Check for updates

Polyelectrolyte pK_a from experiment and molecular dynamics simulation†

Michael S. Bodnarchuk,‡[a] Kay E. B. Doncom,‡[b] Daniel B. Wright,‡[b] David M. Heyes,[a] Daniele Dini*[a] and Rachel K. O'Reilly [iD] *[b]

The pK_a of a polyelectrolyte has been determined experimentally by potentiometric titration and computed using Molecular Dynamics (MD) constant pH (CpH) methodology, which allows the pK_a of each titratable site along the polymer backbone to be determined separately, a procedure which is not possible by current experimental techniques. By using experimental results within the CpHMD method, the simulations show that the protonation states of neighbouring residues are anti-correlated so that the charges are well-separated. As found with previous simulation studies on model polyelectrolytes, the end groups are predicted to be the most acidic. CpHMD is shown to result in distinct polymer conformations, brought about by the range of protonation states changes along the polymer; this can now be used in the design of pH-responsive polymers for, amongst other applications, additive formulation and drug delivery devices.

Introduction

Stimuli-responsive polymers undergo a change in hydrophilicity in response to a particular stimulus. Incorporation of these polymers into self-assembling block copolymer systems leads to stimuli-sensitive self-assembled structures which have the propensity to undergo a morphology transition in response to the particular stimulus. Stimuli which have been investigated within the literature include, but are not limited to, temperature,[1–5] pH[6–8] and light.[4,9] pH-responsive polymers have found uses as potential drug delivery systems,[7,10–13] nanoreactors,[14–18] recyclable catalysts[19–21] and in controlling surface properties.[22,23] Additionally, polyelectrolytes are widely used as water soluble thickeners and deliverers of specific functionality in added value products, where the acid–base equilibria can be exploited to determine performance.[24–27]

For monomers, the acid dissociation constant, or pK_a, is readily measured through titration experiments and is equal to the pH value at which equal amounts of the acid and conjugate base (or base and conjugate acid) are present. It has been reported that the pK_a of the monomers can change upon polymerisation. In one example the pK_a of styrenesulfonic acid was estimated, by solving the Poisson–Boltzmann equation, to increase from −0.53 for the monomer to 2.9 for polymers with

degrees of polymerisation (DP) of 8–12 units.[28] Often the overall pK_a of the polymer is measured by titration and this value is used as a reference point for where the overall shift in amphiphilic balance will occur in some systems. Therefore the hydrophilicity is often treated as an on/off function. However, for polymers the situation is more complicated. Unlike small molecules, where the protonation of individual species occurs independently, the protonation (and hence pK_a) of different sites on a polymer are directly affected by the immediate environment of that group within the molecule. Factors such as hydrophobicity of the environment,[29] distribution of ionisable units[30] and the ionisation state of the neighbouring units[31,32] can all affect the pK_a. Indeed, the pK_a of acidic groups increases with increasing degree of ionisation.[33] Therefore when altering the pH of a solution, in order to induce a hydrophilicity change, the polymer is not being protonated uniformly to the same extent, but rather different sites become ionised first. When considering applications, such effects may be significant. Therefore a more thorough understanding of the pK_a behaviour of units within the polymer structure is advantageous.

As with polymer chains, the pK_a values of the ionisable sites within a protein structure are sensitive to the immediate surrounding environment.[34] Investigating accurate pK_a values for different residues is important because factors such as folding, stabilisation, solubility interactions and reactivity are affected by pK_a values.[35–39] Using computational techniques it is possible to obtain the pK_a value for each individual ionisable residue within the molecule. The software Depth Server can be used to predict the pK_a, taking into account factors such as the depth of the site, accessible surface area, hydrogen bonds and electrostatic interactions.[40] This is useful when considering reactivity of a particular residue at various pHs.

[a]Department of Mechanical Engineering, Imperial College, South Kensington Campus, London SW7 2AZ, UK. E-mail: d.dini@imperial.ac.uk

[b]Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK. E-mail: rachel.oreilly@warwick.ac.uk

‡ These authors contributed equally to this work.

Determining the pK_as of individual units along a polymer backbone can be experimentally challenging. Current experimental potentiometric titration techniques used to determine the pK_a values of different titratable sites on the same molecule are most accurate when the pK_a values are widely separated, and if this is not the case an average value is the best that can usually be achieved. This limits the ability to introduce desired functionality at an early stage of any application of molecule development.

