European Biophysics Journal
https://doi.org/10.1007/s00249-023-01639-y

ORIGINAL ARTICLE

Generation of unfolded outer membrane protein ensembles defined by hydrodynamic properties

Taylor Devlin[1] · Patrick J. Fleming[1] · Nicole Loza[1] · Karen G. Fleming[1]

Received: 31 October 2022 / Revised: 23 January 2023 / Accepted: 20 February 2023

Abstract

Outer membrane proteins (OMPs) must exist as an unfolded ensemble while interacting with a chaperone network in the periplasm of Gram-negative bacteria. Here, we developed a method to model unfolded OMP (uOMP) conformational ensembles using the experimental properties of two well-studied OMPs. The overall sizes and shapes of the unfolded ensembles in the absence of a denaturant were experimentally defined by measuring the sedimentation coefficient as a function of urea concentration. We used these data to model a full range of unfolded conformations by parameterizing a targeted coarse-grained simulation protocol. The ensemble members were further refined by short molecular dynamics simulations to reflect proper torsion angles. The final conformational ensembles have polymer properties different from unfolded soluble and intrinsically disordered proteins and reveal inherent differences in the unfolded states that necessitate further investigation. Building these uOMP ensembles advances the understanding of OMP biogenesis and provides essential information for interpreting structures of uOMP-chaperone complexes.

Keywords Outer membrane protein · Sedimentation velocity · Coarse-grained molecular dynamics · Conformational ensembles

Abbreviations

AUC	Analytical ultracentrifugation
BAM	β-Barrel assembly machine
D_{MAX}	Maximum distance (Å)
IDP	Intrinsically disordered protein
MD	Molecular dynamics
NMR	Nuclear magnetic resonance
OMP	Outer membrane protein
R_G	Radius of gyration (Å)
R_T	Translational hydrodynamic radius (Å)
PPII	Polyproline II
s*	Apparent sedimentation coefficient in Svedbergs (10^{-13} s)
$s_{20,w}$	Sedimentation coefficient corrected to 20 °C in water (units = Svedbergs)
$< s_{20,w} >$	Weight average sedimentation coefficient (units = Svedbergs)
$\{s_{20,w}\}$	Sedimentation coefficient from fitting to a gaussian function in dc/dt + (units = Svedbergs)
$s_{20,w}$(0 M urea)	$s_{20,w}$ extrapolated to 0 M urea (units = Svedbergs)
S	Svedbergs
SV	Sedimentation velocity
uOMP	Unfolded outer membrane protein
VDW	van der Waals
vHW	van Holde–Weischet

Introduction

Outer membrane proteins (OMPs) in Gram-negative bacteria encounter several physical challenges to folding. After cytoplasmic translation and translocation across the inner membrane, these unfolded and hydrophobic proteins must cross the periplasm without misfolding or aggregating. The unfolded polypeptides encounter a large kinetic barrier to folding into the outer membrane (Barral et al. 2004; Gessmann et al. 2014), and the time scale for this process is minutes (Ureta et al. 2007; Costello et al. 2016). To overcome these cellular obstacles, periplasmic chaperones

Special Issue: Analytical Ultracentrifugation 2022.

✉ Karen G. Fleming
karen.fleming@jhu.edu

[1] Thomas C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, MD 21218, USA

Published online: 11 March 2023

suppress unfolded OMP (uOMP) aggregation and promote folding competent conformations before uOMPs interact with the β-barrel assembly machine (BAM), which catalyzes their folding into the outer membrane (Hagan et al. 2010; Gessmann et al. 2014; Ulrich and Rapaport 2015; Plummer and Fleming 2016; Chaturvedi and Mahalakshmi 2017; Konovalova et al. 2017; Tomasek and Kahne 2021). Unfolded OMP tendencies to misfold and aggregate as well as the overall organization of the cell envelope mean that uOMPs must exist in the periplasm in either a free or chaperone-bound state without forming toxic protein aggregates. Insight into the conformations of uOMP ensembles (Krainer et al. 2017) is of great importance for understanding OMP biogenesis in the cell envelope and for obtaining structural models of chaperone-uOMP complexes.

Several methods have been developed to generate and analyze chemically denatured states of classically folded soluble proteins (Fitzkee and Rose 2004; Jha et al. 2005; Curcó et al. 2012) or conformational ensembles of intrinsically disordered proteins (IDPs) (Pelikan et al. 2009; Langridge et al. 2014; Allison 2017; Bonomi et al. 2017; Shrestha et al. 2019; Ahmed et al. 2020; Larsen et al. 2020; Tesei et al. 2021). Coarse-grained and all-atom simulations utilizing various force fields, simulation conditions, and intramolecular restraints have been the computational foundation of such methods. The primary goal is to create a structural ensemble described by calculated properties in agreement with experimental properties (Bernadó et al. 2007; Róycki et al. 2011; Antonov et al. 2016; Shevchuk and Hub 2017; Potrzebowski et al. 2018; Shrestha et al. 2019; Bottaro et al. 2020; Ahmed et al. 2021; Tesei et al. 2021). Experimental properties indicative of the overall size and shape of a collection of unfolded, denatured, or disordered polypeptides include the radius of gyration (R_G) and maximum dimension (D_{MAX}) determined by scattering methods; rotational diffusion determined by nuclear magnetic resonance (NMR) or fluorescence methods; translational diffusion determined by NMR or single-molecule fluorescence methods; and the sedimentation coefficient determined by analytical ultracentrifugation (AUC). To match computational and experimental values, simulation conditions are configured to bias the resulting ensemble toward the experimental value directly, or alternately, the initial unbiased ensemble is trimmed or weighted to obtain agreement with experimental data.

We capitalized upon the overall process described above to develop a simulation procedure that creates uOMP ensembles consistent with experimentally determined hydrodynamic properties. The procedure described here uses a two-step protocol with a coarse-grained molecular dynamics (MD) first step to create a well-sampled conformational ensemble followed by an all-atom MD second step to relax the stereochemistry of the polypeptide.

Unlike typical soluble proteins or IDPs, OMPs contain several hydrophobic segments corresponding to transmembrane strands. Therefore, we reason that hydrophobic interactions play a significant role in dictating the structural properties of the uOMP ensemble in aqueous solutions. To control the degree of hydrophobic collapse during simulations, we used the coarse-grained molecular dynamics simulation software application CafeMol (Kenzaki et al. 2011) for our initial simulations of the unfolded state. CafeMol has an easily configured force field with a tunable hydrophobic potential term and does not require solvent atoms. To experimentally capture the average size and shape of an uOMP, we used sedimentation velocity analytical ultracentrifugation (SV-AUC) to determine sedimentation coefficients as a function of chemical denaturant. Although scattering methods are most frequently employed for defining ensembles of soluble proteins (Riback et al. 2017; Bowman et al. 2020; Ahmed et al. 2021), the high concentrations required for those experiments are not accessible to unfolded membrane proteins. In contrast, sedimentation velocity can be conducted at much lower protein concentrations below the threshold for aggregation. Subsequent comparison of experimental and calculated sedimentation coefficients for the thousands of models generated during a simulation requires a rapid method to calculate hydrodynamic properties. For this purpose, we used HullRad, a fast and accurate program that works with both folded and unfolded protein structural models (Fleming and Fleming 2018). The final ensembles of the two uOMPs described here are more compact than unfolded soluble protein ensembles and reveal variations in the unfolded state properties attributable to either amino acid composition or sequence differences that warrant further investigation. The simplicity of this method allows for a more extensive survey of uOMPs to more deeply understand unanswered questions in the field of OMP biogenesis such as how chaperones recognize and bind their uOMP clients, whether OMP sequences are optimized to prevent aggregation in the unfolded state, the possibility of intrinsic structure in the unfolded state, and the importance of solvent quality in the periplasm.

Materials and methods

Sedimentation velocity as a function of urea concentration

Both unfolded $OmpA_{171}$ ($uOmpA_{171}$, the β-barrel domain of OmpA only) and unfolded OmpX (uOmpX) were diluted to 2 μM in 1, 2, 4, 6, or 8 M urea with either a 20 mM Tris or 20 mM sodium phosphate, pH 8 background buffer for SV-AUC. Samples were prepared and centrifuged in triplicate. All SV-AUC experiments were performed using a

Beckman XL-A ultracentrifuge (Beckman Coulter) and cells with 1.2 cm double-sector epoxy centerpieces and sapphire windows. Each sample was centrifuged at 25 °C using a 4-hole, An-60Ti rotor at a rotor speed of 50,000 rpm. Radial scans at 230 nm were acquired with 0.003 cm radial steps in continuous mode with no delay between scans. Prior to starting each run, the rotor was temperature equilibrated in the instrument for at least 60 min. A protein concentration of 2 μM was chosen for these two uOMPs because this concentration is below the threshold aggregation concentration at 1 M urea (Tan et al. 2010; Danoff and Fleming 2011) and in the linear absorbance range at 230 nm for both. The extinction coefficients at 280 nm of $uOmpA_{171}$ and uOmpX are 45,090 M^{-1} cm^{-1} and 31,860 M^{-1} cm^{-1}, respectively.

All SV-AUC data sets were analyzed using dc/dt + (Philo 2006). Sedimentation coefficient distributions ($g(s^*)$ distributions) were corrected to 20 °C in water using the appropriate densities (ρ), viscosities (η), and partial specific volumes ($\bar{\nu}$) for each buffer and protein calculated using SEDNTERP (Laue et al. 1992). These values are presented in Table S1, which also shows the buffer-corrected experimental weight average sedimentation coefficient ($< s_{20,w} >$), buffer-corrected sedimentation coefficient determined by fitting to a gaussian function in dc/dt + ($\{s_{20,w}\}$), and the experimentally determined molar masses. Plots of $\{s_{20,w}\}$ versus urea concentration were fit to a line ($y = mx + y_0$) to extrapolate the sedimentation coefficient of the uOMP in 0 M urea ($s_{20,w}$(0 M urea)) from the y-intercept. Errors reported on the y-intercept represent the 95% confidence intervals from globally fitting all sedimentation coefficients collected in triplicate for each protein at each urea concentration.

To ensure that the SV-AUC datasets represent the sedimentation of a single, monomeric species, data of $uOmpA_{171}$ and uOmpX sedimenting in 1 M urea was also analyzed in SEDANAL (Stafford and Sherwood 2004), by the $c(s)$ method in sedfit (Schuck 2000), and using the van Holde–Weischet (vHW) method in UltraScan III (Demeler and Van Holde 2004; Demeler 2005). All four analysis methods indicate that the uOMP ensembles of both $uOmpA_{171}$ and uOmpX behave as a single species whose molar masses match those expected for monomeric protein (molar masses are 18.9 kDa and 16.5 kDa for OmpA and OmpX, respectively) and whose sedimentation coefficients are consistent across analysis methods (Figure S1).

