



Cite this: *Soft Matter*, 2024, 20, 8455

Effective patchiness from critical points of a coarse-grained protein model with explicit shape and charge anisotropy

Jens Weimar,  Frank Hirschmann  and Martin Oettel *

Colloidal model systems are successful in rationalizing emergent phenomena like aggregation, rheology and phase behaviour of protein solutions. Colloidal theory in conjunction with isotropic interaction models is often employed to estimate the stability of such solutions. In particular, a universal criterion for the reduced second virial coefficient at the critical point B_2^* is frequently invoked which is based on the behavior of short-range attractive fluids (Noro–Frenkel rule, $B_2^* \approx -1.5$). However, if anisotropic models for the protein–protein interaction are considered, e.g. the Kern–Frenkel (KF) patchy particle model, the value of the B_2^* criterion is shifted to lower values and explicitly depends on the number of patches. If an explicit shape anisotropy is considered, as e.g. in a coarse-grained protein model, the normalization of B_2^* becomes ambiguous to some extent, as no unique exclusion volume can be defined anymore. Here, we investigate a low-resolution, coarse-grained model for the globular protein bovine serum albumin (BSA) and study effects of charge-anisotropy on the phase diagram (determined by simulations) at the isoelectric point. We present methods of assigning an “effective patchiness” to our protein model by comparing its critical properties to the KF model. We find that doubling the native charges increases the critical temperature T_c by $\approx 14\%$ and that our BSA model can be compared to a 3 to 5 patch KF model. Finally, we argue that applying existing B_2^* criteria from colloidal theory should be done with care, due to multiple, physically plausible ways of how to assign effective diameters to shape-anisotropic models.

Received 15th July 2024,
Accepted 7th October 2024

DOI: 10.1039/d4sm00867g

rsc.li/soft-matter-journal

1 Introduction

Liquid–liquid phase transitions (LLPS) in protein solutions are a common phenomenon, they are related to biological function,¹ but they also play an important role in the pathogenesis of some human diseases^{2,3} and also affect pathways to crystallization.^{4,5} Concepts from soft matter science (liquid state and colloidal theory) have been useful to see general and universal trends in protein phase behavior.⁶ This has been possible despite the fact that proteins are in general very anisotropic in shape and mutual interactions. Furthermore, interactions can be specific and non-specific and may sensitively depend on solvent conditions, and additionally proteins are flexible to a highly varying degree. In the simplest approach, isotropic colloidal models (hard spheres with short-ranged attractions) are used to explain features of protein phase diagrams. Notably these are the metastability of liquid–liquid demixing with respect to crystallization and an overall criterion for the necessary strength of attraction for the onset of this

demixing. The latter has been proposed in ref. 7 and 8 and is based on an approximately universal phase behavior of short-range attractive colloidal models. It is formulated in terms of a reduced second virial coefficient B_2^* at the critical temperature which takes the approximate value $B_2^* \approx -1.5$ for the investigated colloidal models and has been suggested also to hold for protein solutions. In the following we refer to this finding as “Noro–Frenkel rule” or “law of corresponding states”.

A comprehensive study from the group of the late Stefan Egelhaaf⁹ finds a rather good agreement with the Noro–Frenkel rule if one scales the low-concentration part of the protein binodal with the corresponding part from a short-range square well fluid. On the other hand, the anisotropy of the attractive interactions between proteins frequently gives rise to clustering behavior which is reminiscent of the one seen in patchy particle models. For the latter (provided the interactions are only patchy and hard core), however, the location of the gas–liquid critical point is not consistent with the Noro–Frenkel rule and mainly the number of patches determine the reduced second virial coefficient, e.g. $B_2^* \approx -30$ for three patches, $B_2^* \approx -4 \dots -5$ for four patches *etc.*¹⁰ Thus one would expect also a deviation of B_2^* from the Noro–Frenkel value for protein solutions if there is

Institute for Applied Physics, University of Tübingen, Auf der Morgenstelle 10, 72076 Tübingen, Germany. E-mail: martin.oettel@uni-tuebingen.de



enough anisotropy in the interactions. Real protein–protein interactions are neither purely isotropic nor purely patchy, but always a combination of both. A study about patchy square well fluids, which represent such a combination, did not observe a law of corresponding states, *i.e.* in the form of B_2^* vs. patch number.¹¹ This implies that using universal B_2^* criteria should only be applicable if the interactions of the model at hand are considered either mostly isotropic or mostly patchy.

Anisotropy is of course fully present if an atomistic protein model in explicit solvent is used. From a simulation point of view, it is still difficult to investigate phase separation in all-atom models for biomolecules and currently is restricted to small proteins and peptides.¹² Here it is better to resort to coarse-grained (CG) models in which the protein is built from pseudo-atoms that unite multiple atoms into a single, bigger one. There exist numerous coarse-graining approaches (for reviews see *e.g.* ref. 12–14), with each of them operating at a certain resolution level. High-resolution CG models like the well-known MARTINI¹⁵ force field are often used with explicit water (*e.g.* for intrinsically disordered proteins¹⁶ or nanoparticles¹⁷), whereas low-resolution models in which only a few effective beads represent the molecule interactions are modeled by implicit-solvent potentials, as *e.g.* of DLVO type. These low-resolution models have been widely employed in the literature, *e.g.* in studies on antibodies^{18–21} but also for more globular proteins.^{22,23}

Here we investigate the phase behavior of a coarse-grained model for the globular protein bovine serum albumine (BSA) consisting of 6 beads, using grand canonical Monte Carlo (GCMC) simulations. Interactions between beads are of DLVO type, the total BSA charge is zero but there is a fixed distribution of relative bead charges with one multiplicative parameter scaling the individual bead charges. In this model, the shape anisotropy (coming from the arrangement of the beads) determines also an asymmetry of the van der Waals attractions. Additionally, the bead charges add a further charge asymmetry to these interactions. We study the influence of charge on the critical point and the bead–bead correlations and attempt to quantify an effective patchiness from the B_2^* -value for the investigated system.

The motivation for studying a BSA model is linked to the existence of LLPS of BSA in solutions with trivalent cations.²⁴ On the one hand, these LLPS have been rationalized using patchy particle models,^{25,26} but on the other hand, scattering studies near the critical point in these systems show only small deviations from the Noro–Frenkel rule, thereby suggesting a small influence of patchiness on the phase diagram. Furthermore, BSA is well characterized by scattering methods in terms of its shape, dynamics and diffusion properties,^{27–31} it appears to be a rather rigid molecule suitable for modelling.

The paper is structured as follows: we first introduce our coarse-grained BSA model and its intermolecular interactions in Section 2, where in total six different parameter combinations are defined. Subsequently, in Section 3 the determination of critical and binodal points *via* GCMC simulation is described. In Section 4 the resulting phase diagrams are shown, along with the effective potentials at criticality, the second virial

coefficients and effective molecule diameters, which are calculated *via* four different approaches. Additionally, we present bead–bead correlation functions and discuss their charge dependence. In Section 5 we discuss how an “effective patchiness” can be deduced from the critical point of our models with paying special attention to the ambiguity in defining effective hard diameters and thus in the normalization of the second virial coefficient (needed for defining B_2^*). Furthermore the relationship of our results to experiments is discussed, before we conclude in Section 6.

2 Model description

The BSA model investigated in this work has previously been derived and analysed *via* Brownian dynamics (BD) simulations in ref. 32. In brief, the authors conducted single-molecule all-atom MD simulations of a crystal-structure of BSA and subsequently used direct Boltzmann inversion to obtain coarse-grained force field parameters, yielding a flexible, six-beaded, low-resolution CG model, in which each bead represents a biological domain, see Fig. 1(a) for illustration. The study assessed the influence of flexibility in self-crowded conditions on static and dynamic solution properties like center-of-mass pair-correlations and diffusion. However, the inter-molecular interactions were taken to be purely repulsive and soft, accounting only for excluded volume of the molecules, corresponding to a high-salt regime, where electrostatic interactions are sufficiently screened.

In this work, we consider more realistic interactions between molecules by incorporating dispersion forces and screened electrostatics which give rise to gas–liquid phase transitions and coexistence. More specifically, we set the interaction potential (bead–bead potential) between two coarse-grained sites on two distinct molecules to:

$$U_{\text{tot}}(r_{ij}) = U_{\text{rep}}(r_{ij}) + U_{\text{vdw}}(r_{ij}) + U_{\text{el}}(r_{ij}) + U_{\text{cont}}(r_{ij}) \quad (1)$$

where indices i and j refer to the i th and j th bead on the first and the second molecule (respectively) and r_{ij} represents the center-to-center distance of these beads. The individual contributions are as follows.

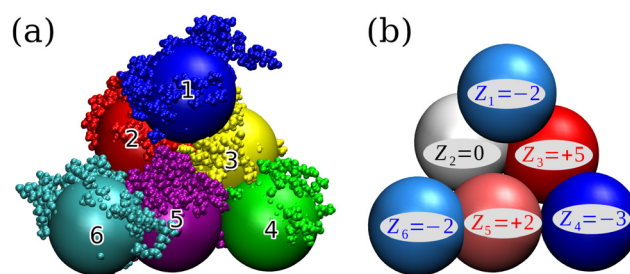


Fig. 1 (a) Atomistic structure of BSA²⁹ superimposed on the domain-based, six-bead coarse-grained model of ref. 32 with bead diameter $\sigma = 29.52$ Å. Missing hydrogens have been added with the pdb2gmx³³ tool. (b) Resulting coarse-grained charge distribution in units of the electron charge e of BSA at the isoelectric point (pH \approx 5.8). Figures have been rendered using the VMD software.³⁴



Volume exclusions are modelled by the repulsive part of the Lennard-Jones potential U_{rep} , which is cut-off at its minimum:

$$U_{\text{rep}}(r_{ij}) = \begin{cases} 4\epsilon \left[(\sigma/r_{ij})^{12} - (\sigma/r_{ij})^6 \right] & r_{ij} \leq r_{\text{rep}}^{\text{cut}} \\ 0 & r_{ij} > r_{\text{rep}}^{\text{cut}} \end{cases}, \quad (2)$$

where ϵ is the energy scale. Dimensionless energies U^* and temperatures T^* can be defined by $U^* = U/\epsilon$ and $T^* = k_{\text{B}}T/\epsilon$, where k_{B} is the Boltzmann constant and T is the temperature. The asterisks are dropped immediately and for the remainder of this work, energies and temperatures are to be understood in their respective reduced units. The cutoff distance is $r_{\text{rep}}^{\text{cut}} = 2^{1/6}\sigma$ and the bead diameter $\sigma = 29.52 \text{ \AA}$ has been derived in ref. 32.

