MABS
2021, VOL. 13, NO. 1, e1907882 (10 pages)
https://doi.org/10.1080/19420862.2021.1907882

Taylor & Francis
Taylor & Francis Group

REPORT

Calculation of therapeutic antibody viscosity with coarse-grained models, hydrodynamic calculations and machine learning-based parameters

Pin-Kuang Lai, James W. Swan, and Bernhardt L. Trout

Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

ABSTRACT

High viscosity presents a challenge for manufacturing and drug delivery of therapeutic antibodies. The viscosity is determined by protein–protein interactions among many antibodies. Molecular simulation is a promising method to study protein–protein interactions; however, all-atom models do not allow the simulation of multiple molecules, which is necessary to compute viscosity directly. Coarse-grained models, on the other hand can do this. In this work, a 12-bead coarse-grained model based on Swan and coworkers (J. Phys. Chem. B 2018, 122, 2867–2880) was applied to study antibody interactions. Two adjustable parameters related to the short-range interactions on the variable and constant regions were determined by fitting experimental data of 20 IgG1 monoclonal antibodies at 150 mg/mL. The root-mean-square deviation improved from 1 to 0.68, and the correlation coefficient improved from 0.63 to 0.87 compared to that of a previous model that assumed the short-range interactions were the same for all the beads. Our model is also able to calculate the viscosity over a wide range of concentrations without additional parameters. A tabulated viscosity based on our model is provided to facilitate antibody screening in early-stage design.

ARTICLE HISTORY
Received 14 January 2021
Revised 17 March 2021
Accepted 20 March 2021

KEYWORDS
Antibody viscosity; coarse-grained models; molecular dynamics simulations; hydrodynamic calculations

Introduction

Predicting the viscosity of antibodies at high concentrations is important for the manufacturing and delivery of therapeutic drugs in development.[1–4] However, the cost of producing a large amount of any antibody in early-stage design is high. Therefore, computational methods that can calculate viscosity at high concentrations are desired. Since antibody viscosity is primarily governed by protein–protein interactions,[5–8] molecular simulations are promising tools to study these interactions.[9–11] Nevertheless, all-atom simulations are too computationally expensive to study high concentration systems for large biomolecules.[12] An alternative way is to apply coarse-grained (CG) models to improve the computational efficiency.[13,14]

CG models have been applied to study the self-association of therapeutic monoclonal antibodies (mAbs).[15–17] The all-atom antibody models are represented by a few domains (beads) in the CG models. Among different CG models, a 10-bead and a 12-bead model have been applied to study the viscosity of mAb solutions.[18–20] Although these works showed progress in computational methods for viscosity calculation, there are some disadvantages of these methods. The method of Chowdhury et al.[20] requires fitting to experimental viscosity data using the cluster size distribution. This is not actually a predictive model for viscosity. The method of Izadi et al.[19] assumes an inverse relationship between calculated diffusion coefficients and experimental viscosity based on the Stokes–Einstein equation. However, it remains uncertain how well the Stokes–Einstein equation applies to the antibody molecules with nonspherical shape.[21] The method of Wang et al.[18] provides a direct way to calculate viscosity from hydrodynamic calculations. Nevertheless, it assumed all the CG beads have the same short-range interaction parameters. This may not be applicable to the variable region where high sequence diversity exists for each mAb. This could also be the reason that Wang's model failed to describe the large viscosity of one antibody they studied. Moreover, all these methods have only been tested for a small number of distinct antibodies (2 or 3 mAbs). Larger datasets are needed to validate the model performance.

In this work, we develop and improve a computational model based on Wang et al.[18] The goal of this project is to determine the short-range interaction parameters, the Hamaker constants (A_H), of the CG model for different antibodies. We propose that the Hamaker constants are divided into two contributions, from the variable regions (A_H^v) and the constant regions of the antibodies (A_H^c). The A_H^v values are determined by a high viscosity index (HVI), previously developed from a machine learning approach.[22] The HVI values only depend on the sequence on the variable fragment (Fv) region. We assume that $A_H^v = \alpha \times \mathrm{HVI}$ and α is a scaling parameter. The A_H^c values are the same for all the antibodies. The two parameters (α and A_H^c) are determined by fitting to 20 experimental data at 150 mg/mL. The new model can calculate antibody viscosity over a wide range of concentrations without additional fitting. This method will facilitate drug development and assist the understanding of the mechanism of antibody viscosity behavior. Finally, a tabulated viscosity table based on our new model is provided. It requires only the sequence

CONTACT Bernhardt L. Trout trout@mit.edu Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
Supplemental data for this article can be accessed on the publisher's website

information to estimate viscosity, which allows fast antibody screening in early-stage design.

Results

Sensitivity analysis of the viscosity model

In the CG model, the viscosity depends on the charges and the short-range interaction parameters, the Hamaker constants. It is informative to analyze the sensitivity of these parameters. Figure 1 shows a heatmap of viscosity as a function of the charges in the VH region and the Hamaker constants on the variable region (A_H^v).

Effect of the A_H^c values on viscosity

The viscosity is also sensitive to interactions with the constant region. Figure 2 depicts the calculated viscosity as a function of A_H^c for $A_H^v = 0.4$ kcal/mol and $A_H^v = 1.0$ kcal/mol.

Effect of the system size on viscosity

The simulation system can be adjusted by the number of molecules and box size to reach a target concentration. It is imperative to examine the effect of system size on viscosity because the viscosity calculations depend on the cluster size distribution. Figure 3 reports the viscosity at 150 mg/mL using three different numbers of molecules in the system as a function of A_H^v.

Determination of the best model for viscosity

The characterization in the previous sections facilitates the parameter tuning to calculate viscosity of therapeutic mAbs. We selected 20 immunoglobulin G1 (IgG1) mAbs from our previous work with experimental data in a wide range of concentrations.[22] As a benchmark, the model developed by Wang et al. was used to calculate the viscosity.[18] In the original model, all the beads had the same Hamaker constants (0.4 kcal/mol). Figure 4 (a) shows the result of the original model.

In this study, we attempted to find a mapping function (or the scaling parameter α) to relate HVI to the short-range interaction parameters on the Fv region (orange spheres in Figure 5).

$$A_H^v = \alpha \times \mathrm{HVI} \quad (1)$$

Another goal was to find a short-range interaction parameter for the constant region A_H^c that applies to all the mAbs (gray spheres in Figure 5).

We performed a grid search approach and assigned α from 0.02 kcal/mol to 0.06 kcal/mol with a step size of 0.01 kcal/mol, and A_H^c from 0.1 kcal/mol to 0.5 kcal/mol with a step size of

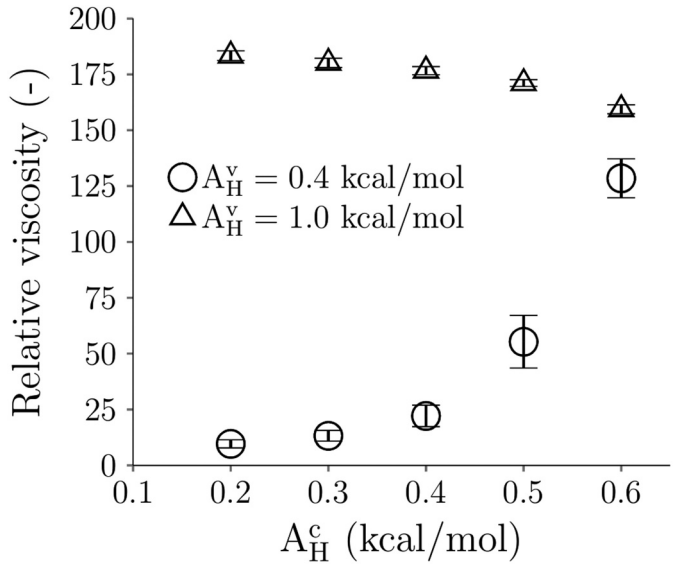

Figure 2. Viscosity as a function of the Hamaker constants on the constant region (A_H^c) for small and large A_H^v values at 150 mg/mL. The error bars indicate standard deviation. mAb4 is used as a template.

