Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

View Article Online
View Journal | View Issue

Cite this: *Chem. Sci.*, 2025, **16**, 6780

All publication charges for this article have been paid for by the Royal Society of Chemistry

Received 23rd December 2024
Accepted 9th March 2025

DOI: 10.1039/d4sc08672d

rsc.li/chemical-science

Flipping out: role of arginine in hydrophobic interactions and biological formulation design†

Jonathan W. P. Zajac, [ae] Praveen Muralikrishnan,[be] Idris Tohidian,[c] Xianci Zeng,[d] Caryn L. Heldt, [c] Sarah L. Perry [d] and Sapna Sarupria [*ae]

Arginine has been a mainstay in biological formulation development for decades. To date, the way arginine modulates protein stability has been widely studied and debated. Here, we employed a hydrophobic polymer to decouple hydrophobic effects from other interactions relevant to protein folding. While existing hypotheses for the effects of arginine can generally be categorized as either direct or indirect, our results indicate that direct and indirect mechanisms of arginine co-exist and oppose each other. At low concentrations, arginine was observed to stabilize hydrophobic polymer folding *via* a sidechain-dominated direct mechanism, while at high concentrations, arginine stabilized polymer folding *via* a backbone-dominated indirect mechanism. Upon introducing partially charged polymer sites, arginine destabilized polymer folding. Further, we found arginine-induced destabilization of a model virus similar to direct-mechanism destabilization of the charged polymer and concentration-dependent stabilization of a model protein similar to the indirect mechanism of hydrophobic polymer stabilization. These findings highlight the modular nature of the widely used additive arginine, with relevance in the information-driven design of stable biological formulations.

1 Introduction

Maintaining native protein structures in biological formulations poses a challenge, and is commonly addressed by strategic additive incorporation.[1–3] Arginine stands out as a frequently employed additive in such formulations, spanning both therapeutic proteins[4] and vaccines.[5,6] Arginine has been widely used as an aggregation suppressor, an agent for protein refolding, a cryoprotectant during lyophilization, and in protein purification.[7–9] Once hailed as a universal stabilizer, emerging studies paint a foggier picture of the effects of arginine. In some settings, the presence of arginine has accelerated the aggregation,[10–12] denaturation,[13–15] and inactivation[16,17] of certain proteins and viruses. Additional studies have found that arginine mechanisms are dependent on concentration.[18–21] Hence, the existing literature on arginine reveals a lack of a cohesive understanding of its multi-faceted effects on protein stability.

In general, additive effects on protein stability are thought to be either direct or indirect. Direct mechanisms involve direct protein–additive interactions, while indirect mechanisms influence protein stability by modulating the surrounding solvent structure. It remains debated whether arginine acts primarily *via* a direct or indirect mechanism. Several studies called attention to direct interactions between arginine and aromatic residues,[7,22,23] acidic residues,[12,20,24] and hydrophobic moieties.[25,26] Other studies have proposed clusters of free arginine molecules in solution enable the crowding out of protein–protein interactions,[20,23,27,28] or alteration of hydration shell water dynamics.[29,30] The wide range of observations related to the role of arginine on protein stability suggests that arginine harbors diverse, context-dependent mechanisms.

To elucidate the mechanisms through which arginine influences protein stability, this study focuses on its effects on hydrophobic interactions. Hydrophobic interactions are key in several biologically relevant phenomena, including protein folding and stability.[31–39] Additives in solutions are known to modulate the strength of hydrophobic interactions, and in turn, the stability of proteins.[40–46] For example, simulation studies have shown trimethylamine N-oxide (TMAO) has a negligible effect on or strengthens hydrophobic interactions,[42,47–51] while these interactions are weakened in urea solutions.[41,43,44,52–56] Indeed, the effects of TMAO and urea on hydrophobic

[a]*Department of Chemistry, University of Minnesota, Minneapolis, MN 55455, USA. E-mail: sarupria@umn.edu*

[b]*Department of Chemical Engineering and Materials Science, University of Minnesota, Minneapolis, MN 55455, USA*

[c]*Department of Chemical Engineering, Michigan Technological University, Houghton, MI 49931, USA*

[d]*Department of Chemical Engineering, University of Massachusetts Amherst, MA 01003, USA*

[e]*Chemical Theory Center, University of Minnesota, Minneapolis, MN 55455, USA*

† Electronic supplementary information (ESI) available: Simulation details, error estimations, and additional analysis. See DOI: https://doi.org/10.1039/d4sc08672d

interactions are consistent with their experimentally-observed roles as a protein stabilizer and denaturant, respectively.[45,57–60]

Molecular dynamics (MD) simulations have provided valuable insights towards understanding these mechanisms as they relate to hydrophobicity. Several studies have used a hydrophobic polymer model for describing the role of solvent and additives on protein-like folding.[43,44,61–64] The use of a hydrophobic polymer model enables the decoupling of additive effects on hydrophobic *vs.* other interactions, which is challenging to do in experiments. Additionally, comparing trends for purely hydrophobic interactions with experimental observations allows for inferences about the underlying physics. In the present study, we utilize MD simulations of a hydrophobic polymer to characterize the effects of arginine on many-body hydrophobic interactions pertinent to protein stability and contextualize these findings within larger-scale models for biological formulations.

Overall, we found arginine stabilizes hydrophobic polymer folding at all concentrations under study. We discovered that arginine sits on the edge of a mechanistic flip, balancing between direct- and indirect-dominated effects. As a consequence of this balance, we found subtle modulation of the polymer chemistry (*via* partial charge incorporation) changes arginine from a stabilizing additive to a destabilizer of polymer folding. Consequently, in practical examples of formulation design, we observed arginine has variable effects on a model virus and protein.

2 Methods

2.1 Hydrophobic polymer system setup and molecular dynamics simulations

We simulated a hydrophobic polymer in arginine solutions at different concentrations (Table S1†). All simulations were performed using GROMACS 2021.4 (ref. 65 and 66) with the PLUMED 2.8.0 (ref. 67 and 68) patch applied. The hydrophobic polymer was modeled as a linear coarse-grained chain with 26 monomers, where each monomer represents a CH_2 unit with Lennard-Jones parameters $\sigma = 0.373$ nm and $\varepsilon = 0.5856$ kJ mol^{-1}.[42] Box dimensions were defined such that 1.5 nm of space separated the fully elongated polymer from the nearest box edge. For simulations with no polymer, the same box dimensions were used. Arginine was modeled in accordance with a pH of 7, resulting in a protonated sidechain (Fig. S1†). An equal number of arginine molecules and Cl^- atoms were added to the box until the desired concentration was reached. The TIP4P/2005 (ref. 69) model was used to describe water, and the CHARM22 force field was used for arginine and Cl^-.[70]

All simulations were initially subject to energy minimization using the steepest descent algorithm. *NVT* equilibration was carried out for 1 ns at 300 K, followed by a 1 ns *NPT* equilibration at 300 K and 1 atm. During equilibration, temperature was controlled according to the V-rescale thermostat,[71] while pressure was controlled *via* the Berendsen barostat.[72] Following equilibration, *NPT* production runs were completed using the Nosé–Hoover thermostat[73] and Parrinello–Rahman barostat.[74] Production runs were completed for 20 ns for arginine/water systems, and between 50–250 ns per window for arginine/polymer/water replica exchange umbrella sampling (REUS) runs. This represents total simulation time of 1.8–9 μs for each system studied depending on the convergence time (see Table S2† for further details). In all simulations, the Particle Mesh Ewald (PME) algorithm was used for electrostatic interactions with a cut-off of 1 nm. A reciprocal grid of $42 \times 42 \times 42$ cells was used with 4th order B-spline interpolation. A single cut-off of 1 nm was used for van der Waals interactions. The neighbor search was performed every 10 steps. Lorentz–Berthelot mixing rules[75,76] were used to calculate non-bonded interactions between different atom types, except for polymer–water oxygen interactions (Fig. S2 and Table S2†).

2.2 Replica exchange umbrella sampling

REUS[77] simulations were completed to sample the hydrophobic polymer conformational landscape in arginine solutions. The radius of gyration (R_g) of the hydrophobic polymer was used as the reaction coordinate to describe polymer folding/unfolding. 12 umbrella potential windows were centered between $R_g = 0.3$ and $R_g = 0.9$ nm, with a spacing of 0.05 nm. A force constant of $K = 5000$ kJ mol^{-1} nm^{-2} was used in all windows, with the exception of the window centered at $R_g = 0.45$, which used $K = 1000$ kJ mol^{-1} nm^{-2} (Fig. S3†).

The potential of mean force (PMF) along the radius of gyration of the polymer was calculated as $W(R_g) = -k_B T \ln(P(R_g))$. Biased probability distributions were reweighted according to the Weighted Histogram Analysis Method (WHAM).[78] The free energy of polymer unfolding (ΔG_u) was calculated according to:

$$\Delta G_u = -k_B T \ln\left(\frac{\int_{R_{g,cut}}^{R_{g,max}} \exp\left(\frac{-W(R_g)}{k_B T}\right) dR_g}{\int_{R_{g,min}}^{R_{g,cut}} \exp\left(\frac{-W(R_g)}{k_B T}\right) dR_g}\right) \quad (1)$$

where $R_{g,cut}$ was determined as the point between the folded and unfolded states where $\frac{\partial W(R_g)}{\partial R_g} = 0$. The convergence of our REUS runs was assessed by examining the evolution of PMFs from three replicate simulations (Fig. S4†).

We decomposed the PMF into individual components to further investigate the role of arginine in polymer folding. Following the methods outlined by several others,[46,79–81] the PMF was separated into intrapolymer degrees of freedom in vacuum, W_{vac}, and a solvent contribution, W_{solv} as:

$$W(R_g) = W_{vac}(R_g) + W_{solv}(R_g) \quad (2)$$

$W_{vac}(R_g)$ is obtained from simulations of the polymer in vacuum, and $W_{solv}(R_g)$ is calculated as $[W(R_g) - W_{vac}(R_g)]$.

