RSC Advances

PAPER

View Article Online
View Journal | View Issue

Cite this: *RSC Adv.*, 2025, **15**, 49374

The viscosity of protein and nucleic acid solutions and their folded structures explored using the free-volume concept and Eyring's rate process theory

Tian Hao *

This article aims to unify the understanding of protein and nucleic acid solution viscosity by integrating the free-volume concept and Eyring's rate process theory. The importance of controlling protein and nucleic acid solution viscosity in therapeutic formulations and manufacturing cannot be overstated, as numerous empirical and semi-empirical equations/models have been proposed to fit experimental data in the literature. These models are intended to extrapolate viscosity predictions at higher concentrations based on low-concentration data or provide guidance on how to reduce viscosity by adjusting pH and adding salt. However, none of these models can be universally applied to all systems, providing reasonable interpretations of experimental results. We borrow the reptation-tube concept from polymer science to treat the molecules of proteins and nucleic acids, and introduce the aspect ratio parameter to describe the fibrousness of the molecular shapes of the proteins and nucleic acids. The obtained equations can adequately correlate the viscosity with protein and nucleic acid volume fraction, salt concentration, zeta potential, pH, and temperature, and fit many experimental data very well. They show that the viscosity increases almost linearly with the volume fraction in low-volume-fraction regions, but increases dramatically with the volume fraction in high-volume-fraction regions; increases gradually with both zeta potential and the aspect ratio of the molecular chains; decreases with the square root of the ionic strength; reaches a minimum point with pH; and generally decreases with temperature, except in DNA solutions due to the transition from double-stranded to single-stranded molecules, *etc.* The viscosity of several protein and DNA solutions are regressed with our equations and very good agreements are obtained. Our work deepens the physical understanding of critical parameters, and provides clues for lowering viscosity in pharmaceutical formulations.

Received 14th October 2025
Accepted 3rd November 2025

DOI: 10.1039/d5ra07868g

rsc.li/rsc-advances

1. Introduction

Protein and nucleic acids are vital biomolecules that have a profound effect on biological processes, with their varied functions being closely tied to their three-dimensional structures. The sequence of amino acids and the spatial arrangement of these components determine how they interact and assemble into single-stranded RNA or double-stranded DNA molecules or aggregates, which, in turn, influence the complex properties of proteins and nucleic acids. Additionally, the viscosity of protein solutions has been found to be a crucial physical parameter that can significantly impact protein behavior, stability, and functionality. Understanding the factors that govern protein viscosity is essential for developing innovative applications across various fields, including pharmaceutical formulation, biotechnology, and food science.

Protein solutions rely heavily on the structural integrity of proteins to maintain their functional capabilities and regulate viscosity. Current methods for determining protein structures, such as X-ray crystallography, NMR spectroscopy, and cryo-electron microscopy, have been widely used after proteins are crystallized under specific conditions that require significant effort. Nevertheless, these approaches can be time-consuming and may not always provide accurate information on the dynamic behaviors of proteins in solution.

Advances in computational methods have greatly enhanced the accuracy of predicting protein structures and dynamics. Novel approaches such as molecular dynamics simulations and machine-learning algorithms enable researchers to gain valuable insights into the folding mechanisms of proteins and how their structures adapt to environmental conditions. The Google AlphaFold2 and AlphaFold3 models have demonstrated exceptional precision in predicting protein structures, providing relevant information on the dynamic behaviors of proteins in solution.[1,2] These models employ diffusion-based neural networks that are constrained by chemical bond lengths and angles, including the stereochemistry of amino acids, their stereochemical properties, or the free space/volume availability in the system.

15905 Tanberry Dr., Chino Hills, CA 91709, USA. E-mail: haotian9@gmail.com

This correlation enables researchers to link folded protein structures with viscosity and the flow behaviors of protein solutions.

The excluded-volume concept has also been used to elucidate protein and nucleic acid conformational structures.[3,4] The excluded-volume effect refers to the spatial constraints imposed by protein molecules in a solution that, in turn, modulate their interactions with neighboring molecules. This phenomenon can result in alterations to protein conformation and dynamics, thereby influencing viscosity. The excluded-volume effect is particularly significant in crowded environments, such as cellular interiors, where proteins are densely packed. The concepts of excluded volume and free volume have been extensively utilized in the literature to explain glass transitions and flow behaviors of various materials.[5–13] While these concepts share similarities, subtle distinctions do exist, as discussed in this article.[14] The free-volume concept, which describes the unoccupied space within a material, may offer valuable insights into protein behaviors in a solution. Eyring's rate process theory,[15] which originally relates chemical reaction rates to energy barriers, may be applied to understand protein folding and the corresponding changes in viscosity with temperature and pressure. The free-volume concept posits that the viscosity of a protein solution is inversely related to the amount of unoccupied space available for proteins to move in. When the temperature increases, the free volume expands, facilitating increased protein mobility and reduced viscosity. Conversely, elevated pressure diminishes the free volume, leading to enhanced viscosity.

The viscosity of protein and nucleic acid solutions is influenced by multiple factors, including protein structure, concentration, molecule size and shape, pH, ionic strength, temperature, and the presence of other solutes. At low concentrations, protein and nucleic acid molecules may behave like free-flowing entities, where their viscosity is primarily dominated by their physical dimensions and morphology. In contrast, as the concentration increases, the intermolecular interactions become more pronounced, leading to non-ideal behaviors and elevated viscosities.[16] The size and shape of proteins and nucleic acids can be influenced by pH, ionic strength, and temperature. Recent theoretical explorations of the crowding effect and the influence of salts on viscosity highlight the limitations of traditional theories and equations.[17] Solvent viscosity influences protein folding rates as well.[18] Salts in general can reduce the viscosity,[19] and a neutral pH may lead to the lowest viscosity.[20] Aggregation sizes of protein molecules should also be taken into account for protein formulation and viscosity control.[21]

Semi-empirical equations based on Arrhenius equations have been utilized to understand the viscosity dependence on temperature to extract insightful activation energy information.[22] Empirical equations based on the excluded-volume concept were proposed to understand the viscosity dependence of protein volume fraction.[22–24] The empirical viscosity exponential growth equation, the modified 3-parameter exponential equation, the modified Ross–Minton equation,[24] Einstein's equation, Krieger–Dougherty's equation,[25] and the Tomar equation,[26] have been evaluated in previous studies.[19,22,27,28] Limitations were identified for all these equations due to their empirical nature. Protein viscosity cannot be fitted with the hard-sphere model,[28] or with many other empirical models,[27] in concentrated solutions. In general, protein solutions are considered pure liquids if they are fully dissolved, or colloidal suspensions if they are partially dissolved. Hao has compared many approaches and equations to describe the viscosities of pure liquids, colloidal suspensions, and polymeric systems, and found that the equations derived based on the free-volume concept and Eyring's rate process theory can give satisfactory fits to experimental data.[29] The free-volume concept and Eyring's rate process theory will be integrated again to describe the viscosity of protein solutions, as this approach has been successfully applied previously in many fields, across a wide range of scales from electrons to the universe.[14,30–37] Although Eyring's rate process theory and excluded-volume quantification have been used to explore protein folding and even collapse from protein solutions,[38] they were applied independently in biological systems,[39] and there are no reports on integrating these two theories to treat biological systems. An attempt is made in this article to describe the viscosities of protein solutions with the integration of Eyring's rate process theory and free-volume concept, building on our previous work. The goal is to create a unified theoretical approach for better describing the viscosity of protein solutions.

2. Theory

At the molecular level, nucleic acids are composed of nucleotides with three components: a phosphate group, a five-carbon sugar (deoxyribose in DNA and ribose in RNA), and a nitrogenous base. The phosphodiester backbone is formed by the interaction between the phosphate groups of two adjacent nucleotides. On the other hand, proteins are made from amino acids with a central carbon atom bonded to an amino group, a carboxyl group, a unique side chain, and hydrogen atoms. Both nucleic acids and proteins have fibrous chain structures with a variety of differently charged regions. Typically, they are highly charged, and their folded structures are therefore influenced by the electrolytes, pH, and other additives that can bind with them. Analogously, they are comparable to fiber-shaped polymer molecules, but have a much lower molecular weight. We then can borrow the ideas used to treat polymer molecules to model the viscosities of proteins and nucleic acids. We will start with the approaches used for theoretically modeling the viscosity of polymers, and expand to charged situations to mimic the behavior of nucleic acids and proteins.

In an article published in 2008,[29] viscosities of pure liquids, colloidal suspensions, and polymeric systems were addressed using various methods, including the free-volume concept and Eyring's rate process theory. The current paper borrows the same ideas/approaches from this article to describe the viscosities of protein and nucleic acid solutions with a few tweaks. The first tweak is that we treat protein and nucleic acid molecules as fibers instead of spherical particles. The second tweak is that we place a significant emphasis on the electrostatic interactions among these fibrous protein molecules and the solvent molecules. Nucleic acid molecules are made from nucleotides, which have different amounts of charge at various

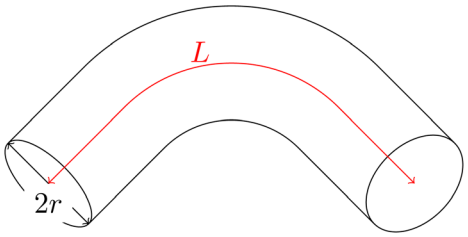


Fig. 1 A single protein molecular chain illustrated as a tube with radius r and length L.

sites, either positive or negative. These charges can interact with each other to form single-stranded or double-stranded structures. Protein and nucleic acid molecules may behave similarly in terms of their interactions with the surrounding solvent molecules, creating various aggregated structures. The third tweak is that, although we use a cylindrical shape to model nucleic acid and protein molecules with an aspect ratio parameter, this cylindrical shape will become a thin fiber when the aspect ratio is higher, typically larger than 10. The same principal can be applied to globular-shaped proteins like bovine serum albumin (BSA) with an aspect ratio of about 1. We then will treat them identically without further clarifications.

