The phase behavior study of human antibody solution using multi-scale modeling

Gang Sun,[1,2] Ying Wang,[3] Aleksey Lomakin,[3] George B. Benedek,[3] H. Eugene Stanley,[2] Limei Xu,[1,4,a)] and Sergey V. Buldyrev[5]
[1]*International Center for Quantum Materials, School of Physics, Peking University, Beijing 100871, China*
[2]*Center for Polymer Studies and Department of Physics, Boston University, Boston, Massachusetts 02215, USA*
[3]*Department of Physics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA*
[4]*Collaborative Innovation Center of Quantum Matter, Beijing, China*
[5]*Department of Physics, Yeshiva University, 500 West 185th Street, New York, New York 10033, USA*

(Received 18 May 2016; accepted 19 October 2016; published online 15 November 2016)

Phase transformation in antibody solutions is of growing interest in both academia and the pharmaceutical industry. Recent experimental studies have shown that, as in near-spherical proteins, antibodies can undergo a liquid-liquid phase separation under conditions metastable with respect to crystallization. However, the phase diagram of the Y-shaped antibodies exhibits unique features that differ substantially from those of spherical proteins. Specifically, antibody solutions have an exceptionally low critical volume fraction (CVF) and a broader and more asymmetric liquid-liquid coexistence curve than those of spherical proteins. Using molecular dynamics simulation on a series of trimetric Y-shaped coarse-grained models, we investigate the phase behavior of antibody solutions and compare the results with the experimental phase diagram of human immunoglobulin G (IgG), one of the most common Y-shape typical of antibody molecules. With the fitted size of spheres, our simulation reproduces both the low CVF and the asymmetric shape of the experimental coexistence curve of IgG antibodies. The broadness of the coexistence curve can be attributed to the anisotropic nature of the inter-protein interaction. In addition, the repulsion between the inner parts of the spherical domains of IgG dramatically expands the coexistence region in the scaled phase diagram, while the hinge length has only a minor effect on the CVF and the overall shape of the coexistence curve. We thus propose a seven-site model with empirical parameters characterizing the exclusion volume and the hinge length of the IgG molecules, which provides a base for simulation studies of the phase behavior of IgG antibodies. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.4966972]

I. INTRODUCTION

Immunoglobulin G (IgG), a type of antibody that is a major component of humoral immunity,[1] is a complex protein composed of four peptide chains—two identical heavy chains and two identical light chains—arranged in a Y-shaped antibody molecule. IgG is the main antibody type found in circulation and it protects the body from infection by binding together many different pathogens, such as viruses, bacteria, and fungi.[2–4] They are also used as drugs to treat diseases, such as autoimmune diseases and cancers.[5] IgG comprises approximately 75% of the serum antibodies in the human body. Most human IgG antibodies are highly soluble. The reason for this is natural selection. If IgG were not highly soluble, it would aggregate in the human body which causes pathological problems. Some IgGs molecules can self-associate to form condensed phases, such as crystals, protein-rich liquid droplets, amorphous aggregates, and gels. Condensation of IgG molecules does occur *in vivo* in the medical condition called cryoglobulinemia. Various types of undesirable condensations of IgG molecules have been observed in antibody drug formulation solutions. It is thus important to understand the phase behavior of IgG solutions underlying these protein condensation phenomena.

Recently the phase diagram of IgG solutions was extensively investigated.[6–8] The experimental phase diagrams of IgG molecules in general share similar features as those of other globular proteins, i.e., they exhibit a metastable liquid-liquid phase separation (LLPS) with a critical point and a parabola-like shape of the LLPS boundary (coexistence curve).[6,7] However, due to the nonspherical molecular geometry, the phase diagram of Y-shaped antibodies also exhibits features that differ substantially from those of spherical proteins. For example, antibodies have an exceptionally low critical volume fraction (CVF), and their LLPS coexistence curves are broader and more asymmetric than those of spherical proteins. In particular, the CVF of IgG is about 6.3%,[6] much smaller than those of spherical particles varying from 13% to 23% as the spatial range of the inter-particle interaction varies from infinity to zero.[9]

There are a number of coarse-grained models, such as the adhesive hard sphere model and the simplified four-site model,[8,10–13] that successfully predict the LLPS in protein solutions. However, these models[8,10–13] are unable to reproduce the smaller CVF and the more asymmetric and

a)Author to whom correspondence should be addressed. Electronic mail: limei.xu@pku.edu.cn.

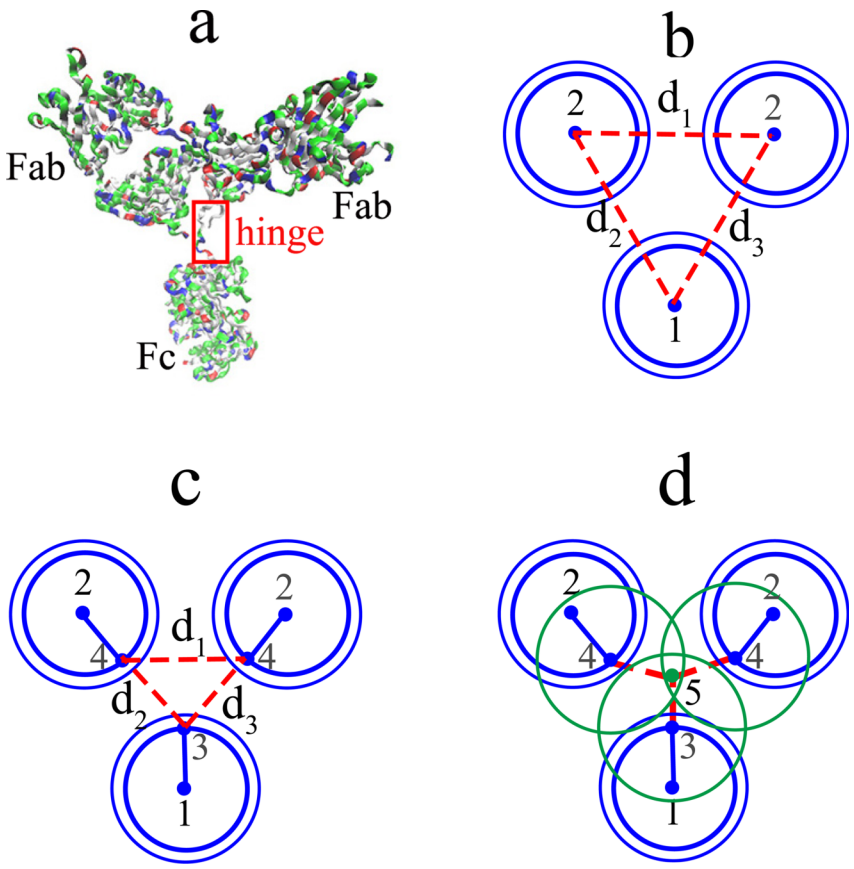

FIG. 1. IgG protein and coarse-grained models. (a) Protein Data Bank (DOI: 10.2210/pdb1hzh/pdb) representation of the IgG molecule with a Fc domain and two Fab domains. PDB 1hzh is rendered by VMD. (b) The 3-site model consisting of one type 1 particle representing the Fc domain and two type 2 particles representing the Fab domains. Dashed lines (d_1, d_2, d_3) between the centers of the domains indicate effective bonds which replace the hinges. The centers of the domains are marked by blue dots. The bold blue circles indicate hard cores of diameter σ of the domains and the thin blue circles indicate the interaction ranges of attractive forces. (c) The 6-site model with hinges. Point 3 is located at the surface of the Fc domain. Points 4 are located at the surfaces of the Fab domains. Dashed lines (d_1, d_2, d_3) between points 4 and 4 and points 3 and 4 on the surfaces of the domains indicate effective bonds which replace the hinges. (d) 7-site mode: Point 5 indicates the branching point of the hinge. Large green circles around points 3 and 4 indicate possible ranges of repulsion in the patchy model. Dashed red lines are effective bonds which replace the hinges.

broader shape of the coexistence curve with respect to those of near-spherical proteins. The primary goal of the work reported here is to obtain a simple model [see Fig. 1] of the Y-shaped trimetric IgG molecules for computational and theoretical studies of the phase behavior of antibody solutions that can address the above issues.