In accordance with perturbation theory approaches applied to solvation phenomena,[37,39,82–87] the solvent contribution can be described as involving two steps: (i) creating a polymer-sized cavity in solution ($W_{cav}(R_g)$), and (ii) turning on attractive polymer–solvent interactions ($E_{ps}(R_g)$). $W_{solv}(R_g)$ can then be written as:

$$W_{solv}(R_g) = W_{cav}(R_g) + E_{ps}(R_g) \quad (3)$$

where:

$$E_{ps}(R_g) = E_{pw}(R_g) + E_{pa}(R_g) + E_{pc}(R_g) \quad (4)$$

$E_{pw}(R_g)$, $E_{pa}(R_g)$, and $E_{pc}(R_g)$ are ensemble averaged polymer–water, polymer–amino acid, and polymer–counterion interaction energies, respectively, and correspond to the energy associated with a state change from folded to unfolded state. All interaction energies were computed using the *gmx rerun* feature of GROMACS, defining separate energy groups for polymer, amino acid, counterion, and water. Eqn (3) is then used to determine the $W_{cav}(R_g)$ term. This decomposition approach closely parallels the methodology employed by Athawale *et al.*[79] for hydrophobic polymer in water, where $\langle U_{pw} \rangle$ and W_{hyd} from their paper correspond to E_{ps} and W_{cav} in the above formulation.

2.3 Preferential interaction coefficients

Distribution of arginine with respect to the polymer can be described *via* the preferential interaction coefficient (Γ_{PA}),[88–90]

$$\Gamma_{PA} = -\left(\frac{\partial \mu_P}{\partial \mu_A}\right)_{m_P,T,P} = \left(\frac{\partial m_A}{\partial m_P}\right)_{\mu_A,T,P} \quad (5)$$

where μ is the chemical potential, m is the concentration and W, P, and A refer to water, polymer, and an additive, respectively. This parameter is calculated in simulations using the two-domain formula[91–93] given by:

$$\Gamma_{PA} = \left\langle N_A^{local} - \left(\frac{N_A^{bulk}}{N_W^{bulk}}\right) N_W^{local} \right\rangle \quad (6)$$

where N represents the number of molecules of a given species and angular brackets denote an ensemble average. The local and bulk domain was separated by a cutoff distance R_{cut} from the polymer. Γ_{PA} gives a measure of the relative accumulation or depletion of an additive in the local domain of the hydrophobic polymer, with $\Gamma_{PA} > 0$ indicating relative accumulation (preferential interaction) and $\Gamma_{PA} < 0$ indicating relative depletion (preferential exclusion).

2.4 Hydrogen bond analysis

Hydrogen bonds were calculated according to geometric criteria of a donor–acceptor distance of $r \leq 0.35$ nm and donor–hydrogen–acceptor angle of $150° < \theta < 180°$.[94] Hydrogen bond existence correlation functions for water–water and arginine–water interactions were estimated according to:[95–97]

$$C(\tau) = \left\langle \frac{\sum_{i,j} h_{ij}(t_0) h_{ij}(t_0 + \tau)}{\sum_{i,j} h_{ij}(t_0)^2} \right\rangle \quad (7)$$

where $h_{ij}(t_0)$ is equal to 1 if there is a hydrogen bond between groups i and j at time t_0, and 0 if no hydrogen bond is present. An average over all possible values of the time origin t_0 was taken over the last 5 ns production simulations.

2.5 Contact coefficients

Contact coefficients (CC) give a measure of excipient preference for interacting with specific residues. Here, we computed CCs as described by Stumpe and Grubmüller:[98]

$$CC_X = \frac{N_{X-A}}{N_{X-W}} \frac{M_W}{M_A} \quad (8)$$

where N_{X-A} and N_{X-W} are the number of atomic contacts of protein residue X with an additive or water molecules, respectively. Atoms were defined to be in contact if any pair of heavy atoms were within a 0.35 nm cut-off. CCs are normalized by the total number of additive atoms (M_A) and water atoms (M_W) in solution. Values of $CC_X > 1$ indicate preferential interaction with an additive for residue X, while $CC_X < 1$ denotes preferential interaction with water.

2.6 Protein and virus simulation setup

Porcine parvovirus (PPV) and hen egg white lysozyme (HEWL) were used as large-scale models for formulation design. To model PPV, we constructed a surface model using 15 monomers from the published crystal structure (PDB: 1K3V).[99] This was achieved by selecting three reference C_α atoms from the 5-fold pore of the viral surface, and aligning the vector normal to these points with the z-axis of the simulation box. Box dimensions were $24 \times 24 \times 12$ nm, with the capsid surface model extending approximately 9 nm in the z-direction.

During equilibration, a position restraint with a force constant of 1000 kJ mol^{-1} nm^{-2} was applied to all capsid atoms within 1.5 nm from the wall. Capsid atoms between 1.5 nm and 3.0 nm from the wall were restrained with a force constant between 1–1000 kJ mol^{-1} nm^{-2}, with atoms further from the wall scaled to a lesser extent. Capsid atoms further than 3.0 nm from the wall were left fully flexible. PPV simulations were carried out in the *NVT* ensemble – first for 1 ns to reach the target temperature, followed by 50 ns to allow the solvent to relax. 100 ns production runs, with positional restraints on the capsid removed, were then completed and used for further analysis. Equilibration of the PPV surface systems was measured according to C_α RMSD and was observed to reach a steady state within 50 ns (Fig. S6†).

To prevent diffusion of excipient molecules into the capsid interior, a wall was placed at $z = 0$ nm, and the capsid surface was positioned on top of this wall (Fig. S6†). The validity of the surface model for PPV was assessed by measuring residue–residue correlations within capsid proteins relative to a fully assembled capsid (Fig. S7†). We find that the residue–residue correlations are well-conserved in the surface model. Because this surface model lacks the full context of a fully assembled capsid, our approach is understandably limited in reproducing protein–protein fluctuations relative to a fully-assembled capsid (Fig. S8†). We surmise that because the surface model retains its assembled structure and conserves the residue–residue correlations it is a sufficient model for assessing virus–excipient interactions at a fraction of the computational cost required for fully-assembled capsid simulations.

Arginine molecules were added to the exterior of the capsid surface model to reach a target concentration of 0.1 M, and Na^+

and Cl^- atoms were distributed throughout the box to both neutralize the system and reach a concentration of 0.15 mM. Analysis was carried out on the converged portion of the trajectory, which constituted an aggregate sampling time of 750 ns across monomers.

HEWL was modeled from an available crystal structure (PDB: 1LYZ),[100] and simulated in solution with 0.1, 0.2, 1.0, and 2.0 M arginine. HEWL simulations were carried out in the *NVT* ensemble for 1 ns to equilibrate to the target temperature, followed by 1 ns *NPT* simulations to reach the target density. HEWL production runs were completed for 100 ns in the *NPT* ensemble, reaching conformational equilibration within 20 ns (Fig. S5†). For both HEWL and PPV, the force field, thermostats, and barostats matched those used for the hydrophobic polymer systems.

2.7 Experimental temperature stability studies

Liquid samples of PPV in arginine solutions were prepared in triplicates and were put in either a heat block at 60 °C[101] 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.[101] Differential scanning fluorimetry (DSF) was performed to compute the hydrophobic exposure temperature (HET) of HEWL in arginine solutions.[102,103] More details for both experimental methods are included in the ESI.†

3 Results and discussion

3.1 Arginine favors hydrophobic polymer folding

Fig. 1a shows the PMF along the R_g reaction coordinate. In all solutions, free energy minima were observed at approximately 0.4 and 0.8 nm (configurations labeled I and III in Fig. 1g), along with a prominent free energy barrier at ∼0.6 nm representing the transition between folded and unfolded states (Fig. 1a). In pure water, hydrophobic polymer folding is unfavorable, with the unfolded state favored by ∼0.3 kT. In contrast, at all arginine concentrations, the folded state of the polymer is favored relative to pure water, and a monotonic increase in ΔG_u is observed. An additional barrier at ∼0.45 nm was identified separating two folded states, labeled as I and II in Fig. 1a. Representative polymer configurations were identified using the HDBSCAN[104] algorithm (Fig. S9†).

The folded polymer ensemble in arginine solutions exhibits free energy minima corresponding to globular (∼0.4 nm) and hairpin-like (∼0.5 nm) configurations (labeled I and II in Fig. 1g). Arginine clusters encapsulating the hydrophobic polymer are observed in each state (Fig. S10†). We propose that the free energy barrier separating these two states arises from an energetic penalty associated with breaking these encapsulating clusters. Such a mechanism is similar to that observed by Li *et al.*,[26] who observed arginine-mediated suppression of hydrophobic association.

3.2 Decomposition of arginine effects on polymer folding

Fig. 1b–f shows the decomposition of the PMF into various components. The vacuum component, W_{vac}, favors the folded state of the polymer, associated with favorable intrapolymer interactions and configurational entropy upon folding (Fig. 1b). Because this component does not depend on the presence of arginine, a balance of the remaining components dictates the effect of arginine on hydrophobic polymer folding.

Cavity formation is a crucial step in hydrophobic hydration phenomena, and has been shown to play an important role in governing the folding behavior of the hydrophobic polymer.[43,44,61–64] W_{cav} is calculated as $W_{solv}(R_g) - E_{ps}(R_g)$, as described in the Methods section. The cavitation component favors the folded state (Fig. 1c), reflecting a strong hydrophobic driving force for polymer folding.[61,79]

Attractive polymer–water interactions become more favorable with increasing R_g (Fig. 1d), indicating polymer–water interactions oppose polymer folding. It is worth noting that the free energy minima at ∼0.8 nm in the unfolded ensemble is observed in an aqueous environment (Fig. 1a), but not in vacuum (Fig. 1b). This minima arises due to favorable polymer–water interactions, consistent with prior MD simulations that showed water-mediated interactions drive large hydrophobic solutes apart.[62,87,105] Additionally, sufficient dewetting of the hydrophobic polymer is a hypothesized bottleneck to folding,[61,79,106] resulting in the free energy barrier at ∼0.6 nm separating the folded and unfolded states.

