Biochemistry

pubs.acs.org/biochemistry Article

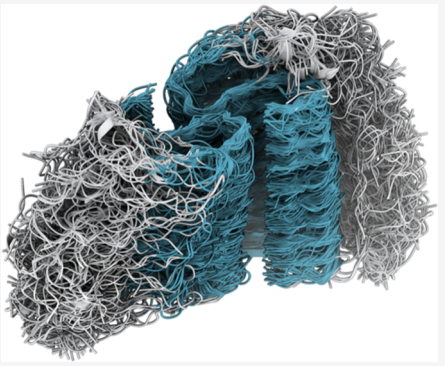
Determination of the Structure and Dynamics of the Fuzzy Coat of an Amyloid Fibril of IAPP Using Cryo-Electron Microscopy

Z. Faidon Brotzakis,[§] Thomas Löhr,[§] Steven Truong, Samuel Hoff, Massimiliano Bonomi, and Michele Vendruscolo*

Cite This: *Biochemistry* 2023, 62, 2407−2416 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: In recent years, major advances in cryo-electron microscopy (cryo-EM) have enabled the routine determination of complex biomolecular structures at atomistic resolution. An open challenge for this approach, however, concerns large systems that exhibit continuous dynamics. To address this problem, we developed the metadynamic electron microscopy metainference (MEMMI) method, which incorporates metadynamics, an enhanced conformational sampling approach, into the metainference method of integrative structural biology. MEMMI enables the simultaneous determination of the structure and dynamics of large heterogeneous systems by combining cryo-EM density maps with prior information through molecular dynamics, while at the same time modeling the different sources of error. To illustrate the method, we apply it to elucidate the dynamics of an amyloid fibril of the islet amyloid polypeptide (IAPP). The resulting conformational ensemble provides an accurate description of the structural variability of the disordered region of the amyloid fibril, known as fuzzy coat. The conformational ensemble also reveals that in nearly half of the structural core of this amyloid fibril, the side chains exhibit liquid-like dynamics despite the presence of the highly ordered network backbone of hydrogen bonds characteristic of the cross-β structure of amyloid fibrils.

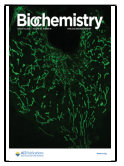


INTRODUCTION

In the last several years, cryo-electron microscopy (cryo-EM) has been pushing the boundaries of structural biology in terms of structural resolution, system complexity, and macromolecular size.[1,2] Imaging single particles by rapid cryo-cooling and vitrification enables structural studies capturing near-native conformations, while offering sample protection from beam radiation.[3−5] Technical advances in electron detectors, computational algorithms accounting for beam-induced motion, and automation of data collection and image analysis have paved the way for a spectacular increase in the resolution of cryo-EM density maps.[6,7] The Electron Microscopy Data Bank (EMDB) currently holds 15,800 single-particle cryo-EM density maps, which offer exquisitely detailed structural information about macromolecular systems of central importance in cell biology.[8]

In standard cryo-EM structure determination, two-dimensional (2D) images of single particles are first classified in conformationally homogeneous classes and then averaged in a computational image processing step, thereby leading to a substantial increase in the signal-to-noise ratio, and thus in structure resolution.[9] However, the continuous dynamics of flexible regions are difficult to detect, therefore complicating the generation of homogenous classes of structures. The resulting low densities cannot be readily used to determine atomistic structures, and are thus often excluded from the final structural model. While methods such as the manifold embedding approach[10] can determine multiple structures from cryo-EM density maps, to account for the conformational dynamics, one should quantitatively and atomistically interpret the cryo-EM density maps as an envelope that corresponds to an averaged conformational ensemble of states with certain populations that interconvert with a characteristic timescale.[10] Such a viewpoint moves away from a single-structure interpretation of the data and links the data to the statistical mechanics concept of free energy landscapes of conformational ensembles.[11,12] Integrative structural ensemble-modeling methods incorporate experimental information into molecular simulations and enable the determination of structural ensembles that maximally conform to the experimental data with atomistic resolution.[13−24] This technique has been applied using nuclear magnetic resonance (NMR) spectroscopy,[18,20,21,25−29] fluorescence resonance energy transfer (FRET) microscopy,[30] small-angle scattering techniques (SAXS/

Received: January 10, 2023
Revised: June 3, 2023
Published: July 21, 2023



ACS Publications

© 2023 The Authors. Published by American Chemical Society

https://doi.org/10.1021/acs.biochem.3c00010
Biochemistry 2023, 62, 2407−2416

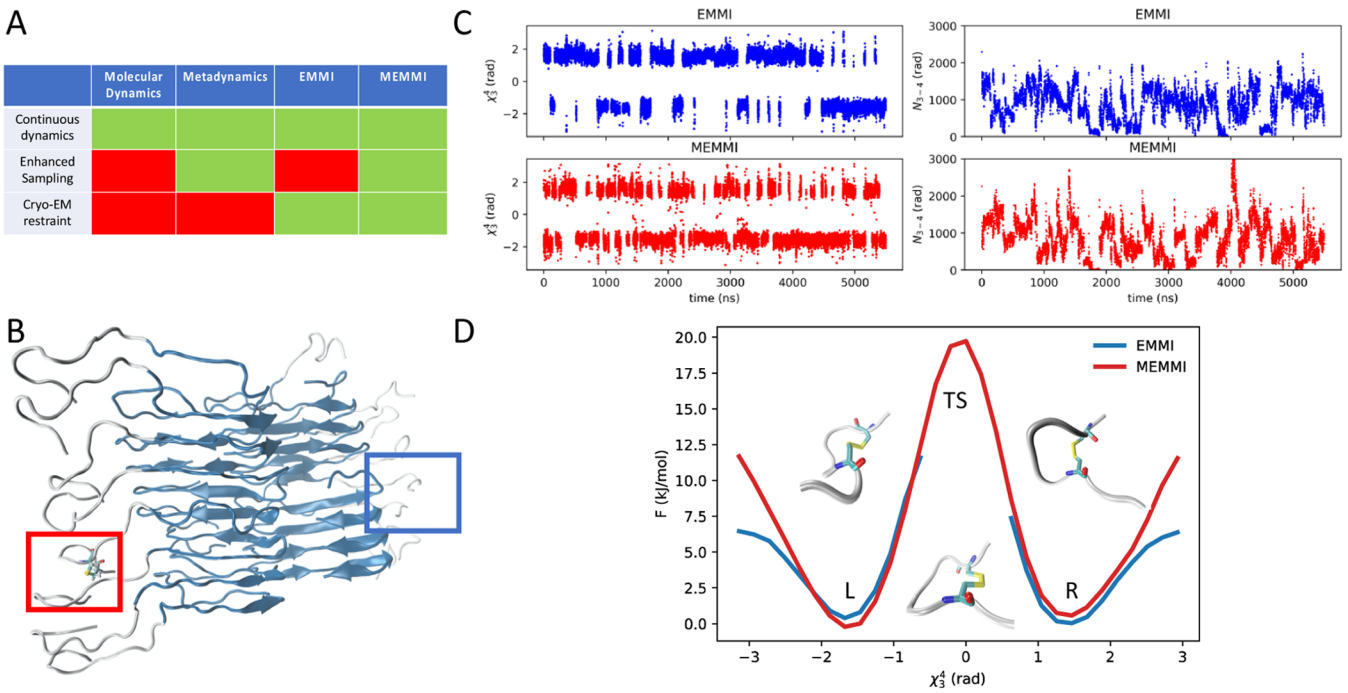


Figure 1. Acceleration of the conformational sampling using MEMMI. (A) Summary of the capability of the different approaches discussed in this work (molecular dynamics, metadynamics, EMMI, and MEMMI) to include continuous dynamics, enhanced sampling, and cryo-EM restraints. (B) Representative configuration of an IAPP amyloid fibril with a biased disulfide bond (red) and an unbiased one (blue). (C) Time trace of the χ_3^4 and $N_{3\text{-}4}$ CVs with EMMI and MEMMI. (D) Free energy profile as a function of χ_3^4 with the corresponding structural states of the disulfide bond: MEMMI (red) EMMI (blue).