Generation of structural ensembles

For each protein, a heavy atom model was built from the UNIPROT sequence of the protein using PyMOL (DeLano 2015). An extended conformation was obtained using backbone torsion angles $\phi = -75°$ and $\psi = 145°$. We performed coarse-grained MD simulations on the extended protein model using CafeMol (Kenzaki et al. 2011), which first converts the protein chain into a random C_α-only chain conformation. Simulations were run at 298 K using Langevin dynamics with residue-specific mass, a flexible local potential, excluded volume repulsive interaction, and a hydrophobic interaction potential. Step size was 0.4, and total simulation steps equaled 2.5×10^7 with a conformation saved every 1000 steps resulting in 2500 frames. A series of different simulations for each protein were run at different coefficients of hydrophobic interaction (described below), and a replicate simulation was run on each protein at the optimal hydrophobic interaction coefficient to ensure the consistency of results.

Every other frame from the last 2000 frames of the saved coarse-grained trajectory was extracted using CATDCD (Humphrey et al. 1996) to ensure non-correlated sampling of the trajectory. VMD (Humphrey et al. 1996) was used to write 1000 PDB files of C_α-only structures, and PULCHRA (Rotkiewicz and Skolnick 2008) was used to rebuild side chains and back map the amino acid residue. All-atom MD simulations in an implicit solvent were carried out on each of the 1000 structures using NAMD (Phillips et al. 2005) to relax van der Waals (VDW) clash and obtain Ramachandran compatible backbone torsion angles. These simulations were controlled with CHARMM 36 (Huang and Mackerell 2013) at 298 K with Langevin temperature control under generalized Born implicit solvent conditions, ion concentration = 0.3, and α cutoff = 12.0. The system was minimized with 1000 steps, and MD continued for 25,000 steps with a 1.0 fs time step. Hydrodynamic properties were calculated using HullRad (Fleming and Fleming 2018) for the final ensemble of structures.

All computer methods described here may be carried out with Mac OS or LINUX machines; the software is freely available.

Results

$uOmpA_{171}$ and uOmpX sedimentation coefficients can be linearly extrapolated to obtain $s_{20,w}$ values in the absence of urea

To simulate unfolded outer membrane protein ensembles, experimentally derived structural properties are required. We performed sedimentation velocity on $uOmpA_{171}$ and uOmpX in several urea concentrations between 1 and 8 M. At urea concentrations below 1 M, these uOMPs begin to aggregate (Tan et al. 2010; Danoff and Fleming 2011, 2015). Both uOMPs were found to be monomeric and monodisperse at protein concentrations of 2 μM in urea concentrations as low as 1 M (Figure S1). Figure 1A, B show the raw $g(s^*)$ distributions of $uOmpA_{171}$ and uOmpX, respectively, at five different urea concentrations. The proteins both

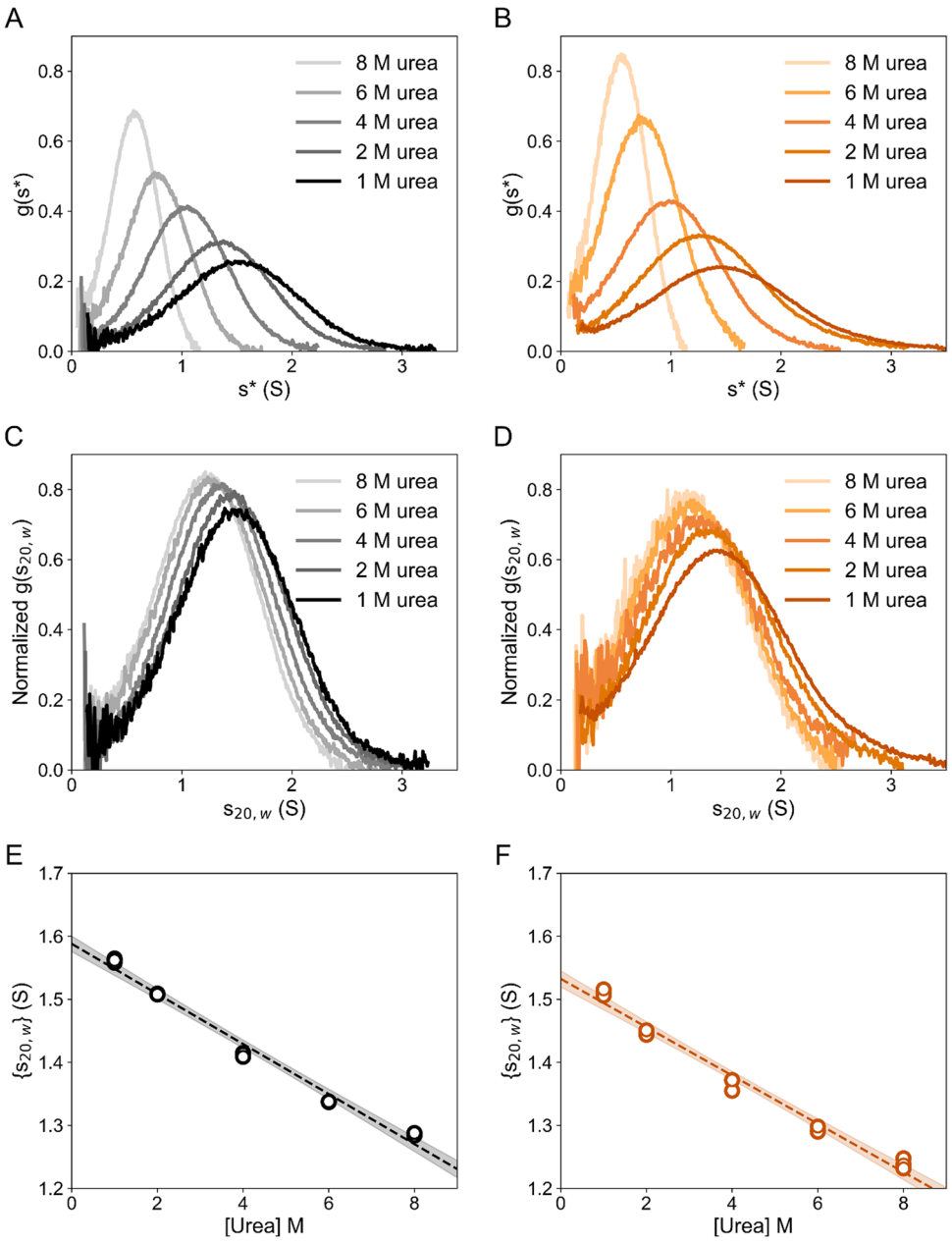


Fig. 1 The sedimentation of $uOmpA_{171}$ and uOmpX depends linearly on the urea concentration. **A** and **B** Representative $g(s^*)$ distributions of **A** $uOmpA_{171}$ and **B** uOmpX in 1, 2, 4, 6, and 8 M urea. **C** and **D** Representative plots of normalized $g(s_{20,w})$ distributions of **A** $uOmpA_{171}$ and **B** uOmpX in 1, 2, 4, 6, and 8 M urea. **E** $uOmpA_{171}$ $\{s_{20,w}\}$ versus urea concentration fits a line with the equation $y = -0.04x + 1.59$. **F** uOmpX $\{s_{20,w}\}$ versus concentration fits a line with the equation $y = -0.04x + 1.53$. For both uOMPs, the protein concentration was 2 μM in a buffer of 20 mM sodium phosphate plus the experimental urea concentration between 1 and 8 M, pH 8. Absorbance was measured at 230 nm while spinning at 50,000 rpm and 25 °C. Data were analyzed using $dc/dt+$ to determine $g(s_{20,w})$ distributions and $\{s_{20,w}\}$ values. Each experiment was performed in triplicate, and all three data points at each urea concentration are included in the liner fit. Shaded regions represent the 95% confidence interval on the fit line

sediment and diffuse more slowly at higher urea concentrations due partly to the increased density and viscosity of the solvent. However, even after correcting to $s_{20,w}$, the sedimentation of the uOMP still linearly depends on the urea concentration, indicating that urea influences the overall expansion or compaction of the uOMP conformational ensemble (i.e. the shape factor of the ensemble) (Fig. 1C, D). We used this linear urea dependence to obtain the sedimentation coefficient in the absence of urea, the intercept, termed $s_{20,w}$(0 M urea). Shown in Fig. 1E, F, the $s_{20,w}$(0 M urea) values for $uOmpA_{171}$ and uOmpX equal 1.59 and 1.53 Svedberg, respectively, and these extrapolated sedimentation coefficients were used to guide the simulated uOMP ensemble calculations. These experimental values agree with previously published results and are independent on the buffer employed (Figure S2) (Danoff and Fleming 2011). All data, fit parameters, and errors for each dataset are listed in Tables 1 and S1.

CafeMol hydrophobic interaction potentials differ for $uOmpA_{171}$ and uOmpX

Unfolded ensembles of the two model OMPs were created as described in the methods using coarse-grained MD simulations targeted to the experimentally determined $s_{20,w}$(0 M urea). The complete force field used during CafeMol

Table 1 Parameters from linearly fitting plots of $s_{20,w}$ versus urea concentration

Protein	Buffer background	Slope Svedberg/M	*Y*-intercept (95% CI) Svedberg
$uOmpA_{171}$	20 mM sodium phosphate	−0.040	1.59 (1.58–1.60)
	20 mM tris	−0.036	1.61 (1.59–1.63)
uOmpX	20 mM sodium phosphate	−0.038	1.53 (1.52–1.54)
	20 mM tris	−0.034	1.53 (1.52–1.54)

Three replicates were performed at each urea concentration, and all data were globally fit to a linear equation. Errors are reported as the 95% confidence interval

simulations consists of four pseudo-energy terms: (1) volume exclusion; (2) backbone angle; (3) backbone torsion; and (4) hydrophobic potential. The backbone angles, backbone torsions, and hydrophobic factors are amino acid specific, and default values were used. The hydrophobic interaction potential is defined as,

$$V_{\mathrm{HP}} = -c_{\mathrm{HP}} \sum_{i \in \mathrm{HP}} \varepsilon_{\mathrm{HP},A(i)} S_{\mathrm{HP}}(\rho_i) \quad (1)$$

where c_{HP} scales the overall strength of the hydrophobic interactions, $\varepsilon_{\mathrm{HP},A(i)}$ is a residue-specific hydrophobicity factor, and $S_{\mathrm{HP}(\rho_i)}$ quantifies the degree of "buriedness" of the residue (Kenzaki et al. 2011). As the targets are membrane proteins, we tuned the strength of the hydrophobic interaction potential, the c_{HP} term, to bias the ensembles during coarse-grained simulation (Fig. 2). We found that c_{HP} values equal to 0.8 and 1.1 reproduce the experimentally-observed data for $uOmpA_{171}$ and uOmpX, respectively.

