.5–3.5 μs and the resulting rmsf distributions are shown in Fig. S1b.

Throughout this work, we will quantify the structural heterogeneity of an ensemble with the 6-quantiles of the rmsf histograms, i.e. the 5 rmsf values Q1 to Q5 that divide the set of rmsf values into subsets of equal size. The rmsf distributions for the different time intervals show that initially the rmsf values decrease with increasing simulation time until 0.3–2.3 μs (Fig. S1b). For later time intervals, the rmsf values increase with a decreasing slope (intervals 0.4–2.4 μs to 1.5–3.5 μs). The observed behaviour is consistent with a slight adaptation of the initial ribosome structure obtained at cryogenic temperatures to near room temperature, with subsequent room temperature fluctuations. For the cooling simulations, we expected a decrease in rmsf values. To observe this effect, it is crucial to start cooling from an ensemble of structures which does show a decrease in rmsf values in the absence of cooling. We therefore chose the ensemble of 41 structures in the 1–3 μs interval (Fig. S1b, red histogram).

To address the question of how the width of the ensemble, measured by the rmsf distribution, depends on the cooling rate, we started T-quench simulations from the 41 structures of the 277.15-K ensemble (Fig. 2a, red points) with the 11 different cooling time spans τ_c (0.1, 0.25, 0.5, 1, 2, 4, 6, 16, 32, 64, and 128 ns) that capture the plunge-freezing cooling rate (Fig. 1c). For each of the starting structures and for each cooling time span τ_c, an MD simulation of length τ_c was started during which the temperature was linearly decreased to reach 77 K at the end of the simulation.

For each cooling time span, we then extracted the final structures of each of the 41 simulations and calculated the rmsf distribution (Fig. 2b, blue histograms). Indeed, the observation that the quantiles from the cooled ensembles (blue lines) are consistently lower than those from the 277.15 K ensemble (red lines) indicates that, first, the cooling on the tested timescales reduces the heterogeneity of the structural ensemble, and, second, that the number of performed simulations is large enough to observe this effect.

To obtain rmsf values during the course of T-quenching, for each of the cooling time spans τ_c, we first extracted conformations from all 41 simulations at 11 time points (0 ns, $0.1\tau_c$, $0.2\tau_c$, ..., τ_c). Next, for each cooling time span and time point, the rmsf values of all atoms were calculated from the 41 conformations. The 6-quantiles of these rmsf distributions are shown in Fig. 2c (Q1–Q5, cyan lines). The light cyan areas denote the standard deviations of the rmsf values obtained from bootstrapping of the conformations.

In addition to the decrease in heterogeneity, we observed a decrease in the size of the ribosome over the course of the cooling simulations, reflecting simple thermal expansion. To quantify this decrease, we

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

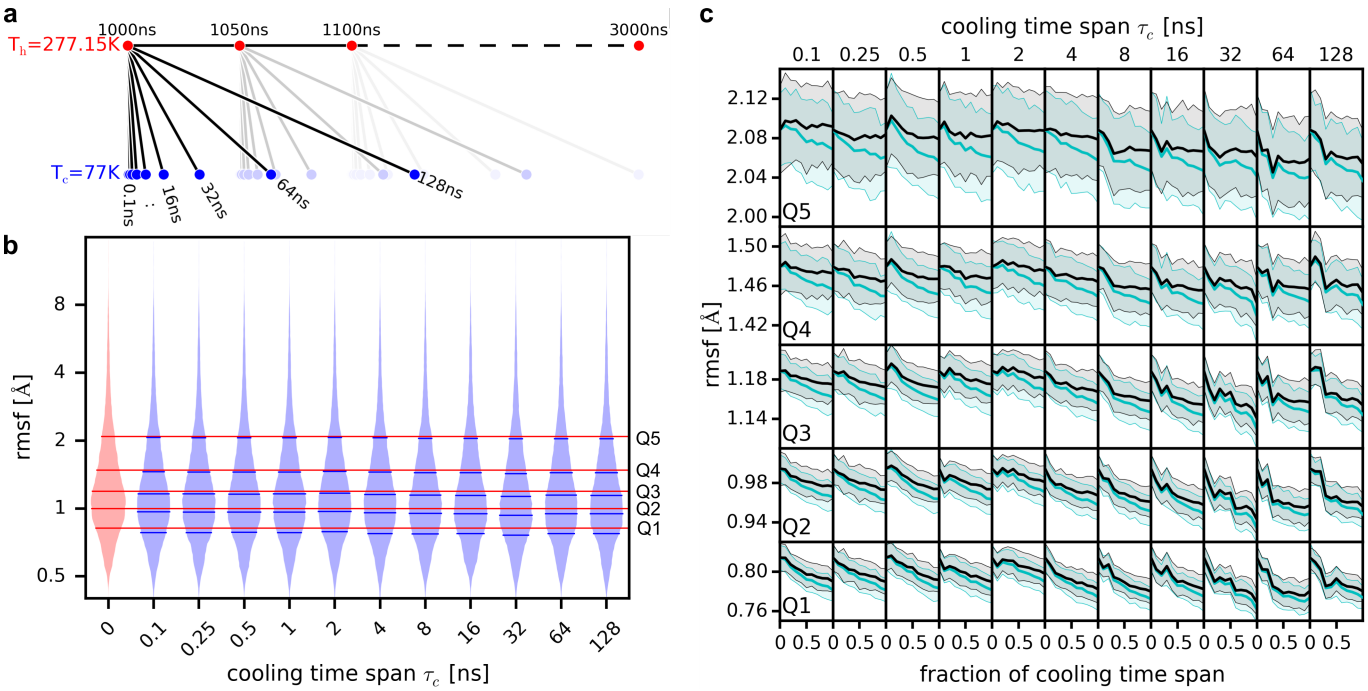

Figure 2: **Effects of different cooling rates in T-quench MD simulations.** (**a**) Schematic of T-quench simulation protocol. From a trajectory of the ribosome-EF-Tu complex at T=277.15 K, 41 snapshots were extracted (1000 ns to 3000 ns, every 50 ns). From these snapshots, T-quench simulations of different lengths (cooling time span τ_c 0.1 ns to 128 ns) were started with linearly decreasing temperature from 277.15 K to 77 K. (**b**) Histograms of the root mean square fluctuations (rmsf) of the heavy atoms of the ribosome·EF-Tu complex obtained from ensembles of 41 snapshots. The histogram for the ensemble before cooling is shown (light red area). Red horizontal lines show the 6-quantiles (Q1–Q5). For each cooling time span τ_c, the rmsf histogram of the ensemble of the final snapshots of the 41 cooling simulations is shown (light blue area) with 6-quantiles (horizontal blue lines). (**c**) Rmsf quantiles are shown as a function the simulation time before (mean value: cyan line, standard deviation: light cyan area) and after rescaling the conformations (black line, gray area).

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

applied a scaling factor s to all of the extracted coordinates (relative to the center of geometry). For each conformation, the scaling factor s that minimized the rmsd from the corresponding starting structure was calculated. For all cooling time spans, we saw the expected linear increase of s as a function of the cooling time (Fig. S2). The scaling factors correspond to a thermal linear expansion coefficient between $4 \cdot 10^{-5}$ K^{-1} and $6 \cdot 10^{-5}$ K^{-1} and a volume decrease between 2.5 % and 3.3 % at the end of the cooling trajectories. This size decrease is in agreement with a volume reduction by 3 % observed in previous MD simulations after switching the temperature from 303 K to 85 K[49]. Further, from crystal structures of ribonuclease-A obtained at nine temperatures from 98 K to 320 K (ref.[35]), we calculated a linear expansion coefficient of $4 \cdot 10^{-5}$ K^{-1} (standard deviation $1 \cdot 10^{-5}$ K^{-1}). Assuming that this heat expansion coefficient is rather independent of the system size, this good agreement suggests that our simulations describe the effect of cooling on the structural ensemble accurately.

To test if the observed changes in the rmsf distributions are largely a result of the temperature-induced shrinkage of the whole ribosome or, if not, to what extent the shrinkage affects the rmsf distributions, we scaled each conformation, which was extracted from the cooling trajectories, with the corresponding value of s. For each cooling time span and each time point, the rmsf quantiles were calculated before and after rescaling. As shown in Figure 2c, the rescaling results in increased rmsf values, but it does not fully compensate the drop in rmsf values during cooling. This result suggests that in addition to the decrease in size, cooling does affect the heterogeneity of the structural ensembles even for very rapid cooling.

