Cracking the Capsid Code: A Computationally-Feasible Approach for Investigating Virus-Excipient Interactions in Biologics Design

Jonathan W. P. Zajac,[†,‡] Idris Tohidian,[¶] Praveen Muralikrishnan,[§,‡] Sarah L. Perry,[‖] Caryn L. Heldt,[¶] and Sapna Sarupria[*,†,‡]

†*Department of Chemistry, University of Minnesota, Minneapolis, MN 55455, USA*

‡*Chemical Theory Center, University of Minnesota, Minneapolis, MN 55455, USA*

¶*Department of Chemical Engineering, Michigan Technological University, Houghton, MI 49931, USA*

§*Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455, USA*

‖*Department of Chemical Engineering, University of Massachusetts Amherst, MA 01003, USA*

E-mail: sarupria@umn.edu

Abstract

The efficacy and equitable distribution of viral biologics—including vaccines and virus-like particles—is hindered due to their inherently low shelf life. To increase the longevity of such products, formulations are typically developed with small molecule additives known as excipients. Incorporating excipients is a costly and time-consuming process due to the large excipient design space and unknown mechanisms underlying excipient-virus interactions. Molecular dynamics simulations are, in theory, well-equipped to efficiently investigate these mechanisms. However, the massive size of fully assembled viral capsids; the protein shell that encapsulates the viral genome; demands computational resources well beyond the requirements of conventional simulations. There exists a need for a novel method that enables high-throughput investigations of virus-excipient interactions at the molecular level. Here, we introduce **CapSACIN** — a computational framework for **Cap**sid **S**urface **A**bstraction and **C**omputationally-**I**nduced **N**anofragmentation. We demonstrate the applicability of this workflow to a model non-enveloped virus, porcine parvovirus (PPV). Through simulations of PPV surface models, we observe that the 2-fold axis of symmetry is significantly weaker at a molecular level than the 3- or 5-fold axes of symmetry. Further, we present results demonstrating excellent agreement with experimentally determined excipient effects on PPV thermal stability.

1 Introduction

Viral-based biologics are designed to trigger an immune response and contain attenuated virulent materials,[1–7] and for drug delivery and gene therapy through the use of virus-like particles (VLPs) devoid of native genetic material.[8–13] The primary ingredient in all such formulations is the virus capsid—a protein shell that protects the infectious viral genome. VLPs are biomimetic materials that retain many inherent properties of virus capsids, including size, biocompatibility, stability, and cell entry behavior.[14,15] Repurposed to encapsulate both natural and synthetic cargo,[8,15] VLPs have been used for intracellular delivery,[16] imaging,[17]

and biocatalytic nanocompartments.[18–20]

Most viral therapeutics have recommended storage temperatures of 2–8 °C and have inherently low shelf lives.[21–23] Potency loss at elevated temperatures is driven by the denaturation of capsid proteins,[24,25] disassembly of the capsid shell,[26,27] or aggregation of viral particles.[28,29] Because of this limitation, the development and distribution of viral biologics is confined within a "cold chain".[30–34] The cold chain preserves viral vectors at refrigerated temperatures, but commonly fails due to improper training and equipment malfunction.[35–37] Breakdowns in the cold chain give rise to loss of product integrity and lead directly to inequitable biologics distribution. Hence, strategies that ease or eliminate the burden of the cold chain are paramount for equitable biologics distribution and reducing waste.

Inspired by the phenomenon of intracellular osmolyte production in extremophiles,[38–43] excipient incorporation into formulations is perhaps the most common strategy for biologics stabilization. During formulation design, excipient selection is high-throughput and empirical[44–47] for two key reasons. First, the excipient design space is extremely large. Most biological formulations include 5 or more unique excipients[48] from a pool of ∼1000 possible excipients.[49] Second, no single excipient can universally stabilize all virus capsids. This gives rise to a trial-and-error process, bottlenecking formulation development.

Excipient-protein interactions (e.g., electrostatic, hydrogen-bonding, or van der Waals interactions) are thought to underlie the stabilization of viruses and VLPs,[50–52] as these interactions are closely linked to changes in protein flexibility, structure, and stability.[45,53–60] Molecular dynamics (MD) simulations offer a powerful means of investigating virus-excipient interactions by capturing their spatial and temporal behavior in atomistic detail. Simulations are also less resource-intensive than wet-lab experiments, reducing reliance on traditional, time-consuming approaches to excipient selection.

Atomistic simulations of individual proteins are well within the reach of modern computational power, though virus simulations are more computationally demanding. Virus capsids are comprised of 60 or more repeating protein subunits, with even the smallest capsids

reaching 10-30 nm in diameter and 200-700 amino acid residues in length.[61] While simulations of isolated capsid proteins are computationally tractable, they lack key features such as curvature,[62,63] asymmetric dynamics,[64,65] protein-protein interactions,[66] and inter-protein binding sites.[67] Hence, an accurate representation of viral capsids requires the full molecular context of the virus. To achieve meaningful atomistic simulations of fully assembled viral capsids, substantial dedicated hardware and computational time is required (Table 1). While methods such as coarse-graining can provide significant computational speed-ups,[68,69] accurate mechanistic insights into virus stability rely on capturing the atomistic details of excipient–water–capsid interactions.

There exists an unmet need in biologics design for high-throughput—yet high-detail—

Table 1: Simulation performance and computational cost associated with recent atomistic virus simulations.

Virus	Size [atoms]	Resources [Cores / GPUs]	Performance [ns/day]	Year	Reference
STMV	1.2M	2,016 / 0 (Cray XE6)	30.0	2012	Larsson *et al.*[70]
PV	2.75M	2048 / 0 (BlueGene/P)	1.27	2012	Roberts *et al.*[71]
AMCV	4M	128 / 0 (CASPUR)	0.3	2014	Arcangeli *et al.*[72]
HIV-1	64M	62,080 / 3,880 (OLCF TITAN)	Not Listed	2017	Perilla and Schulten[73]
EV-1	3.2M	2,048 / 0 (CSC/FGCI/HLRS)	11.0	2017	Pohjolainen *et al.*[74]
IAV	160M	114,688 / 0 (Blue Waters)	4.42	2020	Durrant *et al.*[75]
SARS-CoV-2	305M	180,224 / 24,576 (OLCF Summit)	68.36	2020	Casalino *et al.*[76]
FMDV Subunit	350K	5,360 / 24 (Mole-8.5)	24.9	2021	Li *et al.*[77]
SARS-CoV-2 (CG)	31M	3,920 / 0 (Not Listed)	10.0	2023	Wang *et al.*[69]
SARS-CoV-2 (RA)	1B	180,224 / 24,576, (OLCF Summit)	34.1	2023	Dommer *et al.*[78]
FHV	5M	400 / 20 (IIT New Delhi)	15.5	2024	Shrivastav *et al.*[79]

Abbreviations: STMV – satellite tobacco mosaic virus; PV – poliovirus; AMCV – artichoke mottled crinkle virus; HIV-1 - human immunodificiency virus type 1; EV-1 – echovirus type 1; IAV – influenza virus type A; SARS-CoV-2 – severe acute respiratory syndrome coronavirus 2; FMDV – foot-and-mouth disease virus; FHV – flock house virus; CG – coarse-grained; RA – respiratory aerosol.

investigations of virus-excipient interactions. We address this need by developing a method that simplifies fully assembled capsid structures into models that (i) are computationally feasible, (ii) accurately reproduce capsid structure and dynamics, (iii) retain atomistic resolution, and (iv) can efficiently evaluate excipient stabilizing effects. To this end, we introduce a novel framework: **CapSACIN** — a computational framework for **Cap**sid **S**urface **A**bstraction and **C**omputationally-**I**nduced **N**anofragmentation. This method enables high-throughput identification of virus–excipient interactions at atomistic detail. CapSACIN leverages the symmetry of icosahedral capsids to construct surface models of viruses focused on a region of interest (ROI) embedded in peripheral proteins representing the molecular context of an assembled capsid. We further utilize a computational nanofragmentation approach, inspired by Ghaemi *et al.*,[80] to assess the interfacial stability of protein-protein interfaces spanning the capsid surface. Using the non-enveloped porcine parvovirus (PPV) as a model system, we demonstrate that CapSACIN can accurately predict excipients as stabilizing or destabilizing, as validated by wet-lab experiments. In addition to predictive power, these simulations reveal mechanistic insights into excipient effects on weak interprotein interfaces. By bridging atomistic resolution with high-throughput capability, CapSACIN offers a powerful, information-driven strategy for accelerating formulation design and enhancing the development of stable, globally accessible viral biologics.

2 Methods

2.1 CapSACIN Workflow

CapSACIN is an end-to-end workflow capable of preparing input for and running capsid-excipient simulations (Fig. 1), composed of six steps: *Step 1: Symmetry-Based Alignment* (Fig. 1a), which involves aligning the ROI perpendicular to the z-axis and embedding it within the appropriate context; *Step 2: Surface Abstraction* (Fig. 1b), for abstracting the ROI and contextual proteins in preparation of simulations; *Step 3: Position Restraint Generation*

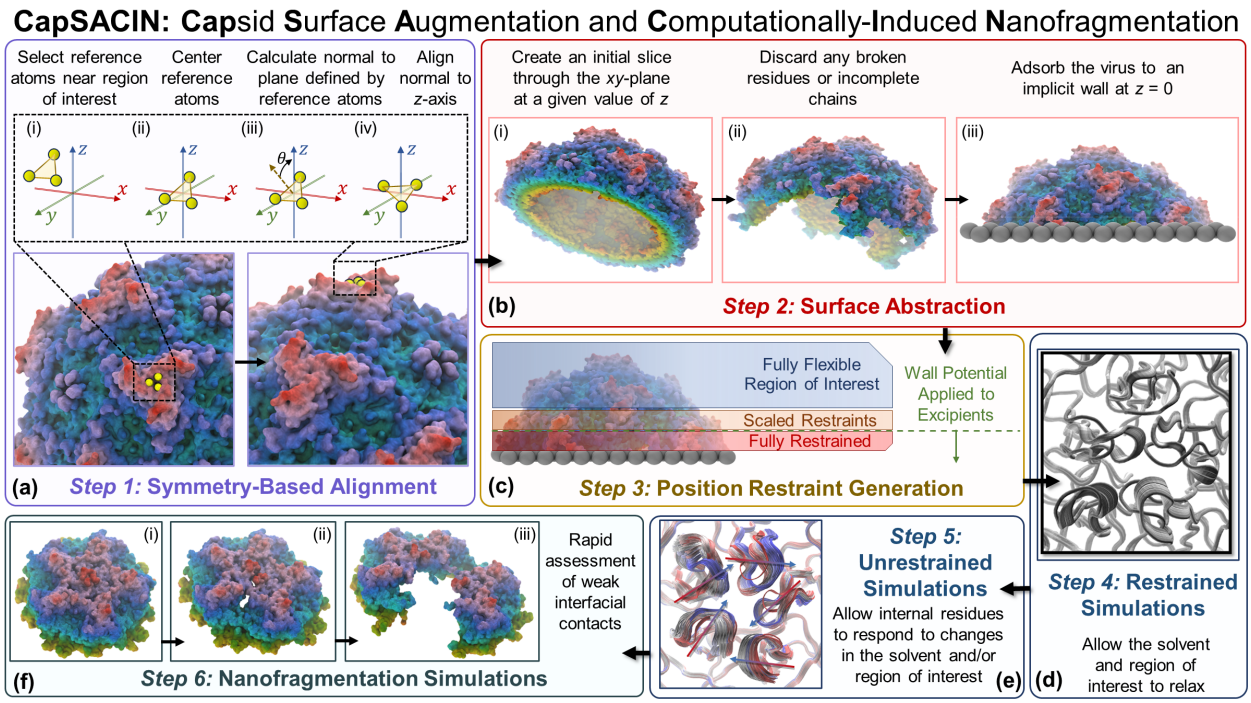

Figure 1: CapSACIN workflow. (a) *Step 1:* Symmetry-Based Alignment. (b) *Step 2:* Surface Abstraction. (c) *Step 3:* Position Restraint Generation. (d) *Step 4:* Restrained Simulations. (e) *Step 5:* Unrestrained Simulations. (f) *Step 6:* Nanofragmentation Simulations. All renders of molecular systems were completed in VMD.[81]

(Fig. 1c), where an implicit wall is placed and position restraints are applied based on distance from the wall (to preserve capsid structural integrity and prevent excipient diffusion into the capsid interior); *Step 4: Restrained Simulations* (Fig. 1d), where applied restraints allow the solvent to equilibrate while preserving capsid structure; *Step 5: Unrestrained Simulations* (Fig. 1e), which allows the capsid surface to equilibrate in response to the solvent; and *Step 6: Nanofragmentation Simulations* (Fig. 1f), which enables prediction of stabilizing/destabilizing excipient effects on capsid protein interfaces.

The CapSACIN workflow depends only on an input PDB file, the selection of at least one reference atom within the ROI, and the target axis of symmetry. Details of each step in the CapSACIN workflow are provided below. Validations of a proof-of-concept surface model, including equilibration, structural integrity, and dynamics, were compared against the fully assembled PPV capsid.