We note that the short all-atom refinement step following PULCHRA is required to obtain models with realistic molecular properties. Elimination of this step results in some atomic VDW clash as well as unfavorable backbone torsion angles. Figure 3 (green dots) shows this unfavorable backbone dihedral angles for alanine residues in the initial $uOmpA_{171}$ 1000-member ensemble. These clashes are relieved by the short all-atom MD step (Fig. 3, black dots). Figure S3 shows that the Ramachandran plots of all residue types also agree with backbone ϕ and ψ values from an unbiased coil library following the short molecular dynamics step (Beck et al. 2008). The simulations for both proteins reached equilibrium (Figure S4).

Unfolded state ensembles include a wide range of conformations

Figure 4 shows the distributions of calculated sedimentation coefficients ($s_{20,w}$) for $uOmpA_{171}$ and uOmpX. $uOmpA_{171}$

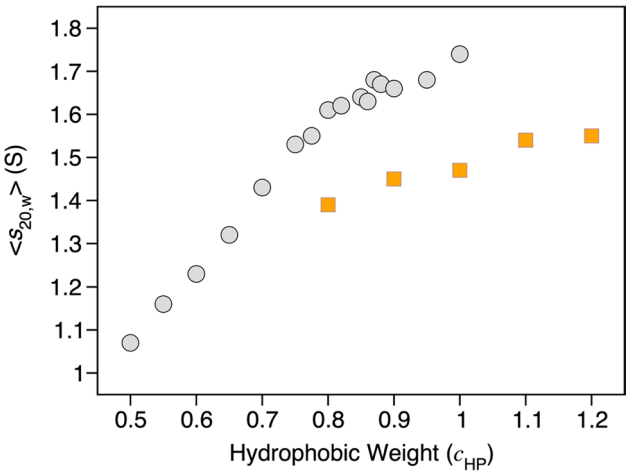


Fig. 2 The effect of the hydrophobic coefficient on the CafeMol modeled sedimentation coefficient. The value of c_{HP} during coarse-grained simulation with CafeMol varied between 0.5 and 1.2, and the average sedimentation coefficients were calculated from each ensemble of 1000 structures using HullRad and after back mapping with PULCRA. Grey data points, $uOmpA_{171}$; orange data points, uOmpX

has a non-symmetrical distribution skewed toward more expanded conformations with smaller sedimentation coefficients, whereas uOmpX exhibits an approximately normal distribution. Despite having the same number of β-strands and similar molar masses, these two uOMPs present different

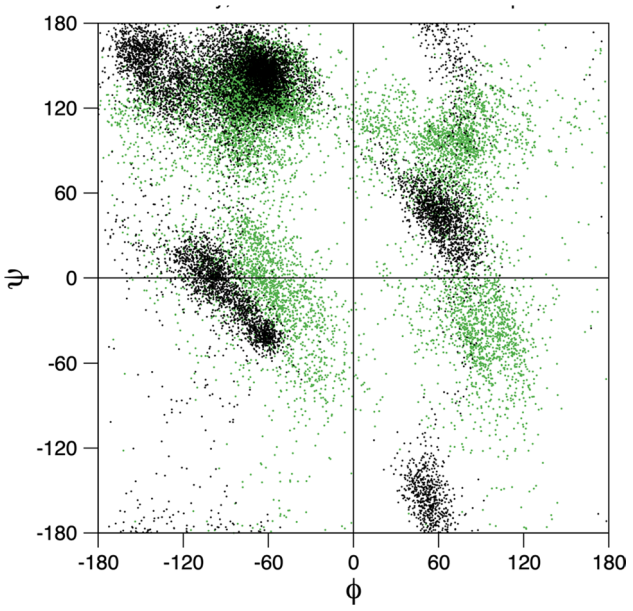


Fig. 3 Non-favorable backbone dihedral angles in initial back-mapped structures are relieved by all-atom simulation. Ramachandran plot of ALA backbone angles in a 1000-member uOMP ensemble before (green) and after (black) refinement using a short all-atom MD simulation. The black data are consistent with the ALA ϕ, ψ angle distribution observed in an unbiased coil library (Beck et al. 2008)

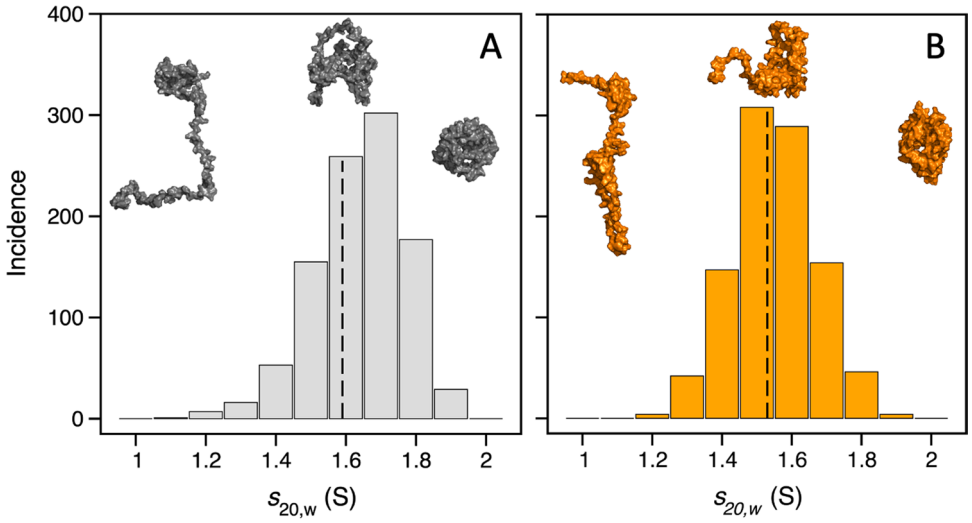


Fig. 4 The sedimentation coefficients of unfolded outer membrane protein ensembles reflect a wide range of conformations. Panels **A** and **B** are $uOmpA_{171}$ and uOmpX, respectively. Three representative atomic models from across the distribution are shown for each uOMP. The average calculated sedimentation coefficient from the distribution of $uOmpA_{171}$ is 1.59 S; a vertical dashed line marks the experimental sedimentation coefficient of 1.59 S. The average calculated sedimentation coefficient from the distribution of uOmpX is 1.50 S; a vertical dashed line marks the experimental sedimentation coefficient of 1.53 S

distributions, indicating that properties of the ensemble derive from amino acid composition or sequence. Duplicate simulations confirm this finding (data not shown). A more extensive survey of several uOMPs will be required to reveal the basis of these observed sequence-specific differences.

All models of the same protein have the same molar mass, which means that the spread of the calculated sedimentation coefficients reflects a wide range of expansion or contraction represented in the final ensembles. Figure 4 shows examples of conformations across the full range of the distribution as molecular surface models. Conformations include compact globules, slightly expanded globules, and "tadpole-like" shapes similar to those observed in IDP ensembles (Das and Pappu 2013). Supplementary movies OmpA171.mp4 (Supplementary File 2) and OmpX.mp4 (Supplementary File 3) illustrate this wide range of conformational states, and Fig. 5 displays the histograms of R_G, D_{MAX}, and *N*-terminal to *C*-terminal distance for the ensembles. The standard deviations found here are consistent with larger conformational sampling afforded by the current procedure in comparison to the sparse ensemble in Marx et al. (2020).

The computed and experimental $s_{20,w}$ *distribution widths* cannot be directly compared. The widths of the computed distributions derive from the range of accessible uOMP conformations, whereas the widths of the experimental $g(s_{20,w})$ distributions result from diffusion effects. Any contribution from conformational heterogeneity to the width of the experimental distribution is not observable due to the rapid interconversion between unfolded conformations, and the sedimentation coefficient is a weighted average of the ensemble of isomeric conformations (Scott and Winzor 2015). Thus, the computed distribution is much narrower than the experimental $g(s_{20,w})$ distribution. Future development of methods to model the diffusional broadening of molecular ensembles would be an important advance in the field.

Discussion

During OMP biogenesis in Gram-negative bacteria, unfolded OMPs traverse the aqueous periplasm before folding into the outer membrane. Periplasmic chaperones bind uOMPs during this process to prevent aggregation and further facilitate folding. To fully understand the folding pathways of these bacterial membrane proteins, we developed a method to delineate the conformational states of free uOMPs so that we may determine how binding to periplasmic chaperones alters these states. Knowledge about uOMP ensembles may even explain how the nascent proteins are recognized by periplasmic chaperones in the first place. Additionally, access to uOMP ensembles is helpful for building structural models of a chaperone-uOMP complex, as has recently been carried out for SurA binding to $uOmpA_{171}$ (Marx et al. 2020) and for Skp binding to full length uOmpA and uOmpW (Zaccai et al. 2016). Indeed, the average R_G values of $uOmpA_{171}$ and uOmpX from the ensembles in this study are consistent with the R_G of similarly sized uOmpW when bound to the chaperone Skp (Zaccai et al. 2016). This indicates that Skp binds to highly populated conformations in the unfolded state ensemble. In contrast, the unfolded conformation of