Theoretical techniques can usefully be applied to help elucidate these issues. Quantum mechanical (QM) approaches can, in principle, be used to calculate the pK_a of each titratable group in small molecules,[41–45] and oligomers,[46] but they are too costly to apply routinely to large molecules such as polyelectrolytes. Also the large computational demands of QM necessitate that it is used to calculate the pK_a of one group at a time. This procedure does not include the effects of simultaneous protonation and deprotonation across the polyelectrolyte, and therefore limits the reliability of any such calculation.

In contrast to separated monomers in solution, the intramolecular electrostatic interactions within the polyelectrolyte molecule can restrict further ionisation unless the distance between two sites are greater than the Bjerrum length.[47] When the polyelectrolyte experiences restricted ionisation, its conformation can be biased towards more elongated structures. Metropolis Monte Carlo (MC) computer simulation methods were developed in the 1990s to mimic the titration procedure using a coarse-grained representation of the polyelectrolyte molecule to determine the backbone site specific pK_a values.[48–50] Mean-field calculations were also carried out by Castelnovo *et al.* to predict the site-resolved pK_a distribution along the backbone.[51] The advantage of the MC method is that it performs a computationally-efficient virtual titration in which the protonated state depends on the local electrostatic environment within the molecule. However, the stochastic nature of the ionisation event and polymer chain fluctuations can limit the physical accuracy of the MC procedure unless rare-event MC moves are included. Historically, MC-based approaches could be limited in regard to the adequate sampling of new conformations arising from changing the protonation state of the polyelectrolyte. Methods such as coarse-graining[52] and the use of the Wang–Landau reaction ensemble[53] show great promise in circumventing such sampling issues, and their application to polymer pK_a determination has yet to be fully exploited.

In this study the more recently developed classical molecular simulation method of CpHMD is used as it overcomes many of the limitations of standard MC. The protonation state changes continuously during the simulation, and the coupled time-dependent structural and conformational fluctuations are better accounted for in the process of determining the local pK_a values. There are some applications in the literature of the application of this approach to polyelectrolytes, for example RNA.[54]

Constant-pH Molecular Dynamics simulations (CpHMD) provide a cost effective way of calculating the pK_a of ionisable groups.[55–60] The sampling of the protonation state of ionisable residues within a CpHMD simulation can be achieved by two methods. The first, based upon λ-dynamics,[61] treats the protonation state (λ_i) as a continuous variable, propagated by the forces acting within the simulation.[59] λ_i can be any value between 0 (protonated) and 1 (deprotonated), with the Hamiltonians describing the energy of the system scaled appropriately for these states. One caveat of this approach is the presence of unphysical states during the simulation (*i.e.* $\lambda_{0.5}$), although biasing schemes, based upon the free energy of deprotonation of a reference compound, can help to push the simulation towards the more realistic states. Treating the protonation state of ionisable residues in a continuous manner allows for the force and energy potential to evolve naturally, although the implementation of such approaches can be challenging.

The second sampling approach treats the protonation state of the ionisable group as a discrete variable, *i.e.* protonated or deprotonated only.[58] Periodically during the CpHMD simulation, an ionisable group is chosen randomly and an attempt is made to change its protonation state by a Monte Carlo procedure based on the pH and difference in Gibbs free energy between the molecule of interest and a reference compound (typically the monomer of the ionisable unit). Whilst such an approach can sometimes lead to a discontinuous energy potential for a large system, it is ideally suited to modelling small systems due to the speed of the method.

Regardless of sampling approach, the CpHMD method allows the pK_a of each group to be determined separately, by fitting the protonation probability of a given site as a function of pH to the Henderson–Hasselbalch equation.[62] The ability to vary simultaneously the protonation state of any one of the ionisable groups during the simulation allows for synergic effects between neighbouring groups to be included explicitly in the molecular simulation. This more balanced procedure offers a significant improvement in computational efficiency over the QM alternative.

Poly(glycerol methacrylate) (pGMA) is a water soluble polymer which has a wide range of applications, from protective coatings,[57] to insulin delivery agents.[58] The acid–base equilibria of pGMA and its related hydrophilicity are therefore of current interest. In this study the pK_a of each titratable group on a reversible addition fragmentation chain transfer RAFT synthesised pGMA, is calculated using CpHMD. These data are compared with the experimental value determined by potentiometric titration.