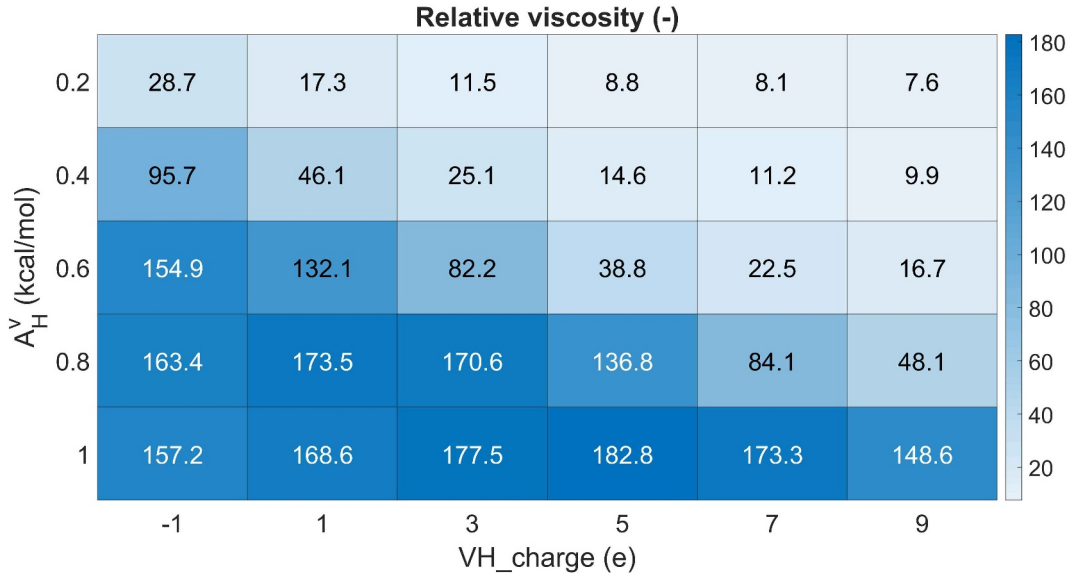

Figure 1. A viscosity heatmap as a function of the charges on the heavy chain variable region (VH) and the Hamaker constants on the Fv region (A_H^v). The Hamaker constant on the constant region is 0.4 kcal/mol. mAb4 at 150 mg/mL is used for analysis.

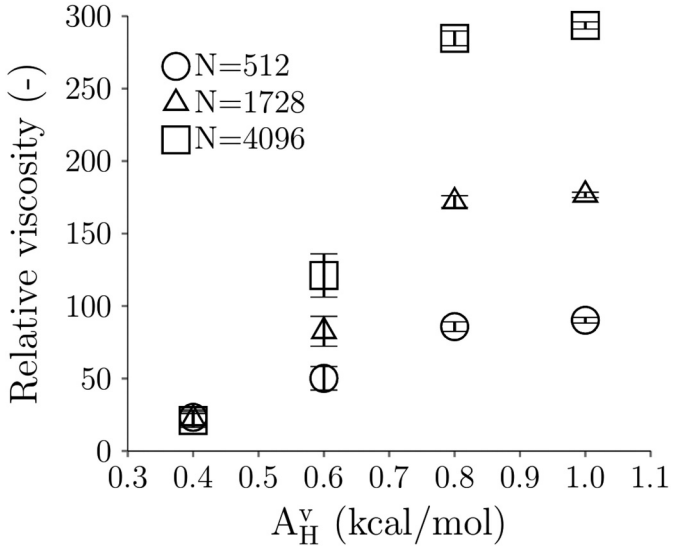

Figure 3. Viscosity as a function of A_H^v for different numbers of molecules in the system at 150 mg/mL. The error bars indicate standard deviation. mAb4 is used as a template. The A_H^c is equal to 0.4 kcal/mol.

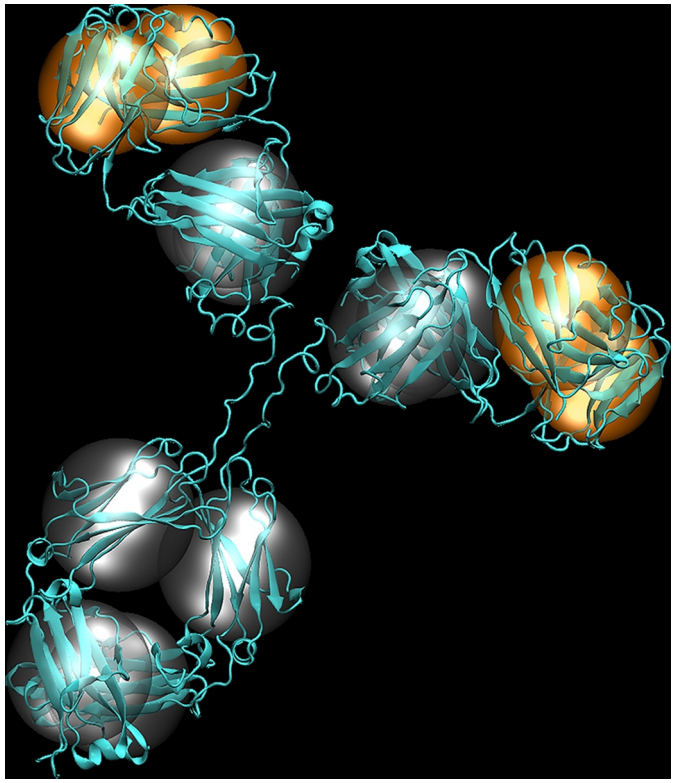

Figure 5. The 12-bead CG model of antibodies. The cartoon representation shows the all-atom model. In the CG model, each antibody domain is represented by one sphere. The orange spheres represent the variable region, and the gray spheres represent the constant region.

0.1 kcal/mol. The objective is to minimize the error of the logarithm of viscosity at 150 mg/mL between models and experiment for the 20 IgG1 mAbs. The best parameter set we found was $\alpha = 0.04$ kcal/mol and $A_H^c = 0.2$ kcal/mol. Table 1 summarizes the key parameters for the CG models of the 20 mAbs.

Figure 4 (b) reports the results of the new model. The same plot on a linear scale is also provided in supporting information (Figure S1). The root-mean-square deviation (rmsd) improved from 1.00 (-) to 0.68 (-), and the linear correlation coefficient improved from 0.63 to 0.87.

