Soft Matter

Thermodynamic non-ideality in charge regulation of weak polyelectrolytes

Journal:	*Soft Matter*
Manuscript ID	SM-ART-06-2021-000848.R2
Article Type:	Paper
Date Submitted by the Author:	25-Aug-2021
Complete List of Authors:	Gallegos, Alejandro; University of California, Riverside, Ong, Gary; University of California, Riverside, Chemical and Environmental Engineering Wu, Jianzhong; University of California, Riverside, Chemical and Environmental Engineering

Thermodynamic non-ideality in charge regulation of weak polyelectrolytes

Alejandro Gallegos, Gary M.C. Ong, and Jianzhong Wu*

Department of Chemical and Environmental Engineering, University of California, Riverside, CA 92521, USA

*To whom correspondence should be addressed. Email: jwu@engr.ucr.edu

Abstract

Polymer ionization differs from that for their monomeric counterparts due to intramolecular correlations. Such effects are conventionally described in terms of the site-binding model that accounts for short-range interactions between neighboring sites. With an apparent equilibrium constant for each ionizable group and the nearest-neighbor energy as adjustable parameters, the site-binding method is useful to correlate experimental titration curves when the site–site interactions are insignificant at long range. This work aims to describe the electrostatic behavior of weak polyelectrolytes in aqueous solutions on the basis of the intrinsic equilibrium constants of the individual ionizable groups and solution conditions underlying the thermodynamic non-ideality. A molecular thermodynamic model is proposed for the protonation of weak polyelectrolytes by incorporating the classical density functional theory into the site-binding model to account for the effects of the local ionic environment on both inter- and intra-chain correlations. By an extensive comparison of theoretical predictions with experimental titration curves, we demonstrate that the thermodynamic model is able to quantify the ionization behavior of weak polyelectrolytes over a broad range of molecular architectures and solution conditions.

Keywords:

Classical density functional theory, site-binding model, titration curves, primitive model

1. Introduction

The functionality of polyelectrolytes is closely affiliated with electrostatic interactions. In contrast to strong polyelectrolytes, weak polyelectrolytes present ionizable groups with the degree of ionization in accordance with the solution pH and the local chemical environment. The pH-regulated charge behavior is advantageous for applications in "smart" polymeric systems to achieve specific functionality such as targeted drug delivery and controlled release in response to the pH variation of intracellular fluids[1, 2]. Such behavior can also be utilized in wastewater treatment[3], thermal energy storage[4], solid-liquid separation[5], and underwater adhesion[6, 7], to name only a few. Clearly, electrostatic interactions play a central role in diverse applications of weak polyelectrolyte systems.

The dependence of site ionization on solution conditions makes electrostatic correlation in weak polyelectrolyte systems more complex than that in strong polyelectrolytes. For weak polyelectrolytes, charge regulation is significantly different from the ionization of monomeric compounds because it is coupled not only with interactions with other chemical species in the solution but also with the ionization of neighboring sites[8]. For weak homo-polyelectrolytes such as polyacids and polybases, the ionization is hindered by the Coulombic repulsion among neighboring charged monomers[9]. On the other hand, polyampholytes are more easily ionized than their monomeric counterparts due to favorable attraction between neighboring acidic and basic groups because of the opposite charge[10]. For many weak polyelectrolytes, the titration curve exhibits a step-like shape caused by the sequential ionization of monomers along the polymer chain. In general, the degree of ionization is lower than that for their monomeric counterparts due to intra-chain correlations. For example, poly(maleic acid), PMA, exhibits a

two-step titration behavior resulting from strong interactions among its high-density carboxyl groups along the polymer backbone[11]. At moderate to high salt concentrations, ionization takes place at alternating sites and must overcome strong electrostatic repulsion with the neighboring charged groups[8]. In weak polyelectrolyte solutions, the charge regulation depends not only on intra-chain correlations but also on intermolecular interactions due to the presence of other chemical species. In that case, we need a thermodynamic model to account for the solution condition such as salt concentration and the chemical details of individual ionic species.

Theoretical modeling of weak polyelectrolytes has a long history and varies greatly in term of complexity[12, 13]. Recent studies of weak polyelectrolyte titration are mostly based on Monte Carlo (MC) simulations, typically within the framework of a coarse-grained representation of the polymer chains[14-17]. For example, Laguecir et al. investigated the ionization of linear poly(acrylic acid) using a freely-jointed hard-sphere-chain model along with a screened Debye-Hückel potential for electrostatic interactions between charged segments[18]. Excellent agreement was achieved in comparison with experimental titration curves for the polyacid in dilute electrolyte solutions. These authors also demonstrated that, at the same pH, the degree of ionization is reduced as the molecular weight (MW) increases due to the long-range correlation effects. A similar model was used by Uhlík et al. to study the ionization behavior of star-like weak polyelectrolytes in a salt-free solution[19]. It was found that, when the number of arms varies from 2 to 10 while fixing the length of each arm at 100 segments, the overall ionization of the polymer is reduced as the number of arms increases due to the increased monomer density near the molecular center. In comparison with analytical methods, one major advantage of MC simulation is that, given a polymer model, it is able to account for both the intra- and inter-molecular correlations exactly. However, molecular simulation of weak polyelectrolyte systems

typically avoids an explicit consideration of electrostatic interactions with protons or hydroxyl ions because the inclusion of such effects would significantly increase the computational cost. By contrast, mean-field methods are able to explicitly account for all ionic species in the solution thus computationally more efficient than simulation methods[20, 21]. Recently, Uhlík et al. compared the predictions of the self-consistent-field theory (SCFT) with simulation data[19]. While SCFT provides a reasonable description of the conformation of weak polyelectrolyte chains, it fails to capture the titration curves for star-like weak polyelectrolytes. Nevertheless, SCFT is able to describe the titration behavior of linear poly(acrylic acid) in good agreement with experimental data. Besides, it can be combined with a two-site model with one site containing a carboxyl group and the other site containing two carboxyl groups (divalent when fully deprotonated)[22].

In this work, we develop a liquid-state method to describe the titration of weak polyelectrolytes by incorporating the classical density functional theory (cDFT) with the conventional site-binding model[23]. By accounting for the free energy of ionization for each segment in a local solution environment, cDFT enables a faithful description of the salt effects on weak polyelectrolyte charging that are typically ignored in conventional methods. With the intrinsic protonation/deprotonation constant as the input, which can be fixed from fitting with experimental data for monomeric systems or predicted by first-principles methods, the thermodynamic model provides a quantitative description of the degree of ionization for individual segments in weak polyelectrolyte systems with arbitrary chain topologies. In the reminder of this article, we will present the theoretical details and numerical results as follows. We first outline the conventional site-binding model for describing the titration of weak polyelectrolyte chains. Next, cDFT is introduced to account for the effect of local solution

environment on the ionization of individual segments by using a coarse-grained model for polymers and electrolyte solutions. The numerical results are then discussed by comparison with experimental titration curves for linear and branched polyelectrolytes. After demonstrating the theoretical capability of the molecular thermodynamic model for diverse weak polyelectrolyte systems, we summarize the main results and offer perspectives on possible further developments and new applications.

2. Thermodynamic model and methods

In this section, we first recapitulate the two-site model for describing the ionization of weak polyelectrolytes with an emphasis on its extension to polymers of arbitrary architecture. Next, we introduce a molecular thermodynamic model to account for the solution effects on the binding energy and site-site interactions that are not included in conventional site-binding models. The theoretical details on the application of classical density functional theory to inhomogeneous ionic systems are given in Supporting Information.

2.1 *Site-binding model*

The site-binding model is built upon the Ising chain model in statistical mechanics. It describes the ionization of weak polyelectrolytes in terms of the one-body potential and nearest-neighbor interactions[8, 24]. Within this model, the ionizable groups are treated as proton-binding sites with their spatial arrangements along the polymer backbone dependent upon the polymer architecture (e.g., linear or branched). Each site may exist in charged or neutral state depending on the solution pH and the apparent protonation constant (*viz*., pK_i'). The interaction between neighboring sites is represented by a screened Coulomb potential that partially accounts for the solution condition[8].

The site-binding model performs well to correlate experimental titration curves with the proton binding constant and nearest-neighbor interaction energy treated as adjustable parameters. One major limitation is that the site-binding model is not able to account for polymer conformations. Theoretical improvements are mostly focused on site-site interactions beyond the immediate neighbors[25]. However, less attention has been directed at the thermodynamic effects on both the one-body potential and intra-chain correlations. In Section 2.2, we will outline a method to account for this thermodynamic non-ideality due to changes in the inter- and intra-molecular correlations with the salt concentration. We first begin with an outline of the original site-binding model and its extension to polymers with arbitrary architectures.