Dispersion interactions are modelled by the van der Waals potential U_{vdW} for two equally sized spheres with:³⁵

$$U_{\text{vdW}}(r_{ij}) = \begin{cases} 0 & r_{ij} < r_{\text{rep}}^{\text{cut}} \\ -\frac{A_{\text{H}}}{12} \left(\frac{\sigma^2}{r_{ij}^2 - \sigma^2} + \frac{\sigma^2}{r_{ij}^2} + 2 \ln \frac{r_{ij}^2 - \sigma^2}{r_{ij}^2} \right) & r_{\text{rep}}^{\text{cut}} \leq r_{ij} \leq r_{\text{vdW}}^{\text{cut}} \\ 0 & r_{ij} > r_{\text{vdW}}^{\text{cut}} \end{cases}, \quad (3)$$

where A_{H} is the Hamaker constant (in units of ϵ) and the cutoff distance has been set to $r_{\text{vdW}}^{\text{cut}} = 2.5\sigma$.

Electrostatic interactions are approximated *via* a screened Yukawa potential U_{el} :

$$U_{\text{el}}(r_{ij}) = \begin{cases} k_{\text{B}}T\lambda_{\text{B}}Z_iZ_j \left(\frac{\exp(\kappa\sigma/2)}{1 + \kappa\sigma/2} \right)^2 \frac{\exp(-\kappa r_{ij})}{r_{ij}} & r_{ij} \leq r_{\text{el}}^{\text{cut}} \\ 0 & r_{ij} > r_{\text{el}}^{\text{cut}} \end{cases}, \quad (4)$$

where λ_{B} is the Bjerrum length, Z_i and Z_j are the electric charges on bead i and j (respectively, in units of the elementary charge e) and κ^{-1} is the Debye screening length. In order to have a noticeable influence of charges, we choose a moderate screening length of $\kappa^{-1} = \sigma$, which corresponds to a monovalent salt concentration of $c \approx 11 \text{ mmol L}^{-1}$. Moreover, for simplicity, we ignore the temperature dependence of the Bjerrum length and set it to the textbook value for water at room temperature, $\lambda_{\text{B}} \approx 0.7 \text{ nm}$. Furthermore, $r_{\text{el}}^{\text{cut}} = r_{\text{vdW}}^{\text{cut}} = 2.5\sigma$. The charge distribution (*i.e.* Z_i in eqn (4)) of all coarse-grained molecules within simulations is identical and determined through processing the atomistic structure of BSA (PDB ID: 4F5S, reported in ref. 29) with the APBS webserver,³⁶ employing the PROPKA model³⁷ to calculate the titration states of amino acids at the isoelectric point at $\text{pH} \approx 5.8$ (experiments determine a slightly deviating value of $\text{pH} \approx 4.7$).³⁸ Subsequently, the partial charges of all amino acid residues belonging to the same coarse-grained site are summed up to calculate the total charge of the bead, see Fig. 1(b) for the resulting charge distribution. We deliberately chose our molecules to have an overall neutral net charge, since in experiments for various globular proteins in solution with multivalent ions, aggregation of molecules and liquid-liquid phase separation (LLPS) is observed around the

isoelectric point.^{39,40} Additionally, in order to artificially alter the charge anisotropy of the CG molecules, we introduce a charge multiplier (CM), which uniformly scales all bead charges up or down according to:

$$Z \rightarrow \text{CM}Z, \quad (5)$$

such that $\text{CM} = 0$ would result in molecules with no electric charges on all beads and $\text{CM} = 2$ would result in a model where all charges depicted in Fig. 1(b) are doubled.

The last term U_{cont} in eqn (1) compensates the usage of cutoffs by making the total potential continuous. It reads:

$$U_{\text{cont}}(r_{ij}) = \begin{cases} -U_{\text{rep}}(r_{\text{rep}}^{\text{cut}}) - U_{\text{vdW}}(r_{\text{vdW}}^{\text{cut}}) & r_{ij} < r_{\text{rep}}^{\text{cut}} \\ -U_{\text{el}}(r_{\text{el}}^{\text{cut}}) + U_{\text{vdW}}(r_{\text{rep}}^{\text{cut}}) & r_{\text{rep}}^{\text{cut}} \leq r_{ij} \leq r_{\text{vdW}}^{\text{cut}} \\ -U_{\text{vdW}}(r_{\text{vdW}}^{\text{cut}}) - U_{\text{el}}(r_{\text{el}}^{\text{cut}}) & r_{ij} > r_{\text{vdW}}^{\text{cut}} = r_{\text{el}}^{\text{cut}} \\ 0 & \end{cases}. \quad (6)$$

For illustration, the complete interaction potential specified by eqn (1) is depicted in Fig. 2 for some typical parameters.

In our simulations the CG molecules are rigid bodies without internal degrees of freedom. Relative bead positions are fixed, corresponding to the average structure present in coarse-grained BD simulations at infinite dilution, as described in ref. 32. In this study, we investigate different parameter compositions, *i.e.* the charge anisotropy is varied by choosing $\text{CM} \in [0, 1, 2]$ and it is either $A_{\text{H}} = 4$ or $T = 1$ fixed (while keeping either T or A_{H} variable). Overall, this yields six different systems for which phase behavior can be studied, see Table 1 for an overview. In the simulations of the $A_{\text{H}} = \text{const.}$ -systems a physical critical temperature is determined. In contrast, in the $T = \text{const.}$ -systems one determines the needed bead-bead attraction strength for criticality at ambient temperatures.

The value of $A_{\text{H}} = 4$ appears to be a reasonable estimate for the bead-bead attraction strength for BSA solutions with

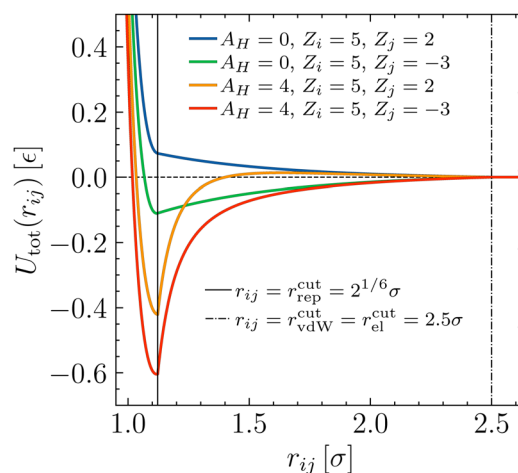


Fig. 2 Total bead-bead potential according to eqn (1) for typical parameters.



Table 1 Summary of investigated system parameters and results for the critical Hamaker constants $A_{H,c}$, temperatures T_c , densities ρ_c (in units of σ^{-3}) and B_2 values (in units of σ^3). Also shown are critical temperatures T_c^{iso} , densities ρ_c^{iso} and B_2 values of the effective, isotropic systems (see Section 4.2). We estimate the relative statistical error for T_c and $A_{H,c}$ to be $\pm 0.1\%$ and for $\rho_c \pm 0.4\%$. The relative total error of B_2 is $\pm 0.7\%$

System	CM	A_H	T	$A_{H,c}$	T_c	ρ_c	$A_{H,c}/k_B T_c$	$B_2(T_c)$	T_c^{iso}	ρ_c^{iso}	$B_2(T_c^{iso})$
I	0	4	Variable	4	0.405	0.059	9.88	-28.9	1.12	0.18	-6.41
II	0	Variable	1	8.59	1	0.061	8.59	-27.5	2.95	0.21	-5.63
III	1	4	Variable	4	0.411	0.059	9.74	-28.6	1.02	0.17	-7.29
IV	1	Variable	1	8.46	1	0.061	8.46	-27.4	2.61	0.19	-6.55
V	2	4	Variable	4	0.470	0.054	8.51	-35.6	1.00	0.12	-10.7
VI	2	Variable	1	7.44	1	0.054	7.44	-35.7	2.25	0.13	-9.9

monovalent salt, since an experimental structure factor at $pH \approx 7.0$, $c_{BSA} = 174 \text{ mg mL}^{-1}$ and $c_{NaCl} = 18 \text{ mmol L}^{-1}$ at room temperature⁴¹ is roughly similar to the center-of-mass structure factor from our simulations at equivalent conditions (the methodology for obtaining experimental structure factors is described in ref. 28). On the other hand in the experiments with BSA solutions with trivalent salts the attraction strength between molecules is a tunable parameter at constant temperature,²⁵ in our model this would be equivalent to a variable bead-bead attraction strength A_H and a variable CM.

3 Simulation methods

The CG molecules were simulated in a grand canonical Monte Carlo scheme with textbook⁴² insertion, deletion and roto-translation moves (with 5%, 5% and 90% attempt probability, respectively). Orientations of molecules were represented *via* quaternions.⁴³ The simulations were performed inside a cubic simulation box of side length $L = 10\sigma$. The main goal of our calculations is to obtain the critical points of the BSA systems. We additionally compute some binodal points to estimate the shape of the phase diagram. In this respect a finite size study was not conducted, since the expected finite size effects on the critical points are estimated not to be significant for the purpose of the present work.⁴² However, we expect that our binodal points are affected by finite size effects to some extent, especially those close to the critical points.

For obtaining the binodal points in the finite system close to T_c , histograms of the instantaneous molecule number N in the simulation box at different temperatures and values for the chemical potential μ were recorded and scanned for the appearance of double peaks. At the coexistence value for the chemical potential the area below these peaks must be equal. In order to obtain equal areas a fitting and optimization process using histogram reweighting^{44–46} was employed. Reweighting towards equal area gave new values for the chemical potential at which new histograms were simulated, whose peaks are possibly modified compared to the reweighted histogram. The procedure was repeated until the area difference below the peaks was below a chosen limit. The final peak positions give the average number of molecules in the coexisting gas and liquid states from which the corresponding number densities $\rho = N/L^3$ are determined (binodal points).

Critical points of the different BSA systems (listed in Table 1) were determined by matching their double-peak histograms to

the universal Ising distribution, following the methodology of ref. 42 and 47. In short, near the critical temperature and the critical chemical potential, where the histograms already looked similar to the universal Ising distribution, more simulations with slightly adjusted μ and T were performed. Second, by performing histogram reweighting on this data for new values μ' , T' and the mixing parameter s (which connects the system energy E and N to the Ising-like order parameter $X = N - sE$) as fit parameters, we determine the critical temperature T_c as the value of T' which gives the best fit result on the universal Ising distribution. In systems where the critical Hamaker constant was determined, the simulation and reweighting process was performed with different values of A_H (estimated by searching for equal area in the histograms) until the fitted value for the critical temperature T_c was very close to unity and then $A_{H,c}$ was computed as $A_{H,c} = A_H/T_c$ (the correction due to $T_c \neq 1$ was always below 0.2%). Fig. 3 shows fits on the universal Ising curve for the six different parameter combinations. The resulting critical densities ρ_c , temperatures T_c and Hamaker constants $A_{H,c}$ are listed in Table 1. Bead-bead correlations (see Section 4.4) were recorded in a larger box with the same GCMC code but with insertion and deletion moves disabled, which corresponds to a canonical NVT ensemble.