SANS),[31−34] transition rate constants,[23,24] and cryo-EM data[35] and AF distance map data.[36]

One of such methods, cryo-EM metainference (EMMI),[35] can accurately model a thermodynamic ensemble by combining prior information on the system, such as physicochemical knowledge (e.g., a force field), with noisy (i.e., subject to systematic and random errors) and heterogeneous (i.e., encoding a conformational ensemble) experimental data, using cryo-EM density maps. EMMI has already been used in a series of complex macromolecular systems, including a CLP protease,[37] microtubules,[38] microtubule-tau complexes,[39] ASCT2 transporter,[40] and SARS-CoV-2 spike protein,[41,42] allowing access to the continuous dynamics of biomolecules with atomistic resolution. The quality of the EMMI structural ensembles, however, is closely related to the exhaustiveness of the conformational sampling, which requires a computational time that scales exponentially with the barriers delimiting individual structural states.[10]

To tackle this rare event problem, several enhanced sampling methods have been developed. Enhanced sampling molecular simulation methods can be classified as trajectory-based[43−47] and collective-variable (CV)-based.[48−52] For detailed reviews, the reader can refer to the recent literature.[53,54] A particularly powerful CV-based enhanced sampling method, which is very efficient once appropriate CVs are chosen, is metadynamics.[52,55] Metadynamics adds a history-dependent bias to the system as a function of microscopic degrees of freedom of the system known as collective variables. With this bias, the simulations can escape deep free energy minima and sample transitions between different states. The choice of the CVs is critical to achieve the desired speed-up of convergence.[53] Recent developments in identifying and automating the search for appropriate CVs have increased the efficiency of this method, thus providing a remedy to the conformational sampling problem,[56−60] which typically affects standard molecular dynamics simulations (Figure 1A). Metainference has been successfully combined with metadynamics[61] and shown significant improvements over unrestrained force fields for many different biophysical systems.[27,29,62,63] However, metadynamics has not yet been combined with EMMI (Figure 1A). Combining EMMI with enhanced sampling methods can lead to accurate and efficient determination of structural ensembles using the large number of datasets in cryo-EM databases, which can in turn provide atomistic insight into a range of biomolecular systems and processes.

In this work, we present the MEMMI method, which incorporates metadynamics into EMMI to accelerate the ability of EMMI to sample structural ensembles that include slowly interconverting states. We illustrate the application of this approach to determine the structural ensemble of an amyloid fibril formed by the full-length (residues 1−37) islet amyloid polypeptide (IAPP), an aberrant assembly associated with the degeneration of pancreatic β-cells in type-2 diabetes (T2D). When functioning correctly, IAPP, together with insulin, contributes to glycaemic control. IAPP and insulin are synthesized and stored together in pancreatic β-cells, but when IAPP aggregates in the extracellular space of the islets of Langerhans, amyloid-induced apoptosis of β-cells may occur.[64] In 95% of T2D patients, IAPP is found as extracellular amyloid deposits,[65−67] which form through surface-catalyzed secondary nucleation.[68]

IAPP fibrils represent a challenging system for structural biology studies since no unique structure can be readily resolved in the low-density regions of the 12-residue long N-terminal tails, due to conformational heterogeneity and

https://doi.org/10.1021/acs.biochem.3c00010
Biochemistry 2023, 62, 2407−2416

associated errors in the measurement. For this reason, although recent cryo-EM experiments determined various amyloid fibril structures of IAPP,[69−71] the structural heterogeneity in the disordered flanking regions, known as fuzzy coat, has so far proved impossible to resolve accurately. It would be desirable to acquire a better understanding of the conformational properties of the fuzzy coat since this region is thought to play a central role in the interactions of amyloid fibrils with other cellular components such as RNA molecules and molecular chaperones.[72,73] Moreover, the fuzzy coat is likely to be involved in cell membrane binding, potentially promoting the catalysis of aggregation and capturing amyloid precursors.[72,73] Recent studies on the tau protein, which is implicated in a family of neurodegenerative diseases known as tauopathies, show that depending on pH conditions, the thick fuzzy coat can change the fibril properties, including mechanical stiffness, and repulsive and adhesive behaviors.[72,73] Here, we detail the dynamics of the fuzzy coat of an IAPP fibril and utilize a thermodynamical theory of melting to characterize the different regions of the fibril to gain insight into its mechanical properties.

MATERIALS AND METHODS

MEMMI Method. *Cryo-EM Forward Model.* A cryo-EM density map resulting from class-averaging and three-dimensional (3D) reconstruction is typically encoded as voxels on a grid, and the map is generally distributed in this form. For computational efficiency, and to enable differentiability and analysis of correlations between data points, the map can be converted to a Gaussian mixture model $\phi_D(x)$ (GMM) consisting of N_D Gaussian components

$$\phi_D(\boldsymbol{x}) = \sum_{i=1}^{N_D} \phi_{D,i}(\boldsymbol{x}) = \sum_{i=1}^{N_D} \omega_{D,i} G(\boldsymbol{x}|\boldsymbol{x}_{D,i}, \Sigma_{D,i}) \tag{1}$$

where $\boldsymbol{x}$ is a vector in Cartesian space, $\omega_{D,i}$ is the scaling factor of the i-th component of the data GMM, and G is a normalized Gaussian function centered at $\boldsymbol{x}_{D,i}$ with covariance matrix $_{D,i}$. The agreement between models generated by molecular dynamics (MD) and the data GMM is calculated by the following overlap function $ov_{MD,i}$

$$ov_{MD,i} = \int \phi_M(\boldsymbol{x}) \phi_{D,i}(\boldsymbol{x}) d\boldsymbol{x} \tag{2}$$

where $\phi_M(\boldsymbol{x})$ corresponds to the model GMM obtained from molecular dynamics. To deal with the heterogeneity of the system, EMMI simulates many replicas, r, of the system. The overlap between model GMM and data GMM is estimated over the ensemble of replicas as an average overlap per GMM component $\overline{ov}_{MD,i}$. This forward model overlap can then be compared to the data GMM self-overlap $ov_{DD,i} = \int \phi_D(\boldsymbol{x}) \phi_{D,i}(\boldsymbol{x}) d\boldsymbol{x}$

Metainference. Metainference is a Bayesian approach for modeling statistical ensembles by combining prior information on a system with experimental data subject to noise or systematic errors.[17] This framework is particularly well suited to structural ensemble determination through molecular dynamics simulations, in which the *prior* (i.e., the force field) is updated with information from experimental methods, such as NMR spectroscopy, SAXS, or cryo-EM data. Metainference is designed to handle systematic errors (such as biases in the force field or forward model), random errors (due to noise in experimental data), and errors due to the limited sample size of the ensemble.[18] The model generation is governed by the metainference energy function, defined as $E_{MI} = -k_B T \log(p_{MI})$, in which k_B is the Boltzmann constant, T is the temperature, and p_{MI} is the metainference posterior probability

$$p_{MI}(\boldsymbol{X}, \sigma^{SEM}, \sigma^{B}|\boldsymbol{D}) = \prod_{r=1}^{N_R} p(X_r) \prod_{i=1}^{N_D} p(d_i|\boldsymbol{X}, \sigma_i^{SEM}, \sigma_{r,i}^{B}) p(\sigma_{r,i}) \tag{3}$$

Here, $\boldsymbol{X}$ is a vector representing the atomic coordinates of the full ensemble, consisting of individual replicas X_r; σ^{SEM} is the error incurred by the limited number of replicas in the ensemble; σ^{B} encodes the random and systematic errors in the prior, forward model, and experiment; and $\boldsymbol{D} = [d_i]$ is the experimental data. Note σ^{SEM} is calculated per data point (σ_i^{SEM}), while σ^{B} is computed per data point i and replica r as $\sigma_{i,r}^{B}$. In the present case, the likelihood $p(d_i|X,\sigma_i^{SEM}, \sigma_{r,i}^{B})$ takes the form of a Gaussian function

$$p(d_i|\boldsymbol{X}, \sigma_i^{SEM}, \sigma_{r,i}^{B}) = \frac{1}{\sqrt{2\pi}\sqrt{(\sigma_{r,i}^{B})^2 + (\sigma_i^{SEM})^2}} \exp\left[-\frac{1}{2}\frac{[ov_{DD,i} - \overline{ov}_{MD,i}]^2}{(\sigma_{r,i}^{B})^2 + (\sigma_i^{SEM})^2}\right] \tag{4}$$

where $\overline{ov}_{MD,i}$ is the ensemble average of the overlap. The metainference energy function for multiple replicas then becomes

$$E_{MI}(\boldsymbol{X}, \sigma) = E_{MD}(\boldsymbol{X}) + \frac{k_B T}{2} \sum_{r,i}^{N_R,N_D} \frac{[ov_{DD,i} - \overline{ov}_{MD,i}]^2}{(\sigma_{r,i}^{B})^2 + (\sigma_i^{SEM})^2} + E_\sigma \tag{5}$$

where E_σ represents the energy associated with the error $\sigma = (\sigma^{B}, \sigma^{SEM})$

$$E_\sigma = k_B T \sum_{r,i}^{N_R,N_D} -\log(\sigma_{r,i}^{B}) + \frac{1}{2}\log[(\sigma_{r,i}^{B})^2 + (\sigma_i^{SEM})^2] \tag{6}$$

E_{MD} represents the molecular dynamics force field. While the space of conformations X_r is sampled by multi-replica molecular dynamics simulations, the error parameters for each data point $\sigma_{r,i}^{B}$ are sampled by a Monte Carlo sampling scheme at each time step. The error parameter related to the limited number of replicas used to estimate the forward model (σ^{SEM}) can be chosen as a constant or estimated on the fly by using a windowed average.[32]