The free volume of a single nucleic acid and protein fiber molecule needs to be estimated once it is dissolved in a solvent. The same approach adopted from Hao's previous work on calculating the free volume of spherically shaped particles dispersed in a continuous medium is used in this article.[40] These fibrous protein and nucleic acid molecules are similar to polymeric chains but with smaller molecular weights and various numbers of charges at different sites. A powerful concept in dealing with the flow properties of polymeric systems is the reptation theory, which was introduced by P. G. de Gennes and expanded on by Doi and Edwards to model the viscoelastic properties of polymer materials.[41,42] The reptation concept of the polymer molecules argues that polymeric molecules behave like a snake in a tube. This concept is based on the assumption that the polymer chains randomly orient and can move freely along their length. Protein and nucleic acid molecules can be modeled similarly to polymer chains in the interpretation of rich, viscous flow behaviors. We therefore would continue to use this reptation concept to model protein and nucleic acid molecules, as illustrated in Fig. 1.

A single protein or nucleic acid molecule is assumed to be a flexible tube with radius r and length L. The aspect ratio of this tube, A_r, is defined as the ratio of its length to its diameter, $A_r = L/2r$. The aspect ratio of a protein molecule can be estimated from the molecular parameters, like molecular weight and gyration radius based on the structure of the protein molecules. The volume of this tube can thus be written as:

$$V_t = 2\pi A_r r^3 \tag{1}$$

According to the original free-volume calculation[40] and the extension to charged particles,[29] the free volume of a protein molecule can be written as:

$$V_{ft} = 64\left[r\left(\frac{\phi_m}{\phi}\right)^{1/3} - \frac{\zeta\epsilon_m\epsilon_0}{d_q}\right]^3 \tag{2}$$

where V_{ft} is the free volume of an individual protein molecule, ϕ_m is the maximum packing fraction of the protein molecules, and ϕ is the volume fraction of the protein molecules in solution, which corresponds to their concentration. ζ is the zeta potential of the charged protein molecules, d_q is the charge density of the protein molecules, and ϵ_0 and ϵ_m are the dielectric constants of vacuum and the dispersing medium (such as water), respectively. Note that the term $\frac{\zeta\epsilon_m\epsilon_0}{d_q}$ has the length unit related to the Debye length in colloidal chemistry. In SI units, the dielectric permittivity of vacuum ϵ_0 has the unit of $A^2\ s^4\ kg^{-1}\ m^{-3}$, the dielectric constant of the dispersed medium ϵ_m is unitless, the zeta potential ζ has the unit of $kg\ m^2\ s^{-3}\ A^{-1}$, and the charge density d_q has the unit of $A\ s\ m^{-2}$; therefore, the term $\frac{\zeta\epsilon_m\epsilon_0}{d_q}$ has the unit of length m, matching the unit of the first term $r\left(\frac{\phi_m}{\phi}\right)^{1/3}$. The total free volume of all protein molecules should be the free volume of an individual molecule V_{ft} multiplied by the number of protein molecules in the solution, *i.e.*:

$$V_{tft} = V_{ft} \times \frac{\phi V_s}{2\pi A_r r^3} \tag{3}$$

$$= \frac{32\phi V_s}{\pi A_r}\left[\left(\frac{\phi_m}{\phi}\right)^{1/3} - \frac{\zeta\epsilon_m\epsilon_0}{rd_q}\right]^3 \tag{4}$$

where V_{tft} is the total free volume of all protein molecules in the solution, and V_s is the volume of the solution. According to Eyring's rate process theory[15,43,44] and Hao's extension of this theory,[29] the viscosities resulting from the crowding effect of protein molecules and the electrostatic interactions can be expressed as:

$$\eta_\phi = \frac{f(2\pi m_p k_B T)^{1/2}}{2\Delta E}\left[\frac{32\phi V_s}{\pi A_r}\left[\left(\frac{\phi_m}{\phi}\right)^{1/3} - \frac{\zeta\epsilon_m\epsilon_0}{rd_q}\right]^3\right]^{-2/3} V_{fm}{}^{-2/3}\exp\left(\frac{E_0}{k_B T}\right) \tag{5}$$

$$= \frac{f(2\pi m_m k_B T)^{1/2}}{2\Delta E}\left[\frac{32\phi}{\pi A_r(1-\phi)}\right]^{-2/3}\left[\left(\frac{\phi_m}{\phi}\right)^{1/3} - \frac{\zeta\epsilon_m\epsilon_0}{rd_q}\right]^{-2} V_{fm}{}^{-2/3}\exp\left(\frac{E_0}{k_B T}\right) \tag{6}$$

$$= \left[\frac{32\phi}{\pi A_r(1-\phi)}\right]^{-2/3}\left[\left(\frac{\phi_m}{\phi}\right)^{1/3} - \frac{\zeta\epsilon_m\epsilon_0}{rd_q}\right]^{-2}\eta_m \tag{7}$$

where V_{fm} is the free volume of an individual molecule of the dispersing medium, f is the shear force applied to the solution, ΔE is the energy gap between the initial shear state to the next equilibrium shear state, E_0 is the activation energy needed

during the shearing process, k_B is the Boltzmann constant, T is the temperature, m_m is the molar mass of the dispersing continuous medium, m_p is the molar mass of the protein, and V_s is the volume of the whole protein or nucleic acid solution. All other parameters are the same as defined earlier. For convenience, all symbols and their units are listed in Table 1. η_m is the viscosity of the dispersing medium or the liquid continuous phase, where $\eta_m = \frac{f(2\pi m_m k_B T)^{1/2}}{2\Delta E} V_{fm}{}^{-2/3} \exp\left(\frac{E_0}{k_B T}\right)$, initially derived by Eyring[15,43] and extended by Hao.[29] As one can see, the viscosity of a protein solution is a function of temperature T, the protein concentration expressed as the volume fraction ϕ, the protein molecule aspect ratio A_r, the zeta potential ζ and the charge density d_q of the protein chains. Both zeta potential and charge density are dependent on the pH, salt concentrations, and the protein's molecular structure and components, *i.e.*, the folded structures and amino acid types and how they connect. Based on Hao's original treatments of pure liquids, colloidal suspensions, and polymeric systems,[29] there are three components contributing to the total viscosity of a colloidal suspension/polymeric system: the viscosities resulting from the continuous liquid medium η_m, the protein molecule crowding effect due to the volume fractions η_{phi}, and the electrostatic interactions between protein molecules that are already taken into account in eqn (7). Obviously, the viscosities of protein solutions should have these three components, too. We therefore may write the viscosity of protein solutions as below:

$$\eta = \eta_m + \eta_\phi \tag{8}$$

$$= \left[1 + \left[\frac{32\phi}{\pi A_r(1-\phi)}\right]^{-2/3}\left[\left(\frac{\phi_m}{\phi}\right)^{1/3} - \frac{\zeta\epsilon_m\epsilon_0}{rd_q}\right]^{-2}\right]\eta_m \tag{9}$$

where η is the viscosity of the protein solution containing protein molecules and the dispersing medium. The relative viscosity, typically defined as the viscosity of a protein solution/suspension divided by that of the dispersing medium, can thus be written as:

$$\eta_{re} = (\eta_m + \eta_\phi)/\eta_m \tag{10}$$

$$= 1 + \left[\frac{32\phi}{\pi A_r(1-\phi)}\right]^{-2/3}\left[\left(\frac{\phi_m}{\phi}\right)^{1/3} - \frac{\zeta\epsilon_m\epsilon_0}{rd_q}\right]^{-2} \tag{11}$$

where η_{re} is the relative viscosity frequently used in the literature. The specific viscosity, defined as $\eta_{re} - 1$, can be written as:

$$\eta_{sp} = \eta_{re} - 1 \tag{12}$$

$$= \left[\frac{32\phi}{\pi A_r(1-\phi)}\right]^{-2/3}\left[\left(\frac{\phi_m}{\phi}\right)^{1/3} - \frac{\zeta\epsilon_m\epsilon_0}{rd_q}\right]^{-2} \tag{13}$$

where η_{sp} is the specific viscosity in the literature. In the equations above, the radius of the protein molecules, r, is present. Due to the strong electrostatic interactions and base-pair attractions among protein molecules, we may not be able to treat it as a constant when the protein volume fraction increases, which is a distinguishing nature of protein molecules in comparison with other entities, such as colloidal particles and polymer chains. It should be dependent on the protein volume fraction. When we evaluate viscosity against the protein volume fraction, we should replace r with a function of the protein volume fraction ϕ. If the number of protein molecules in the solution is n_p, the total volume of the protein molecules in the solution should be the volume of a single protein molecule illustrated in Fig. 1 multiplied by n_p:

$$V_p = 2\pi r^3 A_r n_p \tag{14}$$

where V_p is the total volume of the protein molecules in the solution. So the volume fraction of protein should be:

Table 1 List of symbols and their units

Symbol	Meaning	SI unit
f	Shear force	N
m_p	Mass of the dispersed particles	kg mol^{-1}
m_m	Mass of the dispersed medium	kg mol^{-1}
k_B	Boltzmann constant	J K^{-1}
T	Temperature	K
ΔE	Energy gap between the initial shear state and next equilibrium state	J
E_0	Activation energy during the shearing process	J
ϕ	The volume fraction of the dispersed particles	Unitless
V_s	The volume of the whole protein or nucleic acid solution	m^3
A_r	The aspect ratio of the molecular chain	Unitless
ϕ_m	The maximum packing fraction of the dispersed particles	Unitless
ζ	The zeta potential of the dispersed particles	V
ϵ_m	The dielectric constant of the dispersed medium	Unitless
ϵ_0	The dielectric permittivity of vacuum	F m^{-1}
r	The radius of the reptation tube of a molecule	m
d_q	The charge density of the dispersed particle surfaces	C m^{-2}
V_{fm}	The free volume of an individual molecule of the dispersed medium	m^3
η_m	The viscosity of the dispersing medium	Pa s
η_ϕ	The viscosity resulting from the particle volume fraction and particle interactions	Pa s

$$\phi = \frac{V_p}{V_s} \tag{15}$$

$$= 2\pi r^3 A_r n_p / V_s \tag{16}$$

We then have the relationship between the protein molecule radius and the volume fraction as below:

$$r = \left(\frac{\phi V_s}{2\pi A_r n_p}\right)^{1/3} \tag{17}$$

where V_s is the volume of the protein solution. The charge density, d_q, in eqn (7), should be the total amount of charge, Q, divided by the surface area, S_p, of the protein molecules. This is how the charge density is defined in the electric double layer theory for zeta potential calculations,[45] which is slightly different from the regular density definition where the volume is usually used. So the charge density, d_q, can be expressed as:

$$d_q = \frac{Q}{S_p} \tag{18}$$

$$= \frac{Q}{\pi r^2 (4A_r + 1) n_p} \tag{19}$$

Therefore, the term rd_q in eqn (7), in terms of protein volume fraction, can be expressed as:

$$rd_q = \frac{(2A_r)^{1/3} Q}{(4A_r + 1)\left(\pi^2 {n_p}^2 \phi V_s\right)^{1/3}} \tag{20}$$

This term has a relationship with the protein volume fraction, indeed. Experimental and theoretical studies have shown that the zeta potential is dependent on the volume fraction when the volume fraction exceeds about 0.07, a relatively low volume fraction.[46,47] There is a complicated relationship between zeta potential and volume fraction, dependent on the reduced particle radius defined as the Debye length times the particle radius λr; however, the general trend is that zeta potential decreases with increasing volume fraction.[46,47] For simplicity, we will assume $\zeta = \zeta_0/\phi$, where ζ_0 is the zeta potential at the very low volume fraction. Under this approximation, the term $\frac{\zeta \epsilon_m \epsilon_0}{rd_q}$ can be written as:

$$\frac{\zeta \epsilon_m \epsilon_0}{rd_q} = \frac{(4A_r + 1)\zeta_0 \epsilon_m \epsilon_0 \left(\pi^2 {n_p}^2 V_s\right)^{1/3}}{(2A_r)^{1/3} Q \phi^{2/3}} \tag{21}$$

Using the expression in eqn (21) to replace the term of $\frac{\zeta \epsilon_m \epsilon_0}{rd_q}$ in eqn (9), we may obtain the relationship between the viscosity and the protein volume fraction:

$$\eta = \left[1 + \left[\frac{32\phi}{\pi A_r (1-\phi)}\right]^{-2/3} \left[\left(\frac{\phi_m}{\phi}\right)^{1/3} - \frac{(4A_r + 1)\zeta_0 \epsilon_m \epsilon_0 \left(\pi^2 {n_p}^2 V_s\right)^{1/3}}{(2A_r)^{1/3} Q \phi^{2/3}}\right]^{-2}\right] \eta_m \tag{22}$$

The viscosity of a protein solution is now correlated with the protein concentrations, the protein charge and zeta potential, the protein molecule aspect ratio, the number of protein molecules, the maximum packing fraction of protein molecules, the solution volume, and the viscosity of the dispersing medium. In the next section, eqn (22) will be used for evaluation and regression purposes, as it contains all variables encountered in protein solution formulations, while controlling the viscosity of protein solutions and interpreting experimental data are imperative.

The term $f/\Delta E$ in the viscosity equations of both η_m and η_ϕ contains the shear stress σ and shear rate $\dot{\gamma}$ information. The shear stress σ is the shear force f divided by the shear area A, $\sigma = f/A$. The shear rate $\dot{\gamma}$ is the shear velocity v divided by the shear distance d_p in the perpendicular shear direction, $\dot{\gamma} = v/d_p$. If the shear distance in the shear direction is d_s, then the shear area $A = d_p \times d_s$. If ΔE has some relationship with the shear velocity v, $\Delta E = vE(v)$, where $E(v)$ is a function of the velocity, then we can write the term $\frac{f}{\Delta E}$ as follows:

$$\frac{f}{\Delta E} = \frac{\sigma d_p d_s}{vE(v)} \tag{23}$$

$$= \frac{\sigma d_s}{\dot{\gamma} E(v)} \tag{24}$$

Since the energy by definition is the force times the distance traveled in the direction of the force, our assumption $\Delta E = vE(v)$ is reasonable. In this regard, the viscosity equation can be further written as:

$$\eta = \frac{\sigma d_s}{\dot{\gamma} E(v)} \frac{(2\pi m_m k_B T)^{1/2}}{2} {V_{fm}}^{-2/3} \exp\left(\frac{E_0}{k_B T}\right) \left[1 + \left[\frac{32\phi}{\pi A_r (1-\phi)}\right]^{-2/3} \left[\left(\frac{\phi_m}{\phi}\right)^{1/3} - \frac{(4A_r + 1)\zeta_0 \epsilon_m \epsilon_0 \left(\pi^2 {n_p}^2 V_s\right)^{1/3}}{(2A_r)^{1/3} Q \phi^{2/3}}\right]^{-2}\right] \tag{25}$$

Eqn (26) shows that the viscosity of a protein depends on the shear stress and shear rate, which may show the Newtonian or non-Newtonian shear behaviors we have already seen in the literature.[48]

The surface charge density of protein and nucleic acid molecules d_q is correlated with the salt concentration c_i based on the standard electric double layer theory in colloidal chemistry.[45,49] The Debye length λ can be written as:

$$\lambda = \left(\frac{\epsilon_0 \epsilon_m k_B T}{2e^2 I_0}\right)^{1/2} \tag{27}$$

where e is the elementary charge, I_0 is the ionic strength, $I_0 = \frac{1}{2}\sum_i {z_i}^2 c_i$, e is the elementary charge, z is the charge of the ion, and c_i is the molar concentration of each ion species. When

the surface potential U is low, the surface charge density may be expressed as:[45,49]

$$d_q = \frac{\epsilon_0 \epsilon_m U}{\lambda} \tag{28}$$

$$= eU\left(\frac{2\epsilon_0 \epsilon_m I_0}{k_B T}\right)^{1/2} \tag{29}$$

$$= eU\left(\frac{\epsilon_0 \epsilon_m \sum_i z_i^2 c_i}{k_B T}\right)^{1/2} \tag{30}$$

Eqn (30) shows that the surface charge density is proportional to the square root of ion concentrations. We therefore may correlate the viscosity to the salt concentration *via* eqn (9). If we assume the surface potential U is approximately equal to the zeta potential ζ, then $\zeta/d_q \approx \left(e^2 \frac{\epsilon_0 \epsilon_m \sum_i z_i^2 c_i}{k_B T}\right)^{-1/2}$. We may write the viscosity equation below to correlate with salt concentrations:

$$\eta = \left[1 + \left[\frac{32\phi}{\pi A_r (1-\phi)}\right]^{-2/3}\left[\left(\frac{\phi_m}{\phi}\right)^{1/3} - \frac{(k_B T \epsilon_m \epsilon_0)^{1/2}}{er\left(\sum_i z_i^2 c_i\right)^{1/2}}\right]^{-2}\right]\eta_m \tag{31}$$

Eqn (31) can be used to estimate how the salt concentration would impact the viscosity. Similarly, we may use this equation to evaluate how pH would impact viscosity, as the pH change would change the charge density. When the pH is neutral at 7, the charge density is supposed to be zero in theory. In reality, zero charge density may be achieved at the isoelectric point. By simply assuming $d_q + a = \mathrm{pH}$, where a is a constant, we can then use eqn (31) to predict the viscosity at different pH values.

3. Results

3.1. Theoretical predictions based on the viscosity equations

In this section, we conduct a comprehensive analysis of the relationships between viscosity and its dependence on concentration, zeta potential, charge density, temperature, and folded molecular structures as reflected by the aspect ratio. Subsequently, we compare our theoretical predictions with experimental data from existing literature.