For obtaining the critical points 50 to 60 independent systems have been simulated, each of them with 10^9 GCMC moves in total. For the binodal points and the bead-bead correlations 10 resp. 80 independent systems have been employed, with 1×10^9 to 2×10^9 and 10^9 GCMC moves, respectively. We also

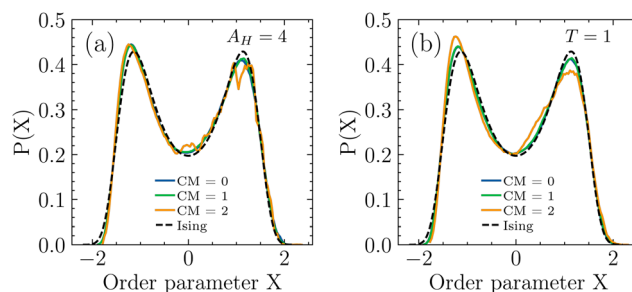


Fig. 3 Fits of the order parameter X on the universal Ising curve for all six different subsystems. X connects particle number and energy through a mixing parameter s as $X = N - sE$ and was scaled on the Ising curve (zero mean, unitary variance, cf. ref. 42 and 47). (a) shows fits for systems with constant $A_H = 4$ (with the critical temperature T_c as fit parameter), and (b) shows fits for the systems with a constant temperature $T = 1$ (with the critical value of the Hamaker constant $A_{H,c}$ as fit parameter).



note that obtaining binodal points proved difficult, especially for low temperatures, due to systems remaining in one phase even for long simulation runs. To this end, more elaborate simulation techniques could be applied, *e.g.* successive umbrella sampling,⁴⁸ which was out of the scope of this study.

4 Results

4.1 Phase diagrams and critical points

Our results for the binodals of the three differently charged systems with CM = 0, 1, 2 and $A_H = 4$, as well as the three systems with $T = 1$ (see Table 1) are depicted in Fig. 4.

In general, the resulting phase diagrams obtained for systems with either variable T or variable A_H are in shape very similar to each other, but not exactly identical, since even for the completely uncharged systems the Hamaker constant with our definition is not just an inverse temperature. Starting from the uncharged molecules (CM = 0) and adding the native charges (CM = 1) to our molecules introduces additional attractive interactions to the systems and shifts T_c upwards by (only) $\approx 1.5\%$ ($A_{H,c}$ downwards by $\approx 1.5\%$), and also inflicts minute changes to the slope of the coexistence curves. However, more pronounced effects are seen if one changes CM from 1 to 2, as this shifts T_c upwards by another $\approx 14\%$ and $A_{H,c}$ downwards by $\approx 12\%$. We attribute these additional attractions to the growing, leading order, electrical dipole moment of our molecules, and note that this physical quantity has been observed to play an important parameter for loci of coexistence regions in protein-like fluids.⁴⁹ Overall, our findings indicate that the stability in our BSA systems for CM = 1 is mainly governed by the van der Waals interactions and not so much by electrostatics.

From the results in Table 1 we infer that the effective temperature-like variable $k_B T/A_H$ is not even approximately the same at the critical point for the different systems ($A_H = \text{const.}$ and $T = \text{const.}$) at the same CM. The two types of systems cannot be mapped onto each other due to the soft repulsions. However the dependence of $k_B T/A_H$ on CM is similar, and indeed it turns out that only CM is a relevant variable (see Section 4.3).

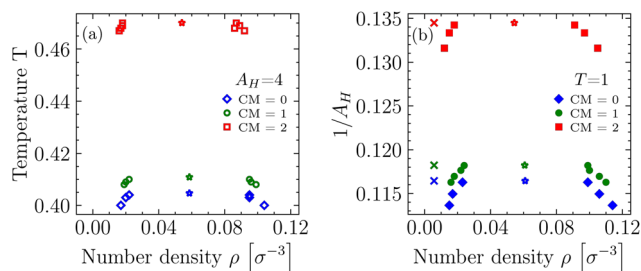


Fig. 4 Binodals of differently charged systems (see Table 1) with fixed Hamaker constant $A_H = 4$ (a) and fixed temperature $T = 1$ (b). Critical points are indicated by colored stars and state points where bead–bead correlation functions $g_{ij}(r_{ij})$ have been recorded are indicated by colored crosses (see Fig. 8 in Section 4.4).

4.2 Effective potentials at the critical points

An isotropic effective potential $U_{\text{eff}}(r)$ between two molecules can be defined in the usual way by the angular average:⁵⁰

$$\exp(-\beta U_{\text{eff}}(r)) = \int d\Omega_1 \int d\Omega_2 \exp(-\beta U(r, \Omega_1, \Omega_2)), \quad (7)$$

where r is the center-of-mass distance between two coarse-grained molecules, Ω_i specifies the orientation of molecule i (*e.g.* by three Euler angles), $U(r, \Omega_1, \Omega_2)$ is the total orientation-dependent pair potential between two molecules (in our case

$U = \sum_i \sum_j U_{\text{tot}}$) and the integral over orientations is normalized ($\int d\Omega_i = 1$). If U does not depend on orientations, then $U = U_{\text{eff}}$ of course. We calculate the six-dimensional integral *via* two methods: first, we employ Monte Carlo sampling with 10^8 uniformly distributed orientations per distance r with a spatial resolution of $\Delta r = 0.01\sigma$. Additionally, we compute eqn (7) *via* Gauss–Legendre quadrature and 20 grid points per angle. The obtained U_{eff} show negligible differences, the resulting B_2 values differ on average by $\approx 0.22\%$ for the two methods. For the Gauss–Legendre method using 15 grid points instead of 20 yields a relative error of obtained B_2 values of $\approx 0.5\%$.

Through examination of molecule–molecule configurations in eqn (7) the lowest-energy state of two molecules can be determined: for the CM = 0, 1 systems this is a very compact state at $r = 1.2$, which maximizes bead contacts. For CM = 2 it is a staggered state at $r = 1.4$, in which one molecule sits on top of the other but slightly rotated, such that bead 4 is near bead 3. These states, however, do not contribute to U_{eff} in a noticeable manner. Consistent with that is the absence of a noticeable amount of compact dimers in simulation snapshots. In these, dimers are on average further apart with varying bonding configurations.

We utilize the obtained effective potentials in additional GCMC simulations, in which isotropic particles interact *via* the six critical $U_{\text{eff}}(r)$ obtained by the previously described procedure. Subsequently, *via* the same protocol as described in Section 3 we determine critical temperatures T_c^{iso} for these systems and the obtained values are listed in Table 1. In the following sections of this paper these systems are referred to as “isotropic systems”. The obtained T_c^{iso} are approximately two to threefold higher as the corresponding T_c of the anisotropic models, pointing to a significant over-representation of attractions and/or under-representation of repulsions in $U_{\text{eff}}(r)$. An example are higher-order multi-body interactions not being present in $U_{\text{eff}}(r)$: it could be favorable for two CG molecules to form a low-energy dimer, which enters the effective potential. In the isotropic system three spheres can easily arrange to a trimer, with a threefold binding energy of the dimer, whereas three CG molecules can not easily arrange to a trimer with an equally low binding energy due to steric incompatibilities.



Table 2 Summary of critical parameters and effective molecule diameters σ_{eff} (in units of σ) at criticality calculated via the different methods (i)–(iv) presented in Section 4.3 for the anisotropic and isotropic systems. We estimate the total relative error for σ_{eff} and B_2^* to be $\pm 0.7\%$ and $\pm 2.6\%$ (respectively) for methods (i) and (ii), $\pm 0.9\%$ and $\pm 3.4\%$ for method (iii) and $\pm 0.4\%$ and $\pm 1.7\%$ for method (iv)

System	CM	T_c	Anisotropic systems										Isotropic systems			
			$A_{H,c}$	(i): $\sigma_{\text{eff}}^{\text{NF}}$		(ii): $\sigma_{\text{eff}}^{\text{rep}}$		(iii): $\sigma_{\text{eff}}^{\text{dry}}$			(iv): $\sigma_{\text{eff}}^{\text{cp}}$		T_c^{iso}	β'_c/β_c	(i): $\sigma_{\text{eff}}^{\text{NF,iso}}$	
				B_2^*	$\sigma_{\text{eff}}^{\text{NF}}$	B_2^*	$\sigma_{\text{eff}}^{\text{rep}}$	B_2^*	$\sigma_{\text{eff}}^{\text{dry}}$	$\sigma_{\text{eff}}^{\text{bead}}$	B_2^*	$\sigma_{\text{eff}}^{\text{cp}}$			B_2^*	$\sigma_{\text{eff}}^{\text{NF,iso}}$
I	0	0.405	4	-5.10	1.39	-1.41	2.14	-2.01	1.90	1.05	-1.70	2.01	1.12	0.36	-1.66	1.23
II	0	1	8.59	-5.09	1.37	-1.41	2.10	-2.09	1.85	1.02	-1.69	1.98	2.95	0.34	-1.61	1.19
III	1	0.411	4	-4.57	1.44	-1.40	2.14	-2.00	1.90	1.04	-1.69	2.01	1.02	0.40	-1.67	1.28
IV	1	1	8.46	-4.54	1.42	-1.41	2.10	-2.08	1.85	1.02	-1.67	1.98	2.61	0.38	-1.62	1.24
V	2	0.470	4	-4.36	1.57	-1.75	2.13	-2.51	1.89	1.04	-1.93	2.06	1.00	0.47	-1.76	1.43
VI	2	1	7.44	-4.50	1.56	-1.83	2.10	-2.71	1.85	1.02	-1.95	2.06	2.25	0.44	-1.73	1.40

4.3 Second virial coefficients and effective molecule diameters at the critical points

Using the effective potential $U_{\text{eff}}(r)$ of eqn (7), the second virial coefficient can be written as:

$$B_2 = 2\pi \int_0^\infty (1 - \exp(-\beta U_{\text{eff}}(r))) r^2 dr. \quad (8)$$

The obtained values at the respective critical points of all models are depicted in Fig. 6(a) and listed in Table 1. Increasing CM from 1 to 2 introduces additional attractions between molecules, as B_2 noticeably decreases, whereas almost no influence is visible if comparing the uncharged molecule to the native CM = 1 state. Within anisotropic systems sharing the same CM no significant difference is visible. The values for the isotropic systems on the other hand show a stronger charge dependence and values of systems with CM = 0, 1 are more separated to systems with CM = 2. Note that the reason for the B_2 values of the isotropic systems being higher than for the anisotropic ones, is that eqn (8) is evaluated at $T = T_c^{\text{iso}}$ and that $T_c^{\text{iso}}/T_c \approx 2 \dots 3$, as pointed out before.