Thermodynamic and kinetic models of the cooling process

The structural heterogeneity, as indicated by the quantiles of the rmsf distributions during the T-quench simulations decreases for all cooling time spans (Fig. 2c). In addition, the decrease appears to be more pronounced with increasing cooling time spans, i.e. with slower cooling rates, which would indicate a kinetic contribution to the decrease. To quantify the effects of cooling and to separate thermodynamic (cooling-rate independent) from kinetic (rate dependent) contributions to the decrease, we used the results from the T-quench simulations to train and validate three different models of the cooling process. First, a thermodynamic model that does not depend on the cooling rate (Fig. 3a, model1) and two kinetic models that do depend on the cooling rate (Fig. 3b,c; model2 and model3).

The cooling-rate independent model1 describes the rmsf of an ensemble of atoms that can be trapped in different positions without the possibility to switch between the conformations during cooling. In this model, the atoms are subjected to harmonic potentials with a force constant c which are uniformly distributed in an interval from $-d$ to d (Fig. 3a). The parameter d determines the spread of the different accessible positions of the atoms.

To test for cooling-rate dependent effects, we used an established kinetic two-state model that was used previously to study of the dynamic behaviour of proteins at different temperatures (Fig. 3b; model2)[30;37;36]. Here, states A and B are located at a distance Δx and the free energy of state B is larger than that of state A by ΔG. The rates of switching between the states k_{AB} and k_{BA} are governed by the barrier height $\Delta G^{\ddagger}$ and the temperature (see Methods). The pre-exponential factor of the modified Arrhenius equation is $\kappa(T(t)/T_h)^{\nu}$, where $T(t)$ is the temperature as a function of time, T_h the temperature before cooling, and κ is a scaling factor. The temperature exponent ν controls how the temperature enters in the pre-exponential factor.

Model3 is a combination of model2 and model3, where, in addition to the two-state kinetics of model2, the probability distribution in the two states is governed by uniformly distributed harmonic potentials (Fig. 3c). This model can describe the kinetic effects of a barrier between states that results in cooling-rate dependent behaviour and the equilibration in multiple conformations.

To obtain probability distributions of the parameters for each model, we used Bayes' theorem and Metropolis sampling[65] (see Methods). As the likelihood function, which describes how well the model reproduces the rmsf values as a function of time during the T-quench simulations, we used a normal distribution. To disentangle the contribution of the size decrease to the decrease in rmsf values from other effects, we used the rmsf values obtained after rescaling the conformations (Fig. 2c, black lines).

Using all model parameters for each quantile separately results in 10, 15, and 25 free parameters for model1, model2, and model3, respectively. The number of free parameters can be reduced by setting parameters to be the same for all quantiles. The different choices of which parameters are the same for all quantiles result in different variants of the model with different numbers of free parameters. To find the optimal model variant, we trained different variants on the rmsf values obtained for cooling time spans 0.1–64 ns

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

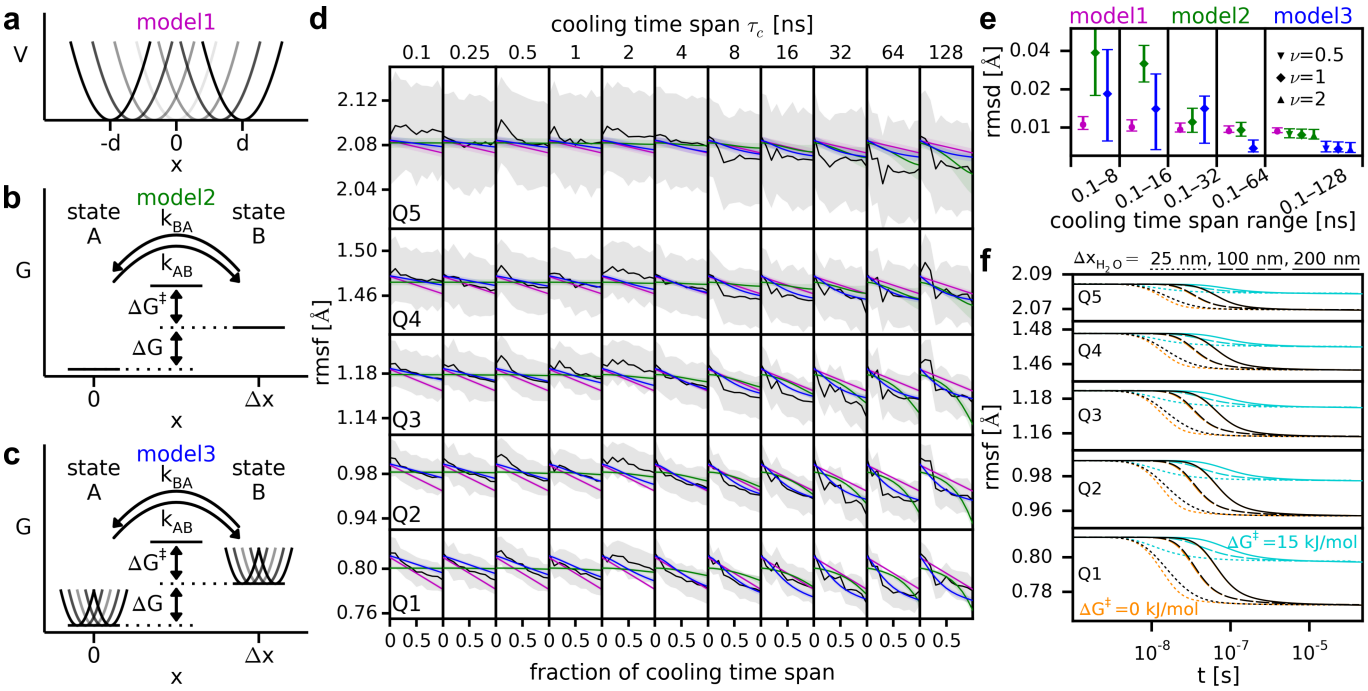

Figure 3: **Models of cooling behaviour.** Schematics of thermodynamic model1 with uniformly distributed harmonic potentials (**a**), kinetic two-state model2 (**b**), and combined model3 (**c**). (**d**) Comparison of the time dependence of the rmsf quantiles obtained from T-quench simulations after rescaling (black lines, compare Fig. 2c) with those obtained from the models (magenta, green, and blue lines). Standard deviations from simulations and the models are shown as grey and colored areas, respectively. (**e**) For each model, the root mean square deviation (rmsd) of the predicted rmsf values from the rmsf values obtained from the T-quench simulations (d, black lines) is shown. To obtain model parameters, the T-quench rmsf values for different ranges of cooling time spans were used (0.1–8 ns, 0.1–16 ns, ...). The symbols denote the mean values and the vertical lines correspond to the 95% confidence interval. (**f**) Rmsf decrease during plunge freezing as a function of time for the quantiles estimated from combining model3 with temperature drops (Fig. 1b) for different water-layer thicknesses ($\Delta x_{\mathrm{H_2O}}$).

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(Fig. S3). The obtained distributions of the parameters were then used to calculate distributions of rmsf values as a function of cooling time for all cooling rates. The deviation of these model-derived rmsf values from the rmsf values used for training shows how well the model is able to reproduce the rmsf decrease in the T-quench simulations (Fig. S3, blue). To control for overfitting, we used cross-validation. To that aim, we first predicted rmsf values for a cooling time span of 128 ns. Next, we calculated their deviation from the corresponding values from the T-quench simulations (Fig. S3, red). For model1, model2, and model3, we chose the variant that shows the lowest cross-validation rmsd for further analysis. When the rmsds of two variants were not significantly different, we chose the variant with fewer free parameters. The chosen model1 variant has 6 free parameters (5 for d and 1 for c), the chosen model2 variant has 11 free parameters (5 for $\Delta G^{\ddagger}$ and Δx, 1 for ΔG), and the chosen model3 variant has 9 parameters (5 for d and 1 for c, $\Delta G^{\ddagger}$, ΔG, Δx). Each model was then trained on the rmsf values of all T-quench simulations (cooling time spans 0.1–128 ns) and the convergence of the Metropolis sampling was assessed by comparing probability densities of two independent calculations, which turned out to be very similar (Fig. S4).