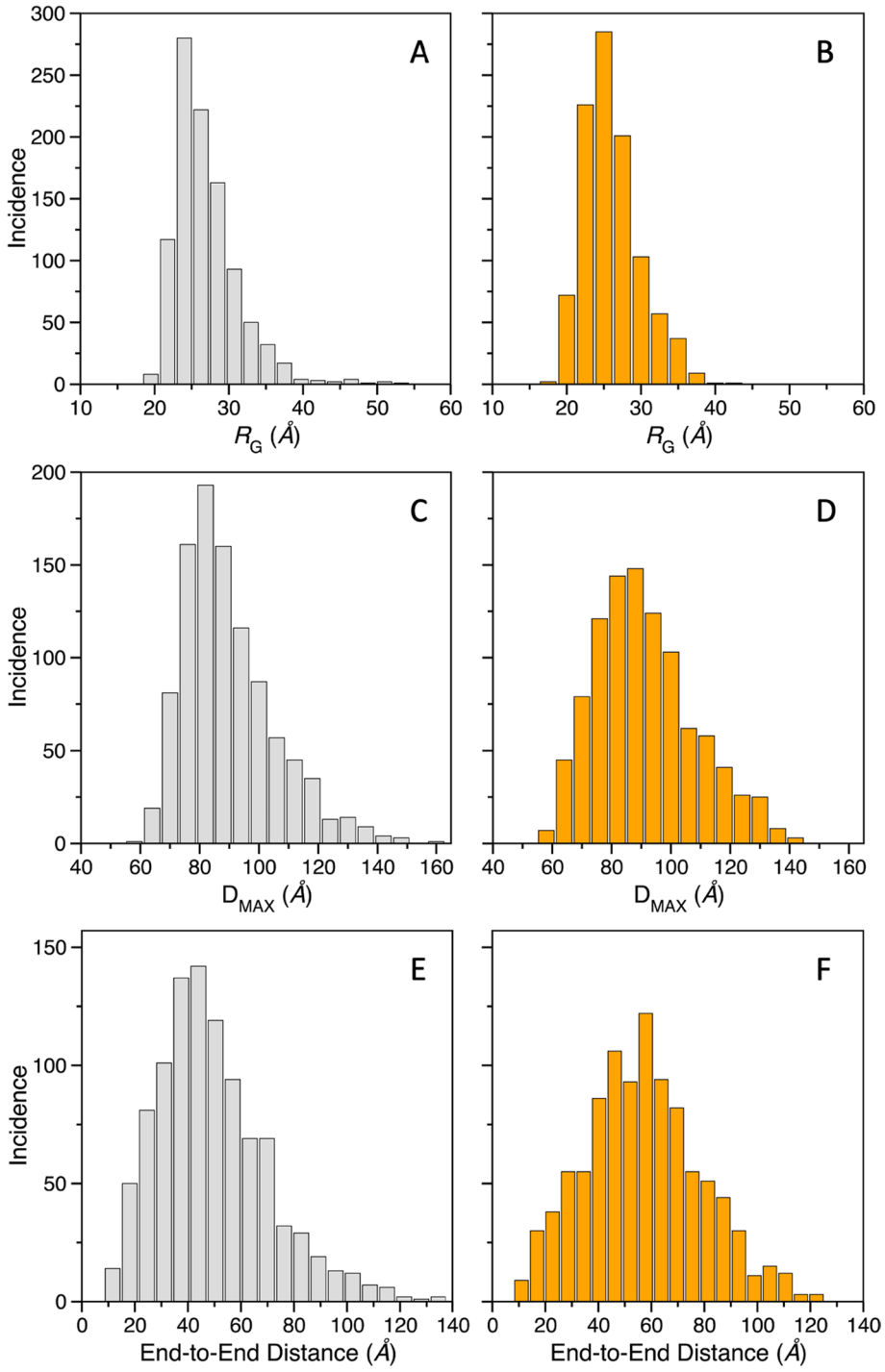


Fig. 5 Unfolded outer membrane protein ensembles defined by sedimentation coefficients show large conformational heterogenerity by several structural metrics. **A** and **B**, radius of gyration (R_G) distributions; **C** and **D**, maximum intramolecular distance (D_{MAX}) distributions; **E** and **F** *N*-terminal to *C*-terminal end-to-end distance distributions. In each panel $uOmpA_{171}$ data is gray and uOmpX data is orange. The means and standard deviations for the distributions are as follows: **A** 25.8 ± 4.4 Å; **B** 24.8 ± 3.8 Å; **C** 86.8 ± 15.8 Å; **D** 88.1 ± 16.9 Å; **E** 45.9 ± 21.3 Å; **F** 54.0 ± 21.9 Å

$OmpA_{171}$ when bound to SurA was found to be much more expanded than the average value of the current ensemble (Marx et al. 2020). One interpretation of this result is that SurA binding provides the energy to expand uOMPs. However, a minor population of extended $uOmpA_{171}$ conformations exist in the computed ensembles, so an alternative explanation is that SurA selectively binds these lowly populated, extended conformations. Thus, it is important to accurately represent the distribution of the conformations in the unfolded state ensemble as it may influence the thermodynamics of binding.

The negative slope of the uOMP sedimentation coefficient as a function of urea concentration indicates that the unfolded ensembles are more expanded in higher urea concentrations and more collapsed in lower urea concentrations. These results are consistent with the expansion of unfolded, disordered, or denatured ensembles in high concentrations of chemical denaturant reported in the literature (Sherman and

Haran 2006; Tezuka-Kawakami et al. 2006; Hofmann et al. 2012; Aznauryan et al. 2016; Zheng et al. 2016; Peran et al. 2019). The urea-dependence of the uOMP global hydrodynamic properties, like their sedimentation coefficient, serves as a reminder that the presence of even low urea concentrations may affect the properties of uOMPs as well as their binding affinities to chaperones.

The urea dependence also raises the question of solvent quality for this class of proteins. The classic experiment to determine a good solvent for a specific polymer is to apply the polymer theory analysis of Flory (Flory 1951), who showed that for a heteropolymer in solution the R_G follows the scaling law $R_G = R_0 N^{\nu}$, where N is the number of residues, R_0 is a constant related to persistence length, and ν is a factor that depends on solvent quality. Values of ν range from 0.33 for a collapsed polymer in a poor solvent, through 0.5 for a theta (or neutral) solvent, to 0.6 in a good solvent that completely "solvates" and expands the polymer. It has been shown that water is a good solvent for unfolded but foldable soluble proteins, resulting in expanded denatured state ensembles (Riback et al. 2017). But this data runs counter to widespread ideas that hydrophobic interactions drive the collapse of unfolded ensembles in water. It is not unexpected, however, that unfolded membrane proteins may experience some degree of hydrophobic collapse, which could explain their linear urea dependence. To compare the state of collapse of the uOMPs described here with other unfolded state ensembles, we plotted the average R_G values for each uOMP as a function of residue number upon the Flory scaling law functions (Fitzkee and Rose 2004). In contrast to the results shown by Riback et al., the average R_G values for both $uOmpA_{171}$ and uOmpX lie on the line defined by the scaling factor equal to 0.5 (Fig. 6). This result is consistent with uOMPs adopting conformations that are more collapsed than unfolded soluble proteins (scaling factors > 0.5) and indicates that water is a neutral solvent for uOMPs.

The comparatively reduced Flory scaling factor for uOMPs also explains a smaller-than-expected translational hydrodynamic radius (R_T) given the proportion of residues in polyproline II (PPII) conformations. The mean ϕ and ψ values for $uOmpA_{171}$ and uOmpX ensembles are plotted in Figure S5. Although the proline content for $OmpA_{171}$ is only 4.7% (OmpX = 2.7%), the fraction of residues with a PPII conformation is 12.8% for $OmpA_{171}$ and 12.1% for OmpX. This relatively large fraction of PPII in an unfolded ensemble is not necessarily unexpected (Mezei et al. 2004; Fleming et al. 2005). However, in an ensemble of IDPs, such a finding would predict expanded conformations with an average R_T of ~ 37 Å for a protein with the same number of residues as $OmpA_{171}$ (English et al. 2020). In contrast, the average HullRad-calculated R_T value for the $uOmpA_{171}$ ensemble here equals 29.0 Å reflecting a degree of compaction

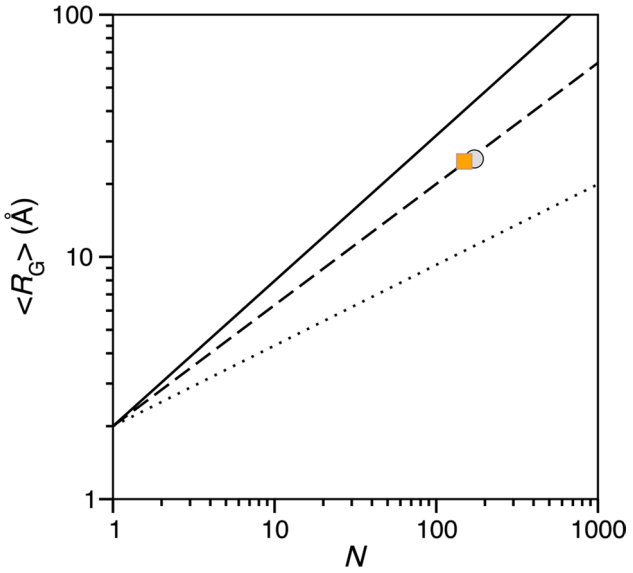


Fig. 6 The calculated radius of gyration (R_G) for uOMPs reflects ideal random-coil behavior in a good solvent. The ensemble average R_G (<R_G>) for $uOmpA_{171}$ (grey circle) and uOmpX (orange square) are plotted versus number of amino acid residues, N. The lines reflect the function $R_G = 2.0N^{\nu}$ where ν equals 0.6 (solid line), 0.5 (dashed line), 0.33 (dotted line)

greater than that found with typical IDPs. Thus, unfolded outer membrane proteins appear to have solution properties distinct from those of IDPs.

The two-part protocol described here, including both a coarse-grained and all-atom simulation, was designed to ensure quick and extensive sampling of available conformational space of the proteins. Even within the confines of allowed ϕ and ψ angles determined by the coarse-grained force field, the refined models still exhibit a wide range of backbone conformations with an average backbone dihedral angle deviation of greater than 50° across the whole protein (excepting prolines) (Fig. 7A, B). Accordingly, the ensemble also displays wide ranges of compaction or expansion and differences in R_G, D_{MAX}, and end-to-end distances (Fig. 5). Additional experimental data on intrinsic structure such as specific amino acid residue distances or backbone dihedral angles would be useful in further validating these types of ensembles.

