Lawrence Berkeley National Laboratory
LBL Publications

Authors

Raza, Saad

Sarkar, Daipayan

Chan, Leanne Jade G

et al.

Comparative Pore Structure and Dynamics for Bacterial Microcompartment Shell Protein Assemblies in Sheets or Shells

Published as part of ACS Omega virtual special issue "3D Structures in Medicinal Chemistry and Chemical Biology".

Saad Raza, Daipayan Sarkar, Leanne Jade G. Chan, Joshua Mae, Markus Sutter, Christopher J. Petzold, Cheryl A. Kerfeld, Corie Y. Ralston, Sayan Gupta, and Josh V. Vermaas*

Cite This: https://doi.org/10.1021/acsomega.4c02406

Read Online

ACCESS | Metrics & More | Article Recommendations | SI Supporting Information

ABSTRACT: Bacterial microcompartments (BMCs) are protein-bound organelles found in some bacteria that encapsulate enzymes for enhanced catalytic activity. These compartments spatially sequester enzymes within semipermeable shell proteins, analogous to many membrane-bound organelles. The shell proteins assemble into multimeric tiles; hexamers, trimers, and pentamers, and these tiles self-assemble into larger assemblies with icosahedral symmetry. While icosahedral shells are the predominant form *in vivo*, the tiles can also form nanoscale cylinders or sheets. The individual multimeric tiles feature central pores that are key to regulating transport across the protein shell. Our primary interest is to quantify pore shape changes in response to alternative component morphologies at the nanoscale. We used molecular modeling tools to develop atomically detailed models for both planar sheets of tiles and curved structures representative of the complete shells found *in vivo*. Subsequently, these models were animated using classical molecular dynamics simulations. From the resulting trajectories, we analyzed the overall structural stability, water accessibility to individual residues, water residence time, and pore geometry for the hexameric and trimeric protein tiles from the *Haliangium ochraceum* model BMC shell. These exhaustive analyses suggest no substantial variation in pore structure or solvent accessibility between the flat and curved shell geometries. We additionally compare our analysis to hydroxyl radical footprinting data to serve as a check against our simulation results, highlighting specific residues where water molecules are bound for a long time. Although with little variation in morphology or water interaction, we propose that the planar and capsular morphology can be used interchangeably when studying permeability through BMC pores.
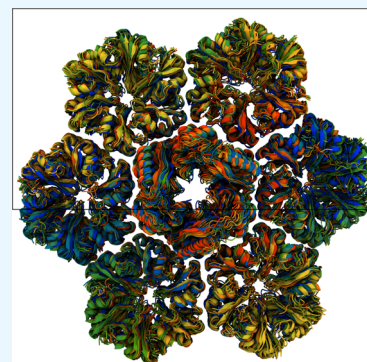
INTRODUCTION

Bacterial microcompartments (BMC) are self-assembling protein-based organelles found in various bacteria.[1,2] BMCs are thought to have evolved to facilitate catalysis in difficult or dangerous reactions. The semipermeable BMC shell protects the bacterial cytosol from the toxic effects of unstable or reactive intermediates that are present in these metabolic pathways by sequestering these intermediate products.[3] Spatially confining these reaction pathways also increases their metabolic efficiency, in part by creating local high concentrations for enzyme substrates.[4] Enzyme compartmentalization increases the rate of catabolism which increases fitness.[5] BMC shells also serve to protect encapsulated enzymes from deleterious metabolites, such as O_2 for oxygen-sensitive enzymes.[6,7] For all of these reasons, BMC shells are emerging as engineering platforms for abiotic and biotic catalysis.[8–11]

A crucial limitation for engineering new catalytic pathways into BMC shells is whether reactants and products permeate across the shell. Natural BMCs are permeable to a wide range of metabolites, such as bicarbonate and Calvin-Benson-Bassham

cycle intermediates to facilitate carbon fixation in cyanobacterial carboxysomes,[12] or reactants and products for propanediol,[13] ethanolamine,[14] and fucose or rhamnose[15] catabolic pathways. It is thought that the permeation occurs through the central pores within the individual tiles identified from molecular structures,[16,17] corroborated by molecular simulation for metabolites through these pores.[18–20]

Molecular simulations can model the permeation for any metabolite across these pores explicitly. A common approach is to use isolated protein tiles in solution as a model for the BMC shell, effectively measuring permeability in the dilute limit.[19,20] More recently, intact shells have been simulated.[18] While intact shells are a more accurate representation for the molecular

nanostructure, simulating these assemblies at the atomic scale substantially raises the cost of determining molecular permeability. The increased computational cost is particularly acute when trimeric shell components are included, as the shell size increases substantially (Figure 1). Since most of the simulation
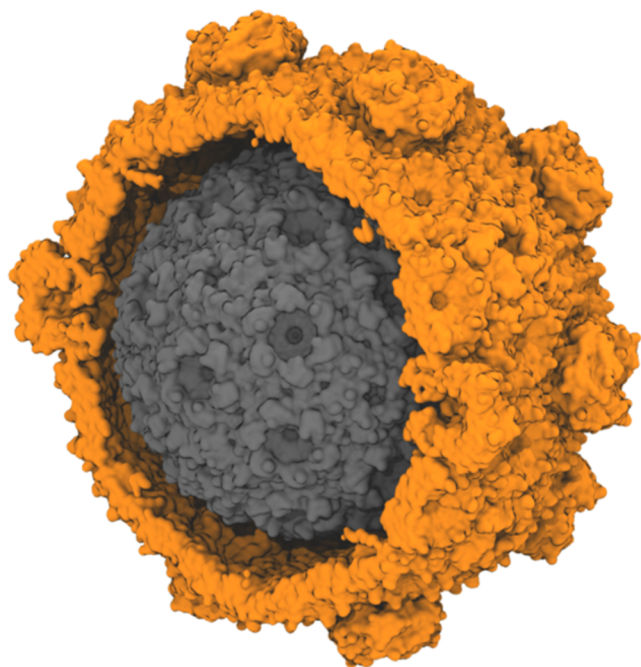

Figure 1. Size comparison for different BMC shell architectures made from *Haliangium ochraceum* shell proteins. The shown structures are determined by cryo-EM, and represent a minimal shell (gray, PDB: 6OWG, approximately 2.4 M atoms when solvated for simulation)[21] and a full HO shell (orange, PDB: 6MZX, approximately 10 M atoms when solvated for simulation).[22] The minimal shell contains only hexameric and pentameric units, while a typical full shell also includes trimeric proteins, some of which form double stacks in the shell. The diameter of the minimal model BMC is around 22 nm, while the full shell diameter is approximately 40 nm.

volume for these intact shells is water, an alternative intermediate system that reflects the symmetry of BMC shell components would be ideal to better balance computational cost and accuracy.

To explore this idea, we leverage the fact that BMC shell proteins do not always form shells. Alternative morphologies such as sheets of hexamers[23] or arrays of trimeric BMC proteins[24] have been experimentally characterized. Cylinders have been reported *in vitro*[25] after modifying the ratio of shell components for the *Haliangium ochraceum* (HO) shell.[26−28] While the HO shell is quite small, other BMCs, such as carboxysomes, form large icosahedral structures with large triangular planar facets connected together with vertices.[29] The vertices for the icosahedron are filled by pentameric shell proteins, but the approximately planar facets are thought to be a mixture of hexameric and trimeric proteins analogous to the hexameric and trimeric proteins in HO.[1,30] In this study, we design a theoretical and computationally feasible mode of the BMC shell protein in sheet conformation and we compare the dynamics and structure of a small periodic BMC sheet model developed *in silico* to an equivalent shell fragment derived directly from experimental structures.[22] Crucially, we find that the dynamics and pore diameter for planar or shell structures are

similar, indicating that the permeability for the developed sheet system is similar to the permeability that would be expected *in vivo*.

■ METHODS

Structure Preparation. Fundamentally, there are two different models prepared in this study: a curved facet from a larger BMC shell and a planar arrangement of the same proteins. The curved shell facet (labeled as shell in Figure 2) starts from
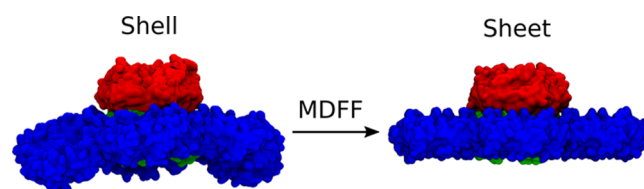

Figure 2. Using MDFF methods, we convert the shell-like curved facet of BMC (left) from a reduced PDB ID:6N0F to a sheet-like conformation (right). Hexamer tiles are in blue, and the stacked trimers are a dimer of two trimers (one red, one green), resulting in the red trimer protruding from the plane of the shell facet, in which the green trimer is embedded.

existing structures determined from cryo-EM, specifically the 6N0F structure that features a single stacked BMC trimer surrounded by BMC hexamers[22] with a closed BMC trimer. The three pentameric units at the edges of the 6N0F structure were removed from the structure prior to the simulation. The six hexamer tiles around a central trimer allow for substantial sampling for six pores simultaneously in the same simulation system. The starting structure is prepared in VMD[31] using the solvate and autoionize plugins to create a $234 \times 203 \times 234$ Å^3 system suitable for further simulation.