For the three models, the distributions of the parameters were then used to calculate the distributions of rmsf values as a function of cooling time (Fig. 3d, magenta, green and blue). Since model1 does not take the cooling rate into account and, therefore, the rmsf only depends on the temperature, the rmsf as a function of the fraction of the cooling time span is the same for all cooling time spans. This property results in an overestimation of the rmsf decrease with short cooling time spans (magenta line below black line) and an underestimation with longer cooling time spans (magenta line above black line).

The rmsf obtained from model2 is almost constant for short cooling time spans, because the temperature drops so rapidly that the barrier is not overcome and the probabilities of being in states A and B almost do not change. This behaviour also results in an underestimation of the rmsf decrease for rapid cooling and an overestimation for slower cooling.

These results indicate that model1 and model2 do not fully capture the underlying physical processes and therefore do not predict the decrease in heterogeneity for longer cooling time spans very well. In contrast, the combined model3 captures both, the rmsf decrease during rapid cooling and the kinetic effect that the decrease is more pronounced during slower cooling.

To test how well the models predict rmsf decrease for a large range of cooling rates based on the information from fast cooling rates only, we trained the models on the rmsf values obtained from T-quench simulations with short cooling time spans, e.g. the range 0.1 ns–8 ns. From the obtained parameter distributions, the rmsf values for all cooling time spans (0.1–128 ns) were predicted and compared to the rmsf values obtained from the corresponding T-quench simulations (Fig. 3e). For the thermodynamic model1, the deviation of the predicted values from the T-quench simulation values did not decrease markedly when rmsf values from longer cooling time spans were used to train the model. In contrast, for model2 and model3, the deviation decreased markedly when more simulation data was used for training. Model3 showed the lowest deviations of all models when cooling time spans ranges of 0.1–64 ns and 0.1–128 ns were used. The choice of the temperature exponent ν does not significantly affect the agreement (Fig. 3e, last column, green symbols) and in the following we will use $\nu = 1$. Interestingly, for model3, although the deviations are larger than for model2 when only short cooling time spans were used for training, model3 predicts large confidence intervals which contain small rmsd values. These results indicate that model3 accurately predicts the rmsf decrease and that more data for training primarily results in less uncertainty of the prediction.

One might argue that a combined model will always be better than the individual two models. However, the fact that the optimal model3 variant has fewer free parameters (9) than the optimal model2 variant (11) indicates that it is the underlying physics of the model and not the number of parameters that result in an improved prediction capability of model3. Taken together, our Bayes approach suggests that a combination of static structural heterogeneity and cooling-rate dependent barrier crossing best describes the narrowing of the structure ensembles observed in all T-quench simulations combined. Therefore, we will subsequently use model3 to describe the effects of cooling on structure ensembles.

Effects of plunge freezing on the ensemble

The main aim of this work is to quantify the effect on plunge-freezing on the heterogeneity of the frozen structural ensemble that is actually used for the cryo-EM measurements. To quantify the effect for realistic cooling protocols, we combined the estimated temperature drop during plunge-freezing (Fig. 1b) with the kinetic model that best describes the rmsf decrease during the T-quench simulations (Fig. 3c). Figure 3f shows the rmsf decrease as a function of the cooling time calculated using the kinetic model with the obtained

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

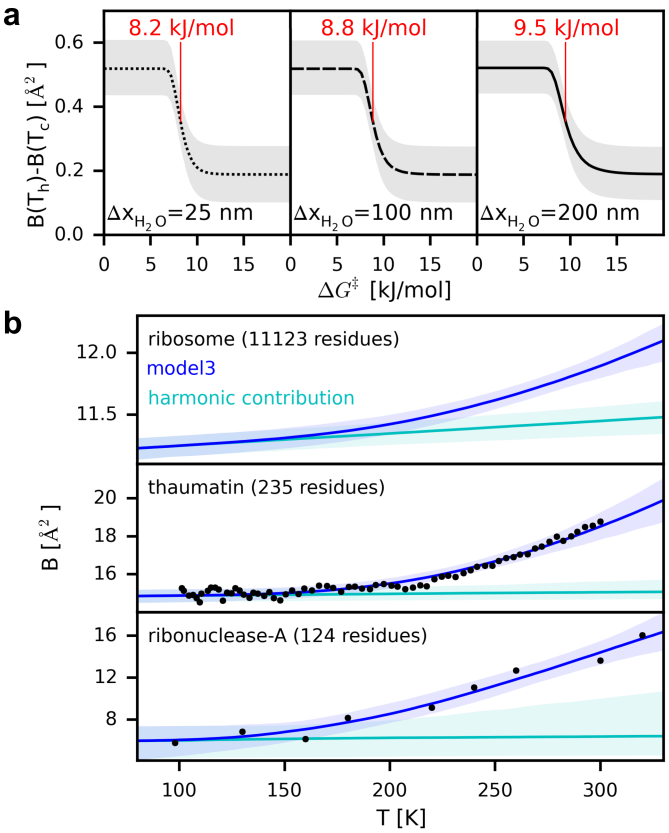

Figure 4: **B-factor dependency on barrier height and temperature.** (**a**) Difference between B-factors at the temperatures before cooling (T_h) and after cooling (T_c) calculated from model3 as a function of barrier height $\Delta G^{\ddagger}$ for different water-layer thicknesses ($\Delta x_{\mathrm{H_2O}}$). The $\Delta G^{\ddagger}$ values at which the B-factor is reduced halfway between the reduction at 0 kJ/mol and 20 kJ/mol is indicated in red. (**b**) B-factors as a function of temperature calculated from model3 under equilibrium conditions (upper panel; expected values: blue lines, 95 % confidence interval: light blue area). The B-factors of the harmonic potentials are shown in cyan. B-factors obtained from X-ray crystallography at different temperatures (black dots) for proteins thaumatin[55] (middle panel) and ribonuclease-A[35] (lower panel). B-factors calculated from model3 applied to the experimental B-factors (blue and cyan).

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

model parameters (Fig. S4c) and the temperature drops for three different water-layer thicknesses (Fig. 1b, 25 nm, 100 nm, and 200 nm). As expected, the thinner water layers which lead to more rapid temperature drops result in more rapid narrowing of the structural ensembles.

However, our model suggests that the final rmsf values are rather independent of the water-layer thickness within the range that we probed, because the barriers are small enough to be overcome during the cooling process such that the final rmsf values do not depend on the barrier height. The final rmsf values are decreased to 95.2% (standard deviation: 0.4%) of their value before cooling for Q1 and to 99.30% (0.06%) for Q5. To put these results into the context of experimental results, we use B-factors, sometimes called temperature factors, atomic displacement parameters or Debye-Waller factors another measure of the displacement of atoms around their mean position that can be obtained from can be obtained from x-ray and neutron scattering data and directly relate to the rsmf, $B = 8\pi/3 \cdot \text{rmsf}^2$ (ref.[66]). The calculated decrease in rmsf corresponds to a decrease in the B-factor of 0.52 Å^2 (standard deviation 0.04 Å^2) which is similar for each quantile.

In the T-quench simulations, only barriers that can be overcome during the atomistic cooling simulations contribute to the rmsf decrease and are therefore described by the model. Accordingly, a high free-energy barrier $\Delta G^{\ddagger}$ of 15 kJ/mol leads to a smaller decrease (Fig. 3f, cyan lines). To probe how the B-factor decrease depends on $\Delta G^{\ddagger}$ and the water-layer thickness, we applied model3 to the estimated plunge-freezing temperature drops with the parameters obtained from our Bayes approach (Fig. S4c) except for $\Delta G^{\ddagger}$ which was set to values between 0 and 20 kJ/mol. The B-factor decrease during cooling as a function of the barrier height $\Delta G^{\ddagger}$ shows a sharp transition between small and large barrier heights (Fig. 4a). The $\Delta G^{\ddagger}$ value where the transition occurs is highlighted in red for the very different cooling rates obtained for different water-layer thicknesses and is shifted to slightly larger values for thicker water layers. This shift to larger barrier heights is expected, because with thicker water layers the temperature drop is slower which results in more time to overcome larger barriers.

In summary, from our analysis, we expect that in addition to the equilibration of atoms in local harmonic potentials, conformational changes can contribute to the decrease in structural heterogeneity during plunge-freezing. For conformational changes with a free-energy barrier below a certain threshold, the barrier can be overcome during cooling and the conformation with lower free energy is predominately occupied, therefore decreasing the structural heterogeneity. A thicker water layer results in slower cooling such that higher barriers can be overcome leading to a more homogeneous ensemble.