2.1.1 *Step 1:* Symmetry-Based Alignment

The first stage of the structure construction module involves a symmetry-based alignment of an input capsid structure (Fig. 1b). In this stage, capsid Cartesian coordinates, $\mathbf{r}_{\text{Capsid}}$, are transformed by a translation vector, $\mathbf{T}$, such that the center-of-geometry of reference atoms $\mathbf{r}_{\text{COG},Ref}$ is at the origin:

$$\mathbf{r}_{\text{COG,Ref}} = \frac{1}{3}(\mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3) \tag{1}$$

$$\mathbf{T}_{\mathbf{r}} = \mathbf{r}_{\text{Capsid}} - \mathbf{r}_{\text{COG,Ref}} \tag{2}$$

Following translation, the normal, $\mathbf{n_{ref}}$, to the plane connecting the three reference atoms is computed as:

$$\mathbf{n_{ref}} = \frac{(\mathbf{r}_2 - \mathbf{r}_1) \times (\mathbf{r}_3 - \mathbf{r}_1)}{|(\mathbf{r}_2 - \mathbf{r}_1) \times (\mathbf{r}_3 - \mathbf{r}_1)|} \tag{3}$$

which is used to define a rotation matrix, $\mathbf{R}$, applied to capsid coordinates such that the normal is aligned with the z-axis, $\mathbf{e}_z$:

$$\theta = \cos^{-1}(\mathbf{n_{ref}} \cdot \mathbf{e}_z) \tag{4}$$

$$R = I + (\sin\theta)[\mathbf{K}] + (1 - \cos\theta)[\mathbf{K}]^2 \tag{5}$$

where I is the 3×3 identity matrix and $\mathbf{K}$ is the skew-symmetric matrix of $\mathbf{n_{ref}}$. This capsid surface construction protocol can, in theory, be extended to any icosahedral (or similarly structured) virus with an experimentally-resolved structure. In demonstrating our approach, we focus on the PPV capsid (PDB: 1K3V),[82] though we have tested this stage of the CapSACIN workflow on a diverse set of capsid architectures (Fig. S1).

2.1.2 *Step 2:* Surface Abstraction

Following alignment, the capsid structure is converted into a surface model (Fig. 1c). In this stage, (i) the capsid is shifted such that its center-of-geometry is at the origin:

$$\mathbf{T_c} = \mathbf{r}_{\text{Capsid}} - \mathbf{r}_{\text{COG,Capsid}} \tag{6}$$

A crude initial slice is made through the xy-plane at a height defined by a weight parameter $\omega \in [0, 1]$, where $\omega = 0$ retains the full capsid and $\omega = 1$ removes all capsid atoms. Atoms with $r_z < \omega m_z$ are deleted, where r_z is the z-coordinate of an atom and m_z is the maximum z-coordinate of the capsid. This operation leaves dangling bonds, broken residues, and incomplete chains, so a clean-up step follows. Residues are removed if their number of atoms is less than the expected count for a given residue type. Similarly, CapSACIN removes chains if they contain fewer residues than expected from the input structure. The remaining capsid is translated upward such that its minimum z-coordinate is 0.

2.1.3 *Step 3:* Position Restraint Generation

The global topology of the capsid surface must not distort while the solvent equilibrates. To this end, we place the capsid surface atop an implicit wall and apply scaled position restraints (Fig. 1c). The wall is constructed using the GROMACS *Walls* function and is placed at $z = 0$ and $z = h_{box}$, where h_{box} is the height of the simulation box in the z dimension. We use the CHARMM36 C_δ parameters as the wall atomtype, with a number density of 5 atoms/nm^2, to prevent diffusion through the wall. Periodic boundaries are applied in the x and y dimensions such that the wall is infinitely repeating perpendicular to the capsid surface normal. We use a stiff 10-4 Lennard-Jones potential integrated over the wall surface. During energy minimization, we apply a linear potential forcing atoms to a separation of 0.1 nm from the wall. This effectively places a constraint that the positions of all atoms in the system must fall between $r_z > 0$ and $r_z < h_{box}$.

Capsid atoms near the wall ($r_z < 1.5$ nm) have their positions restrained according to a harmonic potential ($K = 1000$ kJ mol^{-1} nm^{-2}), while capsid atoms far from the wall ($r_z > 3.0$ nm) are fully flexible ($K = 0$ kJ mol^{-1} nm^{-2}). At intermediate distances, force constants are

weighted from 0 to 1000 kJ mol^{-1} nm^{-2} according to an inverse sigmoid function:

$$s_i = \frac{-1}{1 + \exp\left[-r_{z,i}/\kappa\right]} + 1 \tag{7}$$

where $r_{z,i}$ is the z-coordinate of capsid atom i and κ controls the rate of change of the inverse sigmoid. Force constants are then applied to harmonic restraints of atoms in the intermediate region as $K_i = s_i K$ (Fig. S2a). This allows the ROI to remain fully flexible whereas peripheral capsid regions provide the correct molecular context.

sapna starts here

In addition to providing a stable platform for the capsid surface to adsorb to, the implicit wall provides another crucial function. While viral capsids are typically permeable to water and salt ions,[83,84] larger molecules such as excipients are restricted to the capsid exterior. The wall at $z = h_{box}$ prevents excipients from diffusing across periodic boundaries into the capsid interior. Additionally, we apply flat-bottom position restraints that act only when excipients are ≤ 1.5 nm from the wall at $z = 0$ nm (Fig. S2a). This restraint prevents excipients from interacting with the fully-restrained portion of the capsid surface (Fig. 1d), and takes the form:

$$V_{fb}(\mathbf{r}_i) = \frac{1}{2} K [d_g(\mathbf{r}_i; \mathbf{R}_i) - r]^2 H[d_g(\mathbf{r}_i; \mathbf{R}_i) - r] \tag{8}$$

where $\mathbf{R}_i$ is a reference position for the excipient (Fig. S2), r is the distance from the center of the potential, $\mathbf{r}_i$ is the position of an excipient heavy atom i, $d_g(\mathbf{r}_i; \mathbf{R}_i)$ is the distance from the reference position, and H is the Heaviside step function.

2.1.4 *Steps 4 and 5:* Restrained and Unrestrained Molecular Dynamics Simulations

Excipient positions are initialized as a 1 nm thick slab at the top of the simulation box (Fig. S2a). This slab is used as the reference position for the flat-bottom restraint described in Section 2.1.3. The relative accumulation/depletion of an excipient near the virus surface

is described by the preferential interaction coefficient, Γ_{XV}. This property is calculated in simulations using the two-domain formula[85–87] given by:

$$\Gamma_{XV} = \left\langle N_X^{\text{local}} - \left(\frac{N_X^{\text{bulk}}}{N_W^{\text{bulk}}}\right) N_W^{\text{local}} \right\rangle \tag{9}$$

where V denotes the viral surface, X represents an excipient species, and W indicates water. N represents the number of molecules of a given species in the local or bulk domain of the viral surface. Equilibrium simulations of the capsid surface model with restraints applied are run for 50 ns to monitor Γ_{23} convergence. We observe Γ_{23} convergence within this simulation time (Fig. S7), consistent with previous studies.[87,88]

All simulations were performed using GROMACS,[89,90] with versions 2020.4 (for production runs on SDSC Expanse) or 2023.2 (for equilibration runs and nanofragmentation simulations on MSI). Systems were initially subject to energy minimization using the steepest descent algorithm.[91] An initial NVT equilibration was carried out for 1 ns at 300 K using the V-rescale thermostat,[92] followed by a 50 ns NVT equilibration using the Nosé-Hoover thermostat.[93] During equilibration, capsid positions are restrained as described in Section 2.1.3. Following equilibration, production runs were completed for 50-400 ns with capsid harmonic restraints removed (Fig. 1e) but excipient flat-bottom restraints retained. For the 2- and 3-fold surfaces, 100 ns production runs were carried out; for the 5-fold surface without excipients, 400 ns; and for 5-fold surfaces with excipients, 50 ns.

The Particle Mesh Ewald (PME) algorithm[94] was used for electrostatic interactions with a cut-off of 1.2 nm. A reciprocal grid of 42 x 42 x 42 cells was used with 4th order B-spline interpolation. We applied a correction to the force and potential in the z-dimension to generate a pseudo-2D summation.[95] A force-switching scheme from 1–1.2 nm was used to describe long-range van der Waals interactions. The neighbor search was performed every 10 steps. Lorentz-Berthelot mixing rules[96,97] were used to calculate non-bonded interactions between different atom types.

The CHARMM36 force field[98,99] was used to parameterize virus atoms, arginine (ARG), glutamate (GLU), glycine (GLY), Na^+, and Cl^-, while a modified Kirkwood-Buff based parameter set was used for trehalose (TRE) and sorbitol (SOR).[100] Compared to other traditional force fields, CHARMM36 has been shown to generate secondary structures consistent with capsid protein structure data while also sampling a large conformational space.[101] All systems were solvated with standard TIP3P water[102] and an NaCl concentration of 150 mM. Additional Na^+ and Cl^- atoms were added to neutralize the system, as necessary. The protonation states of titratable virus residues at pH 7 were determined by PROPKA3[103] applied to a PPV monomer (PDB: 1K3V).[82] The same workflow was used for preparing the fully assembled PPV capsid, albeit without restraints. In totality, for capsid protein monomers, we obtained 5.4 μs of restrained and 12.6 μs of unrestrained PPV capsid protein dynamics from surface systems. 25.5 μs of capsid protein dynamics from the fully assembled capsid were obtained to serve as our "ground truth".

Selecting the height of the simulation box, h_{box}, is a non-trivial task. Too small an h_{box} and the wall placed at $z = h_{box}$ will interact with the hydration shell of the capsid. Too large an h_{box} and significant computational cost is incurred. Our criteria for a sufficiently large h_{box} is that (i) interactions between periodic capsid images should be minimal, and (ii) the distance between the capsid surface and wall should be large enough such that bulk properties of water can be recovered. We address (i) by selecting an h_{box} value at least 1.5 nm greater than the height of the capsid surface, ensuring periodic capsid surfaces are at a distance greater than the CHARMM nonbonded cut-off distance of 1.0-1.2 nm. For (ii), we ultimately found $h_{box} = 12$ nm is the minimum height sufficient to recover several bulk properties of TIP3P water (Figs. S3, S4).

2.1.5 *Step 6:* Nanofragmentation Simulations

Ghaemi *et al.*[80] demonstrated that inducing macroscopic cracks across hepatitis B virus (HBV) is an effective strategy for assessing capsid stability. Inspired by this approach,

we investigate PPV capsid stability through induced capsid nanofragmentation using the GROMACS pulling code.[89,90] We apply linear pulling forces along the vectors defined by the center-of-mass of the surface, $X_{COM,surf}$, and the center-of-mass of each capsid protein, $X_{COM,cp}$. We use umbrella pulling with a pulling rate of 0.001 nm ps^{-1} and a force constant of 1000 kJ mol^{-1} nm^{-2} (see SI for more details). All pulling simulations are run for either 3.5 ns or until $\max[|X_{COM,surf} - X_{COM,cp}|] > \frac{1}{2}b_{xy}$, where b_{xy} is the minimum length of the box in the x- and y-dimensions.

We initialize pulling simulations using configurations from equilibrium MD simulations. Specifically, we extract a frame every 5 ns from the final 25 ns of each production run, resulting in 5 independent configurations for pulling. All data arising from pulling simulations are thus shown as the mean and standard error from 5 such simulations. We observed the selection of starting configurations has significant impacts on the outcome of excipient-mediated stability, with configurations taken from equilibrium and with capsid position restraints removed providing the best agreement with experiment (Fig. S14).

2.2 Capsid Dynamics, Structure, and Stability

2.2.1 Dynamic Cross-Correlation

Dynamic cross-correlation (d_{corr}) analysis is useful in quantifying the correlation coefficients of motions between atoms.[104] We compute intraprotein correlations between the C_α atoms of residues i and j, following:

$$d_{corr} = \frac{\langle \Delta\mathbf{r}_i(t) \cdot \Delta\mathbf{r}_j(t) \rangle_t}{\left[\langle ||\Delta\mathbf{r}_i(t)||^2 \rangle_t \right]^{\frac{1}{2}} \left[\langle ||\Delta\mathbf{r}_j(t)||^2 \rangle_t \right]^{\frac{1}{2}}} \tag{10}$$

where $\Delta\mathbf{r}_i(t) = \mathbf{r}_i(t) - \langle \mathbf{r}_i(t) \rangle_t$ and $\langle \cdot \rangle$ denotes the time-averaged value. To assess protein-protein correlations within virus capsids, we calculate d_{corr} using the center-of-mass of capsid proteins.

2.2.2 Residue Interaction Networks

We compute residue interaction networks (RINs) to quantify intraprotein structure.[105–108] To construct these graphs, we consider a pair of residues i, j to be interacting if the following criteria are met: (i) at least one pair of heavy atoms is within 0.4 nm, and (ii) i and j are at least 3 residues away in the amino acid sequence. All residues are then assigned to nodes and edges are created between interacting nodes.

To examine the network of intramolecular interactions within capsid protein monomers, we measure betweenness centrality,[109–111] C_b, of nodes within each RIN. Betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes. This quantity is often thought of as the propensity for a node to act as a "hub" of information transfer. Betweenness centrality is computed as:[109–111]

$$C_b(u) = \sum_{s,t} \frac{np(s,t|u)}{np(s,t)} \tag{11}$$

where $np(s,t)$ is the number of shortest paths between nodes s and t through the connected network, and $np(s,t|u)$ is the number of such paths that pass through node u. We compute this metric for residues in capsid protein monomers, and report $C_b(u)$ values averaged across monomers as:

$$\langle C_b(u) \rangle = \frac{1}{M} \sum_{i}^{M} [C_b(u)]_i \tag{12}$$

where u corresponds to an identical residue in all capsid proteins, M.