Concentration dependence of the viscosity model

The CG model can be applied to calculate mAbs viscosity at different concentrations. The parameters are the same as that at 150 mg/mL without the need for refitting. Figure 6 presents the viscosity for the 20 mAbs at 50, 100, 125 and 150 mg/mL. The same plot on a linear scale is also provided in supporting information (Figure S2).

The overall performance of the original model and the new model for the 20 mAbs at different concentrations is summarized in Table 2.

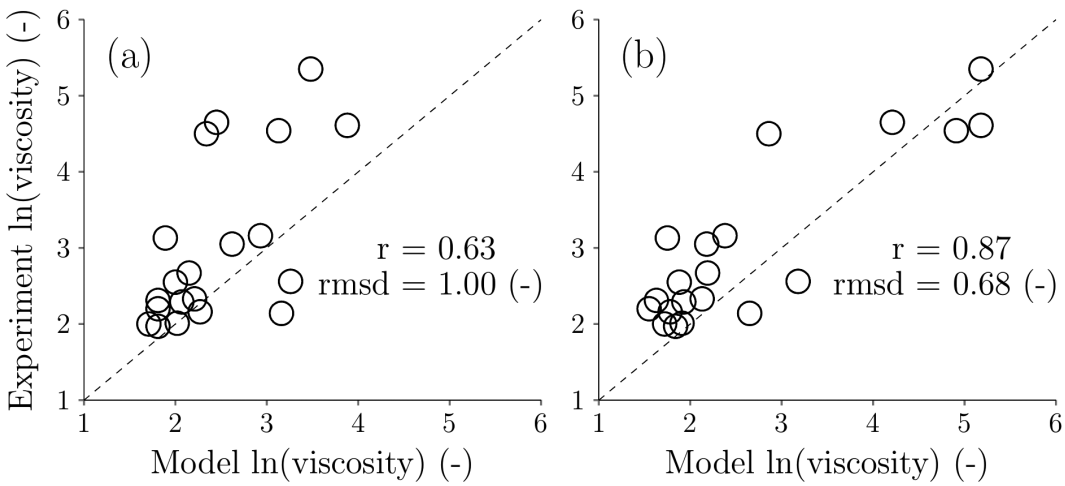

Figure 4. Comparison of the viscosity model with the experimental data at 150 mg/mL for (a) the original model where $A_H^v = A_H^c = 0.4$ kcal/mol and (b) the new model where $A_H^v = 0.04 \times$ HVI (kcal/mol) and $A_H^c = 0.2$ kcal/mol.

Table 1. Parameters for the CG model of the 20 mAbs in this study. CH1, CH2 and CH3 indicate heavy chain constant regions 1, 2 and 3, respectively. CL indicates light chain constant region. $A_H^v = 0.04 \times$ HVI (kcal/mol).

	charge (e)							Hamaker constant (kcal/mol)	
system	VH	CH1	CH2	CH3	VL	CL	Total	A_H^v	A_H^c
mAb1	−1.0	4.5	4.5	0.0	6.0	−1.0	26	0.67	0.2
mAb2	1.0	4.5	4.5	0.0	2.0	−1.0	22	0.50	0.2
mAb4	3.0	4.5	4.5	0.0	0.0	−2.0	20	0.81	0.2
mAb5	1.0	5.5	4.5	0.0	2.0	0.0	26	0.40	0.2
mAb8	−1.0	5.5	4.5	0.0	2.0	0.0	22	0.58	0.2
mAb10	4.0	6.5	4.5	0.0	2.0	0.0	34	0.53	0.2
mAb11	0.0	4.5	4.5	0.0	0.0	−1.0	16	0.95	0.2
mAb12	3.0	5.5	4.5	1.0	2.0	−1.0	30	0.59	0.2
mAb14	3.0	5.5	4.5	−1.0	1.0	−1.0	24	0.54	0.2
mAb15	3.0	5.5	4.5	0.0	3.0	0.0	32	0.67	0.2
mAb16	2.0	4.5	4.5	0.0	2.0	0.0	26	0.61	0.2
mAb17	3.0	4.5	4.5	0.0	−2.0	0.0	20	0.86	0.2
mAb18	3.0	4.5	4.5	0.0	3.0	−1.0	28	0.60	0.2
mAb21	1.0	4.5	4.5	−1.0	3.0	0.0	24	0.69	0.2
mAb22	3.0	5.5	4.5	0.0	3.0	−1.0	30	0.52	0.2
mAb23	4.0	4.5	4.5	0.0	4.0	−2.0	30	0.69	0.2
mAb24	1.0	5.5	4.5	0.0	2.0	0.0	26	0.69	0.2
mAb25	5.0	4.5	4.5	0.0	3.0	0.0	34	0.88	0.2
mAb26	2.0	4.5	4.5	0.0	3.0	−1.0	26	0.69	0.2
mAb27	0.0	5.5	4.5	0.0	3.0	0.0	26	0.77	0.2

Additionally, the analysis of cluster distribution and radial distribution function of high and low viscosity mAbs are discussed in the supporting information.

Tabulated viscosity for antibody screening

Tabulated viscosity values at 150 mg/mL based on our model are calculated in Table 3. The solution condition is at pH 6.0 in 10 mM histidine hydrochloride. In this table, three inputs, VH_charge, VL_charge and A_H^v, are needed. These inputs require only the sequence information of the Fv region. The constant region is based on mAb4, which is an IgG1 antibody. Therefore, this table should be used for IgG1 mAbs. We assumed small sequence variation in the IgG1 constant region does not affect the viscosity.

Discussion

From Figure 1, the general trend is that viscosity increases with more negative charge and higher A_H^v values. For mAbs in low viscosity conditions, the viscosity increases exponentially or follows a power-law function with amount of negative charge or Hamaker constants. The viscosity depends in a more complicated way at lower net charge and higher A_H^v region. For example, the relative viscosity reduces from 173.5 (-) to 163.4 (-) as VH_charge changes from 1 e to −1 e when $A_H^v = 0.8$ kcal/mol. From radial distribution function analysis, the negatively charged beads interact with other positive beads on the constant regions, which disrupt the interactions on the variable regions in high viscosity conditions.

In therapeutic antibody development, the syringeability limit is approximately 30 centipoise (cP) for subcutaneous injection.[23] Figure 1 covers a wide range of net charges, from +12 e to +32 e. It is feasible to reduce the viscosity below the syringeability limit by increasing the net positive charges even when the short-range interactions favor self-association at high-protein concentrations because the repulsive interactions of the whole antibody prevent protein association.