For obtaining reduced second virial coefficients $B_2^* = B_{2,\text{crit}}/B_{2,\text{ex}}$ it is necessary to determine an exclusion volume which represents steric incompatibilities, *i.e.* an effective spherical diameter σ_{eff} of the molecule in terms of which $B_{2,\text{ex}} = (2\pi/3)\sigma_{\text{eff}}^3$. Here, we present four approaches with a certain theoretical preference for the first one and three additional methods for comparison, which however also appear natural from a physical point of view.

(i) $\sigma_{\text{eff}}^{\text{NF}}$ and $\sigma_{\text{eff}}^{\text{NF,iso}}$. In the original Noro–Frenkel (NF) analysis of various short-range attractive, isotropic systems⁸ $\sigma_{\text{eff}}^{\text{NF}}$ is computed from the repulsive part ($v_{\text{rep}}(r)$) of the isotropic pair potential (determined by the WCA rule, with the pair potential $v(r)$ cut-off at the minimum and shifted such that the minimum is zero) via the Barker–Henderson (BH) rule $\sigma_{\text{eff}}^{\text{NF}} = \int_0^\infty dr (1 - \exp(-\beta v_{\text{rep}}(r)))$.⁵¹ In our case, we use the effective potential for the anisotropic systems, *i.e.* $v(r) \rightarrow U_{\text{eff}}(r)$. For the isotropic systems $\sigma_{\text{eff}}^{\text{NF,iso}}$ is calculated exactly the same way, only with $T = T_c^{\text{iso}}$ (see Fig. 6(b) and Table 2 for resulting values).

(ii) $\sigma_{\text{eff}}^{\text{rep}}$. Excluded volume results from the repulsive part of the bead–bead potentials. First we define a bead–bead potential only from repulsive interactions, $U_{\text{tot}}(r_{ij}) = U_{\text{rep}}(r_{ij})$, as defined in eqn (2). Then U_{eff} is computed from eqn (7). The resulting

repulsive U_{eff} is charge-independent (if evaluated at the same T) and it is shown in comparison to the full U_{eff} in Fig. 5 (black, dashed line). Like before, from the BH rule the diameter $\sigma_{\text{eff}}^{\text{rep}}$ of an equivalent hard sphere system can be determined. The values $\sigma_{\text{eff}}^{\text{rep}} \approx 2.1\sigma$ (see Fig. 6(b) and Table 2) are much larger than $\sigma_{\text{eff}}^{\text{NF}}$ from the effective potential, this is also visible in the comparison of the effective potentials in Fig. 5.

(iii) $\sigma_{\text{eff}}^{\text{dry}}$. In ref. 32 the bead diameter $\sigma = 29.52 \text{ \AA}$ was derived by distributing the molecule volume of BSA on six equally sized spheres. The molecule volume was obtained by computing the protein specific volume with the molecular mass of BSA. Here, we conduct the calculation in reverse: first, an effective bead diameter $\sigma_{\text{eff}}^{\text{bead}}$ is determined by applying the BH rule to the repulsive part of the bead–bead potential $U_{\text{rep}}(r_{ij})$ (eqn (2)), its value is always slightly above 1.0σ (see Table 2). Then, the 6 beads of the model constitute the occupied, “dry” volume $6 \cdot (\pi/6)(\sigma_{\text{eff}}^{\text{bead}})^3$ of the protein without hydration shell, which is approximately the BSA volume in protein powder form (and is

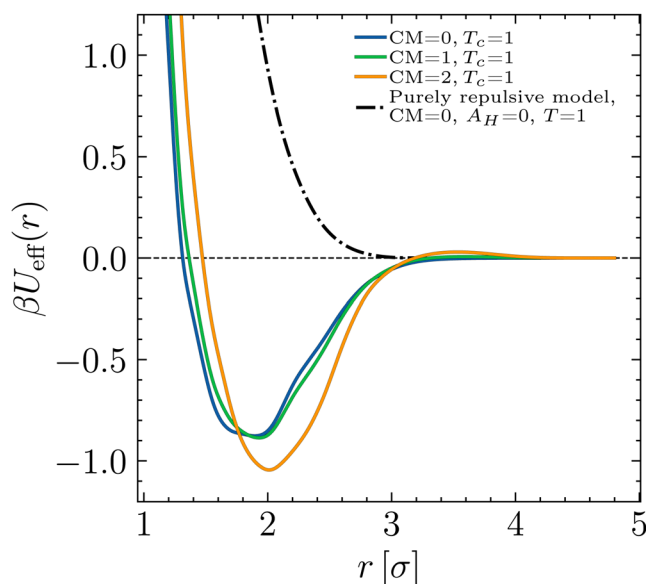


Fig. 5 Effective potential $\beta U_{\text{eff}}(r)$ (eqn (7)) obtained via MC integration for the critical systems with $T_c = 1$ (colored) and for a purely repulsive model for which $A_H = \text{CM} = 0$ and $T = 1$ (black, dashed line). Effective potentials of the critical system with $A_{H,c} = 4$ (not shown) are very similar to the ones depicted with the same CM.



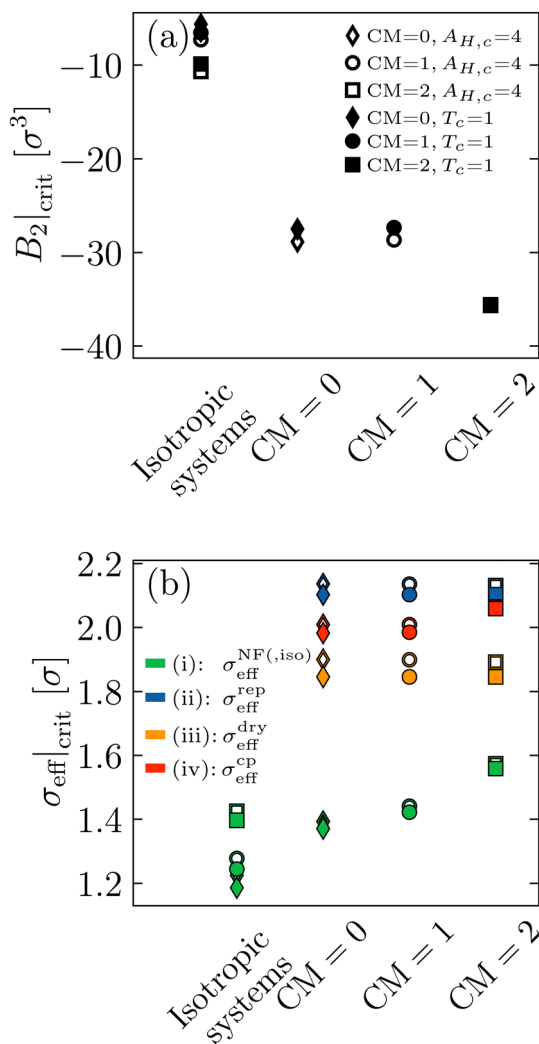


Fig. 6 (a) B_2 values according to eqn (8) of all critical BSA models presented in Table 1, together with B_2 values of the corresponding critical isotropic systems. (b) Effective molecule diameters σ_{eff} at the critical points, calculated via the four different approaches (i)–(iv) discussed in the main text.

nearly identical the molecule volume calculated in ref. 32). The dry volume can be converted to an effective sphere diameter $\sigma_{\text{eff}}^{\text{dry}}$ of an equivalent sphere with the same volume, *i.e.* $\sigma_{\text{eff}}^{\text{dry}} = 6^{1/3} \sigma_{\text{eff}}^{\text{bead}}$. Hence, in contrast to the previous methods no $U_{\text{eff}}(r)$ is computed here.

(iv) $\sigma_{\text{eff}}^{\text{ep}}$. Having calculated the critical molecule number densities ρ_c of our systems (see Table 1), we can define an effective volume packing fraction ϕ_{eff}^c at the critical point *via* $\phi_{\text{eff}}^c = (\pi/6)(\sigma_{\text{eff}}^{\text{ep}})^3 \rho_c$. Subsequently, we equate $\phi_{\text{eff}}^c = \phi_{\text{SW}}^c$, where $\phi_{\text{SW}}^c = 0.249$ is the well defined volume packing fraction at the critical point of a short-range, hard sphere square well fluid (taken from ref. 9, Table 2 in that reference and the system with reduced well width $\delta = 0.05$), which is approximately representative of the short-range potentials to which the Noro–Frenkel rule applies. This analysis yields $\sigma_{\text{eff}}^{\text{ep}} \approx 1.98 \dots 2.06\sigma$, see Fig. 6(b) and Table 2. Furthermore, we calculate the reduced second virial coefficients for the binodal points presented in

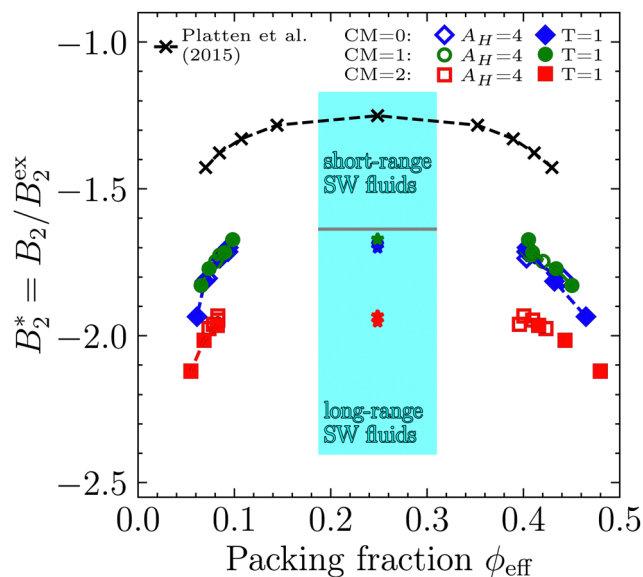


Fig. 7 Reduced second virial coefficient $B_2^* = B_2/B_{2,\text{ex}}$ with $B_{2,\text{ex}} = (2\pi/3)(\sigma_{\text{eff}}^{\text{ep}})^3$ as a function of the volume packing fraction $\phi = \phi_{\text{eff}} = (\pi/6)(\sigma_{\text{eff}}^{\text{ep}})^3 \rho$ for the six different parameter combinations. The cyan shaded area represents the range of critical B_2^* values for short-range ($\delta \leq 0.25$) and long-range ($\delta \geq 0.25$) hard sphere square well fluids calculated by Gazzillo *et al.*⁵² Black crosses represent the reference hard sphere square well binodal taken from Platten *et al.*⁹ (Fig. 2 and Table 2 in that reference), which was used to obtain $\sigma_{\text{eff}}^{\text{ep}}$ as described in Section 4.3 by matching the critical points. Dashed lines are guides to the eye.