As in the Ising model, the Hamiltonian for the site-binding model is a quadratic function of the state variable for the ionizable sites

$$\Gamma\left(s_1,...,s_M\right)=\sum_{i}^{M}\Delta\mu_i s_i+\sum_{i}^{M-1}W^{P}_{i,i+1}s_i s_{i+1} \tag{2}$$

where s_i stands for the microstate of ionizable group i, and M is the number of ionizable sites at each weak polyelectrolyte chain. The state variable is set as $s_i = 0$ if site i is not charged, and $s_i = 1$ otherwise. The first term on the right side of Eq. (1), $\Delta\mu_i$, accounts for the reversible work to convert a monomer from the neutral state to the charged state. The ionization can be achieved by either adding a proton to a basic site or removing a proton from an acidic site. Approximately, the change in the intrinsic free energy, i.e., the free energy of a single site at the single-chain reference state, is the same as that for the monomer at infinite dilution,

$$\beta\Delta\mu_i=\beta\mu_i\left(s_i=1\right)-\beta\mu_i\left(s_i=0\right)=\begin{cases}\left[\mathrm{pK}_{\mathrm{i}}'-\mathrm{pH}\right]\ln 10, & \text{for an acid}\\ \left[\mathrm{pH}-\mathrm{pK}_{\mathrm{i}}'\right]\ln 10, & \text{for a base}\end{cases} \tag{3}$$

where pK_i' is the negative logarithm of the apparent protonation constant for segment *i*, K_i', $pH = -\log a_{H^+}$ with a_{H^+} being the proton activity, and $\beta = 1/k_B T$ where k_B is the Boltzmann constant, and *T* is the absolute temperature. In the conventional site-binding model, the apparent equilibrium constant depends not only on the ionization reactions of individual groups but also on the local solution condition. It should be clarified that the apparent protonation constant is for that of an individual segment and not for the pH at which the polymer is half-charged as it is typically referred to in experiments. In the latter case, the polymer architecture and inter- and intra- molecular correlations dictate the pH value at which the polymer is half-charged. Such monomer-level information is not accessible experimentally. The second term on the right side of Eq. (1), $W_{i,i+1}^{p}$, takes into account an additional work due to the electrostatic (and non-electrostatic) interactions between neighboring sites. This term differentiates the titration of weak polyelectrolytes from that for their monomeric counterparts. Owing to the electrostatic screening, the strength of electrostatic interactions among polymer segments varies with the salt concentration. On the other hand, non-electrostatic interactions, which may result from short-range interactions such as hydrogen bonding and electron sharing, are less sensitive to the salt concentration. The relative strength of the short- and long-range interactions is dependent on the distance between neighboring sites. While the site-binding model is attractive due to its simplicity, it provides little insights in the incorporation of the intermolecular correlations through $\Delta\mu_i$ and the intramolecular correlations through parameter $W_{i,i+1}^{p}$. One major objective of this work is to predict these two phenomenological parameters with a thermodynamic model.

The degree of ionization is obtained from the ensemble average of all possible charge states of the ionizable groups based on the semi-grand canonical partition function

$$\Xi = \sum_{s_1,\ldots,s_M} e^{-\beta\Gamma(s_1,\ldots,s_M)} \quad . \tag{4}$$

The summation extends to all combinations of uncharged and charged sites on a single weak polyelectrolyte chain under consideration. For each ionizable site, the degree of ionization can be determined from

$$\alpha_k = \langle s_k \rangle = -\frac{\partial \ln \Xi}{\partial \beta \mu_k} = \frac{\sum_{\{s_{i\neq k}\}} e^{-\beta\Gamma(\{s_i\},s_k=1)}}{\sum_{\{s_i\}} e^{-\beta\Gamma(\{s_i\})}} \quad . \tag{5}$$

In principle, Eq. (4) can be evaluated by enumerating all possible combinations of the charge states or by applying MC simulation. Whereas such procedures are straightforward for short polymers, direct enumeration of all possible charge states of ionizable groups can quickly become computationally prohibitive as the chain length increases because the number of charge states scales exponentially with the number of segments. From an analytical perspective, the transfer matrix technique provides an elegant approach for evaluating the partition function.

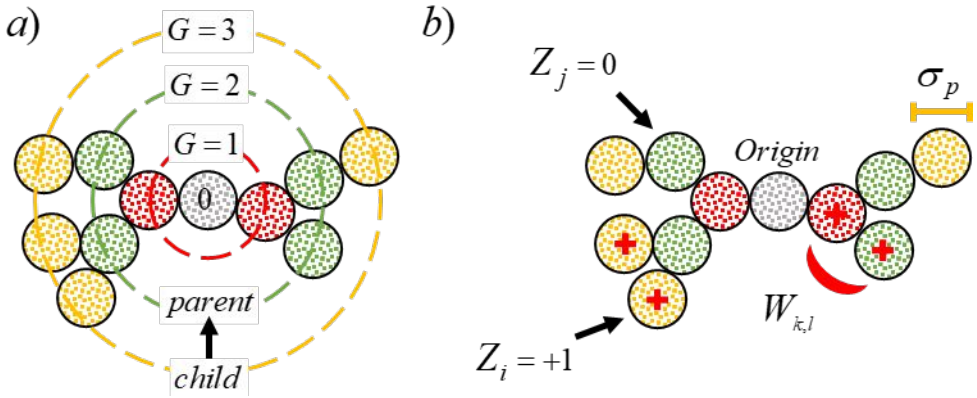

Figure 1. (a) Schematic of polymer architecture where the polymer 'originates' from a segment denoted *origin*. Each generation refers to the number of parent-child relations between the segment and the origin. The origin is designated as generation 0. Any segment without any children is referred to as an endling. (b) Schematic of a weak polymer represented by freely

jointed hard spheres that can be in either neutral of charged states of valence Z=0 and Z=+1 for a basic monomer (or Z=-1 for an acidic monomer), respectively. Here, W stands for the potential of mean force between neighboring segments and σ_p is the size of the monomer.

For linear polyacids or polybases with *M* identical ionizable sites, analytical expressions can be readily derived to describe the ionization behavior[8]. In this case, we are often concerned only with the average degree of ionization, which is calculated from $\alpha = \sum_k \alpha_k / M$. In principle, the transfer matrix technique is applicable to weak polyelectrolytes with more complicated architectures. As shown in Figure 1, we may organize the ionizable groups from a polymer of arbitrary architecture in terms of a "parent-child"-like structure[26]. In this labeling scheme, we first assign a polymer segment (no special requirement on this segment) as the *origin* (i.e., the segment in which all other segments branch from). Then, segments immediately connected to the origin segment are labeled as generation 1 (i.e., the children of generation 0). Thus, the origin segment is also the parent to these segments in generation 1. If the segments in generation 1 have children, then they will be the parents to the segments in generation 3 and so on. However, if the parent-child relation ends (i.e., there is no child for a particular generation path), then this last segment is denoted as an *endling* and is the child of the previous generation. A segment is at a *generation G* if the parent-child relation transfers *G* times between this segment and origin. For consistency, we define the generation of the origin as zero. The key point of this distinction is that we can find one and only one path from the origin to the segment of interest, or from the segment of interest to the origin.

When the ionizable sites in a weak polyelectrolyte are arranged in a parent-child structure, we can define a *restricted* partition function for parent *p* with *k* children each with their own restricted partition functions[27]

$$\Xi'_p\left(s_p\right) = z_p^{s_p} \prod_{k=\text{child}(p)} \left[\sum_{\{s_k\}} u_{pk}^{s_p s_k} \Xi'_k\left(s_k\right) \right]. \tag{6}$$

Eq. (5) corresponds to the partition function of the parent with a fixed (restricted) value of the state variable s_p for the parent segment, where z_p is the reduced activity defined as $z_p = \mathrm{K}'_p a_{H^+}$ and $u_{pk} = \exp[-\beta W_{pk}]$. The product accounts for the contributions of different child structures, and the summation is over the different possible states of child *k*. Recursive application of the *restricted* partition function to the entire polymer leads to an exact partition function for weak polyelectrolytes of arbitrary tree-type architecture. To carry out such calculations, we introduce variable $\tau_p = \Xi'_p\left(1\right) / \Xi'_p\left(0\right)$ and the recursion relation[28]

$$\tau_p = \frac{\Xi'_p\left(1\right)}{\Xi'_p\left(0\right)} = z_p \prod_{k=child(p)} \frac{\left[\Xi'_k\left(0\right) + u_{pk}\Xi'_k\left(1\right)\right]}{\left[\Xi'_k\left(0\right) + \Xi'_k\left(1\right)\right]} = z_p \prod_{k=child(p)} \frac{\left[1 + u_{pk}\tau_k\right]}{\left[1+\tau_k\right]} \ . \tag{7}$$

The total partition function for the entire polymer can then be written in a recursive manner by designating a specific segment as the origin:

$$\Xi = \left(1+\tau_o\right) \prod_{k=\text{child}(o)} \Xi_k \tag{8}$$

where Ξ_k is the partition function for child k who is the parent for its own child *j*

$$\Xi_k = \left(1+\tau_k\right) \prod_{j=\text{child}(k)} \Xi_j \ . \tag{9}$$

The average number of ionized segments, $v_r = \sum_i^M \alpha_i$, can be found from the application of Eq. (4) to Eq. (7), leading to the recursive relation[28]

$$v_i = \frac{b_i \tau_i}{1+\tau_i} + \sum_{k=\text{child}(i)} v_k \tag{10}$$

where

$$b_i = \frac{z_i}{\tau_i}\frac{\partial \tau_i}{\partial z_i} = 1 + \sum_{k=\text{child}(i)} \frac{\left(u_{ik}-1\right)\tau_k b_k}{\left(1+u_{ik}\tau_k\right)\left(1+\tau_k\right)} \quad . \tag{11}$$

The recursion relation is carried out from every endling to the origin generation by generation, which means that we must first calculate the quantities at maximum *generation* G, and then calculate these at levels G-1, G-2, …, down to generation 1. The variable τ_i and b_i are initialized to z_i and 1, respectively, for all endling segments. This results in a straightforward and fast procedure to determine the degree of ionization for any weak polyelectrolyte.

The parent-child resummation scheme discussed above is most useful for polymers with highly branched structures. While this technique is applicable to polymers of any architecture that can be arranged such that there is only one unique parent to a child (i.e., no loops), it is not necessary for linear chains such as poly(acrylic acid) because an analytical expression for the partition function is available. When stereochemistry must be considered and the connection types (i.e., mesomeric and racemic) are chosen according to a prescribed probability distribution, we can also use the parent-child resummation method with one of the end segments of the linear chain assigned as the origin.