From Figure 2, the smaller A_H^v value indicates weak interactions between the variable regions. For $A_H^v = 0.4$ kcal/mol, the viscosity increases exponentially or follows a power-law function with the A_H^c values from 0.2 kcal/mol to 0.6 kcal/mol. On the contrary, for $A_H^v = 1.0$ kcal/mol, the viscosity is high when $A_H^c = 0.2$ kcal/mol and decreases slightly as A_H^c increases. It is noted that the short-range interactions between two beads take a form of $\sqrt{A_H^i \times A_H^j}$. When the short-range interactions between the variable regions are weak, increasing the short-range interactions in the constant region will enhance the association between variable to constant, and constant to constant regions, resulting in large cluster formation and higher viscosity. However, when the short-range interactions between the variable regions are strong, large clusters can form by connecting the two arms of mAbs. Under this circumstance, the increased A_H^c values could lead to the interference of the cluster formation through the variable regions. As A_H^c values increase from 0.2 kcal/mol to 0.6 kcal/mol when $A_H^v = 1.0$ kcal/mol, the average pairwise distance between beads decreases, indicating that the large cluster is more compact. In addition, radial distribution function shows that peaks for light chain variable region (VL)-heavy chain constant region (CH3) and VH-CH3 also increase, indicating head-to-tail bindings, leading to more branched structures and less elongated structures.[24]

In Figure 3, when $A_H^v = 0.4$ kcal/mol, the three systems that exhibit low viscosity give very similar results. As A_H^v increases to 0.6 kcal/mol, a difference between the viscosity is observed. Larger systems exhibit significantly higher viscosity compared to that of the smaller systems. The viscosity gaps between different system sizes further increase with higher A_H^v values. Furthermore, the viscosity only increases marginally when $A_H^v = 1.0$ kcal/mol compared to when $A_H^v = 0.8$ kcal/mol for the three systems.

As far as we know, the system size effect on viscosity has not been investigated. Previous works characterize their viscosity using only one system size. However, the viscosity calculation is based on the cluster size. For example, consider an extreme case where all molecules form a single huge cluster. Equation 18 indicates the hydrodynamic stress is scaled by the system volume; however, the term in the summation does not scale linearly with the cluster size. Therefore, we observe a system size dependence of the viscosity. For strongly interacting molecules, the cluster should extend beyond the system box due to the periodic boundary condition. However, the hydrodynamic calculation can only consider the clusters within the simulation box. For $A_H^v = 0.4$ kcal/mol, the molecular interactions are relatively weak. Most of the clusters are either monomers or dimers, and the cluster size distribution is similar for all system sizes. As a result, there is no system size dependence of the viscosity. In contrast, for larger A_H^v values, there is a single huge cluster occupying the simulation box, so the cluster size distribution is not the same for different system sizes. As a consequence, the viscosity depends on the system sizes.

The goal of this work was to tune the parameters to calculate low and high viscosity of mAbs. The high viscosity mAbs from

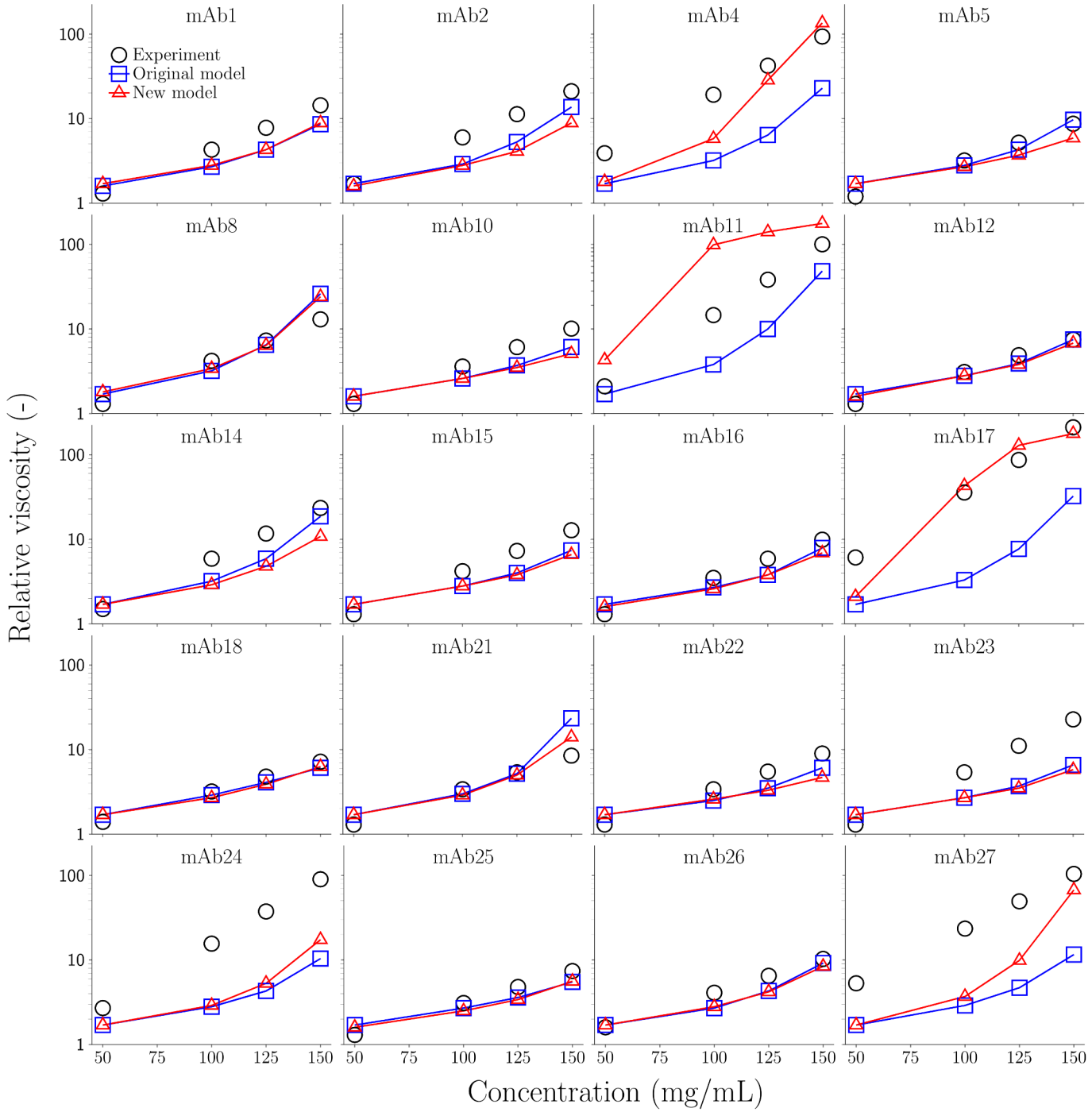

Figure 6. Concentration dependence of the relative viscosity with the experimental data for the original model where $A_H^v = A_H^c = 0.4$ kcal/mol and the new model where $A_H^v = 0.04 \times$ HVI (kcal/mol) and $A_H^c = 0.2$ kcal/mol. Black circles indicate experimental measurement. Blue squares indicate results from the original model. Red triangles indicate results from the new model.

Table 2. RMSD of the logarithm of relative viscosity for the previous work and this work at different concentrations.

	RMSD ln(viscosity) (-)	
Concentration	Previous work	This work
50	0.49	0.48
100	1.01	0.83
125	1.13	0.82
150	1.00	0.68

Table 3. Tabulated relative viscosity values at 150 mg/mL as a function of charges on the heavy and light chain variable regions and A_H^v. The charges have a unit of e, and A_H^v has a unit of kcal/mol. The constant region of mAb4 was used as a template.