Our approach to tackle this problem is based on multi-scale modeling. An IgG molecule consists of three globular domains connected in a characteristic Y-shape by short polypeptide chains called hinges. At small scales, we model the hinges by all-atom simulations and obtain the distribution of their end-to-end distances. In contrast, at large scales, we model the globular domains as hard spheres with short-range attractions linked to each other by linear bonds obeying effective potentials derived from the distribution of the end-to-end distances of the hinges. The large scale simulations also vary in their accuracy from a simple 3-site model with mean-field attraction to more detailed 6-site and 7-site models, including a 7-site model with directional attraction.[14] This array of coarse-grained models allows us to develop a model that combines the ability to accurately reproduce the IgG phase diagram with computational efficiency. In all coarse-grained simulations we apply the discrete molecular dynamics (DMD)[15,16] algorithm, which is well known for its efficiency in simulating colloids with short-range attraction.

In Section II we describe the details of our models and simulation methods. In Section III A we present phase diagrams of 3-site models for several bond lengths and ranges of attractive forces. In Section III B we present a systematic study of the dependence of the phase diagram on the bond length in a 3-site model with the mean-field approximation of the attractive forces and show how the finite range of attraction affects the phase diagrams. In Section III C we present the results of the all-atom simulation of the hinges and use them to define the properties of the bonds in 6-site and 7-site coarse grained models. In Section III D we present the results from the 7-site model with directional attraction between the globular domains. A brief summary is included in Section IV. Our primary conclusion is that the Y-shape geometry of the IgG model alone is insufficient to explain the reduction of the CVF and the widening of the LLSP boundary with respect to the simple globular proteins, but the anisotropy of attractive forces between the domains is crucial to reproduce these unusual features of the IgG phase diagram.

II. METHODS

A. Models

An IgG molecule consists of three globular domains: one domain called fragment crystallizable (Fc) and two identical antigen-binding (Fab) domains as shown in Fig. 1(a). The diameters of these globular domains σ are approximately equal to each other and are estimated to be ≈4.8 nm using the average density of globular proteins.[6,8,9] In our coarse-grained DMD simulations, we model the domains as hard spheres of diameter σ surrounded by attractive square well potentials with diameter R_a and energy $-\epsilon$. The three domains are held together by two identical polypeptide chains (hinges),

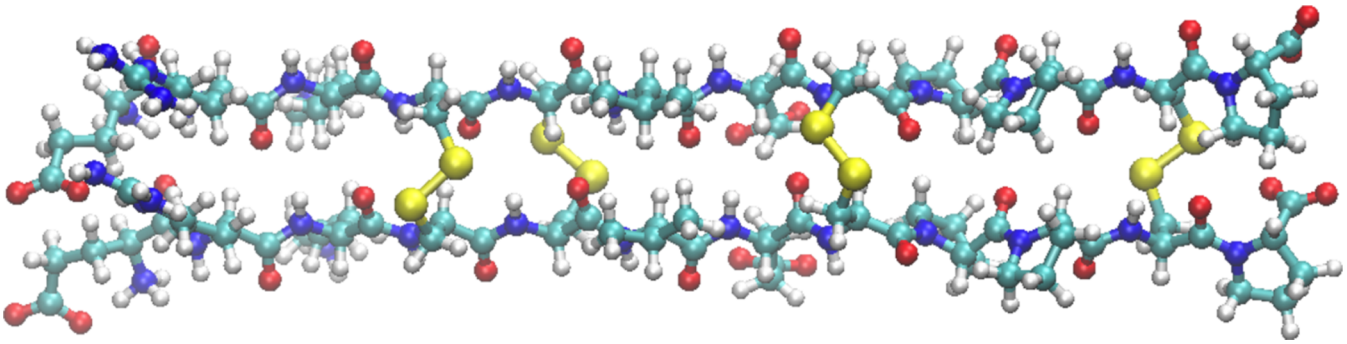

FIG. 2. The initial conformation of the IgG hinges consisting of the two identical polypeptide chains, each consisting of 12 amino acids: GLU ARG LYS CYS CYS VAL GLU CYS PRO PRO CYS PRO (from left to right). Carbon, hydrogen, nitrogen, oxygen, and sulfur atoms are shown in blue, white, green, red, and yellow, respectively. The four disulfide bonds formed between the sulfur atoms of the corresponding cysteines are represented by yellow lines. The maximal elongation of the chains from left to right is 4.6 nm.

which are linked to each other by four disulfide bonds located near their C-termini.[6] Each IgG hinge consists of 12 amino acids. The amino acids of the polypeptide are shown in Fig. 2 in sequence from N-terminus to C-terminus: GLU ARG LYS CYS CYS VAL GLU CYS PRO PRO CYS PRO. The two hinges are connected by disulfide bonds between the corresponding CYS amino acids of the two chains. The N-termini of the hinges are not linked and thus may be separated by a significant distance, since the first disulfide bond between the pair of cysteines is located four amino acids downstream. The C-terminus is attached to the Fc domain, while the N-termini are attached to the Fab domains, thus giving the IgG its characteristic Y-shape in which the branching point coincides with the above-mentioned pair of cysteines. The exact conformation of the hinges is not known but it is believed that in solution they form random coils.[8] Hence the hinges can be represented by three effective bonds with a specific length distribution. The simplest way of representing the hinges is to connect the centers of the globular domains by three linear bonds with a given length distribution, d_1, d_2, d_3 as shown in Fig. 1(b). We call this a 3-site model. It is clear that this model is not an accurate representation of the IgG molecule because in reality the hinges are connected not to the centers but to the surfaces of the domains. We will thus use the 3-site model to study the crude effect of the bond lengths and the attractive potential range on the CVF and LLPS to test the hypothesis that the trimeric structure of the IgG molecule is responsible for the decrease of the CVF and the widening of the LLPS region when compared to monomers, with the same range of attractive forces. Thus, in the 3-site model we use the simplest form of the bonds to connect the spheres. These bonds are modeled by infinite square wells with a minimal distance $d_{\min} = \sigma$ and maximal distance $d_{\max}$, which together with R_a are the only parameters of the 3-site model.

The hinges can be described more accurately using three bonds connecting the auxiliary points on the surfaces of the domains that form a triangle [Fig. 1(c)] or a star [Fig. 1(d)]. These two configurations form the basis for a 6-site model and a 7-site model, respectively. In these models, we use the effective bond potential based on the distribution of the end-to-end distances of the all-atom simulation of the hinges. The distances d_1, d_2, and d_3, defined in Fig. 3(a), are computed differently in the triangular representation of the hinges, which we use in the 6-site model, and the star representation of the hinges [Fig. 3(b)], which we use in the 7-site model. We test the hypothesis that these more accurate models cannot sufficiently change the phase diagram from that of the 3-site model.

We also explore variants of the 3-site, 6-site, and 7-site models without attraction between two monomers. These simplified models take into account only the entropic effects of the hinges and hard spheres. We construct mean-field phase diagrams of these models by adding a van der Waals correction proportional to the square of the number density to the entropic equation of state obtained through simulation. These mean-field phase diagrams can be regarded as a limiting case of the infinitely wide attractive square wells with $R_a \to \infty$ and $\epsilon \to 0$. We refer to these models as mean-field 3-site, 6-site, and 7-site models, respectively.

Finally we bring directionality (patchiness) to the interaction between the domains belonging to different molecules, introducing repulsive step potential wells surrounding the auxiliary points on the surfaces of the domains [Fig. 1(d)]. We call this the most advanced coarse grained representation of the IgG with directional attraction of the patchy 7-site model.

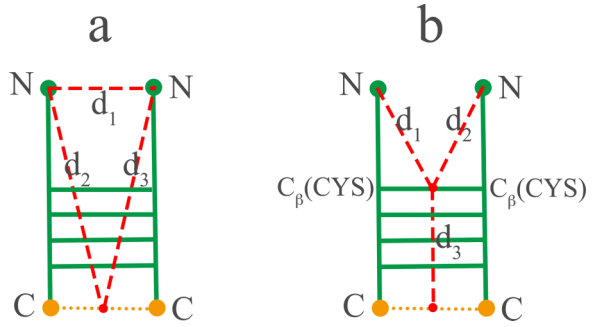

FIG. 3. Sketch of three distances defined between ends in the hinge. (a) Triangular representation of the hinges employed in the six-site model [Fig. 1(c)]: d_1 is the distance between the two N-termini. d_2 and d_3 are the distances between N-termini and the midpoint of the C-termini of the ending prolines. (b) Star representation of the hinges employed in the seven-site model [Fig. 1(d)]: d_1 and d_2 are the distances between N-termini and the middle point of the last disulfide bond, respectively. d_3 is the distance between the midpoint of the disulfide bond and the midpoint of the C-termini of the ending prolines.