We acknowledge that the sample size presented in this paper is small, but these initial observations are intriguing, and it will be interesting to determine whether these observations will hold true for a range of uOMPs with different molar masses. The method we have developed will facilitate future investigations of uOMPs of varying sequence and molar mass to more fully address questions concerning uOMP conformations. A more extensive survey of uOMPs would reveal any length-dependent trends in $s_{20,w}$(0 M urea) and may explore the sequence determinants of unfolded

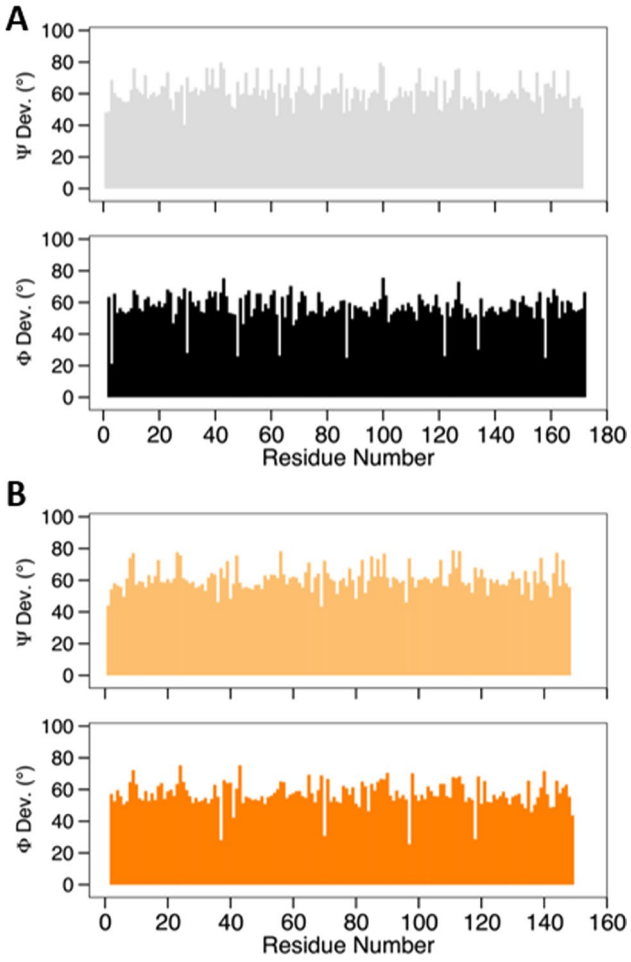


Fig. 7 Refined all-atom simulation models exhibit a wide range of backbone conformations. The average angular standard deviations of backbone ψ and ϕ angles of 1000 refined structures for **A** uOmpA171 and **B** uOmpX are plotted as bar graphs. The very low ϕ angular deviations are observed for proline residues

ensemble properties including the effects of global hydrophobic content, local clustering of hydrophobic residues, and charged residues (Bowman et al. 2020).

The use of coarse-grained simulations for the initial ensembles and HullRad to connect computational models and experimental data makes for a simple, modifiable, and time-efficient method to generate unfolded ensembles. Here we use the sedimentation coefficient to direct ensemble generation, but other hydrodynamic properties such as radius of gyration or translational diffusion coefficient may also be easily implemented in this procedure because HullRad calculates a full suite of hydrodynamic and physical properties. The procedure can also be extended to other unfolded but foldable protein systems such as unfolded membrane proteins in the mitochondrial inner membrane space due to the tunable hydrophobic potential term in CafeMol. The computationally-created ensembles of uOMPs allow for the analysis of intrinsic properties of the unfolded state, for exploring the mechanism of recognition by periplasmic chaperones, and for building structural models of chaperone-uOMP complexes.

Supplementary information

The online version contains supplementary material available as Supplementary Files 1–3.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00249-023-01639-y.

Acknowledgements This work was carried out at the Advanced Research Computing at Hopkins (ARCH) core facility (rockfish.jhu.edu), which is supported by the National Science Foundation (NSF) grant number OAC1920103. We thank members of the Fleming Lab for their helpful feedback.

Author contributions TD, PJF and KGF conceived the experimental project. KGF directed the research. TD and NL carried out laboratory experiments and analyzed the data; PJF conducted MD simulations and computational analysis; TD, PJF and KGF wrote the manuscript. All authors read and approved the final manuscript.

Funding This work was supported by NSF Grant MCB1931211 and NIH Grants R01 GM079440 (KGF). TD was supported by NIH training Grant T32-GM008403.

Availability of data and materials All data will be freely available upon request.

Declarations

Conflict of interest The authors declare that they have no competing interests.

References

Ahmed MC, Crehuet R, Lindorff-Larsen K (2020) Computing, analyzing, and comparing the radius of gyration and hydrodynamic radius in conformational ensembles of intrinsically disordered droteins. In: Kragelund BB, Skriver K (eds) Intrinsically disordered proteins: methods and protocols. Springer Science+Business Media, LLC, Berlin, pp 429–445

Ahmed MC, Skaanning LK, Jussupow A et al (2021) Refinement of α -synuclein ensembles against SAXS data: comparison of force fields and methods. Front Mol Biosci 8:1–13. https://doi.org/10.3389/fmolb.2021.654333

Allison JR (2017) Using simulation to interpret experimental data in terms of protein conformational ensembles. Curr Opin Struct Biol 43:79–87. https://doi.org/10.1016/j.sbi.2016.11.018

Antonov LD, Olsson S, Boomsma W, Hamelryck T (2016) Bayesian inference of protein ensembles from SAXS data. Phys Chem Chem Phys 18:5832–5838. https://doi.org/10.1039/c5cp04886a

Aznauryan M, Delgado L, Soranno A et al (2016) Comprehensive structural and dynamical view of an unfolded protein from the combination of FRET, NMR, and SAXS. Proc Natl Acad Sci U S A 113:E5389–E5398. https://doi.org/10.1073/pnas.1607193113

Barral JM, Broadley SA, Schaffar G, Hartl FU (2004) Roles of molecular chaperones in protein misfolding diseases. Semin Cell Dev Biol 15:17–29. https://doi.org/10.1016/j.semcdb.2003.12.010

Beck DAC, Alonso DOV, Inoyama D, Daggett V (2008) The intrinsic conformational propensities of the 20 naturally occurring amino acids and reflection of these propensities in proteins. Proc Natl Acad Sci U S A 105:12259–12264. https://doi.org/10.1073/pnas.0706527105

Bernadó P, Mylonas E, Petoukhov MV et al (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. J Am Chem Soc 129:5656–5664. https://doi.org/10.1021/ja069124n

Bonomi M, Heller GT, Camilloni C, Vendruscolo M (2017) Principles of protein structural ensemble determination. Curr Opin Struct Biol 42:106–116. https://doi.org/10.1016/j.sbi.2016.12.004

Bottaro S, Bengtsen T, Lindorff-Larsen K (2020) Integrating molecular esimulation and experimental data: a Bayesian/maximum entropy reweighting approach. Methods Mol Biol 2112:219–240. https://doi.org/10.1007/978-1-0716-0270-6_15

Bowman MA, Riback JA, Rodriguez A et al (2020) Properties of protein unfolded states suggest broad selection for expanded conformational ensembles. Proc Natl Acad Sci U S A 117:23356–23364. https://doi.org/10.1073/pnas.2003773117

Chaturvedi D, Mahalakshmi R (2017) Transmembrane β-barrels: evolution, folding and energetics. Biochim Biophys Acta Biomembr 1859:2467–2482. https://doi.org/10.1016/j.bbamem.2017.09.020

Costello SM, Plummer AM, Fleming PJ, Fleming KG (2016) Dynamic periplasmic chaperone reservoir facilitates biogenesis of outer membrane proteins. Proc Natl Acad Sci U S A 113:4794–4800. https://doi.org/10.1073/pnas.1601002113

Curcó D, Michaux C, Roussel G et al (2012) Stochastic simulation of structural properties of natively unfolded and denatured proteins. J Mol Model 18:4503–4516. https://doi.org/10.1007/s00894-012-1456-6

Danoff EJ, Fleming KG (2011) The soluble, periplasmic domain of OmpA folds as an independent unit and displays chaperone activity by reducing the self-association propensity of the unfolded OmpA transmembrane β-barrel. Biophys Chem 159:194–204. https://doi.org/10.1016/j.bpc.2011.06.013

Danoff EJ, Fleming KG (2015) Aqueous, unfolded OmpA forms amyloid-like fibrils upon self-association. PLoS ONE 10:1–11. https://doi.org/10.1371/journal.pone.0132301

Das RK, Pappu RV (2013) Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. Proc Natl Acad Sci U S A 110:13392–13397. https://doi.org/10.1073/pnas.1304749110

DeLano WL (2015) The PyMOL molecular graphics system (2.5.0). Schrodinger, LLC, Vienna

Demeler B (2005) UltraScan: a comprehensive data analysis software package for analytical ultracentrifugation experiments. In: Scott DJ, Harding SE, Rowe AJ (eds) Modern analytical ultracentrifugation: techniques and methods. Royal Society of Chemistry (UK), Cambridge, pp 210–229

Demeler B, Van Holde KE (2004) Sedimentation velocity analysis of highly heterogeneous systems. Anal Biochem 335:279–288. https://doi.org/10.1016/j.ab.2004.08.039

English LR, Voss SM, Tilton EC et al (2020) Impact of heat on coil hydrodynamic size yields the energetics of denatured state conformational bias. J Phys Chem B 123:10014–10024. https://doi.org/10.1021/acs.jpcb.9b09088.Impact

Fitzkee NC, Rose GD (2004) Reassessing random-coil statistics in unfolded proteins. Proc Natl Acad Sci U S A 101:12497–12502. https://doi.org/10.1073/pnas.0404236101

Fleming PJ, Fleming KG (2018) HullRad: fast calculations of folded and disordered protein and nucleic acid hydrodynamic properties. Biophys J 114:856–869. https://doi.org/10.1016/j.bpj.2018.01.002

Fleming PJ, Fitzkee NC, Mezei M et al (2005) A novel method reveals that solvent water favors polyproline II over β-strand conformation in peptides and unfolded proteins: conditional hydrophobic accessible surface area (CHASA). Protein Sci 14:111–118. https://doi.org/10.1110/ps.041047005

Flory PJ (1951) The configuration of real polymer chains. J Chem Phys 19:1315–1316. https://doi.org/10.1063/1.1748031

Gessmann D, Chung YH, Danoff EJ et al (2014) Outer membrane β-barrel protein folding is physically controlled by periplasmic lipid head groups and BamA. Proc Natl Acad Sci U S A 111:5878–5883. https://doi.org/10.1073/pnas.1322473111

Hagan CL, Kim S, Kahne D (2010) Reconstitution of outer membrane protein assembly from purified components. Science (80-) 328:890–892. https://doi.org/10.1126/science.1188919

Hofmann H, Soranno A, Borgia A et al (2012) Polymer scaling laws of unfolded and intrinsically disordered proteins quantified with single-molecule spectroscopy. Proc Natl Acad Sci U S A 109:16155–16160. https://doi.org/10.1073/pnas.1207719109

Huang J, Mackerell AD (2013) CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. J Comput Chem 34:2135–2145. https://doi.org/10.1002/jcc.23354

Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14:33–38

Jha AK, Colubri A, Freed KF, Sosnick TR (2005) Statistical coil model of the unfolded state: resolving the reconciliation problem. Proc Natl Acad Sci U S A 102:13099–13104. https://doi.org/10.1073/pnas.0506078102

Kenzaki H, Koga N, Hori N et al (2011) CafeMol: a coarse-grained biomolecular simulator for simulating proteins at work. J Chem Theory Comput 7:1979–1989. https://doi.org/10.1021/ct2001045