2.2 ***Molecular thermodynamic model***

As discussed above, the conventional site-binding model accounts for the solution effects on weak polyelectrolyte ionization in terms of two adjustable parameters, i.e., the ionization free energy of an isolated monomer and the nearest-neighbor interaction energy. Whereas the theoretical procedure is convenient for fitting the experimental titration curves, a thermodynamic model is needed to relate these parameters to the solution conditions. Towards that end, we employ the primitive model for polyelectrolyte solutions to account for electrostatic interactions, short-range attractions, and molecular excluded volume effects. As shown schematically in

Figure 1, a weak polyelectrolyte is represented by a freely-jointed tangent hard spheres such that each segment is ionizable dependent upon the local solution condition. Unlike typical coarse-grained models used in MC simulation, the salt ions are explicitly considered as charged hard spheres while the solvent is represented by a dielectric background. At room temperature, the dielectric constant of liquid water is set equal to 78.4. The hydrated diameter for the salt ions were taken from Simonin et al.[29]. For simplicity, we assume that the hard-sphere diameter for each segment is not influenced by the ionization status. For all systems considered in this work, the calculated results were relatively insensitive to the ion diameters.

Thermodynamic non-ideality plays a significant role in both one- and two-body potentials of the site-binding model. In general, we may decompose the one-body energy into chemical and physical contributions, i.e., the apparent equilibrium constant can be separated into contributions due to protonation reaction and non-bonding interactions:

$$\mathrm{pK}_{\mathrm{i}}' = \mathrm{pK}_{\mathrm{i}}^{\mathrm{T}} + \frac{1}{k_B T \ln 10}\left[\mu_i^{ex}\left(s_i = 1\right) - \mu_i^{ex}\left(s_i = 0\right)\right] . \quad (12)$$

The first term on the right side of Eq. (11) is affiliated with the thermodynamic equilibrium constant for the ionization of an isolated monomer (i.e., an intrinsic property of the segment independent of charge status of the entire polymer) at a reference state of infinite dilution. In principle, the thermodynamic equilibrium constant can be predicted from the Gibb's energy of formation for individual monomers in the solution by using first-principles methods. However, quantum-mechanical calculations are computationally demanding and the results are sensitive to the theoretical details[30]. Alternatively, the thermodynamic constant can be determined by fitting the thermodynamic model with experimental titration curves for specific polymer systems under consideration.

The second term on the right side of Eq. (11) represents the change in the excess chemical potential of the monomer upon ionization. This term arises from the physical interactions of the ionizable site with all other chemical species in the environment thus changes with the solution condition. Within the thermodynamic framework discussed above, the change in the excess chemical potential can be expressed in terms of the electrostatic correlations (el) and local electric potential (charge)

$$\Delta\mu_i^{ex} = \mu_i^{ex}\left(s_i = 1\right) - \mu_i^{ex}\left(s_i = 0\right) = \Delta\mu_i^{el} + \Delta\mu_i^{charge} \quad . \tag{13}$$

Eq. (12) does not include contributions due to volume exclusion because, as mentioned earlier, the segment size is assumed to be the same regardless of whether the ionizable site is charged or neutral. The neglect of the size effect is obviously a simplification since the hydration structure may change upon protonation or deprotonation[31]. However, the change in the particle size would have substantial influence on the excluded-volume effects only in concentrated electrolyte solutions[32]. Besides, the hydration of a segment within a polymer chain is most likely not the same as that for an isolated monomer. With the assumption that ionization has negligible excluded-volume effects, the change in excess chemical potential is therefore dominated by electrostatic interactions and correlation effects.

By treating the polymer as ionizable hard-sphere chains, we can derive analytical expressions for both the local and bulk thermodynamic properties of weak polyelectrolytes from the classical density functional theory (cDFT)[33]. The excess chemical potential due to electrostatic correlations results from interactions with other ionizable groups in the same polymer chain and with all charged species in the solution. For each polymer segment, the ionization leads to a non-uniform distribution of free ions and other polymer segments in its surrounding environment. Such effects can be described through the mean-spherical

approximation (MSA)[34]. Importantly, MSA reduces to the Debye-Hückel (DH) limiting law at dilute electrolyte concentrations[35].

According to MSA, the change in excess chemical potential due to electrostatic correlations between the ionized monomer and the uncharged monomer is expressed as

$$\beta\Delta\mu_i^{el} = -l_B\left[\frac{Z_i^2\Gamma + 2Z_i\eta\sigma_i}{1+\Gamma\sigma_i}\right] \tag{14}$$

Where $l_B = \beta e^2 / 4\pi\varepsilon_0\varepsilon_r$ is the Bjerrum length (7.14 Å for liquid water at room temperature), e is the elementary charge, and ε_0 and ε_r are the vacuum permittivity and the dielectric constant of the solvent, respectively; σ_i and Z_i are the hard-sphere diameter and valence of the ionic species, η is a parameter related to the asymmetry of the system (i.e., differences in size and valence of the ionic species), and Γ is the MSA screening parameter. Approximately, Γ is proportional to the square root of the ionic strength similar to that for the Debye screening parameter. As a result of this square root dependency, the importance of electrostatics becomes relevant even at dilute conditions. A full set of MSA equations is given in Supporting Information.

The polymer ionization also leads to a change in the segment excess chemical potential affiliated with the local electric potential, ϕ^{local}. For an uncharged segment, the local electric potential is approximately zero unless it is directly connected with a charged segment. Upon ionization, the local electric potential will be dependent on the segment size and the salt concentration. The change in the excess chemical potential due to ionization is given by the difference in the self-energy of the charged segment (i.e., the Born energy)

$$\Delta\mu_i^{charge} = eZ_i\Delta\phi_i^{local} = eZ_i\left[\phi_i^{local}\left(I\right) - \phi_i^{local}\left(I\to 0\right)\right], \tag{15}$$

where Z_i is the valence of the charged segment, I is the ionic strength $I = \sum_k c_k Z_k^2 / 2$, and c_k is the molar concentration of species k. The local electric potential must be taken relative to that corresponding to the infinite dilution (i.e., $I \to 0$) because it is used as the reference state for the thermodynamic equilibrium constant. The contribution to the change in excess chemical potential is always negative since the electrostatic energy is the largest at infinite dilution. In principle, the local electric potential is also dependent upon the position of the ionizable site within the polymer backbone. Because such considerations would be cumbersome from a computational perspective, we treat all segments identical and assume that the local electric potential is the same as that corresponding to an isolated monomer in the polyelectrolyte system. The latter can be estimated from cDFT calculations by fixing a particle of an absolute charge of e and hard-sphere diameter σ at the origin. As demonstrated in our previous work[33], cDFT provides an accurate description of the inhomogeneous distribution of ions in an aqueous environment.

As indicated before, the ionization of weak polyelectrolytes differs from its monomeric counterparts primarily due to the presence of electrostatic interactions between neighboring sites. For homopolymers, the intrachain correlation impedes the charging of ionizable groups in the same polymer chain and gives rise to a step-like titration curve as commonly seen in polyacids[11]. To describe the interaction between adjacent sites, we must account for the extra work to form a pair of charges. This extra work can be separated into non-electrostatic and electrostatic components, $W^P_{i,i+1} = u_{i,i+1} + W^{el}_{i,i+1}$. The non-electrostatic energy originates from specific interactions such as the breakdown of hydrogen bonding between neighboring pairs[36] or the stabilization of charge like that seen in PMA[11]. Whereas a quantitative prediction of the non-electrostatic energy is difficult to achieve from a molecular perspective, we expect that the short-

range interaction is relatively insensitive to the solution conditions. In this work, the non-electrostatic component is estimated by fitting the experimental data across different solution conditions.

The electrostatic interactions between neighboring segments can be estimated from the first-order thermodynamic perturbation theory (TPT)[37].

$$\beta W^{el}_{i,i+1}\left(\sigma_{i,i+1}\right) = -\ln g^{el}_{i,i+1}\left(\sigma_{i,i+1}\right) \tag{16}$$

where $\sigma_{i,j} = (\sigma_i + \sigma_j)/2$, and $g^{el}_{i,j}$ stands for the electrostatic component of the contact value of the radial distribution function between segments *i* and *j*. The latter is often approximated by the radial distribution function for a reference system of charged hard spheres[38]

$$g^{el}_{i,j}\left(\sigma_{i,j}\right) = \exp\left[-\frac{l_B}{\sigma_{i,j}}\frac{Z_i - \eta\sigma_i^2}{1+\Gamma\sigma_i}\frac{Z_j - \eta\sigma_j^2}{1+\Gamma\sigma_j}\right]. \tag{17}$$

At infinite dilution, the radial distribution function reduces to the exponential of the contact value of the electrostatic pair potential between neighboring segments. Intuitively, Eq. (15) can be understood as the electrostatic work to form an ionic bond as represented by the potential of mean force. As the salt concentration increases, two segments of the same charge become less repulsive due to electrostatic screening, leading to an increased contact value for the radial distribution function. It is important to note that TPT takes into account the chain-connectivity effect on ionization only at the nearest-neighbor level. Nevertheless, unlike the original site-binding model (viz., the Ising chain) that ignores the chain conformation completely, the thermodynamic model takes into account the intra-chain correlations through the excess chemical potential of monomers and chain connectivity. Because such correlations are considered at the nearest-neighbor level, our model is not able to describe the conformation

effects adequately. The influence of polymer conformation on ionization is an important problem for future development but beyond the scope of this study.