Metadynamic Cryo-EM Metainference (MEMMI). To accelerate the sampling of the metainference ensemble, one can utilize an enhanced sampling scheme such as metadynamics.[52,61] In this case, we use parallel-bias metadynamics (PBMetaD)[56] with the multiple walkers scheme.[74] Here, V_{PB} is a time-dependent biasing potential acting on a set of N_{CV} collective variables $s(\boldsymbol{X})$, which in turn are functions of the system coordinates

$$V_{PB}(s_j(\boldsymbol{X}), t) = -k_B T \log\left\{\sum_{j=1}^{N_{CV}} \exp\left[\frac{V_G(s_j(\boldsymbol{X}), t)}{k_B T}\right]\right\} \tag{7}$$

In contrast to conventional metadynamics, in PBMetaD, multiple one-dimensional bias potentials V_G are deposited

https://doi.org/10.1021/acs.biochem.3c00010
Biochemistry 2023, 62, 2407−2416

rather than a single high-dimensional one. This alleviates the curse of dimensionality while still allowing an efficient exploration of phase space.[56] Additionally, the use of multiple replicas through the multiple walkers scheme[74] allows the sharing of the bias potential to drastically improve the sampling performance, while at the same time being a natural fit for the replica averaging approach of EMMI. Analogously to well-tempered metadynamics, these bias potentials $V_G(s_j(\boldsymbol{X}), t)$ eventually converge to the free energy $F(s_j(\boldsymbol{X}))$. The MEMMI energy function then becomes

$$E_{\mathrm{MEMMI}}(\boldsymbol{X}, \sigma, t) = E_{\mathrm{EMMI}}(\boldsymbol{X}, \sigma) + \sum_{r=1}^{N_R} V_{\mathrm{PB}}[\boldsymbol{s}(X_r), t] \tag{8}$$

While the PBMetaD bias potential is shared among replicas, each replica may still experience a varying potential depending on its location in phase space. Thus, the arithmetic average over the forward models (i.e., the overlap) no longer presents an unbiased estimate of the ensemble average. It therefore needs to be replaced with a weighted average, utilizing the bias potential of each replica r to unbias the ensemble at time t

$$f_i(\boldsymbol{X}, t) = \sum_{r=1}^{N_R} w(X_r, t) f_i(X_r, t) \tag{9}$$

where the unbiasing weight $w(X_r, t)$ is defined as

$$w(X_r, t) = \exp\left\{\frac{V_{\mathrm{PB}}[\boldsymbol{s}(X_r), t]}{k_B T}\right\} \left[\sum_{j=1}^{N_R} \exp\left\{\frac{V_{\mathrm{PB}}[\boldsymbol{s}(\boldsymbol{X}_{jr}), t]}{k_B T}\right\}\right]^{-1} \tag{10}$$

The unbiasing procedure used here is analogous to the standard umbrella-sampling technique.[48] Now the ensemble average $\overline{ov}_{\mathrm{MD},i}(\boldsymbol{X})$ is given by

$$\overline{ov}_{\mathrm{MD},i}(\boldsymbol{X}) = \sum_{r=1}^{N_R} w(X_r, t) ov_{\mathrm{MD},i}(X_r) \tag{11}$$

The MEMMI energy E_{MEMMI} is thus equal to eq 5.

Initial Fibril Structure. We build the initial structure of the full fibril by starting from a deposited fibril structure (PDB: 6Y1A), which only contains the core of the fibril, and extending each of the 16 polypeptide chains by adding the missing 12-residue N-terminal sequence with guidance from the cryo-EM density map EMD-10669. We use the macromolecular model-building program Coot.[75] We note that since all of the simulations that we present in this work reached convergence, the results that we report are independent of the initial configuration. We also note that the electron density map EMD-10669 used here for the MEMMI approach was generated using a helical reconstruction method, which in principle could lead to a periodicity not only for the fibril core but also for the disordered N-terminal regions. There are indeed periodic densities visible outside of the fibril core in the map, and it is unclear whether these regions should be seen as artifacts of the reconstruction approach or valid densities. The original reconstruction procedure does not mention the use of a mask to discard information on the fuzzy coat,[71] as is often done for amyloid reconstructions.[76] On the other hand, well-defined densities beyond the fibril core often have important structural context,[77] and we thus decided to use the density map as is.

Molecular Dynamics Setup and Equilibration. We continue by creating a 12.34 nm × 12.34 nm × 12.34 nm cubic simulation box, solvating with 58291 water molecules and neutralizing the net charge by adding 48 Cl^- ions. We use the CHARMM22*[78] force field and TIP3P[79] water models. We continue with an energy minimization followed by a 500 ps NPT equilibration at a temperature of 310 K and pressure of 1 atm, followed by an additional 2 ns NVT equilibration at 310 K. The molecular dynamics parameters are the same used previously.[41]

MEMMI Simulations. We first express the experimental voxel map data as a data GMM containing 10,000 Gaussians in total, resulting in a 0.975 correlation to the original voxel map EMD-10669,[64] using the gmmconvert utility.[80] We continue by extracting 32 configurations from the previous NVT equilibration and initiate an MEMMI simulation, consisting of 32 replicas, resulting in an aggregate runtime of 5.49 μs, using PLUMED.2.6.0-dev[81] and gromacs-2020.6.[82] The simulation is performed in the NVT ensemble at 310 K using the same MD parameters as in the equilibration step. Configurations are saved every 10 ps for post-processing. The cryo-EM restraint is updated every 2 MD steps, using neighbor lists to compute the overlaps between model and data GMMs, with a neighbor list cutoff of 0.01 and update frequency stride of 100 steps. The biasing collective variables $s = [s_i]$ in the simulation are shown in Figure S1A, and the biasing scheme is PBMetaD[56] with the well-tempered[83] and multiple walkers[74] protocols. The hill height is set to 0.3 kJ/mol, with a deposition frequency of 200 steps and an adaptive Gaussian diffusion scheme.[84] The biasing collective variables correspond to degrees of freedom of the left-hand-side N-terminal. The respective degrees of freedom of the right-hand-side N-terminal do not feel a metadynamics potential and therefore in the remaining text will be referred to as EMMI degrees of freedom and are also listed in Figure S1. As a post-processing step, we generate the final structural ensemble by resampling the generated configurations based on the converged unbiasing weights for each structure after an equilibration of 7 ns shown in Figure 1A. To establish convergence, we perform a clustering analysis on the structural ensemble based separately on the first and second half of each replica (taking into account the weights) using the GROMOS method,[85] and with metric the root-mean-square deviation (RMSD) calculated on the Cα (CA) atoms (Figure S1B). Time traces of CVs as well as their time-dependent free energy profile are shown in Figures S2−S4. For molecular visualizations and calculating the local correlation of the final structural-ensemble-generated cryo-EM map with the experimental cryo-EM map, we use Chimera and gmconvert.[86] Except otherwise mentioned, all of the structural analysis is performed on the degrees of freedom of the N-tails on the left-hand side, which is the one with MEMMI restraints, and the two central pairs of polypeptide chains in the overall stack of eight pairs, to avoid the finite size effects.

RESULTS AND DISCUSSION

Structure and Dynamics of IAPP Amyloid Fibrils. *Acceleration of the Conformational Sampling.* MEMMI accelerates the conformational sampling by biasing a set of microscopic degrees of freedom of the system, also known as collective variables (CVs, Figure S1). In addition, MEMMI also corrects possible inaccuracies in the force field used in the