Konovalova A, Kahne DE, Silhavy TJ (2017) Outer membrane biogenesis. Annu Rev Microbiol 71:539–556. https://doi.org/10.1146/annurev-micro-090816-093754

Krainer G, Gracia P, Frotscher E et al (2017) Slow interconversion in a heterogeneous unfolded-state ensemble of outer-membrane phospholipase A. Biophys J 113:1280–1289. https://doi.org/10.1016/j.bpj.2017.05.037

Langridge TD, Tarver MJ, Whitten ST (2014) Temperature effects on the hydrodynamic radius of the intrinsically disordered N-terminal region of the p53 protein. Proteins Struct Funct Bioinforma 82:668–678. https://doi.org/10.1002/prot.24449

Larsen AH, Wang Y, Bottaro S et al (2020) Combining molecular dynamics simulations with small-angle X-ray and neutron scattering data to study multi-domain proteins in solution. PLoS Comput Biol 16:1–29. https://doi.org/10.1371/journal.pcbi.1007870

Laue TM, Shah BD, Ridgeway TM, Pelletier SL (1992) Computer-aided interpretation of analytical sedimentation data for proteins. In: Harding S, Rowe A, Hoarton J (eds) Analytical ultracentrifugation in biochemistry and polymer science. Royal Society of Chemistry, Cambridge, pp 90–125

Marx DC, Plummer AM, Faustino AM et al (2020) SurA is a cryptically grooved chaperone that expands unfolded outer membrane proteins. Proc Natl Acad Sci U S A 117:28026–28035. https://doi.org/10.1073/pnas.2008175117

Mezei M, Fleming PJ, Srinivasan R, Rose GD (2004) Polyproline II helix is the preferred conformation for unfolded polyalanine in water. Proteins Struct Funct Genet 55:502–507. https://doi.org/10.1002/prot.20050

Pelikan M, Hura GL, Hammel M (2009) Structure and flexibility within proteins as identified through small angle X-ray scattering. Gen Physiol Biophys 28:174–189. https://doi.org/10.4149/gpb_2009_02_174

Peran I, Holehouse AS, Carrico IS et al (2019) Unfolded states under folding conditions accommodate sequence-specific conformational preferences with random coil-like dimensions. Proc Natl

Acad Sci U S A 116:12301–12310. https://doi.org/10.1073/pnas.1818206116
Phillips JC, Braun R, Wang W et al (2005) Scalable molecular dynamics with NAMD. J Comput Chem 26:1781–1802. https://doi.org/10.1002/jcc.20289
Philo JS (2006) Improved methods for fitting sedimentation coefficient distributions derived by time-derivative techniques. Anal Biochem 354:238–246. https://doi.org/10.1016/j.ab.2006.04.053
Plummer AM, Fleming KG (2016) From chaperones to the membrane with a BAM! Trends Biochem Sci 41:872–882. https://doi.org/10.1016/j.tibs.2016.06.005
Potrzebowski W, Trewhella J, Andre I (2018) Bayesian inference of protein conformational ensembles from limited structural data. PLoS Comput Biol 14:1–27. https://doi.org/10.1371/journal.pcbi.1006641
Riback JA, Bowman MA, Zmyslowski AM et al (2017) Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water. Science (80-) 358:238–241
Rotkiewicz P, Skolnick J (2008) Fast procedure for reconstruction of full-atom protein models from reduced representations. J Comput Chem 29:1460–1465. https://doi.org/10.1002/jcc
Róycki B, Kim YC, Hummer G (2011) SAXS ensemble refinement of ESCRT-III CHMP3 conformational transitions. Structure 19:109–116. https://doi.org/10.1016/j.str.2010.10.006
Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. Biophys J 78:1606–1619. https://doi.org/10.1016/S0006-3495(00)76713-0
Scott DJ, Winzor DJ (2015) Characterization of intrinsically disordered proteins by analytical ultracentrifugation, 1st edn. Elsevier Inc, Amsterdam
Sherman E, Haran G (2006) Coil—globule transition in the denatured state of a small protein. Proc Natl Acad Sci U S A 103:11539–11543
Shevchuk R, Hub JS (2017) Bayesian refinement of protein structures and ensembles against SAXS data using molecular dynamics. PLoS Comput Biol 13:1–27. https://doi.org/10.1371/journal.pcbi.1005800
Shrestha UR, Juneja P, Zhang Q et al (2019) Generation of the configurational ensemble of an intrinsically disordered protein from unbiased molecular dynamics simulation. Proc Natl Acad Sci U S A 116:20446–20452. https://doi.org/10.1073/pnas.1907251116
Stafford WF, Sherwood PJ (2004) Analysis of heterologous interacting systems by sedimentation velocity: curve fitting algorithms for estimation of sedimentation coefficients, equilibrium and kinetic constants. Biophys Chem 108:231–243. https://doi.org/10.1016/j.bpc.2003.10.028
Tan AE, Burgess NK, DeAndrade DS et al (2010) Self-association of unfolded outer membrane proteins. Macromol Biosci 10:763–767
Tesei G, Schulze TK, Crehuet R, Lindorff-Larsen K (2021) Accurate model of liquid-liquid phase behavior of intrinsically disordered proteins from optimization of single-chain properties. Proc Natl Acad Sci U S A. https://doi.org/10.1073/pnas.2111696118
Tezuka-Kawakami T, Gell C, Brockwell DJ et al (2006) Urea-induced unfolding of the immunity protein Im9 monitored by spFRET. Biophys J 91:L42–L44. https://doi.org/10.1529/biophysj.106.088344
Tomasek D, Kahne D (2021) The assembly of β-barrel outer membrane proteins. Curr Opin Microbiol 60:16–23. https://doi.org/10.1016/j.mib.2021.01.009
Ulrich T, Rapaport D (2015) Biogenesis of β-barrel proteins in evolutionary context. Int J Med Microbiol 305:259–264. https://doi.org/10.1016/j.ijmm.2014.12.009
Ureta AR, Endres RG, Wingreen NS, Silhavy TJ (2007) Kinetic analysis of the assembly of the outer membrane protein LamB in *Escherichia coli* mutants each lacking a secretion or targeting factor in a different cellular compartment. J Bacteriol 189:446–454. https://doi.org/10.1128/JB.01103-06
Zaccai NR, Sandlin CW, Hoopes JT, et al. (2016) Deuterium labeling together with contrast variation small-angle neutron scattering suggests how Skp captures and releases unfolded outer membrane proteins. Methods Enzymol 566:159–210. https://doi.org/10.1016/bs.mie.2015.06.041
Zheng W, Borgia A, Buholzer K et al (2016) Probing the action of chemical denaturant on an intrinsically disordered protein by simulation and experiment. J Am Chem Soc 138:11702–11713. https://doi.org/10.1021/jacs.6b05443.Probing

Supplemental Information

Generation of Unfolded Outer Membrane Protein Ensembles Defined by Hydrodynamic Properties

Taylor Devlin[1], Patrick J. Fleming[1], Nicole Loza[1], Karen G. Fleming[1,2]
[1]Thomas C. Jenkins Department of Biophysics, Johns Hopkins University, Baltimore, MD 21218
[2]To whom correspondence may be addressed. Email: karen.fleming@jhu.edu

Table of Contents

Supplemental Methods

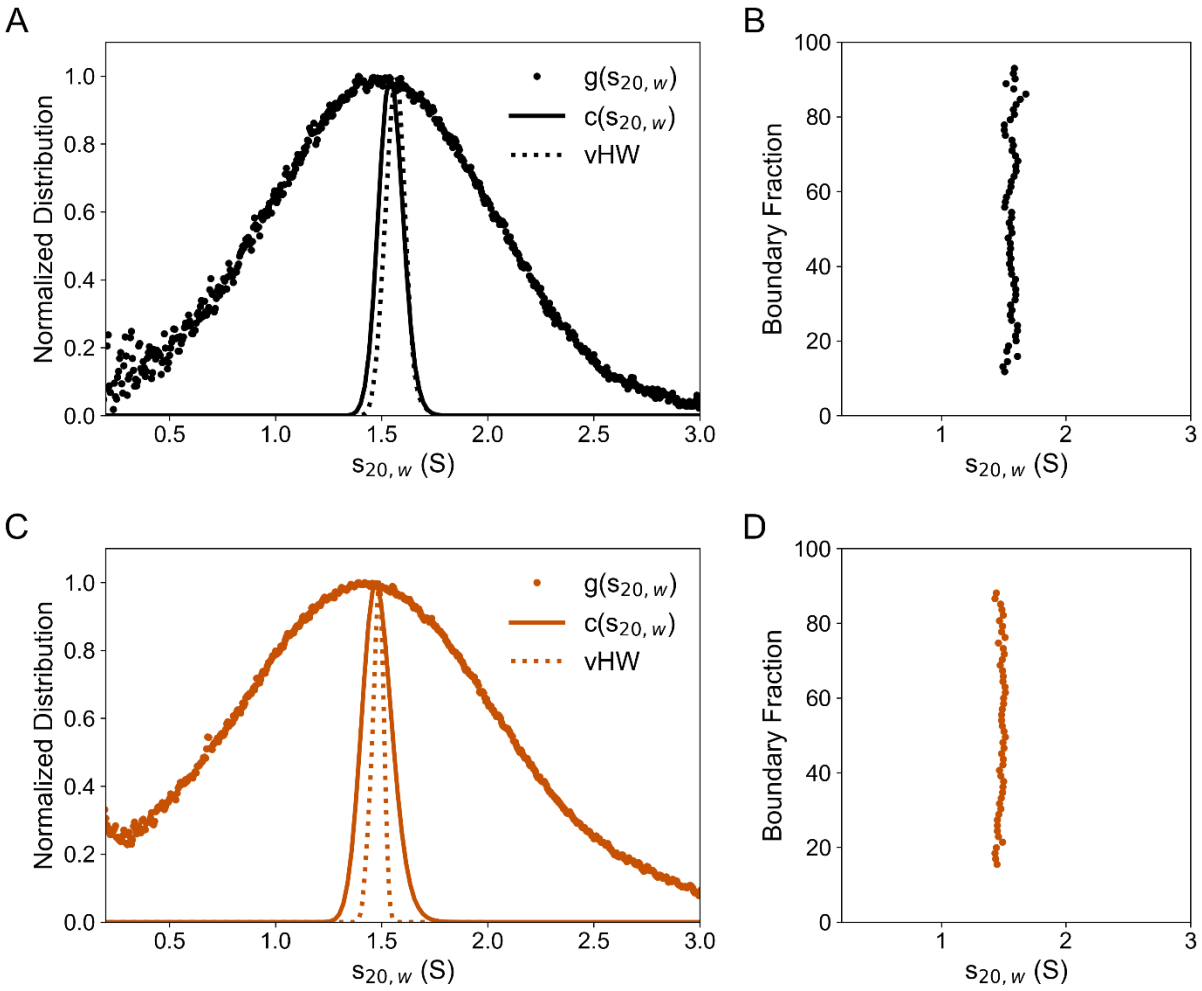
Purification of OMPs from Inclusion Bodies

Cloning of $OmpA_{171}$ into a pET11a vector with ampicillin resistance is described in (Danoff and Fleming 2011). Cloning of OmpX into a pET11a vector with ampicillin resistance is described in (Burgess et al. 2008). Glycerol stocks of plasmid transformed into HMS174(DE3) cells that had been stored at -80 °C were used to start 5 mL terrific broth (TB, Fisher) overnight cultures containing 100 ug/mL ampicillin (Sigma). The following day, overnight cultures were used to inoculate 500 mL of TB containing 100 ug/mL ampicillin, and cultures were grown at 37 °C with shaking until reaching an optical density at 600 nm (OD_{600}) of 1.0. Protein expression was induced by supplementing cultures with 1 mM isopropyl-β-D1-thiogalactopyranoside (IPTG) (ThermoScientific) and incubating them at 37 °C with shaking for an additional 4-6 hours. Cells were then harvested by centrifugation in a Beckman J2-MI centrifuge using a JA-10 rotor at 5,000 rpm and 4 °C for 30 min. Cell pellets were stored at -20 C until lysis.