To create a planar sheet-like structure, we use molecular dynamics flexible fitting (MDFF)[32,33] to flatten the initial structure (Figure 2). Since MDFF requires an electron density, either real or synthetic, as the target, we generated a nominally flat target conformation for the hexamers. Starting with the curved facet, we first determine a vector normal to the trimer pore by creating vectors $\overrightarrow{v_1}$ and $\overrightarrow{v_2}$ from adjacent protein pairs in the upper trimer. The cross-product of these two vectors is normal to the trimer pore and can be brought to a specific axis using the transvecinv routine in VMD.[31] The trimer is moved to the origin, and this procedure is repeated for each hexamer individually to create the sheet arrangement. After transformation, the synthetic target density map was generated at 4 Å resolution via the mdff plugin within VMD,[31] using the combined atomic model of the flattened hexamers. The Tcl scripts for generating the flattened hexamers using rigid body transformation and generating the synthetic density map of the flattened system are available via Zenodo.[34]

In order to create an effectively infinite planar sheet of BMC shell proteins, the flattened facet was arranged in the $X−Y$ plane such that the pore normal is aligned with the Z-axis. Current molecular dynamics (MD) simulation algorithms, particularly the GPU-resident integrator of NAMD,[35] are most performant on simulation boxes with orthogonal box dimensions. We reduce the system size and maintain an orthogonal periodic boundary condition by packing three hexamer and one trimer tile into a rectangular unit cell such that the trimer tile is always surrounded by hexamer tiles as is the case in cryo-EM structure.[22] The $X−Y$ plane dimensions were set based on the
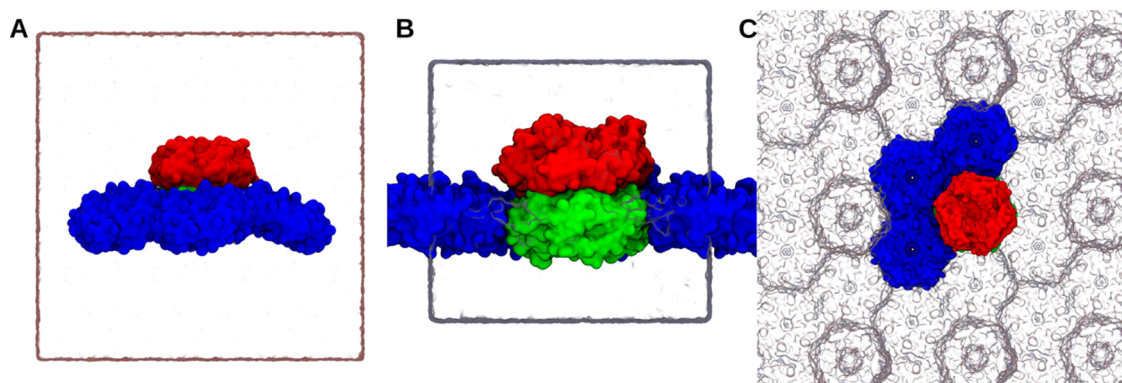
Figure 3. Simulation box of curved shell (A) and planar sheet conformation (B) with water box represented as the white rectangular which also represent the periodic boundary of the system. (C) Top view of sheet simulation system comprising three hexameric units and one trimer dimer which can mimic the planar sheet conformation and keeps the system periodic in a rectangular simulation box. Periodic images of the facets are presented as glass bubble. Hexamers are represented as blue surface representation, trimer upper half as red, and lower half as green surface representation. Planar conformation can be seen mimicking a lipid membrane where the hexameric units are touching the periodic trimeric subunits. Animations for these simulation systems are available as Animations S1 and S2.

distance between repeating units within the structure. Once optimized in this manner and solvated, the unit cell dimensions for the final structure in Figure 3B are 130 Å × 170 Å × 152 Å. Figure 3C places the repeating unit within the larger context, while Figure S1 explicitly shows a surface representation for the repeating unit and how it tiles together.

Simulation Protocol. Classical molecular dynamics simulations were carried out using the CHARMM36m protein force field[36] in explicit TIP3 solvent[37] using NAMD.[35] Minimization and a brief equilibration were performed with NAMD 2.14 prior to unbiased production simulations using the GPU-resident integrator on NAMD 3.0a9 to maximize performance.[35]

Most simulation parameters were shared when running the shell and sheet-like structures shown in Figure 3. Temperature was controlled by using the Langevin thermostat at 298 K with 1 ps^{-1} damping. Hydrogen bonds were handled with SETTLE algorithm to enable 2 fs timesteps.[38] Long-range nonbonded Lennard-Jones (LJ) cutoff was set to 12 Å. Long-range electrostatic interactions were calculated with particle mesh Ewald (PME) grid with 1.2 Å spacing.[39,40] The switching between nonbonded interaction and electrostatic is done after 10 Å. LJ correction is applied to improve energy conservation during switching.[41] Energy minimization of the system was initially performed using the 1000 steps of conjugate gradient in NAMD.[42] Both systems were briefly equilibrated for 50 ps in the NPT ensemble using 5 Å margin to allow the box to adjust after any distortion from minimization prior to transitioning to the GPU-resident integrator. Simulations in the production NPT ensemble were run for 1 μs for shell and 1.5 μs for sheet structures shown in Figure 3.

The difference between the sheet and shell structures is in the pressure control. A Langevin barostat was used to maintain pressure at 1 atm,[43] with the shell fragment using isotropic pressure control and the sheet using anisotropic pressure control. For sheet simulations, the x- and y-dimensions are tied to the expansion and contraction of the individual shell proteins, analogous to a lipid bilayer. Unlike lipid bilayer systems, where the membrane plane would be uncoupled except through the barostat, the protein itself can grow and shrink in different amounts along the axes parallel to the sheet surface. Thus, shell fragment simulations used a flexible simulation box with a fixed aspect ratio, while production simulations for the sheet structure

applied anisotropic pressure control that varied independently in every dimension.

X-ray Footprinting with Mass Spectrometry (XFMS). Two samples were prepared for X-ray Footprinting with Mass Spectrometry (XFMS). The first sample was taken from intact synthetic HO shells, assembled following the steps laid out in prior literature.[44] The second sample was the purified component HO BMC-H protein (hexamer tile), which spontaneously forms uniformly oriented sheets at high concentrations;[28,44] this was diluted to the extent that no sheets formed.

Samples were exposed at the Advanced Light Source beamline 5.3.1 with exposures of 0, 100, 250, 500, 750, 1000, and 2000 s, using a horizontal capillary as previously described.[45] Post exposure, the samples were digested using trypsin enzyme (Promega) overnight at 37 °C at pH 8 in 50 mM ammonium bicarbonate buffer. Liquid chromatography–mass spectrometry (LCMS) was conducted on an Agilent 6550 iFunnel Q-TOF mass spectrometer (Agilent Technologies, Santa Clara, CA) coupled to an Agilent 1290 LC system (Agilent). Peptide samples were loaded onto a Sigma-Aldrich Ascentis Peptides ES-C18 column (2.1 mm × 100 mm, 2.7 μm particle size; Sigma-Aldrich, St. Louis, MO) via an Infinity Autosampler (Agilent) with Buffer A (2% Acetonitrile, 0.1% Formic Acid) with flow rate 0.400 mL/min. Peptides were eluted into the mass spectrometer via a gradient with an initial condition of 5% buffer B (98% acetonitrile, 0.1% formic acid) increasing to 90% B over 15 min. The data were acquired with MassHunter B.05.00 operating in Auto MS/MS mode whereby the three most intense ions (charge states 2–5) within m/z 300–1400 mass range above a threshold of 1000 counts were selected for MS/MS analysis. MS/MS spectra were collected with the quadrupole set to a narrow resolution and collision energy to optimize fragmentation. MS/MS spectra were scanned from m/z 100 to 1700 and were collected until 40000 total counts were collected or for a maximum accumulation time of 333 ms. Parent ions were excluded for 0.1 min following MS/MS acquisition.