Attractive polymer–arginine interactions approach an energetic minima at ∼0.5 nm (Fig. 1e), giving rise to the global minimum observed in the overall PMF. Polymer–Cl^- interactions were observed on the order of thermal fluctuations, consistent with previous observations that Cl^- ions are depleted from the local domain of hydrophobic solutes.[41,107]

Fig. 1h shows the change in each component upon unfolding in arginine solution relative to that observed in water. In Fig. 1h, the first Δ arises from the difference between folded and unfolded states (*e.g.*, $\Delta E = \langle E_u \rangle - \langle E_f \rangle$), and represents the free energy associated with a state change from folded to unfolded. The second Δ arises from the free energy difference between arginine solution (ΔE_{arg}) and water (ΔE_{wat}) (*e.g.*, $\Delta\Delta E = \Delta E_{arg} - \Delta E_{wat}$).

With increasing arginine concentration, we observed an increasingly favorable cavitation component for folding (Fig. 1h). Specifically, we observe a change in sign in the cavitation component from negative (unfavorable) to positive (favorable) with increasing arginine concentration. In other words, formation of a smaller cavity becomes more favorable at high arginine concentration. Concurrently, preferential polymer–arginine interactions are observed to increase with increasing arginine concentration. Thus, we hypothesize that arginine may strengthen hydrophobic interactions associated with hydrophobic polymer folding *via* a surfactant-like mechanism. Such a mechanism has been described in detail for alcohol cosolvents and TMAO, where preferential cosolvent adsorption to the polymer–solvent interface drives hydrophobic polymer folding.[51,108] Our findings strengthen the notion that the effective attraction between hydrophobic solutes in the presence of amphiphilic cosolvents (such as arginine, alcohols, and TMAO) may be a generalizable phenomenon.

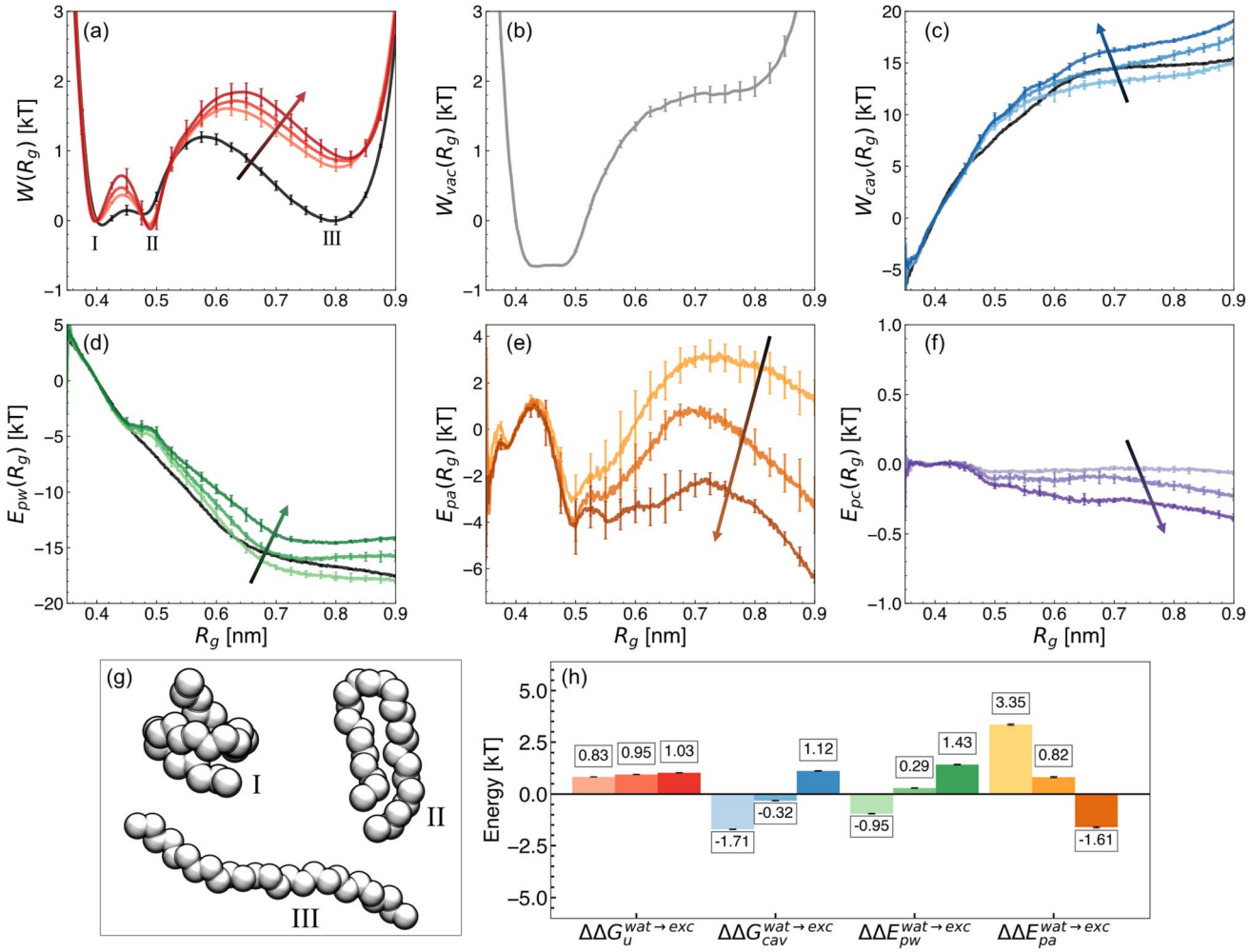

Fig. 1 PMF decomposition in 0.0 M, 0.25 M, 0.5 M, and 1.0 M arginine solutions. (a) The PMF obtained along the R_g reaction coordinate, $W(R_g)$, (b) vacuum component, W_{vac}, (c) cavitation component, W_{cav}, (d) polymer–water interactions, E_{pw}, (e) polymer–arginine interactions, E_{pa}, and (f) polymer–Cl^- interactions, E_{pc}. (g) Representative configurations along the reaction coordinate as denoted in (a) as I, II, and III. (h) Changes in overall free energy of unfolding ($\Delta\Delta G_u$), cavitation contribution ($\Delta\Delta G_{cav}$), polymer–water interactions ($\Delta\Delta E_{pw}$), and polymer–arginine interactions ($\Delta\Delta E_{pa}$). Increasing arginine concentration is denoted by increased shading (light to dark) and is indicated by arrows in (a–f). The polymer in water alone is shown in black, where appropriate. Mean values were estimated from three replicate REUS simulations. Error bars are reported as described in the ESI.† All plots are normalized to 0 at $R_g = 0.4$ nm, where appropriate.

Looking at the trends in $\Delta\Delta E_{pw}$, we found that in 0.25 M arginine solutions, the polymer–water component favors polymer unfolding relative to in pure water (Fig. 1h). This is consistent with prior work that reported the key role of water–solute interactions in hydrophobic behavior.[46,80,85,109–115] At 0.5 M and 1.0 M arginine concentrations, this component favors polymer folding. With arginine present, the local domain of the polymer exhibits a reduction in the average number of water molecules (Fig. S11†), indicating an effective expulsion of water. This, in turn, diminishes polymer–water interactions that resist polymer folding.

The polymer–arginine contribution favors the folded polymer state at 0.25 M and 0.5 M arginine concentrations relative to in pure water (Fig. 1h). However, at 1.0 M arginine concentrations, polymer–arginine interactions promote polymer unfolding. Together, these results indicate that neither direct nor indirect mechanistic hypotheses alone can describe the effects of arginine on hydrophobic polymer folding.

3.3 Mechanistic flip with increasing arginine concentration

To investigate potential competing effects of direct and indirect mechanisms, we combined the components of our PMF decomposition, delineating between those linked to direct effects (polymer–arginine and polymer–Cl^-; $\Delta\Delta G_{dir}$) and indirect effects (cavitation and polymer–water; $\Delta\Delta G_{ind}$) of arginine (Fig. 2a).

We discovered that, with increasing concentration, the mechanism underlying the effects of arginine transitions from direct to indirect dominance. At 0.25 M, cavity formation and polymer–water interactions oppose polymer folding, while arginine–polymer interactions favor folding. The balance of these components gives rise to $\Delta\Delta G_{dir} > \Delta\Delta G_{ind}$, resulting in the

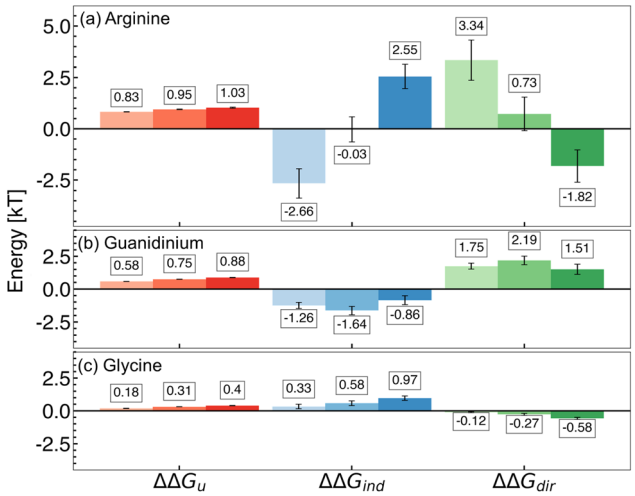

Fig. 2 Components of the free energy of hydrophobic polymer unfolding in 0.25 M, 0.50 M, and 1.0 M (a) arginine, (b) guanidinium, and (c) glycine solutions. Changes in overall free energy of unfolding ($\Delta\Delta G_u$), direct interactions ($\Delta\Delta G_{dir}$), and indirect effects ($\Delta\Delta G_{ind}$) are shown. Increasing additive concentration is denoted by increased shading (light to dark; left to right). Mean values are reported from three replicate REUS simulations. Error bars were estimated *via* error propagation (see ESI for details†).

net stabilization of folded conformations and supporting the direct mechanism hypothesis (Fig. 2a).