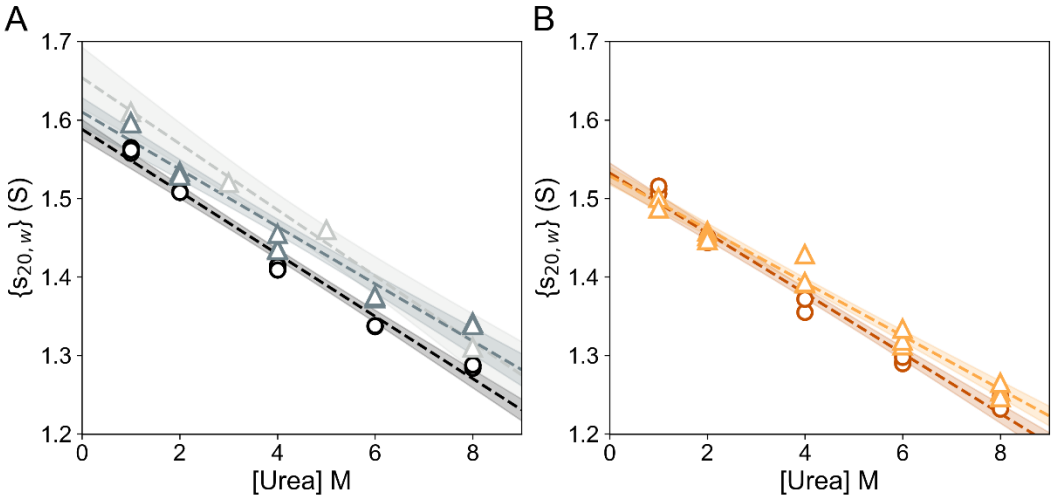
A pellet from a 500 mL growth was resuspended in 25 mL of OMP Lysis Buffer (50 mM Tris, 40 mM ethylenediaminetetraacetic acid (EDTA), pH 8) and subsequently lysed using an Avestin Emulsiflex homogenizer. After lysis, Brij-L23 (Sigma) was added to a final concentration of 0.1%, and full lysates were centrifuged (Beckman J2-MI centrifuge, JA-10 rotor) at 5,000 rpm and 4 °C for 30 minutes. Pellets containing the isolated inclusion bodies were washed twice by resuspending in 25 mL of OMP Wash Buffer (10 mM Tris, 1 mM EDTA, pH 8) and subsequently centrifuging under the same conditions listed above. The purified inclusion bodies were resuspended a final time in 10 mL of OMP Wash Buffer and were aliquoted into 1 mL portions in 1.5 mL Eppendorf tubes before centrifuging a final time at 14,000 rpm and room temperature for 30 minutes using a table-top centrifuge (Eppendorf). The supernatant was discarded, and inclusion body pellets were stored at -20 °C.

An inclusion body aliquot was resuspended in 1 mL of either 20 mM Tris or 20 mM sodium phosphate buffer at pH 8 then pelleted at 14,000 rpm and room temperature for 5 minutes using a table-top centrifuge (Eppendorf). The inclusion body pellet was dissolved in 1.2 mL of 8 M Urea in the chosen buffer background (Tris or phosphate) by incubation at room temperature for >15 minutes. Once the pellet was fully dissolved, contaminating nucleic acids were removed by centrifugation at 14,000 rpm and room temperature for 5 minutes using a table-top centrifuge (Eppendorf). The supernatant was carefully pipetted off, and final sample purity and concentration were assessed by collecting a UV-visible absorbance wavelength spectrum. Samples were deemed pure of nucleic acid contaminants if the $A_{260}/A_{280} \leq 0.6$, and protein concentration was determined using the theoretical extinction coefficients calculated using the Edelhoch method (Edelhoch 1967) in SEDNTERP (Laue et al. 1992) (ε_{280} of $OmpA_{171}$ = 45,090 M^{-1} cm^{-1} and ε_{280} of OmpX = 31,860 M^{-1} cm^{-1}). Protein was diluted with 8 M Urea in the appropriate buffer (Tris or Phosphate) to a final concentration of approximately 80 μM and then aliquoted and stored at -80 °C until use.

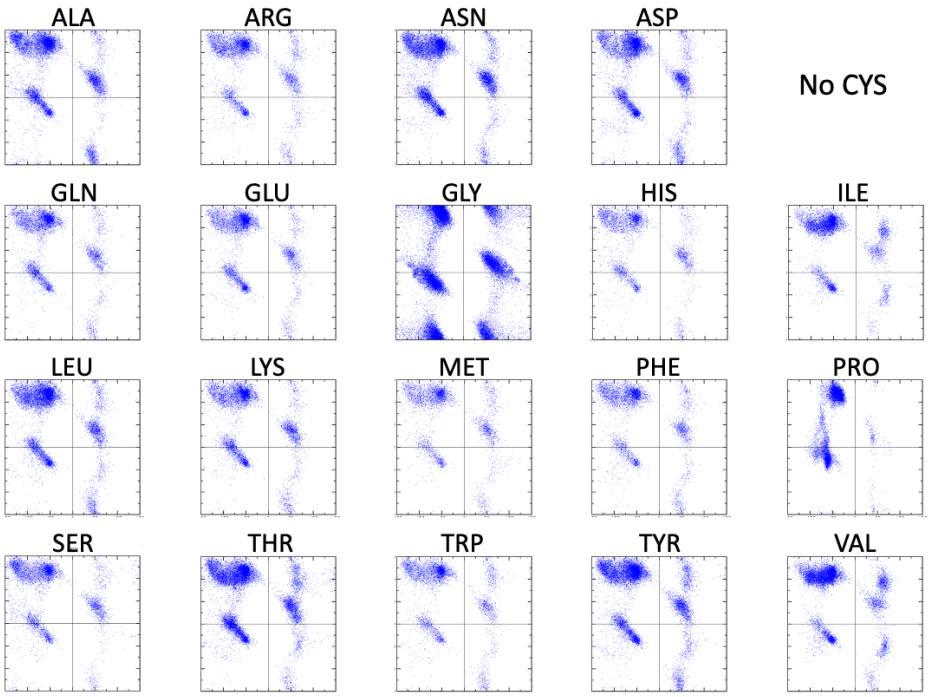
Supplemental Figures



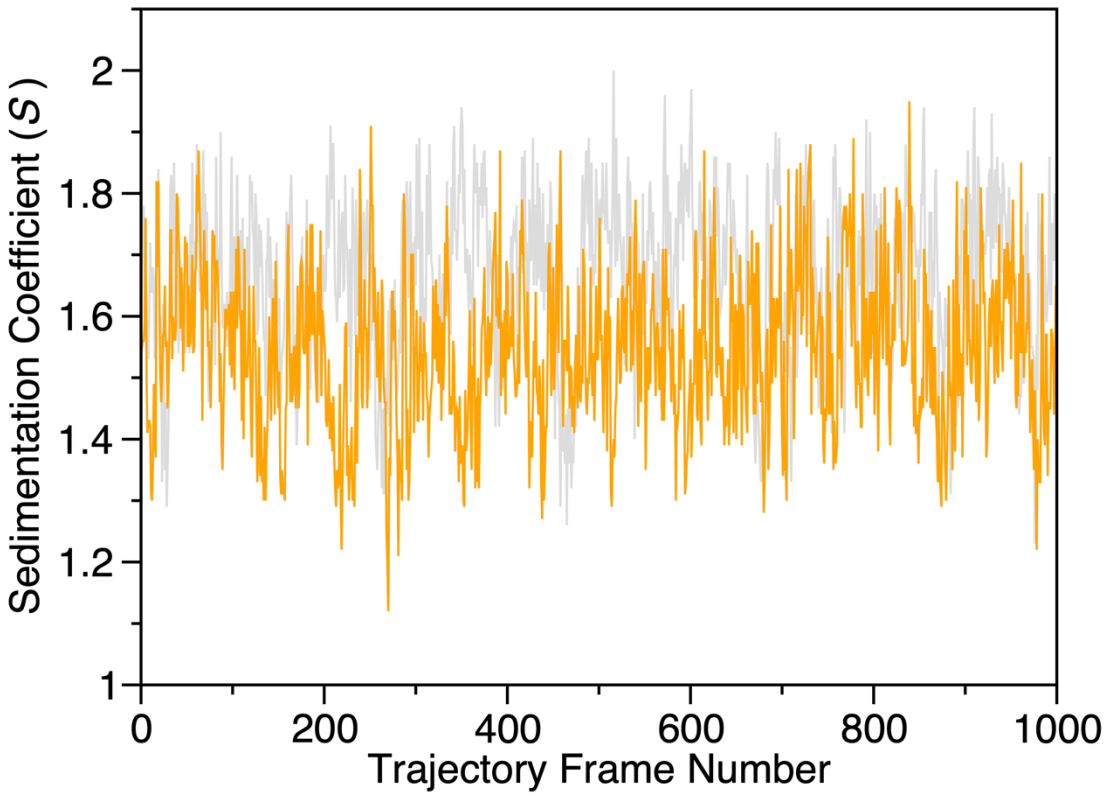
Supplemental Figure 1. uOMPs are monomeric and monodisperse in 1 M urea. (A and C) Comparison of representative $g(s_{20,w})$ (dots), $c(s_{20,w})$ (solid line), and van Holde-Weischet (dashed line) distributions of (A) $uOmpA_{171}$ or (C) uOmpX in 1 M urea analyzed by dc/dt+, sedfit, and UltraScan-III respectively. (B and D) Representative van Holde-Weischet analysis plotted as boundary fraction vs $s_{20,w}$ for (B) $uOmpA_{171}$ or (D) uOmpX. All analyses show the uOMPs behaving as a single ideal species that is monomeric based on molar mass (Table S1).