Temperature and protein volume fraction are two major factors impacting the viscosity. Fig. 2 shows the normalized viscosity against temperature and protein volume fraction for low volume fractions below 0.5 and high volume fractions up to 0.9. The temperature range is assumed to be within −100 to 100 °C, a wide range for most protein solutions for freezing storage and heat processing purposes. With the temperature increasing, the viscosity is predicted to decrease, which is in agreement with the experimental measurements reported in the literature. When the volume fraction is low in Fig. 2(a), the viscosity seems to have a linear relationship with the volume fraction, until the volume fraction is close to the maximum packing fraction where an exponential increase is shown. This dramatic exponential increase is more pronounced when the maximum volume fraction is high, close to 0.9, as shown in Fig. 2(b). Atoms, molecules, and even large particles all need to pack into certain structures, and there is a maximum packing volume fraction for each possible structure.[50] How protein molecules can pack in a solution, *i.e.*, the protein 3D folded structures, would determine what the maximum packing volume fraction is, which is a central concept in free-volume calculation.[40] The maximum volume fraction regulates at what volume fraction point the viscosity would start to dramatically increase. Many experimental data support these predictions of a linear relationship at low volume fractions followed by an exponential increase at high volume fractions.[16,28,51]

The radius and aspect ratio of protein molecules should impact the viscosity. We assume that the volume fraction does not change when the radius and aspect ratio vary, and that they are independent of each other, which is not true in reality but can be used for evaluation purposes of our equations. We can use eqn (22) to see how the viscosity is going to vary with A_r, and use eqn (9) to see how the viscosity is going to change with r, only for pure illustration purposes. In both cases, the viscosity increases with A_r and r, which makes perfect sense as large and long protein/nucleic acid molecules would increase the internal frictional forces during the shear process, leading to higher viscosities. The large and long protein molecules usually have a high molecular weight and relatively poor solubility. In other words, they cannot easily dissolve into the dispersing liquid medium and reach a high volume fraction; the viscosity may already be extremely high at low protein volume fractions. Nonetheless, our focus is on how the aspect ratio and the radius of the protein impact the viscosity. From these predictions, we may expect higher viscosity in aggregated protein solutions where protein molecules are longer and larger, like bundles, which is evidenced in experimental investigation.[52] If protein aggregates create inhomogeneity in solution, a lower viscosity should be expected due to how viscosity is measured experimentally,[53] which has nothing to do with the intrinsic viscosity of a protein solution. The shapes of protein molecules have been experimentally found to have a big impact on viscosities: the viscosities of seven proteins with distinct structural differences from small spherical shapes to fibrous "*Y*" shapes were investigated, and the fiber-shaped proteins exhibited higher viscosities at the same volume fraction.[16] This is possibly due to the larger A_r of fibrous protein molecules, as described with eqn (22) and illustrated in Fig. 3.

Zeta potential and charge density are the two main factors that determine the viscosity of a protein. Both of them reflect the amount of charge that the protein molecules carry, but in different ways: the zeta potential is a measure of the electric potential at the electric mobile layer created by the charge entities and the charge density is a measure of the charge per unit area of the surface. They are the indications of how strong the electrostatic interactions are, and should influence the viscosity *via* so-called electroviscous effects.[52,54] Note that these two parameters only become important in concentrated protein

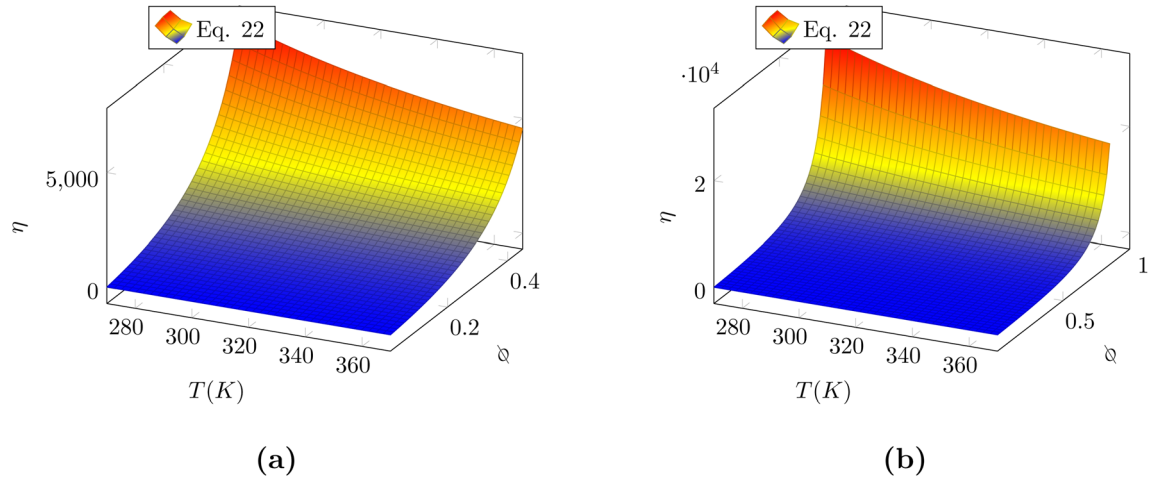


Fig. 2 Illustration of normalized viscosity against temperature and the protein concentration based on eqn (22) for (a) $\phi_m = 0.5$ and (b) $\phi_m = 0.9$.

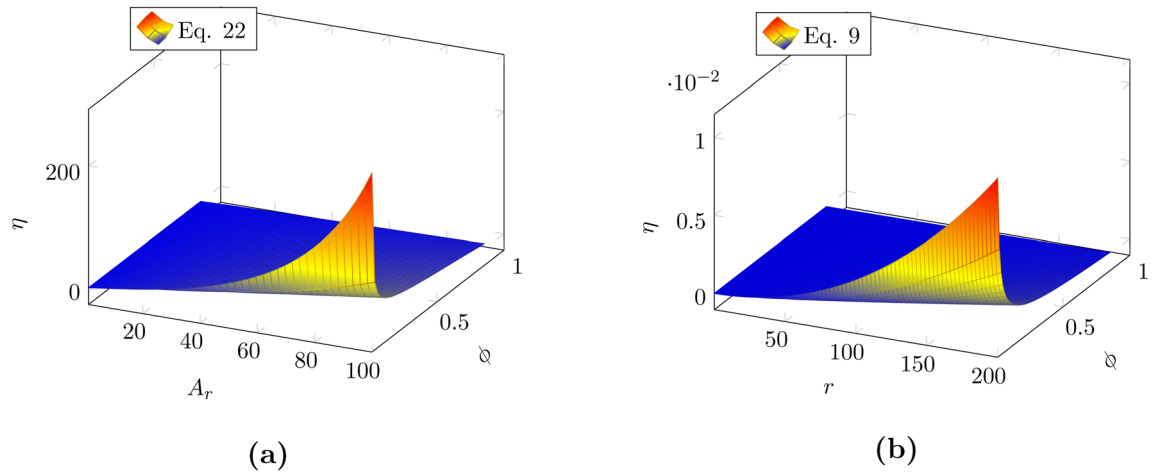
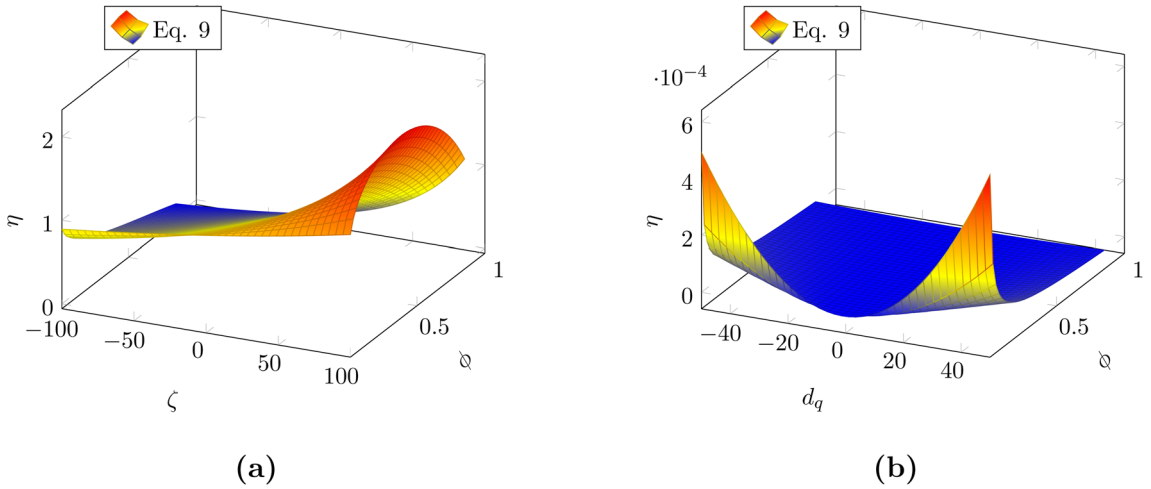


Fig. 3 Illustration of normalized viscosity against A_r (a) and the size of the protein molecules r (b) based on eqn (22) and (9), respectively. The protein volume fraction is also presented in the graphs but considered as an independent variable.

solutions where there is overlapping between the electric double layers around each molecule.[19,54] For illustration purposes, the predicted viscosities from eqn (9) are plotted against zeta potential ζ and charge density d_q in Fig. 4. When the zeta potential changes from −100 to 100 mV at relatively low volume fractions, the predicted viscosities increase at first, and decrease when the volume fraction is extremely high, implying that the electrostatic interactions overtake the crowding effect. These abnormal behaviors may only happen when the proteinmolecules can be heavily charged with charging agents or salts, and in the meantime the aggregations of molecules don't cause protein collapse from the solution – a phase change. For the charge density, there is a "U"-shaped relationship; the viscosity decreases initially, reaches the lowest point, and then increases. Salt concentration and pH variations can cause both the zeta potential and charge density to change, leading to the reduction or enhancement of viscosity profiles, which is evidenced in the literature,[19,54] where maximum viscosities are found when the salts or pH are varied. Four types of electrostatic interactions are proposed: the excluded-volume repulsion, electrostatic repulsion, electrostatic attraction, and hydrophobic attractions.[19] After salts are added to the solutions, the kosmotropic effect, resulting from the solubility decrease of hydrophobes, and the chaotropic effect, resulting from the solubility increase of hydrophobes, could occur. Depending on the intrinsic charge properties of protein molecules, the added salts may neutralize, enhance, or reduce the charge density, creating different viscosity impacts.