Fig. 4, using $\phi_{\text{eff}} = (\pi/6)(\sigma_{\text{eff}}^{\text{ep}})^3 \rho$ and $B_{2,\text{ex}} = (2\pi/3)(\sigma_{\text{eff}}^{\text{ep}})^3$. The resulting phase diagrams in the $\phi_{\text{eff}} - B_2^*$ plane are shown in Fig. 7, along with the phase diagram of the aforementioned, short-range, reference square well system (black crosses). The packing fractions of our binodals approximately match the binodals of the reference, which is in effect similar to the construction⁹ used by the group of Stefan Egelhaaf mentioned in the introduction and further discussed in Section 5.4. Note that if a longer ranged hard sphere square well fluid with $\delta \approx 0.25$ would be used as reference, then $\phi_{\text{SW}}^c \approx 0.2$,⁹ and therefore $\sigma_{\text{eff}}^{\text{ep}}$ would decrease by approx. 7%.

Fig. 7 also shows $B_2/B_{2,\text{ex}}$ -ranges in which short-range (SR, $\delta \leq 0.25$) and long-range (LR, $\delta \geq 0.25$) hard sphere square well systems are critical (cyan shaded area).⁵² This total range is nearly identical to the range of the isotropic Noro–Frenkel fluids depicted in Fig. 9(a) (also cyan shaded area), but allows for a distinction between different ranges of attraction.

All previously discussed effective diameters are listed in Table 2 and shown in Fig. 6(b). Evidently, method (i) yields the smallest σ_{eff} compared to the other methods, whereas methods (ii)–(iv) yield higher values, which are generally relatively close to each other for all investigated systems. Systems with constant T roughly yield identical results as systems with constant A_H for the same CM. Adding charges to the molecules significantly increases their effective diameter only for method (i), as the increased electrostatic repulsion at short distances is directly present in U_{eff} , whereas methods (ii) and (iii) are (for the systems with variable temperature) only indirectly sensitive



to the parameter CM, *i.e.* solely *via* the shifted inverse temperature β_c which enters the BH rule. The latter is also responsible for all $\sigma_{\text{eff}}^{\text{NF,iso}}$ being smaller than their respective $\sigma_{\text{eff}}^{\text{NF}}$, as $T_c^{\text{iso}}/T_c \approx 2 \dots 3$.

Note that methods (i) and (ii) are in spirit similar to each other, as they derive the effective diameter only from the two-body interaction and the pair potential we attribute to the model. Method (iii) is based on an experimentally measured quantity, while method (iv) is based on the critical point, a collective property.

4.4 Bead-bead correlations

In order to evaluate how charge-anisotropy influences the preferred bead contacts in our simulations we compute bead-bead correlation functions g_{ij} in Fig. 8 (note that $g_{ij} = g_{ji}$) for our coarse-grained molecules. We choose a relatively low density $\rho = 5.75 \times 10^{-3} \sigma^{-3}$ outside the coexistence region in the gas region of the phase diagram (colored crosses in Fig. 4(b)) to mitigate packing effects, which would complicate a separation with effects stemming from possible patchy interactions.

Starting from the uncharged molecule state CM = 0 one notices first-peak contacts are most pronounced for combinations (i, j) of exposed beads at the edges of the molecule, *i.e.* beads 1, 4 and 6 and least pronounced for beads that are closer to the center of the molecule, *i.e.* beads 2, 3 and 5 (compare to Fig. 1). Adding the native charges to the BSA models (CM = 1) does not introduce new (i, j) combinations to the top first-peak bead-bead correlations, indicating that for this system the van der Waals and not the electrostatic interaction determines association of the molecules. This is additionally supported by the observation that only bead combinations with like charge are within the top correlation functions. However, increasing CM to 2 finally switches the dominant interaction mechanism from van der Waals to

electrostatics, as now the strongest charged beads exhibit either the highest first-peak correlations, *i.e.* (3, 6) and (3, 4) if oppositely charged, or the lowest first-peak correlations, *i.e.* (3, 3) and (3, 5) if like charged. Intriguingly, despite the electrostatic dominance the bead-bead combinations (4, 4) and (1, 4) favoured by the molecule's geometry combined with the dispersion force still exhibit the third and fourth highest correlation functions. We thus conclude that for our model there is a crossing-over between CM = 1 and CM = 2, where the mechanism of preferred contact selection between beads switches from van der Waals to electrostatic interactions, where the former's influence still persist to some extent. Additionally, we infer that this analysis provides arguments that our BSA model effectively exhibits three patches (within a KF picture), due to the three sites at the edges of the CG molecule (beads 1, 4 and 6), that show preferred contact propensity in g_{ij} .

5 Discussion: effective patchiness from critical points

5.1 Reduced second virial coefficient B_2^* *via* method (i)

In their seminal paper,⁸ Noro and Frenkel have shown that systems with isotropic, short-range attractive pair potentials can be mapped onto hard sphere square well systems (with the same B_2 and the effective hard sphere diameter $\sigma_{\text{eff}}^{\text{NF}}$ determined with the Barker-Henderson rule), with the effect that the reduced virial coefficient B_2^* at the critical temperature is approximately constant over quite a range of pair potentials and given by $B_2^* = B_{\text{NF}}^* \approx -1.5$. The Noro-Frenkel rule does not hold for patchy particles, though. Foffi and Sciortino¹⁰ have shown that for Kern-Frenkel (KF) particles⁵³ (hard spheres with attractive surface patches defined *via* narrow cones with opening angle θ_0 and with reduced attraction range δ_{KF}) with $M = 3$,

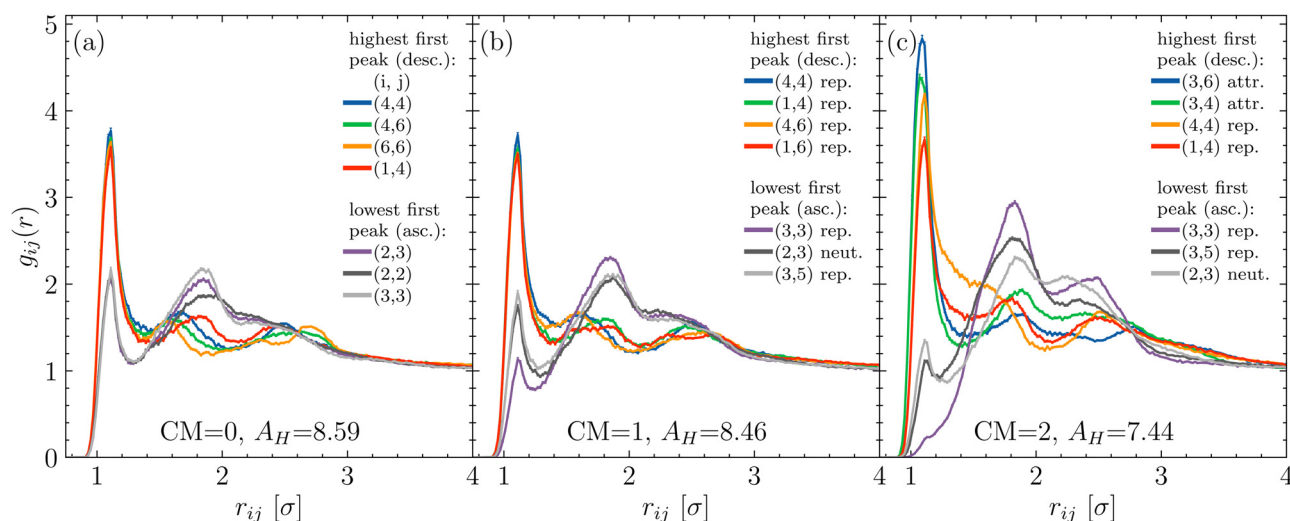


Fig. 8 Bead-bead correlation functions $g_{ij}(r_{ij})$ with r_{ij} representing the center-to-center distance of bead i to bead j , for systems with constant temperature $T = 1$ and variable charge ((a): CM = 0, (b): CM = 1, (c): CM = 2) at a dilute molecule number density $\rho = 5.75 \times 10^{-3} \sigma^{-3}$ (for the locations in the phase diagram see the colored crosses in Fig. 4(b)). Only selected (i, j) combinations that exhibit the highest and lowest contact peak values are shown. Attractive, repulsive and neutral charge configurations are indicated in the legend. Error bars show the standard deviation of the mean.



4, 5 patches B_2^* is approximately a function of M only, but the actual values are very different from B_{NF}^* (≈ -30 for $M = 3$, $\approx -4 \dots -5$ for $M = 4$ and $\approx -2 \dots -3$ for $M = 5$).

In the KF model, let ε_{KF} be the attractive interaction strength of two interacting KF patches. The corresponding effective potential is a square well potential with well depth ε' , which is related to ε_{KF} by:

$$\exp(\beta\varepsilon') = \chi^2(\exp(\beta\varepsilon_{\text{KF}}) - 1) + 1, \quad (9)$$

where $\chi = (M/2)(1 - \cos \theta_0) < 1$. It is interesting to ask for the relation of critical temperatures in the orientation-dependent KF system vs. in the corresponding effective square well system. For the inverse critical temperature $\beta_{\text{c,KF}}$ of KF we use the data taken from ref. 10. The inverse critical temperature β'_c of the effective square well system can be safely estimated using the Noro–Frenkel rule and the reduced B_2 in the square well system, i.e.:

$$\begin{aligned} B_2/B_{2,\text{ex}}|_{\text{crit}} &= \left(1 - \exp(\beta'_c\varepsilon')\right) \left((1 + \delta)^3 - 1\right) + 1 \\ &\approx B_{\text{NF}}^*, \end{aligned} \quad (10)$$

where δ is the range of the square well attraction in units of the hard sphere diameter. Here, ε' is determined from ε_{KF} from eqn (9), with $\beta = \beta_{\text{c,KF}}$ (the inverse critical temperature of the KF system) and we set $\delta = \delta_{\text{KF}}$. From the data in ref. 10 for various KF systems we find a ratio $\beta'_c/\beta_{\text{c,KF}} \approx 0.46 \dots 0.56$ ($M = 3$) and $\beta'_c/\beta_{\text{c,KF}} \approx 0.69 \dots 0.87$ ($M = 4$) and $\beta'_c/\beta_{\text{c,KF}} \approx 0.79 \dots 0.94$ ($M = 5$). It is clear that the critical temperature of the effective system (particles interacting with U_{eff}) is always noticeably higher than the one of the KF system, but the ratio is not only a function of the number M of patches. There is a clear

separation between the values for all $M = 3$ systems from the rest, but there is a considerable overlap in the $\beta'_c/\beta_{\text{c,KF}}$ -ranges for $M = 4$ and 5, see Fig. 9(b). All values are, however, significantly smaller than 1 which is the limit of isotropic potentials.