https://doi.org/10.1021/acs.biochem.3c00010
Biochemistry 2023, 62, 2407−2416

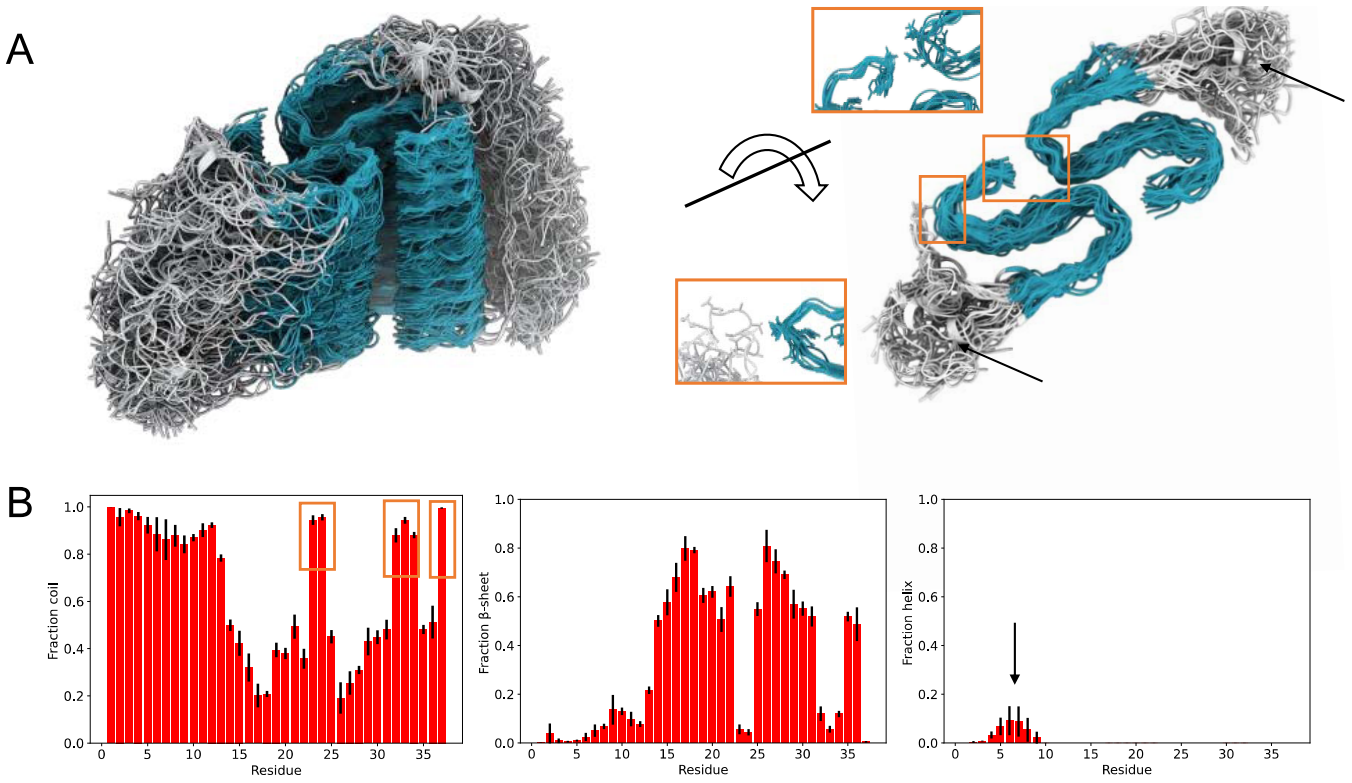


Figure 2. Structural ensemble of an IAPP amyloid fibril. (A) Core fibril residues (13−37) are shown in cyan, while the N-terminal tail residues (1−12) are shown in gray (fuzzy coat). This representation of the structural ensemble was generated by extracting 50 conformations from the final structural ensemble. Close-ups of interfaces are shown from 10 conformations. Examples of individual conformations are shown in Figure S6. (B) Structural analysis reporting on the fraction of coil, β-sheet, and α-helix formed per residue in the ensemble obtained from MEMMI. Error bars are calculated as standard deviation between the first and second simulation halves.

molecular dynamics simulations through the use of experimental restraints. To demonstrate the performance of this approach, we visualize the time trace of the collective variable χ_3^4, which corresponds to the disulfide bond dihedral of the 4th polypeptide in the eight-layer stack of the fibril. This polypeptide is thus representative of a buried monomer with little interaction with the fibril ends. Note that, due to C2 helical symmetry, there are two of these dihedrals, one corresponding to the right-hand-side N-terminal tail and one to the left-hand-side N-terminal tail (Figure 1B). The sampling of one side is accelerated by a biasing potential of eq 7 (MEMMI), while the other is not (EMMI). Compared to the dihedral, the biased disulfide χ_3^4 (MEMMI) shows an increased transition rate, and thus more efficient conformational sampling (Figure 1C). While both methods characterize the wells of the stable states (L, R), MEMMI is able to provide access to higher free energy transition state (TS) regions (Figure 1D). Monitoring a nonperiodic CV, such as the number of contacts between N-tail three and four, shows diffusion along high−low contact regions in both MEMMI and EMMI cases but is somewhat more frequent in the MEMMI case. The combination of diffusion shown in the traces of CVs and the free energy profiles as a function of simulation time of all collective variables biased by EMMI and MEMMI are shown in Figures S2−S4 and indicate that our simulations are well converged. Taken together, these results show that both the MEMMI and EMMI simulations are converged in the low free energy regions, while MEMMI enables visiting high free energy regions.

Conformational Heterogeneity of the Fuzzy Coat. A structure of the IAPP fibril core (residues 13−37) has been previously published (PDB: 6Y1A) using the cryo-EM data used also in the present work (EMD-10669). Here, we determine a structural ensemble of the whole IAPP fibril (residues 1−37). We model the system as a stack of eight polypeptides per side (Figure 2A). While the core of the fibril maintains a parallel β-sheet structure, the flanking region (residues 1−12) exhibits a large conformational heterogeneity (Figure 2B). While we find the cores residues 12−37 to be largely in a β-sheet conformation, we also note significant heterogeneity for residues 23−24, 32−34, and 37. As shown in Figure 2B, residues 23−24 interact with TYR37 and maintain mostly a coil structure, while residues 32−34 interact with the fuzzy coat and maintain mostly a coiled structure. We also detected a small population of α-helical conformations in the region of residues 5−9.

Correlation between Experimental and Calculated Cryo-EM Maps. We estimate the correlation of the MEMMI structural ensemble with the experimental cryo-EM map (Figure 3). We find that using a structural ensemble, IAPP (residues 1−37) correlates better with the experimental cryo-EM map than a single structure (PDB: 6Y1A) (Figure 3A,B). The coefficient of correlation of the structural ensemble to the experimental electron density map is on average 0.92. Furthermore, an important feature of MEMMI is its ability to estimate the error in the experimental electron density map (Figures 3C and S5). We find that the relative error per Gaussian data point is on average 0.09, where the relative error is the error of each Gaussian data point with respect to the

https://doi.org/10.1021/acs.biochem.3c00010
Biochemistry 2023, 62, 2407−2416

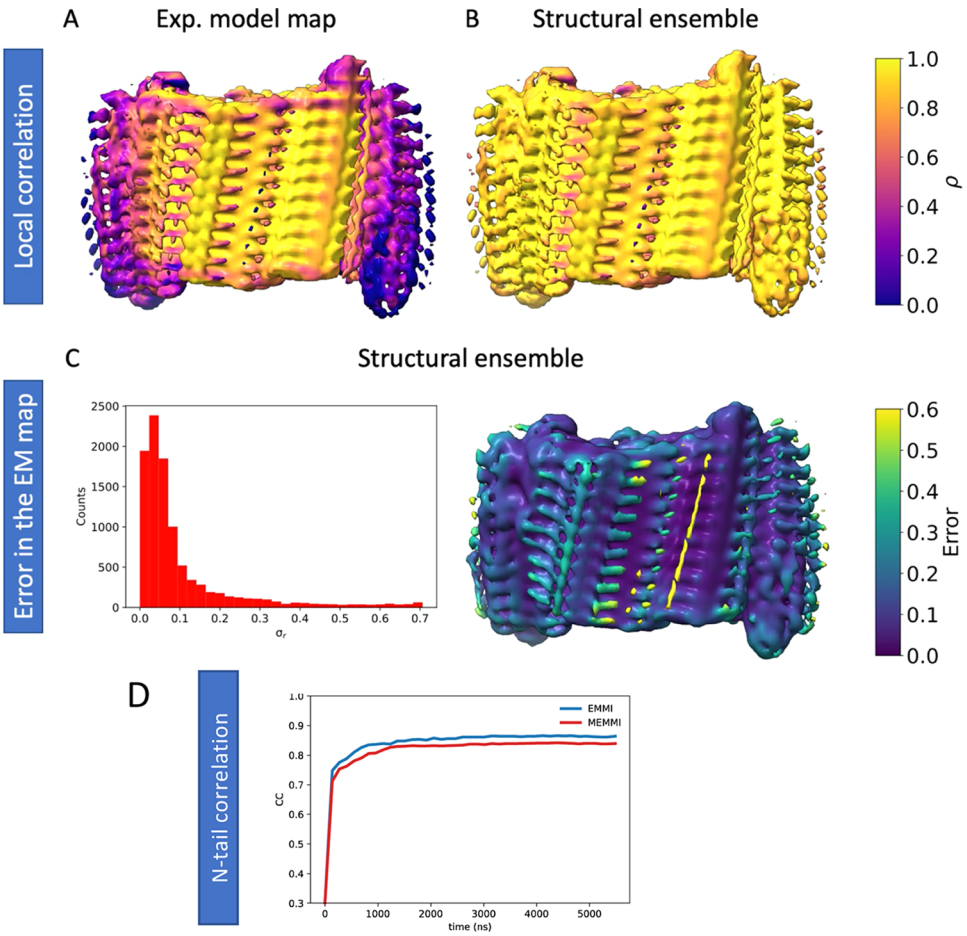


Figure 3. Local correlation and error in the data from MEMMI. (A, B) Assessment of the local correlation between the cryo-EM map (EMD-EM-10669) with cryo-EM maps generated by a previously determined single structure (PDB: 6Y1A) (A), and the MEMMI structural ensemble (B). (C) Histogram of the error in the GMM data as obtained by the structural ensemble (left). Error in the data, projected in the EMD-EM-10669 (right). (D) Time-dependent correlation of the left/right-side N-tail structural ensemble (MEMMI/EMMI) with the corresponding region in the cryo-EM map.

total overlap between all data GMM and the *i*th component of the data GMM.[35] The low-density, high-error volume around residues S34, N35, and T36 can likely be attributed to the lack of MES/NaOH buffer in our MEMMI simulations, which is present in the experimental setup. Both MEMMI and EMMI exhibit good correlation (about 0.83) in the respective N-tail region with the cryo-EM map (Figure 3D).