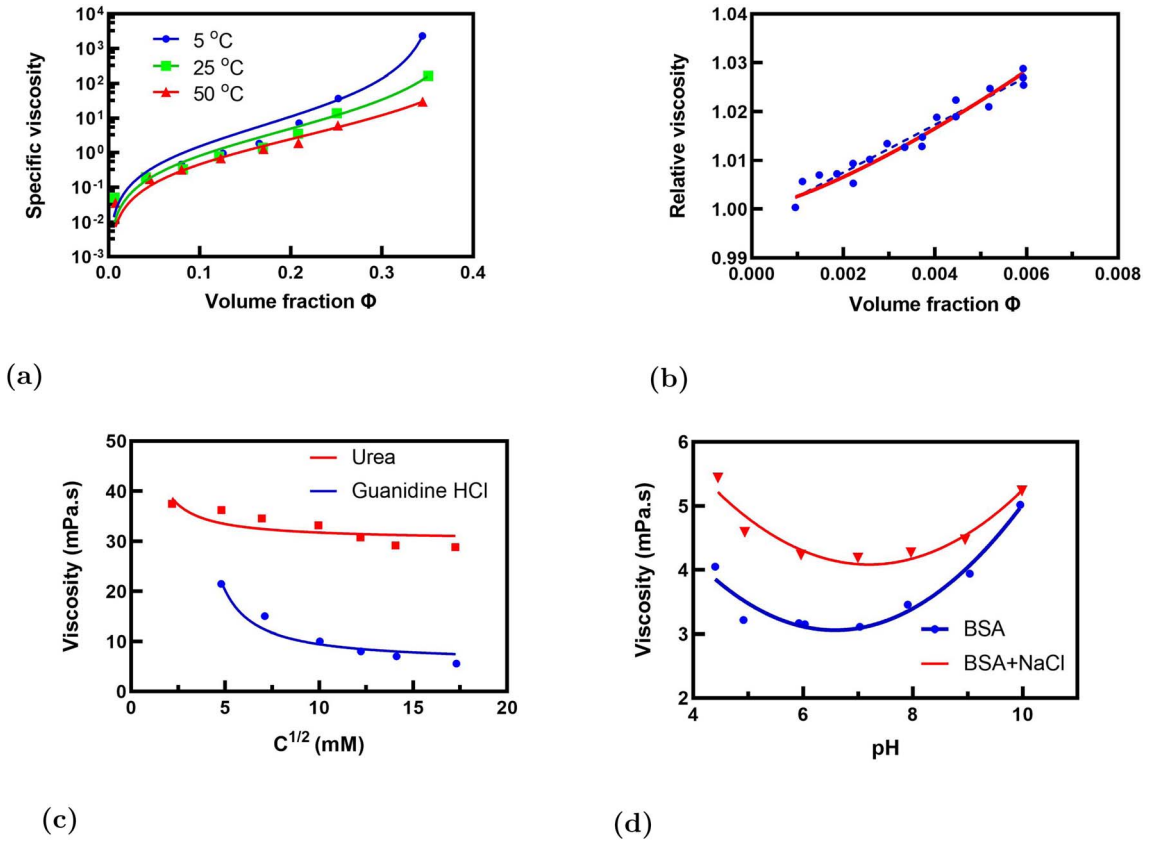
3.2. Comparisons with experimental data

The viscosities of lysozyme solutions with a wide range of volume fractions from 0.01 to about 0.35 measured at three different temperature of 5, 25, and 50 °C were extracted from the article[55] and regressed with eqn (7). The data and the regressions are shown in Fig. 5(a). The fitting curves were



(a) (b)

Fig. 4 Illustration of normalized viscosity against zeta potential ζ (a) and the charge density of the protein molecules d_q (b) based on eqn (9). The protein volume fraction is also presented in the graphs but considered as a independent variable.



(a) (b) (c) (d)

Fig. 5 Viscosities of several protein solutions are fitted with the equations derived in this article. The data points are extracted from the literature. (a) The specific viscosity of lysozyme solutions at three different temperatures, 5 °C, 25 °C, and 50 °C, *vs.* protein volume fraction. The data points are extracted from ref. 55 and the lines are regressed with eqn (22). Viscosities decrease with increasing temperature, consistent with the predictions. (b) The relative viscosity of bovine serum albumin and human serum albumin solutions at different pH values from 5.1 to 8.5 and zeta potentials from 0 to −50 mV *vs.* protein volume fraction. The data points are extracted from ref. 51. The line is regressed with eqn (22). The dotted line is a linear fit for comparison purposes only. (c) Viscosity of 125 mg per ml mAb1 in 30 mM histidine buffer at pH 6.0 *vs.* the square root of additive or salt concentrations. Data points are extracted from ref. 56, and the lines are regressed with eqn (31). The protein concentration data were converted to $c^{1/2}$ for the data-fitting purpose required by eqn (31). (d) Viscosity of 125 mg per ml mAb1 in 30 mM histidine buffer at pH 6.0 *vs.* the square root of additive or salt concentrations. Data points are extracted from ref. 56, and the lines are regressed with eqn (31). The protein concentration data were converted to $c^{1/2}$ for the data-fitting purpose required by eqn (31).

generated with eqn (7) and expressed in solid lines. As one can see, the predicted curves match perfectly with experimental data. The viscosities decrease with temperature: higher temperatures lead to lower viscosities. Such a wide volume fraction range does not show any linear regions, though in the very low volume fraction region, it is hard to tell if that is linear. Many articles claim that a linear relationship exists between the viscosity and the volume fractions at extremely low volume fractions. Fig. 5(b) shows the viscosities of two globular protein solutions, bovine serum albumin (BSA) and human serum albumin (HSA), against protein volume fractions below 0.008, *i.e.*, very diluted solutions. The viscosities linearly increase with the volume fraction below 0.006. The viscosity data points contain the samples when the pH changes from 5.1 to 8.5 and the zeta potential changes from almost zero to −50 mV.[51] These data indicate that pH and zeta potential do not impact the viscosities in such a low volume fraction region. As explained in the previous section, the viscosities calculated with eqn (22) and presented in Fig. 2 show the linear relationship at low volume fractions, and the viscosities would dramatically increase at higher volume fractions close to the maximum packing fraction, ϕ_m. The zeta potential and charge density terms in eqn (9), $\frac{\zeta\epsilon_m\epsilon_0}{rd_q}$, only become important at high volume fractions, where overlapping between electric double layers occurs, and the electrostatic interactions become comparable to the crowding effect caused by the high volume fractions. Otherwise, the viscosities would be solely dominated by the volume fraction term, and we would not see that viscosities vary with zeta potential, pH, salts, *etc.*, the parameters related to the charge properties of protein molecules.

For the purpose of evaluating our equations with other protein systems, we plot the viscosity against the protein volume fraction of recombinant human albumin (rAlbumin) solution measured with a micro-viscometer/rheometer in Fig. 6(a). Within a relatively wide concentration range from 0 to 35% by volume, our equation fits the data very well. The viscosity dramatically increases at a protein volume fraction of around 30%, implying that the system becomes very crowded and does not have much free volume for protein molecules to move around in. Therefore, a huge shear resistance, *i.e.*, a larger viscosity, is observed in higher volume fractions. Fig. 6(b) shows the viscosity against the protein volume fraction calculated with the atomistic molecular dynamics simulation method for ubiquitin (UBQ), the third IgG-binding domain of protein G (GB3), and hen egg white lysozyme (LYZ). A cluster model is used to describe the protein structures in the solution, and the viscosity is calculated from the pressure tensor fluctuations.[57] The data points are imported from the figure of the viscosity *vs.* the protein volume fraction *via* the WebPlotDigitizer software. Amazingly, the fitting lines regressed with eqn (22) are consistent with the calculated values, implying that their method really works. Of course, as stated in the original article,[57] their calculation method is calibrated with real systems and the data are reliable.

Viscosity was found to decrease with the addition of anions[56] but increase with the addition of cations,[55] which can be easily explained with eqn (9). With anions, the term $\frac{\zeta\epsilon_m\epsilon_0}{rd_q}$ becomes negative and the term $\left[\left(\frac{\phi_m}{\phi}\right)^{1/3} - \frac{\zeta\epsilon_m\epsilon_0}{rd_q}\right]^{-2}$ becomes more positive. The viscosity becomes much smaller due to the power of −2. In contrast, with cations, the term $\frac{\zeta\epsilon_m\epsilon_0}{rd_q}$ remains positive, and the term $\left[\left(\frac{\phi_m}{\phi}\right)^{1/3} - \frac{\zeta\epsilon_m\epsilon_0}{rd_q}\right]^{-2}$ becomes larger, leading to higher viscosities. These trends are demonstrated in Fig. 4 as well. The lowest viscosity occurs at the point where the charge density is zero, neither positive nor negative. The pH, similar to salt, impacts the viscosity as it changes the charge density. In the literature, the viscosity of protein solutions is