The thermodynamic procedure discussed above allows us to predict the effects of solution conditions on the one-body energy and nearest-neighbor interaction parameters used in the site-binding model. The one-body energy includes a contribution affiliated with the change in the excess chemical potential of each *monomer* due to ionization. This term accounts for the reversible work to place the monomer at the position of ionization from its ideal reference state. In other words, the excess chemical potential arises from the interactions of each ionizable site with all other species in the solution. Because the excess chemical potential neglects the intra-chain correlation among neighboring sites, TPT is used at the nearest-neighbor level to account for the additional work to bring charged monomers together in the same polymer chain. There is no double-counting in the one-body and nearest-neighbor interactions.

3. Results and Discussion

In the following, we provide three case studies to illustrate applications of our thermodynamic model to describe the effects of solution conditions on ionization of weak polyelectrolytes with different molecular architectures. We first consider the effects of salt types and ionic strength on ionization of poly(acrylic acid), a linear polyacid with a relatively low density of ionizable groups. In this case, the titration behavior has been well studied and can be described with various mean-field methods owing to the weak interactions between neighboring sites[8, 11, 20]. Next, we consider two polyacids with a higher linear charge density, poly(maleic acid) and poly(fumaric acid), which differ in their stereochemical arrangement of the carboxyl groups. Unlike poly(acrylic acid), these polyacids exhibit step-like titration behavior that cannot be described by mean-field methods[11]. Lastly, we investigate the titration properties of

dendrimers at different generations that are composed of titratable amine groups. Normally, the site-binding description of dendrimers requires many adjustable parameters due to large changes in the local chemical environment from the outer to inner regions of the dendrimer. We find that the application of our thermodynamic model to predicting the titration of weak polyelectrolytes is most successful at moderate to high salt conditions. As low salt concentration, a discrepancy emerges between the theoretical results and experimental data due to interactions beyond immediate neighbors that are not explicitly accounted for in our thermodynamic model.

3.1 *Poly(acrylic acid)*

Poly(acrylic acid) (PAA) is a linear weak polyelectrolyte that shows no apparent step-like ionization in the titration curve. Conventional models are able to describe the titration curves provided that both the binding constant and nearest-neighbor energy are allowed to change with solution conditions.

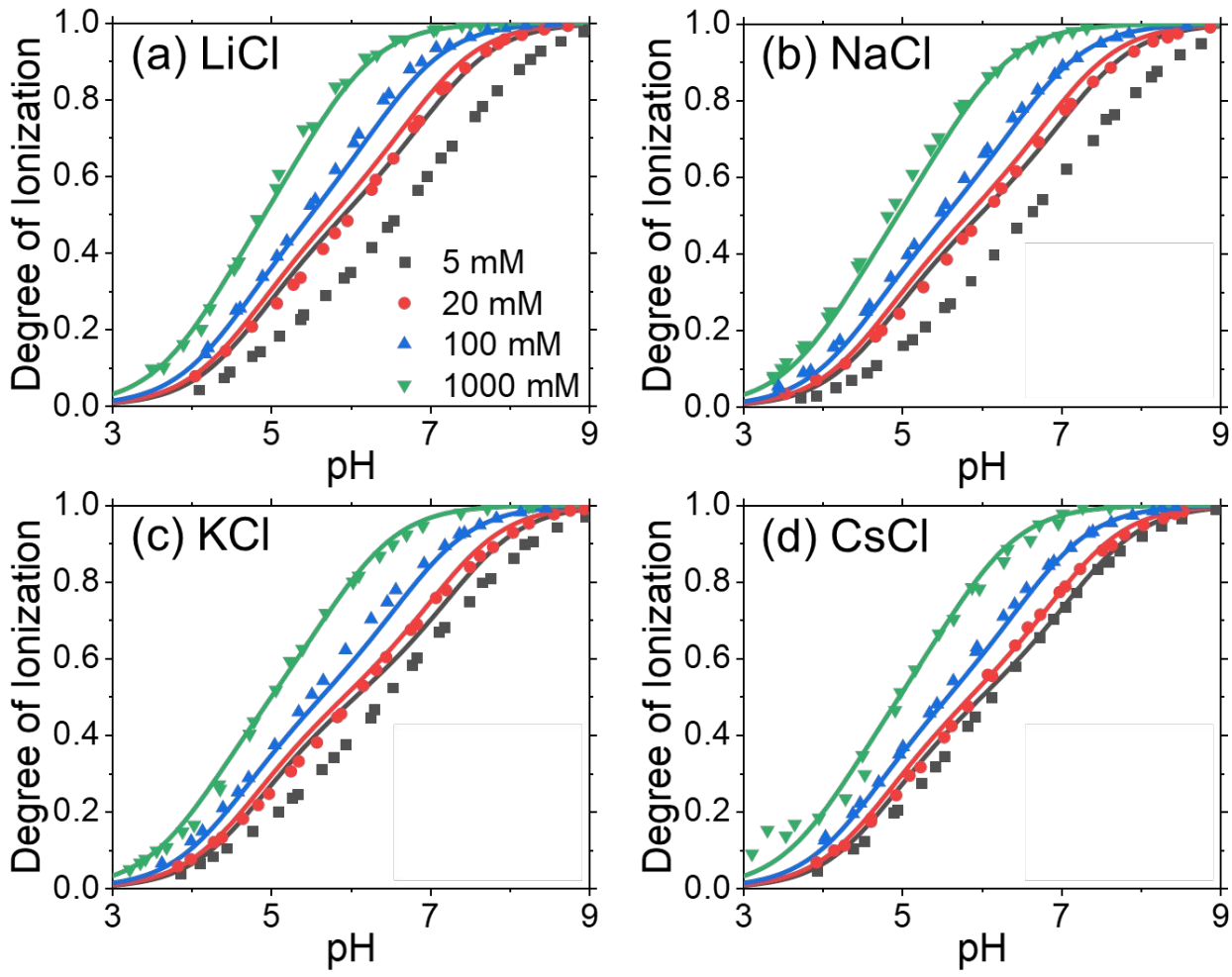

Figure 2. Titration curves for poly(acrylic acid) at 25 °C in aqueous solutions of (a) lithium chloride, (b) sodium chloride, (c) potassium chloride, and (d) cesium chloride solutions from experiment[20] (symbols) and theoretical correlations (lines). As marked in panel (a), different symbols stand for different salt concentrations. The theoretical correlations determined a thermodynamic equilibrium constant of 5.22 and the monomer size was 3.56 Å, 3.37 Å, 3.05 Å, and 3.25 Å, respectively, for the alkali salts.

Figure 2 shows the titration curves of PAA (MW=88 kg/mol[20]) in four types of alkali aqueous solutions (lithium, sodium, potassium, and cesium) at different salt concentrations. The ionization behavior is not significantly influenced by the salt type except at the lowest salt

concentration (e.g., 5 mM). Except for the case of 5 mM salt concentration, our thermodynamic model successfully reproduces the experimental data for all alkali metals. As the salt concentration increases, the polyacid begins to ionize at a lower pH. Due to the electrostatic screening effect, a narrowed pH window is observed between the neutral and fully charged polymer chain. The changes in the position and shape of the titration curves are reflected by the decrease in both the apparent equilibrium constant and the work to ionize two adjacent sites together as determined from the conventional site-binding model (see Table S1-S4), respectively.

At the lowest salt concentration (5 mM), the large discrepancy between theory and experiment is unsurprising because the site-binding model neglects intra-chain correlations beyond the nearest-neighbor level. When interactions beyond adjacent sites become more significant as the salt concentration falls, the monomer ionization is subject to an additional potential due to the presence of non-adjacent sites. For homopolymers, the degree of ionization is reduced due to the additional repulsion. Since the nearest-neighbor model neglects long-range intra-chain correlations, it overpredicts the ionization weak polyelectrolytes in comparison with experiment and the deviation is most significant at low salt concentration. With the apparent binding constant and nearest-neighbor energy treated as adjustable parameters, the site-binding model can capture the titration curve at 5 mM salt concentration (see Figures S1-S4). However, the good agreement should be ascribed to the weak correlation between these sites that allows for the overall titration curve to be approximated at the two-body level when using two adjustable parameters. In particular, it misses the initial ionization of the polymer, which can be attributed to the overestimation of pK' to compensate for interactions between non-neighboring segments neglected in the site-binding model. The mean-field approximation neglects the explicit charge of individual sites. Instead, it treats the energy of each site as an average of the overall

interactions[20]. By contrast, our coarse-grained model demonstrates the importance of long-range interactions that must be considered for a fully predictive model of the titration curves even at low, but not infinitely dilute, salt concentrations.

In correlation of the experimental titration curves for poly(acrylic acid), we use only the thermodynamic equilibrium constant and the monomer size (i.e., the distance between sites) as the adjustable parameters. Unlike the conventional site-binding model, these parameters are intrinsic properties of the thermodynamic system independent of the solution conditions. It was found that the addition of a non-electrostatic pair interaction did not improve the quality of the fitting results (see Figure S3). For the results shown in Figure 2, the experimental data for the salt concentration of 5 mM were not included in the fitting procedure because our model does not adequately account for intra-chain correlations between next-nearest neighbors and beyond. The thermodynamic equilibrium constant was constrained to be the same for all alkali metals because it is an intrinsic property of the polymer (i.e., independent on the specific salt present in the solution). The pK_a value, 5.22 ± 0.02, is greater than that corresponding to an isolated monomer, $pK_a = 4.756$ because of the influence of neighboring segments[39]. The elevated pK_a value agrees with previous quantum-mechanical calculations[40, 41]. The polymerization leads to an increased equilibrium constant for deprotonation, i.e., the ionizable sites are less acidic in the polymer than the corresponding isolated monomeric counterparts. The apparent equilibrium constant shows a stronger dependence on salt concentration than found for simple carboxylic acids (e.g., $\Delta pK' = pK' - pK$ is approximately -0.6 for PAA versus -0.2 for acetic acid[42]). This can likely be attributed to the high local charge of a polymer compared to that of a simple acid such that the screening effect is more significant.