		VH_charge							VH_charge				
	A_H^v=0.4	−2	0	2	4	6		A_H^v=0.5	−2	0	2	4	6
VL_charge	−2	71.4	38.3	24.4	16.1	10.7	VL_charge	−2	104.4	71.9	41.5	25.4	16.1
	0	34.2	19.4	11.7	8.2	6.4		0	60.3	32.3	16.4	9.9	7.3
	2	18.7	11.5	7.3	5.7	5.4		2	30.4	14.7	8.4	6.2	5.8
	4	11.9	6.6	5.4	4.8	4.8		4	16.1	8.0	5.6	5.2	4.9
	6	7.7	5.8	4.9	4.8	4.8		6	10.1	6.1	5.2	4.9	4.7
		VH_charge							VH_charge				
	A_H^v=0.6	−2	0	2	4	6		A_H^v=0.7	−2	0	2	4	6
VL_charge	−2	127.3	118.4	84.3	46.9	25.2	VL_charge	−2	145.6	151.0	145.2	112.6	64.6
	0	106.1	72.0	30.3	14.2	8.6		0	150.6	141.4	84.9	23.7	12.4
	2	72.6	25.5	10.7	6.8	5.9		2	138.1	66.5	16.4	8.6	6.5
	4	28.8	10.9	6.7	5.4	5.0		4	77.9	17.4	7.8	5.9	5.1
	6	14.4	7.0	5.5	5.0	4.9		6	28.7	8.7	5.8	5.3	5.2
		VH_charge							VH_charge				
	A_H^v=0.8	−2	0	2	4	6		A_H^v=0.9	−2	0	2	4	6
VL_charge	−2	155.5	168.4	174.9	171.1	142.5	VL_charge	−2	154.4	169.1	175.8	180.8	181.9
	0	167.4	176.0	160.1	71.1	21.3		0	170.7	177.1	183.3	158.0	59.7
	2	172.4	150.4	43.9	11.8	7.4		2	176.6	182.0	136.4	19.4	9.6
	4	162.1	46.3	10.4	6.4	5.5		4	182.8	141.1	17.3	7.4	6.2
	6	92.4	12.7	6.6	5.3	5.2		6	168.6	28.1	7.7	5.6	5.4

the experimental data exhibit viscosity up to 200 cP. Therefore, we chose to use N = 1728 for all the characterization.

As can be seen in Figure 4 (a), the original model underestimates the high viscosity mAbs significantly. This is consistent with the observations from Wang et al.[18] There are different types of short-range interactions of proteins, including hydrogen-bonding, dipole–dipole interactions, and hydrophobic interactions.[20] However, the original model assumed that all beads have the same short-range interactions, so the viscosity only depends on the charge distribution on the beads. Assigning the short-range interaction parameters for the CG model is still challenging. Based on our previous work,[22] we found a key parameter, HVI, which is a low/high viscosity classifier and it only depends on the sequence on the Fv region. It is noted that the CG model cannot model all short-range interactions explicitly. The van der Waals form is used to model the effective short-range interactions.

The new model for the high viscosity mAbs is considerably enhanced. The A_H^v values range from 0.4 kcal/mol to 0.95 kcal/mol (Table 1). In the original model, the assignment of the Hamaker constant to 0.4 kcal/mol for the variable region apparently underestimates the short-range interactions.

In Figure 6, there are five high viscosity mAbs (mAb4, mAb11, mAb17, mAb24, mAb27). The original model significantly underestimates the viscosity at all concentrations. For mAb4 and mAb17, the new model agrees with the experimental viscosity very well. For mAb11, the new model overestimates the viscosity at concentrations >100 mg/mL. It is worth noting that mAb11 has the lowest net charge (+16 e). From the sequence analysis, we found that mAb11 exhibits many negative charges on the framework regions rather than on the surface-exposed complementarity-determining regions. The 12-bead model does not capture the charge distribution within each region. This may overestimate the negative charge contribution on the bead. Future work should address the effect of different charge distribution on the CG models. It is noted that the negative curvature for mAb11 and mAb17 at high concentrations is due to a finite size effect that clusters exceed the box size. For mAb27, the new model slightly underestimates the viscosity; however, it still performs better than that of the original model. For mAb24, the new model did not describe the high viscosity. It is noted that mAb24 has the smallest A_H^v among the high viscosity mAbs, and the short-range interactions are not strong enough to overcome the charge repulsion.

For low viscosity mAbs, they generally have smaller A_H^v values. One exception is mAb25, which has a high A_H^v value (0.88 kcal/mol). However, it also has a high net charge (+34 e). The strong charge repulsion prevents the protein association despite favorable short-range interactions. Interestingly, our previous work suggests filtering out the mAbs with high net charges first and apply HVI to classify low/high viscosity mAbs. The CG model takes into account both effects.

The rmsd at concentrations of 50, 100, 125 and 150 mg/mL of the new model were all better than that of the original model (Table 2). The improvement is more significant at higher concentrations.

The computational time for CG simulations is still too long to allow screening efficiently for over hundreds of antibody drug candidates in early-stage design. We provided tabulated viscosity values at 150 mg/mL based on our model. Taking mAb4 and mAb5 for example, the parameters (VH_charge, VL_charge, A_H^v) are (3,0,0.81) and (1,2,0.40), respectively (Table 1). By linear interpolation from Table 3, the estimated relative viscosities are 115.6 (-) and 9.4 (-), respectively. The experimental results are 93.7 (-) and 8.7 (-), respectively. The tabulated relative viscosity from our model will facilitate antibody screening in the early-stage design. The CG simulations can be extended to other formulation and solution conditions to build new tables for screening.

In conclusion, here we combined a 12-bead CG model for antibodies, hydrodynamic calculations and the HVI parameters to calculate antibody viscosity over a wide range of concentrations. Two adjustable parameters were determined by fitting to 20 IgG1 viscosity data. The comparison of a large number of experimental data provides credibility to this model. This model is able to calculate viscosity of new mAbs without

additional fitting parameters. A viscosity table was provided that requires only three sequence-based inputs to estimate the high concentration viscosity. This approach facilitates fast screening in early-stage design. For future work, this model can be extended to other solution conditions, isotypes and antibody formats to investigate the viscosity behavior and study protein–protein interactions at the molecular level.

Materials and methods

Homology modeling of mAbs

The sequences of the 20 mAbs used in this study were obtained from a previous work.[22] The mAb molecules were constructed followed the protocol proposed by Brandt et al.[25] The antigen-binding fragment (Fab) structures were obtained from either available crystal structures or homology models built from RosettaAntibody.[26–28] The structures of the Fab regions were superimposed on a template structure of a full-length IgG1 model. The IgG1 template was obtained from the KOL/Padlan structure.[29,30] The glycan structure was G0F. The homology models were energy-minimized to remove steric clashes in protein structures using NAMD2.[31]

All-atom molecular dynamics simulation

Molecular dynamics (MD) simulations were performed using all-atom structures with explicit solvent using the TIP3P water model.[32] Simulation boxes were set up using VMD[33] to place a single antibody in a water box extending 12 Å beyond the protein surface. Simulations were performed at 300 K and 1 atm in the NPT ensemble, using the NAMD2[31] software package and the CHARMM36m force field.[34] The system pH was set to 6.0 to match the experimental pH by adjusting the protonation states of the histidine residues using the PROPKA3 protocol.[35] Electrostatic interactions were treated with the Particle Mesh Ewald method and van der Waals interactions were calculated using a switching distance of 10 Å and a cutoff of 12 Å. The integration time step was set to 2 fs. Each mAb system was pre-equilibrated for 10 ns, followed by 50 ns production runs.