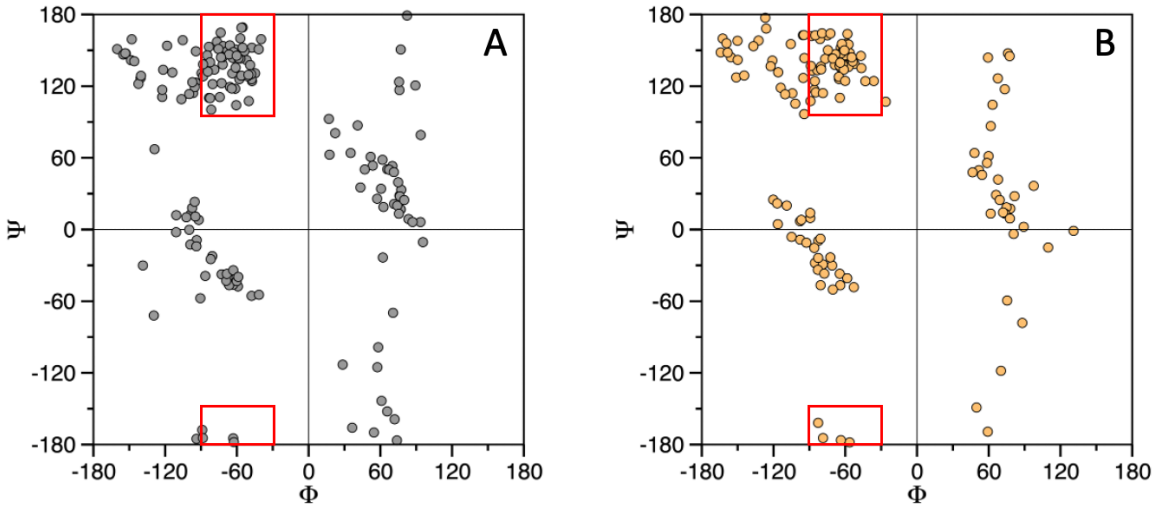
Supplemental Figure 2. The buffer background negligibly affects the sedimentation of a uOMP ensemble. For both uOMPs, the protein concentration was 2 µM in a buffer background of 20 mM sodium phosphate (dark color, open circles) or 20 mM Tris (light color, open triangles). Previously published data for $uOmpA_{171}$ in a Tris background is reproduced from Danoff and Fleming (Danoff and Fleming 2011) in the lightest grey color. In phosphate, $s_{20,w}$(0 M urea) = 1.59 (1.58-1.59) Svedbergs for (A) $uOmpA_{171}$ and $s_{20,w}$(0 M urea) = 1.53 (1.52-1.54) Svedbergs for (B) uOmpX. In tris, $s_{20,w}$(0 M urea) = 1.61 (1.59-1.63) Svedbergs for (A) $uOmpA_{171}$ and $s_{20,w}$(0 M urea) = 1.53 (1.52-1.54) Svedbergs for (B) uOmpX. Absorbance was measured at 230 nm while spinning at 50,000 rpm and 25 °C. Data were analyzed using dc/dt+ to determine $s_{20,w}$ values. Each experiment was performed in triplicate, and all three data points at each urea concentration are included in the liner fit. Shaded regions represent the 95% confidence interval on the fit line.



Supplemental Figure 3. All residue types in refined uOMP structures have favorable backbone dihedral angles. Plots of ψ versus ϕ angles for an all-atom relaxed ensemble of 1000 uOMP structures. Each plot has axes of -180° to +180°. $uOmpA_{171}$ and uOmpX do not contain cysteine. The angle distributions are qualitatively consistent with ψ, ϕ angle distributions observed in an unbiased coil library (Beck et al. 2008).



Supplemental Figure 4. Calculated sedimentation coefficient as a function of simulation time. The sedimentation coefficients calculated by HullRad for uOmpA$_{171}$ with c_{HP}=0.80 (gray lines) and uOmpX with c_{HP}=1.1 (orange lines) are plotted versus saved trajectory frame number.



Supplemental Figure 5. Ensembles of uOMPs contain well-sampled polyproline II backbone conformations. Plots of mean ψ versus mean ϕ angles for $uOmpA_{171}$ (A) and uOmpX (B) ensembles. Red boxes indicate polyproline II conformations defined by LINUS (Srinivasan et al. 2004).

Supplemental Tables

Supplemental Table 1. Summary of experimental SV-AUC values. All experiments were performed at 25 °C and analyzed in dc/dt+. All densities (ρ), viscosities (η), and partial specific volumes ($\bar{v}$) reported here assume a temperature of 25 °C and were calculated in SEDNTERP.

Protein	**Buffer**	**[Urea]** (M)	**ρ** (g/mL)	**η** (cP)	$\bar{v}$ (mL/g)	**<s>** Svedbergs	**<$s_{20,w}$>** Svedbergs	**Fit {$s_{20,w}$}** Svedbergs	**MW** (kDa)
uOmpA	Phosphate	1	1.01527	0.93316	0.7229	1.516	1.487	1.558	18.71
						1.525	1.498	1.565	18.81
						1.518	1.489	1.562	18.65
		2	1.03108	0.97648	0.7217	1.355	1.452	1.509	19.87
						1.351	1.450	1.508	19.76
						1.356	1.454	1.508	19.36
		4	1.06210	1.0856	0.7193	1.045	1.362	1.415	20.34
						1.043	1.359	1.414	20.45
						1.042	1.359	1.409	20.12
		6	1.09147	1.2459	0.7169	0.766	1.256	1.337	20.20
						0.765	1.252	1.338	20.56
						0.768	1.259	1.338	20.00
		8	1.12146	1.4854	0.7145	0.557	1.200	1.287	20.45
						0.555	1.196	1.284	20.59
						0.558	1.203	1.288	20.59
	Tris	1	1.01299	0.92900	0.7229	1.565	1.520	1.596	18.85
						1.557	1.512	1.597	19.21
						1.552	1.506	1.596	18.77
		2	1.02881	0.97232	0.7217	1.399	1.487	1.534	20.26
						1.376	1.461	1.530	20.06
						1.421	1.508	1.529	20.19
		4	1.05982	1.0815	0.7193	1.062	1.371	1.436	21.07
						1.055	1.360	1.434	20.80
						1.078	1.390	1.455	20.90
		6	1.08919	1.2418	0.7168	0.795	1.286	1.373	21.54
						0.793	1.284	1.376	21.53
						0.801	1.297	1.375	21.01
		8	1.11936	1.4813	0.7144	0.590	1.258	1.342	20.84
						0.581	1.238	1.341	20.62
						0.586	1.250	1.339	21.01
uOmpX	Phosphate	1	1.01527	0.93316	0.7155	1.522	1.491	1.512	15.21
						1.508	1.477	1.506	15.24
						1.535	1.504	1.516	15.14
		2	1.03108	0.97648	0.7142	1.340	1.432	1.448	16.21
						1.322	1.414	1.444	16.07
						1.329	1.420	1.451	16.18
		4	1.06210	1.0856	0.7116	1.016	1.315	1.355	16.44
						1.016	1.316	1.372	17.08
						1.021	1.322	1.372	17.05
		6	1.09147	1.2459	0.7090	0.752	1.216	1.292	17.36
						0.754	1.221	1.290	16.90
						0.757	1.224	1.298	17.12
		8	1.12146	1.4854	0.7064	0.547	1.159	1.248	17.50
						0.538	1.141	1.238	17.24
						0.532	1.126	1.232	16.83
	Tris	1	1.01299	0.92900	0.7155	1.463	1.420	1.501	16.50

						1.445	1.401	1.501	16.28
						1.424	1.381	1.487	16.51
		2	1.02881	0.97232	0.7142	1.302	1.375	1.458	17.37
						1.301	1.373	1.452	17.01
						1.276	1.353	1.447	16.99
		4	1.05982	1.0815	0.7115	1.054	1.352	1.429	15.62
						1.014	1.302	1.392	17.83
						1.013	1.298	1.393	17.53
		6	1.08919	1.2418	0.7088	0.755	1.210	1.313	17.62
						0.745	1.195	1.319	17.89
						0.760	1.217	1.334	17.10
		8	1.11936	1.4813	0.7062	0.554	1.161	1.254	18.02
						0.544	1.140	1.247	17.70
						0.559	1.171	1.265	17.82

Supplemental References

Beck DAC, Alonso DOV, Inoyama D, Daggett V (2008) The intrinsic conformational propensities of the 20 naturally occurring amino acids and reflection of these propensities in proteins. Proc Natl Acad Sci U S A 105:12259–12264. https://doi.org/10.1073/pnas.0706527105

Burgess NK, Dao TP, Stanley AM, Fleming KG (2008) β-Barrel proteins that reside in the Escherichia coli outer membrane in vivo demonstrate varied folding behavior in vitro. J Biol Chem 283:26748–26758. https://doi.org/10.1074/jbc.M802754200

Danoff EJ, Fleming KG (2011) The soluble, periplasmic domain of OmpA folds as an independent unit and displays chaperone activity by reducing the self-association propensity of the unfolded OmpA transmembrane β-barrel. Biophys Chem 159:194–204. https://doi.org/10.1016/j.bpc.2011.06.013

Edelhoch H (1967) Spectroscopic Determination of Tryptophan and Tyrosine in Proteins. Biochemistry 6:1948–1954. https://doi.org/https://doi.org/10.1021/bi00859a010

Laue TM, Shah BD, Ridgeway TM, Pelletier SL (1992) Computer-aided interpretation of analytical sedimentation data for proteins. In: Harding S, Rowe A, Hoarton J (eds) Analytical Ultracentrifugation in Biochemistry and Polymer Science. Royal Society of Chemistry, Cambridge, UK, pp 90–125

Srinivasan R, Fleming PJ, Rose GD (2004) Ab Initio Protein Folding Using LINUS. Methods Enzymol 383:48–66. https://doi.org/10.1016/S0076-6879(04)83003-9