According to our thermodynamic model, the distance between adjacent sites was dependent upon the cation and was found to be 3.56 Å, 3.37 Å, 3.05 Å, and 3.25 Å with a 95% confidence interval of 0.12 Å for lithium, sodium, potassium, and cesium cation, respectively. The salt-type dependence of the distance between the neighboring ionizable groups could be attributed to the differences in hydration of the alkali ions which influences the hydration of the polymer sites when the cations approach the negatively charged monomers in PAA. The specific chemical details that lead to this difference are not captured by the primitive model. For example, there is a noticeable difference between the titration curves for lithium chloride versus cesium chloride at the same salt concentration of 5 mM. One consideration is that even though both cations are monovalent, they differ significantly in hydrated diameters (4.76 Å versus 1.89 Å)[29]. The larger hydration shell for lithium ions is partially responsible for the larger distance between adjacent sites due to their condensation at the polymer backbone. At high salt concentrations, the specific ion effect becomes less important because the electrostatic interactions are dominated by the overall screening of all ionic species in the system. Further improvement of our model would be possible by adopting an effective potential to distinguish different alkali ions besides size and valence[32]. Nonetheless, our results demonstrate that the incorporation of thermodynamic non-ideality into the nearest-neighbor site-binding framework can correctly account for the titration behavior of poly(acrylic acid) in different aqueous solutions.

3.2 *Poly(maleic acid) and poly(fumaric acid)*

Both poly(maleic acid), PMA, and poly(fumaric acid), PFA, have ionizable groups with a linear density significantly larger than that of poly(acrylic acid). As a result, their titration curves exhibit a step-like behavior which is characteristic for many weak polyelectrolytes. The steps in

each titration curve result from alternating ionization of adjacent sites and can be attributed to the Coulombic repulsion and non-electrostatic effects such as hydrogen bonding or electron sharing between functional groups[11, 43]. However, PFA shows significant asymmetry in its titration curve below and above 50% of ionization and a less prominent step-like feature. Interestingly, PMA and PFA differ only in the stereochemical arrangement of their carboxyl groups. For PMA, every second bond is racemic (i.e., neighboring carboxyl groups are on the same side along the C-C bond) and about one-third from the remaining ones (i.e., the odd bonds) are mesomeric (i.e., neighboring carboxyl groups are on the opposite side across the C-C bond). By contrast, in PFA every second bond is mesomeric and only about half of the remaining ones are racemic[43, 44]. Due to the closer distance between adjacent sites for a racemic bond compared to that for a mesomeric bond, we expect that, on average, the interaction between neighboring sites in PMA will be stronger than that in PFA. It is important to clarify that whether the odd bond is racemic or mesomeric can only be given in terms of probability. For example, the probability of the odd bond being racemic is 0.655 for PMA and that for PFA is 0.54[45]. Thus, there is no regular pattern in the PMA backbone structure. Previous application of the site-binding model with regularity in the structure found the resulting titration curves to be unsatisfactory[43]. We expect that the poor agreement is due to the absence of stretches in racemic bonds and mesomeric bonds in PMA and PFA, respectively.

The site-binding model can be applied to describe the titration behavior of PMA and PFA by accounting for the different possible connections (i.e., racemic or mesomeric bonds). The assumption of different connections between adjacent sites on the chain means different strengths for the pair interactions, dependent upon if the connection is by a racemic or mesomeric bond. Following the method proposed by Groot et al.[45], we may select the odd bonds

to be either racemic or mesomeric by drawing pseudorandom numbers using the propagation rate P_r as a decision threshold level. The degree of dissociation converges after evaluating for approximately 8000 consecutive sites on the chain. Since the two weak polyelectrolytes are different only in the sequence and the numbers of racemic and mesomeric bonds, we assume that the distance between neighboring sites depends only on the type of bonds but not the polymer. In that case, the same set of model parameters can be used for both PMA and PFA. We obtained these parameters by fitting the thermodynamic model with the titration data at different solution conditions[11, 43, 45].

Unfortunately, due to the use of a hard-sphere chain model, the different distances between racemic and mesomeric bonds introduces an inconsistency. Because the hard-sphere diameter of the individual segments cannot account for this difference, we make the monomer size equal to the distance between two sites connected by a racemic bond. The work to charge two adjacent sites connected by a mesomeric bond is then considered through an analytical approximation by Henderson et al.[46]

$$\beta W^{el}_{i,i+1}\left(d_{i,i+1}\right) = -\ln g^{el}_{i,i+1}\left(d_{i,i+1}\right) = \frac{l_B}{d_{i,i+1}} \frac{Z_i - \eta\sigma_i^2}{1+\Gamma\sigma_i} \frac{Z_{i+1} - \eta\sigma_{i+1}^2}{1+\Gamma\sigma_{i+1}} \exp\left[-\kappa\left(d_{i,i+1} - \sigma_{i,i+1}\right)\right] \qquad (18)$$

where $d_{i,i+1}$ is the distance between two sites connected by a mesomeric bond and $\kappa = \sqrt{8\pi l_B I}$. The pK_a values for PMA and PFA were 4.47 and 4.16 with a 95% confidence interval of ± 0.04, respectively, and the distances between adjacent sites characterizing the racemic and mesomeric bonds were determined to be 0.69 ± 0.01 Å and 1.95 ± 0.06 Å, respectively. Because PMA and PFA have a higher linear density of ionizable sites compared with PAA, the distances between neighboring sites are much smaller than that for PAA (e.g., 3.57 Å for LiCl). At such short distances, the assumption that the Coulomb interaction is governed by the bulk dielectric

constant of 78.5 is likely not valid[47]. If a lower dielectric constant were assumed to govern this interaction, the distance parameter would need to be larger in order to reproduce the same interaction energy. Considering that the description of inhomogeneous dielectric behavior cannot be well described through conventional methods, we use the small distances as a zeroth-order approximation. The lower dielectric constant of the interchain interaction may play a role in PAA as well, but to a lesser degree owing to the larger distance between adjacent sites.

In addition to the electrostatic component of the pair interaction, our thermodynamic model includes a non-electrostatic pair potential of -4.53 ± 0.18 k_BT between adjacent sites connected by each racemic bond in PMA. It is important to note that this energy will not vary with salt concentration and counteracts some of the electrostatic repulsion ($W^{el} \sim +10\ k_BT$). It has been suggested that the nearest-neighbor attraction originates from the local stabilization of charge where a proton can be shared between an uncharged carboxyl group and a neighboring ionized group connected by a racemic bond[36, 48]. Such interaction does not exist between adjacent sites connected by a mesomeric bond since it is sterically impossible. Due to the stereochemistry of PFA, two racemic bonds never occur in a row. On the other hand, the racemic bonds in PMA often appear in series leading to significant charge stabilization by proton sharing. Thus, we see that stereochemistry can make an important difference in the titration behavior of weak polyelectrolytes, despite otherwise sharing the same chemical composition.

Figure 3 compares the theoretical and experimental titration curves for PMA (MW=17.6 kg/mol[43]) and PFA (MW=6.6 kg/mol[11]) in pure water and in aqueous solutions of LiCl and NaCl. Our thermodynamic model performs quite well at moderate to high salt concentrations. Interestingly, the ionization of PMA and PFA extends over 9 pH units owing to the strong intra-chain repulsion. Due to the difference in stereochemistry between the two polymers, there is a

noticeable contrast between the titration curves. While PMA shows a short plateau in the titration curve when it is half ionized, no such feature is observed in the titration curve for PFA. The difference between the two titration curves can be understood by the larger proportion of the weakly repulsive mesomeric bonds in PFA compared to that in PMA. Two charged sites connected by a racemic bond in PFA show more significant repulsion than that of PMA, leading to an extended titration curve in the late stage of ionization.

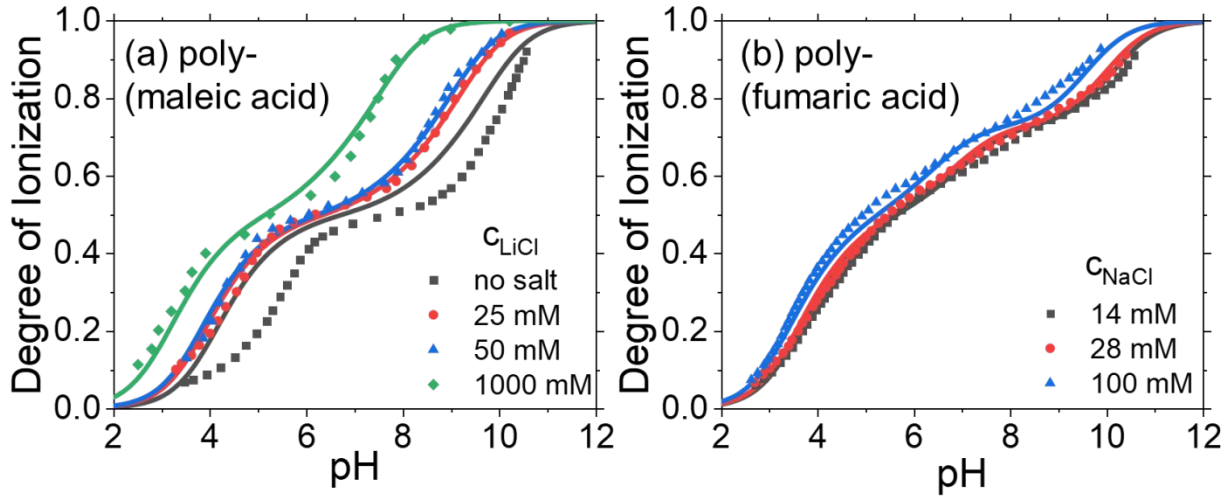

Figure 3. The degree of ionization for (a) poly(maleic acid) in pure water and in three lithium chloride solutions and (b) poly(fumaric acid) in three sodium chloride solutions according to experiments[11, 43] (symbols) and theory (lines).