In the following we specify the details of all the models we study.

1. 3-site model

Each IgG molecule is represented by three spheres of two distinct types, type 1 representing the Fc domain and type 2 representing the Fab domains [Fig. 1(b)]. For simplicity we assume that particles of type 1 and 2 are identical, which are characterized by the hard core diameter σ and the attractive square well diameter R_a of depth ϵ. The particle centers are connected by three bonds: d_1 connecting particles of type 2, d_2 and d_3 connecting particles of type 1 and type 2. We assume that all bonds are identical and are modeled by the square well with infinitely high walls $\sigma < d_i < d_{\max}$.

2. 6-site model

The domains are modeled the same way as in the 3-site model. We also introduce three auxiliary particles: one particle of type 3 and two particles of type 4 indicating the hinge attaching points to the domains of type 1 and 2, respectively. The particle of type 3 corresponds to the midpoint between the C-termini of the hinges. The particles of type 4 correspond to the N-termini of the hinges as shown in Fig. 1. The auxiliary particles are attached to the centers of the corresponding domains by rigid bonds that are modeled by infinite square wells $0.49\sigma < d < 0.51\sigma$. Bond d_1 links auxiliary particles of type 4, and bonds d_2 and d_3 link the auxiliary particles of type 3 and 4. The bonds are modeled by the effective step potentials based on the distribution of hinge distances obtained in all-atom simulations in a triangular representation defined in Fig. 3(a). The auxiliary particles do not interact with other particles other than those to which they are linked by bonds.

3. 7-site model

The domains of type 1 and 2 and the auxiliary particles of type 3 and 4 are modeled the same way as in the 6-site model [Fig. 1(d)]. In addition to the six existing particles of the 6-site model, we introduce one auxiliary particle of type 5 which represents the branching point of the hinges, i.e., the center of the disulfide bond between the last pair of cysteines before the N-termini of the hinges as shown in Fig. 3(b). A particle of type 5 is linked by two bonds d_1 and d_2 to particles of type 4 and by bond d_3 to particle of type 3. The bonds are modeled by the effective step potentials based on the distribution of distances obtained in all-atom simulations of the hinges in a star representation shown in Fig. 3(b). Thus bond d_1 between particles of type 5 and 3 is the double stranded part of the hinges from their C-termini to the branching point, while bonds d_2 and d_3 between particles of type 5 and type 4 simulate single stranded parts of the hinges connecting the branching point and N-termini. The auxiliary particles do not interact with other particles except those to which they are linked by bonds. In the patchy 7-site model, the auxiliary particles of type 3 and 4 are surrounded by a repulsive shoulder of diameter R_a and height ϵ. This means that if the distance between any particle of type 3 or 4 and any other particle of type 1, 2, 3, or 4 is less than R_a the system gains a potential energy $+\epsilon$ for any such pair.

B. Simulation details for coarse-grained models

For all coarse-grained models, DMD simulations are performed for a system of N_m molecules at constant volume V with periodic boundary conditions and constant temperature T (NVT-ensemble). The temperature T is kept constant by the Berendsen thermostat.[18] We express the results of the DMD simulations in dimensionless quantities normalized by the corresponding units. The distance in DMD simulations is measured in units of σ, the volume, V, is measured in units of σ^3, the temperature, T, is measured in units of ϵ/k_B, where k_B is the Boltzmann constant, the pressure, P, is measured in units of ϵ/σ^3, and the time is measured in units of $\sigma\sqrt{m/\epsilon}$, where m is the mass of the particle. We assume that all particles have the same mass m. This assumption does not affect the phase diagrams, but equal masses minimize the equilibration time. The volume fraction, η, a dimensionless quantity representing the protein volume fraction, is defined as

$$\eta \equiv \frac{\pi}{6}\sigma^3\rho = \frac{3\pi N_m\sigma^3}{6V} = \frac{\pi N_m\sigma^3}{2V}, \tag{1}$$

where ρ is the number density of the monomers, σ is the diameter of the monomer model, N_m is the number of molecules, and V is the volume of the system.

C. Calculation of coexistence lines

For the 3-site and 7-site models with monomer attraction, the liquid-gas and crystal-gas coexistence curves are determined as the loci of the state points at which two-phases are at equilibrium for a given temperature. The gas and the liquid phases of the model correspond to the low-concentration and high-concentration solutions of the IgG proteins, respectively. We first obtain the position of the critical point (η_c, T_c, and P_c) using NVT simulations with $N_m = 500$ and $N_m = 1000$ in a cubic box of volume V corresponding to various volume fractions between $0.01 < \eta < 0.25$ and several values of T for 10^6 time units. For each simulation, we compute pressure $P(\eta,T)$ and construct the isotherms. We then find the critical temperature T_c as the temperature separating the monotonically decreasing isotherms $P(V)$ for $T > T_c$ from the isotherms exhibiting the van der Waals loops for $T < T_C$. The van der Waals loops are observed in the MD simulations due to periodic boundary conditions as explained in Ref. 19. The values of η_c and P_c are obtained from the inflection point of the critical isotherm. The error-bars of the critical parameters are estimated by comparing results of simulations with different numbers of particles. In particular, we estimate the relative error bar of T_c is 0.2%, while the relative error bar of η_c is 5%.

For temperatures relatively close to the critical point $0.95T_c \le T \le T_c$, we obtain the coexistence values of volume fractions in liquid phase η_l and in gaseous phase η_g by the Maxwell construction on the isotherms. The gas and the liquid phases of the model correspond to the low-concentration and

high-concentration solutions of the IgG proteins, respectively. The error bars of these values can be determined by fitting the isotherms obtained from several independent runs with polynomials of different power. The relative errors of both η_l and η_g are estimated as 5%.

For temperatures far below the critical point $T < 0.98T_c$, the correlation length of fluctuations becomes much smaller than the system size. We then determine η_l and η_g by simulating a system with a flat interface between the two phases at constant volume and constant temperature. The system consists of $N_m = 3000$ molecules in an elongated rectangular box of volume $V = L_x L_y L_z = N_m \pi\sigma^3/(2\eta_c)$, corresponding to critical volume fraction with $L_x = L_y = L_z/3$, respectively. Periodic boundary conditions are applied in each direction. After initial equilibration time, the system spontaneously phase segregates along the xy plane to liquid and gaseous phases with parallel interfaces in between and each phase occupying approximately half of the system. We then divide the system into $n_s = 30$ slabs of equal width $\Delta z = L_z/n_s$ orthogonal to the z-axis and count the number of molecules in each slab, which gives us a density profile along the z-axis. We determine the *gas*/*liquid* phase consisting of such slabs that in any of them and in any of n_b neighboring slabs on both sides the density is less/greater than the critical density. Exclusion of $n_b = 5$ slabs is done in order to eliminate the effects of phase boundary. For each snapshot of the system taken each 1000 time units, we define η_g and η_l as the volume fractions of the gaseous and liquid phases and the components P_{xx}, P_{yy}, and P_{zz} of the stress tensor. We then average the volume fractions over approximately 200 independent snapshots and define the 95% confidence error bars by the standard statistical techniques. This method of determining the equilibrium volume fractions is in excellent agreement with the Maxwell construction method in the range of the temperatures $0.95T_c < T < 0.98T_c$, where both methods are applicable. The difference between the volume fractions determined for the systems of different sizes is much smaller than the error bars. The equilibrium volume fractions of the crystal-gas coexistence curves are defined by the same procedure for the systems in the elongated box in which the liquid spontaneously crystallizes and a flat interface forms between the crystal and gas phases. In addition, we determine the equilibrium vapor pressure, $P_e(T) = P_{zz}$, and the liquid-gas surface tension by the KirkwoodBuff method:[20] $\gamma = (2P_{zz} - P_{xx} - P_{yy})L_z/4$.