In contrast, for the high-concentration regime (0.5 M and 1.0 M), cavity formation and polymer–water attractive interactions favor polymer folding, while attractive arginine–polymer interactions favor unfolding of the hydrophobic polymer. In this case, indirect components dominate the free energy difference ($\Delta\Delta G_{dir} < \Delta\Delta G_{ind}$), stabilizing polymer folding and supporting the indirect hypothesis (Fig. 2a).

Thus, within the range of concentrations studied, we have uncovered that arginine exists at the edge of a mechanistic flip between direct- and indirect-dominated stabilization of many-body hydrophobic interactions. The identification of this mechanistic switch may explain the variety of hypotheses in the existing arginine literature. Because arginine is situated on this razor's edge, small changes associated with the chemistry of a protein surface, the addition of cosolvents to solution, or differences in sample preparation may cause significant changes in the modulation of hydrophobic interactions due to arginine.

3.4 Distinct roles of arginine's sidechain and backbone

Arginine is comprised of a polar backbone and an aliphatic sidechain characterized by a guanidinium group. To investigate the roles of these components on hydrophobic polymer folding, we completed an additional PMF decomposition in guanidinium and glycine solutions (see ESI for simulation details†). At all concentrations under study, we observed that guanidinium favors the hydrophobic polymer folding primarily *via* a direct mechanism, while glycine stabilizes polymer folding primarily *via* an indirect mechanism (Fig. 2b and c).

In the case of guanidinium, stabilization is driven entirely by attractive polymer–guanidinium interactions that favor folding, while polymer–water interactions and cavity formation oppose polymer folding (Fig. 2b). In glycine solutions, however, stability is driven by the inverse mechanism; polymer–water interactions and the cavitation component favor folding, while folding is opposed by attractive polymer–glycine interactions (Fig. 2c). Based on these findings, we characterize arginine as exhibiting a guanidinium-like mechanism at low concentrations and a glycine-like mechanism at high concentrations.

While glycine is known to be an effective stabilizer of proteins,[21,116,117] our observations obtained for guanidinium are somewhat surprising due to its common role as a protein denaturant.[118–120] Several studies have stressed the importance of direct interactions in guanidinium-induced denaturation, primarily *via* breaking salt bridges, competing for intra-protein hydrogen bonds, and interacting with aromatic moieties *via* cation–pi stacking.[119,121,122] Usually, this occurs at high concentrations of guanidinium salts. Our findings suggest that while guanidinium may stabilize hydrophobic interactions at low concentrations, this is outweighed by denaturing mechanisms at high concentrations.

3.5 Resolving polymer–arginine–water interactions

Thermodynamic analyses of arginine effects on polymer folding discussed above indicate that direct and indirect mechanisms co-exist and compete at all concentrations under study. To probe this further, we characterize the molecular interactions between arginine, water, and the polymer. Specifically, we look at hydrogen bonding between arginine and water to describe hydration patterns of the arginine sidechain and backbone groups. Preferential interactions are used to elucidate the balance of polymer–arginine–water interactions.

To describe how arginine interacts with water, we considered hydrogen bonding interactions between water and backbone (COO^-, NH_3+) or sidechain (Gdm^+) atoms of arginine. Overall, the number of backbone-water hydrogen bonds is greater than sidechain-water hydrogen bonds (Fig. 3a). We further observe the fraction of occupied hydrogen bonding sites to be higher for backbone groups than the sidechain (Fig. S12†). The hydrogen bond existence autocorrelation function for the 0.25 M arginine solution revealed hydrogen bonds formed between backbone-water atoms are, on average, longer-lived than sidechain-water hydrogen bonds (Fig. 3b). As arginine concentration increases, hydrogen bond lifetimes also increase for both arginine–water and water–water interactions (Fig. S13†). In principle, this may be considered a stabilizing property of arginine, as a growing body of literature has supported the role of stabilizing osmolytes in increasing hydrogen bond lifetimes and reducing water dynamics.[123–129]

Pairwise radial distribution functions (RDFs) were computed between the water oxygen (OW) and either the alpha carbon (CA) or guanidinium carbon (CZ) of arginine to quantify the local structure of water around arginine molecules (Fig. 3c). The first peak in the OW–CA RDF was observed to increase slightly with concentration. This indicates preferential hydration of the

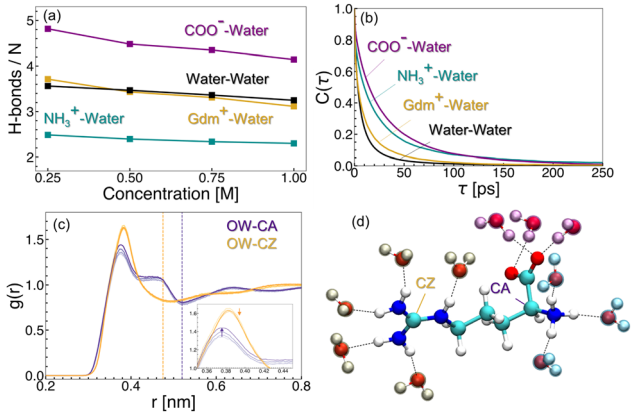

Fig. 3 Arginine–water interactions. (a) Hydrogen bonds observed between backbone groups (COO^-, NH_3+) and the guanidinium (Gdm^+) sidechain with water. (b) Hydrogen bond existence correlation functions for water–water, COO^-–water, Gdm^+–water, and NH_3+–water in 0.25 M arginine solution. (c) Radial distribution function between OW and either CA (purple) or CZ (gold). (c, inset) Arrows denote trends observed with increasing arginine concentration. (d) A representative snapshot of hydrogen-bonding interactions involving arginine and water. Water molecules interacting with the Gdm^+ sidechain are highlighted in yellow, while those interacting with NH_3+ and COO^- are shaded in blue and purple, respectively.

backbone group as more arginine molecules are introduced to the solution. There is, however, no such change observed in the OW–CZ RDF with concentration. A representative snapshot of water interactions with a single arginine molecule is shown in Fig. 3d (2D representation is shown in Fig. S14†). Together, these results indicate that the backbone of arginine is the primary site for interaction with water.

While we found arginine preferentially interacts with water *via* its backbone, we hypothesized arginine interacts with the polymer *via* its sidechain. It has been reported elsewhere that dehydration of the planar guanidinium face is important in forming face–face stacking interactions in aqueous guanidinium solutions.[130] In our case, the dehydrated face of guanidinium is expected to play a key role in direct arginine–polymer interactions, similar to interactions observed between guanidinium and hydrophobic/aromatic protein residues.[7,22,23,131]

Wyman–Tanford theory relates the dependence of any equilibrium process (such as protein folding) and preferential interaction as:[132–134]

$$-\left(\frac{\partial \Delta G_u}{\partial \mu_A}\right) = \Gamma^u_{PA} - \Gamma^f_{PA} \quad (9)$$

where Γ^u_{PA} represents the preferential interaction coefficient in the unfolded state, while Γ^f_{PA} represents the folded state. As a result, denaturants are expected to have a greater preferential interaction coefficient in the unfolded ensemble, while stabilizing osmolytes have a greater preferential interaction coefficient in the folded ensemble.[45,63,135,136] In 0.25 M arginine concentration, we observed greater preferential interactions with the folded state relative to the unfolded state (Fig. 4a and S15†). With increasing concentration, we observe a diminishing difference between Γ^u_{PA} and Γ^f_{PA}. We hypothesize that this reflects a reduced preference for arginine to interact with the folded hydrophobic polymer, which may give rise to the change in sign observed in the polymer–excipient component from favoring the folded state to favoring the unfolded state.

Preferential interaction coefficients for guanidinium and glycine solutions are shown in Fig. 4b and c, respectively. We observed that at all concentrations, guanidinium preferentially interacts with the hydrophobic polymer (Fig. 4b). This is consistent with experimental evidence, as well as a prior simulation study that observed attractive guanidinium–polymer interactions with a model hydrophobic polymer.[43,80] Glycine, meanwhile, was found to be preferentially excluded from the local domain of the hydrophobic polymer (Fig. 4c). This finding is consistent with the observed preference for the backbone of arginine to hydrogen bond with water, relative to the sidechain. Elsewhere, glycine has been observed to deplete from the surface of several model miniproteins, consistent with our findings.[136]

To explore whether the preferential interactions of arginine with polymer and water are accompanied by preferential orientations, we computed an orientation parameter inspired by Shukla and Trout.[137] This parameter was computed as the angle formed in three-dimensional space between polymer–CZ and CZ–CA vectors (Fig. 5a). The monomer closest to CZ is taken for the polymer–CZ vector. Angles where $\theta > 90°$ indicate the

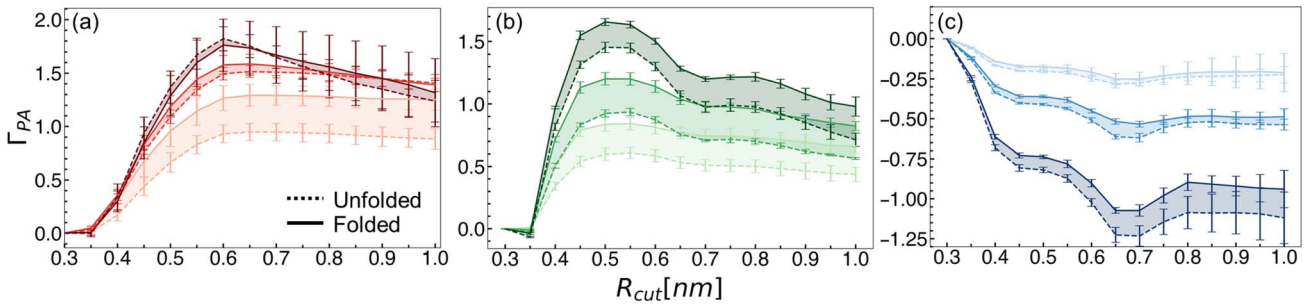

Fig. 4 Preferential interaction coefficient values as a function of the cut-off distance for the local domain of the hydrophobic polymer for (a) arginine, (b) guanidinium, and (c) glycine. Dashed lines indicate values for the unfolded state, while solid lines indicate the folded state. Increasing concentration is denoted by increased shading (light to dark). Mean values and errors were estimated from three replicate simulations. Errors are reported as standard deviations from mean values.