Construction of coarse-grained models

A 12-bead CG model derived from all-atom MD simulations was used to construct the structure of all the mAbs in this work (Figure 5).[15] The 12 beads include two heavy chains (VH, CH1, CH2 and CH3) and two light chains (VL and CL). It is noted that although the glycans are not incorporated in the CG structure. The presence of the glycans in the all-atom model will affect the interactions and distance between each domain in the MD simulations, which will subsequently affect the structure and conformation of the CG model. The intramolecular interaction parameters were obtained from the dynamic averages of the MD simulations. The intramolecular interactions consist of bond, angle, Urey-Bradley (UB) and dihedral interactions.

$$E_{bond} = \frac{1}{2}k_{bond}(r - r_0)^2 \quad (2)$$

$$E_{angle} = \frac{1}{2}k_{angle}(\theta - \theta_0)^2 + \frac{1}{2}k_{UB}(r - r_{UB})^2 \quad (3)$$

$$E_{dihedral} = \frac{1}{2}k_{dihedral}(1 - \cos(\phi)) \quad (4)$$

The equilibrium values for bond, angle and UB were obtained from the mean of the coordinate displacement and their spring constants were calculated from the variance of the coordinate fluctuation.

$$k_{spring} = \frac{k_B T}{2\sigma^2} \quad (5)$$

where k_B, T and σ^2 are the Boltzmann constant, temperature and the variance of the coordinate fluctuation, respectively. The spring constants for the dihedrals are 200 (kcal/mol).

The charges on the beads are the sum of all the atomic charges on the corresponding domains. The radii of the beads were calculated from the average radius of gyration from MD simulations. The radius of gyration (R_g) is calculated using the formula

$$R_g = \sqrt{\frac{\sum_{i=1}^{N} m_i r_i^2}{\sum_{i=1}^{N} m_i}} \quad (6)$$

where m_i is the mass of the atom i, r_i is the distance of atom i from the center of mass of the bead. The radius of the bead averaged from simulations is summarized in Table S1.

Brownian dynamics simulation and hydrodynamic interactions

The CG models were treated as colloidal particles as applied in a previous work.[18] Brownian dynamics simulation was used to describe the dynamic behaviors of the CG beads. For any CG particle, its motion satisfies the Langevin equation

$$m\frac{dU}{dt} = F^H + F^B + F^P \quad (7)$$

where m is the mass of the bead, and U is the translational velocity vector. F^H is the hydrodynamic force, generated by the drag of solvents on the particles. F^B is the stochastic force, caused by the Brownian motion of the particles. F^P is the deterministic nonhydrodynamic force, produced by the interactions between particles.

The hydrodynamic force F^H in the overdamped regime is linear in the bead velocity, $F^H = -\mathcal{R} \cdot U$, where $\mathcal{R}$ is the

Table 4. Parameters used in CG simulation.

parameter	value
hydrodynamic radius	2 nm
time step	2.17 ps
simulation time	2.17 μs
number of CG antibodies in a box	512, 1728, 4096
relative permeability	80
antibody concentration	50, 100, 125, 150 mg/mL
Debye length	3.08 nm

hydrodynamic resistance tensor. Neglecting the inertia term, the Langevin equation reads as

$$0 = -\mathcal{R} \cdot \mathrm{U} + \mathrm{F^B} + \mathrm{F^P} \tag{8}$$

$$\frac{\mathrm{dx}}{\mathrm{dt}} = \mathcal{M} \cdot \left(\mathrm{F^B} + \mathrm{F^P}\right) \tag{9}$$

where $\mathcal{M} = \mathcal{R}^{-1}$ is the mobility tensor, and x is the vector of particle positions.

Following a previous work, the mobility tensor is approximated by the Rotne-Prager-Yamakawa tensor.[36] A recently developed algorithm, positively split Ewald method,[37] was applied to solve the Langevin equation coupled with hydrodynamic calculations with the RPY mobility tensor. The PSE method is implemented as a plugin to HOOMD-blue,[38] a general-purpose particle simulation toolkit. The box length was adjusted to reach different concentrations. A list of parameters used in the CG model and the hydrodynamic calculations is summarized in Table 4 and Table S1 to Table S4.

Force fields for coarse-grained models

The force fields for the CG models were modified from a previous work.[18] The electrostatic interactions between the beads were approximated as a form of the Yukawa potential

$$\mathrm{E_{elec}} = \left(\frac{\mathrm{q_1 q_2}}{4\pi\epsilon_\mathrm{r}\epsilon_0}\right)\left(\frac{\exp(\kappa \mathrm{a_1})}{1+\kappa \mathrm{a_1}}\right)\left(\frac{\exp(\kappa \mathrm{a_2})}{1+\kappa \mathrm{a_2}}\right)\frac{\exp(-\kappa \mathrm{r})}{\mathrm{r}} \tag{10}$$

where $\mathrm{q_1}$, $\mathrm{q_2}$ and $\mathrm{a_1}$, $\mathrm{a_2}$ are the charges and radii of the two beads, respectively. κ^{-1} is the Debye length, which is 3.08 nm in 10 mM histidine-HCl buffer. ϵ_0 is the vacuum permeability. ϵ_r is the relative permeability, which is 80 in this work.

The dispersion interactions were described as van der Waals interactions[39]

$$\mathrm{E_{vdW}} = -\frac{\mathrm{A_H}}{6}\left[\frac{2\mathrm{a_1 a_2}}{\mathrm{r^2}-(\mathrm{a_1}+\mathrm{a_2})^2} + \frac{2\mathrm{a_1 a_2}}{\mathrm{r^2}-(\mathrm{a_1}-\mathrm{a_2})^2} + \ln\frac{\mathrm{r^2}-(\mathrm{a_1}+\mathrm{a_2})^2}{\mathrm{r^2}-(\mathrm{a_1}-\mathrm{a_2})^2}\right] \tag{11}$$

where $\mathrm{A_H}$ is the Hamaker constant, and $\mathrm{a_1}$ and $\mathrm{a_2}$ are the radii for the two particles. The Hamaker constant for all the 12 beads was set to 0.4 kcal/mol in the previous work.[18] In this study, $\mathrm{A_H}$ is further divided into $\mathrm{A_H^v}$ and $\mathrm{A_H^c}$ to account for variable and constant regions of the antibodies. The cross-term for different types of beads takes the form of a geometric mean

$$\mathrm{A_{H,ij}} = \sqrt{\mathrm{A_{H,i}} \times \mathrm{A_{H,j}}} \tag{12}$$

The primary goal of this study is to determine the values of $\mathrm{A_H^v}$ and $\mathrm{A_H^c}$. For $\mathrm{A_H^c}$, it is the same for all mAbs of the IgG1 isotype. For $\mathrm{A_H^v}$, we applied a mapping function to assign the values based on HVI previously proposed

$$\mathrm{A_H^v} = \alpha \times \mathrm{HVI} \tag{13}$$

$$\mathrm{HVI} = \frac{\mathrm{N\_hydrophilic\ residues\ in\ Fv} - \mathrm{N\_hydrophobic\ residues\ in\ Fv}}{\mathrm{N\_residues\ in\ Fv}} \times 100 \tag{14}$$

The scaling parameter α is determined by minimizing the error of the viscosity between the model and experiment at 150 mg/mL for 20 IgG1 mAbs.[22,23] The details are discussed in the Results and Discussion sections.