Note that in the KF analysis above the anisotropy only arose from the patchy attractions, while repulsions are isotropic hard sphere in both the KF and the effective system. This is different in our system where excluded volume interactions are very anisotropic and the effective isotropic repulsions are very soft.

These considerations can be applied to our model since U_{eff} and the corresponding critical temperatures in the effective systems have been determined in Section 4.2, we refer to the results for $\sigma_{\text{eff}}^{\text{NF,iso}}$ and B_2^* in Fig. 6(b) and 9(a) (respectively) and also in Table 2. For all investigated isotropic systems, $B_2^* \approx -1.6 \dots -1.7$ which is different from $B_{\text{NF}}^* \approx -1.5$. This is due to the rather long range in the attractions of U_{eff} and matches very well results in Noro and Frenkel's work⁸ for wide square wells: B_2^* changes from -1.65 to -2.26 if the reduced radial range δ of the attractive square well changes from 0.5 to 1. The effective diameters $\sigma_{\text{eff}}^{\text{NF,iso}} \approx 1.2 \dots 1.4\sigma$ for all systems (see Fig. 6(b)), and it is interesting to note that these diameters are only $\approx 20 \dots 40\%$ larger than the bead diameter σ .

For the assessment of “patchiness”, the values of β'_c/β_c , as well as B_2^* and $\sigma_{\text{eff}}^{\text{NF}}$ at the critical temperature of the anisotropic system are relevant. As can be seen in Fig. 9(b) and Table 2, the values of β'_c/β_c are around 0.34...0.47 with a moderate spread and all values of B_2^* are between -4 and -5 with a likewise moderate spread. Regarding charge dependence, $\sigma_{\text{eff}}^{\text{NF}}$ exhibits a systematic increase for a rising CM value (see Fig. 6(b),

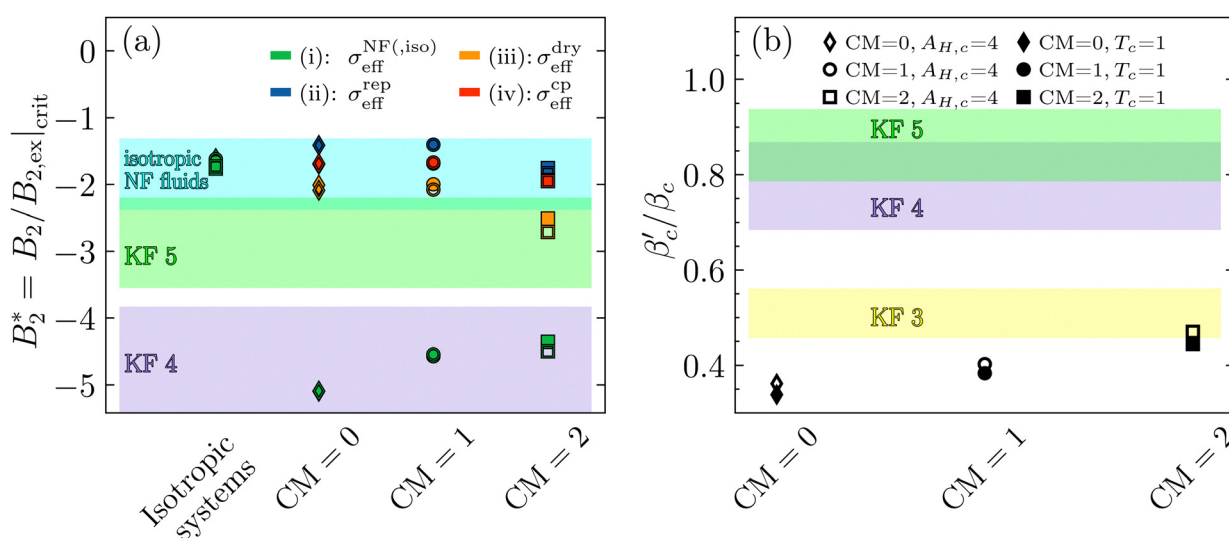


Fig. 9 (a) Reduced second virial coefficients $B_2^* = B_2/B_{2,\text{ex}}$, with $B_{2,\text{ex}} = 2\pi\sigma_{\text{eff}}^3/3$ and where σ_{eff} is taken as $\sigma_{\text{eff}}^{\text{NF,(iso)}}$, $\sigma_{\text{eff}}^{\text{rep}}$, $\sigma_{\text{eff}}^{\text{dry}}$ or $\sigma_{\text{eff}}^{\text{CP}}$ (see Section 4.3 and Fig. 6(b)) for the anisotropic, critical BSA systems together with B_2^* values of the corresponding critical isotropic systems, as well as regions in which isotropic fluids (Table 1 in ref. 8, cyan shaded area) and Kern–Frenkel particles with $M = 4, 5$ patches are critical (Table 1 in ref. 10, violet and green shaded areas). (b) Ratios β'_c/β_c for all critical BSA systems together with regions in which $\beta'_c/\beta_{\text{c,KF}}$ falls for different numbers of patches $M = 3, 4, 5$ in the Kern–Frenkel model (yellow, violet and green shaded areas). Latter regions have been calculated as explained in the main text with values taken from ref. 8 and 10.



corresponding to the observation of a larger repulsive core in U_{eff}). Hence, although the magnitude of B_2 is growing with increasing CM (see Fig. 6(a)), the charge dependence of $\sigma_{\text{eff}}^{\text{NF}}$ (and consequently also of $B_{2,\text{ex}}$) leads to the monotonic and moderate charge dependence in B_2^* (similar to β'_c/β_c).

Comparing this to the KF results, these values correspond to 4 patches (from B_2^*) or 3 patches for the CM = 2 systems (from β'_c/β_c), which points to a considerable effective anisotropy in the BSA model.

5.2 Alternative determinations of $B_{2,\text{ex}}$ via methods (ii)–(iv) and their influence on B_2^*

The normalization of B_2 with $B_{2,\text{ex}} = 2\pi\sigma_{\text{eff}}^3/3$, the second virial coefficient for the excluded volume interaction of the equivalent sphere system, leaves a certain ambiguity in the definition of $B_{2,\text{ex}}$ which becomes manifest in a system where the repulsive interactions are explicitly anisotropic. It is interesting to explore resulting B_2^* if the additional σ_{eff} of Section 4.3 are used.

The values for $\sigma_{\text{eff}}^{\text{rep}}$, $\sigma_{\text{eff}}^{\text{dry}}$ and $\sigma_{\text{eff}}^{\text{CP}}$ are around 1.9...2.1 σ (approximately charge-independent) and relatively close to each other, see also Fig. 6(b) and Table 2. All three effective diameters are considerably larger than $\sigma_{\text{eff}}^{\text{NF}}$ with the immediate consequence that B_2^* is reduced in magnitude and they all exhibit a stronger charge dependence than before: for method (ii) (with $\sigma_{\text{eff}} = \sigma_{\text{eff}}^{\text{rep}}$) we find $B_2^* \approx -1.4$ for the systems with CM = 0 and 1, and $B_2^* \approx -1.8$ for the systems with CM = 2. For method (iii) (with $\sigma_{\text{eff}} = \sigma_{\text{eff}}^{\text{dry}}$) the difference between CM = 0, 1 on the one side and CM = 2 on the other side can be seen as well: $B_2^* \approx -2.0 \dots -2.1$ (CM = 0, 1) and $\approx -2.5 \dots -2.7$ (CM = 2). Method (iv) yields approximately values that fall in the middle of (ii) and (iii): $B_2^* \approx -1.7$ (CM = 0, 1) and $B_2^* \approx -1.9$ (CM = 2). Compared with the values from the KF model, we would argue that according to methods (ii)–(iv) our BSA model is fairly isotropic for CM = 0, 1. However, the deviating values of B_2^* for the CM = 2 systems are intriguing. Here method (iii) yields results compatible with a 5 patch KF model.

The ambiguity in defining $B_{2,\text{ex}}$ drops out if we form the ratio of critical B_2 -values

$$r = \frac{B_2(\text{CM} = 2)}{B_2(\text{CM} = 0)} \quad (11)$$

which yields $\approx 1.24 \dots 1.29$. Using the results for different systems in Table 1 of ref. 10, we estimate for the KF model $r \approx 4.12 \dots 7.35$ for the ratio between a 3-patch and a 4-patch model, and $r \approx 1.09 \dots 2.56$ for the ratio between a 4-patch and a 5-patch model. The latter range includes our results, such that one could conclude that our BSA model is moderately patchy, around 4 patches for CM = 2 and 5 patches for CM = 0.

As a remark, we note that the term “patchiness” used in this study refers solely the kind present in the KF model, *i.e.* a purely attractive one. There are other kinds of interesting model systems without uniform van der Waals attraction exhibiting patches that repel each other, for example inverse patchy colloids⁵⁴ (IPCs), lattice charge models⁵⁵ or continuous

charge patchy models.⁵⁶ Such repulsive patches are obviously also present in our CG model, due to same-charge interactions, and they are overall less dominant than the opposite-charge interactions (otherwise T_c would not increase by increasing CM). By the choice of our KF/NF reference systems the effects of these two types of patches are mixed to an effective, attractive patchiness. Here, a systematic study that separates these effects would be intriguing, however, no law of corresponding states that would enable a investigation similar to this one has been observed for IPCs so far.⁵⁷

5.3 Critical density ρ_c

Also the critical density ρ_c is sensitive to the number of patches within the KF model, but in contrast to B_2^* the range of critical densities for long-range isotropic fluids overlaps with the corresponding range for KF fluids with 4 and 5 patches (see Fig. 10). The corresponding quantity in our model systems is the critical density (Table 1) evaluated in units of $1/\sigma_{\text{eff}}^3$, using the various definitions from Sections 5.1 and 5.2 (Fig. 6(b) and Table 2), see the symbols in Fig. 10. The analysis confirms the conclusions from Sections 5.1 and 5.2 from the analysis of B_2^* , especially regarding the difference between $\sigma_{\text{eff}}^{\text{NF}}$ on the one hand and the alternative definitions (ii)–(iv) from Section 5.2. If the latter are used, the critical densities for the CM = 0 systems are in the range of isotropic short-range square well fluids and decrease slightly with increasing the charge. Only for $\sigma_{\text{eff}}^{\text{dry}}$, the highly charged CM = 2 systems appear slightly patchy (5 KF patches), however, overlapping with the range of isotropic long-range square well fluids. Overall, this is comparable to the slight decrease of B_2^* with increasing charge seen in Fig. 9(a) for the alternative definitions of σ_{eff} . The relative decrease of ρ_c (CM = 2) *vs.* ρ_c (CM = 0) of about 10% roughly corresponds to