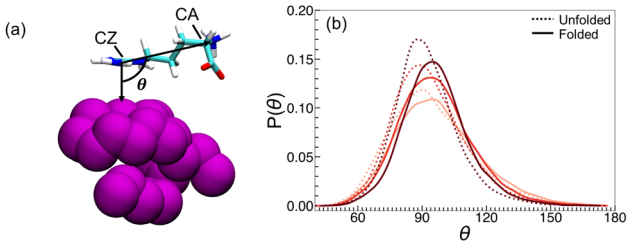

Fig. 5 Preferential orientation of arginine relative to the hydrophobic polymer. (a) Representation of the three-body angle, θ. (b) $P(\theta)$ is shown for 0.25 M, 0.5 M, and 1.0 M arginine concentrations. Solid lines denote the probability distribution for folded conformations, while dashed lines indicate the unfolded state. Increased concentration is denoted by increased shading.

arginine backbone orients towards the bulk solvent, while $\theta <$ 90° indicates the arginine backbone orients towards the polymer. We observed that for all concentrations, the probability $P(\theta)$ is skewed towards angles greater than 90° (Fig. 5b).

Further, for all arginine concentrations, the mean value of θ for the folded state is greater than that observed for unfolded configurations (Fig. 5b). This preferential orientation of arginine enables the hydrophobic face of the guanidinium sidechain to interact with the hydrophobic polymer, while extension of the backbone towards the bulk enables additional interactions with either water or other free arginine molecules. The greater ability for arginine to adopt preferred orientations in the folded state, particularly at 0.25 M, may partially explain the favorable Γ_{PA} values described previously.

4 Contextualizing the mechanistic flip of arginine

4.1 Arginine destabilizes a charged polymer

To explore whether we could shift arginine further towards a direct-dominated mechanism, we modified our hydrophobic polymer to include four beads with opposing partial charges (Fig. 6a). In pure water, charged polymer folding was observed to be favorable, resulting in a prominent free energy minimum at R_g ~0.5 nm (Fig. 6b). Relative to the hydrophobic polymer, this is indicative of a conformational preference of the polymer to adopt hairpin-like, rather than globule-like, configurations (II and I in Fig. 6a, respectively).

Upon addition of either 0.25 M or 1.0 M arginine, folding of the charged polymer becomes less favorable (Fig. 6c). We attribute this to a preference for arginine to interact with the charged sites, which are more accessible to arginine in the unfolded state. Indeed, at both concentrations, attractive polymer–arginine interactions dominate (Fig. 6d), driving polymer unfolding. This model demonstrates that even subtle changes to the chemistry of a macromolecule can re-balance arginine mechanisms.

4.2 Implications for formulation design

Arginine-induced destabilization of a charged polymer illustrates that, when hydrophobic interactions compete with other

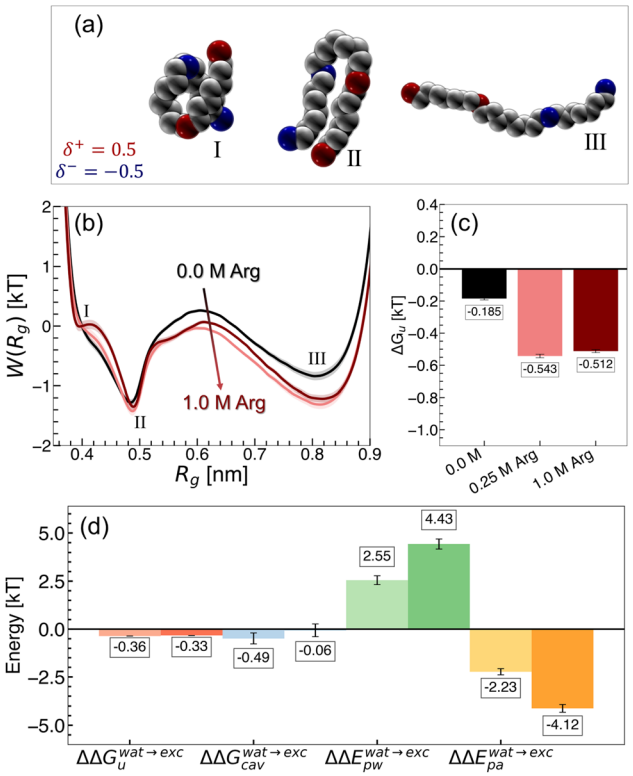

Fig. 6 Charged polymer simulations. (a) Locations of the negative (blue) and positive (red) partial charges. (b) PMFs associated with the charged polymers. (c) Free energy difference of polymer unfolding as a function of arginine concentration. (d) Change in PMF components in Arg solution relative to pure water, for the charged polymer. Increasing concentration is denoted by increased shading.

effects, the effectiveness of arginine as a stabilizer can be altered. To better understand this feature of arginine, we explored the temperature stability of two models for formulation design: PPV and HEWL.

PPV is a non-enveloped virus with a single stranded DNA genome.[138] The PPV capsid is a spherical shell comprised of 60 copies of viral proteins (VP) VP1, VP2, and VP3 in a 1 : 10 : 1 ratio, arranged in an icosahedral symmetry.[99] 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.[101,139–141]

HEWL, meanwhile, is a relatively small protein with a well-defined fold (PDB: 1LYZ) and high stability.[100,142] Due to these features, it has been widely used as a model for protein folding[143–145] and exploring osmolyte effects.[116,146,147] HEWL has also been used to investigate aggregation-suppressing effects of arginine. Several studies have proposed that arginine interacts favorably with aromatic and acidic residues of HEWL, which limits solvent exposure of aggregation-prone patches.[7,20,22,23,148] Arginine has also been observed to enhance the heat-induced aggregation of bovine serum albumin and β-lactoglobulin, but not HEWL – highlighting its context-dependent effects.[11]

To understand the effect of arginine on these biomolecules, temperature stability assays were carried out at different concentrations of arginine. For PPV, we completed an infectivity

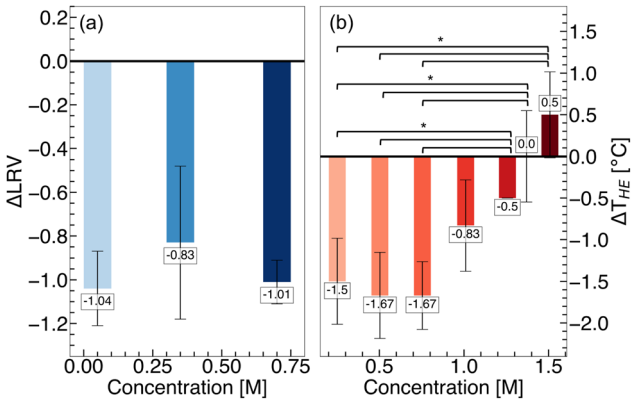

Fig. 7 Temperature stability of (a) PPV and (b) HEWL as a function of arginine concentration. ΔLRV is reported as $LRV_{Buffer} - LRV_{Arg}$, while $\Delta T_{HE} = T_{HE,Arg} - T_{HE,buffer}$. Increasing concentration is denoted by increased shading. An asterisk (*) denotes $p < 0.05$, and error bars are the standard deviation of the reported mean.

assay (see ESI for Experimental details†) following virus incubation at 60 °C for 3 days – sufficiently long to observe a significant decrease in the infectious titer.[139] Log reduction values (LRV) describe the decrease in the infectious titer of heat-treated virus, relative to the initial virus solution. We found, at all investigated arginine concentrations, more infectious PPV is lost relative to in buffer alone, resulting in negative ΔLRV values (Fig. 7a). Such a finding indicates reduced temperature stability of the virus in arginine solutions.

For HEWL, we completed a thermal shift assay (see ESI for details†) to quantify the hydrophobic exposure temperature (T_{HE}) of the protein.[102,103] At low concentrations (0.2–1.0 M) of arginine, a decrease in T_{HE} is observed relative to buffer alone, indicating destabilization of HEWL (Fig. 7b). With increasing arginine concentration (1.5 M), the temperature stability of HEWL improves. Such a finding closely resembles the temperature stability of ovalbumin and lysozyme in arginine solutions reported by Vagenende *et al.*[20] The findings at low concentrations of arginine are in contrast to the stabilization of hydrophobic polymer folding observed in the MD simulations. Thus, we infer that at these concentrations, the destabilizing effects of arginine on HEWL arise from interactions beyond hydrophobic interactions. We surmise that with increasing arginine concentrations, the strengthening of hydrophobic interactions dominates over the destabilizing effects, resulting in the stabilization of HEWL.

4.3 Molecular-level investigation of formulation models

The lack of consensus regarding the effects of arginine – further accentuated by our investigation of PPV and HEWL temperature stability – may be rationalized by the positioning of arginine at the edge of the mechanistic flip described in the present work. Additionally, we have shown that when charged beads are added to a hydrophobic polymer, this subtle change results in complete arginine-induced destabilization. To better understand the connection between the mechanistic flip of arginine and the relevance of differing molecular contexts, we completed straightforward MD simulations of HEWL and PPV systems in arginine solution.

The PPV major capsid protein, VP2, is known to self-assemble into virus-like particles that are non-pathogenic, but are morphologically similar to native PPV virions.[149] Experimental identification of this structure[99] enables PPV for molecular-level investigations *via* molecular simulations. On the capsid surface, prominent structural features include a canyon surrounding a pore at the 5-fold axis of symmetry, protruding spikes located at the 3-fold axis of symmetry, and a dimple on the 2-fold axis of symmetry.[99] We found that the use of 15 VP2 proteins is the minimum system size required to capture these structural features, reducing the system size from ~3 000 000 to ~750 000 atoms. With this truncation, we were able to reach conformational equilibrium in a computationally-feasible amount of time, which was not possible for a fully-assembled capsid system (Fig. S5†). This surface model also accurately reproduced residue–residue cross-correlations of a fully-assembled capsid (Fig. S7†), reflecting accurate local protein dynamics. We note, however, that global protein–protein dynamics are not perfectly captured (Fig. S8†). Together, this approach enables simulations of virus–excipient interactions at timescales accessible to MD without sacrificing atomistic resolution.