2.2.3 Sphericity

Capsid sphericity, Ψ, is obtained according to Wadell's definition:[112]

$$\Psi = \frac{\pi^{1/3}(6V_p)^{2/3}}{A_p} \tag{13}$$

where V_p is the volume of the capsid and A_p is the surface area. $\Psi = 1$ for spheres, while all non-spheres have $\Psi < 1$. We estimate V_p and A_p from the convex hull defined by the point cloud of capsid atomic coordinates, constructed using Trimesh.[113]

2.2.4 Residue-Level Capsid Fragmentation

We define the fraction of native interfacial contacts, Q_{IF}, to monitor fragmentation progression during pulling simulations:

$$Q_{IF}(X) = \frac{1}{N}\sum_{i,j}\delta[r_{ij}(X) - r_{ij}^0] \tag{14}$$

where δ is the Dirac delta function, the sum is over all N pairs of native contacts (i, j), $r_{ij}(X)$ is the distance between i and j in configuration X, and r_{ij}^0 is the distance between i and j in a reference state. We define any pair of heavy atoms ij as a native interfacial contact if the following criteria are met: (i) i and j do not belong to the same capsid protein, and (ii) r_{ij}^0 < 0.45 nm in the reference configuration. We use the first frame of the pulling simulation as the reference configuration, though this choice is arbitrary. The change in Q_{IF} over the course of the pulling simulation is reported as:

$$\Delta Q_{IF}(t) = Q_{IF}(t) - Q_{IF}(0) \tag{15}$$

To measure the difference in $\Delta Q_{IF}(t)$ curves in excipient and reference solutions, we follow the following protocol: (i) First, $\Delta Q(t)_{IF}$ curves are normalized such that $\Delta Q(\tau)_{IF} = 1$. We select τ as 10 ps, which is enough time for transient interfacial contacts to break prior to

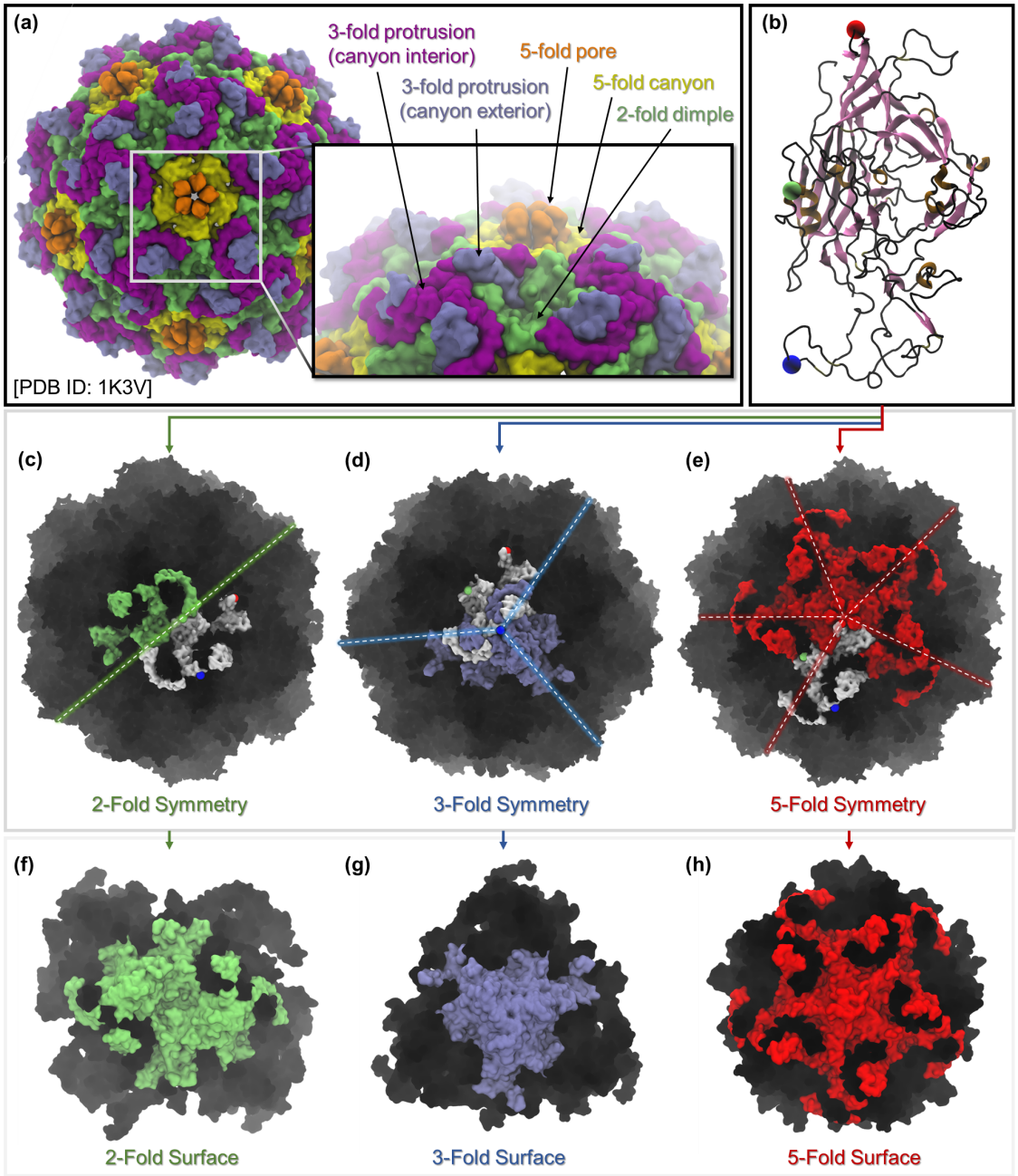

Figure 2: Construction of model PPV surfaces. (a) PPV capsid structure, including key structural ROIs: 5-fold pore (orange), 5-fold canyon (yellow), 3-fold protrusions (blue, purple), and 2-fold dimple (green). (b) PPV monomeric capsid protein, with reference atoms for surfaces shown as spheres (2-fold, green; 3-fold, blue; 5-fold, red). (c) PPV 2-fold axis of symmetry. (d) PPV 3-fold axis of symmetry. (e) PPV 5-fold axis of symmetry. In (c-e), the reference capsid protein from the deposited PDB entry is colored in white, symmetrical proteins are green, blue, or red, and the rest of the capsid is black. Surface models are shown for (f) 2-fold, (g) 3-fold, and (h) 5-fold axes of symmetries. In (f-h), the 2-, 3- and 5-fold ROIs are colored in green, blue, and red, respectively, while contextual proteins are black.

$\Delta Q(t)$ leveling off. (ii) For each value of t, we compute $\Delta Q_{IF}(t)^{Exc} - \Delta Q_{IF}(t)^{Wat}$. (iii) The final value reported is the sum of differences:

$$\Delta\Delta Q_{IF} = \sum_{t=\tau}^{T} \left(\Delta Q_{IF}(t)^{Exc} - \Delta Q_{IF}(t)^{Wat}\right) \tag{16}$$

Mean ΔQ_{IF} values of an individual residue i are computed by considering all observed interfacial contacts ij involving i. Errors are reported as standard errors from the mean of five replicate simulations.

2.2.5 Temperature Stability Assays

Wet-lab experiments were performed to measure the thermal stability of PPV. Liquid samples of PPV in excipient solutions were prepared in triplicate and were put in either a heat block at 60 °C[114] or in a fridge at 4 °C for 72 hours. The titer of PPV was determined by 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) colorimetric cell viability assay.[115] Additional experimental details, including materials and methods, are provided in the Supporting Information.

3 Results and Discussion

3.1 Construction and Performance of PPV Surface Models

PPV is a non-enveloped virus with a single stranded DNA genome.[116] The PPV capsid is an icosahedral shell comprised of 60 copies of viral proteins (VP) VP1, VP2, and VP3 in a 1:10:1 ratio.[82] Due to the relatively small size and structural simplicity of PPV, recent studies have employed the virus as a model for investigating virus purification and thermostabilization techniques.[114,115,117,118] PPV, and other similar parvoviruses, are often used in virus removal studies for biotherapeutic manufacturing.[119–121]

The major capsid protein of PPV, VP2, is known to self-assemble into virus-like particles

Table 2: Simulation performance and computational cost associated with fully assembled and surface models of PPV. Error estimates of simulation performance, where reported, are the standard deviation from three replicate simulations.

System	System Size [atoms]	CPU Cores	Performance [ns/day]
Assembled PPV Capsid	3,146,527	128	<1
		256	3.881 ± 0.1
		512	7.484 ± 0.07
		1024	12.279 ± 0.29
		2048	13.886
PPV 5-Fold Surface	745,274	128	11.031 ± 0.01
		256	17.450 ± 1.56
		512	29.537 ± 0.92
		1024	54.188 ± 0.96
		2048	86.031 ± 11.78
PPV 3-Fold Surface	635,743	128	11.370 ± 0.56
		256	17.417 ± 2.88
		512	34.719 ± 1.56
		1024	54.677 ± 9.36
		2048	72.638 ± 8.42
PPV 2-Fold Surface	692,773	128	11.033 ± 0.42
		256	18.766 ± 0.60
		512	30.275 ± 4.73
		1024	50.925 ± 1.36
		2048	57.330 ± 13.34

(VLPs) that are non-pathogenic, but are morphologically similar to native PPV virions.[122] On the PPV surface, prominent structural regions of interest (ROIs) include a canyon surrounding a pore at the 5-fold axis of symmetry, protrusions located at the 3-fold axis of symmetry, and a dimple at the 2-fold axis of symmetry (Fig. 2a).[82] To capture the PPV ROIs, we utilize the CapSACIN workflow to generate PPV surface models centered on the 2-, 3-, and 5-fold axes of symmetry (Fig. 2).

Each capsid surface model includes a complete ROI (5-fold pore, 3-fold protrusion, or 2-fold dimple), as well as a variable number of contextual proteins at the periphery of each surface. The number of contextual proteins to include depends on the complexity and number of proteins within the ROI, and at the very least, should be sufficiently high such that no ROI atoms are included in the positional restraining scheme. With a slicing weight of $w = 0.7$, each of our PPV surface systems had a 2:1 ratio of contextual to ROI proteins. The resulting surface systems are between 20–23% the size of a fully assembled capsid simulation (Table S1).

Since the PPV 5-fold surface is the largest ROI (Fig. 2c-h) we studied, we focus the majority of the discussion on this surface. Similar analyses were performed for the other surfaces and are reported in the SI.

The 5-fold surface model consists of nearly 3 million fewer atoms relative to the fully-assembled capsid. Atomistic MD simulations for the 5-fold surface are between 5–18x faster than the assembled capsid depending on the available resources and computing hardware (Table 2). We also observe the 5-fold surface scales with available CPU cores roughly 5x more efficiently than the assembled capsid (Fig. S8). When incrementing by 128 cores, the 5-fold surface improves by ∼4.638 ns/day, while the assembled capsid is limited to ∼0.811 ns/day improvements. The speed-up we observe is comparable to popular coarse-graining strategies,[123,124] yet the virus surface models presented in this study can be carried out at atomistic resolution. Below, we demonstrate that these PPV surface models accurately reproduce, locally and globally, structural and dynamic features of the fully assembled capsid.

3.2 Intraprotein Structure and Dynamics: Local Similarity

For PPV surface models to accurately reflect the structure and dynamics of a fully assembled capsid, capsid proteins must behave identically at the monomer level between the two representations. In the fully assembled PPV capsid, individual capsid protein monomers have $RMSD_{C_\alpha}$ fluctuations between 0.125–0.24 nm for the majority of the production run

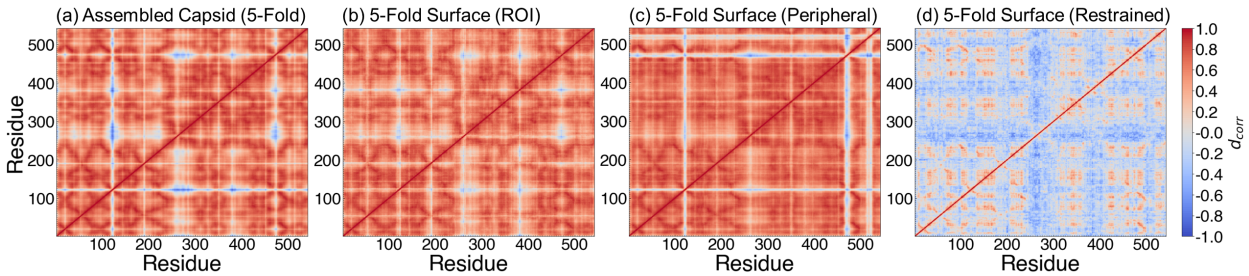

Figure 3: Dynamic cross-correlation matrices of residue-residue pairs within PPV 5-fold capsid proteins. (a) 5-fold ROI subset from the assembled capsid, (b) 5-fold surface ROI, (c) 5-fold surface peripheral proteins, and (d) 5-fold surface ROI from the restrained simulation. d_{corr} values are colored from blue to white to red, with blue indicating strong negative correlation, red indicating strong positive correlation, and white indicating weak or no correlation.

(Fig. S6a). In the PPV 5-fold surface model, ROI monomers have $RMSD_{C_\alpha}$ fluctuations within a similar range (0.1-0.215 nm), while peripheral monomers are subject to larger conformational fluctuations (Fig. S6b). To further assess intraprotein dynamics of PPV capsid proteins, we compute dynamic cross-correlations between all residue-residue pairs at the 5-fold axis of symmetry (Fig. 3), as well as for 3- and 2-fold surface models (Fig. S9).