As was seen for PAA, the truncation of the intrachain correlations to only those between nearest neighbors becomes problematic at low to dilute salt conditions. Recently, an efficient approach was demonstrated by Garcés and coworkers using the Gibbs-Bogoliubov variational principle to incorporate intra-chain correlations via an effective one-body potential[49, 50]. The agreement between the exact results and this new approach was quantitative when applied to weak polyelectrolytes with rod-like configurations. While this new method is promising, its application to realistic systems has not yet been achieved. Alternatively, Ghasemi and Larson

utilized the random phase approximation (RPA) to account for the intramolecular electrostatic correlations of polymers with a specified structure (e.g., coil-like)[51]. The RPA method incorporates the intramolecular correlations through an effective one-body potential which utilizes the average behavior of the ionizable segments. Thus, unlike our model employed here, it cannot capture the step-like titration since the explicit charge status of neighboring sites (viz. paired charge states of neighbors) is not considered. Nonetheless, this method is a key step towards a universal description of weak polyelectrolytes by accounting for contributions beyond nearest neighbors in a convenient way. Unfortunately, the use of a fixed structure of the polymer is problematic as the conformation is dependent upon the polymer charge. The coupling of conformation and ionization is not easily described with analytical models because both electrostatic and non-electrostatic (e.g., hydrophobic) interactions determine the polymer configuration. While the polymer tends to expand upon ionization to reduce the repulsive Coulomb interactions, the hydrophobic interactions cause the polymer to collapse upon itself. Future work in these directions could lead to a more accurate description of weak polyelectrolytes across all solution conditions.

The choice of counterions is of particular importance for the ionization of weak polyelectrolytes with a high linear charge density. We show the titration curve of poly(maleic acid) in three different 50 mM salt solutions in Figure 4. While the charging of the polymer with pH is relatively independent of the salt type at a low degree of ionization, the choice of salt cation plays a major role above 50% ionization where nearest-neighbor interactions become important. While lithium chloride shows only a small plateau when it is half charged, tetramethylammonium (TMA) chloride exhibits a plateau in ionization spanning across nearly 5 pH units. Interestingly, even at pH 11, the polymer is still only 75% charged when TMA is used

as the counterion. To model the corresponding titration curves, we used the pK_a and distance parameters for the racemic and mesomeric bond determined previously for PMA in lithium chloride solution and adjusted the racemic and mesomeric non-electrostatic energies to account for the salt type effect.

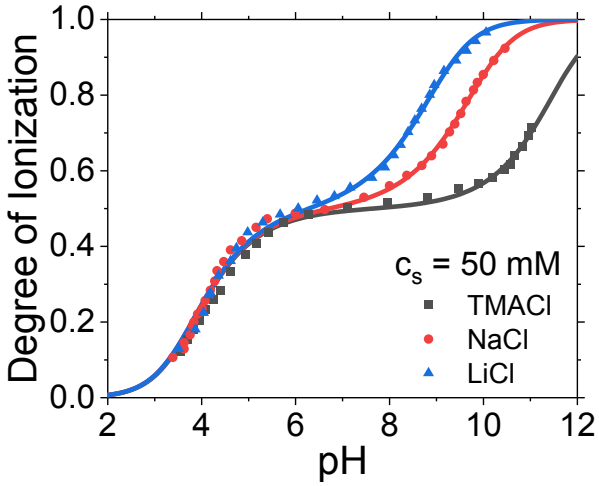

Figure 4. The ionization behavior of poly(maleic acid) in tetramethylammonium chloride, sodium chloride, and lithium chloride at 100 mM salt concentration according to experiments[11, 43] (symbols) and theory (lines).

While some variation in the distance between adjacent sites may exist for different salt types as was seen previously for PAA, we expect that the influence of salt type is mostly non-electrostatic in origin. This has been suggested previously by the conventional site-binding model which found the pair interaction for TMACl to unexpectedly increase with salt concentration for both the racemic and mesomeric bonds[45]. For TMACl, NaCl, and LiCl, the interaction energies were determined to be -1.90, -3.72, and -4.53 k_BT, respectively, for the racemic bond. For NaCl and TMACl, it was necessary to include a non-electrostatic energy of +1.99 k_BT and +4.86 k_BT between the adjacent sites connected by a mesomeric bond as well. As expected from the titration curves, the repulsion between neighboring sites increases as the

counterion is changed from lithium to tetramethylammonium. This can be understood by the differences in the hydration of the counterions since lithium is well solvated while TMA is poorly solvated, leading to a modulation of the local water structure around the polymer. The importance of salt type in the two-body interaction reflects the need for a more comprehensive understanding of the solvation effects.

3.3 *Weak polyelectrolytes with branched architecture*

Branched polyelectrolytes behave quite differently from their linear counterparts because of additional ionizable neighbors at the branching points. For example, a linear poly(ethyleneimine) chain is fully charged below pH=2 but a substantial number of amine groups remain uncharged when the polymer has a branched architecture[28]. While the titration behavior is affected by the chemical nature of the amine groups with different coordination numbers, such effects were considered as relatively minor in comparison with that due to the topological arrangement of the ionizable sites[8].

In comparison with linear chains of similar molecular weight, dendrimers have a more compact internal structure thereby more interactions among ionizable sites in the same polymer. A dendrimer is typically described as an onion with each shell from the center containing double the number of sites within the previous shell. As illustrated in Figure 1, we may label the total number of shells as generations, designated by letter *G*. If all sites are ionizable, the total number of ionizable groups is 2^{G+2}-2; and the number of ionizable groups in the outermost shell is 2^{G+1}. The ionization of a dendrimer proceeds first through the outmost and every other shells, which constitute two-thirds of the total number of ionizable groups, then followed by the ionization of the remaining shells. However, to ionize the remaining shells is more difficult than in their linear counterparts because there are three pair interactions in the dendrimer case versus the two

pair interactions in the linear case. Previous investigations based on the site-binding model, MC simulation and SCFT calculations demonstrated that the ionization of groups in the outmost shell is easier than that for those groups in the inner shells due to the increased density of ionizable sites[19, 52]. To better understand the ionization behavior of weak polyelectrolyte dendrimers, we must be able to account for not only the topological effects but also the influence of solution conditions.

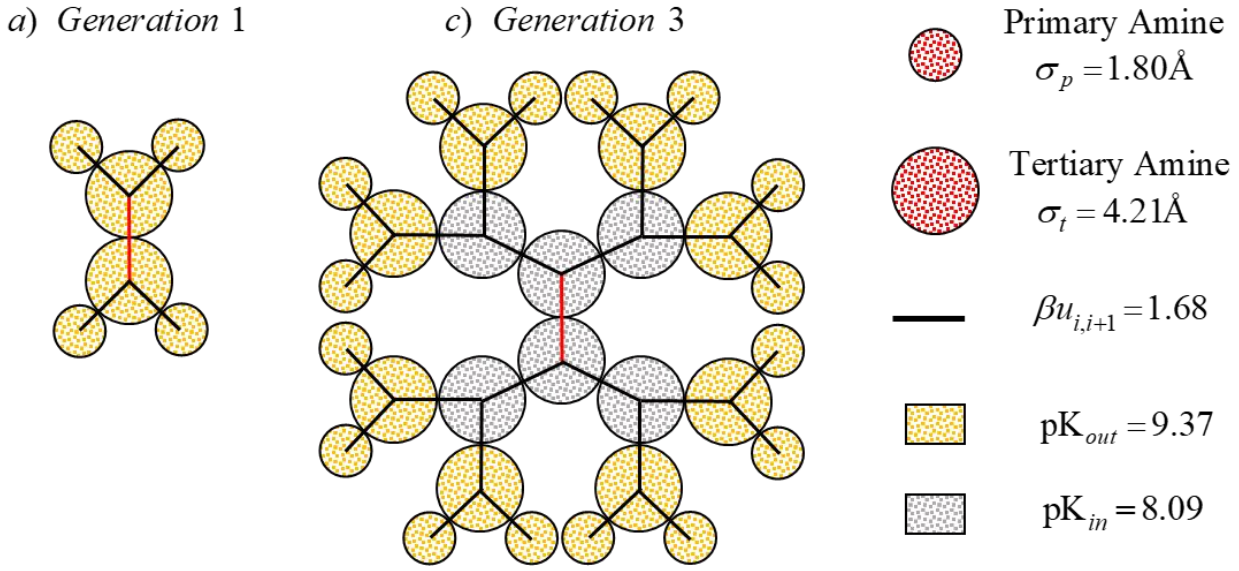

Figure 5. Coarse-grained dendrimer of different generations employed in our molecular thermodynamic model. The equilibrium constant for the two outermost shells of any dendrimer is given by pK_{out} and all other (i.e., inner) shells are given by pK_{in}. For generation 1, both shells are given by pK_{out}. All consecutive charged amines, besides the two in the center, experience a non-electrostatic repulsion of $\beta u_{i,i+1}$.