XFMS peptide identification and analysis was performed using the Byos (Protein Metrics, Inc.) integrated software platform at the Molecular Foundry as previously described.[46] Briefly, the abundances of the identified unmodified and modified peptides at each irradiation time point area were measured from their respective extracted ion chromatogram of
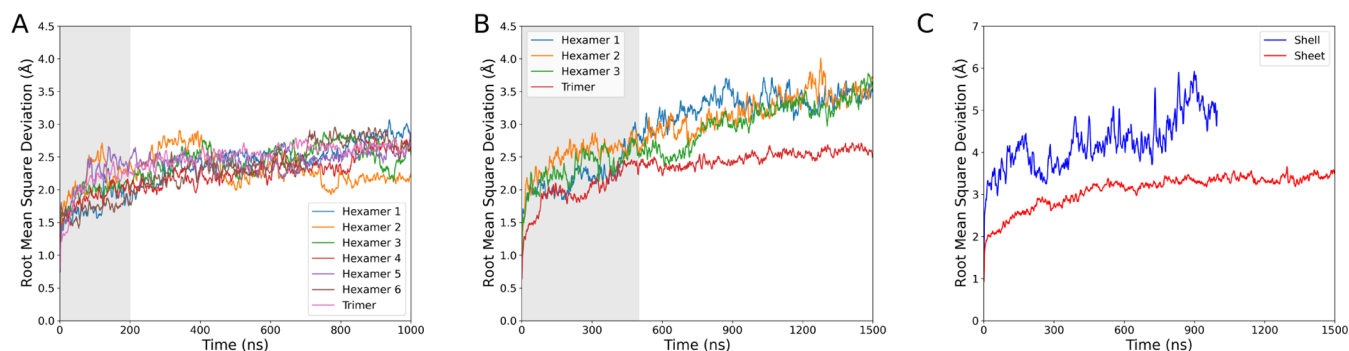
Figure 4. Root mean square deviation (RMSD) of carbon α for each hexamer and trimer when taken from the (A) shell fragment or (B) sheet simulation systems from Figure 3. (C) RMSD of the entire facet in the shell and sheet conformation. The reference structure to align to when assessing the RMSD was a tile from the initial 6N07 structure,[22] and is plotted individually for each individual tile within the system. RMSD has been smoothed with a rolling window average over 5 consecutive frames. The shaded gray area is the simulation equilibration time set for the shell (200 ns) and sheet (500 ns) conformations.

the mass spectrometry data collected in the precursor ion mode. The fraction unmodified for each peptide was calculated as the ratio of the integrated peak area of the unmodified peptide to the sum of the integrated peak areas from the modified and unmodified peptides. The dose—response curves (fraction unmodified vs X-ray exposure) were fitted to single-exponential functions, producing a k-value (s^{-1}). The ratio of k-values provided the relative change in solvent accessibility between the sheet and shell forms.

Solvent Accessibility Analysis. Structure files and MD simulation trajectories were visualized and analyzed using Python-enabled VMD 1.9.4a58.[31] Python-enabled VMD provides an interface to apply the numpy numerical library[47] and plotting tools like matplotlib.[48] System stability was assessed first by computing the root-mean-square deviation (RMSD) for the entire trajectory. Through RMSD analysis, we determined the equilibration period for each simulation system. The RMSD for the shell fragment stabilized in approximately 200 ns, while that for the sheet stabilized more slowly. For subsequent analysis, only the last 800 ns for shell and 1000 ns of sheet configuration were used. The fluctuation of each atom in the hexamer and trimer monomers was calculated using root-mean-square fluctuation (RMSF). The solvent-accessible surface area (SASA) was computed residue-wise, accelerated by a modified analysis routine that has been committed upstream to the VMD developers. Water contacts with the BMC hexamers and trimer were calculated using the contact function in VMD to track the number of unique water molecules within 5 Å of a given residue. Beyond computing water contacts, we also calculated the water retention time around the BMC hexamers and trimers. The residence time was determined by tracking frame by frame if a water molecule was initially within 5 Å of a given residue and stopping the clock when the water molecule was further than 8 Å away from the residue.

To compare directly with experimental observations based on hydroxyl radical footprinting data, we subdivided the shell fragment hexamers based on their water accessibility. For the purposes of comparison with intact shells, we evaluate hexamer monomers within the shell fragment system that are interfacing directly with the trimer tile (red in Figure S2). When compared to dilute hexameric tiles in solution, we compare with hexamer monomers in the shell fragment system that are solvent-exposed (blue in Figure S2). This facilitates a direct comparison with the companion experiment.

Pore Analysis. The primary analysis of interest is determining the pore size within a BMC protein tile. Borrowing from membrane protein studies, we used the HOLE program[49] to determine pore radius along the channel formed at the center of BMC hexamer and trimers tiles. The HOLE algorithm works by finding a maximum sphere fitting inside the cavities of protein along the z axis of the protein. The HOLE program was written to analyze a single conformation, so for full trajectory analysis, an additional wrapper is required. While other tools such as MDAnalysis have such wrappers already built-in,[50,51] optional parameters were essential to guiding HOLE along the pore of interest. In this vein, we wrote HoleHelper, a Tcl plugin to VMD that facilitates using HOLE for our specific systems with the VMD atomselection language. HoleHelper is available on github for public download and use (https://github.com/joshua-mae/HoleHelper).[52] The simulation snapshots that were judged to be equilibrated by RMSD were used to determine the pore size probability distribution.

■ RESULTS

Molecular simulation provides a unique perspective to address specific mechanical and structural questions at the nanoscale and has been called a "computational microscope".[53,54] Turning this microscope to BMC shell protein assemblies, the key question is whether the pores respond at all to the environment, similar to the opening and closing mechanism for mechanosensitive channels.[55] By also checking for stability and comparing our structures to experimental observables, we are confident that the BMC shell components are closer in nature to aquaporins and do not uniquely depend on external pressure or their environment to govern pore dynamics.

Structural Stability Considerations. Prior to any pore geometry comparison, we use the root-mean-square deviation (RMSD) over time to assess general protein stability within our simulation environment. Since the resolution for the original cryo-EM structure is 3.6 Å,[22] we anticipate an RMSD similar in magnitude to this resolution, as this relationship has been noted previously for membrane proteins.[56] That is indeed what we see in Figure 4A,B, with extended simulation only yielding RMSDs that occasionally exceed the solved structure resolution. The shell fragment routinely has lower RMSD than the sheet, suggesting that there are subtle structural changes that have occurred, as the reference structure for each tile is identical between both states. Since the RMSD change is so small, and
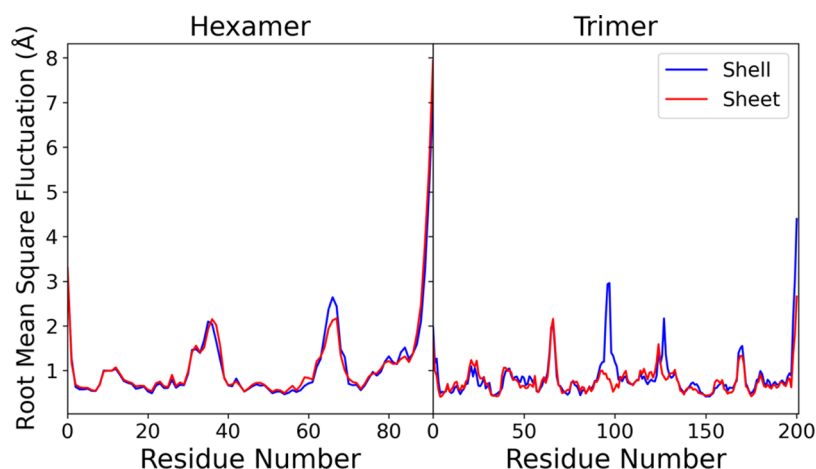
Figure 5. Root mean square fluctuation (RMSF) of C_α for each hexamer and trimer monomer when taken from the shell fragment and sheet simulation systems from Figure 3. The reference structure for alignment when assessing the RMSF was the initial simulation frame.
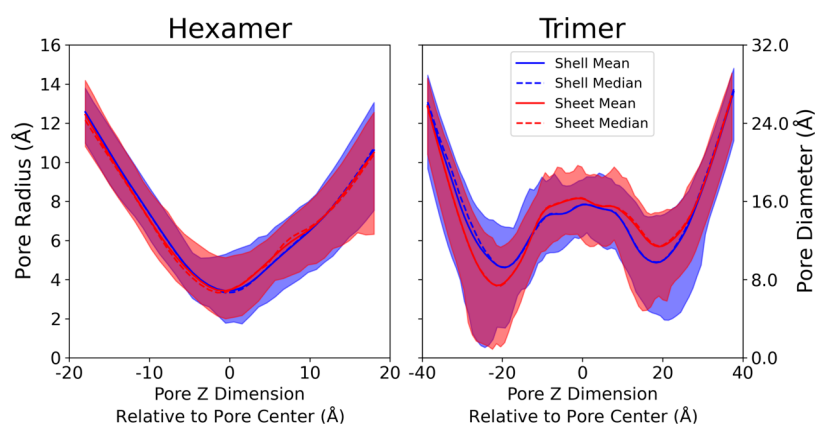

Figure 6. Distribution of pore radii along the central pore for hexameric and trimeric protein tiles in either a shell fragment or sheet morphology. In this representation, the midpoint Z is at the geometric center of the hexamers or trimers that make up the pore, which is not where the bottlenecks occur. The shaded area represents the maximum and minimum distribution of pore sizes observed during the simulation. The solid line and dashed line represent the average pore size and median pore size, respectively, observed in the simulation trajectory, across all 6 hexamer tiles (shell fragment) or 3 hexamer tiles (sheet).

largely confined to the hexamers, the overall secondary structure is consistent between states.