Intraprotein correlations from 5-fold ROI proteins in the assembled capsid (Fig. 3a) and 5-fold surface (Fig. 3b) are in excellent agreement. In these proteins, residues 115-125 and 470-475 are negatively correlated, while residues 0-100 show a high degree of positive correlation. Similar results are obtained for the 3- and 2-fold surfaces (Fig. S9), demonstrating that internal dynamics of capsid proteins are preserved even in distinct ROIs of capsid surface models. Greater disagreement with assembled capsid proteins is observed for peripheral proteins from the 5-fold surface (Fig. 3c), where a marked decrease in correlated residue-residue fluctuations is observed from residues ∼510-530 (Fig. 3c). The missing band in Fig. 3c for 5-fold ROI proteins shows that including surrounding proteins is essential to capture the full intraprotein dynamics of the assembled capsid. Residue-residue correlations are almost completely lost during the restrained portion of the capsid surface simulations (Fig. 3d), highlighting that removal of positional restraints is a necessary step to restore intraprotein dynamics.

In addition to characterizing intraprotein dynamics, we assess intraprotein structure by constructing residue interaction networks (RINs) for the PPV systems under study (Fig. 4). RINs provide graph-based representations of pairwise residue-residue interaction networks within proteins and have been useful in providing insights into factors that drive protein structure, function, and stability.[105–108] Within each RIN, we compute betweenness centrality, C_b, as a measure for how often a given residue acts as a bridge connecting all other residues. Several key residues with high C_b values are identified (Fig. 4a-c): Tyr185, Pro358, Asn370, Arg404, Thr456, and Gln493. These residues consistently have relatively large C_b, regardless of whether they are extracted from the assembled capsid, 5-fold surface ROI, or peripheral

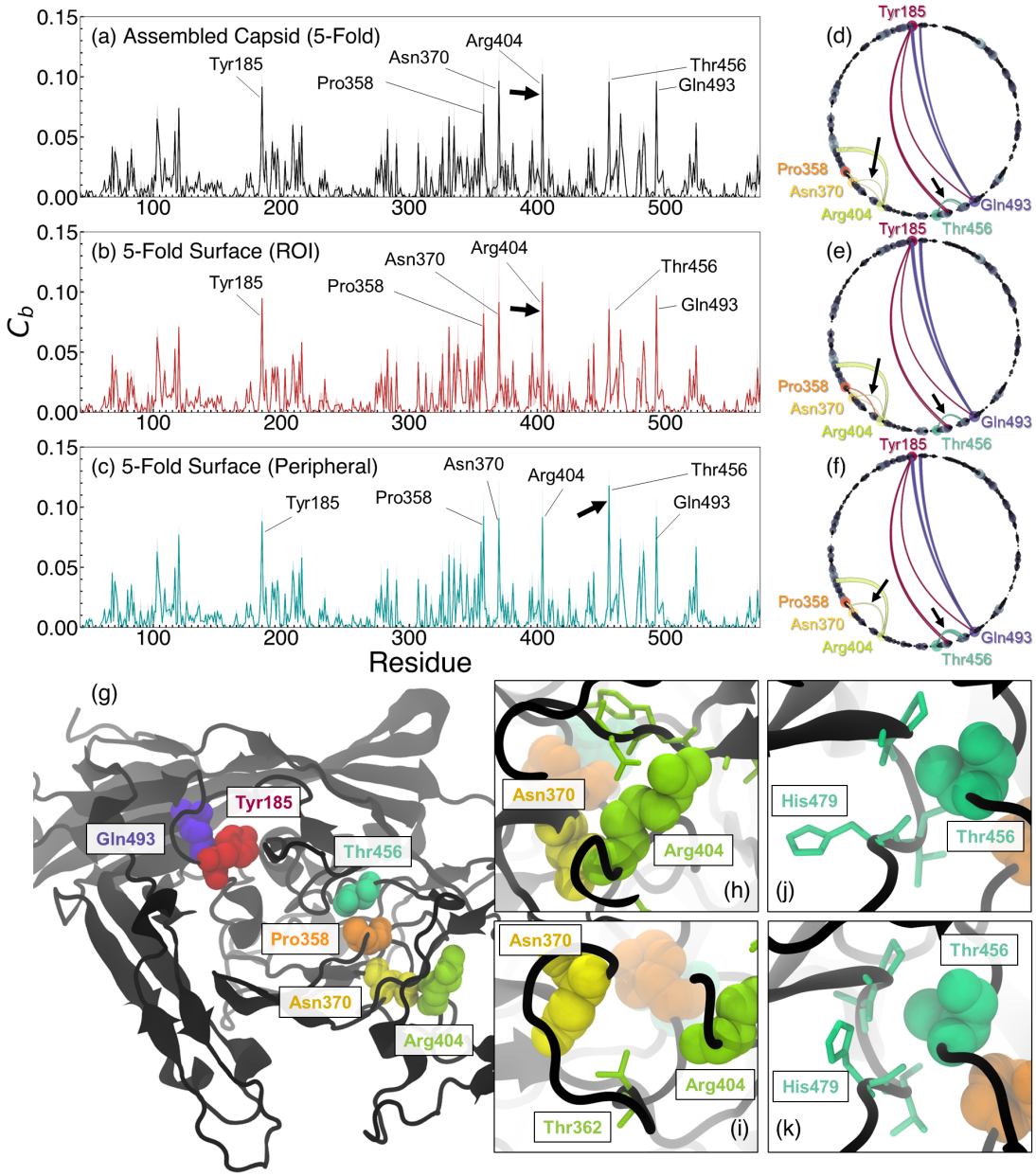

Figure 4: Residue interaction network analysis of PPV capsid proteins. (a-c) Betweenness centrality (C_b) values for capsid protein residues. Arrows are drawn to indicate residues with the highest average C_b value. (d-f) Residue-residue interactions with contact frequency >80% during production runs. Edge widths are proportional to the total amount of time two residues are in contact ($r_{ij} < 0.4$ nm). Node sizes are proportional to residue betweenness centrality. Highest betweenness residues are colored as Tyr185, red; Pro358, orange; Asn370, gold; Arg404, yellow-green; Thr456, turquoise; Gln493, violet. Capsid proteins extracted from (a,d) assembled capsid, (b,e) 5-fold surface ROI, and (c,f) peripheral 5-fold surface. (g) Representative configuration showing the relative locations of Tyr185, Pro358, Asn370, Arg404, Thr456, and Gln493. (h) Asn370–Arg404 contact in the assembled capsid and 5-fold surface ROI proteins. (i) Broken Asn370–Arg404 contact in peripheral 5-fold surface proteins. (j) Broken Thr456–His479 contact in the assembled capsid and 5-fold surface ROI proteins. (k) Thr456–His479 contact in peripheral 5-fold surface proteins.

5-fold surface. These same key residues are identified 3-fold and 2-fold proteins (Fig. S11), indicating critical pathways within the residue interaction network do not vary with protein location in the assembled capsid.

In the assembled capsid and 5-fold ROI, Arg404 has the highest average betweenness centrality in the residue interaction network,(Fig. 4a,b) while for 5-fold peripheral proteins, Thr456 becomes the most central residue (Fig. 4c). Arg404 betweenness is driven by persistent contacts between Arg404 and the stretch of residues Val335, Gly336, Tyr337, and Asn338 (Fig. 4d-f). In the assembled capsid and 5-fold surface ROI, Arg404 additionally forms persistent interactions with Asn370 (Fig. 4d,e,h), while this contact is disrupted in peripheral 5-fold proteins (Fig. 4f,i). Thr456 betweenness is driven by interactions with Val480, Ala482, and Pro483 (Fig. 4d-f). In peripheral 5-fold proteins, these interactions are more persistent (Fig. 4f) relative to assembled capsid and 5-fold ROI proteins (Fig. 4d,e). Thr456 interactions with this stretch of residues is strengthened in peripheral 5-fold proteins by a rotation of the His479 side chain towards Thr456 (Fig. 4k), while His479 is primarily rotated away from Thr456 in assembled capsid and 5-fold ROI proteins (Fig. 4j). Thus, the residue interaction networks of the 5-fold surface ROI and assembled capsid converge on Arg404 as the central hub, whereas peripheral 5-fold proteins deviate by shifting centrality to Thr456, underscoring a distinct change in the local interaction landscape. Together, these shifts in centrality highlight how removing the proper molecular context of capsid proteins can give rise to subtle conformational rearrangements, rewire residue interaction networks, and reshaping local communication pathways.

3.3 Interprotein Structure and Dynamics: Global Similarity

Dynamic cross-correlations reveal how capsid proteins move in concert, uncovering collective motions that shape the global architecture of the assembled virus. To quantify these motions, we compute the dynamic cross-correlation of monomer-monomer capsid protein pairs (Fig. 5). The 5-fold ROI (denoted by black boxes in Fig. 5a-c) extracted from the assembled capsid

(Fig. 5a) shows strong similarity with the 5-fold surface ROI (Fig. 5b). Within the ROI, proteins exhibit highly correlated fluctuations, consistent with previous studies demonstrating concerted 5-fold motions—such as twists, expansions, and contractions—of the poliovirus capsid.[125]

By contrast, contextual proteins (outside of the black boxes in Fig. 5) are in poor agreement between the assembled capsid (Fig. 5a) and the 5-fold surface model (Fig. 5b). In the surface model, 5-fold contextual proteins show larger positive correlations than the same proteins in the assembled capsid. The restrained 5-fold surface (Fig. 5c) captures similar intrapentamer correlations among ROI proteins, though the strength of these correlations are dampened

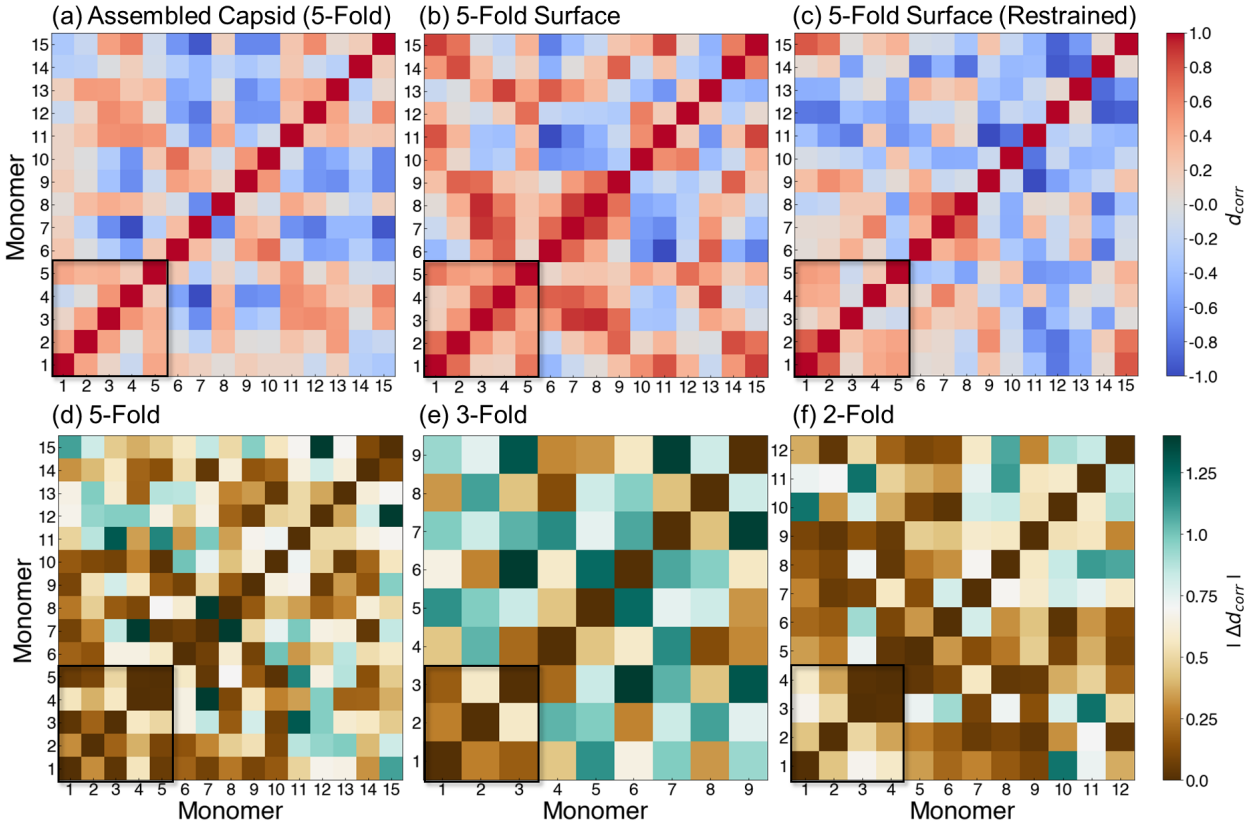

Figure 5: Dynamic cross-correlation matrices of monomer-monomer pairs within PPV capsids. 5-fold capsid proteins taken from (a) the assembled capsid, (b) the 5-fold surface model production run, and (c) the restrained portion of the 5-fold surface model trajectory. d_{corr} values are colored from blue to white to red, with blue indicating strong negative correlation, red indicating strong positive correlation, and white indicating weak or no correlation. In (d-f), the magnitude in differences between the assembled capsid and capsid surfaces are plotted. Brown indicates a high degree of similarity, while blue indicates greater dissimilarity. In all panels, a box is drawn around monomers that form the respective ROI.