To demonstrate the application of our molecular thermodynamic model to weak polyelectrolytes with branched architecture, we consider the acid-base equilibrium of 1,4-diaminobutane poly(propylene imine) (PPI) dendrimers at different salt concentrations. van Duijvenbode et al. studied the titration curves for PPI dendrimers of generation one to five in

three sodium chloride solutions of 0.1 M, 0.5 M, and 1.0 M using experimental techniques and a site-binding model[53]. The original site-binding model employed four equilibrium constants and three pair interaction energies to account for different locations of amine groups (i.e., the shell position) within the dendrimer. The large number of parameters are required because of significant differences in the chemical environment from the outer shell to the inner shell of the dendrimer. As shown in Figure 5, we use only two equilibrium constants to represent the ionizable behavior of the dendrimers. Besides, all parameters are independent of salt concentrations. Specifically, our model employs two size parameters for the amine groups: one for the primary amine (those in outermost shell) and one for the tertiary amines (all others). Their hard-sphere diameters are given by $\sigma_p = 1.80 \pm 0.22$Å and $\sigma_t = 4.21 \pm 0.47$Å, respectively. One additional parameter, $\beta u_{i,i+1} = 1.68 \pm 0.13$, is used to account for non-electrostatic repulsion between the ionizable groups. For the generation 1 dendrimer, the equilibrium constant of all amines is the same, $\mathrm{pK}_{out} = 9.37 \pm 0.01$. For higher generation dendrimers, we find that all ionizable groups except those in the two outermost shells can be described by $\mathrm{pK}_{in} = 8.09 \pm 0.06$. The different equilibrium constants reflect significant changes in the local chemical environment of the dendrimer in its inner region from the outer region. We see that our coarse-grained model greatly simplifies the depiction of a dendrimer employed in the original site-binding model and provides physical meaning to the correlated parameters.

Figure 6 compares the theoretical and experimental titration curves for PPI dendrimers of generations one and five at 0.1, 0.5, and 1.0 M sodium chloride concentrations. The titration results for dendrimers of total generations two, three, and four can be found in Figure S7. Overall, the titration behavior of dendrimers is well accounted for through the coupling of our thermodynamic model with the site-binding description of ionization in weak polyelectrolytes.

In contrast to a polyacid, a polybase reduces its charge as the pH increases because it is charged in the protonated form. The increase in salt concentration shifts the titration curve to higher pH because of electrostatic screening.

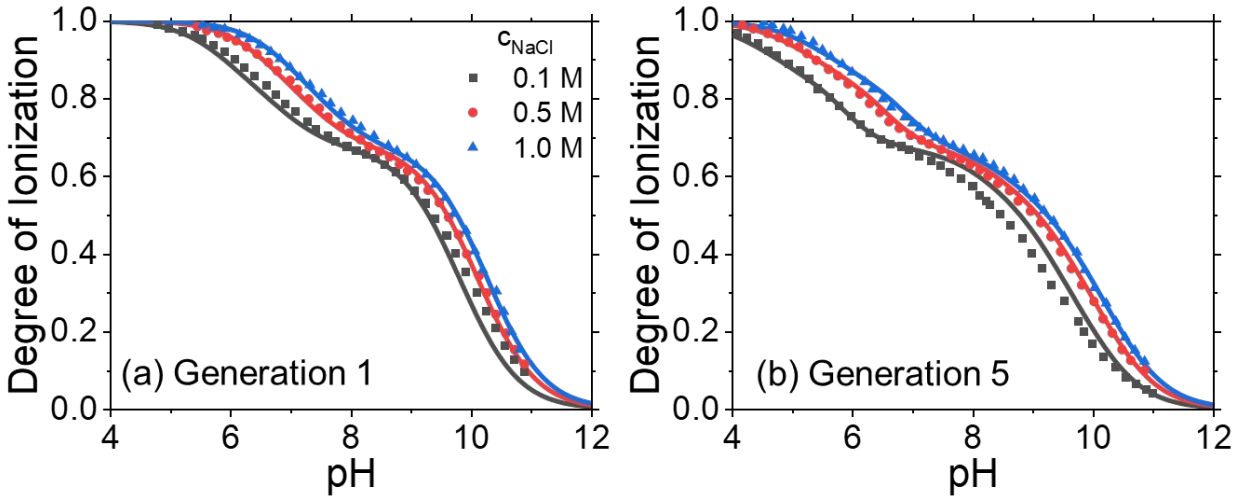

Figure 6. The titration curves for 1,4-diaminobutane poly(propylene imine) (PPI) dendrimers in aqueous solutions at three concentrations of sodium chloride according to experiments[53] (symbols) and theory (lines).

In the conventional site-binding model, the parameters must vary with the solution condition (e.g., salt concentration) in order to reproduce the experimental titration data. However, with the assistance of the thermodynamic model, we use only the size of the titratable group (i.e., primary and tertiary amines) to account for the effects of the solution condition on the apparent equilibrium constant and pair interaction energies. We fit the theoretical results to the titration curves for different dendrimer generations at 0.5 and 1.0 M sodium chloride concentrations. The interactions with non-neighboring segments seem to be relevant at 0.1 M as indicated by the overestimation in the ionization when the generation five is weakly charged. The excellent agreement between the theoretical results and experimental data indicates that the physics of weak dendrimer ionization is well described by incorporating the electrostatic inter-

and intra-correlations as well as the influence of changes in local electric potential into the site-binding model.

The theoretical prediction shows more noticeable deviation from the experimental data at a small degree of ionization while it is more quantitative in the later stages of ionization. In the early stage of charging, the interactions of segments with non-neighboring segments are of more significance as there are not sufficient ions drawn into the dendrimer to screen these interactions at low salt concentrations. This is why our results overestimate the ionization at 0.1 M for the generation 5 dendrimer since only the correlation with neighboring segments is considered explicitly. For low generation dendrimers, the discrepancy between experiments and theory is reduced as the dendrimer is less compact so the influence of non-neighboring segments is diminished. One exception is the generation two dendrimer (Figure S7b) which shows a consistent overestimation of the ionization for all three salt concentrations beyond the condition when the polymer is two-thirds ionized. The reason for this is that the equilibrium constant of the middle (i.e., second outermost) shell is not well represented by pK_{out}, but by a value ~0.3 units lower (viz. pK = 9.1). This likely reflects generation two as a transition from a small molecule (G=1) where the local environment is approximately the same for all amines to a large molecule ($G \geq 3$) which shows two clear regions (viz. inner and outer) as reflected by their own pK values.

4. Conclusions

We developed a molecular thermodynamic model for predicting the ionization of weak polyelectrolytes by incorporating the site-binding model with classical density functional theory (cDFT). The new model accounts for the local thermodynamic condition for individual ionizable sites through the separation of the apparent equilibrium constant into two components: a

chemical contribution, i.e., the thermodynamic equilibrium constant for proton binding/dissociation of ionizable functional groups, and physical contributions due to non-bonding interactions (e.g., electrostatics) with all chemical species in the solution. The influence of solution conditions on the one-body term is affiliated with the electrostatic correlations as determined by the mean spherical approximation (MSA) and with the change in the local electric potential upon ionization of the monomer as determined by cDFT. In addition, we account for the non-ideality due to electrostatic correlations between neighboring monomers within the polymer. The electrostatic work to form a pair of charged sites can be related to the negative natural logarithm of the radial distribution function which is determined from the reference charged hard-sphere system. The incorporation of the thermodynamic non-ideality into the site-binding model allows us to describe the titration behavior of weak polyelectrolytes with different arrangements of ionizable sites and configurations.

The molecular thermodynamic model provides a theoretical basis to estimate the parameters employed in previous applications of the site-binding model. While quantities like the intrinsic binding constants (viz., pK_a values) must be determined by fitting with experimental data, they maintain physically realistic values and are invariant with the solution conditions. We demonstrate the *predictive* application of our thermodynamic model by extensive comparison to experimental titration behavior of weak polyelectrolytes across different solution conditions and polymer architectures. Specifically, we first consider the linear weak polyacid, poly(acrylic acid) (PAA), in various solutions of alkali chloride ranging in salt concentration from 5 mM to 1 M. It is found that the titration behavior at moderate to high salt concentration (i.e., greater than or equal to 20 mM) can be described through two parameters: the thermodynamic equilibrium constant and the size of the monomer. The monomer size varies with the types of alkali cations

present in solution likely owing to differences in their hydration behavior. At the lowest salt concentration considered in this work, the importance of intra-chain correlations beyond nearest neighbors cannot be ignored and noticeable deviation is seen between the theoretical values and experimental data. A similar trend was found for the linear weak polyacids, poly(maleic acid) (PMA) and poly(fumaric acid) (PFA), which have a much greater density of carboxyl groups along their backbones in comparison with PAA. Despite the strong correlations between neighboring monomers in PMA and PFA, our model agrees well with the experimental titration data at moderate to high salt concentrations. As a final example for the application our model to branched weak polyelectrolytes, we considered the titration behavior of poly(propylene imine) dendrimers. In spite of the complex structure of dendrimers, the influence of solution conditions on the degree of ionization was quantitatively captured by our thermodynamic model. Thus, the inclusion of inter- and intra-chain correlations into the one- and two-body terms of the conventional site-binding model, respectively, provides a feasible method to account for the salt concentration effect on the titration behavior of different weak polyelectrolytes.

According to the site-binding model, the influence of chain length on the ionization behavior of weak polyelectrolytes is significant when the end-effects are relevant (e.g., $M<20$). However, recent experiments (e.g., Laguecir et al.[18]) indicate that the chain length can be important even for long polymers if there is a strong coupling between ionization and polymer conformation. Such effect is not captured in this work because our calculations are based on the site-binding model but with a predictive one-body energy and the two-body interaction parameter. Indeed, the description of weak polyelectrolytes using our thermodynamic model provides only a first step towards a fully predictive model for weak polyelectrolyte ionization. A major shortcoming of the site binding model is the truncation of the multi-body expansion of the

free energy to only include up to the nearest neighbor two-body term. While previous applications of the site-binding model with nearest neighbor interactions have been used successfully to describe the titration behavior of weak polyelectrolytes at low salt concentrations, our model demonstrates that correlations beyond adjacent monomers do become relevant at such conditions. Therefore, the success of the conventional method must be contributed to the weak correlation between non-consecutive sites. While further improvements can be done by accounting for contributions beyond nearest-neighbor interactions systematically, the theoretical description of these interactions is not well understood, and the incorporation of multi-body correlations is computationally challenging. Future work incorporating such correlations may glean useful insights from the charging behavior of weak polyelectrolytes in dilute aqueous solutions.