At the PPV surface in 0.25 M arginine solution, we observed preferential arginine accumulation at various sites across the surface (Fig. 8a and c). Specifically, *via* calculation of contact coefficients of different residue types distributed across the PPV capsid surface, we identified significant accumulation of arginine near negatively charged glutamate and aspartate residues.

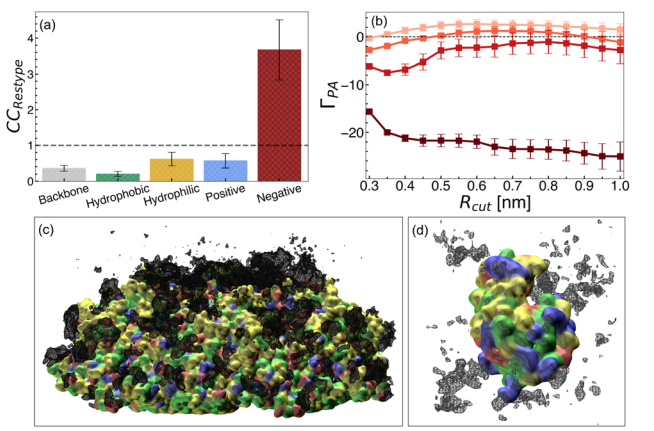

Fig. 8 Arginine distributions from PPV and HEWL systems. (a) Arginine contact coefficients at 0.1 M concentration at the PPV 5-fold surface. Bars are shown for different residue types, hydrophobic (green), hydrophilic (yellow), positive (blue), and negative (red), and was computed separately for backbone atoms (gray). (b) Preferential interaction coefficient for arginine/HEWL in 0.1 M, 0.2 M, 1.0 M, and 2.0 M arginine solutions. Increasing concentration is denoted by increased shading (light to dark). Representative snapshots and volumetric density plots of the (c) PPV 5-fold surface and (d) HEWL in arginine solutions are shown. PPV and HEWL are shown in a surface representation, and protein residues are colored according to residue type as in panel a. Arginine density at an isovalue of 0.1 atoms per $Å^3$ is shown in black mesh representation.

From these findings, we hypothesize that direct interactions between arginine and charged residues at the PPV surface drive the instability observed from experiments. In this case, the dominating mechanism may be similar to that observed for arginine-induced unfolding of a charged polymer.

For HEWL, we completed simulations of 0.1 M, 0.2 M, 1.0 M, and 2.0 M arginine solutions. With increasing concentration, we observed an increase in preferential exclusion of arginine and reduction of arginine density local to the protein (Fig. 8b and d). This trend matches the concentration-dependent exclusion of arginine from various proteins reported elsewhere.[20,27,93,134,150] This finding, along with the increase in HEWL melting temperature observed at high arginine concentration, suggests that arginine imparts stabilization of HEWL *via* indirect effects. On the other hand, direct arginine–HEWL interactions may be destabilizing, which explains the decreased melting temperature of HEWL while arginine is less excluded from the local domain of the protein.

5 Conclusions

Overall, our findings illuminate the intricate mechanisms underlying the multi-faceted effects of arginine on hydrophobic polymer folding. Arginine was observed to increase the favorability of many-body hydrophobic interactions, a key factor in protein stabilization. Our observations reveal a nuanced interplay in the impact of arginine on hydrophobic interactions, teetering on the edge of a mechanistic flip. At low concentrations, direct sidechain-driven interactions dominate, shifting to indirect backbone-driven effects at high concentrations.

The simultaneous presence of competing direct and indirect effects implies that changes in the chemistry of a protein surface, the addition of co-additives to solution, or differences in sample preparation, may cause significant changes in the mechanism of action of arginine. A shift towards the direct mechanism risks guanidinium-like denaturation of native proteins by disrupting electrostatic and hydrogen-bonding interactions. Conversely, a shift towards the indirect mechanism may yield glycine-like stabilization of native proteins through preferential hydration.

We illustrated examples of the context-dependent effects of arginine through models of protein (HEWL) and virus (PPV) stability. We observed that arginine destabilizes PPV across a wide concentration range, which we attribute to destabilizing direct arginine–PPV interactions. This mechanism resembles the effects of arginine on a charged polymer model, suggesting that in both cases, arginine-induced destabilization *via* charge–charge interactions outweighs arginine-induced stabilization of hydrophobic interactions. We further found that arginine destabilizes HEWL at low concentrations and stabilizes HEWL at high concentrations, which we found is associated with an increased exclusion of arginine from the HEWL surface. This mechanism closely resembles the concentration-dependence of the indirect mechanism of arginine in a hydrophobic polymer model. Hence, at high concentrations, HEWL destabilization due to arginine charge–charge interactions may be opposed by an overall stabilization of hydrophobic interactions.

Classical models may underestimate weak hydrogen-bonding interactions, such as CH–O hydrogen bonds, that have been found to be important in some protein-solvent systems such as in relative temperature-dependent stability of hydrogen bonds in an α-helix *versus* in a 3_{10}-helix.[151–154] While osmolyte–protein and osmolyte–water interactions are generally consistent between classical and quantum mechanics,[155–157] some discrepancies have also been reported.[158] Thus, it will be interesting to further probe the effects reported here using *ab initio* simulations that explicitly account for the impact of electronic structure on stability. However, the computational cost of such simulations remains prohibitive in terms of high-throughput investigations of excipient effects. For this reason, the acceleration of *ab initio* simulations through methods such as multiple time-stepping[159] or machine learning frameworks[160] is critical for the application of such methods towards biological formulation design.

Our results suggest overall that arginine is uniquely situated for use in formulations due to the tunable, context-dependent properties of its effects. Hence, while arginine may not be considered the universal stabilizer it once was, its balance between direct- and indirect-driven stabilization of hydrophobic interactions solidifies its significance in formulation design.

Data availability

Data for this article, including simulation parameter files and analysis scripts, are available at **https://github.com/SAMPEL-Group/Arginine-Hydrophobic-Polymer**.

Author contributions

JWPZ: investigation – computational, validation, formal analysis, visualization, writing – original draft, writing – review and editing; PM: investigation – computational, validation, writing – review and editing; IT: conceptualization, investigation – experimental, writing – review and editing; XZ: investigation – experimental, writing – review and editing; CLH: conceptualization, funding acquisition, resources, writing – review and editing; SLP: conceptualization, funding acquisition, resources, writing – review and editing; SS: conceptualization, methodology, validation, supervision, project administration, funding acquisition, resources, writing – review and editing.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This material is based upon work supported by the National Science Foundation under DMREF Grant No. 2118788, 2118693, and 2118638. Computational resources were provided by the Minnesota Supercomputing Institute at the University of Minnesota – Twin Cities. We thank computing resources through the ACCESS program for providing additional resources at the

San Diego Supercomputing Cluster at the University of California – San Diego.