Kinetic cooling model shows glass transition

The dynamic behaviour of proteins undergoes a glass transition at temperatures of around 200 K[56]. Above the transition temperature which has been found to be similar for different proteins[29;31;37;41;42;43;35;44;45;46;38;56;48;57], the slope of the averaged B-factor as a function of temperature is markedly higher than below the transition temperature. This observation provides an opportunity to test if our model3 (Fig. 3c) reproduces this ubiquitous behaviour. To that aim, we calculated the B-factors from the model with the parameters obtained from the Bayes approach (Fig. S4c) and compared them to B-factors from x-ray crystallography. Since the temperature drop for crystals is much slower than for thin water layers[50], we assumed the two states to be in equilibrium at any temperature, i.e. the probability of being in state A or B is determined only by the free-energy difference ΔG and the temperature T. The B-factors as a function of temperature calculated from the model for the Q3 quantile are shown in Fig. 4b (upper panel, blue line) and for the other quantiles in Fig. S5a. Additionally, we show the contribution of the harmonic potentials (cyan line) by setting parameters ΔG and Δx to 0 kJ/mol and 0 Å, respectively. At low temperatures, the probability of being in the state with higher free energy (state B) is negligible. Therefore the increase of the B-factors is governed by the harmonic potentials as indicated by the overlap of the B-factors from the harmonic contribution (cyan) and from the full model (blue). Around a temperature of 200 K, our model predicts a markedly steeper increase of the B-factors which is in agreement with the experimentally observed protein glass transition temperatures and results from the non-harmonic contributions. For higher temperatures, the occupancy of conformational states with higher free energies is increased (state B in our model) giving rise to a steeper B-factor increase.

The temperature dependency of the B-factors predicted by our model trained on the ribosome T-quench simulations is similar to that of the B-factors obtained from X-ray crystallography of proteins thaumatin[55] and ribonuclease-A[35] (Fig. 4b, black dots). However, the B-factors at high temperatures as well as the B-factor difference between high and low temperatures are different for the three cases as expected for macromolecules with different sizes and structures. In addition, the finding that these values differ for the

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

two protein crystal structures suggest that the differences for the three cases are not a result of the different methods used to obtain these values, but rather that they are specific for the chosen macromolecules. To further compare the B-factor dependency of the three macromolecules, we trained the equilibrium model3 on the x-ray B-factors (Fig. 4b). Interestingly, the probability densities of the model parameters Δx, ΔG, and c for model3 trained on the T-quench simulations and on the two x-ray structures overlap. These overlaps show that the same values of these parameters can describe the B-factor dependencies for these very different macromolecules. Only the probability densities of the parameter d, which determines the absolute B-factor at low temperatures, do not overlap. Taken together, these results further suggest that our simple model captures the effects of cooling on the B-factors of a broad range of biomolecules.

Discussion

In addition to achieving high resolution, the recent revolution in cryo-EM facilitated the determination of ensembles of structures in different states[18;67;19;20;21]. These structural ensembles are obtained from EM images of samples that were rapidly cooled down to cryogenic temperatures by plunge-freezing[13]. Biomolecules can thermodynamically access more conformations at room temperature than at the cryogenic temperature and rates of conformational changes are determined by free-energy barriers and temperatures[22]. For very rapid cooling, low temperatures that prevent barrier crossing are quickly reached and the conformations prior to cooling are expected to be kinetically trapped. In contrast, during very slow cooling, most barriers are expected to be overcome and low free-energy conformations are increasingly populated resulting in a narrower ensemble that is more homogeneous such that the room-temperature ensemble is preserved. However, where on the spectrum between the two extremes typical plunge-freezing experiments are actually located and, accordingly, if an how much the rapid cooling perturbs the ensembles is currently unknown.

Here, we address this question and quantify the effects of plunge-freezing on structural ensembles of biomolecules by combining continuum model calculations of the temperature drop during plunge-freezing, extensive MD simulations of the bacterial ribosome before and during cooling, and several kinetic models. Our results suggest that three effects contribute to the narrowing of the structural ensembles, resulting in smaller B-factors: first, the thermal contraction of the biomolecules, second, the equilibration within local potential wells, and third, the equilibration into lower free-energy conformations by overcoming the barriers that separate the conformations. The first two effects appear to be largely cooling-rate independent in contrast to the third effect where the cooling rate determines the barrier heights that are overcome during cooling. Our combined approach suggests that this barrier-height threshold is between 8 and 10 kJ/mol for realistic cooling rates and that the B-factor is reduced by 0.5 $Å^2$ in the ensemble imaged by cryo-EM compared to the ambient-temperature ensemble.

Our approach is divided into three major steps. First, we used a continuum model to estimate the temperature drop during plunge-freezing with different experimentally observed water-layer thicknesses. Second, we trained and tested kinetic models on T-quench MD simulations of a ribosome complex with different cooling rates. Finally, we applied the kinetic model which reproduced and predicted the T-quench results best to the plunge-freezing temperature drop to predict the narrowing of structural ensembles.

In our continuum model calculation of the water-layer temperature drop during plunge-freezing, we assumed that the water layer comes into contact with the liquid ethane instantaneously and the heat exchange is via conduction only. Our calculations yielded cooling rates consistent with cooling rates required for vitrification ($>10^6$ K/s)[13]. By high-speed photographic imaging it was observed that the grid, when plunged into the liquid ethane, initially displaces the ethane and only a fraction of the circumference of the grid comes into direct contact with the ethane immediately[68]. From this observation it was concluded that vitrification might instead be achieved by heat conduction through the grid. This mechanism might result in cooling rates different from our estimated rates but also not slower than the vitrification limit of 10^6 K/s. We tested this possibility with our continuum model applied to a water layer with a width of 3 mm (corresponding to the cryo-EM grid diameter) with a liquid-ethane layer on one side and air on the other side and found that after one second, temperatures below 200 K are only reached within a distance of 0.2 mm from the circumference. Within this time the grid would be fully submerged in liquid ethane, such that, except for a small region close the circumference of the grid, the heat conduction through the grid can not markedly contribute to the cooling. This result (data not shown) suggests that our assumption that the cooling is dominated by direct contact between the water-layer and the ethane is likely to be valid.

We are not aware of any attempt to directly measure the temperature drop during plunge-freezing. Should new data become available, any measured cooling rate could be included within our kinetic model in

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

a straightforward way, thus refining our estimate of how plunge-freezing affects structural ensembles.

To disentangle the contributions of thermal contraction, equilibration within local potential wells, and equilibration into lower free-energy conformations to the narrowing of structural ensembles, we performed T-quench MD simulations with different cooling rates and subsequently trained and cross-validated thermodynamic and kinetic models. For the T-quench simulations, we used cooling rates that are realistic for the onset of cooling, where the temperature is still high enough to enable conformational changes, and therefore the expected effect on the structural ensemble is largest. Combined with the kinetic models, this approach allowed us to overcome the problem of simulating the slow cooling 'tails' at low temperatures, which would be computationally intractable.

Our approach permitted to separate the narrowing of structural ensembles from simple thermal expansion. Indeed, the thermal expansion coefficient obtained from the T-quench simulations of a ribosome complex seems to be largely independent of the cooling rate and agrees with coefficients obtained from x-ray crystallography and MD simulations of proteins[35;49]. This observation now allows one to quantify the contribution of thermal contraction to B-factors and to determine the size of biomolecules at room temperature from the cryo-EM structures of different types of biomolecules.

To quantify if and how much the two effects, equilibration within local potential wells and the equilibration into lower free-energy conformations by overcoming the barriers, contribute to the narrowing of structural ensembles, we used the results of the T-quench simulations to train and validate three different models of the cooling process. We found that the model which combines both effects matches and predicts the T-quench simulations significantly better than the two models which considers only one of the two effects. The combined model allowed us to disentangle the contributions of the two effects and to address the question of how different temperature drops and different barrier heights affect the B-factor decrease (Fig. 3f, Fig. 4a). The observation that our model predicts a glass transition temperature within the range found with very different methods for other biomolecules[29;31;37;41;42;43;35;44;45;46;38;56;48;57] provides an independent confirmation of our approach.

In cryo-EM experiments, the B-factors result from several effects. Apart from thermal motion of the atoms, kinetic trapping in different conformations during cooling and the thermal contraction, also detector resolution, inaccurate sorting and alignment errors contribute to the obtained B-factors. Obtaining cryo-EM reconstructions from biomolecules that were cooled at different cooling rates would allow to quantify which contribution arises from the kinetic trapping and to determine the distribution of free-energy barrier heights. Further, cryo-EM reconstructions from biomolecules that were cooled down in the same way, but imaged at different temperatures (below the glass transition temperature), would allow to quantify the contribution of thermal motion. Effects that are independent of both, the cooling rate and the temperature of imaging, could then be attributed to the other effects.