When computing the RMSD for the entire system, rather than of an individual tile within the assembly, the trend is reversed. The RMSD of the protein components for the full system is higher in the shell fragment as individual subunits move relative to one another during the simulation (Figure 4C). While the larger number of protein units within the shell fragment undoubtedly increase the RMSD somewhat, the overall conclusion that the structure is consistent across the simulation is unchanged.

We identified from RMSD analysis that there is still a slight upward trend in the RMSD values in sheet conformation just after simulation begins. Thus, when we calculate equilibrium properties, we do so only on the time points after 200 ns in the shell and after 500 ns in the sheet when we judge the structure to be equilibrated (Figure 4). When evaluating the RMSF on these states (Figure 5), we see only slight changes on a per-residue basis, suggesting that the dynamics are similar in both states. The largest RMSF is observed at the N and C terminal loops within the hexamer monomers. These are adjacent to regions that are not resolved by cryo-EM, and may just be naturally flexible and have high RMSF in general. The largest changes in the RMSF are

at residues 66−70 in the hexamer, which is a connecting loop between the secondary structure elements. Similarly in the trimer, the high fluctuation regions in the interior are also connecting loops between structural elements far from the central pore. The larger variation between conditions in the trimer shown in Figure 5 likely results from reduced sampling, as there are only three monomers in each system compared with many more copies of the hexamer monomer.

It is worth noting at this point that while individual proteins are reasonably intact, we also observed transient gap openings in our simulation within the assembled sheets. These gaps are large enough to let water and gases pass through the interfaces but likely still hinder the product and reactant movement through these transient channels. Thus, larger metabolites would likely be confined to the standard pores within the BMC. Although we do not see any opening in shell conformation from our simulation trajectory, prior simulations have observed water and gas permeation through transient interfaces in a complete carboxysome shell,[18] and thus this is not unexpected. We anticipate that the curvature imposed as the shell fragment was flattened into a sheet may have loosened the tight packing that typically characterizes a HO shell.
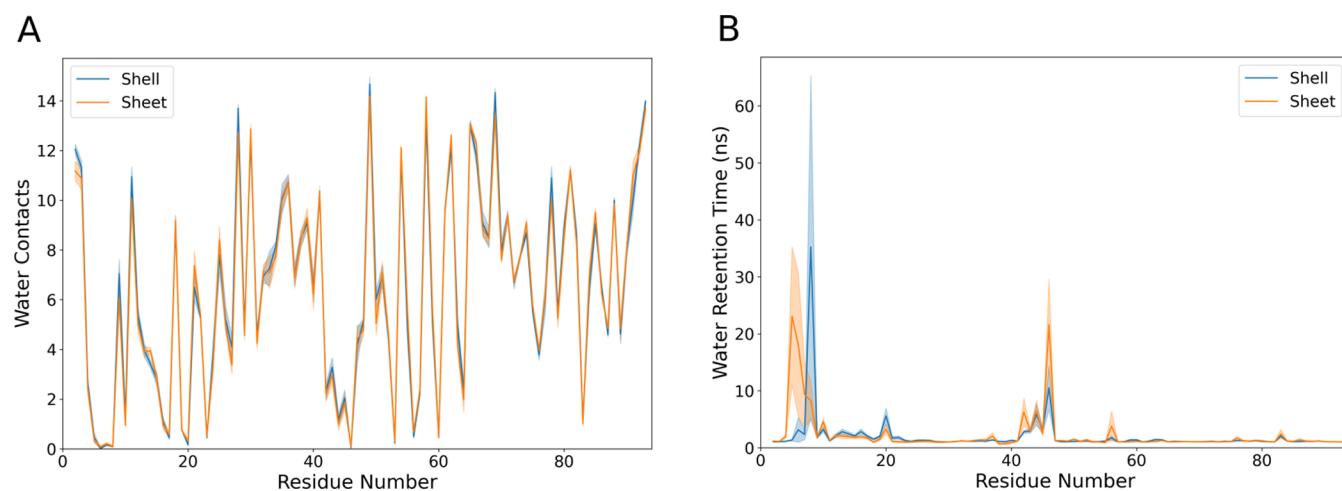
A

B

Figure 7. Hexamer tile interactions with water. (A) Quantification of the water contacts on a per-residue basis, counting the average number of pairwise contacts between amino acid residue atoms and water atoms that are under 5 Å in separation over the total trajectory. To maintain consistency with the semi-infinite sheet, the shell values reported here are averaged only over the three hexamer monomers nearest to the central trimer. (B) Measurement of the water retention times, based on how long on average a given water molecule remains within 5 Å of a given residue. The shaded area is the standard error of the mean for SASA and water retention for each residue in a hexamer monomer. Note that both quantities within the hexamer are averaged over the 18 monomers that are not solvent-exposed for the shell fragment system, and over all 18 monomers in the sheet system.

Pore Dynamics in BMC Shell Fragments and Sheets. The small variations in RMSD from Figure 4 leave open the possibility that the individual pores may change their structure when shell proteins are exposed to different local environments. Since pore size and dynamics can alter the metabolite transport, monitoring pore fluctuations over time is essential. On average, we find that the pore radii, both in their ranges and their average, are highly consistent between simulations run in either condition (Figure 6). On average, the hexamer tile has a central pore with a bottleneck diameter of 6.9 Å in the shell or 7.1 Å in the sheet, with a distribution whose mean is approximately equal to the median. This minimal change indicates that the hexamer pore is invariant to the local protein environment and exhibits similar variation in size across the simulations. The relative distribution also has a similar pattern and no large changes are seen in the shell and sheet conformations (Figure S3).

The double-stacked trimer pores exhibit substantially greater variation in the range of possible pore diameters at the bottleneck, at ±20 Å. In some conformations, particularly at the beginning of the simulation where the structure has not diverged very far from the closed starting point created by the 6N07 starting structure, the trimeric pore is effectively closed at the bottleneck. This closed pore would likely represent a large barrier to permeation of all but the smallest molecules. However, in other conformations, the trimeric pore has a substantially larger diameter. On average, the trimers expose a larger pore for metabolites to transit across. Thus, we anticipate that the trimer may be the preferred path for some molecules to permeate that cannot be accommodated by the smaller hexamer.

For small-molecule permeation, pore dynamics are essential. As the starting structure is taken from a closed starting structure, our initial structures exhibit a closed conformation for the hexamer and trimer complexes. The minimum hexamer opening often occurs at time zero, starting with a bottleneck radius of only 2 Å (Animations S1−S4). The bottleneck expands quickly as the pore hydrates and side chains rearrange to a typical radius of 5−6 Å (Figure 6). The trimeric complex exhibits even stronger dynamics, with an effectively closed pore in the initial

structure.[22] The HOLE output highlights an expanding pore over time (Animations S3 and S4).

From the starting structure, the amino acid side chains that line the pore effectively close the pore and would likely bar many small molecules from transiting the BMC shell. However, during simulation, these pore-lining residues can readily move without large secondary structure variations to a more open state, as shown in Figure S4. In our simulations, we find that once the pore is opened, the pore will not dehydrate and close again for a long time during the simulation, although side chains may sporadically occlude the pore. These briefly closed conformations are visible in the supporting animations and have included a still image of one such closed conformation as Figure S5. Shell and sheet assemblies open up in a similar manner, and both systems demonstrate sporadic pore closure. While classical simulation water models are unusual in many respects, we anticipate that in this case the pore structure observed in cryoEM was more a result of the low temperature at which the samples were collected. It is possible that the pore may reclose fully if the simulation were to be extended; however, given how few trimers we have in our simulation systems, the cost in computer time to observe reclosing events was thought to be prohibitive to depend on stochastic sampling alone to reclose the pore.

From the pore diameter distribution, we note an asymmetric distribution in the top and bottom halves of the trimer (Figure S3). BMC trimers are thought to have a gating role,[2,22,24,30,57] and so some asymmetry can be readily expected. While we are limited in our sampling for the trimer, with only a single copy in both scenarios, the observed asymmetry in Figure S3 may be real, driven by the need to gate substrate translocation.