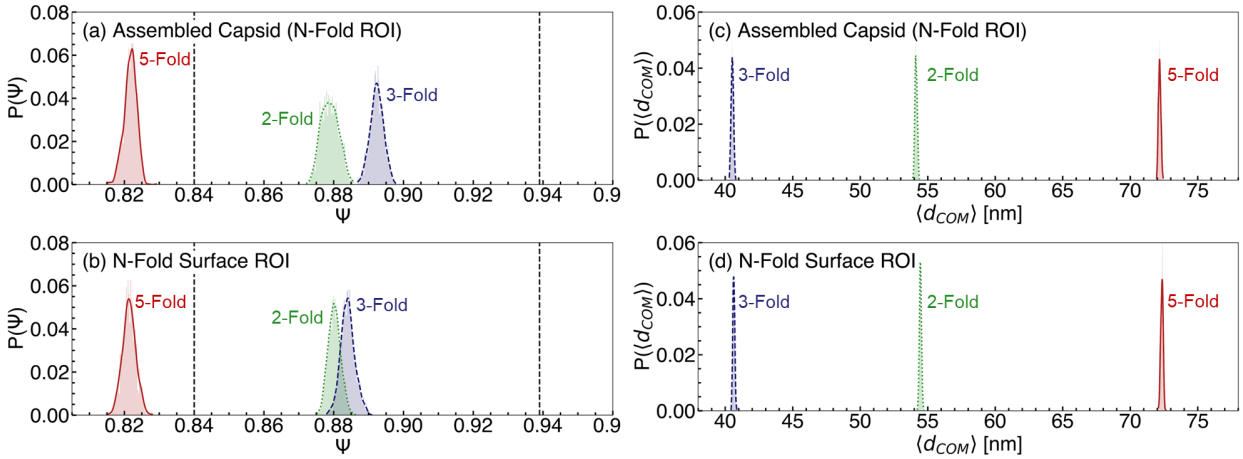

Figure 6: Global structure of the 5-, 3-, and 2-fold ROIs from the assembled capsid and capsid surface model. (a) Probability distribution of sphericity, Ψ, from assembled capsid ROIs. (b) $P(\Psi)$ from capsid surface ROIs. (c) Probability distribution of protein-protein center-of-mass distances $\langle d_{COM} \rangle$, from assembled capsid ROIs. (d) $P(\langle d_{COM} \rangle)$ from capsid surface ROIs. The 5-fold ROI is denoted by a solid red line, the 3-fold ROI by a dashed blue line, and the 2-fold ROI by a dotted green line.

relative to the unrestrained surface. Thus, while contextual proteins do not capture assembled capsid dynamics, the 5-fold ROI in the surface model faithfully reflects the 5-fold dynamics of the fully assembled capsid.

Comparable trends emerge for the 3-fold and 2-fold surfaces. In Fig. 5d-f, we plot the difference in correlation strength ($|\Delta d_{corr}|$) between assembled capsid and capsid surface models. Across all three surfaces, d_{corr} is similar for ROI proteins in both the capsid surface and assembled capsid. The magnitude of $|\Delta d_{corr}|$ is much larger for peripheral, context-providing proteins. This underscores the importance of molecular context in virus simulations, and explains why truncated models (e.g., isolated capsid monomers) often fail to capture the dynamics of fully assembled capsids.

Having established that the 5-fold ROI preserves concerted dynamics of the assembled capsid, we next asked whether the capsid surface model also maintains its overall architecture. In particular, we sought to ensure that no large-scale conformational changes—such as collapse of the surface, distortions in curvature, or unreasonable protein-protein separations—arise after removal of position restraints. To this end, we compared the global structure of the 5-fold surface model to the fully assembled PPV capsid using sphericity measurements, Ψ,

and protein-protein center-of-mass distances, d_{COM}.

Sphericity measures the deviation of an object from a perfect sphere. For assembled icosahedral viruses, Ψ approaches 1, which we observe for PPV (Fig. S12) and Hadden *et al.* observed for HBV.[64] For the 5-fold PPV surface model (Fig. 6a,b), we obtain Ψ values distributed between 0.815–0.825, which is in good agreement with the Ψ values of the 5-fold ROI extracted from the fully assembled PPV capsid. For the 2- and 3-fold ROIs, $P(\Psi)$ distributions are likewise consistent between surface models and fully assembled PPV (Fig. 6a,b). The 5-fold ROI $P(\Psi)$ distributions are centered at 0.822 ± 0.0018 (assembled capsid) and 0.821 ± 0.0019 (surface model); the 3-fold ROI at 0.892 ± 0.0019 and 0.884 ± 0.0020; and the 2-fold ROI at 0.879 ± 0.0025 and 0.880 ± 0.0018. Further, sphericity follows the trend 5-fold $<$ 2-fold $<$ 3-fold, consistent across both surface models and assembled capsid.

We measure protein-protein spacing using the center-of-mass distance between all protein pairs. For each frame, we compute an average distance $\langle d_{COM} \rangle$, with $P(\langle d_{COM} \rangle)$ shown in Fig. 6c,d. $P(\langle d_{COM} \rangle)$ distributions agree closely between assembled capsid and surface ROIs: The 5-fold ROI distributions are centered at 72.183 ± 0.088 nm (assembled capsid) and 72.350 ± 0.088 nm (surface model); the 3-fold ROI at 40.531 ± 0.087 nm and 40.637 ± 0.071 nm; and the 2-fold ROI at 54.141 ± 0.086 nm and 54.466 ± 0.078 nm. These global structure metrics confirm that truncating the PPV capsid into surface models does not significantly alter the overall architecture.

3.4 Nanofragmentation Along Weak Capsid Protein Interfaces

Simulations have recently been used to explore the mechanical stability of virus capsids,[80,126–128] an attractive approach due to direct comparison with experimental methods such as nanoindentation or atomic force microscopy.[129–131] Inspired by this, we utilize pulling simulations supported by GROMACS to induce nanofragmentation in the three capsid surface models under study. Using this approach, we identify weak interfaces in the PPV capsid

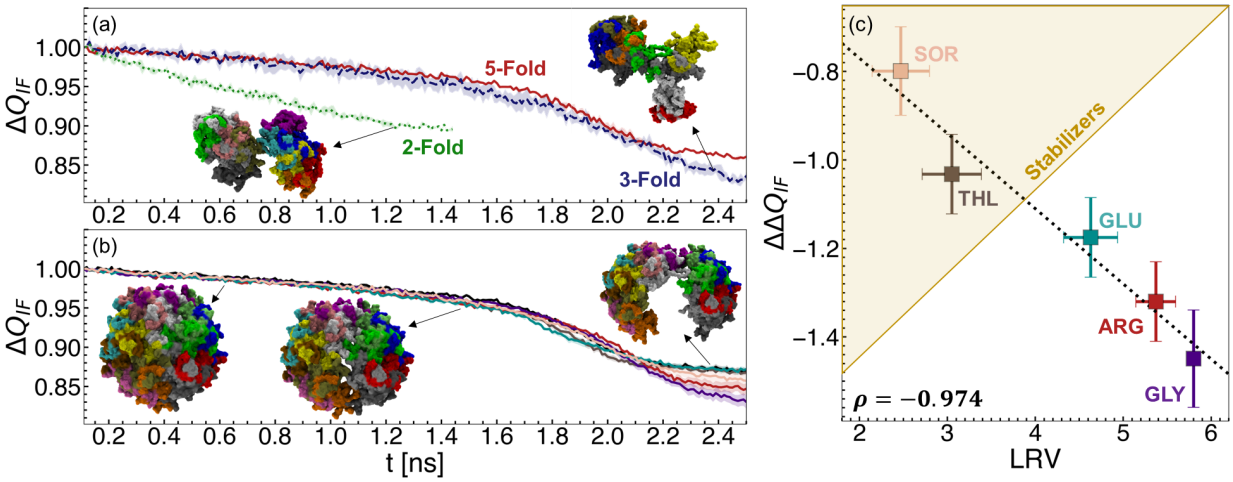

Figure 7: Change in the fraction of native interfacial contacts (ΔQ_{IF}) during nanofragmentation simulations of PPV 5-fold, 3-fold, and 2-fold surface models. (a) Change in ΔQ_{IF} for different surface ROIs. Colors indicate 5-fold, red; 3-fold, blue; and 2-fold, green capsid surface in water/0.15 M NaCl. (b) Change in ΔQ_{IF} for the PPV 5-fold surface in 0.1 M sorbitol (SOR; pink), 0.1 M trehalose (TRE; brown), 0.1 M glutamate (GLU; cyan), 0.1 M arginine (ARG; red), and 0.1 M glycine (GLY; purple). In (a) and (b), representative snapshots of 2-, 3-, and 5-fold surface fragmentation are shown, with arrows indicating the approximate points along the fragmentation curve configurations were taken. (c) Correlation between $\Delta\Delta Q_{IF}$ and LRV for a given excipient solution. Excipients known to be stabilizing experimentally are highlighted by the yellow shaded region. The Pearson correlation coefficient, ρ, is shown in the bottom left of the panel.

and explore how excipient molecules modulate the resistance/susceptibility of PPV surface fragmentation.

We first carried out nanofragmentation simulations of the 5-, 3-, and 2-fold PPV surface models in 150 mM NaCl solution (Fig. 7a). At the 2-fold surface, the onset of crack formation begins almost immediately, and the surface has a large macroscopic crack within ∼1.4 ns. The 3- and 5-fold surfaces are more resistant to fragmentation, with the rate of crack formation significantly increasing after ∼1.8 ns and surface-spanning cracks forming after 2.5 ns. By the end of the simulation, the 3-fold surface undergoes slightly more extensive fragmentation than the 5-fold surface. Together, these data point to the 5-fold surface as most resistant to fragmentation, followed by the 3-fold surface, while the 2-fold surface is the most susceptible to fragmentation. In other words, a logical progression of capsid disassembly may proceed as follows: (i) pentamer-pentamer interactions at the 2-fold axis of symmetry break apart, resulting in trimers of pentamers, (ii) pentamer-pentamer contacts begin to separate, leaving behind individual pentamers, which (iii) are the most stable intermediate unit. Such a view

supports hierarchical assembly pathways of simple icosahedral viruses wherein pentamers or hexamers form early in the process and nucleate assembly.[132,133] A susceptible 2-fold interface is consistent with several experimental studies that point to weak 2-fold interactions in T=1 icosahedral viruses,[134–138] such as PPV. Meanwhile, a stable 3-fold interface supports a previous study that found certain intertrimer mutations in a closely related parvovirus causes assembly incompetence.[139]

3.5 Excipient Effects on PPV Fragmentation are Consistent with Experimental Stability Data

We carried out similar nanofragmentation simulations for 0.1 M solutions of arginine (ARG), glutamate (GLU), glycine (GLY), trehalose (TRE), or sorbitol (SOR) with the 5-fold PPV surface model under study. These excipients were selected due to their prominence in many biological formulations.[140–143] Based on the results shown in Fig. 7a, we hypothesized that nanofragmentation of the 5-fold surface should be the rate-limiting step for total capsid disassembly.

Changes in ΔQ_{IF} during nanofragmentation simulations in excipient solutions are shown in Fig. 7b. Significant deviations in ΔQ_{IF} do not begin to occur until ∼1.4 ns. Some excipient solutions, including ARG and GLY, delay the rate of fragmentation from ∼1.4–1.8 ns, but ultimately destroy more native interfacial contacts from ∼2.0–2.5 ns than any other excipient. Others, like TRE and GLU, form larger cracks early in the simulation (∼1.4–2.0 ns) but result in roughly the same amount of broken interfacial contacts by the end of the trajectory as in water. SOR results in a ΔQ_{IF} curve that is nearly overlapping with the behavior of PPV in 0.15 M NaCl for the duration of the simulation.

Experimentally, we determine excipient effects on PPV stability through temperature stability assays. For PPV, we completed an infectivity assay (see SI for experimental details) following virus incubation at 60 °C for 3 days – sufficiently long to observe a significant decrease in the infectious titer.[114] Log reduction values (LRV) describe the decrease in the

infectious titer of heat-treated virus relative to the initial virus solution (Fig. 7c). Lower LRV relative to buffer indicates more infectious virus is present following heat treatment, indicating stability. We found that SOR (2.47 ± 0.326) and TRE (3.05 ± 0.337) are stabilizing relative to the reference solution with buffer only (4.36 ± 0.260), while GLU (4.63 ± 0.307), ARG (5.37 ± 0.226), and GLY (5.8 ± 0.010) are destabilizing at elevated temperatures.

We derive $\Delta\Delta Q_{IF}$ as a measure of excipient-mediated stability in nanofragmentation simulations, which describes ΔQ_{IF}^{Exc} relative to ΔQ_{IF}^{Wat}. The more negative $\Delta\Delta Q_{IF}$, the more susceptible the surface is to fragmentation in excipient solution relative to in 0.15 M NaCl. All excipient solutions result in $\Delta\Delta Q_{IF} < 0$ (Fig. 7c), implying the PPV 5-fold surface is more susceptible to fragmentation in excipient solutions than in 0.15 M NaCl. While this is at odds with results from heat treatment of PPV experimentally, the rank-ordering of excipients according to $\Delta\Delta Q_{IF}$ values is in excellent agreement with experimental LRV data (Fig. 7c). Specifically, SOR (-0.80 ± 0.035 LRV) and TRE (-1.03 ± 0.031 LRV) result in the smallest changes to ΔQ_{IF} relative to in 0.15 M NaCl, GLU (-1.17 ± 0.036 LRV) and ARG (-1.32 ± 0.034 LRV) are the next most destabilizing excipients, and GLY (-1.45 ± 0.037 LRV) is the most destabilizing excipient in simulations. The strong correlation (Pearson correlation coefficient, $\rho = -0.974 \pm 0.0226$) between change in stability from experiment and simulation gives us confidence that our approach can effectively identify stabilizing/destabilizing excipients.