D. Mean-field phase diagrams

Obtaining accurate phase diagrams of the 3-site model [Fig. 1(b)] with the finite range of attraction is a time consuming process. To quickly estimate how the CVF depends on $d_{\max}$, we employ the mean-field model of the attractive forces and assume that the monomers are attracting each other with an infinitely wide square-well of infinitesimally small depth. Analogous to the behavior of hard spheres,[21] we see that the pressure for the 3-site, 6-site, and 7-site models without attraction is

$$P = P_{HS}(T,\eta) = \frac{k_B T \rho}{n} Z(\eta), \tag{2}$$

where Z is the compressibility factor and n (= 3) is the number of monomers in one molecule. For these models with the mean-field attraction between monomers, the pressure is given by

$$P(\eta,T) = P_{HS}(T,\eta) - a\rho^2/n, \tag{3}$$

where a is the attraction parameter in the van der Waals correction. Using a Maxwell construction, the equilibrium pressure $P_e(T)$ of the gas-liquid coexistence obeys

$$\int_{V_l}^{V_g} P(T,\eta(V))dV = P_e(T)(V_g - V_l), \tag{4}$$

where $\eta(V) = \frac{\pi N_m \sigma^3}{2V}$, $V_l = \frac{\pi N_m \sigma^3}{2\eta_l(T)}$, and $V_g = \frac{\pi N_m \sigma^3}{2\eta_g(T)}$ are the equilibrium specific volumes of the liquid and gas, respectively. V_g and V_l satisfy the equation of state

$$P(\eta_g,T) = P(\eta_l,T) = P_e. \tag{5}$$

The critical point (T_c, η_c, P_c) is determined by $\eta_l(T_c) = \eta_g(T_c) = \eta_c$, and $P_c = P_e(T_c)$. For monomers, $Z(\eta)$ can be well approximated by the Carnahan-Starling formula[21]

$$Z(\eta) = \frac{1 + \eta + \eta^2 - \eta^3}{(1-\eta)^3}. \tag{6}$$

We perform DMD simulations for various systems of molecules without attraction and fit $Z(\eta)$ by a polynomial

$$Z(\eta)(1-\eta)^3 = 1 + \sum_{i=1}^{k} z_i \eta^i \tag{7}$$

for $0.025 \le \eta = 0.2$ with $k = 4$ or $k = 5$. As a consistency check, we obtain $z_1 = z_2 = 1$, $z_3 = -0.584\,22$, and $z_4 = -1.0127$ for monomers with $k = 4$, which is in very good agreement with the Carnahan-Starling approximation. Accordingly, we obtain the critical values for monomers, $T_c = 0.1799a/k_B\sigma^3$, $\eta_c = 0.1302$, and $P_c = 0.016\,04a/\sigma^6$.

E. All-atom simulations of the hinges

The initial structure of the IgG hinges is obtained from the Protein Data Bank (PDB). The all-atom simulations are performed using the MD software package GROMACS.[22] The non-bonded interactions for the all-atom simulations are described by OPLS/AA force field.[23] This force field has been optimized and widely tested for organic molecules, and it has been broadly applied to peptides.[23] The Particle-Mesh Ewald sum (PME) for electrostatics has a real-space cut-off of 1.2 nm. The interpolation order of PME is six and the relative strength of the electrostatic interaction at cut-off is 1.0×10^{-5}. The velocity Verlet algorithm is used to integrate the equation of motion. For all atom simulations, the IgG hinges with the initial structure placed in an empty box and then water molecules (SPC model) are added. Minimization is performed for protein with explicit water. The NVT ensemble is used in the all-atom simulations of hinges in vacuum and the NPT ensemble is used for simulation of the hinges in water (SPC model). We do not apply the long-range corrections for pressure and energy in our simulations. In GROMACS, temperature and pressure are controlled by a Nose-Hoover algorithm.[24,25] The integration time step is 2.0 fs, and the time

TABLE I. The critical parameters of T_c, η_c, and P_c for the 3-site model.

	$d_{max}=0$	$d_{max}=1.0$	$d_{max}=1.64$	$d_{max}=2.0$
$R_a=1.06$	$T_c=0.3916$	$T_c=0.4384$	$T_c=0.4258$	$T_c=0.439$
	$P_c=0.0519$	$P_c=0.0159$	$P_c=0.0129$	$P_c=0.0128$
	$\rho_c=0.52$	$\rho_c=0.48$	$\rho_c=0.40$	$\rho_c=0.37$
	$\eta_c=0.274$	$\eta_c=0.249$	$\eta_c=0.210$	$\eta_c=0.194$
$R_a=1.1$	$T_c=0.4792$	$T_c=0.5515$	$T_c=0.5305$	$T_c=0.5483$
	$P_c=0.0592$	$P_c=0.0195$	$P_c=0.0156$	$P_c=0.0159$
	$\rho_c=0.48$	$\rho_c=0.44$	$\rho_c=0.36$	$\rho_c=0.34$
	$\eta_c=0.251$	$\eta_c=0.234$	$\eta_c=0.189$	$\eta_c=0.174$
$R_a=1.25$	$T_c=0.7612$	$T_c=0.9540$	$T_c=0.8875$	$T_c=0.9115$
	$P_c=0.0782$	$P_c=0.0289$	$P_c=0.0224$	$P_c=0.0221$
	$\rho_c=0.41$	$\rho_c=0.38$	$\rho_c=0.30$	$\rho_c=0.29$
	$\eta_c=0.213$	$\eta_c=0.197$	$\eta_c=0.160$	$\eta_c=0.152$

constant for coupling of temperature and pressure is 1 ps. Periodic boundary conditions are employed for simulations of a protein with explicit water, while non-periodic boundary conditions are used for simulations of a protein in vacuum. Conformations of the hinges are recorded every 1 ps and 50 ps for simulations in vacuum and water, respectively. For each temperature and pressure, the simulations last for at least 100 ns, which is sufficient to obtain a reliable distance distribution of the hinge geometry in vacuum. However, in water, much longer simulation times are needed to explore all partially folded conformations of the hinges. The results of the GROMACS simulations are expressed in SI units.

F. Four-bead DMD model of the hinges

As a fast consistency check of the GROMACS simulations, we also perform the DMD simulations of the hinges using the four-bead model,[16] which reproduces backbone conformations with a realistic distribution of the Ramachandran angles and excluded volume, assuming that all amino acids are identical. The disulfide bonds are modeled by narrow square well potentials between the C_β atoms of the amino acids corresponding to cysteines from the different chains. We perform simulation on the 4-bead model for 10^7 time units and measure the distances d_1, d_2, and d_3 at $T = 0.1\epsilon_{HB}/k_B$, where ϵ_{HB} is the energy of the effective hydrogen bond forming between C and N atoms in the four-bead model. At this temperature, the system never forms more than 8 hydrogen bonds and is able to explore all its conformational space during the simulation time. Since the distance in the four-bead model is measured in Angstroms and the energy of the hydrogen bond is 8 kJ/mol, thus the simulation time corresponds to approximately 10^{-6} s.

III. RESULTS

A. The effects of hinge length and attractive range on the phase diagram

We use the 3-site model to investigate the effect of bond length $d_{\max}$ and attractive range R_a on the phase diagram of IgG. We perform simulations for 12 points in the parameter space with R_a = 1.06, 1.10, and 1.25 and $d_{\max}$ = 0, 1.00, 1.64, and 2.00, where $d_{\max}$ = 0 corresponds to simulations of free monomers (see Table I). Figure 4(a) shows the phase diagrams of this model for R_a = 1.1 and