Beyond pore dynamics, the sheet simulations in particular show small gaps that appear between tiles and these gaps are substantially larger than what is seen in the shell simulations (Animations S1 and S2). In our own internal testing, we could not identify barostat conditions that would disallow these gaps to form, suggesting that either the sheet is not packed perfectly during the flattening process or these sporadic openings are real when the highly curved HO shell is flattened. In model

carboxysomes, water has been observed to transit the shell through the interface made between the edges of BMC hexamers.[18] In our view, this suggests that these small gaps are also present in native HO shells, perhaps compounded by the flattening process to which we subjected them here.

Water Interaction Analysis. Beyond the proteins themselves, our simulation systems feature water to fill in the rest of the simulation volume. While the pore dynamics are likely most critical to permeability, observing changes in water interactions may be another avenue by which we can tease apart the subtleties of structural differences between sheet-like and shell-like structures. The per-residue water contacts vary minimally between the two tested conformations for the hexamer (Figure 7A), suggesting that even structural details are largely conserved at a global scale between the curved shell fragments and a larger sheet. This total picture of conserved contacts is retained if we expand our view to include the trimer. By mapping water contacts onto the structure, as in Figure S6, we visually see the same water contact patterning across both structures. Contacts are naturally highest on the protein periphery that are solvent-exposed, with the residues at the central bottleneck having roughly half the number of water contacts due to protein occluding many potential water interaction sites. In general, this suggests that both structures are facing the same solvent environment regardless of the exact geometry at play.

Quantifying contacts alone is only one metric of interest. With an eye toward comparing to hydroxy radical footprinting data, where the timespan individual water molecules spend near specific protein residues is of interest, we quantify the residence time for water molecules near individual residues (Figure 7B). We see some increased retention of water molecules near specific residues, with substantial differences in the range of residues 5−11 (LGMIEVR), 20 (A), 41−48 (YVTAVRGD), 50−52 (VAA) and 83 (V) (Figure 7B). The strongest difference in water retention between sheets and shells occurs around G6, where water in the sheet conformation is retained for over 30 ns on average. This residue is on the border of an interstitial water site near β sheets within an individual hexamer (Figures 8 and 9). The residues lining the pores have much shorter water interactions, as water at the interface readily exchanges with the bulk.

Correlating Simulation to XFMS. Our *in silico* work so far has emphasized that the differences between water access and pore formation are really minimal. Is it possible to use an experimental measure like XFMS, where amino acid hydroxylation induced by water ionization through X-ray exposure can be measured by mass spectrometry,[58,59] to provide experimental support for these findings? Data for the hexamer in solution or as part of a shell are provided in Table 1. The *k*-values measure the hydroxyl replacement rate at specific residues where hydroxylation is possible. The replacement rate in solution can be greatly accelerated when compared to the complete shell, such as at the M7 residue, which has a very high ratio of hydroxylation rate in solution compared to when it is in the shell. This indicated that the M7 residue has a substantially lower availability to water when it is in the shell compared with the solution state. We did not directly simulate an isolated hexameric tile in solution, and not all hexamer monomers within our curved shell fragment are a good facsimile for the environment of a complete shell. Thus, we segment our data into two populations based on hexamer identity (Figure S2) to establish comparisons with Table 1. After this segmentation process, we made comparisons with structural metrics.

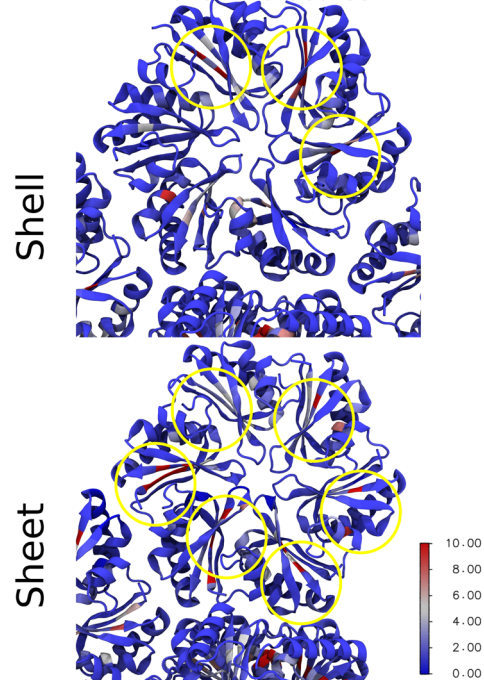## Beta Sheets Accommodating Water Molecules

Figure 8. Water retention mapped to each residue on the hexamers and trimers in shell-like and sheet morphologies. Proteins are drawn in a cartoon representation where each residue is color-coded on the blue-white-red spectrum to represent the water retention time. The color bar measures the water retention time in nanoseconds.
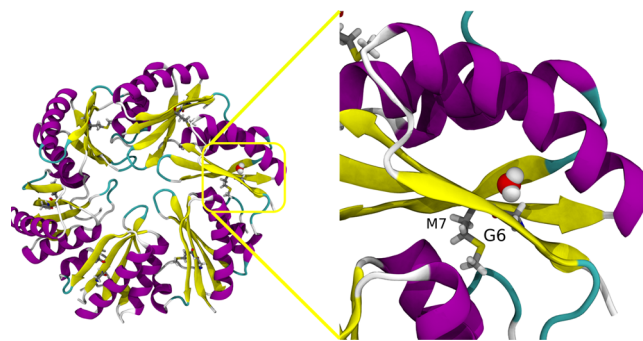

Figure 9. Retention of water in a pocket around the β sheet in hexamers near residues G6 and M7. Proteins are represented in cartoon representation, residues G6 and M7 in licorice, and water in VDW. The left panel represents the whole hexamer unit, and the right panel represents the water pocket.

Table 1. XFMS *k*-Values at Different Positions on the BMC Hexamer for a Single Hexamer Tile in Solution and as Part of a BMC Shell Assembly[a]

residues	solution *k*-value	shell *k*-value	ratio
M7	$(242.0 \pm 9.6) \times 10^{-6}$	$(29.8 \pm 8.2) \times 10^{-6}$	8.12
M16, M23	$(65.2 \pm 2.1) \times 10^{-6}$	$(51.2 \pm 1.4) \times 10^{-6}$	1.27
Y34	$(14.7 \pm 1.9) \times 10^{-6}$	$(15.50 \pm 0.71) \times 10^{-6}$	0.95
Y41	$(6.0 \pm 1.7) \times 10^{-6}$	$(3.90 \pm 0.55) \times 10^{-6}$	1.53
K54	$(14.00 \pm 0.84) \times 10^{-6}$	$(6.50 \pm 0.31) \times 10^{-6}$	2.16
P77, P79, P88	$(477.0 \pm 9.1) \times 10^{-6}$	$(187.0 \pm 8.9) \times 10^{-6}$	2.54

[a]The ratio of these two rates tells us something about how water accessibility changes based on BMC protein morphology.
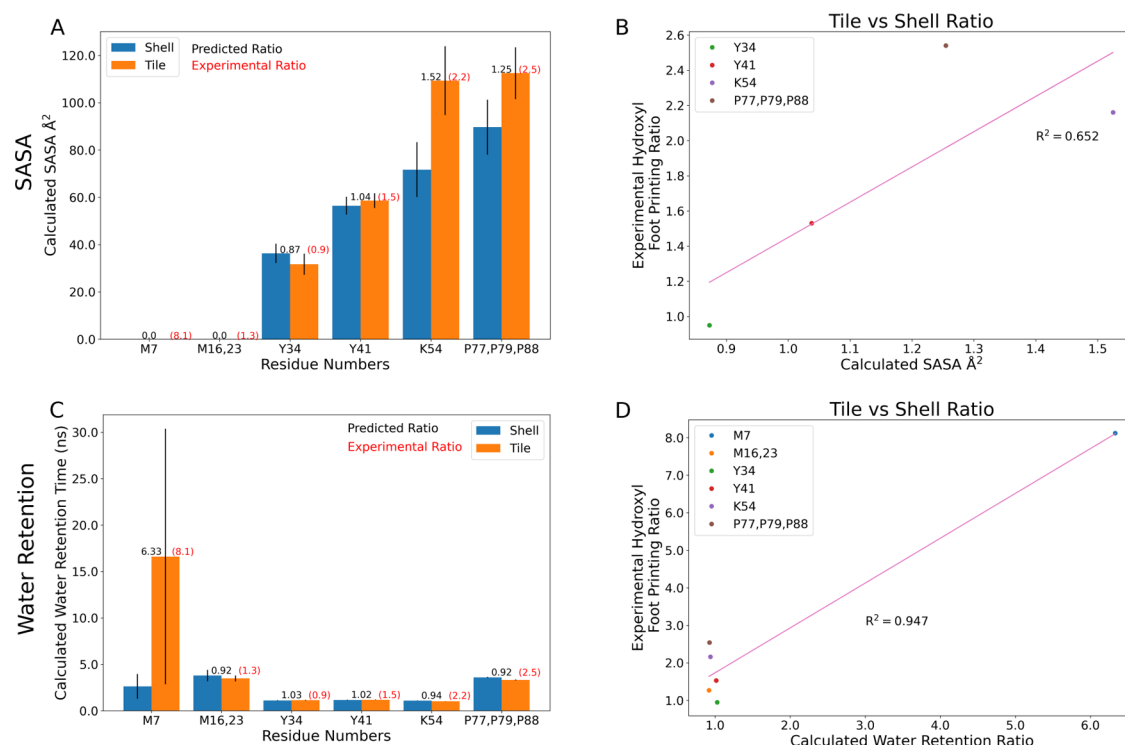
Figure 10. Measured solvent-accessible surface area (SASA) and water retention times across the residues where we have XFMS comparison data (Table 1), directly comparing the equivalent residues within our models that belong to exposed or buried monomers (Figure S2). The ratios of these quantities determined from simulation are reported in black above the histograms for (A) SASA or (C) water retention times, while the experimental equivalents are written in red. The error bar represents the standard error of the mean in SASA and water retention for the residues. The scatter plot comparing the reaction rate ratios to the ratio of (B) SASA or (D) water retention times has a line of best fit along with a correlation coefficient given in pink. For SASA correlation measures, residues M7 and M16,23 are not shown, as they have zero SASA values and thus a poorly defined ratio.