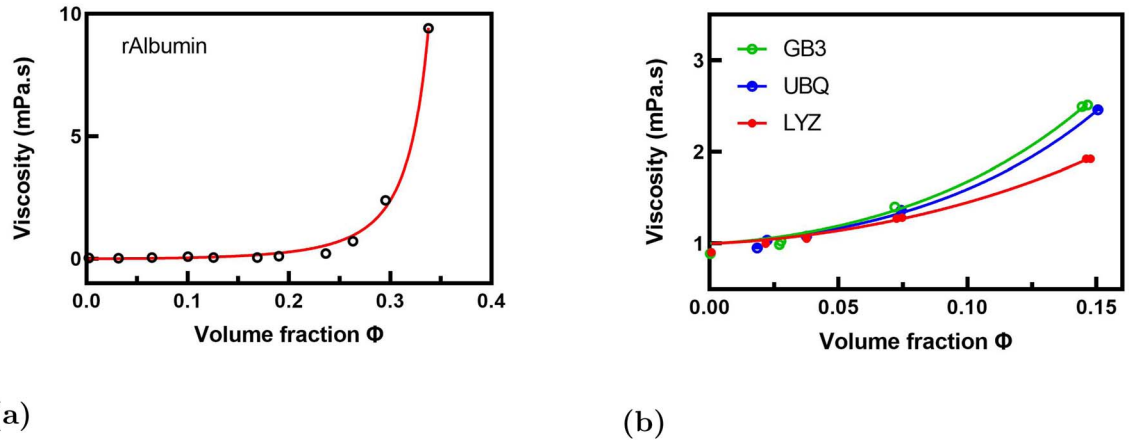


Fig. 6 Viscosities of several protein solutions are fitted with the equations derived in this article. The data points are extracted from the literature. UBQ represents ubiquitin, GB3 is the third IgG-binding domain of protein G, and LYZ is the hen egg white lysozyme. (a) The viscosity of recombinant albumin solution is plotted against the protein volume fraction, the data points are extracted from ref. 58. The line is regressed with eqn (22). (b) The viscosity of several proteins solutions is plotted against the protein volume fractions. The data points are taken from ref. 57, and the lines are regressed with eqn (22).

either independent of pH, or it increases initially and then decreases, or it decreases initially and then increases.[19,51] All of these phenomena are dependent on the intrinsic charges of protein molecules, and the added salts and adjusted pH may neutralize, intensify, or reduce the intrinsic charge properties. Fig. 5(c) shows the viscosities of mAb1 solutions against the square root of the additive or salt concentrations in two cases: one is added with urea and another is added with guanidine hydrochloride. The data points are fitted with eqn (31). Satisfactory agreements are reached between the data points and regression lines, implying that our equations reflect the underlying mechanisms of how viscosities vary with various factors.

Viscosity can change dramatically with pH, as shown in the literature.[19] It can increase with pH initially, reach a peak, and then decrease, or decrease first, reach a minimum and then increase, or be independent of pH. Fig. 5(d) shows the viscosities of bovine serum albumin (BSA) with and without NaCl at different pH values extracted from ref. 19. The viscosities show a strong pH dependence with a minimum point. The regressed lines generated with eqn (9) are also shown in the figure, by assuming that pH would simply change the charge density d_q with a shift of the isoelectric point value. Again, good agreements with experimental data further demonstrate that our theory and the derived equations are sound and can explain many rich viscosity phenomena observed in protein solutions experimentally. This non-monotonicity has recently been explored using Metropolis Monte Carlo simulations by assuming that a critical charge density is required for the electrostatically driven adsorption–desorption process.[59] This charge-regulation effect should be universal across a wide range of protein structures, based on the electrical double-layer theory.[60] Assuming different relationships between pH and d_q, we may use the same equation to fit the experimental data of the other two scenarios in ref. 19.

The viscosity of DNA solutions is found to increase with the volume fraction of DNA,[61] similar to protein solutions. Fig. 7(a) shows the viscosity of salmon testes double-stranded DNA solutions plotted against the DNA volume fraction. The density of DNA is assumed to be 1 g ml^{-1} when converting the concentrations from mg ml^{-1} to the volume fraction. The trend between the viscosity and the DNA concentrations is very similar to that of proteins. Initially, the impact of the volume fraction on the viscosity is small, but becomes dominant when the volume fraction is higher. A dramatic increase in viscosity is observed with high volume fractions. In the response to temperature, a melting point where the double-stranded DNA transitions to the single-stranded DNA is expected. Since the concentration of DNA is doubled after this transition, higher viscosity is anticipated.[62] Based on eqn (7), we may write the viscosity resulting from the single-stranded DNA, η_{ss}, as follows:

$$\eta_{ss} = \frac{f(2\pi m_m k_B T)^{1/2}}{2\Delta E}\left[\frac{64\phi}{\pi A_r(1-2\phi)}\right]^{-2/3}$$

$$\left[\left(\frac{\phi_m}{2\phi}\right)^{1/3} - \frac{\zeta\epsilon_m\epsilon_0}{rd_q}\right]^{-2} V_{fm}{}^{-2/3}\exp\left(\frac{E_0}{k_B T}\right) \quad (32)$$

$$= A\left[\frac{64\phi}{\pi A_r(1-2\phi)}\right]^{-2/3}\left[\left(\frac{\phi_m}{2\phi}\right)^{1/3} - \frac{\zeta\epsilon_m\epsilon_0}{rd_q}\right]^{-2}\exp\left(\frac{E_0}{k_B T}\right) \quad (33)$$

where A is a constant. Note that the parameter ϕ is changed to 2ϕ as the double-stranded DNA changes to single-stranded DNA after the melting. The viscosity of these DNA systems will be a mixture of these two types of DNA, plus the dispersing medium. If the percentage of the double-stranded DNA is p, then that of the single-stranded DNA should be $1 - p$. So the viscosity of the DNA solutions may be expressed as follows:

$$\eta_{DNA} = p\eta_{ds} + (1-p)\eta_{ss} + \eta_{extra} \quad (34)$$

where η_{ds} is the viscosity of the double-stranded DNA, and can be expressed with eqn (7). η_{extra} is the viscosity caused by the tertiary effects among the double-stranded DNA, single-

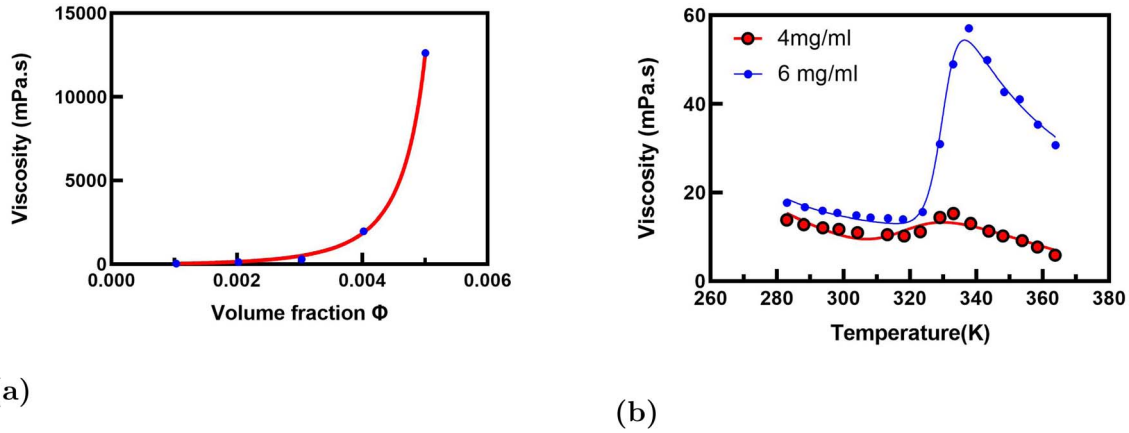


Fig. 7 Viscosities of DNA solutions are fitted with the equations derived in this article. The data points are extracted from the literature. (a) The viscosity of salmon testes double-stranded DNA water solution at 25 °C is plotted against the DNA volume fraction. The data points are extracted from ref. 66. The line is regressed with eqn (22). (b) The viscosity of genomic calf-thymus DNA water solutions is plotted against temperature. The data points are taken from ref. 62 and the lines are regressed with eqn (34).

stranded DNA, and the dispersing medium. As shown earlier, for binary systems containing only one dispersed entity and one dispersing medium, η_{extra} is assumed to be η_m, the viscosity of the medium (see eqn (22)). The parameter p is temperature dependent, and can be expressed as shown below based on the thermodynamics of DNA melting and the logistic function:[62–64]

$$p = \frac{1}{1 + \exp[a(T - T_m)]} \quad (35)$$

where a is a constant dependent on the concentration, and T_m is the melting point, *i.e.*, the temperature at which the double-stranded DNA transitions to single-stranded DNA. Fig. 7(b) shows the viscosity of calf-thymus DNA solutions at two concentrations, 4 and 6 mg ml^{-1}, against the temperature. In both concentrations, a sudden increase of viscosity is observed at the temperature about 330 K, implying a melting of the double-stranded DNA happens. The solid lines are fitted with eqn (34). Amazing agreements between the experimental data and fittings are achieved using our equation. Sensing viruses by detecting the viscosity/mechanical tension changes after viruses changes from double-strand to a single-strand structures is proposed,[65] which should be a very good tool in the future.

We select several examples to demonstrate how our equations can be used to fit the experimental data and predict the viscosities of protein and nucleic acid solutions. These equations should be applicable to a wide range of protein solutions, including globular proteins, fibrous proteins, and protein aggregates. In addition to examples of how viscosities vary with temperature, protein or nucleic acid volume fraction, the ion concentrations, and pH, they can also be used to predict how the viscosities would change with zeta potential, charge density, and other factors listed in our equations. Readers are encouraged to perform additional regressions based on their interests using our equations.