Comparison of the Dynamical Properties of the Fibril Core and Fuzzy Coat. The conformational properties of the fuzzy coat and core region have been shown to be relevant in modulating the properties of amyloid fibrils, including their ability to interact with various cellular components.[72,73] To investigate this phenomenon in the case of the present IAPP amyloid fibril, we take inspiration from a thermodynamic theory of melting and characterize the residue-dependent Lindemann parameter Δ_L (Figure 4), which encodes information on solid-like and liquid-like behavior.[87] At the backbone level, we find that the fibril core (23−30) is solid-like ($\Delta_L < 0.15$), while the flanking region (1−12) is on the verge of a liquid-like behavior ($\Delta_L \geq 0.15$). The Lindemann parameters of the side chains indicate more mobility and are liquid-like outside the region of residues 20−32. These results reveal that about half of the structural core of this amyloid fibril remains rather disordered at the side-chain level, a phenomenon observed also for folded native states, even in otherwise rigid systems such as ubiquitin.[87] The fuzzy coat thus exhibits a degree of conformational heterogeneity that is much lower than that of monomeric disordered proteins such as amyloid-β, which generally exhibit Lindemann parameters $\Delta_L \geq 1.0$ (Figure S7).

Residue-Specific Solubility of the Amyloid Fibril Surface. To investigate whether or not the surface of the amyloid fibril is soluble, we calculated the solubility per residue using the structure-corrected CamSol solubility score[88] (Figure 5). We thus found three insoluble regions: (i) residues 5−7 (ATC) in the fuzzy coat, (ii) residues 26−29 (ILSS) at the end of the fibril fragment, and (iii) residues 14−18 (NFLVH) on the fibril surface, stretched along the helical axis. The latter result is consistent with experimental evidence of residues N14, H18, and S20 being implicated in IAPP aggregation[66,89] via secondary nucleation. We note that the solvent-exposed and aggregation-prone residues 5−7 and 18−20 might present attractive targets for structure-based drug discovery. This strategy for choosing targets was recently demonstrated experimentally in the case of α-synuclein fibrils.[90]

CONCLUSIONS

We have presented the MEMMI method for the simultaneous determination of the structure and dynamics of large and conformationally heterogeneous biomolecular structures from

https://doi.org/10.1021/acs.biochem.3c00010
Biochemistry 2023, 62, 2407−2416

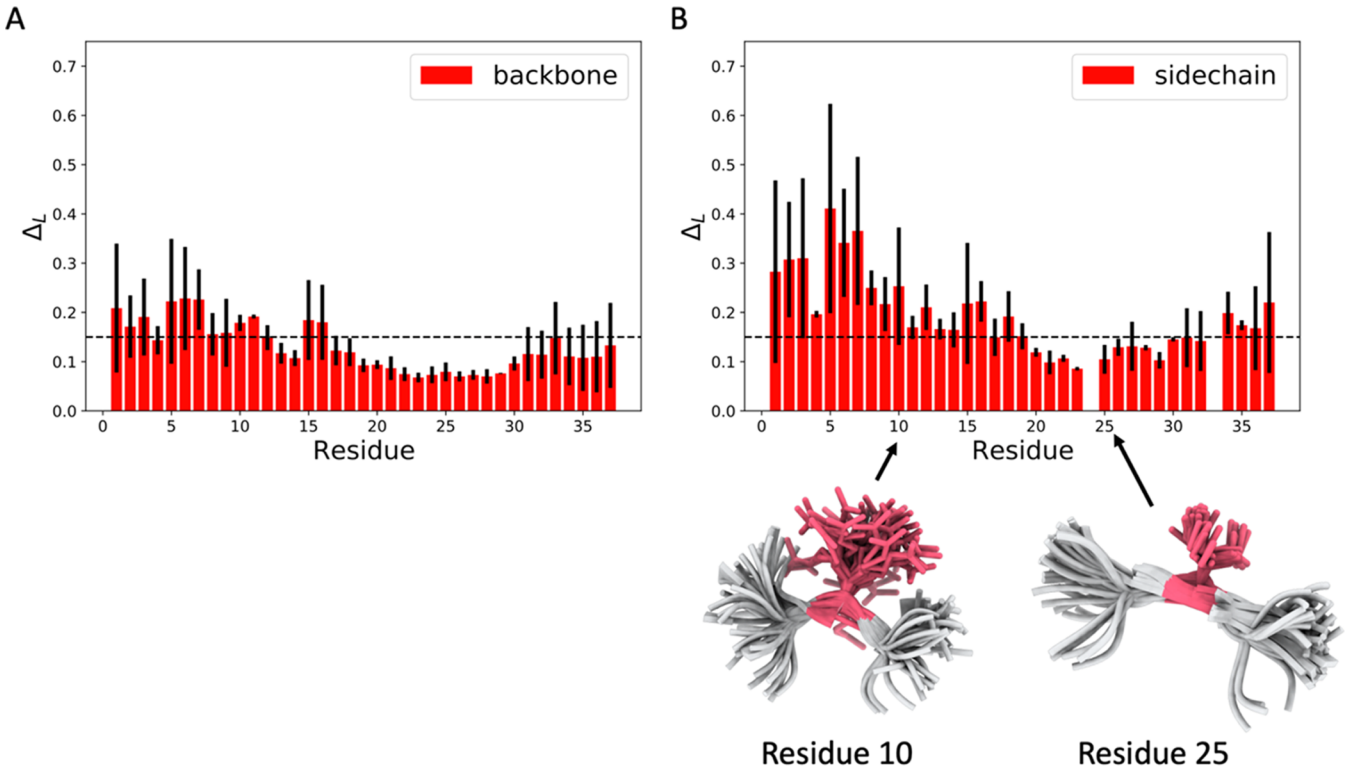


Figure 4. Analysis of the liquid-like and solid-like behaviors of the backbone and side chains. (A, B) Per-residue Lindemann parameter (Δ_L) for backbone (A) and side-chain (B) atoms. The dashed line at $\Delta_L = 0.15$ marks the threshold value of the Lindemann parameter that distinguishes the solid-like and liquid-like behaviors. The behavior of the side chains is solid-like only in the region of residues 20−32, which is about half of the fibril core (residues 13−37). Error bars are calculated as standard deviation between the first and second simulation halves.

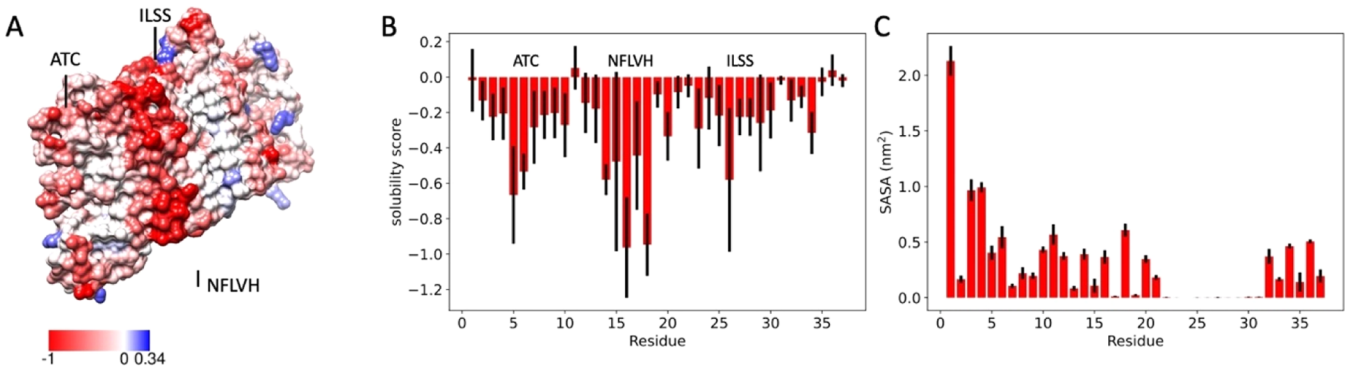


Figure 5. Analysis of the solubility of the fibril surface from the MEMMI ensemble. (A) Representative structure with CamSol solubility scores indicated on the surface (right). (B) Per-residue CamSol solubility score. (C) Per-residue solvent-accessible surface area. Error bars are calculated as standard deviation between the first and second simulation halves.

cryo-EM density maps. To illustrate the information that can be extracted from this type of approach, we have reported a structural ensemble of an amyloid fibril formed by IAPP. The analysis of the structural ensembles has revealed the conformational and aggregation-related properties of the fuzzy coat of the amyloid fibril, and that many of the side chains in the structural core of the amyloid fibril exhibit a liquid-like behavior. Since this phenomenon has also been observed for native states of proteins, these results reveal a similarity in the structural behavior of proteins upon folding in their native and amyloid states.