4 Conclusions

In the development of stable viral biologics, excipient selection is a critical step. This stage of development, however, is generally time-consuming and expensive due to the lack of a fundamental understanding of excipient–virus interactions. We have tackled this issue by developing a high-throughput, high-detail workflow—CapSACIN—suitable for examining excipient–virus interactions through molecular simulations.

In the present study, we have shown that this method is capable of generating surface models that capture any region of interest from a viral capsid, such as canyons, protrusions, or dimples. We demonstrated that for PPV, inter- and intraprotein structure and dynamics are similar between a fully assembled PPV capsid model and surface models representing the 5-, 3-, and 2-fold axes of symmetries. This approach enables efficient simulations of viral systems when a region of interest is known, resulting in a $\sim 5\times$ speed-up relative to fully assembled non-enveloped virus capsids and virus-like particles (VLPs).

Further, we have shown that by applying radial pulling forces to induce nanofragmentation, we are able to identify weak protein–protein interfaces that are consistent with experimental data. Specifically, we showed that the 2-fold axis of symmetry is the capsid region most prone to nanofragmentation, while the 5-fold and 3-fold regions are more resistant to nanofragmentation. Through simulations, we are also able to elucidate the effects of excipients on capsid stability. Our simulations reveal a rank-ordering of excipient effects on PPV stability that matches experimental data. Specifically, SOR and TRE are the most stabilizing excipients of PPV, followed by GLU and ARG, while GLY is the most destabilizing excipient.

Qualitatively, our simulations agree with several previous studies that identified TRE as a broad virus stabilizer,[50,140,143–146] while ARG is capable of viral inactivation[147–149] and instability.[150] Similar to PPV, SOR has been shown to be a highly effective stabilizer of a live attenuated *Salmonella* Enteriditis vaccine.[151] On the other hand, SOR was not an effective stabilizer of live attenuated influenza vaccine.[152] The dependence of virus identity on excipient-mediated effects underscores the need for a reliable approach to predict excipient behavior prior to expending costly experimental resources.

Overall, we present a method for high-throughput investigations into excipient–virus interactions at the atomistic level. The level of detail provided by molecular simulations has the potential to enable viral biologics development from a more information-driven approach, reducing both the time and cost associated with excipient selection. The cost-effective design of viral biologics in this manner can aid in the design of formulations that are stable, effective,

and equitably distributed. Importantly, by capturing the essential dynamics of virus-like particles without requiring the simulation of entire capsids, our approach provides a practical and scalable strategy for accelerating VLP-based vaccine and gene therapy development. Because VLPs are increasingly used as safe and versatile delivery vehicles, improving their stability through rational excipient selection directly supports broader deployment of next-generation biologics. Our approach further enables future studies aiming to explore different capsid architectures (i.e., icosahedral vs. helical capsid symmetry, T>1 triangulation numbers, or empty vs. full capsids), an expanded library of excipients (i.e., surfactants, peptoids, or oligosaccharides), or exploring different denaturation mechanisms (i.e., temperature, pressure, or chemical denaturation of capsid proteins). Additionally, the generalizability of our approach to viruses that are closely related to porcine parvovirus (i.e., adeno-associated viruses, canine parvovirus, or minute virus of mice) is a clear next step for this work.

Author Contributions

J.W.P.Z.: Investigation - Computational, Conceptualization, Validation, Formal Analysis, Visualization, Writing - Original Draft, Writing - Review and Editing; I.T.: Investigation - Experimental, Writing - Review and Editing; P.M.: Conceptualization; C.L.H.: Conceptualization, Funding Acquisition, Resources, Writing - Review and Editing; S.L.P.: Conceptualization, Funding Acquisition, Resources, Writing - Review and Editing; S.S.: Conceptualization, Methodology, Validation, Supervision, Project Administration, Funding Acquisition, Resources, Writing - Review and Editing

Conflicts of interest

There are no conflicts to declare.

Data availability

Data for this article, including simulation parameter files and analysis scripts, and the complete CapSACIN workflow are available at https://github.com/SAMPEL-Group/CapSACIN.

Acknowledgments

This work is supported by U.S. National Science Foundation DMREF grants #2118788, #2118693, and #2325392. Computational resources were provided by the Minnesota Supercomputing Institute at the University of Minnesota - Twin Cities. This work additionally used computing resources from SDSC Expanse at the University of California - San Diego through allocation CHE240118 from the Advanced Cyberinfrastructure Coordination Ecosystem: Services & Support (ACCESS) program, which is supported by U.S. National Science Foundation grants #2138259, #2138286, #2138307, #2137603, and #2138296.

References

(1) D'Amico, C.; Fontana, F.; Cheng, R.; Santos, H. A. Development of vaccine formulations: past, present, and future. Drug Deliv. and Transl. Res. **2021**, 11, 353–372.

(2) Zhang, C.; Maruggi, G.; Shan, H.; Li, J. Advances in mRNA Vaccines for Infectious Diseases. Front. Immunol. **2019**, 10, 594.

(3) Minor, P. D. Live attenuated vaccines: Historical successes and current challenges. Virology **2015**, 479-480, 379–392.

(4) Rappuoli, R.; De Gregorio, E.; Costantino, P. On the mechanisms of conjugate vaccines. Proc. Natl. Acad. Sci. U.S.A. **2019**, 116, 14–16.

(5) Bull, J. J.; Nuismer, S. L.; Antia, R. Recombinant vector vaccine evolution. PLoS Comput Biol **2019**, 15, e1006857.

(6) Mettu, R.; Chen, C.-Y.; Wu, C.-Y. Synthetic carbohydrate-based vaccines: challenges and opportunities. J Biomed Sci **2020**, 27, 9.

(7) Delany, I.; Rappuoli, R.; De Gregorio, E. Vaccines for the 21st century. EMBO Mol Med **2014**, 6, 708–720.

(8) Aljabali, A. A.; Hassan, S. S.; Pabari, R. M.; Shahcheraghi, S. H.; Mishra, V.; Charbe, N. B.; Chellappan, D. K.; Dureja, H.; Gupta, G.; Almutary, A. G. et al. The Viral Capsid As Novel Nanomaterials for Drug Delivery. Future Sci. OA **2021**, 7, FSO744.

(9) Comisel, R.-M.; Kara, B.; Fiesser, F. H.; Farid, S. S. Lentiviral vector bioprocess economics for cell and gene therapy commercialization. Biochemical Engineering Journal **2021**, 167, 107868.

(10) Xiang, Y. S.; Hao, G. G. Biophysical characterization of adeno-associated virus capsid through the viral transduction life cycle. Journal of Genetic Engineering and Biotechnology **2023**, 21, 62.

(11) Eid, F.-E.; Chen, A. T.; Chan, K. Y.; Huang, Q.; Zheng, Q.; Tobey, I. G.; Pacouret, S.; Brauer, P. P.; Keyes, C.; Powell, M. et al. Systematic multi-trait AAV capsid engineering for efficient gene delivery. Nat Commun **2024**, 15, 6602.

(12) Carneiro, A. D.; Schaffer, D. V. Engineering novel adeno-associated viruses (AAVs) for improved delivery in the nervous system. Current Opinion in Chemical Biology **2024**, 83, 102532.

(13) Nisanov, A. M.; Rivera De Jesús, J. A.; Schaffer, D. V. Advances in AAV capsid

engineering: Integrating rational design, directed evolution and machine learning. Molecular Therapy **2025**, 33, 1937–1945.

(14) Lua, L. H.; Connors, N. K.; Sainsbury, F.; Chuan, Y. P.; Wibowo, N.; Middelberg, A. P. Bioengineering virus-like particles as vaccines. Biotech & Bioengineering **2014**, 111, 425–440.

(15) McNeale, D.; Dashti, N.; Cheah, L. C.; Sainsbury, F. Protein cargo encapsulation by virus-like particles: Strategies and applications. WIREs Nanomed Nanobiotechnol **2023**, 15, e1869.

(16) Dashti, N. H.; Abidin, R. S.; Sainsbury, F. Programmable *In Vitro* Coencapsidation of Guest Proteins for Intracellular Delivery by Virus-like Particles. ACS Nano **2018**, 12, 4615–4623.

(17) Das, S.; Zhao, L.; Crooke, S. N.; Tran, L.; Bhattacharya, S.; Gaucher, E. A.; Finn, M. G. Stabilization of Near-Infrared Fluorescent Proteins by Packaging in Virus-like Particles. Biomacromolecules **2020**, 21, 2432–2439.

(18) Wilkerson, J. W.; Yang, S.-O.; Funk, P. J.; Stanley, S. K.; Bundy, B. C. Nanoreactors: Strategies to encapsulate enzyme biocatalysts in virus-like particles. New Biotechnology **2018**, 44, 59–63.

(19) Jordan, P. C.; Patterson, D. P.; Saboda, K. N.; Edwards, E. J.; Miettinen, H. M.; Basu, G.; Thielges, M. C.; Douglas, T. Self-assembling biomolecular catalysts for hydrogen production. Nature Chem **2016**, 8, 179–185.

(20) Wang, Y.; Uchida, M.; Waghwani, H. K.; Douglas, T. Synthetic Virus-like Particles for Glutathione Biosynthesis. ACS Synth. Biol. **2020**, 9, 3298–3310.

(21) Organization, W. H. Temperature sensitivity of vaccines. 2006.

(22) Hanson, C. M.; George, A. M.; Sawadogo, A.; Schreiber, B. Is freezing in the vaccine cold chain an ongoing issue? A literature review. Vaccine **2017**, 35, 2127–2133.

(23) Pambudi, N. A.; Sarifudin, A.; Gandidi, I. M.; Romadhon, R. Vaccine cold chain management and cold storage technology to address the challenges of vaccination programs. Energy Reports **2022**, 8, 955–972.

(24) Price, W. Advances in Virus Research; Elsevier, 1964; Vol. 10; pp 171–217.

(25) Epand, R. M.; Epand, R. F. Thermal denaturation of influenza virus and its relationship to membrane fusion. Biochemical Journal **2002**, 365, 841–848.

(26) Ausar, S. F.; Foubert, T. R.; Hudson, M. H.; Vedvick, T. S.; Middaugh, C. R. Conformational Stability and Disassembly of Norwalk Virus-like Particles. Journal of Biological Chemistry **2006**, 281, 19478–19488.

(27) Qiu, X. Heat Induced Capsid Disassembly and DNA Release of Bacteriophage λ. PLoS ONE **2012**, 7, e39793.

(28) McHugh, C. A.; Tammariello, R. F.; Millard, C. B.; Carra, J. H. Improved stability of a protein vaccine through elimination of a partially unfolded state. Protein Science **2004**, 13, 2736–2743.

(29) Dumpa, N.; Goel, K.; Guo, Y.; McFall, H.; Pillai, A. R.; Shukla, A.; Repka, M. A.; Murthy, S. N. Stability of Vaccines. AAPS PharmSciTech **2019**, 20, 42.

(30) Kumru, O. S.; Joshi, S. B.; Smith, D. E.; Middaugh, C. R.; Prusik, T.; Volkin, D. B. Vaccine instability in the cold chain: Mechanisms, analysis and formulation strategies. Biologicals **2014**, 42, 237–259.

(31) Lloyd, J.; Cheyne, J. The origins of the vaccine cold chain and a glimpse of the future. Vaccine **2017**, 35, 2115–2120.

(32) Yu, Y. B.; Briggs, K. T.; Taraban, M. B.; Brinson, R. G.; Marino, J. P. Grand Challenges in Pharmaceutical Research Series: Ridding the Cold Chain for Biologics. Pharm Res **2021**, 38, 3–7.

(33) Dadari, I. K.; Zgibor, J. C. How the use of vaccines outside the cold chain or in controlled temperature chain contributes to improving immunization coverage in low- and middle-income countries (LMICs): A scoping review of the literature. J Glob Health **2021**, 11, 04004.

(34) Fahrni, M. L.; Ismail, I. A.-N.; Refi, D. M.; Almeman, A.; Yaakob, N. C.; Saman, K. M.; Mansor, N. F.; Noordin, N.; Babar, Z.-U.-D. Management of COVID-19 vaccines cold chain logistics: a scoping review. J of Pharm Policy and Pract **2022**, 15, 16.

(35) Hibbs, B. F.; Miller, E.; Shi, J.; Smith, K.; Lewis, P.; Shimabukuro, T. T. Safety of vaccines that have been kept outside of recommended temperatures: Reports to the Vaccine Adverse Event Reporting System (VAERS), 2008–2012. Vaccine **2018**, 36, 553–558.

(36) Thielmann, A.; Puth, M.-T.; Weltermann, B. Visual inspection of vaccine storage conditions in general practices: A study of 75 vaccine refrigerators. PLoS ONE **2019**, 14, e0225764.

(37) Thielmann, A.; Puth, M.-T.; Weltermann, B. Improving knowledge on vaccine storage management in general practices: Learning effectiveness of an online-based program. Vaccine **2020**, 38, 7551–7557.