Notes and references

1 S. H. Jeong, *Arch. Pharmacal Res.*, 2012, **35**, 1871–1886.
2 W. Wang, *Int. J. Pharm.*, 2005, **289**, 1–30.
3 S. Y. Patro, E. Freund and B. S. Chang, *Biotechnol. Annu. Rev.*, 2002, **8**, 55–84.
4 T. Arakawa, K. Tsumoto, Y. Kita, B. Chang and D. Ejima, *Amino Acids*, 2007, **33**, 587–605.
5 J. Hamborsky, A. Kroger and C. Wolfe, *Epidemiology and prevention of vaccine-preventable diseases*, Centers for Disease Control and Prevention, Atlanta, GA, 13th edn, 2015.
6 M. J. Mistilis, J. C. Joyce, E. S. Esser, I. Skountzou, R. W. Compans, A. S. Bommarius and M. R. Prausnitz, *Drug Delivery Transl. Res.*, 2017, **7**, 195–205.
7 K. Tsumoto, M. Umetsu, I. Kumagai, D. Ejima, J. Philo and T. Arakawa, *Biotechnol. Prog.*, 2004, **20**, 1301–1308.
8 K. Tsumoto, D. Ejima, Y. Kita and T. Arakawa, *Protein Pept. Lett.*, 2005, **12**, 613–619.
9 P. Stärtzel, *J. Pharm. Sci.*, 2018, **107**, 960–967.
10 E. Smirnova, I. Safenkova, B. Stein-Margolina, V. Shubin and B. Gurvits, *Amino Acids*, 2013, **45**, 845–855.
11 D. Shah, A. R. Shaikh, X. Peng and R. Rajagopalan, *Biotechnol. Prog.*, 2011, **27**, 513–520.
12 T. B. Eronina, N. A. Chebotareva, N. N. Sluchanko, V. V. Mikhaylova, V. F. Makeeva, S. G. Roman, S. Y. Kleymenov and B. I. Kurganov, *Int. J. Biol. Macromol.*, 2014, **68**, 225–232.
13 Q. Xie, T. Guo, J. Lu and H.-M. Zhou, *Int. J. Biochem. Cell Biol.*, 2004, **36**, 296–306.
14 B. Anumalla and N. P. Prabhu, *Appl. Biochem. Biotechnol.*, 2019, **189**, 541–555.
15 T. Arakawa and N. K. Maluf, *Int. J. Biol. Macromol.*, 2018, **107**, 1692–1696.
16 C. Meingast and C. L. Heldt, *Biotechnol. Prog.*, 2020, **36**, e2931.
17 C. L. Meingast, P. U. Joshi, D. G. Turpeinen, X. Xu, M. Holstein, H. Feroz, S. Ranjan, S. Ghose, Z. J. Li and C. L. Heldt, *Biotechnol. J.*, 2021, **16**, 2000342.
18 R. J. Falconer, C. Chan, K. Hughes and T. P. Munro, *J. Chem. Technol. Biotechnol.*, 2011, **86**, 942–948.
19 S. V. Thakkar, S. B. Joshi, M. E. Jones, H. A. Sathish, S. M. Bishop, D. B. Volkin and C. Russell Middaugh, *J. Pharm. Sci.*, 2012, **101**, 3062–3077.
20 V. Vagenende, A. X. Han, M. Mueller and B. L. Trout, *ACS Chem. Biol.*, 2013, **8**, 416–422.
21 L. Platts and R. J. Falconer, *Int. J. Pharm.*, 2015, **486**, 131–135.
22 T. Arakawa, D. Ejima, K. Tsumoto, N. Obeyama, Y. Tanaka, Y. Kita and S. N. Timasheff, *Biophys. Chem.*, 2007, **127**, 1–8.
23 D. Shukla and B. L. Trout, *J. Phys. Chem. B*, 2010, **114**, 13426–13438.
24 Y. K. Ng and L. Konermann, *J. Am. Chem. Soc.*, 2024, **146**, 8394–8406.
25 U. Das, G. Hariprasad, A. S. Ethayathulla, P. Manral, T. K. Das, S. Pasha, A. Mann, M. Ganguli, A. K. Verma, R. Bhat, S. K. Chandrayan, S. Ahmed, S. Sharma, P. Kaur, T. P. Singh and A. Srinivasan, *PLoS One*, 2007, **2**, e1176.
26 J. Li, M. Garg, D. Shah and R. Rajagopalan, *J. Chem. Phys.*, 2010, **133**, 054902.
27 C. P. Schneider and B. L. Trout, *J. Phys. Chem. B*, 2009, **113**, 2050–2058.
28 C. P. Schneider, D. Shukla and B. L. Trout, *J. Phys. Chem. B*, 2011, **115**, 7447–7458.
29 S. Santra, S. Dhurua and M. Jana, *J. Chem. Phys.*, 2021, **154**, 084901.
30 S. Santra and M. Jana, *J. Phys. Chem. B*, 2022, **126**, 1462–1476.
31 W. Kauzmann, *Adv. Protein Chem.*, 1959, **14**, 1–63.
32 C. Tanford, *J. Am. Chem. Soc.*, 1962, **84**, 4240–4247.
33 C. Tanford, *Science*, 1978, **200**, 1012–1018.
34 K. A. Dill, *Biochemistry*, 1990, **29**, 7133–7155.
35 H. J. Savage, C. J. Elliott, C. M. Freeman and J. L. Finney, *Farad. Trans.*, 1993, **89**, 2609–2617.
36 G. Hummer, S. Garde, A. Garcıa and L. Pratt, *Chem. Phys.*, 2000, **258**, 349–370.
37 L. R. Pratt and A. Pohorille, *Chem. Rev.*, 2002, **102**, 2671–2692.
38 A. Ben-Naim, *Hydrophobic interactions*, Plenum Press, New York, 2005.
39 D. Chandler, *Nature*, 2005, **437**, 640–647.
40 S. N. Timasheff, *Annu. Rev. Biophys. Biomol. Struct.*, 1993, **22**, 67–97.
41 T. Ghosh, A. Kalra and S. Garde, *J. Phys. Chem. B*, 2005, **109**, 642–651.
42 M. V. Athawale, J. S. Dordick and S. Garde, *Biophys. J.*, 2005, **89**, 858–866.
43 M. V. Athawale, S. Sarupria and S. Garde, *J. Phys. Chem. B*, 2008, **112**, 5661–5670.
44 R. Zangi, R. Zhou and B. J. Berne, *J. Am. Chem. Soc.*, 2009, **131**, 1535–1541.
45 D. R. Canchi and A. E. García, *Annu. Rev. Phys. Chem.*, 2013, **64**, 273–293.
46 N. F. A. van der Vegt and D. Nayar, *J. Phys. Chem. B*, 2017, **121**, 9986–9998.
47 S. Paul and G. N. Patey, *J. Phys. Chem. B*, 2008, **112**, 11106–11111.
48 R. D. Macdonald and M. Khajehpour, *Biophys. Chem.*, 2013, **184**, 101–107.
49 P. Ganguly, N. F. A. Van Der Vegt and J.-E. Shea, *J. Phys. Chem. Lett.*, 2016, **7**, 3052–3059.
50 Z. Su, G. Ravindhran and C. L. Dias, *J. Phys. Chem. B*, 2018, **122**, 5557–5566.
51 A. Folberth, S. Bharadwaj and N. F. A. Van Der Vegt, *Phys. Chem. Chem. Phys.*, 2022, **24**, 2080–2087.
52 A. Wallqvist, D. G. Covell and D. Thirumalai, *J. Am. Chem. Soc.*, 1998, **120**, 427–428.
53 M. Ikeguchi, S. Nakamura and K. Shimizu, *J. Am. Chem. Soc.*, 2001, **123**, 677–682.
54 N. F. A. Van Der Vegt, M.-E. Lee, D. Trzesniak and W. F. Van Gunsteren, *J. Phys. Chem. B*, 2006, **110**, 12852–12855.

55 M.-E. Lee and N. F. A. Van Der Vegt, *J. Am. Chem. Soc.*, 2006, **128**, 4948–4949.
56 T. A. Shpiruk and M. Khajehpour, *Phys. Chem. Chem. Phys.*, 2013, **15**, 213–222.
57 A. Wang and D. W. Bolen, *Biochemistry*, 1997, **36**, 9101–9108.
58 Q. Zou, B. J. Bennion, V. Daggett and K. P. Murphy, *J. Am. Chem. Soc.*, 2002, **124**, 1192–1202.
59 B. J. Bennion and V. Daggett, *Proc. Natl. Acad. Sci. U.S.A.*, 2003, **100**, 5142–5147.
60 P. Ganguly, D. Bubák, J. Polák, P. Fagan, M. Draćínský, N. F. A. Van Der Vegt, J. Heyda and J.-E. Shea, *J. Phys. Chem. Lett.*, 2022, **13**, 7980–7986.
61 P. R. ten Wolde and D. Chandler, *Proc. Natl. Acad. Sci. U.S.A.*, 2002, **99**, 6539–6543.
62 D. Nayar and N. F. A. van der Vegt, *J. Phys. Chem. B*, 2018, **122**, 3587–3595.
63 J. Mondal, G. Stirnemann and B. J. Berne, *J. Phys. Chem. B*, 2013, **117**, 8723–8732.
64 S. N. Jamadagni, R. Godawat and S. Garde, *Langmuir*, 2009, **25**, 13092–13099.
65 D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark and H. J. C. Berendsen, *J. Comput. Chem.*, 2005, **26**, 1701–1718.
66 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1–2**, 19–25.
67 M. Bonomi, D. Branduardi, G. Bussi, C. Camilloni, D. Provasi, P. Raiteri, D. Donadio, F. Marinelli, F. Pietrucci, R. A. Broglia and M. Parrinello, *Comput. Phys. Commun.*, 2009, **180**, 1961–1972.
68 G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni and G. Bussi, *Comput. Phys. Commun.*, 2014, **185**, 604–613.
69 J. L. F. Abascal and C. Vega, *J. Chem. Phys.*, 2005, **123**, 234505.
70 B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York and M. Karplus, *J. Comput. Chem.*, 2009, **30**, 1545–1614.
71 G. Bussi, D. Donadio and M. Parrinello, *J. Chem. Phys.*, 2007, **126**, 014101.
72 H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**, 3684–3690.
73 D. J. Evans and B. L. Holian, *J. Chem. Phys.*, 1985, **83**, 4069–4074.
74 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.
75 H. A. Lorentz, *Ann. Phys.*, 1881, **248**, 127–136.
76 D. Berthelot, *C. R. Acad. Sci.*, 1898, **126**, 1703–1855.
77 Y. Sugita, A. Kitao and Y. Okamoto, *J. Chem. Phys.*, 2000, **113**, 6042–6051.
78 F. Zhu and G. Hummer, *J. Comput. Chem.*, 2012, **33**, 453–465.
79 M. V. Athawale, G. Goel, T. Ghosh, T. M. Truskett and S. Garde, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 733–738.
80 R. Godawat, S. N. Jamadagni and S. Garde, *J. Phys. Chem. B*, 2010, **114**, 2246–2254.
81 S. Dasetty and S. Sarupria, *J. Phys. Chem. B*, 2021, **125**, 2644–2657.
82 L. R. Pratt and D. Chandler, *J. Chem. Phys.*, 1977, **67**, 3683–3704.
83 K. Lum, D. Chandler and J. D. Weeks, *J. Phys. Chem. B*, 1999, **103**, 4570–4577.
84 P. R. T. Wolde, *J. Phys.: Condens. Matter*, 2002, **14**, 9445–9460.
85 D. M. Huang and D. Chandler, *J. Phys. Chem. B*, 2002, **106**, 2047–2053.
86 S. Garde, A. E. Garcıa, L. R. Pratt and G. Hummer, *Biophys. Chem.*, 1999, **78**, 21–32.
87 D. Ben-Amotz, *Annu. Rev. Phys. Chem.*, 2016, **67**, 617–638.
88 G. Scatchard, *J. Am. Chem. Soc.*, 1946, **68**, 2315–2319.
89 E. F. Casassa and H. Eisenberg, *Adv. Prot. Chem.*, Elsevier, 1964, vol. 19, pp. 287–395.
90 J. A. Schellman, *Biopolymers*, 1987, **26**, 549–559.
91 H. Inoue and S. N. Timasheff, *Biopolymers*, 1972, **11**, 737–743.
92 M. Record and C. Anderson, *Biophys. J.*, 1995, **68**, 786–794.
93 D. Shukla, C. Shinde and B. L. Trout, *J. Phys. Chem. B*, 2009, **113**, 12546–12554.
94 M. Ferrario, M. Haughney, I. R. McDonald and M. L. Klein, *J. Chem. Phys.*, 1990, **93**, 5156–5166.
95 A. Luzar and D. Chandler, *J. Chem. Phys.*, 1993, **98**, 8160–8173.
96 A. Luzar and D. Chandler, *Nature*, 1996, **379**, 55–57.
97 A. Luzar, *J. Chem. Phys.*, 2000, **113**, 10663–10675.
98 M. C. Stumpe and H. Grubmüller, *J. Am. Chem. Soc.*, 2007, **129**, 16126–16131.
99 A. A. Simpson, B. Hébert, G. M. Sullivan, C. R. Parrish, Z. Zádori, P. Tijssen and M. G. Rossmann, *J. Mol. Biol.*, 2002, **315**, 1189–1198.
100 R. Diamond, *J. Mol. Biol.*, 1974, **82**, 371–391.
101 P. U. Joshi, C. Decker, X. Zeng, A. Sathyavageeswaran, S. L. Perry and C. L. Heldt, *Biomacromolecules*, 2024, **25**, 741–753.
102 S. A. Seabrook and J. Newman, *ACS Comb. Sci.*, 2013, **15**, 387–392.
103 S. Shi, A. Semple, J. Cheung and M. Shameem, *J. Pharm. Sci.*, 2013, **102**, 2471–2483.
104 L. McInnes, J. Healy and S. Astels, *J. Open Source Softw.*, 2017, **2**, 205.
105 M. Makowski, C. Czaplewski, A. Liwo and H. A. Scheraga, *J. Phys. Chem. B*, 2010, **114**, 993–1003.
106 D. Dhabal, Z. Jiang, A. Pallath and A. J. Patel, *J. Phys. Chem. B*, 2021, **125**, 5434–5442.
107 A. Kalra, N. Tugcu, S. M. Cramer and S. Garde, *J. Phys. Chem. B*, 2001, **105**, 6380–6386.
108 S. Bharadwaj, D. Nayar, C. Dalgicdir and N. F. A. Van Der Vegt, *Commun. Chem.*, 2020, **3**, 165.