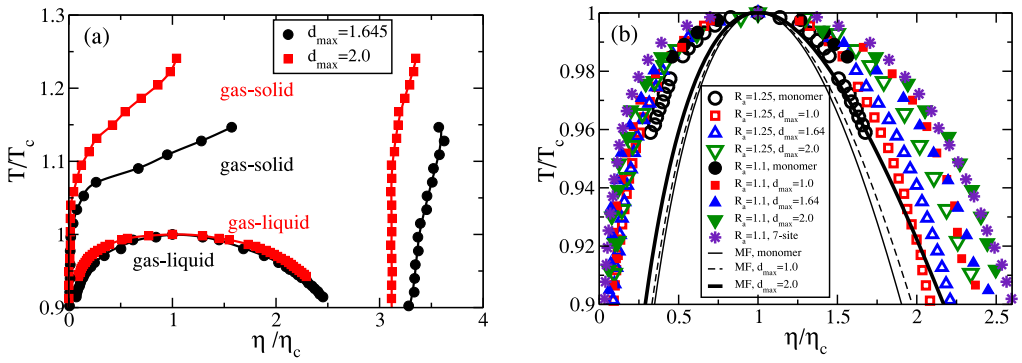

FIG. 4. (a) The phase diagram of 3-site model with radius of attraction R_a = 1.1 and two different bond distances, $d_{\max}$ = 1.64 (black circle) and 2 (red square). The equilibrium coexistent lines of gas-liquid and gas-solid are shown in this figure. For $d_{\max}$ = 1.64 or 2, the middle curve is the gas-liquid equilibrium coexistent line, and the maximum point is the critical point (T_c, η_c) of gas-liquid phase transition. The left curve and the right curve are gas phase and solid phase (hcp crystal), respectively. We find that the equilibrium coexistent lines for both gas-liquid and gas solid are both shifted to higher temperatures from $d_{\max}$ = 1.64 to $d_{\max}$ = 2.0. (b) Scaled phase diagram of the 3-site models with $d_{\max}$ = 1.0 (small squares), $d_{\max}$ = 1.64 (medium triangles), $d_{\max}$ = 2.0 (large triangles) and monomers (circles) with R_a = 1.1 (solid symbols), R_a = 1.25 (empty symbols), and mean-field attraction (lines). One can see systematic broadening of the coexistence regions when R_a decreases. The scaled coexistence region for the 7-site patchy model (stars) is also shown, the broadest among all the model studied in this paper. The error-bars in all cases are less than the symbol size.

two different hinge lengths: $d_{max} = 1.64$ and $d_{max} = 2.0$ in terms of (T, η). The critical volume fraction η_c decreases with d_{max}, and the critical temperature T_c increases with d_{max}, i.e., when $d_{max} = 1.64$, $\eta_c = 0.189$, and $T_c = 0.5305$, and when $d_{max} = 2.00$, $\eta_c = 0.174$, and $T_c = 0.5483$. The shape of the coexistence curve for the three-site model is asymmetric, which is consistent with the experimental phase diagram of IgG antibody proteins.[7]

Previous experimental studies by Wang *et al.*[6] found that the critical temperature T_c of the IgG solution increases with the net inter-protein interactions. Our study confirms that the increase of the range of attraction between particles, R_a, leads to the dramatic increase of the critical temperature. The effect of d_{max} on the critical temperature is weak and non-monotonic. The critical temperature of the 3-site model with d_{max} is greater than that for monomers with the same R_a, than it slightly decreases from 1 to 1.64 and increases again from 1.64 to 2.

We also find that for gas-liquid coexistence, the dependence of equilibrium pressure on temperature follows the Arrhenius law and the surface tension of the liquid decreases almost linearly with temperature. The critical point of the gas-liquid transition can also be determined as the temperature where surface tension $\gamma = 0$.

We find that all 3-site models with a short range of attraction spontaneously crystallize at sufficiently low T into a hexagonal close-packed crystal. Figure 4(a) also shows the gas-solid coexistence line, which is above the metastable gas-liquid coexistence curve in the phase diagram for each d_{max}, indicating that the liquid-liquid phase separation is metastable against crystallization. The volume fractions of the crystal and gas phase for a given temperature decrease with d_{max}. In addition, the increase in d_{max} causes the LLPS region to submerge more deeply into the LLPS region below the gas-solid coexistence line.

The increase of attraction radius, R_a, and the increase of the maximal hinge length, d_{max}, both reduce the CVF [Fig. 5(a)]. Note that this statement includes monomers, which can be regarded as 3-site model with $d_{max} = 0$. Moreover, CVF as a function of R_a and d_{max} can be well approximated as

$$\eta_c = f(R_a) + g(d_{max}), \tag{8}$$

where $f(R_a)$ and $g(d_{max})$ are monotonically decreasing functions.

In contrast, R_a and d_{max} have opposite effect on the width of the gas-liquid coexisting region in the scaled phase diagrams [Fig. 4(b)]. The increase of R_a always decreases the width of the scaled coexistence region, while the increase of d_{max}, in general, increases the width of the coexistence region. Moreover, the scaled coexisted region for the 3-site model is always broader than the scaled coexistence region for monomers with the same R_a. The most significant widening of the coexistence region with d_{max} is observed for wide R_a, while for narrow $R_a = 1.1$, the effect becomes non-monotonic: the width slightly decreases when d_{max} is between 1.0 and 1.64, but significantly increases from 1.64 to 2.0. Thus the three site model qualitatively reproduces experimental findings that the coexistence region for the IgG is wider than that of the globular proteins.

In conclusion, our simulations of the 3-site model show that the trimeric structure of the IgG may contribute to the experimentally observed decrease of the CVF comparatively to the globular proteins. However, quantitatively, the effect is not sufficiently strong to produce the twofold reduction of the CVF in the IgG. For example, for $R_a = 1.1$ which is a realistic attractive range for proteins, monomers have $\eta_c = 0.251$, while η_c in the 3-site model with $d_{max} = 1.0$ is 0.234, which gives only a 1.1 reduction in CVF. Even for $d_{max} = 2.0$, for which $\eta_c = 0.174$, the CVF reduction is only 1.4.

B. Comparison with the mean-field model

The scaled coexistence region of the 3-site mean field model is significantly narrower than that for the 3-site model with finite range of attraction, and has a parabolic top corresponding to the mean field critical exponent $\beta = 0.5$ [Fig. 4(b)]. The coexistence regions for models with finite range of attraction have much flatter tops suggesting smaller

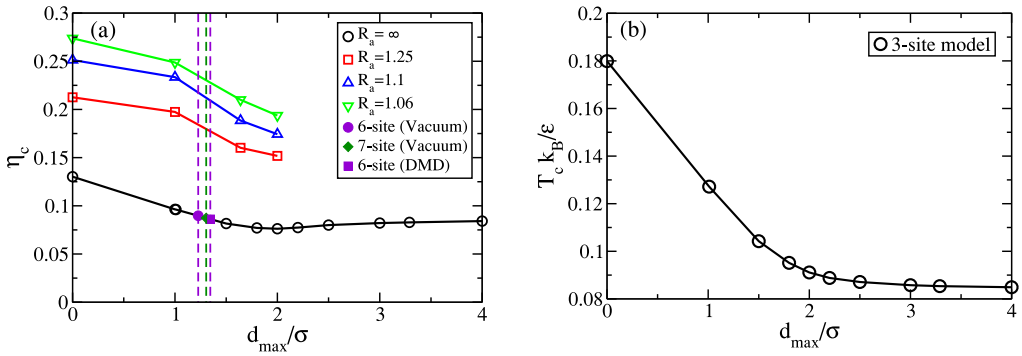

FIG. 5. (a) Critical volume fraction η_c as function of d_{max} for the mean-field model for a 3-site model of a trimer ($d_{min} = \sigma$) in comparison to the monomer model for which we assume $d_{max} = d_{min} = 0$ (black circle). We also show η_c for several d_{max} and finite ranges of attraction $R_a = 1.25$ (red square), $R_a = 1.1$ (blue triangle), and $R_a = 1.06$ (green triangle). Filled symbols indicate η_c for the mean-field approximation for 6 and 7 site models, for which distances d_1, d_2, and d_3 are governed by the effective potential obtained from the distribution of the end-to-end distances of the all-atom simulations of the hinges in vacuum at $T = 3000$ K and the DMD simulations of the four bead model. From the values of η_c, we can define the effective value of d_{max} (vertical dashed lines) and finally we can obtain approximate values of η_c for the models with finite range attraction by interpolation. (b) Critical temperature T_c as a function of d_{max} for the mean-field approximation of the 3-site model.

β, which for the Ising universality class is approximately 0.326 45.[17] The increase of $d_{\max}$ in the 3-site mean-field model also causes widening of the scaled coexistence region, but to a smaller extent than in the model with finite range of attraction.

Figure 5 shows the results for various 3-site, 6-site, and 7-site models. The main finding is that the CVF for the mean field 3-site model as a function of $d_{\max}$ reduces to $\eta_c = 0.0762$ for $d_{\max} = 2$, slowly increases, and then stabilizes at $\eta_c = 0.0814$ for $d_{\max} \to \infty$. The maximum reduction in the CVF compared to the mean field model of monomers is 1.7 times. Comparing the mean-field results with the results for the 3-site model with finite range attraction suggests an almost parallel shift upwards of the CVF for the finite range of attraction with respect to the mean-field calculation predictions. Thus, for the narrow range of attraction that is relevant for protein globules, the maximal reduction of the CVF of the 3-site model compared to the monomer is less significant. For example, when $R_a = 1.06$, the maximal reduction of the CVF with respect to the monomer is only 1.34. We hypothesize that the same shift in the CVF would be observed in the 6- and 7-site models when we replace the mean-field attraction by the same finite attractive range.