(38) Gillett, M. B.; Suko, J. R.; Santoso, F. O.; Yancey, P. H. Elevated levels of trimethylamine oxide in muscles of deep-sea gadiform teleosts: A high-pressure adaptation? J. Exp. Zool. **1997**, 279, 386–391.

(39) Selvarajan, R.; Sibanda, T.; Tekere, M. Thermophilic bacterial communities inhabiting

the microbial mats of "indifferent" and chalybeate (iron-rich) thermal springs: Diversity and biotechnological analysis. MicrobiologyOpen **2018**, 7, e00560.

(40) Wen, X.; Wang, S.; Duman, J. G.; Arifin, J. F.; Juwita, V.; Goddard, W. A.; Rios, A.; Liu, F.; Kim, S.-K.; Abrol, R. et al. Antifreeze proteins govern the precipitation of trehalose in a freezing-avoiding insect at low temperature. Proc. Natl. Acad. Sci. U.S.A. **2016**, 113, 6683–6688.

(41) Arakawa, T.; Timasheff, S. The stabilization of proteins by osmolytes. Biophysical Journal **1985**, 47, 411–414.

(42) Burg, M. B.; Ferraris, J. D. Intracellular Organic Osmolytes: Function and Regulation. Journal of Biological Chemistry **2008**, 283, 7309–7313.

(43) Khan, S. H.; Ahmad, N.; Ahmad, F.; Kumar, R. Naturally occurring organic osmolytes: From cell physiology to disease prevention. IUBMB Life **2010**, 62, 891–895.

(44) Arakawa, T.; Tsumoto, K.; Kita, Y.; Chang, B.; Ejima, D. Biotechnology applications of amino acids in protein purification and formulations. Amino Acids **2007**, 33, 587–605.

(45) Kamerzell, T. J. Protein–excipient interactions: Mechanisms and biophysical characterization applied to protein formulation development. Advanced Drug Delivery Reviews **2011**, 42.

(46) Bongioanni, A.; Bueno, M. S.; Mezzano, B. A.; Longhi, M. R.; Garnero, C. Amino acids and its pharmaceutical applications: A mini review. International Journal of Pharmaceutics **2022**, 613, 121375.

(47) Jeong, S. H. Analytical methods and formulation factors to enhance protein stability in solution. Arch. Pharm. Res. **2012**, 35, 1871–1886.

(48) Ionova, Y.; Wilson, L. Biologic excipients: Importance of clinical awareness of inactive ingredients. PLoS ONE **2020**, 15, e0235076.

(49) Kerlin, R. L.; Li, X. Haschek and Rousseaux's Handbook of Toxicologic Pathology; Elsevier, 2013; pp 725–750.

(50) Kissmann, J.; Ausar, S. F.; Foubert, T. R.; Brock, J.; Switzer, M. H.; Detzi, E. J.; Vedvick, T. S.; Middaugh, C. Physical Stabilization of Norwalk Virus-Like Particles. Journal of Pharmaceutical Sciences **2008**, 97, 4208–4218.

(51) Zhang, Y.; Zhang, H.; Ghosh, D. The Stabilizing Excipients in Dry State Therapeutic Phage Formulations. AAPS PharmSciTech **2020**, 21, 133.

(52) Pan, L.; Liu, X.; Fan, D.; Qian, Z.; Sun, X.; Wu, P.; Zhong, L. Study of Oncolytic Virus Preservation and Formulation. Pharmaceuticals **2023**, 16, 843.

(53) Kim, Y.-S.; Jones, L. S.; Dong, A.; Kendrick, B. S.; Chang, B. S.; Manning, M. C.; Randolph, T. W.; Carpenter, J. F. Effects of sucrose on conformational equilibria and fluctuations within the native-state ensemble of proteins. Protein Sci. **2003**, 12, 1252–1261.

(54) Kamerzell, T. J.; Russell Middaugh, C. The Complex Inter-Relationships Between Protein Flexibility and Stability. Journal of Pharmaceutical Sciences **2008**, 97, 3494–3517.

(55) Jorgensen, L.; Hostrup, S.; Moeller, E. H.; Grohganz, H. Recent trends in stabilising peptides and proteins in pharmaceutical formulation – considerations in the choice of excipients. Expert Opinion on Drug Delivery **2009**, 6, 1219–1230.

(56) Ohtake, S.; Kita, Y.; Arakawa, T. Interactions of formulation excipients with proteins in solution and in the dried state. Advanced Drug Delivery Reviews **2011**, 63, 1053–1073, Number: 13.

(57) Barata, T.; Zhang, C.; Dalby, P.; Brocchini, S.; Zloh, M. Identification of Pro-

tein–Excipient Interaction Hotspots Using Computational Approaches. IJMS **2016**, 17, 853.

(58) Stärtzel, P. Arginine as an Excipient for Protein Freeze-Drying: A Mini Review. Journal of Pharmaceutical Sciences **2018**, 107, 960–967.

(59) Castañeda Ruiz, A. J.; Shetab Boushehri, M. A.; Phan, T.; Carle, S.; Garidel, P.; Buske, J.; Lamprecht, A. Alternative Excipients for Protein Stabilization in Protein Therapeutics: Overcoming the Limitations of Polysorbates. Pharmaceutics **2022**, 14, 2575.

(60) Santra, S.; Jana, M. Influence of Aqueous Arginine Solution on Regulating Conformational Stability and Hydration Properties of the Secondary Structural Segments of a Protein at Elevated Temperatures: A Molecular Dynamics Study. J. Phys. Chem. B **2022**, 126, 1462–1476.

(61) Lošdorfer Božič, A.; Šiber, A.; Podgornik, R. Statistical analysis of sizes and shapes of virus capsids and their resulting elastic properties. J Biol Phys **2013**, 39, 215–228.

(62) Ni, T.; Gerard, S.; Zhao, G.; Dent, K.; Ning, J.; Zhou, J.; Shi, J.; Anderson-Daniels, J.; Li, W.; Jang, S. et al. Intrinsic curvature of the HIV-1 CA hexamer underlies capsid topology and interaction with cyclophilin A. Nat Struct Mol Biol **2020**, 27, 855–862.

(63) Zhao, G.; Perilla, J. R.; Yufenyuy, E. L.; Meng, X.; Chen, B.; Ning, J.; Ahn, J.; Gronenborn, A. M.; Schulten, K.; Aiken, C. et al. Mature HIV-1 capsid structure by cryo-electron microscopy and all-atom molecular dynamics. Nature **2013**, 497, 643–646.

(64) Hadden, J. A.; Perilla, J. R.; Schlicksup, C. J.; Venkatakrishnan, B.; Zlotnick, A.; Schulten, K. All-atom molecular dynamics of the HBV capsid reveals insights into biological function and cryo-EM resolution limits. eLife **2018**, 7, e32478.

(65) Quinn, C. M.; Wang, M.; Fritz, M. P.; Runge, B.; Ahn, J.; Xu, C.; Perilla, J. R.; Gronenborn, A. M.; Polenova, T. Dynamic regulation of HIV-1 capsid interaction with the restriction factor TRIM5α identified by magic-angle spinning NMR and molecular dynamics simulations. Proc. Natl. Acad. Sci. U.S.A. **2018**, 115, 11519–11524.

(66) Ceres, P.; Zlotnick, A. Weak Protein-Protein Interactions Are Sufficient To Drive Assembly of Hepatitis B Virus Capsids. Biochemistry **2002**, 41, 11525–11531.

(67) Perilla, J. R.; Hadden, J. A.; Goh, B. C.; Mayne, C. G.; Schulten, K. All-Atom Molecular Dynamics of Virus Capsids as Drug Targets. J. Phys. Chem. Lett. **2016**, 7, 1836–1844.

(68) Hagan, M. F.; Zandi, R. Recent advances in coarse-grained modeling of virus assembly. Current Opinion in Virology **2016**, 18, 36–43.

(69) Wang, D.; Li, J.; Wang, L.; Cao, Y.; Kang, B.; Meng, X.; Li, S.; Song, C. Toward atomistic models of intact severe acute respiratory syndrome coronavirus 2 via Martini coarse-grained molecular dynamics simulations. Quant. Biol. **2023**, 11, 421–433.

(70) Larsson, D. S. D.; Liljas, L.; Van Der Spoel, D. Virus Capsid Dissolution Studied by Microsecond Molecular Dynamics Simulations. PLoS Comput Biol **2012**, 8, e1002502.

(71) Roberts, J. A.; Kuiper, M. J.; Thorley, B. R.; Smooker, P. M.; Hung, A. Investigation of a predicted N-terminal amphipathic α-helix using atomistic molecular dynamics simulation of a complete prototype poliovirus virion. Journal of Molecular Graphics and Modelling **2012**, 38, 165–173.

(72) Arcangeli, C.; Circelli, P.; Donini, M.; Aljabali, A. A.; Benvenuto, E.; Lomonossoff, G. P.; Marusic, C. Structure-based design and experimental engineering of a plant virus nanoparticle for the presentation of immunogenic epitopes and as a drug carrier. Journal of Biomolecular Structure and Dynamics **2014**, 32, 630–647.

(73) Perilla, J. R.; Schulten, K. Physical properties of the HIV-1 capsid from all-atom molecular dynamics simulations. Nat Commun **2017**, 8, 15959.

(74) Pohjolainen, E.; Malola, S.; Groenhof, G.; Häkkinen, H. Exploring Strategies for Labeling Viruses with Gold Nanoclusters through Non-equilibrium Molecular Dynamics Simulations. Bioconjugate Chem. **2017**, 28, 2327–2339.

(75) Durrant, J. D.; Kochanek, S. E.; Casalino, L.; Ieong, P. U.; Dommer, A. C.; Amaro, R. E. Mesoscale All-Atom Influenza Virus Simulations Suggest New Substrate Binding Mechanism. ACS Cent. Sci. **2020**, 6, 189–196.

(76) Casalino, L.; Dommer, A.; Gaieb, Z.; Barros, E. P.; Sztain, T.; Ahn, S.-H.; Trifan, A.; Brace, A.; Bogetti, A.; Ma, H. et al. AI-Driven Multiscale Simulations Illuminate Mechanisms of SARS-CoV-2 Spike Dynamics. 2020.

(77) Li, C.; Chen, W.; Lin, X.; Zhang, S.; Wang, Y.; He, X.; Ren, Y. Molecular dynamics study on the stability of foot-and-mouth disease virus particle in salt solution. Molecular Simulation **2021**, 47, 1104–1111.

(78) Dommer, A.; Casalino, L.; Kearns, F.; Rosenfeld, M.; Wauer, N.; Ahn, S.-H.; Russo, J.; Oliveira, S.; Morris, C.; Bogetti, A. et al. #COVIDisAirborne: AI-enabled multi-scale computational microscopy of delta SARS-CoV-2 in a respiratory aerosol. The International Journal of High Performance Computing Applications **2023**, 37, 28–44.

(79) Shrivastav, G.; Borkotoky, S.; Dey, D.; Singh, B.; Malhotra, N.; Azad, K.; Jayaram, B.; Agarwal, M.; Banerjee, M. Structure and energetics guide dynamic behaviour in a T = 3 icosahedral virus capsid. Biophysical Chemistry **2024**, 305, 107152.

(80) Ghaemi, Z.; Gruebele, M.; Tajkhorshid, E. Molecular mechanism of capsid disassembly in hepatitis B virus. Proc. Natl. Acad. Sci. U.S.A. **2021**, 118, e2102530118.

(81) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. Journal of Molecular Graphics **1996**, 14, 33–38.

(82) Simpson, A. A.; Hébert, B.; Sullivan, G. M.; Parrish, C. R.; Zádori, Z.; Tijssen, P.; Rossmann, M. G. The structure of porcine parvovirus: comparison with related viruses. Journal of Molecular Biology **2002**, 315, 1189–1198.

(83) Roos, W. H.; Ivanovska, I. L.; Evilevitch, A.; Wuite, G. J. L. Viral capsids: Mechanical characteristics, genome packaging and delivery mechanisms. Cell. Mol. Life Sci. **2007**, 64, 1484.

(84) Selivanovitch, E.; LaFrance, B.; Douglas, T. Molecular exclusion limits for diffusion across a porous capsid. Nat Commun **2021**, 12, 2903.

(85) Inoue, H.; Timasheff, S. N. Preferential and absolute interactions of solvent components with proteins in mixed solvent systems. Biopolymers **1972**, 11, 737–743.

(86) Record, M.; Anderson, C. Interpretation of preferential interaction coefficients of nonelectrolytes and of electrolyte ions in terms of a two-domain model. Biophysical Journal **1995**, 68, 786–794.

(87) Shukla, D.; Shinde, C.; Trout, B. L. Molecular Computations of Preferential Interaction Coefficients of Proteins. J. Phys. Chem. B **2009**, 113, 12546–12554.

(88) Vagenende, V.; Yap, M. G. S.; Trout, B. L. Molecular Anatomy of Preferential Interaction Coefficients by Elucidating Protein Solvation in Mixed Solvents: Methodology and Application for Lysozyme in Aqueous Glycerol. J. Phys. Chem. B **2009**, 113, 11743–11753.

(89) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. J. Comput. Chem. **2005**, 26, 1701–1718.

(90) Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX **2015**, 1-2, 19–25.

(91) Wardi, Y. A stochastic steepest-descent algorithm. J Optim Theory Appl **1988**, 59, 307–323.