The solvent-accessible surface area (SASA) for individual residues has been previously demonstrated to correlate somewhat with hydroxyl radical footprinting data derived from XFMS.[60] There certainly is a trend when evaluating the SASA overall for specific residues (Figure 10A,B). The fit improves if M7 is excluded from consideration, which might be reasonable as our SASA determination algorithm cannot find a water-accessible surface near this buried residue. However, from Figures 7 and 8, we know that water can access these buried residues near the β-sheet. Moreover, if water accesses these residues, then they may be present near these amino acids for a considerable duration (Figure 7B). Thus, the initial outlier of M7 may be explained by tracking how long a water molecule is present as this ratio can change substantially for buried residues (Figure 10C). Indeed, taking the ratio yields a very strong correlation coefficient (Figure 10D), although this is almost entirely due to the differential water retention around M7 previously noted (Figures 7B and 8). We also mapped the residues back to the hexamer tiles, noting that only Y41 is near the pore bottleneck and might impact critical dynamics in this region (Figure S7). The water retention and SASA ratio of Y41 are around 1, both in simulation and experiment, which strongly suggests that the environment around Y41 is unaffected by the change in environment between isolated tiles and assembled shells. As mentioned previously, M7, M16, and M23 are buried in the secondary structure, and Y34 is exposed to the surface but still away from the narrowest part of the pore. K54, P77, P79, and P88 are at the interface between hexamer tiles. Although some of these residues have differences in the water retention

and SASA ratio, these residues are likely not actively engaged in pore dynamics or structure.

DISCUSSION AND CONCLUSIONS

From the outset, the primary question we were seeking to answer was if future calculations aimed at determining permeability at the molecular scale could assume that the pores are similar, irrespective of their local environment within either shell or a tile within a larger sheet. The direct evidence indicates that this is a reasonable assumption, with Figure 6 showing little difference in the pore diameter, regardless of whether the hexamer and trimer tiles are arranged as they would in the HO shell, or if they are instead tiled into a planar sheet. Given the substantial reduction in system size that a planar sheet-like arrangement provides, we anticipate using these results as the foundation for the computational simplification of determining permeability through the individual pores at the center of the abundant trimer and hexamer tiles found in many BMC shells.

Indeed, the depth at which we had to look to find any differences between shells and sheets is quite remarkable. The water contacts are the same (Figure 7A), and while we do not show it, the SASA analysis is also highly similar between sheets and shells when one looks at monomers without a solvent-exposed edge. The only difference we find is that there are sporadic trapped waters whose lifetimes are a bit longer in the sheet rather than in the shell (Figure 7B). Admittedly, the number of trapped water molecules that contribute to these long lifetimes is not large compared with the total number of water molecules that interact with BMC components. However, since

we have multiple copies of the hexamer within our system and can average over 18 monomers where waters may be trapped when conducting our analysis, we are confident that the effect is real.

Our confidence is increased by comparisons to XFMS data for isolated hexamer tiles relative to hexamers within an assembled shell (Table 1). Inferring as we do in Figure S2 that our shell fragment simulation has components that are in similar environments to both experimental systems, with a solvent-exposed edge and a buried edge to individual hexamers, the SASA correlates with XFMS data (Figure 10), analogous to prior results in other systems.[60] However, for buried residues such as M7, M16, and M23, SASA no longer corresponds to XFMS data, as the SASA is uniformly zero. In this case, water retention times, which more directly measure nanoscale interactions with water, yield better correlations (Figure 10). In particular, the correlation between long-lived water molecules near buried sites such as M7 and the change in the rate of hydroxylation is far stronger than we had anticipated. Initially, the M7 result was thought to be an outlier, but only by using molecular simulation to visualize trapped water molecules can we develop a rational basis for this result.

Zooming out, we think it is helpful to analogize how these BMC protein pores compare with those of typical membrane channels and transporters. Membrane transporters often must go through a conformational change to fulfill their function.[61] While we do see pore dynamics over our simulation, with the bottleneck radius increasing and decreasing over time (Animations S3 and S4), we find that these dynamics are primarily driven by side chain rearrangements, such as what is highlighted in Figure S4, rather than large scale conformational change as might occur in a membrane transporter. Metabolite-driven gating has been postulated for other trimer pores[24,62,63] and may well be what occurs in HO shells as well.

Thus, the closest membrane protein analogy for these BMC shell components appears to be that of a channel. Despite starting from a closed state, the trimer opens spontaneously during our simulations, suggesting that the trimer can be gated depending on conditions, analogous to gated ion channels. This was also considered as an explanation for the two particle classes observed cryo-EM studies of the synthetic HO shell.[22] The hexameric assembly has a smaller variation in pore diameter and is more analogous to a constitutively open channel, such as some aquaporins.[64,65]

While the border between channels and transporters is often murky, channels typically have higher conductances,[66] which would benefit reactant and product exchange across the BMC shell. With the membrane channel analogy in mind, we anticipate that many small molecules may transit through the central pores with a high permeability. The limiting factor will be molecular size, as we anticipate that sufficiently large molecules will be unable to transit the pore through these tiled BMC shell proteins arrangements. Now armed with a computationally efficient planar arrangement of BMC shell components, we are well positioned to test the high permeability hypothesis explicitly. When these products are tested over multiple metabolic pathways featuring different substrates and products, we hope to develop general rules for transport across BMC shells.

■ ASSOCIATED CONTENT

Data Availability Statement

ⓈI Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.4c02406 with four animations and a supporting PDF.

> Molecular dynamics simulation for the sheet system featuring hexameric and trimeric BMC shell tiles (Animation S1) (MP4)
>
> Molecular dynamics simulation of shell fragment from a larger BMC (Animation S2) (MP4)
>
> Animation depicting pore dynamics in BMC trimer and hexamer pores when taken from a shell fragment (Animation S3) (MP4)
>
> Animation depicting pore dynamics in BMC trimer and hexamer pores when taken from a planar shell sheet (Animation S4) (MP4)
>
> Reduced simulation system consisting of three hexameric units and one trimer dimer which can mimic the planar sheet conformation and keeps the system periodic in a conventional orthogonal unit cell (Figure S1); exposed and buried hexamer monomers from our shell fragment simulation (Figure S2); relative distribution of pore radii along the central pore for hexameric and trimeric proteins tiles in either a shell fragment or sheet morphology (Figure S3); top view for open and closed hexamers and stacked trimer observed during the simulation (Figure S4); partially closed conformation predicted by hole program and molecular view of the partially closed conformation of the trimer in sheet conformation (Figure S5); water contacts mapped to each residue of the shell fragment and sheet assemblies (Figure S6); and residue from XFMS analysis mapped onto the hexamer tiles (Figure S7) (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Josh V. Vermaas − *MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, Michigan 48824, United States; Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824, United States;* ⓞ orcid.org/0000-0003-3139-6469; Phone: +1 (517) 884-6937; Email: vermaasj@msu.edu

Authors

Saad Raza − *MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, Michigan 48824, United States;* ⓞ orcid.org/0000-0003-4842-2192

Daipayan Sarkar − *MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, Michigan 48824, United States;* ⓞ orcid.org/0000-0002-4167-2108

Leanne Jade G. Chan − *Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States;* Present Address: Calico Life Sciences LLC, South San Francisco, California 94080, United States

Joshua Mae − *MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, Michigan 48824, United States*

Markus Sutter − *MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, Michigan 48824, United States; Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States*

Christopher J. Petzold − *Biological Systems and Engineering Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States;* ⓘ orcid.org/0000-0002-8270-5228

Cheryl A. Kerfeld − *MSU-DOE Plant Research Laboratory, Michigan State University, East Lansing, Michigan 48824, United States; Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States; Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, Michigan 48824, United States;* ⓘ orcid.org/0000-0002-9977-8482

Corie Y. Ralston − *Molecular Foundry Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States;* ⓘ orcid.org/0000-0002-7899-0951

Sayan Gupta − *Molecular Biophysics and Integrated Bioimaging Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States;* ⓘ orcid.org/0000-0003-4752-068X

Complete contact information is available at:
https://pubs.acs.org/10.1021/acsomega.4c02406

Notes

The authors declare no competing financial interest.