To this aim, slower cooling rates could be achieved by increasing the temperature that the cryogen is kept at. However, the feasible cryogen temperature range is limited because the temperature as to be low enough to prevent boiling of the cryogen which would result in insulation of the sample from to cryogen via the Leidenfrost effect. Other cryogenic liquids with increased thermal diffusivities would result in a more rapid temperature drop. A promising way to alter the cooling rate in a controlled manner is provided by microfluidic cryofixation[69]. The geometry of the microfluidic channel determines the cooling rate such that different cooling rates should be achievable.

Our results bear several implications for ways to optimize and interpret cryo-EM results. To obtain high-resolution structures, where a homogeneous ensemble of structures in required, one would start cooling from a low temperature ensemble (above the melting point of water) and then cool down as slow as the vitrification limit allows. In contrast, to achieve the best representation of a room-temperature ensemble, one would cool down as rapidly as possible to avoid relaxation into nearby free-energy minima. Our observed reduction in B-factors can be used to rescale measures B-factors and thus to obtain better estimates for room temperature ensembles. This simple scaling procedure rests of course on the assumption — supported by the glass transition studies mentioned above — that the scaling factor is similar for different proteins and protein complexes. Application of our combined approach, including the atomistic simulations, to any molecular system of interest also allows one to drop or test this assumption.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Methods

Cooling rate estimation

To estimate the temperature drop of a film of water plunged into liquid ethane, we numerically solved the heat equation for a layer of water starting at T=277.15 K between two layers of ethane starting at T=90 K. The temperature at the outer boundaries was fixed to $T_b = 90$ K. To characterize the effect of the thicknesses of the layers, we used water-layer thicknesses of 25 nm, 100 nm, and 200 nm as well as ethane layer thicknesses of 100 nm, 200 nm, 400 nm, 800 nm, 1.6 μm, and 3.2 μm. For ethane, we used the thermal diffusivity $\alpha = 0.173$ mm^{2}s^{-1} calculated from the thermal conductivity, the specific heat capacity, and the density at a temperature of $T = 90.35$ K[70]. For water, we used $\alpha = 0.133$ mm^{2}s^{-1} calculated from the thermal conductivity[71], the specific heat capacity[72], and the density[71] at a temperature of $T = 273.15$ K. The heat equation is $\partial_t T(x,t) = \partial_x \left[\alpha(x)\partial_x T(x,t)\right]$, where x is the 1-d coordinate and $\alpha(x)$ is the thermal diffusivity at position x. This equation was solved numerically and more details can be found in the Supplementary Methods (*Numerical solution of the heat equation for a water layer*).

Structural ensemble before cooling

To generate an ensemble of structures of the ribosome·EF-Tu·kirromycin complex at the temperature before cooling (277.15 K), we used all-atom explicit-solvent molecular dynamics (MD) simulations starting from a high-resolution cryo-EM structure[25]. We used the same system setup and pre-equilibration protocol as described earlier for MD simulations of the same system[64], except that the temperature was set to 277.15 K and that the production simulation was extended to 3.5μs. Briefly, the simulations were carried out using GROMACS 5.1[73], with the amber99sb force field[74], the SPC/E water model[75], and K^+Cl^- ion parameters by Joung et al.[76]. Bond lengths were constrained using the LINCS algorithm[77]. Virtual site constraints for hydrogens[78] allowed a 4-fs integration step. For more details, refer to Warias et al.[64].

Convergence with regards to structural heterogeneity

We used the root mean square fluctuation (rmsf) of atomic positions calculated for an ensemble of structures as a measure for the structural heterogeneity of the ensemble. To assess convergence of the 277.15-K simulation with regards to this observable, we first extracted structures from the trajectory at intervals of 50 ns. The extracted structures were aligned via rigid-body fitting using the Cα-atoms of the amino acids and the P-atoms of the nucleotides. Then we grouped all 41 structures between 0 μs and 2 μs into one ensemble (Fig. S1a, top) and calculated the rmsf for all atoms resolved in the cryo-EM structure[25] (Fig. S1b). This calculation was repeated for the ensembles consisting of structures between 0.1 μs and 2.1 μs, between 0.2 μs and 2.2 μs, up to between 1.5 μs and 3.5 μs.

T-quench simulations

To estimate the effect of different cooling rates on the ensemble, we carried out MD simulations with different cooling time spans $\tau_c \in [0.1\text{ns}, 0.25\text{ns}, 0.5\text{ns}, 1\text{ns}, 2\text{ns}, 4\text{ns}, 8\text{ns}, 16\text{ns}, 32\text{ns}, 64\text{ns}, 128\text{ns}]$. These T-quench simulations were started from 41 structures extracted from the 277.15-K trajectory at 1000 ns, 1050 ns, ..., 3000 ns (Fig. 2a). For each cooling time span τ_c, 41 simulations of length τ_c were carried out with the temperature linearly decreasing from 277.15 K to 77 K. The temperatures of solute and solvent were controlled independently using velocity rescaling[79] with a coupling constant of τ_T=0.1 ps.

Structural heterogeneity during T-quench simulations

To quantify the effects of cooling on the structural heterogeneity of the simulated ensemble, we calculated rmsf values for each atom in the ensembles before, during, and at the end of cooling. To that aim, we first extracted the coordinates of all atoms of the ribosome complex which were resolved in the cryo-EM structure from the frames of the trajectories obtained from the T-quench simulations. For each cooling time span τ_c, the 41 T-quench trajectories were analysed as an approximation to the non-equilibrium ensemble. From each trajectory, the coordinates were extracted at 11 time points t_j with $t_j \in [0\text{ns}, 0.1\tau_c, 0.2\tau_c, \ldots, \tau_c]$. Subsequently, for each cooling time span τ_c and each time point t_j, we then grouped the 41 structures from

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the 41 simulations in an ensemble. Next, for each ensemble the structures were aligned via rigid-body fitting using the Cα- and the P-atoms. Then for each atom the rmsf was calculated (Fig. 2b).

For each cooling time span τ_c, the rmsf values for all N atoms are collected in the matrix

$$\boldsymbol{R}_{\tau_c} = \begin{bmatrix} r_{1,\tau_c}(0\text{ns}) & r_{1,\tau_c}(0.1\tau_c) & r_{1,\tau_c}(0.2\tau_c) & \cdots & r_{1,\tau_c}(\tau_c) \\ r_{2,\tau_c}(0\text{ns}) & r_{2,\tau_c}(0.1\tau_c) & r_{2,\tau_c}(0.2\tau_c) & \cdots & r_{2,\tau_c}(\tau_c) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{N,\tau_c}(0\text{ns}) & r_{N,\tau_c}(0.1\tau_c) & r_{N,\tau_c}(0.2\tau_c) & \cdots & r_{N,\tau_c}(\tau_c) \end{bmatrix}.$$

Each row of $\boldsymbol{R}_{\tau_c}$ contains all rmsf values of a certain time point t_j for a certain cooling time span τ_c. To describe the distributions of rmsf values and to quantify changes during cooling, we calculated the 6-quantiles, i.e. the 5 rmsf values that divide the set of atoms into 6 equally sized subsets of atoms. The quantiles are collected in the matrix

$$\boldsymbol{Q}_{\tau_c} = \begin{bmatrix} Q_{1,\tau_c}(0\text{ns}) & Q_{1,\tau_c}(0.1\tau_c) & Q_{1,\tau_c}(0.2\tau_c) & \cdots & Q_{1,\tau_c}(\tau_c) \\ Q_{2,\tau_c}(0\text{ns}) & Q_{2,\tau_c}(0.1\tau_c) & Q_{2,\tau_c}(0.2\tau_c) & \cdots & Q_{2,\tau_c}(\tau_c) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ Q_{5,\tau_c}(0\text{ns}) & Q_{N,\tau_c}(0.1\tau_c) & Q_{N,\tau_c}(0.2\tau_c) & \cdots & Q_{N,\tau_c}(\tau_c) \end{bmatrix},$$

where $Q_{i,\tau_c}(t_j)$ is the i-th quantile of the distribution of rmsf values at time point t_j for cooling time span τ_c. To estimate the uncertainty of the 6-quantiles, we carried out bootstrapping of the structures in each ensemble resulting in the matrices $\boldsymbol{S}_{\tau_c}$ containing the standard deviations $\sigma_{i,\tau_c}(t_j)$ of each of the $Q_{i,\tau_c}(t_j)$ value.