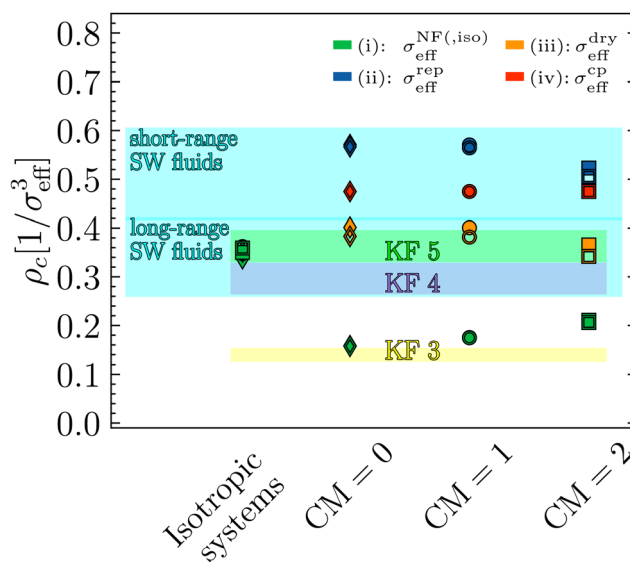


Fig. 10 Critical densities ρ_c (Table 1) in units of $1/\sigma_{\text{eff}}^3$, where σ_{eff} is taken as $\sigma_{\text{eff}}^{\text{NF(iso)}}$, $\sigma_{\text{eff}}^{\text{rep}}$, $\sigma_{\text{eff}}^{\text{dry}}$ or $\sigma_{\text{eff}}^{\text{CP}}$ (see Section 4.3 and Fig. 6(b)). The shaded areas correspond to the critical densities for KF particles reported in ref. 10 and to short- ($\delta \leq 0.25$) and long-range ($\delta \geq 0.25$) hard sphere square well systems reported in ref. 9.



the relative decrease in ρ_c of KF 5-patch to 4-patch models (and would constitute an estimate independent of σ_{eff}). If however method (i) is used for the anisotropic systems, the critical densities for the $\text{CM} = 0$ systems are compatible with a 3-patch KF system and the charges tend to increase the critical density towards the one of a 4-patch KF system. This effect of charge is similar to the increase of B_2^* for this definition of the effective diameter, see Fig. 9(a), the only difference being that according to the critical density analysis the CG model has an overall more pronounced patchy nature. The critical densities of the isotropic systems are in the range of isotropic long-range fluids (in agreement with the long-range character of U_{eff} in Fig. 5).

5.4 Relation to experiment

Scattering experiments on BSA solutions in monovalent salt indicated stability of the BSA solution.²⁸ Available experimental data for B_2 in sodium chloride solution at $\text{pH} = 7.4$ and ambient temperatures were fitted in ref. 23 using a 20-bead model with DLVO forces, similar to ours. The resulting B_2 values (in our unit σ^3) in the range $-5 \dots .25$ are considerably larger than the B_2 values at the critical point for our models (which are in the range $-28 \dots -37$, see Fig. 6(a)). Since the fit of ref. 23 was for high salt concentrations and therefore effective charge screening, one should compare to the maximal value for the chargeless model ($\text{CM} = 0$), $B_2 \approx -28$. Thus one would conclude that a hypothetical critical temperature in a sodium chloride solution is way below the freezing point of water which is outside of physical interest.

In contrast to sodium chloride solutions, liquid–liquid phase separations have been observed in BSA solutions with trivalent cations.²⁴ The necessary attraction has been proposed to be due to cation bridges at specific places on the BSA surface, *i.e.* patchy attractions in essence.²⁵ The phase separation occurs near the point of zero charge. Measurements using small angle X-ray scattering (SAXS) of $B_2/B_{2,\text{ex}}$ ^{40,58–60} have indicated though that $B_2^* \approx -1.5 \dots -1.8$, which is of the order of $B_{\text{NF}}^* \approx -1.5$, and thus one would ascribe no patchiness to BSA. However, in these works the effective hard sphere diameter σ_{eff} used in $B_{2,\text{ex}}$ was fitted from an ellipsoidal form factor and gave $\sigma_{\text{eff}} \approx 9.2$ nm, which is larger than our σ_{eff} by a factor $1.5 \dots 2.6$. This particular ellipsoidal form factor fit is probably due to clusters. Note that a similar form factor fit in the SAXS measurements of ref. 28 (with monovalent salt, less attraction and thus less propensity for cluster formation) gave $\sigma_{\text{eff}} \approx 6.7$ nm which is much closer to our $\sigma_{\text{eff}}^{\text{iso}} \approx 6.3$ nm which can be considered similar in spirit. Thus, using our range of σ_{eff} , B_2^* in the BSA-trivalent cation system is estimated to be around $-4.6 \dots -32$ which indicates much more patchiness and also a larger patchiness than found in our model (ranging between $-1.4 \dots -5.1$). Our model has not been designed with very specific bridging attractions, rather we have put the effect of the attraction into an increase of A_{H} and the attraction between positive and negative charges at the BSA surface. We would expect that B_2^* attains larger negative values if specific bridging sites are included in our model.

A somewhat different perspective on the validity of the Noro–Frenkel rule for protein solutions has been given in ref. 9, discussing the phase diagrams of globular proteins, mainly of lysozyme. The determination of a suitable effective diameter and thus of B_2^* uses the liquid–liquid binodal, which is in effect similar to method (iv) of Section 4.3. Specifically, the argument in ref. 9 concerning the validity of the Noro–Frenkel rule proceeds as follows: from an extended body of simulation data for binodals of isotropic fluids a “Noro–Frenkel band” of binodals is defined in the $\phi - B_2^*$ plane (packing fraction *vs.* reduced second virial coefficient). Considering now a protein solution, an effective diameter is defined by fitting the gas part of its binodal to the binodal of a square well fluid with reduced well width $\delta = 0.05$ (this binodal is located approximately in the middle of the “Noro–Frenkel band”). As a result, the investigated lysozyme binodals nicely fall into this band and $B_2^* \approx B_{\text{NF}}^*$. Thus, ref. 9 argues, that the Noro–Frenkel rule is valid and would presumably also hold for other globular proteins, *i.e.* from critical liquid–liquid demixing no patchiness can be inferred. As a result, a such defined diameter is considerably larger than values derived from an effective, isotropic DLVO potential (as also seen in ref. 9). This is in agreement with our findings for $\sigma_{\text{eff}}^{\text{CP}} \approx 2.0$ (whose derivation is in effect similar to the construction of ref. 9), which is larger than $\sigma_{\text{eff}}^{\text{NF(iso)}} \approx 1.2 \dots 1.6$ (from the orientation-averaged effective potential). Consequently, the critical points of all systems with varying CM are well within the region where isotropic fluids are critical, see Fig. 7. If using a finer distinction of $B_2^*(T_c)$ values, that discriminates short and long-range hard sphere square well systems, the $\text{CM} = 2$ system clearly corresponds to longer ranged wells ($\delta \geq 0.25$), whereas the $\text{CM} = 0, 1$ systems fall in an intermediate region where $\delta \approx 0.25$.

6 Conclusion

We have investigated the phase diagram and critical points of an anisotropic coarse-grained model for the globular protein BSA. Besides the shape anisotropy through the spatial arrangement of the beads, we investigated also the dependence of the critical point on the charge asymmetry, realized by different bead charges at the isoelectric point. We have attempted to quantify the anisotropy effects on the critical point through an “effective patchiness” which is built on the suggestion by Foffi and Sciortino¹⁰ for an extended Noro–Frenkel rule⁸ for patchy particles. A difficulty arises from the fact that in the analysis of Foffi and Sciortino the repulsive core of the patchy particles are isotropic hard spheres whereas in our case the repulsive core is soft and anisotropic.

We have argued that the analysis of the effective, orientation-averaged pair potential U_{eff} provides a useful link: first, it allows to define an effective hard sphere diameter $\sigma_{\text{eff}}^{\text{NF}}$ and a repulsive part $B_{2,\text{ex}}$ of the second virial coefficient and secondly, the ratio of critical temperatures of an isotropic model interacting with U_{eff} and of the fully anisotropic model contains a definite patch dependence. The reduced second virial coefficient at the critical



point $B_2^* = B_2/B_{2,\text{ex}}(T_c)$ points to an effective patchiness of 4 Kern-Frenkel patches ($B_2^* \approx -4 \dots -5$), whereas the ratio of critical temperatures points to a 3-patch model for $\text{CM} = 2$, see Fig. 9. Through computing the ratio r of B_2 (eqn (11)), the dependence on $B_{2,\text{ex}}$ drops out and for our BSA model we see that the $\text{CM} = 0$ and $\text{CM} = 2$ charges lead to an effective anisotropy of approximately 4 ($\text{CM} = 2$) to 5 ($\text{CM} = 0$) Kern-Frenkel patches.

Through its dependence on $B_{2,\text{ex}}$, the B_2^* -criterion does not appear to be unique. We have investigated two other sensible criteria for defining $B_{2,\text{ex}}$ and σ_{eff} directly related to the molecular properties (through the repulsive part of the bead-bead potentials and the dry protein volume, yielding $\sigma_{\text{eff}}^{\text{rep}}$ and $\sigma_{\text{eff}}^{\text{dry}}$) and obtain larger effective diameters and consequently also values $B_2^* \approx -2$ (and larger) which are mostly compatible with the Noro-Frenkel rule for isotropic particles, except for $\sigma_{\text{eff}} = \sigma_{\text{eff}}^{\text{dry}}$ and $\text{CM} = 2$, see Fig. 9(a).

The analysis presented here shows that care is needed in the definition of an effective hard sphere diameter when working with a B_2^* -criterion. This has become evident when (as a third alternative) applying an analysis similar to the one by Platten *et al.*⁹ on the phase behavior of lysozyme, yielding $\sigma_{\text{eff}}^{\text{cp}}$. Here an effective hard sphere diameter was defined through matching the critical points of the proteins to a square well system, which again resulted in an approximate validity of the Noro-Frenkel rule.