■ REFERENCES

(1) Kerfeld, C. A.; Aussignargues, C.; Zarzycki, J.; Cai, F.; Sutter, M. Bacterial microcompartments. *Nat. Rev. Microbiol.* **2018**, *16*, 277−290.

(2) Sutter, M.; Melnicki, M. R.; Schulz, F.; Woyke, T.; Kerfeld, C. A. A catalog of the diversity and ubiquity of bacterial microcompartments. *Nat. Commun.* **2021**, *12*, No. 3809.

(3) Kerfeld, C. A.; Heinhorst, S.; Cannon, G. C. Bacterial microcompartments. *Annu. Rev. Microbiol.* **2010**, *64*, 391−408.

(4) Turmo, A.; Gonzalez-Esquer, C. R.; Kerfeld, C. A. Carboxysomes: metabolic modules for CO2 fixation. *FEMS Microbiol. Lett.* **2017**, *364*, No. fnx176.

(5) Jakobson, C. M.; Tullman-Ercek, D. Dumpster diving in the gut: bacterial microcompartments as part of a host-associated lifestyle. *PLoS Pathog.* **2016**, *12*, No. e1005558.

(6) Kirst, H.; Ferlez, B. H.; Lindner, S. N.; Cotton, C. A.; Bar-Even, A.; Kerfeld, C. A. Toward a glycyl radical enzyme containing synthetic bacterial microcompartment to produce pyruvate from formate and acetate. *Proc. Natl. Acad. Sci. U.S.A.* **2022**, *119*, No. e2116871119.

(7) Li, T.; Jiang, Q.; Huang, J.; Aitchison, C. M.; Huang, F.; Yang, M.; Dykes, G. F.; He, H.-L.; Wang, Q.; Sprick, R. S.; et al. Reprogramming bacterial protein organelles as a nanoreactor for hydrogen production. *Nat. Commun.* **2020**, *11*, No. 5448.

(8) Kennedy, N. W.; Mills, C. E.; Nichols, T. M.; Abrahamson, C. H.; Tullman-Ercek, D. Bacterial microcompartments: tiny organelles with big potential. *Curr. Opin. Microbiol.* **2021**, *63*, 36−42.

(9) Kirst, H.; Kerfeld, C. A. Bacterial microcompartments: catalysis-enhancing metabolic modules for next generation metabolic and biomedical engineering. *BMC Biol.* **2019**, *17*, 79.

(10) Rose, S. M.; Radhakrishnan, A.; Sinha, S. Innate and engineered attributes of bacterial microcompartments for applications in bio-materials science. *J. Mater. Chem. B* **2023**, *11*, 4842−4854.

(11) Kerfeld, C. A.; Sutter, M. Engineered bacterial microcompartments: apps for programming metabolism. *Curr. Opin. Biotechnol.* **2020**, *65*, 225−232.

(12) Raven, J. A.; Beardall, J. The ins and outs of CO2. *J. Exp. Bot.* **2016**, *67*, 1−13.

(13) Bobik, T. A.; Havemann, G. D.; Busch, R. J.; Williams, D. S.; Aldrich, H. C. The propanediol utilization (pdu) operon of Salmonella enterica serovar Typhimurium LT2 includes genes necessary for formation of polyhedral organelles involved in coenzyme B12-dependent 1, 2-propanediol degradation. *J. Bacteriol.* **1999**, *181*, 5967−5975.

(14) Kofoid, E.; Rappleye, C.; Stojiljkovic, I.; Roth, J. The 17-gene ethanolamine (eut) operon of Salmonella typhimurium encodes five homologues of carboxysome shell proteins. *J. Bacteriol.* **1999**, *181*, 5317−5329.

(15) Petit, E.; LaTouf, W. G.; Coppi, M. V.; Warnick, T. A.; Currie, D.; Romashko, I.; Deshpande, S.; Haas, K.; Alvelo-Maurosa, J. G.; Wardman, C.; et al. Involvement of a bacterial microcompartment in the metabolism of fucose and rhamnose by Clostridium phytofermentans. *PLoS One* **2013**, *8*, No. e54337.

(16) Sutter, M.; Greber, B.; Aussignargues, C.; Kerfeld, C. A. Assembly principles and structure of a 6.5-MDa bacterial microcompartment shell. *Science* **2017**, *356*, 1293−1297.

(17) Kalnins, G.; Cesle, E.-E.; Jansons, J.; Liepins, J.; Filimonenko, A.; Tars, K. Encapsulation mechanisms and structural studies of GRM2 bacterial microcompartment particles. *Nat. Commun.* **2020**, *11*, No. 388.

(18) Sarkar, D.; Maffeo, C.; Sutter, M.; Aksimentiev, A.; Kerfeld, C.; Vermaas, J. Atomic View of Photosynthetic Metabolite Permeability Pathways and Confinement in Cyanobacterial Carboxysomes, *ChemRxiv*, 2024. This content is a preprint and has not been peer-reviewed. DOI: 10.26434/chemrxiv-2024-kbcdf.

(19) Mahinthichaichan, P.; Morris, D. M.; Wang, Y.; Jensen, G. J.; Tajkhorshid, E. Selective permeability of carboxysome shell pores to anionic molecules. *J. Phys. Chem. B* **2018**, *122*, 9110−9118.

(20) Faulkner, M.; Szabó, I.; Weetman, S. L.; Sicard, F.; Huber, R. G.; Bond, P. J.; Rosta, E.; Liu, L.-N. Molecular simulations unravel the molecular principles that mediate selective permeability of carboxysome shell protein. *Sci. Rep.* **2020**, *10*, No. 17501.

(21) Sutter, M.; Laughlin, T. G.; Sloan, N. B.; Serwas, D.; Davies, K. M.; Kerfeld, C. A. Structure of a synthetic β-carboxysome shell. *Plant Physiol.* **2019**, *181*, 1050−1058.

(22) Greber, B. J.; Sutter, M.; Kerfeld, C. A. The plasticity of molecular interactions governs bacterial microcompartment shell assembly. *Structure* **2019**, *27*, 749−763.

(23) Garcia-Alles, L. F.; Lesniewska, E.; Root, K.; Aubry, N.; Pocholle, N.; Mendoza, C. I.; Bourillot, E.; Barylyuk, K.; Pompon, D.; Zenobi, R.; et al. Spontaneous non-canonical assembly of CcmK hexameric components from β-carboxysome shells of cyanobacteria. *PLoS One* **2017**, *12*, No. e0185109.

(24) Cai, F.; Sutter, M.; Cameron, J. C.; Stanley, D. N.; Kinney, J. N.; Kerfeld, C. A. The structure of CcmP, a tandem bacterial micro-compartment domain protein from the β-carboxysome, forms a subcompartment within a microcompartment. *J. Biol. Chem.* **2013**, *288*, 16055−16063.

(25) Noël, C. R.; Cai, F.; Kerfeld, C. A. Purification and characterization of protein nanotubes assembled from a single bacterial microcompartment shell subunit. *Adv. Mater. Interfaces* **2015**, *3*, No. 1500295.

(26) Hagen, A. R.; Plegaria, J. S.; Sloan, N.; Ferlez, B.; Aussignargues, C.; Burton, R.; Kerfeld, C. A. In vitro assembly of diverse bacterial microcompartment shell architectures. *Nano Lett.* **2018**, *18*, 7030−7037.

(27) Trettel, D. S.; Winkler, W. C. In Vitro Analysis of Bacterial Microcompartments and Shell Protein Superstructures by Confocal Microscopy. *Microbiol. Spectrum* **2023**, *11*, No. e03357-22.

(28) Sutter, M.; Faulkner, M.; Aussignargues, C.; Paasch, B. C.; Barrett, S.; Kerfeld, C. A.; Liu, L.-N. Visualization of bacterial microcompartment facet assembly using high-speed atomic force microscopy. *Nano Lett.* **2016**, *16*, 1590−1595.