Size decrease during cooling

We observed a decrease in the size of the ribosome over the course of the cooling simulations. To quantify the decrease, we introduced a scaling factor s that is multiplied to all the extracted coordinates after placing the center of geometry in the origin. For each t_j >0 ns and each τ_c, we calculated the scaling factor s that minimizes the average root mean square deviation (rmsd) of all 41 structures from to their respective starting structure (t_j=0 ns). Fig. S2 shows the scaling factor s for all t_j and τ_c. To disentangle the effect of size decrease from the reduced structural heterogeneity, we then calculated the rmsf quantiles $\boldsymbol{Q}_{\tau_c}$ and standard deviations $\boldsymbol{S}_{\tau_c}$ from the scaled structures and used those in the following analysis (Fig. 2c).

Thermodynamic and kinetic models of the cooling process

To check if the decrease in conformational heterogeneity during cooling depends on the rate of cooling, we considered one thermodynamic and two kinetic models of the cooling process. The first model describes the rmsf of atoms in uniformly distributed harmonic potentials in equilibrium and is therefore independent of the cooling rate. The second model describes the rmsf of atoms by jumps between two states at different free energies separated by a free-energy barrier. The third model is a combination of the first two models.

Model1: uniformly distributed harmonic potentials

The probability density $p_h(a,b,x)$ for an atom in harmonic potentials whose mean values are uniformly distributed between a and b is given by

$$p_h(a,b,x) = \int_a^b \frac{1}{b-a} \cdot \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) d\mu \quad (1)$$

$$= \frac{1}{2(a-b)} \left[\text{erf}\left(\frac{x-b}{\sigma\sqrt{2}}\right) - \text{erf}\left(\frac{x-a}{\sigma\sqrt{2}}\right)\right] \quad (2)$$

with $\sigma = \sqrt{\frac{k_B T}{c}}$, where k_B is the Boltzmann constant, and c is the force constant of the harmonic potentials. The rmsf of the atom is then given by

$$\text{rmsf}_h(a,b,\bar{x}) = \sqrt{\int_{-\infty}^{\infty} (x-\bar{x})^2 p_h(a,b,x) dx} \quad (3)$$

$$= \sqrt{\sigma^2 + \frac{a^2+ab+b^2}{3} - (a+b)\bar{x} + \bar{x}^2}, \quad (4)$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

where $\bar{x}$ is the expectation value of x. For model1, we set a to $-d$ and b to $+d$ (compare Fig. 3a). Therefore, $\bar{x} = 0$ and

$$\text{rmsf}_{\text{model1}} = \text{rmsf}_h(-d, d, 0) = \sqrt{\frac{k_B T}{c} + \frac{d^2}{3}}. \tag{5}$$

Model2: kinetic two-state model

In the two-state model, each atom can visit two states A and B that are separated by a distance Δx along a 1-d coordinate x (Fig. 3b). The free energy of state B is chosen to be larger than that of state A by ΔG. Additionally, the two states are separated by a free-energy barrier $\Delta G^{\ddagger}$.

Under the assumption that the two states are in equilibrium at temperature T_h before cooling, the probability of being in states A and B is $P_A = 1/(1 + \exp(-\Delta G/(k_b T_h)))$ and $P_B = 1/(1 + \exp(\Delta G/(k_b T_h)))$, respectively. For the non-equilibrium time evolution during cooling, the rates between the two states contribute to the probabilities and are given by a modified Arrhenius equation,

$$\begin{aligned} k_{AB}(t) &= \kappa \left(\frac{T(t)}{T_h}\right)^{\nu} \exp\left(-\frac{\Delta G + \Delta G^{\ddagger}}{k_b T(t)}\right), \\ k_{BA}(t) &= \kappa \left(\frac{T(t)}{T_h}\right)^{\nu} \exp\left(-\frac{\Delta G^{\ddagger}}{k_b T(t)}\right), \end{aligned}$$

where $\kappa \left(\frac{T(t)}{T_h}\right)^{\nu}$ is the pre-exponential factor with temperature T depending on time t. ν is an exponent of the temperature and κ a scaling factor. In our T-quench simulations the temperature was linearly decreased from T_h at $t = 0$ to T_c at $t = \tau_c$. Hence, $T(t) = T_h - \alpha t$ with $\alpha = (T_h - T_c)/\tau_c$. The Master equation for the probability of being in state A as a function of time $P_A(t)$ is then

$$\frac{dP_A(t)}{dt} = k_{BA}(t)\,(1 - P_A(t)) - k_{AB}(t) P_A(t).$$

Inserting the rates $k_{AB}(t)$ and $k_{BA}(t)$ results in the differential equation

$$\frac{dP_A(t)}{dt} = \kappa \left(\frac{T(t)}{T_h}\right)^{\nu} \left[\exp\left(-\frac{\Delta G^{\ddagger}}{k_b(T_h - \alpha t)}\right)(1 - P_A(t)) - \exp\left(-\frac{\Delta G + \Delta G^{\ddagger}}{k_b(T_h - \alpha t)}\right) P_A(t)\right]. \tag{6}$$

For given values of κ, ΔG, $\Delta G^{\ddagger}$, and ν, the differential equation for $P_A(t)$ was solved numerically using the function odeint of SciPy[80]. The rmsf of an atom described by the model as a function of time can then be calculated by

$$\text{rmsf}_{\text{model2}}(t) = \Delta x \sqrt{P_A(t)(1 - P_A(t))}. \tag{7}$$

Model3: kinetic two-state model with uniformly distributed harmonic potentials

To capture possible kinetics during cooling and to include the uniformly distributed harmonic potentials, we combined the two models described above (Fig. 3c). Here, as in the model2, $P_A(t)$ and $P_B(t) = 1 - P_A(t)$ are the probabilities of the atom being in in states A and B obtained from numerically solving equation 6. Extending the two-state model, the probability in states A and B is governed by harmonic potentials uniformly distributed from $-d$ to d and from $\Delta x - d$ to $\Delta x + d$, respectively. The probability density as a function of time t is then given by $p(x,t) = P_A(t) p_h(-d, d, x) + P_B(t) p_h(\Delta x - d, \Delta x + d, x)$, where p_h is given by equation 2. The probability density results in a mean position of the atom $\bar{x}(t) = P_B(t)\Delta x$. The rmsf is then obtained from

$$\text{rmsf}_{\text{model3}}(t) = \sqrt{\int_{\infty}^{\infty} (x - \bar{x}(t))^2\, p(x,t) dx} = \sqrt{P_A(t) r_A(t) + P_B(t) r_B(t)} \tag{8}$$

with

$$\begin{aligned} r_A(t) &= \int_{\infty}^{\infty} (x - \bar{x}(t))^2 p_h(-d, d, \bar{x}(t)) dx \\ &= [\text{rmsf}_h(-d, d, \bar{x}(t))]^2 \end{aligned}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and

$$\begin{aligned} r_B(t) &= \int_{\infty}^{\infty} (x - \bar{x}(t))^2 p_h(\Delta x - d, \Delta x + d, \bar{x}(t)) dx \\ &= [\mathrm{rmsf}_h(\Delta x - d, \Delta x + d, \bar{x}(t))]^2, \end{aligned}$$

where rmsf_h is given by equation 4.

Finding the parameters of the models

The three models of the cooling process have different sets of parameters c and d for model1, Δx, ΔG, and $\Delta G^{\ddagger}$ for model2 and all five parameters for the combined model3. We considered different variants of the models where each parameter can either be the same value for all five quantiles or one value for each quantile. To obtain probability densities of the parameters, we used a Bayes approach; the details can be found in the Supplementary Methods (*Metropolis sampling with Bayesian inference*).

To obtain the optimal combination of numbers of parameters, we trained models with different numbers of parameters (Fig. S3) and compared how well they reproduced the rmsf values they were trained on and how well they predict rmsf values not included in the training; the details can be found in the Supplementary Methods (*Finding optimal model variants*).