Viscosity calculation

The total viscosity η is contributed from hydrodynamic (η_h), interparticle (η_p) and solvent (η_s) interactions[18]

$$\eta = \eta_\mathrm{h} + \eta_\mathrm{p} + \eta_\mathrm{s} \tag{15}$$

In this work, all the viscosity values are expressed as the relative viscosity (η/η_s).

$$\eta/\eta_\mathrm{s} = \eta_\mathrm{h}/\eta_\mathrm{s} + \eta_\mathrm{p}/\eta_\mathrm{s} + 1 \tag{16}$$

In the hydrodynamic calculations, the solvent viscosity (η_s) was scaled so that the drag coefficient $6\pi\eta_\mathrm{s}\mathrm{a_h} = 1$.

The interparticle viscosity (η_p) was calculated from the Green-Kubo relation.[40,41]

$$\eta_\mathrm{p} = \frac{\mathrm{V}}{\mathrm{k_B T}}\int_0^\infty \langle \sigma_{\mathrm{xy}}(0)\sigma_{\mathrm{xy}}(\mathrm{t})\rangle \mathrm{dt} \tag{17}$$

where V is the volume of the simulation box. σ_{xy} is the off-diagonal stress. The ensemble average of the auto-correlation function was averaged from the three off-diagonal components (xy, xz and yz).

The hydrodynamic viscosity (η_h) was calculated from static structures of antibody rigid clusters. Two antibody molecules are identified as connected if the nearest distance between any of their beads is shorter than a cutoff length of $1.65\mathrm{a_h}$. The cutoff distance was determined from the first minimum after the highest peak of the radial distribution function.[18] This includes molecules within the first ring of nearest neighbors. The viscosity was obtained by measuring the hydrodynamic stress (σ^H) when the rigid bodies are immersed in an imposed linear flow with the strain rate tensor e.

$$\sigma^\mathrm{H} = -\frac{1}{\mathrm{V}}\sum_\mathrm{N}\left(\begin{bmatrix}\mathrm{R_{SU}} & \mathrm{R_{S\Omega}}\end{bmatrix}\cdot\begin{bmatrix}\mathrm{R_{FU}} & \mathrm{R_{F\Omega}}\\ \mathrm{R_{TU}} & \mathrm{R_{T\Omega}}\end{bmatrix}^{-1}\cdot\begin{bmatrix}\mathrm{R_{FE}}\\ \mathrm{R_{TE}}\end{bmatrix} - \mathrm{R_{SE}}\right)\cdot \mathrm{e} \tag{18}$$

The resistance tensors couple force (F), torque (T), translational velocity (U), angular velocity (Ω), stresslet (S) and strain (E). The solution for this equation is detailed in the previous work.[18] The ensemble averaged stress tensor in an isotropic medium is related to the strain rate tensor by the hydrodynamic contribution to the equation

$$<\sigma^\mathrm{H}> = 2\eta_\mathrm{h}\mathrm{e} \tag{19}$$

The hydrodynamic viscosity (η_h) was obtained by averaging the shear stress produced by different antibody configurations exposed to a simple shear flow via Equation 18.

Abbreviations

CG	coarse-grained
CH1	heavy chain constant region 1
CH2	heavy chain constant region 2
CH3	heavy chain constant region 3
CL	light chain constant region
cP	centipoise
Fab	antigen-binding fragment
Fv	variable fragment
HVI	high viscosity index
IgG1	immunoglobulin G1
mAbs	monoclonal antibodies
MD	molecular dynamics
RMSD	root-mean-square deviation
UB	Urey-Bradley
VH	heavy chain variable region

Acknowledgments

This work was partially funded by Merck. We thank Dr Gang Wang for providing the codes for hydrodynamic and viscosity calculation. We also appreciate the insightful discussion with members of Dr Swan's group.

Funding

This work was supported by the Merck.