C. Effects of IgG hinge elasticity on the CVF

As can be seen from simulations of the 3-site model, the maximal CVF reduction is achieved in the 3-site model for $d_{\max} = 2$. It is unlikely that the elasticity of the hinges can produce a larger CVF reduction. To test this hypothesis, we include the elasticity of the hinges to the 6-site and 7-site models [see Figs. 1(c) and 1(d)]. To compute the elasticity of the hinges, we perform all-atom GROMACS simulations of the hinges in vacuum and in water and measure the histogram $P_i(d_i)$ of the end-to-end distances d_1, d_2, and d_3 of the hinges in triangular and star representations as shown in Figs. 3(a) and 3(b), respectively. We also perform DMD simulations of the four-bead model of the hinges. For all the models we use a 0.1 nm bin to construct the histograms.

The most accurate predictions of the effects of the hinge elasticity on the CVF would be produced by the all-atom simulations of the hinges in water at ambient conditions,

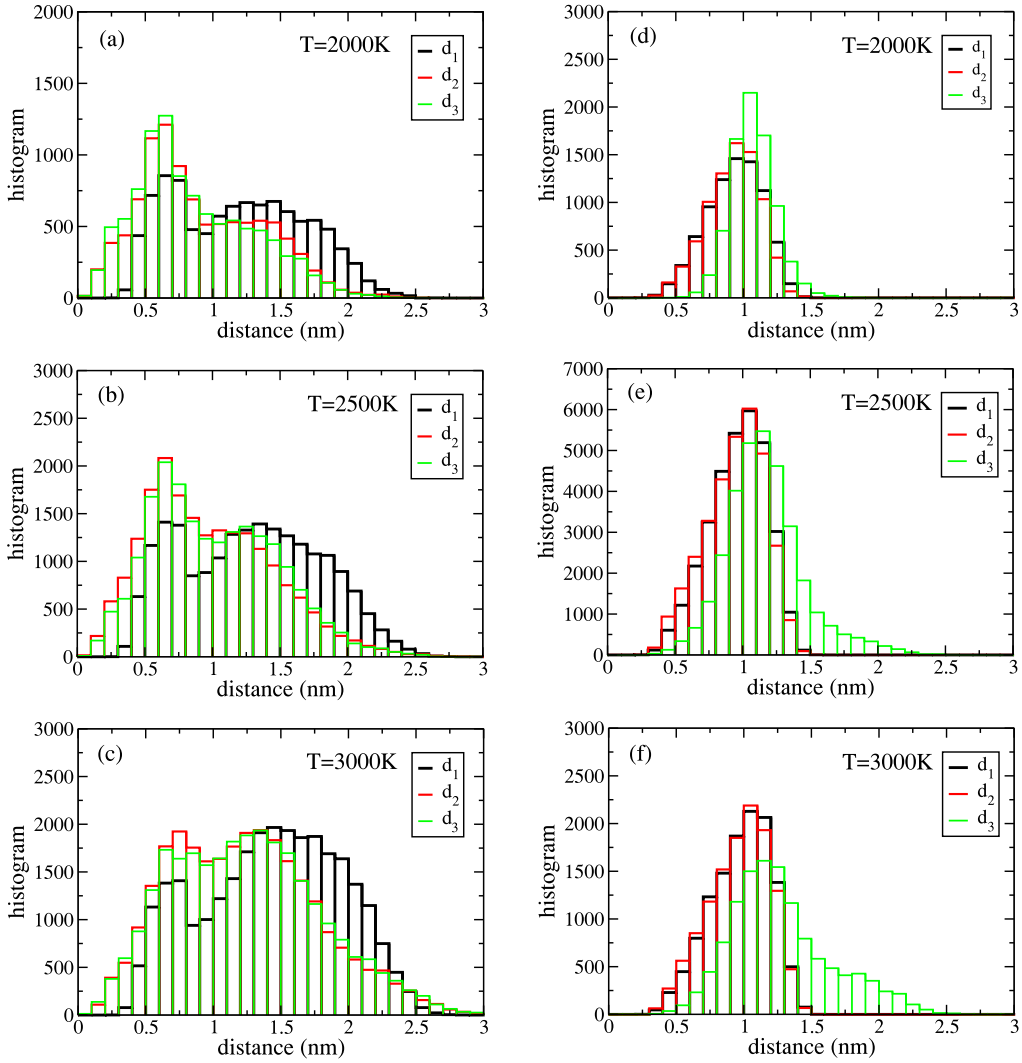

FIG. 6. The distribution of three distances, d_1, d_2, and d_3 for the triangle and star representations of the hinges in Fig. 3 obtained by all atom simulations in vacuum for three temperatures: T = 2000 K, 2500 K, and 3000 K for both the triangle ((a)-(c)) representation and the star representation ((d)-(f)). We can see that the distributions are slightly broadening with temperature. For each temperature, we run simulations for at least 100 ns and save conformations every 1 ps.

provided that the histograms obtained in such simulations are statistically representative of all possible hinge conformations. Since the hinges are symmetrical, the distances d_1 and d_2 in the star-representation of the hinges and distances d_2 and d_3 in the triangular representation must have, in the limit of infinitely long runs, identical histograms. The differences between these histograms indicate the statistical errors in the distributions. Another test of the statistical significance of the histograms is the comparison among different runs obtained using different initial velocities of the atoms. We note that the simulations of the hinges in water at ambient conditions for 100 ns always produce totally different distributions of the symmetrical end-to-end distances and differ dramatically in different runs. This indicates spontaneous folding of hinges in various folded conformations, separated by very high free energy barriers which cannot be overcome within reasonable computational times (see the supplementary material). The same is true for the simulations of hinges in vacuum for $T < 2000$ K. Only for $T \geq 2000$ K the all atom simulations in vacuum produce statistically representative histograms.

Figure 6 shows the histograms $P(d_i)$ of the distribution of the distances d_1, d_2, and d_3 for the all-atom triangle and star representation of the hinges in the vacuum (Fig. 3). Above 2000 K, the histograms for star representation are unimodal as one would expect for random coil conformations. Moreover, the distance d_3 corresponding to a long double-stranded part of the hinges has a significantly wider distribution than the distances d_1 and d_2, which correspond to shorter single strands of the hinges. In contrast, the distributions of the distances in the triangular representation obtained from the same simulations are bimodal. The bimodal histograms indicate the presence of two types of configurations of the hinges: one with hinge ends being attracted to each other and the other with ends of hinges independently distributed in space. The histograms of the hinges in water at ambient conditions show multiple erratic peaks in both triangle and star representations in the range between 0.5 and 2.5 nm with a tendency to be more compact than the simulations in vacuum at high temperature. As a consistency check of the GROMACS simulations, we also perform simulations of hinges using the DMD 4-bead model. The peaks of the distribution in the triangular representation are located at $d_2 = d_3 = 2.8$ nm and $d_1 = 0.8$ nm in contrast to the all-atom simulations in vacuum at $T = 3000$ K in which the highest peaks are located at $d_2 = d_3 = 1.5$ nm. In general, this model results in more stretched conformations of the hinges than all-atom simulations and thus can serve to establish the upper bound of the effect of the hinge length on the CVF.

Using the values of $P_i(d_i)$ in each bin, we define stepwise effective potentials,

$$U_i(x_i) = -k_B T \ln(P_i(x_i\sigma)/x_i^2), \tag{9}$$

where $x_i = d_i/\sigma$ and σ is the diameter of the actual IgG monomer globules, which is estimated to be $\sigma = 4.8$ nm from the average density of the globular proteins. These stepwise potentials with the step width $\Delta x_i = \sigma/48$ corresponding to the bins of the histograms can be directly used in the DMD simulations of the 6-site and 7-site models to specify the elasticity of the auxiliary bonds.