■ ASSOCIATED CONTENT

Data Availability Statement

Input files for the simulations can be found at PLUMED-NEST (plumID: 22.023). The ensemble can be found at https://zenodo.org/record/6518554.

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.biochem.3c00010.

Additional simulations including convergence of the MEMMI and EMMI simulations, assessment of the local correlation between experimental and calculated cryo-EM map, examples of conformations of the final structural ensemble, and comparison of Lindemann parameters reference systems (PDF)

https://doi.org/10.1021/acs.biochem.3c00010
Biochemistry 2023, 62, 2407−2416

AUTHOR INFORMATION

Corresponding Author

Michele Vendruscolo − *Centre for Misfolding Diseases, Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, U.K.;* orcid.org/0000-0002-3616-1610; Email: mv245@cam.ac.uk

Authors

Z. Faidon Brotzakis − *Centre for Misfolding Diseases, Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, U.K.*

Thomas Löhr − *Centre for Misfolding Diseases, Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, U.K.*

Steven Truong − *Centre for Misfolding Diseases, Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, U.K.;* orcid.org/0000-0001-9624-4461

Samuel Hoff − *Department of Structural Biology and Chemistry, Institut Pasteur, Université Paris Cité CNRS UMR 3528, 75015 Paris, France*

Massimiliano Bonomi − *Department of Structural Biology and Chemistry, Institut Pasteur, Université Paris Cité CNRS UMR 3528, 75015 Paris, France;* orcid.org/0000-0002-7321-0004

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.biochem.3c00010

Author Contributions

§Z.F.B. and T.L. contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Z.F.B. acknowledges the Federation of European Biochemical Societies (FEBS) for financial support (LTF). M.B. and S.E.H. acknowledge the support of the French Agence Nationale de la Recherche (ANR), under grant ANR-20-CE45-0002 (project EMMI).

REFERENCES

(1) Kühlbrandt, W. The resolution revolution. *Science* **2014**, *343*, 1443−1444.

(2) Bai, X.-C.; McMullan, G.; Scheres, S. H. How cryo-em is revolutionizing structural biology. *Trends Biochem. Sci.* **2015**, *40*, 49−57.

(3) Dubochet, J.; McDowall, A. Vitrification of pure water for electron microscopy. *J. Microsc.* **1981**, *124*, 3−4.

(4) Dubochet, J.; Lepault, J.; Freeman, R.; Berriman, J.; Homo, J. C. Electron microscopy of frozen water and aqueous solutions. *J. Microsc.* **1982**, *128*, 219−237.

(5) Henderson, R. Realizing the potential of electron cryo-microscopy. *Q. Rev. Biophys.* **2004**, *37*, 3−13.

(6) Renaud, J.-P.; Chari, A.; Ciferri, C.; Liu, W.-t.; Rémigy, H.-W.; Stark, H.; Wiesmann, C. Cryo-EM in drug discovery: Achievements, limitations and prospects. *Nat. Rev. Drug Discovery* **2018**, *17*, 471−492.

(7) Nakane, T.; Kotecha, A.; Sente, A.; McMullan, G.; Masiulis, S.; Brown, P. M.; Grigoras, I. T.; Malinauskaite, L.; Malinauskas, T.; Miehling, J.; et al. Single-particle cryo-em at atomic resolution. *Nature* **2020**, *587*, 152−156.

(8) Lawson, C. L.; Patwardhan, A.; Baker, M. L.; Hryc, C.; Garcia, E. S.; Hudson, B. P.; Lagerstedt, I.; Ludtke, S. J.; Pintilie, G.; Sala, R.; et al. Emdatabank unified data resource for 3dem. *Nucleic Acids Res.* **2016**, *44*, D396−D403.

(9) Van Heel, M.; Frank, J. Use of multivariates statistics in analysing the images of biological macromolecules. *Ultramicroscopy* **1981**, *6*, 187−194.

(10) Bonomi, M.; Vendruscolo, M. Determination of protein structural ensembles using cryo-electron microscopy. *Curr. Op. Struct. Biol.* **2019**, *56*, 37−45.

(11) Bonomi, M.; Heller, G. T.; Camilloni, C.; Vendruscolo, M. Principles of protein structural ensemble determination. *Curr. Op. Struct. Biol.* **2017**, *42*, 106−116.

(12) Fraser, J. S.; Lindorff-Larsen, K.; Bonomi, M. What will computational modeling approaches have to say in the era of atomistic cryo-EM data? *J. Chem. Inf. Model.* **2020**, *60*, 2410−2412.

(13) Pitera, J. W.; Chodera, J. D. On the use of experimental observations to bias simulated ensembles. *J. Chem. Theory Comput.* **2012**, *8*, 3445−3451.

(14) Cavalli, A.; Camilloni, C.; Vendruscolo, M. Molecular dynamics simulations with replica-averaged structural restraints generate structural ensembles according to the maximum entropy principle. *J. Chem. Phys.* **2013**, *138*, No. 03B603.

(15) Roux, B.; Weare, J. On the statistical equivalence of restrained-ensemble simulations with the maximum entropy method. *J. Chem. Phys.* **2013**, *138*, No. 02B616.

(16) Boomsma, W.; Ferkinghoff-Borg, J.; Lindorff-Larsen, K. Combining experiments and simulations using the maximum entropy principle. *PLoS Comput. Biol.* **2014**, *10*, No. e1003406.

(17) Hummer, G.; Köfinger, J. Bayesian ensemble refinement by replica simulations and reweighting. *J. Chem. Phys.* **2015**, *143*, No. 12B634_631.

(18) Bonomi, M.; Camilloni, C.; Cavalli, A.; Vendruscolo, M. Metainference: A Bayesian inference method for heterogeneous systems. *Sci. Adv.* **2016**, *2*, No. e1501177.

(19) Marinelli, F.; Faraldo-Gómez, J. D. Ensemble-biased metadynamics: A molecular simulation method to sample experimental distributions. *Biophys. J.* **2015**, *108*, 2779−2782.

(20) Olsson, S.; Wu, H.; Paul, F.; Clementi, C.; Noé, F. Combining experimental and simulation data of molecular processes via augmented markov models. *Proc. Natl. Acad. Sci. U.S.A.* **2017**, *114*, 8265−8270.

(21) Cesari, A.; Reißer, S.; Bussi, G. Using the maximum entropy principle to combine simulations and solution experiments. *Computation* **2018**, *6*, No. 15.

(22) Bottaro, S.; Bengtsen, T.; Lindorff-Larsen, K. Integrating Molecular Simulation and Experimental Data: A Bayesian/Maximum Entropy Reweighting Approach. In *Methods in Molecular Biology*; Springer, 2020; Vol. *2112*, pp 219−240.

(23) Brotzakis, Z. F.; Vendruscolo, M.; Bolhuis, P. G. A method of incorporating rate constants as kinetic constraints in molecular dynamics simulations. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, No. e2012423118.

(24) Bolhuis, P. G.; Brotzakis, Z. F.; Vendruscolo, M. A maximum caliber approach for continuum path ensembles. *Eur. Phys. J. B* **2021**, *94*, No. 188.

(25) Camilloni, C.; Vendruscolo, M. A tensor-free method for the structural and dynamical refinement of proteins using residual dipolar couplings. *J. Phys. Chem. B* **2015**, *119*, 653−661.

(26) Bottaro, S.; Lindorff-Larsen, K. Biophysical experiments and biomolecular simulations: A perfect match? *Science* **2018**, *361*, 355−360.

(27) Heller, G. T.; Aprile, F. A.; Michaels, T. C. T.; Limbocker, R.; Perni, M.; Ruggeri, F. S.; Mannini, B.; Lohr, T.; Bonomi, M.; Camilloni, C.; De Simone, A.; Felli, I. C.; Pierattelli, R.; Knowles, T. P. J.; Dobson, C. M.; Vendruscolo, M. Small-molecule sequestration of amyloid-β as a drug discovery strategy for Alzheimer's disease. *Sci. Adv.* **2020**, *6*, No. eabb5924.

(28) Löhr, T.; Kohlhoff, K.; Heller, G. T.; Camilloni, C.; Vendruscolo, M. A kinetic ensemble of the Alzheimer's Aβ peptide. *Nat. Comput. Sci.* **2021**, *1*, 71−78.