(92) Bussi, G.; Donadio, D.; Parrinello, M. Canonical sampling through velocity rescaling. J. Chem. Phys. **2007**, 126, 014101.

(93) Evans, D. J.; Holian, B. L. The Nose–Hoover thermostat. J. Chem. Phys. **1985**, 83, 4069–4074.

(94) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An $N \bullet \log(N)$ method for Ewald sums in large systems. The Journal of Chemical Physics **1993**, 98, 10089–10092.

(95) Yeh, I.-C.; Berkowitz, M. L. Ewald summation for systems with slab geometry. The Journal of Chemical Physics **1999**, 111, 3155–3162.

(96) Lorentz, H. A. Ueber die Anwendung des Satzes vom virial in der Kinetischen Theorie der Gase. Annalen der Physik **1881**, 248, 127–136.

(97) Berthelot, D. Sur le mélange des gaz. Compt. Rendus **1898**, 126.

(98) Brooks, B. R.; Brooks, C. L.; Mackerell, A. D.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S. et al. CHARMM: The biomolecular simulation program. J. Comput. Chem. **2009**, 30, 1545–1614.

(99) Huang, J.; MacKerell, A. D. CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. J. Comput. Chem. **2013**, 34, 2135–2145.

(100) Cloutier, T.; Sudrik, C.; Sathish, H. A.; Trout, B. L. Kirkwood–Buff-derived alcohol parameters for aqueous carbohydrates and their application to preferential interaction

coefficient calculations of proteins. The Journal of Physical Chemistry B **2018**, 122, 9350–9360.

(101) Teo, R. D.; Tieleman, D. P. Evaluation of all-atom force fields in viral capsid simulations and properties. RSC Adv. **2022**, 12, 216–220.

(102) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of simple potential functions for simulating liquid water. The Journal of Chemical Physics **1983**, 79, 926–935.

(103) Olsson, M. H. M.; Søndergaard, C. R.; Rostkowski, M.; Jensen, J. H. PROPKA3: Consistent Treatment of Internal and Surface Residues in Empirical p K_{a} Predictions. J. Chem. Theory Comput. **2011**, 7, 525–537.

(104) McCammon, J. A. Protein dynamics. Rep. Prog. Phys. **1984**, 47, 1–46.

(105) Grewal, R.; Roy, S. Modeling proteins as residue interaction networks. PPL **2015**, 22, 923–933.

(106) Franke, L.; Peter, C. Visualizing the Residue Interaction Landscape of Proteins by Temporal Network Embedding. J. Chem. Theory Comput. **2023**, 19, 2985–2995.

(107) Dasetty, S.; Zajac, J. W. P.; Sarupria, S. Exploitation of active site flexibility-low temperature activity relation for engineering broad range temperature active enzymes. Mol. Syst. Des. Eng. **2023**, 8, 1355–1370.

(108) Yehorova, D.; Di Geronimo, B.; Robinson, M.; Kasson, P. M.; Kamerlin, S. C. Using residue interaction networks to understand protein function and evolution and to engineer new proteins. Current Opinion in Structural Biology **2024**, 89, 102922.

(109) Freeman, L. C. Centrality in social networks conceptual clarification. Social Networks **1978**, 1, 215–239.

(110) Brandes, U. A faster algorithm for betweenness centrality. The Journal of Mathematical Sociology **2001**, 25, 163–177.

(111) Brandes, U. On variants of shortest-path betweenness centrality and their generic computation. Social Networks **2008**, 30, 136–145.

(112) Wadell, H. Volume, Shape, and Roundness of Quartz Particles. The Journal of Geology **1935**, 43, 250–280.

(113) Trimesh [Computer software]. (2019). Retrieved from https://github.com/mikedh/trimesh.

(114) Mi, X.; Blocher McTigue, W. C.; Joshi, P. U.; Bunker, M. K.; Heldt, C. L.; Perry, S. L. Thermostabilization of viruses *via* complex coacervation. Biomater. Sci. **2020**, 8, 7082–7092.

(115) Joshi, P. U.; Decker, C.; Zeng, X.; Sathyavageeswaran, A.; Perry, S. L.; Heldt, C. L. Design Rules for the Sequestration of Viruses into Polypeptide Complex Coacervates. Biomacromolecules **2024**, 25, 741–753.

(116) Streck, A. F.; Truyen, U. Porcine Parvovirus. Current Issues in Molecular Biology **2020**, 33–46.

(117) Gencoglu, M. F.; Pearson, E.; Heldt, C. L. Porcine parvovirus flocculation and removal in the presence of osmolytes. Journal of Biotechnology **2014**, 186, 83–90.

(118) Tafur, M. F.; Vijayaragavan, K. S.; Heldt, C. L. Reduction of porcine parvovirus infectivity in the presence of protecting osmolytes. Antiviral Research **2013**, 99, 27–33.

(119) Agarwal, H.; Wang, X.; Kulkarni, N. R.; Tao, S.; Demers, C. Application of machine learning in ensuring viral safety of biotherapeutics: Case study demonstrating prediction and optimization of viral clearance performance of anion exchange chromatography. Current Research in Biotechnology **2023**, 6, 100140.

(120) Nowak, T.; Popp, B.; Roth, N. J. Choice of parvovirus model for validation studies influences the interpretation of the effectiveness of a virus filtration step. Biologicals **2019**, 60, 85–92.

(121) Miesegaes, G.; Lute, S.; Brorson, K. Analysis of viral clearance unit operations for monoclonal antibodies. Biotech & Bioengineering **2010**, 106, 238–246.

(122) Rueda, P.; Fominaya, J.; Langeveld, J. P.; Bruschke, C.; Vela, C.; Casal, J. Effect of different baculovirus inactivation procedures on the integrity and immunogenicity of porcine parvovirus-like particles. Vaccine **2000**, 19, 726–734.

(123) Marrink, S. J.; Tieleman, D. P. Perspective on the Martini model. Chem. Soc. Rev. **2013**, 42, 6801.

(124) Marrink, S. J.; Monticelli, L.; Melo, M. N.; Alessandri, R.; Tieleman, D. P.; Souza, P. C. T. Two decades of Martini: Better beads, broader scope. WIREs Comput Mol Sci **2023**, 13, e1620.

(125) Van Vlijmen, H. W.; Karplus, M. Normal Mode Calculations of Icosahedral Viruses with Full Dihedral Flexibility by Use of Molecular Symmetry. Journal of Molecular Biology **2005**, 350, 528–542.

(126) Cieplak, M.; Robbins, M. O. Nanoindentation of 35 Virus Capsids in a Molecular Model: Relating Mechanical Properties to Structure. PLoS ONE **2013**, 8, e63640.

(127) Boyd, K. J.; Bansal, P.; Feng, J.; May, E. R. Stability of Norwalk Virus Capsid Protein Interfaces Evaluated by in Silico Nanoindentation. Front. Bioeng. Biotechnol. **2015**, 3.

(128) Gibbons, M. M.; Klug, W. S. Influence of Nonuniform Geometry on Nanoindentation of Viral Capsids. Biophysical Journal **2008**, 95, 3640–3649.

(129) Michel, J. P.; Ivanovska, I. L.; Gibbons, M. M.; Klug, W. S.; Knobler, C. M.; Wuite, G. J. L.; Schmidt, C. F. Nanoindentation studies of full and empty viral capsids and the

effects of capsid protein mutations on elasticity and strength. Proc. Natl. Acad. Sci. U.S.A. **2006**, 103, 6184–6189.

(130) Mi, X.; Heldt, C. L. Single-Particle Chemical Force Microscopy to Characterize Virus Surface Chemistry. BioTechniques **2020**, 69, 363–370.

(131) Mi, X.; Bromley, E. K.; Joshi, P. U.; Long, F.; Heldt, C. L. Virus Isoelectric Point Determination Using Single-Particle Chemical Force Microscopy. Langmuir **2020**, 36, 370–378.

(132) Zlotnick, A.; Aldrich, R.; Johnson, J. M.; Ceres, P.; Young, M. J. Mechanism of Capsid Assembly for an Icosahedral Plant Virus. Virology **2000**, 277, 450–456, Publisher: Elsevier BV.

(133) Mateu, M. G. Assembly, stability and dynamics of virus capsids. Archives of Biochemistry and Biophysics **2013**, 531, 65–79.

(134) Du, C.; Cleary, S. P.; Kostelic, M. M.; Jones, B. J.; Kafader, J. O.; Wysocki, V. H. Combining Surface-Induced Dissociation and Charge Detection Mass Spectrometry to Reveal the Native Topology of Heterogeneous Protein Complexes. Anal. Chem. **2023**, 95, 13889–13896.

(135) Sherman, M. B.; Smith, H. Q.; Smith, T. J. The Dynamic Life of Virus Capsids. Viruses **2020**, 12, 618.

(136) Dong, Y.; Liu, Y.; Jiang, W.; Smith, T. J.; Xu, Z.; Rossmann, M. G. Antibody-induced uncoating of human rhinovirus B14. Proc. Natl. Acad. Sci. U.S.A. **2017**, 114, 8017–8022.

(137) Bostina, M.; Levy, H.; Filman, D. J.; Hogle, J. M. Poliovirus RNA Is Released from the Capsid near a Twofold Symmetry Axis. J Virol **2011**, 85, 776–783.

(138) Rayaprolu, V.; Kruse, S.; Kant, R.; Venkatakrishnan, B.; Movahed, N.; Brooke, D.; Lins, B.; Bennett, A.; Potter, T.; McKenna, R. et al. Comparative Analysis of Adeno-Associated Virus Capsid Stability and Dynamics. J Virol **2013**, 87, 13150–13160.

(139) Reguera, J.; Carreira, A.; Riolobos, L.; Almendral, J. M.; Mateu, M. G. Role of interfacial amino acid residues in assembly, stability, and conformation of a spherical virus capsid. Proc. Natl. Acad. Sci. U.S.A. **2004**, 101, 2724–2729.

(140) Sun, Y.; Shen, Z.; Zhang, C.; Yi, Y.; Zhu, K.; Xu, F.; Kong, W. Development of a Stable Liquid Formulation for Live Attenuated Influenza Vaccine. Journal of Pharmaceutical Sciences **2019**, 108, 2315–2322.

(141) Ohtake, S.; Martin, R.; Saxena, A.; Pham, B.; Chiueh, G.; Osorio, M.; Kopecko, D.; Xu, D.; Lechuga-Ballesteros, D.; Truong-Le, V. Room temperature stabilization of oral, live attenuated Salmonella enterica serovar Typhi-vectored vaccines. Vaccine **2011**, 29, 2761–2771.

(142) Tonnis, W.; Amorij, J.-P.; Vreeman, M.; Frijlink, H.; Kersten, G.; Hinrichs, W. Improved storage stability and immunogenicity of hepatitis B vaccine after spray-freeze drying in presence of sugars. European Journal of Pharmaceutical Sciences **2014**, 55, 36–45.

(143) Pachauri, R. Stability of live attenuated classical swine fever cell culture vaccine virus in liquid form for developing an oral vaccine. **2020**,

(144) Li, J.; Chang, Y.; Yang, S.; Zhou, G.; Wei, Y. Formulation enhanced the stability of Foot-and-mouth virus and prolonged vaccine storage. Virol J **2022**, 19, 207.

(145) Wiggan, O.; Livengood, J. A.; Silengo, S. J.; Kinney, R. M.; Osorio, J. E.; Huang, C. Y.-H.; Stinchcomb, D. T. Novel formulations enhance the thermal stability of live-attenuated flavivirus vaccines. Vaccine **2011**, 29, 7456–7462.

(146) Schlehuber, L. D.; McFadyen, I. J.; Shu, Y.; Carignan, J.; Duprex, W. P.; Forsyth, W. R.; Ho, J. H.; Kitsos, C. M.; Lee, G. Y.; Levinson, D. A. et al. Towards ambient temperature-stable vaccines: The identification of thermally stabilizing liquid formulations for measles virus using an innovative high-throughput infectivity assay. Vaccine **2011**, 29, 5031–5039.

(147) Ohtake, S.; Arakawa, T.; Koyama, A. H. Arginine as a Synergistic Virucidal Agent. Molecules **2010**, 15, 1408–1424.

(148) Meingast, C.; Heldt, C. L. Arginine-enveloped virus inactivation and potential mechanisms. Biotechnology Progress **2020**, 36, e2931.

(149) Meingast, C. L.; Joshi, P. U.; Turpeinen, D. G.; Xu, X.; Holstein, M.; Feroz, H.; Ranjan, S.; Ghose, S.; Li, Z. J.; Heldt, C. L. Physiochemical properties of enveloped viruses and arginine dictate inactivation. Biotechnology Journal **2021**, 16, 2000342.

(150) Zajac, J. W. P.; Muralikrishnan, P.; Tohidian, I.; Zeng, X.; Heldt, C. L.; Perry, S. L.; Sarupria, S. Flipping out: role of arginine in hydrophobic interactions and biological formulation design. Chem. Sci. **2025**, 16, 6780–6792.

(151) Barbour, E.; Abdelnour, A.; Jirjis, F.; Faroon, O.; Farran, M. Evaluation of 12 stabilizers in a developed attenuated Salmonella Enteritidis vaccine. Vaccine **2002**, 20, 2249–2253.

(152) White, J. A.; Estrada, M.; Flood, E. A.; Mahmood, K.; Dhere, R.; Chen, D. Development of a stable liquid formulation of live attenuated influenza vaccine. Vaccine **2016**, 34, 3676–3683.