We run the DMD simulations for 6-site and 7-site models with hard spheres of diameter σ without attraction connected by the auxiliary bonds obeying the effective potential $U_i(x_i)$ for different volume fractions in the 0.025 to 0.2 range at $T = 1$ and fit the pressure as a function of volume fraction by Eq. (7) [Fig. 7(b)]. Using Maxwell construction for this fit, we find the mean-field phase diagrams for all these models.

As an example, Figure 7(a) shows the effective potential defined by Eq. (9) for the 6-site model derived from the histograms of the distances in the triangular representation of the DMD 4-bead model of the hinges. Figure 7(b) shows the fit of $P(\eta)$ obtained in the 6-site model simulations at $T = 1$ with the bond potential displayed in Fig. 7(a) and without monomer attraction. Finally, Figure 7(c) shows the mean-field coexistence $T(\eta)$ line based on this fit. Accordingly, we find $\eta_c = 0.862$ for the 6-site model based on the 4-bead DMD simulations of the hinges. Calculations using 6- and 7-site models for the effective bond potentials derived for triangular and star representation of the all-atom simulations in vacuum are qualitatively the same. Using $P_i(d_i)$ obtained for all-atom simulations in vacuum at $T = 3000$ K, we find that $\eta_c = 0.0873$ for the 7-site model, and $\eta_c = 0.0896$ for the 6-site model. These values are slightly larger than η_c for the DMD 4-bead model of the hinges, since the latter model results in more expanded conformations than all-atom simulations in vacuum.

Comparing the mean-field values of η_c obtained for the 6- and 7-site models with the bond potential derived from the simulation of hinges and the graph of $\eta_c(d_{\max})$ of the 3-site model mean-field simulations with different values of $d_{\max}$ [Fig. 5(a)], we can estimate the effective $d_{\max}$ of the

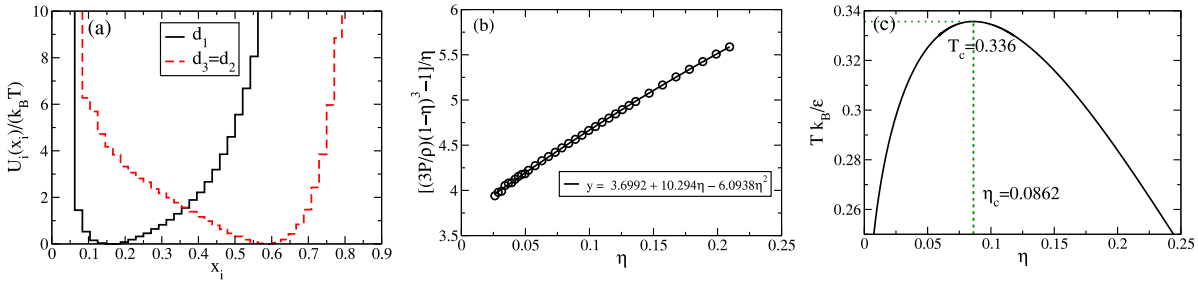

FIG. 7. (a) Effective bond potential obtained from the distribution of the distance d_1, d_2, and d_3 for the triangular representation [Fig. 3(b)] in the 4-bead DMD model of the hinges. (b) Fit of the function $[Z(\eta)(1-\eta)^3-1]/\eta]$ obtained for the 6-site model shown in Fig. 1(c) without attraction simulated for the bond potential shown in panel (a) with $\sigma = 4.8$ nm. (c) Coexistence curve obtained by the Maxwell construction using the fit in panel (b) and Eq. (3). The analogous graphs for the effective potential and coexistence curve obtained for the all-atom simulations for the 6 and 7-site models are very similar.

3-site model corresponding to the 6- and 7-site models that we study in this section. As a result, we find that for the 6-site model based on the all-atom simulations in vacuum $d_{\max} = 1.227\sigma$, while for the 7-site model based on the all-atom simulations of the hinges in vacuum $d_{\max} = 1.305\sigma$. For the 6-site model based on the 4-bead DMD simulations of the hinges $d_{max} = 1.345\sigma$, which provides the upper bound for the effective $d_{\max}$. Thus, we can conclude that the effective $d_{\max}$ is quite insensitive to the details of the hinge simulations, but cannot be greater than $d_{\max} = 1.345\sigma$.

We assume that η_c for the model with finite range attraction R_a can be found by the vertical shift of η_c obtained for the corresponding mean-field model. In practice, we find these values of η_c by interpolating the graphs $\eta_c(d_{\max})$ for a given R_a using the effective $d_{\max}$ found for these 6- and 7- site models as shown in Fig. 5(a). As one can see, the maximum CVF reduction of the 6-site and 7-site models based on the hinge simulations in vacuum for $R_a = 1.06$ is only 1.17 with respect to the monomer model. For the all-atom simulations of the hinges in water, the distances d_1, d_2, and d_3 are shorter than those in vacuum. Thus we expect the effective $d_{\max}$ for the hinges in water to be even smaller, but it still must be larger than σ. Thus the CVF reduction cannot be smaller than that of the closely packed 3-site model with $d_{\max} = \sigma$. Accordingly, for $R_a = 1.06$ the CVF reduction compared to monomers must be greater than 1.13. To conclude, for all hinge models studied we expect the CVF reduction compared to the monomer to be in the 1.13 to 1.17 range for $R_a = 1.06$. Even when $R_a = 1.25$ the CVF reduction cannot be greater than 1.26.

We thus conclude that the trimeric structure of the IgG alone cannot explain the reduction of the CVF by a factor of 2 compared to that of the monomers observed in experiments. One possible explanation could be that the effective radius of the IgG globules is larger than the one estimated for the average protein density. It could be due to the non-spherical shape of the actual domains or due to the existence of the empty spaces inside the domains. Assuming that $\sigma > \sigma_0 = 4.8$ nm, the scaled end-to-end distances x_i of the hinges will be reduced by a factor σ_0/σ, but the effective $d_{\max}$ cannot be smaller than σ, hence the effective reduction of the CVF will still be at least 1.1. However, the actual reduction of the CVF should be at least $1.1(\sigma/\sigma_0)^3$. Thus, by choosing $\sigma = 2^{1/3}\sigma_0 = 6$ nm one can achieve the experimental reduction of the volume fraction of a factor of 2.2.

D. Effects of the patchiness of the attractive potential

Another explanation for the reduction of the CVF of IgG compared to a typical globular protein can be the patches of the attractive potential.[14,26] For example, it is natural to assume that the inner parts of the globules close to the points of the hinge attachment are not attractive. To verify whether this is a sufficient condition for significantly reducing the CVF, we design a 7-site patchy model of IgG [see Fig. 1(d)]. Based on the end-to-end distances of the hinges obtained from all-atom simulations using GROMACS, the bonds between particles of type 5 and 3 in the seven-site model fluctuate between 0.0625σ and 0.563σ, and the bonds between particles of type 5 and 4 fluctuate between 0.0417σ and 0.333σ. These ranges are approximately equal to the fluctuation ranges of the end-to-end distances of the atomistic hinge models. Finally we add the repulsive shoulders of radius $R_a = 1.1\sigma$ and height ϵ to the auxiliary particles of type 3 and 4. These repulsive soft cores of the auxiliary particles prevent the monomers of the other molecules from being attracted to the central part of the molecule.

Simulations of the 7-site model with repulsion (Fig. 8) give the $\eta_c = 0.11$, while the model of disconnected monomers with attraction range $R_a = 1.1$ has η_c of 0.254. Thus the CVF reduction factor for the 7-site patchy model is 2.3. The width of the scaled phase diagram for the patchy model is the same as for the 3-site with $d_{\max} = 2$ for the same radius of attraction $R_a = 1.1$ and it is significantly wider than the phase diagram of the 3-site model with $d_{\max} \approx 1.3$ which is the most realistic value predicted from analyzing the all-atom models of the hinges (Fig. 4). We assume that the scaled coexistence region for this $d_{\max}$ should be almost the same as the coexistence regions for $d_{\max} = 1.0$ and $d_{\max} = 1.64$, which practically coincide with each other. Thus the 7-site patchy model can account for both the experimentally observed reduction of the CVF and for the widening of the scaled phase diagram. Note that this model does not spontaneously crystallize near the liquid-gas critical point. The radial distribution function indicates that the liquid phase is a typical net-forming liquid in which each molecule has 9 neighboring molecules in the first coordination shell. Thus, in a typical configuration, the outer surface of each monomer of a given molecule is attached to three monomers, each of which belongs to a different neighboring molecule.