109 S. Gómez, N. Rojas-Valencia, S. A. Gómez, C. Cappelli, G. Merino and A. Restrepo, *Chem. Sci.*, 2021, **12**, 9233–9245.
110 C. A. Cerdeiriña and P. G. Debenedetti, *J. Chem. Phys.*, 2016, **144**, 164501.
111 H. S. Ashbaugh, T. M. Truskett and P. G. Debenedetti, *J. Chem. Phys.*, 2002, **116**, 2907–2921.
112 N. Choudhury and B. M. Pettitt, *J. Am. Chem. Soc.*, 2005, **127**, 3556–3567.
113 M. V. Athawale, S. N. Jamadagni and S. Garde, *J. Chem. Phys.*, 2009, **131**, 115102.
114 N. Choudhury and B. M. Pettitt, *J. Phys. Chem. B*, 2005, **109**, 6422–6429.
115 H. S. Ashbaugh and M. E. Paulaitis, *J. Am. Chem. Soc.*, 2001, **123**, 10721–10728.
116 T. Arakawa and S. Timasheff, *Biophys. J.*, 1985, **47**, 411–414.
117 K. C. Bauer, S. Suhm, A. K. Wöll and J. Hubbuch, *Int. J. Pharm.*, 2017, **516**, 82–90.
118 D. B. Watlaufer, S. K. Malik, L. Stoller and R. L. Coffin, *J. Am. Chem. Soc.*, 1964, **86**, 508–514.
119 H. Meuzelaar, M. R. Panman and S. Woutersen, *Angew. Chem.*, 2015, **127**, 15470–15474.
120 S. K. Jha and S. Marqusee, *Proc. Natl. Acad. Sci. U.S.A.*, 2014, **111**, 4856–4861.
121 C. E. Dempsey, P. E. Mason, J. W. Brady and G. W. Neilson, *J. Am. Chem. Soc.*, 2007, **129**, 15895–15902.
122 J. Heyda, H. I. Okur, J. Hladílková, K. B. Rembert, W. Hunn, T. Yang, J. Dzubiella, P. Jungwirth and P. S. Cremer, *J. Am. Chem. Soc.*, 2017, **139**, 863–870.
123 M. Hishida, R. Anjum, T. Anada, D. Murakami and M. Tanaka, *J. Phys. Chem. B*, 2022, **126**, 2466–2475.
124 G. S. Jas, E. C. Rentchler, A. M. Słowicka, J. R. Hermansen, C. K. Johnson, C. R. Middaugh and K. Kuczera, *J. Phys. Chem. B*, 2016, **120**, 3089–3099.
125 A. Diaz and V. Ramakrishnan, *Comput. Biol. Chem.*, 2023, **105**, 107883.
126 G. Saladino, M. Marenchino, S. Pieraccini, R. Campos-Olivas, M. Sironi and F. L. Gervasio, *J. Chem. Theory Comput.*, 2011, **7**, 3846–3852.
127 I. Jahan and S. M. Nayeem, *RSC Adv.*, 2020, **10**, 27598–27614.
128 J. Zeman, C. Holm and J. Smiatek, *J. Chem. Eng. Data*, 2020, **65**, 1197–1210.
129 R. Gazi, S. Maity and M. Jana, *ACS Omega*, 2023, **8**, 2832–2843.
130 M. Vazdar, J. Heyda, P. E. Mason, G. Tesei, C. Allolio, M. Lund and P. Jungwirth, *Acc. Chem. Res.*, 2018, **51**, 1455–1464.
131 P. Gund, *J. Chem. Educ.*, 1972, **49**, 100.
132 J. Wyman, *Adv. Protein Chem.*, Elsevier, 1964, vol. 19, pp. 223–286.
133 S. N. Timasheff, *Proc. Natl. Acad. Sci. U.S.A.*, 2002, **99**, 9721–9726.
134 D. Shukla, C. P. Schneider and B. L. Trout, *Adv. Drug Delivery Rev.*, 2011, **63**, 1074–1085.
135 J. Mondal, D. Halverson, I. T. S. Li, G. Stirnemann, G. C. Walker and B. J. Berne, *Proc. Natl. Acad. Sci. U.S.A.*, 2015, **112**, 9270–9275.
136 M. Mukherjee and J. Mondal, *J. Phys. Chem. B*, 2020, **124**, 6565–6574.
137 D. Shukla and B. L. Trout, *J. Phys. Chem. B*, 2011, **115**, 1243–1253.
138 A. F. Streck and U. Truyen, *Curr. Issues Mol. Biol.*, 2020, 33–46.
139 X. Mi, W. C. Blocher McTigue, P. U. Joshi, M. K. Bunker, C. L. Heldt and S. L. Perry, *Biomater. Sci.*, 2020, **8**, 7082–7092.
140 M. F. Gencoglu, E. Pearson and C. L. Heldt, *J. Biotechnol.*, 2014, **186**, 83–90.
141 M. F. Tafur, K. S. Vijayaragavan and C. L. Heldt, *Antiviral Res.*, 2013, **99**, 27–33.
142 B. Matthews, *Advances in Protein Chemistry*, Elsevier, 1995, vol. 46, pp. 249–278.
143 C. M. Dobson, P. A. Evans and S. E. Radford, *Trends Biochem. Sci.*, 1994, **19**, 31–37.
144 T. Yoshimura, H. Noguchi, T. Inoue and N. Saitô, *Biophys. Chem.*, 1991, **40**, 277–291.
145 G. Merlini and V. Bellotti, *Clin. Chim. Acta*, 2005, **357**, 168–172.
146 M. M. Santoro, Y. Liu, S. M. A. Khan, L. X. Hou and D. W. Bolen, *Biochemistry*, 1992, **31**, 5278–5283.
147 B. Adamczak, M. Wieczór, M. Kogut, J. Stangret and J. Czub, *Biochem. J.*, 2016, **473**, 3705–3724.
148 S. Brudar and B. Hribar-Lee, *Int. J. Mater. Sci.*, 2023, **24**, 1197.
149 P. Rueda, J. Fominaya, J. P. Langeveld, C. Bruschke, C. Vela and J. Casal, *Vaccine*, 2000, **19**, 726–734.
150 T. K. Cloutier, C. Sudrik, N. Mody, S. A. Hasige and B. L. Trout, *mAbs*, 2020, **12**, 1816312.
151 S. Horowitz and R. C. Trievel, *J. Biol. Chem.*, 2012, **287**, 41576–41582.
152 S. Scheiner, *Phys. Chem. Chem. Phys.*, 2011, **13**, 13860.
153 A. Tkatchenko, M. Rossi, V. Blum, J. Ireta and M. Scheffler, *Phys. Rev. Lett.*, 2011, **106**, 118102.
154 O. T. Unke, M. Stöhr, S. Ganscha, T. Unterthiner, H. Maennel, S. Kashubin, D. Ahlin, M. Gastegger, L. Medrano Sandonas, J. T. Berryman, A. Tkatchenko and K.-R. Müller, *Sci. Adv.*, 2024, **10**, eadn4397.
155 V. Spiwok, *Molecules*, 2017, **22**, 1038.
156 V. Spiwok, P. Lipovová, T. Skálová, E. Vondráčková, J. Dohnálek, J. Hašek and B. Králová, *J. Comput.-Aided Mol. Des.*, 2006, **19**, 887–901.
157 C. J. Sahle, M. A. Schroer, J. Niskanen, M. Elbers, C. M. Jeffries and C. Sternemann, *Phys. Chem. Chem. Phys.*, 2020, **22**, 11614–11624.
158 M. K. Tiwari and R. K. Murarka, *Phys. Chem. Chem. Phys.*, 2021, **23**, 5527–5539.
159 R. Kar, S. Mandal, V. Thakkur, B. Meyer and N. N. Nair, *J. Chem. Theory Comput.*, 2023, **19**, 8351–8364.
160 V. Botu and R. Ramprasad, *Int. J. Quantum Chem.*, 2015, **115**, 1074–1083.