It appears difficult to straightforwardly apply rules for isotropic systems with reasonably well defined excluded volume to anisotropic systems. Our work is an example how much the results for B_2^* can vary when applying different methods which all appear reasonable. In this respect, computing the isotropic potential U_{eff} offers two advantages. First, the definition of a hard sphere diameter *via* Barker-Henderson integration (as applied for soft isotropic systems in ref. 8) and second the possibility of using the ratio of critical temperatures between the anisotropic and the effective isotropic system as a variable to infer patchiness.

Author contributions

Jens Weimar and Frank Hirschmann: data curation, formal analysis, investigation, methodology, software, visualization, writing – original draft, writing – review & editing. Martin Oettel: conceptualization, funding acquisition, methodology, project administration, resources, supervision, validation, writing – original draft, writing – review & editing.

Data availability

The data supporting our results is included in Table 2. The code and algorithms yielding our results are standard computational techniques (multi-dimensional integration, grand canonical Monte Carlo simulation and histogram reweighting) which are described in the references provided in our manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors acknowledge support by the state of Baden-Württemberg through bwHPC and thank Dr Fajun Zhang for insightful discussions on this manuscript and for providing experimental structure factors.

Notes and references

- 1 P.-H. Peng, K.-W. Hsu and K.-J. Wu, *Am. J. Cancer Res.*, 2021, **11**, 3766–3776.
- 2 A. Pande, J. Pande, N. Asherie, A. Lomakin, O. Ogun, J. King and G. B. Benedek, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 6116–6120.
- 3 X. Zhang, Y. Lin, N. A. Eschmann, H. Zhou, J. N. Rauch, I. Hernandez, E. Guzman, K. S. Kosik and S. Han, *PLoS Biol.*, 2017, **15**, 1–28.
- 4 P. R. ten Wolde and D. Frenkel, *Science*, 1997, **277**, 1975–1978.
- 5 F. Zhang, G. Zoicher, A. Sauter, T. Stehle and F. Schreiber, *J. Appl. Crystallogr.*, 2011, **44**, 755–762.
- 6 A. Stradner and P. Schurtenberger, *Soft Matter*, 2020, **16**, 307–323.
- 7 G. A. Vliegthart and H. N. W. Lekkerkerker, *J. Chem. Phys.*, 2000, **112**, 5364–5369.
- 8 M. G. Noro and D. Frenkel, *J. Chem. Phys.*, 2000, **113**, 2941–2944.
- 9 F. Platten, N. E. Valadez-Pérez, R. Castañeda-Priego and S. U. Egelhaaf, *J. Chem. Phys.*, 2015, **142**, 174905.
- 10 G. Foffi and F. Sciortino, *J. Phys. Chem. B*, 2007, **111**, 9702–9705.
- 11 H. Liu, S. K. Kumar, F. Sciortino and G. T. Evans, *J. Chem. Phys.*, 2009, **130**, 044902.
- 12 M. A. Blanco, *mAbs*, 2022, **14**, 2044744.
- 13 J. Jin, A. J. Pak, A. E. P. Durumeric, T. D. Loose and G. A. Voth, *J. Chem. Theory Comput.*, 2022, **18**, 5759–5791.
- 14 S. Kmiecik, D. Gront, M. Kolinski, L. Wieteska, A. E. Dawid and A. Kolinski, *Chem. Rev.*, 2016, **116**, 7898–7936.
- 15 S. J. Marrink, H. J. Risselada, S. Yefimov, D. P. Tieleman and A. H. de Vries, *J. Phys. Chem. B*, 2007, **111**, 7812–7824.
- 16 E. Fagerberg and M. Skepö, *J. Chem. Inf. Model.*, 2023, **63**, 4079–4087.
- 17 L. Monticelli, *J. Chem. Theory Comput.*, 2012, **8**, 1370–1378.
- 18 A. Chaudhri, I. E. Zarraga, S. Yadav, T. W. Patapoff, S. J. Shire and G. A. Voth, *J. Phys. Chem. B*, 2013, **117**, 1269–1279.
- 19 G. Wang, Z. Varga, J. Hofmann, I. E. Zarraga and J. W. Swan, *J. Phys. Chem. B*, 2018, **122**, 2867–2880.
- 20 S. Izadi, T. W. Patapoff and B. T. Walters, *Biophys. J.*, 2020, **118**, 2741–2754.
- 21 H. Shahfar, J. K. Forder and C. J. Roberts, *J. Phys. Chem. B*, 2021, **125**, 3574–3588.



- 22 A. Grünberger, P.-K. Lai, M. A. Blanco and C. J. Roberts, *J. Phys. Chem. B*, 2013, **117**, 763–770.
- 23 S. Pusara, P. Yamin, W. Wenzel, M. Krstić and M. Kozłowska, *Phys. Chem. Chem. Phys.*, 2021, **23**, 12780–12794.
- 24 F. Zhang, M. W. A. Skoda, R. M. J. Jacobs, S. Zorn, R. A. Martin, C. M. Martin, G. F. Clark, S. Weggler, A. Hildebrandt, O. Kohlbacher and F. Schreiber, *Phys. Rev. Lett.*, 2008, **101**, 148101.
- 25 F. Roosen-Runge, F. Zhang, F. Schreiber and R. Roth, *Sci. Rep.*, 2014, **4**, 7016.
- 26 M. R. Fries, D. Stopper, M. W. A. Skoda, M. Blum, C. Kertzscher, A. Hinderhofer, F. Zhang, R. M. J. Jacobs, R. Roth and F. Schreiber, *Sci. Rep.*, 2020, **10**, 10349.
- 27 A. K. Gaigalas, J. B. Hubbard, M. McCurley and S. Woo, *J. Phys. Chem.*, 1992, **96**, 2355–2359.
- 28 F. Zhang, M. W. A. Skoda, R. M. J. Jacobs, R. A. Martin, C. M. Martin and F. Schreiber, *J. Phys. Chem. B*, 2007, **111**, 251–259.
- 29 A. Bujacz, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2012, **68**, 1278–1289.
- 30 F. Ameseder, R. Biehl, O. Holderer, D. Richter and A. M. Stadler, *Phys. Chem. Chem. Phys.*, 2019, **21**, 18477–18485.
- 31 C. Beck, M. Grimaldo, M. K. Braun, L. Bühl, O. Matsarskaia, N. H. Jalarvo, F. Zhang, F. Roosen-Runge, F. Schreiber and T. Seydel, *Soft Matter*, 2021, **17**, 8506–8516.
- 32 F. Hirschmann, H. Lopez, F. Roosen-Runge, T. Seydel, F. Schreiber and M. Oettel, *J. Chem. Phys.*, 2023, **158**, 084112.
- 33 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1–2**, 19–25.
- 34 W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33–38.
- 35 H. Hamaker, *Physica*, 1937, **4**, 1058–1072.
- 36 E. Jurrus, D. Engel, K. Star, K. Monson, J. Brandi, L. E. Felberg, D. H. Brookes, L. Wilson, J. Chen, K. Liles, M. Chun, P. Li, D. W. Gohara, T. Dolinsky, R. Konecny, D. R. Koes, J. E. Nielsen, T. Head-Gordon, W. Geng, R. Krasny, G.-W. Wei, M. J. Holst, J. A. McCammon and N. A. Baker, *Protein Sci.*, 2018, **27**, 112–128.
- 37 C. R. Søndergaard, M. H. M. Olsson, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 2284–2295.
- 38 A. Salis, M. Boström, L. Medda, F. Cugia, B. Barse, D. F. Parsons, B. W. Ninham and M. Monduzzi, *Langmuir*, 2011, **27**, 11597–11604.
- 39 F. Zhang, S. Weggler, M. J. Ziller, L. Ianeselli, B. S. Heck, A. Hildebrandt, O. Kohlbacher, M. W. A. Skoda, R. M. J. Jacobs and F. Schreiber, *Proteins: Struct., Funct., Bioinf.*, 2010, **78**, 3450–3457.
- 40 O. Matsarskaia, F. Roosen-Runge, G. Lotze, J. Möller, A. Mariani, F. Zhang and F. Schreiber, *Phys. Chem. Chem. Phys.*, 2018, **20**, 27214–27225.
- 41 F. Zhang, personal communication, 2024.
- 42 N. Gnan, F. Sciortino and E. Zaccarelli, in *Patchy Particle Models to Understand Protein Phase Behavior*, ed. J. J. McManus, Springer New York, New York, NY, 2019, pp. 187–208.
- 43 K. Shoemake, *SIGGRAPH Comput. Graph.*, 1985, **19**, 245–254.
- 44 A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.*, 1988, **61**, 2635–2638.
- 45 A. M. Ferrenberg and R. H. Swendsen, *Phys. Rev. Lett.*, 1989, **63**, 1195–1198.
- 46 R. H. Swendsen, *Phys. A*, 1993, **194**, 53–62.
- 47 N. B. Wilding, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1995, **52**, 602–611.
- 48 P. Virnau and M. Müller, *J. Chem. Phys.*, 2004, **120**, 10925–10930.
- 49 M. A. Blanco and V. K. Shen, *J. Chem. Phys.*, 2016, **145**, 155102.
- 50 J. K. G. Dhont, *An Introduction to Dynamics of Colloids*, Elsevier, 1996.
- 51 J. A. Barker and D. Henderson, *J. Chem. Phys.*, 1967, **47**, 4714–4721.
- 52 D. Gazzillo and D. Pini, *J. Chem. Phys.*, 2013, **139**, 164501.
- 53 N. Kern and D. Frenkel, *J. Chem. Phys.*, 2003, **118**, 9882–9889.
- 54 E. Bianchi, G. Kahl and C. N. Likos, *Soft Matter*, 2011, **7**, 8313–8323.
- 55 S. Kohara, Y. Matsuzawa and Y. Kuroda, *Chem. Phys. Lett.*, 2022, **802**, 139767.
- 56 A. Lošdorfer Božič, *Soft Matter*, 2018, **14**, 1149–1161.
- 57 D. Notarmuzi and E. Bianchi, *Soft Matter*, 2024, **20**, 7601–7614.
- 58 M. K. Braun, M. Wolf, O. Matsarskaia, S. Da Vela, F. Roosen-Runge, M. Sztucki, R. Roth, F. Zhang and F. Schreiber, *J. Phys. Chem. B*, 2017, **121**, 1731–1739.
- 59 F. Surfaro, R. Maier, K.-F. Pastryk, F. Zhang, F. Schreiber and R. Roth, *J. Chem. Phys.*, 2023, **158**, 164902.
- 60 M. K. Braun, A. Sauter, O. Matsarskaia, M. Wolf, F. Roosen-Runge, M. Sztucki, R. Roth, F. Zhang and F. Schreiber, *J. Phys. Chem. B*, 2018, **122**, 11978–11985.