References

1. Shire SJ. Formulation and manufacturability of biologics. Current Opinion in Biotechnology. 2009;20(6):708–14. doi:10.1016/j.copbio.2009.10.006.
2. Goswami S, Wang W, Arakawa T, Ohtake S. Developments and challenges for mAb-based therapeutics. Antibodies. 2013;2(3):452–500. doi:10.3390/antib2030452.
3. Zhang Z, Liu Y. Recent progresses of understanding the viscosity of concentrated protein solutions. Current Opinion in Chemical Engineering. 2017;16:48–55. doi:10.1016/j.coche.2017.04.001.
4. Tilegenova C, Izadi S, Yin J, Huang CS, Wu J, Ellerman D, Hymowitz SG, Walters B, Salisbury C, Carter PJ. Dissecting the molecular basis of high viscosity of monospecific and bispecific IgG antibodies. mAbs. 2020;12(1):1692764. doi:10.1080/19420862.2019.1692764.
5. Connolly BD, Petry C, Yadav S, Demeule B, Ciaccio N, Moore JMR, Shire SJ, Gokarn YR. Weak interactions govern the viscosity of concentrated antibody solutions: high-throughput analysis using the diffusion interaction parameter. Biophysical Journal. 2012;103(1):69–78. doi:10.1016/j.bpj.2012.04.047.
6. Woldeyes MA, Qi W, Razinkov VI, Furst EM, Roberts CJ. How well do low- and high-concentration protein interactions predict solution viscosities of monoclonal antibodies? Journal of Pharmaceutical Sciences. 2019;108(1):142–54. doi:10.1016/j.xphs.2018.07.007.
7. Hung JJ, Dear BJ, Karouta CA, Chowdhury AA, Godfrin PD, Bollinger JA, Nieto MP, Wilks LR, Shay TY, Ramachandran K, et al. Protein–protein interactions of highly concentrated monoclonal antibody solutions via static light scattering and influence on the viscosity. The Journal of Physical Chemistry B. 2019;123(4):739–55. doi:10.1021/acs.jpcb.8b09527.
8. Dear BJ, Chowdhury A, Hung JJ, Karouta CA, Ramachandran K, Nieto MP, Wilks LR, Sharma A, Shay TY, Cheung JK, et al. Relating collective diffusion, protein–protein interactions, and viscosity of highly concentrated monoclonal antibodies through dynamic light scattering. Industrial & Engineering Chemistry Research. 2019;58(50):22456–71. doi:10.1021/acs.iecr.9b03432.
9. Lai P-K, Kaznessis YN. Free energy calculations of microcin j25 variants binding to the fhua receptor. Journal of Chemical Theory and Computation. 2017;13(7):3413–23. doi:10.1021/acs.jctc.7b00417.
10. Pan AC, Jacobson D, Yatsenko K, Sritharan D, Weinreich TM, Shaw DE. Atomic-level characterization of protein–protein association. Proceedings of the National Academy of Sciences. 2019;116(10):4244–49. doi:10.1073/pnas.1815431116.
11. Jost Lopez A, Quoika PK, Linke M, Hummer G, Köfinger J. Quantifying Protein–Protein Interactions in molecular simulations. The Journal of Physical Chemistry B. 2020;124(23):4673–85. doi:10.1021/acs.jpcb.9b11802.
12. Hollingsworth SA, Dror RO. Molecular dynamics simulation for all. Neuron. 2018;99(6):1129–43. doi:10.1016/j.neuron.2018.08.011.
13. Grünberger A, Lai P-K, Blanco MA, Roberts CJ. Coarse-grained modeling of protein second osmotic virial coefficients: sterics and short-ranged attractions. The Journal of Physical Chemistry B. 2013;117(3):763–70. doi:10.1021/jp308234j.
14. Baaden M, Marrink SJ. Coarse-grain modelling of protein-protein interactions. Current Opinion in Structural Biology. 2013;23(6):878–86. doi:10.1016/j.sbi.2013.09.004.
15. Chaudhri A, Zarraga IE, Kamerzell TJ, Brandt JP, Patapoff TW, Shire SJ, Voth GA. Coarse-grained modeling of the self-association of therapeutic monoclonal antibodies. The Journal of Physical Chemistry B. 2012;116(28):8045–57. doi:10.1021/jp301140u.
16. Chaudhri A, Zarraga IE, Yadav S, Patapoff TW, Shire SJ, Voth GA. The role of amino acid sequence in the self-association of therapeutic monoclonal antibodies: insights from coarse-grained modeling. The Journal of Physical Chemistry B. 2013;117(5):1269–79. doi:10.1021/jp3108396.
17. Calero-Rubio C, Saluja A, Roberts CJ. Coarse-grained antibody models for "weak" protein–protein interactions from low to high concentrations. The Journal of Physical Chemistry B. 2016;120(27):6592–605. doi:10.1021/acs.jpcb.6b04907.
18. Wang G, Varga Z, Hofmann J, Zarraga IE, Swan JW. Structure and Relaxation in Solutions of Monoclonal Antibodies. The Journal of Physical Chemistry. B. 2018;122(11):2867–80. doi:10.1021/acs.jpcb.7b11053.
19. Izadi S, Patapoff TW, Walters BT. Multiscale coarse-grained approach to investigate self-association of antibodies. Biophysical Journal. 2020;118(11):2741–54. doi:10.1016/j.bpj.2020.04.022.
20. Chowdhury A, Bollinger JA, Dear BJ, Cheung JK, Johnston KP, Truskett TM. Coarse-grained molecular dynamics simulations for understanding the impact of short-range anisotropic attractions on structure and viscosity of concentrated monoclonal antibody solutions. Molecular Pharmaceutics. 2020;17(5):1748–56. doi:10.1021/acs.molpharmaceut.9b00960.
21. Von BS, Siggel M, Linke M, Hummer G. Dynamic cluster formation determines viscosity and diffusion in dense protein solutions. Proceedings of the National Academy of Sciences. 2019;116(20):9843–52. doi:10.1073/pnas.1817564116.
22. Lai P-K, Fernando A, Cloutier TK, Gokarn Y, Zhang J, Schwenger W, Chari R, Calero-Rubio C, Trout BL. Machine learning applied to determine the molecular descriptors responsible for the viscosity behavior of concentrated therapeutic antibodies. Molecular Pharmaceutics. 2021;18(3):1167–75. doi:10.1021/acs.molpharmaceut.0c01073.
23. Kingsbury JS, Saini A, Auclair SM, Fu L, Lantz MM, Halloran KT, Calero-Rubio C, Schwenger W, Airiau CY, Zhang J, et al. A single molecular descriptor to predict solution behavior of therapeutic antibodies. Science Advances. 2020;6(32):eabb0372. doi:10.1126/sciadv.abb0372.
24. Ramallo N, Paudel S, Schmit J. Cluster formation and entanglement in the rheology of antibody solutions. J Phys Chem B. 2019;123(18):3916–23. doi:10.1021/acs.jpcb.9b01511.
25. Brandt JP, Patapoff TW, Aragon SR. Construction, MD simulation, and hydrodynamic validation of an all-atom model of a monoclonal IgG antibody. Biophysical Journal. 2010;99(3):905–13. doi:10.1016/j.bpj.2010.05.003.
26. Sivasubramanian A, Sircar A, Chaudhury S, Gray JJ. Toward high-resolution homology modeling of antibody Fv regions and application to antibody-antigen docking. Proteins. 2009;74(2):497–514. doi:10.1002/prot.22309.

27. Weitzner BD, Kuroda D, Marze N, Xu J, Gray JJ. Blind prediction performance of RosettaAntibody 3.0: grafting, relaxation, kinematic loop modeling, and full CDR optimization. Proteins. 2014;82(8):1611–23. doi:10.1002/prot.24534.
28. Weitzner BD, Jeliazkov JR, Lyskov S, Marze N, Kuroda D, Frick R, Adolf-Bryfogle J, Biswas N, Dunbrack RL, Gray JJ. Modeling and docking of antibody structures with Rosetta. Nature Protocols. 2017;12(2):401–16. doi:10.1038/nprot.2016.180.
29. Padlan EA. Anatomy of the antibody molecule. Molecular Immunology. 1994;31(3):169–217. doi:10.1016/0161-5890(94)90001-9.
30. Boehm MK, Woof JM, Kerr MA, Perkins SJ. The Fab and Fc fragments of IgA1 exhibit a different arrangement from that in IgG: a study by X-ray and neutron solution scattering and homology modelling. Journal of Molecular Biology. 1999;286(5):1421–47. doi:10.1006/jmbi.1998.2556.
31. Phillips JC, Hardy DJ, Maia JDC, Stone JE, Ribeiro JV, Bernardi RC, Buch R, Fiorin G, Hénin J, Jiang W, et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. The Journal of Chemical Physics. 2020;153(4):044130. doi:10.1063/5.0014475.
32. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. The Journal of Chemical Physics. 1983;79(2):926–35. doi:10.1063/1.445869.
33. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. Journal of Molecular Graphics. 1996;14(1):33–38. doi:10.1016/0263-7855(96)00018-5.
34. Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, De Groot BL, Grubmüller H, MacKerell AD. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nature Methods. 2017;14(1):71–73. doi:10.1038/nmeth.4067.
35. Olsson MHM, Søndergaard CR, Rostkowski M, Jensen JH. PROPKA3: consistent treatment of internal and surface residues in empirical pKa predictions. Journal of Chemical Theory and Computation. 2011;7(2):525–37. doi:10.1021/ct100578z.
36. Rotne J, Prager S. Variational Treatment of Hydrodynamic Interaction in Polymers. The Journal of Chemical Physics. 1969;50(11):4831–37. doi:10.1063/1.1670977.
37. Fiore AM, Balboa Usabiaga F, Donev A, Swan JW. Rapid sampling of stochastic displacements in Brownian dynamics simulations. The Journal of Chemical Physics. 2017;146(12):124116. doi:10.1063/1.4978242.
38. Anderson JA, Glaser J, Glotzer SC. HOOMD-blue: a Python package for high-performance molecular dynamics and hard particle Monte Carlo simulations. Computational Materials Science. 2020;173:109363. doi:10.1016/j.commatsci.2019.109363.
39. Hamaker HC. The London—van der Waals attraction between spherical particles. Physica. 1937;4(10):1058–72. doi:10.1016/S0031-8914(37)80203-7.
40. Green MS. Markoff random processes and the statistical mechanics of time-dependent phenomena. ii. irreversible processes in fluids. The Journal of Chemical Physics. 1954;22(3):398–413. doi:10.1063/1.1740082.
41. Statistical-Mechanical KR. Theory of irreversible processes. i. general theory and simple applications to magnetic and conduction problems. Journal of the Physical Society of Japan. 1957;12(6):570–86. doi:10.1143/JPSJ.12.570.