(29) Iancu, C. V.; Ding, H. J.; Morris, D. M.; Dias, D. P.; Gonzales, A. D.; Martino, A.; Jensen, G. J. The structure of isolated Synechococcus strain WH8102 carboxysomes as revealed by electron cryotomography. *J. Mol. Biol.* **2007**, *372*, 764−773.

(30) Roberts, E. W.; Cai, F.; Kerfeld, C. A.; Cannon, G. C.; Heinhorst, S. Isolation and characterization of the Prochlorococcus carboxysome reveal the presence of the novel shell protein CsoS1D. *J. Bacteriol.* **2012**, *194*, 787−795.

(31) Humphrey, W.; Dalke, A.; Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **1996**, *14*, 33−38.

(32) Trabuco, L. G.; Villa, E.; Mitra, K.; Frank, J.; Schulten, K. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* **2008**, *16*, 673−683.

(33) Sarkar, D.; Lee, H.; Vant, J. W.; Turilli, M.; Vermaas, J. V.; Jha, S.; Singharoy, A. Adaptive Ensemble Refinement of Protein Structures in High Resolution Electron Microscopy Density Maps with Radical Augmented Molecular Dynamics Flexible Fitting. *J. Chem. Inf. Model.* **2023**, *63*, 5834−5846.

(34) Raza, S.; Sarkar, D.; Chan, L. J. G.; Mae, J.; Gupta, S.; Petzold, C. J.; Ralston, C. Y.; Vermaas, J. V. Comparative Pore Structure and Dynamics for Bacterial Microcompartment Shell Protein Assemblies in Sheets or Shells. DOI: 10.5281/zenodo.10800570.

(35) Phillips, J. C.; Hardy, D. J.; Maia, J. D.; Stone, J. E.; Ribeiro, J. V.; Bernardi, R. C.; Buch, R.; Fiorin, G.; Hénin, J.; Jiang, W.; et al. Scalable Molecular Dynamics on CPU and GPU Architectures with NAMD. *J. Chem. Phys.* **2020**, *153*, 044130.

(36) Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; De Groot, B. L.; Grubmüller, H.; MacKerell Jr, A. D. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **2017**, *14*, 71−73.

(37) 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.

(38) Miyamoto, S.; Kollman, P. A. Settle: An Analytical Version of the SHAKE and RATTLE Algorithm for Rigid Water Models. *J. Comput. Chem.* **1992**, *13*, 952−962.

(39) Darden, T.; York, D.; Pedersen, L. Particle Mesh Ewald: An N Log (N) Method for Ewald Sums in Large Systems. *J. Chem. Phys.* **1993**, *98*, 10089−10092.

(40) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.

(41) Shirts, M. R.; Mobley, D. L.; Chodera, J. D.; Pande, V. S. Accurate and Efficient Corrections for Missing Dispersion Interactions in Molecular Simulations. *J. Phys. Chem. B* **2007**, *111*, 13052−13063.

(42) Fletcher, R. *Numerical Analysis*; Springer, 1976; pp 73−89.

(43) Feller, S. E.; Zhang, Y.; Pastor, R. W.; Brooks, B. R. Constant Pressure Molecular Dynamics Simulation: The Langevin Piston Method. *J. Chem. Phys.* **1995**, *103*, 4613−4621.

(44) Lassila, J. K.; Bernstein, S. L.; Kinney, J. N.; Axen, S. D.; Kerfeld, C. A. Assembly of robust bacterial microcompartment shells using building blocks from an organelle of unknown function. *J. Mol. Biol.* **2014**, *426*, 2217−2228.

(45) Gupta, S.; Celestre, R.; Petzold, C. J.; Chance, M. R.; Ralston, C. Development of a microsecond X-ray protein footprinting facility at the Advanced Light Source. *J. Synchrotron Radiat.* **2014**, *21*, 690−699.

(46) Rosi, M.; Russell, B.; Kristensen, L. G.; Farquhar, E. R.; Jain, R.; Abel, D.; Sullivan, M.; Costello, S. M.; Dominguez-Martin, M. A.; Chen, Y.; et al. An automated liquid jet for fluorescence dosimetry and microsecond radiolytic labeling of proteins. *Commun. Biol.* **2022**, *5*, 866.

(47) Harris, C. R.; Millman, K. J.; van der Walt, S. J.; et al. Array programming with NumPy. *Nature* **2020**, *585*, 357−362.

(48) Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90−95.

(49) Smart, O. S.; Neduvelil, J. G.; Wang, X.; Wallace, B.; Sansom, M. S. HOLE: a program for the analysis of the pore dimensions of ion channel structural models. *J. Mol. Graphics* **1996**, *14*, 354−360.

(50) Michaud-Agrawal, N.; Denning, E. J.; Woolf, T. B.; Beckstein, O. MDAnalysis: a toolkit for the analysis of molecular dynamics simulations. *J. Comput. Chem.* **2011**, *32*, 2319−2327.

(51) Gowers, R. J.; Linke, M.; Barnoud, J.; Reddy, T. J. E.; Melo, M. N.; Seyler, S. L.; Domanski, J.; Dotson, D. L.; Buchoux, S.; Kenney, I. M.et al. MDAnalysis: a Python package for the rapid analysis of molecular dynamics simulations. In *Proceedings of the 15th Python in Science Conference*; Benthall, S.; Rostrup, S., Eds.; SciPy: Austin, TX, 2016; pp 98−105. DOI: 10.25080/majora-629e541a-00e.

(52) Mae, J. HoleHelper 2023 https://github.com/joshua-mae/HoleHelper.

(53) Dror, R. O.; Dirks, R. M.; Grossman, J.; Xu, H.; Shaw, D. E. Biomolecular simulation: a computational microscope for molecular biology. *Annu. Rev. Biophys.* **2012**, *41*, 429−452.

(54) Lee, E. H.; Hsin, J.; Sotomayor, M.; Comellas, G.; Schulten, K. Discovery through the computational microscope. *Structure* **2009**, *17*, 1295−1306.

(55) Kefauver, J. M.; Ward, A.; Patapoutian, A. Discoveries in structure and physiology of mechanically activated ion channels. *Nature* **2020**, *587*, 567−576.

(56) Law, R. J.; Capener, C.; Baaden, M.; Bond, P. J.; Campbell, J.; Patargias, G.; Arinaminpathy, Y.; Sansom, M. S. Membrane protein structure quality in molecular dynamics simulation. *J. Mol. Graphics Model.* **2005**, *24*, 157−165.

(57) Mallette, E.; Kimber, M. S. A complete structural inventory of the mycobacterial microcompartment shell proteins constrains models of global architecture and transport. *J. Biol. Chem.* **2017**, *292*, 1197−1210.

(58) McKenzie-Coe, A.; Montes, N. S.; Jones, L. M. Hydroxyl radical protein footprinting: a mass spectrometry-based structural method for studying the higher order structure of proteins. *Chem. Rev.* **2022**, *122*, 7532−7561.

(59) Gupta, S.; Feng, J.; Chan, L. J. G.; Petzold, C. J.; Ralston, C. Y. Synchrotron X-ray footprinting as a method to visualize water in proteins. *J. Synchrotron Radiat.* **2016**, *23*, 1056−1069.

(60) Huang, W.; Peng, Y.; Kiselar, J.; Zhao, X.; Albaqami, A.; Mendez, D.; Chen, Y.; Chakravarthy, S.; Gupta, S.; Ralston, C.; et al. Multidomain architecture of estrogen receptor reveals interfacial cross-talk between its DNA-binding and ligand-binding domains. *Nat. Commun.* **2018**, *9*, No. 3520.

(61) Drew, D.; Boudker, O. Shared molecular mechanisms of membrane transporters. *Annu. Rev. Biochem.* **2016**, *85*, 543−572.

(62) Klein, M. G.; Zwart, P.; Bagby, S. C.; Cai, F.; Chisholm, S. W.; Heinhorst, S.; Cannon, G. C.; Kerfeld, C. A. Identification and

structural analysis of a novel carboxysome shell protein with implications for metabolite transport. *J. Mol. Biol.* **2009**, *392*, 319−333.

(63) Larsson, A. M.; Hasse, D.; Valegård, K.; Andersson, I. Crystal structures of *β*-carboxysome shell protein CcmP: ligand binding correlates with the closed or open central pore. *J. Exp. Bot.* **2017**, *68*, 3857−3867.

(64) Hadidi, H.; Kamali, R.; Binesh, A. Investigation of the aquaporin-2 gating mechanism with molecular dynamics simulations. *Proteins* **2021**, *89*, 819−831.

(65) Wang, Y.; Tajkhorshid, E. Molecular mechanisms of conduction and selectivity in aquaporin water channels. *J. Nutr.* **2007**, *137*, 1509S−1515S.

(66) Ashcroft, F.; Gadsby, D.; Miller, C. Introduction. The blurred boundary between channels and transporters. *Philos. Trans. R. Soc. B* **2009**, *364*, 145−147.