Decrease of structural heterogeneity during plunge-freezing

To estimate the decrease in structural heterogeneity during the plunge-freezing, we applied model3 with the obtained parameters (Fig. S4c) to the temperature drops estimated for different water-layer thicknesses (Fig. 1b). After omitting the first 20% of steps, from 1000 randomly chosen Metropolis steps, parameters were extracted and 1000 sets of rmsf curves were calculated from the model. To investigate the influence of the barrier height $\Delta G^{\ddagger}$, from the same parameters rmsf curves were calculated after setting $\Delta G^{\ddagger}$ to 0 kJ/mol or to 15 kJ/mol. The median and 95% confidence intervals of the rmsf curves are shown in Fig. 3f.

B-factors as a function of temperature and barrier height

To compare the results of model3 to B-factors obtained from experiments, e.g. x-ray crystallography at different temperatures[35], we calculated B-factors from the model at different temperatures with the parameters from 1000 randomly chosen Metropolis steps (Fig. 4a, blue lines). The contribution of the harmonic potentials to the B-factors was calculated by evaluating $\mathrm{rmsf}_{\mathrm{model1}}$ (equation 5) with the chosen c and d parameters (Fig. S5a, cyan lines). Next, we trained model3 using the B-factor values for ribonuclease-A obtained from x-ray crystallography[35] to obtain probability densities for parameters Δx, ΔG, c, and d (Fig. S5b). Since cooling of crystals is much slower than for the thin water layers used in cryo-EM[50], we set the $\Delta G^{\ddagger}$ parameter to 0 kJ/mol such that there is no kinetic contribution. We calculated B-factor values as a function of temperature (Fig. 4b) as described above (Fig. 4a).

To study the effect of the barrier height $\Delta G^{\ddagger}$ on the reduction of B-factors, we applied model3 to the temperature drops (Fig. 1b) with 1000 sets of parameters chosen as described above. Here, however, we set the barrier height $\Delta G^{\ddagger}$ to values between 0 kJ/mol and 20 KJ/mol. The median and confidence intervals of the differences between the B-factors before cooling (T_h) and after cooling (T_c) are shown as a function of $\Delta G^{\ddagger}$ (Fig 4c).

Acknowledgements

The work was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2067/1- 390729940. Computer time has been provided by the Max Planck Computing and Data Facility and the Leibniz Rechenzentrum.

Author contributions

L.V.B. performed and analyzed the continuum model calculations, the MD simulations and the Metropolis sampling of the kinetic models. Both authors conceived the project, interpreted the results, and wrote the paper.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

[1] Bai, X.-c., McMullan, G. & Scheres, S. H. How cryo-EM is revolutionizing structural biology. *Trends in Biochemical Sciences* **40**, 49–57 (2015).

[2] Wu, S., Armache, J. P. & Cheng, Y. Single-particle cryo-EM data acquisition by using direct electron detection camera. *Microscopy (Oxford, England)* **65**, 35–41 (2016).

[3] Cheng, Y. Single-particle cryo-EM-How did it get here and where will it go. *Science* **361**, 876–880 (2018).

[4] Yip, K. M., Fischer, N., Paknia, E., Chari, A. & Stark, H. Atomic-resolution protein structure determination by cryo-EM. *Nature* **587**, 157–161 (2020).

[5] Nakane, T. *et al.* Single-particle cryo-EM at atomic resolution. *Nature* **587**, 152–156 (2020).

[6] Adrian, M., Dubochet, J., Lepault, J. & McDowall, A. W. Cryo-electron microscopy of viruses. *Nature* **308**, 32–36 (1984).

[7] Razinkov, I. *et al.* A new method for vitrifying samples for cryoEM. *Journal of Structural Biology* **195**, 190–198 (2016).

[8] Feng, X. *et al.* A Fast and Effective Microfluidic Spraying-Plunging Method for High-Resolution Single-Particle Cryo-EM. *Structure* **25**, 663–670.e3 (2017).

[9] Dandey, V. P. *et al.* Spotiton: New features and applications. *Journal of Structural Biology* **202**, 161–169 (2018).

[10] Wei, H. *et al.* Optimizing "self-wicking" nanowire grids. *Journal of Structural Biology* **202**, 170–174 (2018).

[11] Parisi, G. Physics of the glass transition. *Physica A: Statistical Mechanics and its Applications* **280**, 115–124 (2000).

[12] Debenedetti, P. G. & Stillinger, F. H. Supercooled liquids and the glass transition. *Nature* **410**, 259–267 (2001).

[13] Dubochet, J. *et al.* Cryo-electron microscopy of vitrified specimens. *Quarterly Reviews of Biophysics* **21**, 129 (1988).

[14] Tivol, W. F., Briegel, A. & Jensen, G. J. An improved cryogen for plunge freezing. *Microscopy and Microanalysis* **14**, 375–379 (2008).

[15] Baker, L. A. & Rubinstein, J. L. Radiation damage in electron cryomicroscopy. In *Methods in Enzymology*, vol. 481, 371–388 (Academic Press Inc., 2010).

[16] Nogales, E. & Scheres, S. H. Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. *Molecular Cell* **58**, 677–689 (2015).

[17] Fischer, N., Konevega, A. L., Wintermeyer, W., Rodnina, M. V. & Stark, H. Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy. *Nature* **466**, 329–33 (2010).

[18] Zhao, J., Benlekbir, S. & Rubinstein, J. L. Electron cryomicroscopy observation of rotational states in a eukaryotic V-ATPase. *Nature 2015 521:7551* **521**, 241–245 (2015).

[19] Petrychenko, V. *et al.* Structural mechanism of GTPase-powered ribosome-tRNA movement. *bioRxiv* 2021.06.01.446629 (2021).

[20] Carbone, C. E. *et al.* Time-resolved cryo-EM visualizes ribosomal translocation with EF-G and GTP. *bioRxiv* 2021.05.31.446434 (2021).

[21] Rundlet, E. J. *et al.* Structural basis of early translocation events on the ribosome. *Nature 2021* 1–5 (2021).

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[22] Frauenfelder, H., Petsko, G. A. & Tsernoglou, D. Temperature-dependent x-ray diffraction as a probe of protein structural dynamics. *Nature* **280**, 558–563 (1979).

[23] Moffat, K. & Henderson, R. Freeze trapping of reaction intermediates. *Current Opinion in Structural Biology* **5**, 656–663 (1995).

[24] Chen, C. Y., Chang, Y. C., Lin, B. L., Huang, C. H. & Tsai, M. D. Temperature-Resolved Cryo-EM Uncovers Structural Bases of Temperature-Dependent Enzyme Functions. *Journal of the American Chemical Society* **141**, 19983–19987 (2019).

[25] Fischer, N. *et al.* Structure of the E. coli ribosome–EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. *Nature* (2015).

[26] Berriman, J. & Unwin, N. Analysis of transient structures by cryo-microscopy combined with rapid mixing of spray droplets. *Ultramicroscopy* **56**, 241–252 (1994).

[27] Frank, J. Time-resolved cryo-electron microscopy: Recent progress. *Journal of Structural Biology* **200**, 303–306 (2017).

[28] Klebl, D. P. *et al.* Need for Speed: Examining Protein Behavior during CryoEM Grid Preparation at Different Timescales. *Structure* **28**, 1238–1248.e4 (2020).

[29] Parak, F. *et al.* Evidence for a correlation between the photoinduced electron transfer and dynamic properties of the chromatophore membranes from Rhodospirillum rubrum. *FEBS Letters* **117**, 368–372 (1980).

[30] Parak, F., Frolov, E. N., Mössbauer, R. L. & Goldanskii, V. I. Dynamics of metmyoglobin crystals investigated by nuclear gamma resonance absorption. *Journal of Molecular Biology* **145**, 825–833 (1981).

[31] Parak, F., Knapp, E. W. & Kucheida, D. Protein dynamics. Mössbauer spectroscopy on deoxymyoglobin crystals. *Journal of Molecular Biology* **161**, 177–194 (1982).

[32] Frolov, E. N., Gvosdev, R., Goldanskii, V. I. & Parak, F. G. Differences in the dynamics of oxidized and reduced cytochrome c measured by Mössbauer spectroscopy. *JBIC Journal of Biological Inorganic Chemistry* **2**, 710–713 (1997).

[33] Hartmann, H. *et al.* Conformational substates in a protein: structure and dynamics of metmyoglobin at 80 K. *Proceedings of the National Academy of Sciences of the United States of America* **79**, 4967–4971 (1982).

[34] Rasmussen, B. F., Stock, A. M., Ringe, D. & Petsko, G. A. Crystalline ribonuclease a loses function below the dynamical transition at 220 K. *Nature* **357**, 423–424 (1992).