(29) Löhr, T.; Jussupow, A.; Camilloni, C. Metadynamic metainference: Convergence towards force field independent

structural ensembles of a disordered peptide. *J. Chem. Phys.* **2017**, *146*, No. 165102.
(30) Bonomi, M.; Pellarin, R.; Kim, S. J.; Russel, D.; Sundin, B. A.; Riffle, M.; Jaschob, D.; Ramsden, R.; Davis, T. N.; Muller, E. G.; Sali, A. Determining protein complex structures based on a Bayesian model of in vivo förster resonance energy transfer (fret) data. *Mol. Cell. Proteom.* **2014**, *13*, 2812−2823.
(31) Hermann, M. R.; Hub, J. S. Saxs-restrained ensemble simulations of intrinsically disordered proteins with commitment to the principle of maximum entropy. *J. Chem. Theory Comput.* **2019**, *15*, 5103−5115.
(32) Bengtsen, T.; Holm, V. L.; Kjølbye, L. R.; Midtgaard, S. R.; Johansen, N. T.; Tesei, G.; Bottaro, S.; Schiøtt, B.; Arleth, L.; Lindorff-Larsen, K. Structure and dynamics of a nanodisc by integrating NMR, saxs and sans experiments with molecular dynamics simulations. *eLife* **2020**, *9*, No. e56518.
(33) Paissoni, C.; Jussupow, A.; Camilloni, C. Determination of protein structural ensembles by hybrid-resolution saxs restrained molecular dynamics. *J. Chem. Theory Comput.* **2020**, *16*, 2825−2834.
(34) Jussupow, A.; Messias, A. C.; Stehle, R.; Geerlof, A.; Solbak, S. M.; Paissoni, C.; Bach, A.; Sattler, M.; Camilloni, C. The dynamics of linear polyubiquitin. *Sci. Adv.* **2020**, *6*, No. eabc3786.
(35) Bonomi, M.; Pellarin, R.; Vendruscolo, M. Simultaneous determination of protein structure and dynamics using cryo-electron microscopy. *Biophys. J.* **2018**, *114*, 1604−1613.
(36) Brotzakis, Z. F.; Zhang, S.; Vendruscolo, M. AlphaFold Prediction of Structural Ensembles of Disordered Proteins. *bioRxiv* **2023**, DOI: 10.1101/2023.01.19.524720.
(37) Vahidi, S.; Ripstein, Z. A.; Bonomi, M.; Yuwen, T.; Mabanglo, M. F.; Juravsky, J. B.; Rizzolo, K.; Velyvis, A.; Houry, W. A.; Vendruscolo, M.; et al. Reversible inhibition of the clpp protease via an N-terminal conformational switch. *Proc. Natl. Acad. Sci. U.S.A.* **2018**, *115*, E6447−E6456.
(38) Eshun-Wilson, L.; Zhang, R.; Portran, D.; Nachury, M. V.; Toso, D. B.; Löhr, T.; Vendruscolo, M.; Bonomi, M.; Fraser, J. S.; Nogales, E. Effects of α-tubulin acetylation on microtubule structure and stability. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116*, 10366−10371.
(39) Brotzakis, Z. F.; Lindstedt, P. R.; Taylor, R. J.; Rinauro, D. J.; Gallagher, N. C.; Bernardes, G. J.; Vendruscolo, M. A structural ensemble of a tau-microtubule complex reveals regulatory tau phosphorylation and acetylation mechanisms. *ACS Cent. Sci.* **2021**, *7*, 1986−1995.
(40) Garibsingh, R.-A. A.; Ndaru, E.; Garaeva, A. A.; Shi, Y.; Zielewicz, L.; Zakrepine, P.; Bonomi, M.; Slotboom, D. J.; Paulino, C.; Grewer, C.; Schlessinger, A. Rational design of ASCT2 inhibitors using an integrated experimental-computational approach. *Proc. Natl. Acad. Sci. U.S.A.* **2021**, *118*, No. e2104093118.
(41) Brotzakis, Z. F.; Löhr, T.; Vendruscolo, M. Determination of intermediate state structures in the opening pathway of SARS-CoV-2 spike using cryo-electron microscopy. *Chem. Sci.* **2021**, *12*, 9168−9175.
(42) Mikolajek, H.; Weckener, M.; Brotzakis, Z. F.; Huo, J.; Dalietou, E. V.; Le Bas, A.; Sormanni, P.; Harrison, P. J.; Ward, P. N.; Truong, S.; Moynie, L.; Clare, D. K.; Dumoux, M.; Dormon, J.; Norman, C.; Hussain, N.; Vogirala, V.; Owens, R. J.; Vendruscolo, M.; Naismith, J. H. Correlation between the binding affinity and the conformational entropy of nanobody SARS-CoV-2 spike protein complexes. *PNAS* **2022**, *119* (31), e2205412119.
(43) Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. Transition path sampling and the calculation of rate constants. *J. Chem. Phys.* **1998**, *108*, 1964−1977.
(44) Van Erp, T. S.; Moroni, D.; Bolhuis, P. G. A novel path sampling method for the calculation of rate constants. *J. Chem. Phys.* **2003**, *118*, 7762−7774.
(45) Allen, R. J.; Warren, P. B.; Ten Wolde, P. R. Sampling rare switching events in biochemical networks. *Phys. Rev. Lett.* **2005**, *94*, No. 018104.
(46) Faradjian, A. K.; Elber, R. Computing time scales from reaction coordinates by milestoning. *J. Chem. Phys.* **2004**, *120*, 10880−10889.
(47) Brotzakis, Z. F.; Bolhuis, P. G. Approximating free energy and committor landscapes in standard transition path sampling using virtual interface exchange. *J. Chem. Phys.* **2019**, *151*, No. 174111.
(48) Torrie, G. M.; Valleau, J. P. Nonphysical sampling distributions in monte carlo free-energy estimation: Umbrella sampling. *J. Comput. Phys.* **1977**, *23*, 187−199.
(49) Grubmüller, H. Predicting slow structural transitions in macromolecular systems: Conformational flooding. *Phys. Rev. E* **1995**, *52*, 2893.
(50) Wang, F.; Landau, D. P. Efficient, multiple-range random walk algorithm to calculate the density of states. *Phys. Rev. Lett.* **2001**, *86*, 2050.
(51) Darve, E.; Pohorille, A. Calculating free energies using average force. *J. Chem. Phys.* **2001**, *115*, 9169−9183.
(52) Laio, A.; Parrinello, M. Escaping free-energy minima. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 12562−12566.
(53) Valsson, O.; Tiwary, P.; Parrinello, M. Enhancing important fluctuations: Rare events and metadynamics from a conceptual viewpoint. *Annu. Rev. Phys. Chem.* **2016**, *67*, 159−184.
(54) Bolhuis, P. G.; Swenson, D. W. Transition path sampling as markov chain monte carlo of trajectories: Recent algorithms, software, applications, and future outlook. *Adv. Theory Simul.* **2021**, *4*, No. 2000237.
(55) Bonomi, M.; Bussi, G.; Camilloni, C.; Tribello, G. A.; Banas, P.; Barducci, A.; Bernetti, M.; Bolhuis, P. G.; Bottaro, S.; Branduardi, D.; Capelli, R.; Carloni, P.; Ceriotti, M.; Cesari, A.; Chen, H.; Chen, W.; Colizzi, F.; De, S.; De La Pierre, M.; Donadio, D.; Drobot, V.; Ensing, B.; Ferguson, A. L.; Filizola, M.; Fraser, J. S.; Fu, H.; Gasparotto, P.; Gervasio, F. L.; Giberti, F.; Gil-Ley, A.; Giorgino, T.; Heller, G. T.; Hocky, G. M.; Iannuzzi, M.; Invernizzi, M.; Jelfs, K. E.; Jussupow, A.; Kirilin, E.; Laio, A.; Limongelli, V.; Lindorff-Larsen, K.; Lohr, T.; Marinelli, F.; Martin-Samos, L.; Masetti, M.; Meyer, R.; Michaelides, A.; Molteni, C.; Morishita, T.; Nava, M.; Paissoni, C.; Papaleo, E.; Parrinello, M.; Pfaendtner, J.; Piaggi, P.; Piccini, G.; Pietropaolo, A.; Pietrucci, F.; Pipolo, S.; Provasi, D.; Quigley, D.; Raiteri, P.; Raniolo, S.; Rydzewski, J.; Salvalaglio, M.; Sosso, G. C.; Spiwok, V.; Sponer, J.; Swenson, D. W. H.; Tiwary, P.; Valsson, O.; Vendruscolo, M.; Voth, G. A.; White, A. Promoting transparency and reproducibility in enhanced molecular simulations. *Nat. Methods* **2019**, *16*, 670−673.
(56) Pfaendtner, J.; Bonomi, M. Efficient sampling of high-dimensional free-energy landscapes with parallel bias metadynamics. *J. Chem. Theory Comput.* **2015**, *11*, 5062−5067.
(57) Tiwary, P.; Berne, B. Spectral gap optimization of order parameters for sampling complex molecular systems. *Proc. Natl. Acad. Sci. U.S.A.* **2016**, *113*, 2839−2844.
(58) Sultan, M.; Pande, V. S. Tica-metadynamics: Accelerating metadynamics by using kinetically selected collective variables. *J. Chem. Theory Comput.* **2017**, *13*, 2440−2447.
(59) Ribeiro, J. M. L.; Bravo, P.; Wang, Y.; Tiwary, P. Reweighted autoencoded variational bayes for enhanced sampling (rave). *J. Chem. Phys.* **2018**, *149*, No. 072301.
(60) Brotzakis, Z. F.; Parrinello, M. Enhanced sampling of protein conformational transitions via dynamically optimized collective variables. *J. Chem. Theory Comput.* **2019**, *15*, 1393−1398.
(61) Bonomi, M.; Camilloni, C.; Vendruscolo, M. Metadynamic metainference: Enhanced sampling of the metainference ensemble using metadynamics. *Sci. Rep.* **2016**, *6*, No. 31232.
(62) Ahmed, M. C.; Skaanning, L. K.; Jussupow, A.; Newcombe, E. A.; Kragelund, B. B.; Camilloni, C.; Langkilde, A. E.; Lindorff-Larsen, K. Refinement of α-synuclein ensembles against saxs data: Comparison of force fields and methods. *Front. Mol. Biosci.* **2021**, *8*, No. 654333.
(63) Achour, A.; Broggini, L.; Han, X.; Sun, R.; Santambrogio, C.; Buratto, J.; Visentin, C.; Barbiroli, A.; De Luca, C. M. G.; Sormanni, P.; Moda, F.; De Simone, A.; Sandalova, T.; Grandori, R.; Camilloni, C.; Ricagno, S. Biochemical and biophysical comparison of human and mouse beta-2 microglobulin reveals the molecular determinants of low amyloid propensity. *FEBS J.* **2020**, *287*, 546−560.