4. Discussion

The folded structures of proteins and nucleic acids can be described using three key parameters: radius, aspect ratio, and maximum packing fraction. The radius represents the size of the gyration, while the aspect ratio indicates how elongated the protein molecules are compared to their base diameter. The maximum packing fraction measures the tightness with which protein molecules can be packed in a 3D space. By using our equations, we can fit experimental data and extract this information to better characterize folded protein structures.

The viscosity equations derived in this article are based on Eyring's rate process theory and free-volume concept. The estimation of the free volume of a protein or nucleic acid molecule is based on the same approach as previous applications in pure liquids, colloidal suspensions, and polymeric systems. The distinct nature of protein and nucleic acid molecules is that their molecular sizes can change substantially with the volume fractions, pH, and salt concentrations, *etc.* In addition, the electrostatic interactions become prominent and cannot be neglected in most cases. Secondly, the protein and nucleic acid molecules are not spherical, and the aspect ratio parameter, A_r, must be taken into account in modelling the viscosities. The idea behind the self-diffusion process, which is restricted by crowders and responsible for the viscosity, is consistent with the free-volume concept in general.[67]

A unified theory describing the viscosities of protein and nucleic acid solutions, based on our previous theoretical treatments of pure liquids, colloidal suspensions, and polymeric systems, is provided with a few tweaks and expansions. We also provide a set of numerical examples to demonstrate the applicability of our equations. More examples could be tested with the equations once the experimental conditions and parameters are known. Please note that the theoretical treatments presented in this article are not bound to any specific protein or nucleic acid. The theory should be universally applicable. Our equations contain all important variables, like shear stress, shear rate, temperature, volume fraction, zeta potential, charge density, salt charge, and concentrations, *etc.* As demonstrated in our previous work, these equations based on Eyring's rate process theory and free-volume concept may provide much better fits with experimental data in comparison with other models, like Einstein's viscosity equation, Mooney's equation, Krieger–Dougherty's equation.[29] This kind of comparison is thus not repeated in this article.

When visualizing our equations graphically, it is essential to acknowledge that certain parameters within these equations are often assumed to be independent of each other, even though this may not accurately reflect real-world conditions. Furthermore, the graphs themselves might be mathematically accurate but fail to make physical sense. For instance, when the aspect ratio of molecules shifts from low to high, it is logical that the volume fraction should also change in response, yet we typically maintain the volume fraction as a separate variable. This highlights the importance of exercising caution when drawing conclusions solely based on these graphical representations.

Another important point is that the radius of the protein and nucleic acid molecules is assumed to be a constant when we evaluate other parameters, which may not be true for nucleic acids due to the strong base-pair interactions and strong electrostatic interactions. This is an unavoidable assumption in the evaluation.

5. Conclusions

A hypothesis that the molecules of proteins and nucleic acids would behave like reptation tubes is successfully employed in this article to formulate a unified theory of viscosity. A set of new equations are therefore derived based on previous work using Eyring's rate process theory and free-volume concept. This approach is the same in studies on pure liquids, colloidal suspensions, and polymeric systems, but with three modifications to address the unique characteristics of protein and nucleic acid solutions: (1) protein molecules are treated as fibers rather than spherical particles, allowing for consideration of their aspect ratio. (2) Electrostatic interactions are emphasized in the model. (3) The size variations of protein and nucleic acid molecules with various conditions are taken into account, including non-constant radii. The derived equations can fit the experimental viscosity data across a range of parameters,

including temperature, volume fraction, zeta potential, pH, and salt concentrations. The main findings are summarized below:

(1) At low volume fractions, a linear relationship is observed between viscosity and volume fraction, whereas at higher volumes, the viscosity profile exhibits a dramatic increase as the volume fraction reaches its maximum value. This packing-related parameter and aspect ratio can be used to characterize protein shapes at the molecular level. Both proteins and nucleic acids behave very similarly in this regard.

(2) The effects of pH and added salts or additives may change the molecules' intrinsic charging nature, depending on their type and concentration. These substances can either neutralize, intensify, or reduce the charging properties of proteins and nucleic acids, which in turn affects their viscosity in a different manner. Our equations can accurately handle these complex scenarios and provide reliable predictions that align with experimental results, especially the viscosity dip with pH, and the non-linearity with the square root of the salt concentrations.

(3) Our equations can reasonably describe the temperature dependence of viscosity without using the empirical Arrhenius equation. In particular, it can fit the viscosity data of the double-stranded DNA solutions before and after melting into the single-stranded DNA chains.

(4) All these predictions can come from a single-source equation, providing a unified approach to understand the viscosity of proteins and nucleic acids.

Our research provides insight into the mechanisms governing protein and nucleic acid viscosities, which could lead to the creation of a predictive model for the viscosities of therapeutic formulations under diverse conditions, including low temperatures, high volumes, and limited quantities. The derived equations would be particularly valuable when developing expensive proteins with short supply chains and tight deadlines. By examining viscosity data at room temperature and diluted conditions, we can refine therapeutic formulations to achieve optimal viscosities at preferred temperatures, pH levels, salt concentrations, and desired concentrations. This approach could enable the creation of various formulations without the need for additional experimental work, thereby significantly accelerating the development process and saving time and effort.

Conflicts of interest

There are no conflicts to declare.

Data availability

This study was carried out using publicly available data from the literature. Citations of these articles are shown in the manuscript. Thank you for your consideration.

Acknowledgements

The author sincerely appreciates colleagues' and reviewers' feedback and comments for substantially improving the readability and rationality of this article.

References

1 J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, *et al.*, Highly accurate protein structure prediction with AlphaFold, *Nature*, 2021, **596**, 583–589.
2 J. Abramson, J. Adler, J. Dunger, R. Evans, T. Green, A. Pritzel, O. Ronneberger, L. Willmore, A. J. Ballard, J. Bambrick, *et al.*, Accurate structure prediction of biomolecular interactions with AlphaFold 3, *Nature*, 2024, **630**, 493–500.
3 E. Kussell, J. Shimada and E. I. Shakhnovich, Excluded volume in protein side-chain packing, *J. Mol. Biol.*, 2001, **311**, 183–193.
4 T. Skóra, F. Vaghefikia, J. Fitter and S. Kondrat, Macromolecular crowding: how shape and interactions affect diffusion, *J. Phys. Chem. B*, 2020, **124**, 7537–7543.
5 J. F. Kincaid and H. Eyring, Free volumes and free angle ratios of molecules in liquids, *J. Chem. Phys.*, 1938, **6**, 620–629.
6 R. Furth, On the theory of the liquid state, *Proc. Cambridge Philos. Soc.*, 1941, 252–290.
7 C. D'Agostino, Hole theory as a prediction tool for Brownian diffusive motion in binary mixtures of liquids, *RSC Adv.*, 2017, **7**, 51864–51869.
8 M. Stone and I. Sanchez, Excluded volume scaling for polymers, *Comput. Theor. Polym. Sci.*, 1999, **9**, 117–123.
9 R. P. White and J. E. Lipson, Polymer free volume and its connection to the glass transition, *Macromolecules*, 2016, **49**, 3987–4007.
10 J. Sharma, K. Tewari and R. K. Arya, Diffusion in polymeric systems-A review on free volume theory, *Prog. Org. Coat.*, 2017, **111**, 83–92.
11 D. Bamford, G. Dlubek, A. Reiche, M. A. Alam, W. Meyer, P. Galvosas and F. Rittig, The local free volume, glass transition, and ionic conductivity in a polymer electrolyte: A positron lifetime study, *J. Chem. Phys.*, 2001, **115**, 7260.
12 K. J. Rao, The phenomenon of glass transition, *Bull. Mater. Sci.*, 1979, **3**, 181–193.
13 S. Lübeck, Universal scaling behavior of non-equilibrium phase transitions, *Int. J. Mod. Phys. B*, 2004, **18**, 3977–4118.
14 T. Hao, The empty world-a view from the free volume concept and Eyring's rate process theory, *Phys. Chem. Chem. Phys.*, 2024, **26**, 26156–26191.
15 S. Glasstone, K. Laidler, and H. Eyring, *The Theory of Rate Processes*, McGraw Hill, 1st edn, 1941.
16 Z. Tian, X. Jiang, Z. Chen, C. Huang and F. Qian, Quantifying protein shape to elucidate its influence on solution viscosity in high-concentration electrolyte solutions, *Mol. Pharm.*, 2024, **21**, 1719–1728.
17 M. Heinen, F. Zanini, F. Roosen-Runge, D. Fedunová, F. Zhang, M. Hennig, T. Seydel, R. Schweins, M. Sztucki, M. Antalík, *et al.*, Viscosity and diffusion: crowding and salt effects in protein solutions, *Soft Matter*, 2012, **8**, 1404–1419.