[35] Tilton, R. F., Dewan, J. C. & Petsko, G. A. Effects of Temperature on Protein Structure and Dynamics: X-ray Crystallographic Studies of the Protein Ribonuclease-A at Nine Different Temperatures from 98 to 320 K. *Biochemistry* **31**, 2469–2481 (1992).

[36] Halle, B. Biomolecular cryocrystallography: Structural changes during flash-cooling. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 4793–4798 (2004).

[37] Doster, W., Cusack, S. & Petry, W. Dynamical transition of myoglobin revealed by inelastic neutron scattering. *Nature* **337**, 754–756 (1989).

[38] Réat, V. *et al.* Solvent dependence of dynamic transitions in protein solutions. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 9961–9966 (2000).

[39] Iben, I. E. *et al.* Glassy behavior of a protein. *Physical Review Letters* **62**, 1916–1919 (1989).

[40] Lee, A. L. & Wand, A. J. Microscopic origins of entropy, heat capacity and the glass transition in proteins. *Nature* **411**, 501–504 (2001).

[41] Wong, C. F., Zheng, C. & McCammon, J. A. Glass transition in SPC/E water and in a protein solution: A molecular dynamics simulation study. *Chemical Physics Letters* **154**, 151–154 (1989).

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[42] Smith, J., Kuczera, K. & Karplus, M. Dynamics of myoglobin: Comparison of simulation results with neutron scattering spectra. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 1601–1605 (1990).

[43] Loncharich, R. J. & Brooks, B. R. Temperature Dependence of dynamics of hydrated myoglobin: Comparison of force field calculations with neutron scattering data. *Journal of Molecular Biology* **215**, 439–455 (1990).

[44] Steinbach, P. J. & Brooks, B. R. Protein hydration elucidated by molecular dynamics simulation. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 9135–9139 (1993).

[45] Steinbach, P. J. & Brooks, B. R. Protein simulation below the glass-transition temperature. Dependence on cooling protocol. *Chemical Physics Letters* **226**, 447–452 (1994).

[46] Norberg, J. & Nilsson, L. Glass transition in DNA from molecular dynamics simulations. *Proceedings of the National Academy of Sciences of the United States of America* **93**, 10173–10176 (1996).

[47] Vitkup, D., Ringe, D., Petsko, G. A. & Karplus, M. Solvent mobility and the protein 'glass' transition. *Nature Structural Biology* **7**, 34–38 (2000).

[48] Tournier, A. L., Xu, J. & Smith, J. C. Translational hydration water dynamics drives the protein glass transition. *Biophysical Journal* **85**, 1871–1875 (2003).

[49] Mehra, R., Dehury, B. & Kepp, K. P. Cryo-temperature effects on membrane protein structure and dynamics. *Physical Chemistry Chemical Physics* **22**, 5427–5438 (2020).

[50] Teng, T. Y. & Moffat, K. Cooling Rates during Flash Cooling. *Journal of Applied Crystallography* **31**, 252–257 (1998).

[51] Kriminski, S., Kazmierczak, M. & Thorne, R. E. Heat transfer from protein crystals: Implications for flash-cooling and X-ray beam heating. *Acta Crystallographica - Section D Biological Crystallography* **59**, 697–708 (2003).

[52] Berejnov, V., Husseini, N. S., Alsaied, O. A. & Thorne, R. E. Effects of cryoprotectant concentration and cooling rate on vitrification of aqueous solutions. *Journal of Applied Crystallography* **39**, 244–251 (2006). 0604109.

[53] Kuntz, I. D., Brassfield, T. S., Law, G. D. & Purcell, G. V. Hydration of macromolecules. *Science* **163**, 1329–1331 (1969).

[54] Sartor, G., Hallbrucker, A., Hofer, K. & Mayer, E. Calorimetric glass-liquid transition and crystallization behavior of a vitreous, but freezable, water fraction in hydrated methemoglobin. *Journal of Physical Chemistry* **96**, 5133–5138 (1992).

[55] Warkentin, M. & Thorne, R. E. Slow cooling of protein crystals. *Journal of Applied Crystallography* **42**, 944–952 (2009).

[56] Ringe, D. & Petsko, G. A. The 'glass transition' in protein dynamics: What it is, why it occurs, and how to exploit it. *Biophysical Chemistry* **105**, 667–680 (2003).

[57] Doster, W. The protein-solvent glass transition. *Biochimica et Biophysica Acta - Proteins and Proteomics* **1804**, 3–14 (2010).

[58] Bock, L. V. *et al.* Energy barriers and driving forces in tRNA translocation through the ribosome. *Nature structural & molecular biology* **20**, 1390–6 (2013).

[59] Stagg, S. M. *et al.* Automated cryoEM data acquisition and analysis of 284 742 particles of GroEL. *Journal of Structural Biology* **155**, 470–481 (2006).

[60] Rice, W. J. *et al.* Routine determination of ice thickness for cryo-EM grids. *Journal of Structural Biology* **204**, 38–44 (2018).

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.08.463658; this version posted October 8, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[61] Noble, A. J. *et al.* Routine single particle CryoEM sample and grid characterization by tomography. *eLife* **7** (2018).

[62] Noble, A. J. *et al.* Reducing effects of particle adsorption to the air–water interface in cryo-EM. *Nature Methods* **15**, 793–795 (2018).

[63] Fraser, J. S., Lindorff-Larsen, K. & Bonomi, M. What Will Computational Modeling Approaches Have to Say in the Era of Atomistic Cryo-EM Data? *Journal of Chemical Information and Modeling* **60**, 2410–2412 (2020). `2002.00151`.

[64] Warias, M., Grubmüller, H. & Bock, L. V. tRNA Dissociation from EF-Tu after GTP Hydrolysis: Primary Steps and Antibiotic Inhibition. *Biophysical Journal* **118**, 151–161 (2020).

[65] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. & Teller, E. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics* **21**, 1087–1092 (1953).

[66] Sun, Z., Liu, Q., Qu, G., Feng, Y. & Reetz, M. T. Utility of B-Factors in Protein Science: Interpreting Rigidity, Flexibility, and Internal Motion and Engineering Thermostability. *Chemical Reviews* (2019).

[67] Bonomi, M. & Vendruscolo, M. Determination of protein structural ensembles using cryo-electron microscopy. *Current Opinion in Structural Biology* **56**, 37–45 (2019).

[68] Kasas, S., Dumas, G., Dietler, G., Catsicas, S. & Adrian, M. Vitrification of cryoelectron microscopy specimens revealed by high-speed photographic imaging. *Journal of Microscopy* **211**, 48–53 (2003).

[69] Mejia, Y. X., Feindt, H., Zhang, D., Steltenkamp, S. & Burg, T. P. Microfluidic cryofixation for correlative microscopy. *Lab on a chip* **14**, 3281–4 (2014).

[70] Younglove, B. A. & Ely, J. F. Thermophysical Properties of Fluids. II. Methane, Ethane, Propane, Isobutane, and Normal Butane. *Journal of Physical and Chemical Reference Data* **16**, 577–798 (1987).

[71] Sengers, J. V. & Watson, J. T. Improved International Formulations for the Viscosity and Thermal Conductivity of Water Substance. *Journal of Physical and Chemical Reference Data* **15**, 1291–1314 (1986).

[72] Marsh, K. N. & Marsh, K. N. *Recommended reference materials for the realization of physicochemical properties* (Blackwell Scientific Publications Oxford, 1987).

[73] Pronk, S. *et al.* Gromacs 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).

[74] Hornak, V. *et al.* Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins* **65**, 712–725 (2006).

[75] Poole, P. H., Sciortino, F., Essmann, U. & Stanley, H. E. Phase behaviour of metastable water. *Nature* **360**, 324 (1992).

[76] Joung, I. S. & Cheatham III, T. E. Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. *J. Phys. Chem. B* **112**, 9020–9041 (2008).

[77] Hess, B. P-LINCS: A parallel linear constraint solver for molecular simulation. *J. Chem. Theory Comput.* **4**, 116–122 (2008).

[78] Feenstra, K. A., Hess, B. & Berendsen, H. J. C. Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. *J. Comput. Chem.* **20**, 786–798 (1999).

[79] Bussi, G., Donadio, D. & Parrinello, M. Canonical sampling through velocity rescaling. *Journal of Chemical Physics* **126** (2007). `arXiv:0803.4060v1`.

[80] Virtanen, P. *et al.* SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods* **17**, 261–272 (2020).