The modified site-binding model discussed in this work can be readily extended to weak polyelectrolytes with different types of ionizable groups. In comparison with that for a monomer, the ionization of a polymer segment can be either prohibited or promoted in a heterogeneous weak polyelectrolyte because the extra work to ionize adjacent sites is positive for like-charges and negative for unlike-charges. In addition to electrostatic interactions and excluded volume effects, it can be expected that other non-electrostatic interactions may also be relevant. For example, polyacids such as poly(fumaric acid) exhibit nearest-neighbor interactions that are dependent upon the salt cations in the solution[45]. It was suggested that the variation in the nearest-neighbor energy originates from differences in cation hydration. In addition, the loss of hydrogen bonding mediated by the surrounding water molecules can also be a contributing factor[53]. We expect that, by accounting for the non-ideality in both the electrostatic and non-electrostatic components of the reversible work of ionization, the

thermodynamic model will be able to describe the titration behavior of weak polyelectrolytes with molecular parameters independent of the solution conditions. Moving beyond the site-binding model, polymer density functional theory (PDFT) provides a convenient avenue to analyze the microscopic structure of polymers in solution and near surfaces. We are currently pursuing further development along these directions.

Author contributions: J. W. designed and supervised the research; A. G. performed the calculations; A.G., G. O. and J.W. analyzed the data; and A.G. and J. W. wrote the paper.

Acknowledgements: This work is financially supported by the National Science Foundation Harnessing the Data Revolution Big Idea under Grant No. NSF 1940118. Additional support is provided by the NSF Graduate Research Fellowship under Grant No. DGE-1326120.

Supporting Information:

Full set of equations for determining the excess chemical potential due to electrostatic correlations and changes in the local electric potential. It also includes results from application of the site-binding model to the titration of poly(acrylic acid).

References:

1. H. Priya James, R. John, A. Alex and K. R. Anoop, *Acta Pharmaceutica Sinica B*, 2014, **4**, 120-127.
2. S. Jaganathan, *Artificial Cells, Nanomedicine, and Biotechnology*, 2016, **44**, 1080-1097.
3. B. Bolto and J. Gregory, *Water Research*, 2007, **41**, 2301-2324.
4. S. Seitz and H. Ajiro, *Solar Energy Materials and Solar Cells*, 2019, **190**, 57-64.
5. J. Ng, I. Osborn, D. Harbottle, Q. Liu, J. H. Masliyah and Z. Xu, *Environmental Science & Technology*, 2019, **53**, 6436-6443.
6. M. Kobayashi, M. Terada and A. Takahara, *Soft Matter*, 2011, **7**, 5717-5722.
7. C. Li, Y. Gu and N. S. Zacharia, *ACS Applied Materials & Interfaces*, 2018, **10**, 7401-7412.
8. G. J. M. Koper and M. Borkovec, *Polymer*, 2010, **51**, 5649-5662.
9. T. N. Nekrasova, Y. V. Anufriyeva, A. M. Yel'yashevich and O. B. Ptitsyn, *Polymer Science U.S.S.R.*, 1965, **7**, 1008-1018.
10. S. Ulrich, M. Seijo and S. Stoll, *The Journal of Physical Chemistry B*, 2007, **111**, 8459-8467.
11. T. Kitano, S. Kawaguchi, K. Ito and A. Minakata, *Macromolecules*, 1987, **20**, 1598-1606.

12. E. Gonzalez Solveyra, R. J. Nap, K. Huang and I. Szleifer, *Polymers*, 2020, **12**, 2282.
13. J. Landsgesell, L. Nová, O. Rud, F. Uhlík, D. Sean, P. Hebbeker, C. Holm and P. Košovan, *Soft Matter*, 2019, **15**, 1155-1185.
14. V. S. Rathee, H. Sidky, B. J. Sikora and J. K. Whitmer, *Polymers*, 2019, **11**, 183.
15. C. Hofzumahaus, P. Hebbeker and S. Schneider, *Soft Matter*, 2018, **14**, 4087-4100.
16. V. S. Rathee, A. J. Zervoudakis, H. Sidky, B. J. Sikora and J. K. Whitmer, *The Journal of Chemical Physics*, 2018, **148**, 114901.
17. F. Carnal and S. Stoll, *The Journal of Physical Chemistry B*, 2011, **115**, 12007-12018.
18. A. Laguecir, S. Ulrich, J. Labille, N. Fatin-Rouge, S. Stoll and J. Buffle, *European Polymer Journal*, 2006, **42**, 1135-1144.
19. F. Uhlík, P. Košovan, Z. Limpouchová, K. Procházka, O. V. Borisov and F. A. M. Leermakers, *Macromolecules*, 2014, **47**, 4004-4016.
20. A. Sadeghpour, A. Vaccaro, S. Rentsch and M. Borkovec, *Polymer*, 2009, **50**, 3950-3954.
21. K. N. Witte, S. Kim and Y.-Y. Won, *The Journal of Physical Chemistry B*, 2009, **113**, 11076-11084.
22. J. Lützenkirchen, J. van Male, F. Leermakers and S. Sjöberg, *Journal of Chemical & Engineering Data*, 2011, **56**, 1602-1612.
23. M. Borkovec, G. J. M. Koper and C. Piguet, *Current Opinion in Colloid & Interface Science*, 2006, **11**, 280-289.
24. M. Borkovec and G. J. M. Koper, *The Journal of Physical Chemistry*, 1994, **98**, 6038-6045.
25. R. G. Smits, G. J. M. Koper and M. Mandel, *The Journal of Physical Chemistry*, 1993, **97**, 5745-5751.
26. X. Xu, D. Cao, X. Zhang and W. Wang, *Physical Review E*, 2009, **79**, 021805.
27. D. Chandler, *Mechanics. Oxford University Press, Oxford, UK*, 1987, **40**.
28. M. Borkovec and G. J. M. Koper, *Macromolecules*, 1997, **30**, 2151-2158.
29. J.-P. Simonin, L. Blum and P. Turq, *The Journal of Physical Chemistry*, 1996, **100**, 7704-7709.
30. R. Gómez-Bombarelli, M. González-Pérez, M. T. Pérez-Prior, E. Calle and J. Casado, *The Journal of Physical Chemistry A*, 2009, **113**, 11423-11428.
31. A. K. Mishra and J. C. Ahluwalia, *The Journal of Physical Chemistry*, 1984, **88**, 86-92.
32. A. Gallegos and J. Wu, *Journal of Chemical & Engineering Data*, 2020, **65**, 5630-5642.
33. Y.-X. Yu, J. Wu and G.-H. Gao, *The Journal of Chemical Physics*, 2004, **120**, 7223-7233.
34. L. Blum, *Molecular Physics*, 1975, **30**, 1529-1535.
35. B. Maribo-Mogensen, G. M. Kontogeorgis and K. Thomsen, *Industrial & Engineering Chemistry Research*, 2012, **51**, 5353-5363.
36. S. Kawaguchi, T. Kitano and K. Ito, *Macromolecules*, 1992, **25**, 1294-1299.
37. J. Jiang, H. Liu, Y. Hu and J. M. Prausnitz, *The Journal of Chemical Physics*, 1998, **108**, 780-784.
38. Q. Xu, K. Wu, J. Mi and C. Zhong, *The Journal of Chemical Physics*, 2008, **128**, 214508.
39. D. R. Lide, *CRC handbook of chemistry and physics*, CRC press, 2004.
40. H. Dong, H. Du, S. R. Wickramasinghe and X. Qian, *The Journal of Physical Chemistry B*, 2009, **113**, 14094-14101.

41. H. Dong, H. Du and X. Qian, *The Journal of Physical Chemistry B*, 2009, **113**, 12857-12859.
42. J. Reijenga, A. van Hoof, A. van Loon and B. Teunissen, *Analytical Chemistry Insights*, 2013, **8**, ACI.S12304.
43. S. Kawaguchi, T. Kitano, K. Ito and A. Minakata, *Macromolecules*, 1990, **23**, 731-738.
44. J. de Groot, J. G. Hollander and J. de Bleijser, *Macromolecules*, 1997, **30**, 6884-6887.
45. J. de Groot, G. J. M. Koper, M. Borkovec and J. de Bleijser, *Macromolecules*, 1998, **31**, 4182-4188.
46. D. Henderson and W. R. Smith, *Journal of Statistical Physics*, 1978, **19**, 191-200.
47. J. T. Edward, P. G. Farrell and J. L. Job, *The Journal of Chemical Physics*, 1972, **57**, 5251-5256.
48. S. Kawaguchi, T. Kitano and K. Ito, *Macromolecules*, 1991, **24**, 6030-6036.
49. P. M. Blanco, S. Madurga, F. Mas and J. L. Garcés, *Polymers*, 2018, **10**, 811.
50. J. L. Garcés, S. Madurga, C. Rey-Castro and F. Mas, *Journal of Polymer Science Part B: Polymer Physics*, 2017, **55**, 275-284.
51. M. Ghasemi and R. G. Larson, *Progress in Polymer Science*, 2021, **112**, 101322.
52. J. Klein Wolterink, J. van Male, M. A. Cohen Stuart, L. K. Koopal, E. B. Zhulina and O. V. Borisov, *Macromolecules*, 2002, **35**, 9176-9190.
53. R. C. van Duijvenbode, M. Borkovec and G. J. M. Koper, *Polymer*, 1998, **39**, 2657-2664.