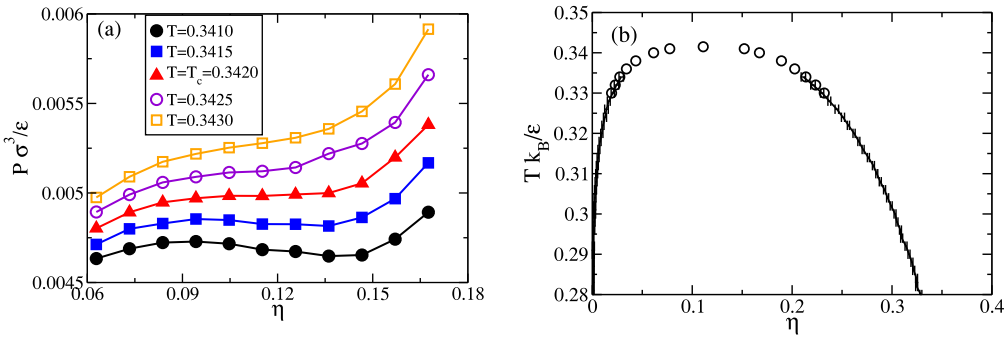

FIG. 8. Results of the simulations of the 7-site patchy model. (a) Isotherms near the critical temperature. (b) Coexistence curve obtained by simulations of the gas-liquid boundary in the elongated box. Symbols show results obtained by the Maxwell construction on the isotherms. Lines with error bars show the results obtained by direct simulations of the coexisting phases.

This loose-packed structure with a very low volume fraction can collapse under high pressure, suggesting the existence of a density anomaly and even a second liquid-liquid phase transition.[26–31] Note that the seven-point model has a number of parameters that can be adjusted to better simulate the actual IgG antibody proteins.

IV. CONCLUSION

We perform multiscale simulations of the human IgG in order to explain large differences between the phase diagram of Y-shape IgG proteins and those of globular proteins. Our results show that if we assume that the IgG domains are spheres with an average protein density connected by bonds of appropriate length obtained from the simulations of the IgG hinges, the reduction of CVF with respect to the globular proteins is insignificant. There are two possible ways to explain the discrepancy. The first is simply to assume that the effective diameters of the domains are greater (6 nm) than one would predict from the average protein density (4.8 nm). A second possible explanation is to assume that the inner parts of the IgG molecule near the points of attachment to the hinges do not attract the domains of other IgG molecules. The results of the simulations of a 7-site model with repulsion of the inner parts of the IgG not only reproduce the reduction of the CVF but also display a dramatical expansion of the coexistence region in the scaled $\eta - T$ phase diagram [Fig. 4(b)]. Thus, our simulations of the 7-site patchy model provide a general framework for future modeling of the phase diagrams of human antibody solutions.

SUPPLEMENTARY MATERIAL

See supplementary material for the distribution of distances defined in Fig. 3(a). In supplementary material Fig. S1, we show the distribution of distances of IgG hinges in vacuum. Fig. S2 shows the distribution of distances of IgG hinges in water. In Fig. S3, we indicate the dependence of the surface tension on temperature and the Arrhenius relationship for temperature and pressure in a gas-liquid equilibrium system for trimers with finite range of attraction.

ACKNOWLEDGMENTS

G. Sun and L. Xu acknowledge the support from the National Basic Research Program of China (973 Program, Grant No. 2015CB856801), the National Key Research and Development Program of China (Grant No. 2016YFA0300901), and the National Science Foundation of China (NSFC Grant Nos. 11525520 and 11290162). S.V.B. acknowledges the support of DTRA (Grant HDTRA1-14-1-0017) and the partial support of this research through Dr. Bernard W. Gamson computational Science Center at Yeshiva College. We are also grateful for computational resources provided by the supercomputer TianHe-1A in Tianjin, China. The Boston University work was supported by NSF Grants PHY 1444389, PHY 1505000, CMMI 1125290, and CHE-1213217, by DTRA Grant HDTRA1-14-1-0017, and by DOE Contract DE-AC07-05Id14517.

[1]T. J. Kindt, R. A. Goldsby, B. A. Osborne, and J. Kuby, *Kuby Immunology*, 6th ed. (W.H. Freeman, New York, 2007).
[2]A. R. Neurath, Encycl. Virol. **25**, 56-70 (2008).
[3]M. Welschof, P. Terness, S. M. Kipriyanov, D. Stanescu, F. Breitling, H. Dörsam, S. Dübel, M. Little, and G. Opelz, Proc. Natl. Acad. Sci. U. S. A. **94**, 1902 (1997).
[4]A. E. Stewart and D. H. Hunsaker, Otolaryngol.–Head Neck Surg. **127**, 324-332 (2002).
[5]S. Dübel, *Handbook of Therapeutic Antibodies: Technologies, Emerging Developments and Approved Therapeutics* (John Wiley & Sons, 2010).
[6]Y. Wang, A. Lomakin, R. F. Latypov *et al.*, J. Chem. Phys. **139**, 121904 (2013).
[7]Y. Wang, A. Lomakin, R. F. Latypov, and G. B. Benedek, Proc. Natl. Acad. Sci. U. S. A. **108**, 16606 (2011).
[8]J. Li, R. Rajagopalan, and J. Jiang, J. Chem. Phys. **128**, 205105 (2008).
[9]A. Lomakin, N. Asherie, and G. B. Benedek, J. Chem. Phys. **104**, 1646-1656 (1996).
[10]M. H. J. Hagen and D. Frenkel, J. Chem. Phys. **101**, 4093 (1994).
[11]D. L. Pagan and J. D. Gunton, J. Chem. Phys. **122**, 184515 (2005).
[12]J. F. Lutsko and G. Nicolis, J. Chem. Phys. **122**, 244907 (2005).
[13]S. Brandon, P. Katsonis, and P. G. Vekilov, Phys. Rev. E **73**, 061917 (2006).
[14]N. Kern and D. Frenkel, J. Chem. Phys. **118**, 9882-9889 (2003).
[15]D. C. Rapaport, *The Art of Molecular Dynamics Simulations* (Cambridge University Press, New York, 1995).
[16]S. V. Buldyrev, *Application of Discrete Molecular Dynamics to Protein Folding and Aggregation*, Lecture Notes in Physics Vol. 752 (Springer, Berlin Heidelberg, 2008), pp. 97-131.
[17]S. El-Showk, M. F. Paulos, D. Poland, S. Rychkov, D. Simmons-Duffin, and A. Vichi, J. Stat. Phys. **157**, 869 (2014).
[18]H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren *et al.*, J. Chem. Phys. **81**, 3684-3690 (1984).
[19]K. Binder, B. J. Block, P. Virnau, and A. Tröster, Am. J. Phys. **80**, 1099-1109 (2012).
[20]J. Wang and X. C. Zeng, J. Theor. Comput. Chem. **08**, 733-763 (2009).
[21]N. F. Carnahan and K. E. Starling, J. Chem. Phys. **51**, 635-636 (1969).
[22]B. Hess, C. Kutzner, D. van der Spoel, and E. Lindahl, J. Chem. Theory Comput. **4**, 435 (2008).
[23]W. L. Jorgensen, D. S. Maxwell, and J. Tirado-Rives, J. Am. Chem. Soc. **118**, 11225-11236 (1996).
[24]S. Nose, J. Chem. Phys. **81**, 511-519 (1984).
[25]William G. Hoover, Phys. Rev. A **31**, 1695-1697 (1985).
[26]F. Romano, E. Sanz, and F. Sciortino, J. Chem. Phys. **134**, 174502 (2011).
[27]P. H. Poole, I. Saika-Voivod, and F. Sciortino, J. Phys.: Condens. Matter **17**, L431-L437 (2005).
[28]Y. Tu and H. Fang, Phys. Rev. E **79**, 016707 (2009).
[29]Y. Tu, S. V. Buldyrev, Z. Liu, H. Fang, and H. E. Stanley, Europhys. Lett. **97**, 56005 (2012).
[30]S. V. Buldyrev and G. Franzese, J. Non-Cryst. Solids **407**, 392 (2015).
[31]T. A. Kesselring, E. Lascaris, G. Franzese, S. V. Buldyrev, H. J. Herrmann, and H. E. Stanley, J. Chem. Phys. **138**, 244506 (2013).