18 B. Zagrovic and V. Pande, Solvent viscosity dependence of the folding rate of a small protein: distributed computing study, *J. Comput. Chem.*, 2003, **24**, 1432–1436.
19 T. Hong, K. Iwashita and K. Shiraki, Viscosity control of protein solution by small solutes: a review, *Curr. Protein Pept. Sci.*, 2018, **19**, 746–758.
20 B. Holma and P. O. Hegg, pH-and protein-dependent buffer capacity and viscosity of respiratory mucus. Their interrelationships and influence of health, *Sci. Total Environ.*, 1989, **84**, 71–82.
21 J. Zarzar, T. Khan, M. Bhagawati, B. Weiche, J. Sydow-Andersen and A. Sreedhara, High concentration formulation developability approaches and considerations, *mAbs*, 2023, 2211185.
22 M. A. Woldeyes, W. Qi, V. I. Razinkov, E. M. Furst and C. J. Roberts, Temperature dependence of protein solution viscosity and protein-protein interactions: insights into the origins of high-viscosity protein solutions, *Mol. Pharmaceutics*, 2020, **17**, 4473–4482.
23 P. D. Ross and A. P. Minton, Hard quasispherical model for the viscosity of hemoglobin solutions, *Biochem. Biophys. Res. Commun.*, 1977, **76**, 971–976.
24 S. Yadav, S. J. Shire and D. S. Kalonia, Factors affecting the viscosity in high concentration solutions of different monoclonal antibodies, *J. Pharm. Sci.*, 2010, **99**, 4812–4829.
25 I. M. Krieger and T. J. Dougherty, A mechanism for non-Newtonian flow in suspensions of rigid spheres, *Trans. Soc. Rheol.*, 1959, **3**, 137–152.
26 D. S. Tomar, S. Kumar, S. K. Singh, S. Goswami and L. Li, Molecular basis of high viscosity in concentrated antibody solutions: Strategies for high concentration drug product development, *mAbs*, 2016, **8**, 216–228.
27 G. B. Armstrong, A. Roche, W. Lewis and Z. Rattray, Reconciling predicted and measured viscosity parameters in high concentration therapeutic antibody solutions, *mAbs*, 2024, **16**, 2438172.
28 Z. Zhang and Y. Liu, Recent progresses of understanding the viscosity of concentrated protein solutions, *Curr. Opin. Chem. Eng.*, 2017, **16**, 48–55.
29 T. Hao, Viscosities of liquids, colloidal suspensions, and polymeric systems under zero or non-zero electric field, *Adv. Colloid Interface Sci.*, 2008, **142**, 1–19.
30 T. Hao, Electrical conductivity equations derived with the rate process theory and free volume concept, *RSC Adv.*, 2015, **5**, 48133–48146.
31 T. Hao, Analogous viscosity equations of granular powders based on Eyring's rate process theory and free volume concept, *RSC Adv.*, 2015, **5**, 95318–95333.
32 T. Hao, Y. Xu and T. Hao, Conductivity equations of protons transporting through 2D crystals obtained with the rate process theory and free volume concept, *Chem. Phys. Lett.*, 2018, **698**, 67–71.
33 T. Hao, Defining temperatures of granular powders analogously with thermodynamics to understand jamming phenomena, *AIMS Mater. Sci.*, 2018, **5**, 1–33.
34 T. Hao, Infection dynamics of coronavirus disease 2019 (Covid-19) modeled with the integration of the Eyring's rate process theory and free volume concept, *medRxiv*, 2020, preprintDOI: 10.1101/2020.02.26.20028571.
35 T. Hao, Prediction of coronavirus disease (COVID-19) evolution in USA with the model based on the Eyring's rate process theory and free volume concept, *medRxiv*, 2020, preprintDOI: 10.1101/2020.04.16.20068692.
36 T. Hao, Theoretical exploration of external pressure impact on superconducting transition temperatures, *Chem. Phys. Lett.*, 2022, **802**, 139792.
37 T. Hao, Y. Xu and T. Hao, Exploring the inflation and gravity of the universe with Eyring's rate process theory and free volume concept, *Phys. Essays*, 2018, **31**, 177–187.
38 D. Rajendran, S. Mitra, H. Oikawa, K. Madhurima, A. Sekhar, S. Takahashi and A. N. Naganathan, Quantification of entropic excluded volume effects driving crowding-induced collapse and folding of a disordered protein, *J. Phys. Chem. Lett.*, 2022, **13**, 3112–3120.
39 H.-X. Zhou, Rate theories for biologists, *Q. Rev. Biophys.*, 2010, **43**, 219–293.
40 T. Hao and R. E. Riman, Calculation of interparticle spacing in colloidal systems, *J. Colloid Interface Sci.*, 2006, **297**, 374–377.
41 P. G. de Gennes, Reptation of a Polymer Chain in the Presence of Fixed Obstacles, *J. Chem. Phys.*, 1971, **55**, 572–579.
42 M. Doi and S. F. Edwards, Dynamics of concentrated polymer systems. Part 1. Brownian motion in the equilibrium state, *J. Chem. Soc., Faraday Trans. 2*, 1978, (74), 1789–1801.
43 H. Eyring and J. Hirschfelder, The theory of the liquid state, *J. Phys. Chem.*, 1937, **41**, 249–257.
44 J. J. Rooney, Eyring transition-state theory and kinetics in catalysis, *J. Mol. Catal. A: Chem.*, 1995, **96**, L1–L3.
45 P. C. Hiemenz, and R. Rajagopalan, *Principles of Colloid and Surface Chemistry, Revised and Expanded*, CRC press, 2016.
46 M. Kaszuba, J. Corbett, F. M. Watson and A. Jones, High-concentration zeta potential measurements using light-scattering techniques, *Philos. Trans. R. Soc., A*, 2010, **368**, 4439–4451.
47 H. Ohshima, Dynamic electrophoretic mobility of spherical colloidal particles in concentrated suspensions, *J. Colloid Interface Sci.*, 1997, **195**, 137–148.
48 A. Malhotra and J. N. Coupland, The effect of surfactants on the solubility, zeta potential, and viscosity of soy protein isolates, *Food Hydrocolloids*, 2004, **18**, 101–108.
49 A. D. Ratschow, H.-J. Butt, S. Hardt and S. A. Weber, Liquid slide electrification: advances and open questions, *Soft Matter*, 2025, **21**, 1251–1262.
50 J. Cumberland, and R. J. Crawford, *The Packing of Particles*, Elsevier, Amsterdam, 1987.
51 B. Jachimska, M. Wasilewska and Z. Adamczyk, Characterization of globular protein solutions by dynamic light scattering, electrophoretic mobility, and viscosity measurements, *Langmuir*, 2008, **24**, 6866–6872.
52 G. V. Barnett, W. Qi, S. Amin, E. N. Lewis and C. J. Roberts, Aggregate structure, morphology and the effect of

aggregation mechanisms on viscosity at elevated protein concentrations, *Biophys. Chem.*, 2015, **207**, 21–29.

53 L. Nicoud, M. Lattuada, A. Yates and M. Morbidelli, Impact of aggregate formation on the viscosity of protein solutions, *Soft Matter*, 2015, **11**, 5513–5522.

54 J. Li, Y. Cheng, X. Chen and S. Zheng, Impact of electroviscous effect on viscosity in developing highly concentrated protein formulations: Lessons from non-protein charged colloids, *Int. J. Pharm.:X*, 2019, **1**, 100002.

55 P. D. Godfrin, S. D. Hudson, K. Hong, L. Porcar, P. Falus, N. J. Wagner and Y. Liu, Short-time glassy dynamics in viscous protein solutions with competing interactions, *Phys. Rev. Lett.*, 2015, **115**, 228302.

56 S. Kanai, J. Liu, T. W. Patapoff and S. J. Shire, Reversible self-association of a concentrated monoclonal antibody solution mediated by Fab-Fab interaction that impacts solution viscosity, *J. Pharm. Sci.*, 2008, **97**, 4219–4227.

57 S. von Bülow, M. Siggel, M. Linke and G. Hummer, Dynamic cluster formation determines viscosity and diffusion in dense protein solutions, *Proc. Natl. Acad. Sci. U. S. A.*, 2019, **116**, 9843–9852.

58 A. D. Gonçalves, C. Alexander, C. J. Roberts, S. G. Spain, S. Uddin and S. Allen, The effect of protein concentration on the viscosity of a recombinant albumin solution formulation, *RSC Adv.*, 2016, **6**, 15143–15154.

59 D. A. Anhesini, D. L. Caetano, I. P. Caruso, A. G. Cherstvy and S. J. de Carvalho, Critical adsorption of polyelectrolytes onto patchy particles with a low-dielectric interior, *Polymers*, 2025, **17**, 2205.

60 D. L. Caetano, R. Metzler, A. G. Cherstvy and S. J. de Carvalho, Adsorption of lysozyme into a charged confining pore, *Phys. Chem. Chem. Phys.*, 2021, **23**, 27195–27206.

61 S. E. Álvarez, G. T. González, J. F. A. S. Martínez, E. B. F. Ochoa, G. L. Gómez, A. Saint-Jalmes, and L. M. Bravo-Anaya, Effect of DNA's Molecular Weight on their Solution Viscosity, *Critical Concentrations, and Liquid Crystals Formation*, Macromolecular Symposia, 2024, p. 2400120.

62 A. Laesecke and J. L. Burger, Viscosity measurements of DNA solutions with and without condensing agents, *Biorheology*, 2014, **51**, 15–28.

63 J. Gibbs and E. DiMarzio, Statistical Mechanics of Helix-Coil Transitions in Biological Macromolecules, *J. Chem. Phys.*, 1959, **30**, 271–282.

64 S. Mochrie, The Boltzmann factor, DNA melting, and Brownian ratchets: topics in an introductory physics sequence for biology and premedical students, *Am. J. Phys.*, 2011, **79**, 1121–1126.

65 J. Shin, A. G. Cherstvy and R. Metzler, Sensing viruses by mechanical tension of DNA in responsive hydrogels, *Phys. Rev. X*, 2014, **4**, 021002.

66 I. Elkin, A. K. Weight and A. M. Klibanov, Markedly lowering the viscosity of aqueous solutions of DNA by additives, *Int. J. Pharm.*, 2015, **494**, 66–72.

67 J. Shin, A. G. Cherstvy and R. Metzler, Self-subdiffusion in solutions of star-shaped crowders: non-monotonic effects of inter-particle interactions, *New J. Phys.*, 2015, **17**, 113028.