(64) Pilkington, E. H.; Xing, Y.; Wang, B.; Kakinen, A.; Wang, M.; Davis, T. P.; Ding, F.; Ke, P. C. Effects of protein corona on iapp amyloid aggregation, fibril remodelling, and cytotoxicity. *Sci. Rep.* **2017**, *7*, No. 2455.
(65) Zhang, J.; Tan, J.; Pei, R.; Ye, S. Acidic environment significantly alters aggregation pathway of human islet amyloid polypeptide at negative lipid membrane. *Langmuir* **2020**, *36*, 1530−1537.
(66) Santos, J.; Iglesias, V.; Santos-Suárez, J.; Mangiagalli, M.; Brocca, S.; Pallarès, I.; Ventura, S. pH-dependent aggregation in intrinsically disordered proteins is determined by charge and lipophilicity. *Cells* **2020**, *9*, 145.
(67) Nguyen, P. H.; Ramamoorthy, A.; Sahoo, B. R.; Zheng, J.; Faller, P.; Straub, J. E.; Dominguez, L.; Shea, J. E.; Dokholyan, N. V.; De Simone, A.; Ma, B.; Nussinov, R.; Najafi, S.; Ngo, S. T.; Loquet, A.; Chiricotto, M.; Ganguly, P.; McCarty, J.; Li, M. S.; Hall, C.; Wang, Y.; Miller, Y.; Melchionna, S.; Habenstein, B.; Timr, S.; Chen, J.; Hnath, B.; Strodel, B.; Kayed, R.; Lesné, S.; Wei, G.; Sterpone, F.; Doig, A. J.; Derreumaux, P. Amyloid oligomers: A joint experimental/computational perspective on Alzheimer's disease, Parkinson's disease, type ii diabetes, and amyotrophic lateral sclerosis. *Chem. Rev.* **2021**, *121*, 2545−2647.
(68) Rodriguez Camargo, D. C.; Chia, S.; Menzies, J.; Mannini, B.; Meisl, G.; Lundqvist, M.; Pohl, C.; Bernfur, K.; Lattanzi, V.; Habchi, J.; Cohen, S. I.; Knowles, T. P. J.; Vendruscolo, M.; Linse, S. Surface-catalyzed secondary nucleation dominates the generation of toxic IAPP aggregates. *Front. Mol. Biosci.* **2021**, *8*, No. 1037.
(69) Cao, Q.; Boyer, D. R.; Sawaya, M. R.; Ge, P.; Eisenberg, D. S. Cryo-EM structure and inhibitor design of human iapp (amylin) fibrils. *Nat. Struct. Mol. Biol.* **2020**, *27*, 653−659.
(70) Cao, Q.; Boyer, D. R.; Sawaya, M. R.; Abskharon, R.; Saelices, L.; Nguyen, B. A.; Lu, J.; Murray, K. A.; Kandeel, F.; Eisenberg, D. S. Cryo-EM structures of hiapp fibrils seeded by patient-extracted fibrils reveal new polymorphs and conserved fibril cores. *Nat. Struct. Mol. Biol.* **2021**, *28*, 724−730.
(71) Röder, C.; Kupreichyk, T.; Gremer, L.; Schäfer, L. U.; Pothula, K. R.; Ravelli, R. B.; Willbold, D.; Hoyer, W.; Schröder, G. F. Cryo-EM structure of islet amyloid polypeptide fibrils reveals similarities with amyloid-β fibrils. *Nat. Struct. Mol. Biol.* **2020**, *27*, 660−667.
(72) Ulamec, S. M.; Brockwell, D. J.; Radford, S. E. Looking beyond the core: The role of flanking regions in the aggregation of amyloidogenic peptides and proteins. *Front. Neurosci.* **2020**, *14*, No. 611285.
(73) Wegmann, S.; Medalsy, I. D.; Mandelkow, E.; Müller, D. J. The fuzzy coat of pathological human tau fibrils is a two-layered polyelectrolyte brush. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, E313−E321.
(74) Raiteri, P.; Laio, A.; Gervasio, F. L.; Micheletti, C.; Parrinello, M. Efficient reconstruction of complex free energy landscapes by multiple walkers metadynamics. *J. Phys. Chem. B* **2006**, *110*, 3533−3539.
(75) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486−501.
(76) Scheres, S. H. W. Amyloid structure determination in relion-3.1. *Acta Crystallogr., Sect. D: Struct. Biol.* **2020**, *76*, 94−101.
(77) Fitzpatrick, A. W. P.; Falcon, B.; He, S.; Murzin, A. G.; Murshudov, G.; Garringer, H. J.; Crowther, R. A.; Ghetti, B.; Goedert, M.; Scheres, S. H. Cryo-EM structures of tau filaments from Alzheimer's disease. *Nature* **2017**, *547*, 185−190.
(78) Piana, S.; Lindorff-Larsen, K.; Shaw, D. E. How robust are protein folding simulations with respect to force field parameterization? *Biophys. J.* **2011**, *100*, L47−L49.
(79) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(80) Kawabata, T. Multiple subunit fitting into a low-resolution density map of a macromolecular complex using a gaussian mixture model. *Biophys. J.* **2008**, *95*, 4643−4658.
(81) Tribello, G. A.; Bonomi, M.; Branduardi, D.; Camilloni, C.; Bussi, G. Plumed 2: New feathers for an old bird. *Comput. Phys. Commun.* **2014**, *185*, 604−613.
(82) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015**, *1*−*2*, 19−25.
(83) Barducci, A.; Bussi, G.; Parrinello, M. Well-tempered metadynamics: A smoothly converging and tunable free-energy method. *Phys. Rev. Lett.* **2008**, *100*, No. 020603.
(84) Branduardi, D.; Bussi, G.; Parrinello, M. Metadynamics with adaptive gaussians. *J. Chem. Theory Comput.* **2012**, *8*, 2247−2254.
(85) Daura, X.; Gademann, K.; Jaun, B.; Seebach, D.; Van Gunsteren, W. F.; Mark, A. E. Peptide folding: When simulation meets experiment. *Angew. Chem., Int. Ed.* **1999**, *38*, 236−240.
(86) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(87) Lindorff-Larsen, K.; Best, R. B.; DePristo, M. A.; Dobson, C. M.; Vendruscolo, M. Simultaneous determination of protein structure and dynamics. *Nature* **2005**, *433*, 128−132.
(88) Sormanni, P.; Aprile, F. A.; Vendruscolo, M. The camsol method of rational design of protein mutants with enhanced solubility. *J. Mol. Biol.* **2015**, *427*, 478−490.
(89) Cao, P.; Tu, L.-H.; Abedini, A.; Levsh, O.; Akter, R.; Patsalo, V.; Schmidt, A. M.; Raleigh, D. P. Sensitivity of amyloid formation by human islet amyloid polypeptide to mutations at residue 20. *J. Mol. Biol.* **2012**, *421*, 282−295.
(90) Horne, R. I.; Possenti, A.; Chia, S.; Brotzakis, F.; Staats, R.; Habchi, J.; Vendruscolo, M. A machine learning approach to identify specific small molecule inhibitors of secondary nucleation in α-synuclein aggregation. *bioRxiv* **2021**, No. 2021-11.

https://doi.org/10.1021/acs.biochem.3c00010
Biochemistry 2023, 62, 2407−2416