Experimental

Materials

AIBN (2,2′-azo-bis(isobutyronitrile)) was recrystallised from methanol and stored in the dark at 4 °C. All other materials were used as received from Aldrich, Fluka, and Acros. All aqueous solutions were prepared using nanopure water (deionised water, resistance > 18 MΩ). HCl (1 M) and NaOH (1 M) were calibrated and standardised using tris(hydroxymethyl) aminomethane and potassium hydrogen phthalate, respectively.

Instrumentation

^{1}H NMR spectra were recorded on a Bruker DPX-400 spectrometer in CDCl$_3$. Chemical shifts are given in ppm downfield from TMS.

Size exclusion chromatography (SEC) measurements were performed with HPLC grade solvents (Fisher), N,N-dimethylformamide (DMF) containing LiCl (1.06 gL^{-1}) as an eluent at 40 °C and a flow rate of 1 mL min^{-1}. The molecular weights of the synthesised polymers were calculated relative to poly(methyl methacrylate) (PMMA) standards.

Potentiometric titration[30,33,63] was performed at room temperature with an automatic titrator (Mettler Toledo G20) controlled by LabX software. 40 mL of solution (approximately 1 gL^{-1}) was used for each potentiometric titration experiment. The polymers were first dissolved at $\alpha = 1$ with 1.1 excess of HCl 1 M and then back-titrated with 0.1 M NaOH. The addition of NaOH 0.1 M titrant was added at volume increments of 5–50 μl and spaced with 180 s intervals. From the raw titration data the total number of titratable units was calculated.

Synthetic procedures

Synthesis of GMA monomer.[64] Glycidyl methacrylate (2.48 g) in water (22.4 g) was added to a round bottomed flask, 10% w/w solution. The emulsion was stirred for 8 h at 80 °C. ^{1}H NMR showed >99% conversion and no further purification was used. ^{1}H NMR (400 MHz, CDCl$_3$, 298 K): δ 1.97 (s, 3H, –CH$_3$), 3.55–3.65 (m, 2H, –CH$_2$), 3.70–3.75 (m, 2H, –CH$_2$), 3.85–3.95 (m, 1H, –CH), 4.10–4.25 (m, 2H, –CH$_2$), 4.90–5.0 (m, 1H, –CH), 5.60–5.70 (m, 1H, =CH$_2$), 6.10–6.20 (m, 1H, =CH$_2$).

Synthesis of pGM.[65] Glycidyl methacrylate (1.60 g), AIBN (7.4 mg) and 2-cyano-2-propyl 4-cyanobenzodithioate (CPDB) (50 mg) in 1,4-dioxane (1 : 4 volume compared to monomer) were added to a dry ampoule containing a stirrer bar. The solution was degassed using at least 3 freeze–pump–thaw cycles, back filled with nitrogen, sealed and placed in a pre-heated oil bath at 70 °C. After 4 hours the polymerisation was quenched by liquid nitrogen. The viscous crude reaction medium was dissolved in the minimum amount of tetrahydrofuran (THF), and the polymer was precipitated into diethyl ether and filtered. The precipitation process was repeated three times to afford a bright pink polymer. $M_{n\,NMR}$ = 5.0 kDa $M_{n\,SEC}$ = 5.8 kDa, $Đ_M$ = 1.10. ^{1}H NMR (400 MHz, CDCl$_3$, 298 K): δ 0.85–1.30 (br, 3H, CH$_2$–C–CH$_3$), 1.90–2.20 (br, 2H, –CH$_2$–C–CH$_3$), 1.94 (s, 6H, end group), 2.80–2.90 (br, 2H, –CH$_2$–CH–O–CH$_2$), 3.20–3.30 (br, 2H, –CH$_2$–CH–O–CH$_2$), 3.95–4.20 (br, 1H, –CH$_2$–CH–O–CH$_2$), 7.41 (t, J = 8.1 Hz, 2H, end group), 7.55 (t, J = 7.4 Hz, 2H end group), 7.85 (d, J = 8.2 Hz, 1H end group).

Hydrolysis of pGM to pGMA.[65] Poly(glycidyl methacrylate) (1 g) added to a 10% w/w solution of H$_2$SO$_4$ and stirred for 48 hours at room temperature. After 48 hours the solution was neutralised with sodium bicarbonate. The solution was transferred to a dialysis membrane tube with the appropriate molecular weight cut-off (MWCO) and dialysed against nanopure water (1.5 L) incorporating 5 water changes. Lyophilisation resulted in a pink polymer. $M_{n\,NMR}$ = 5.0 kDa, $M_{n\,SEC}$ = 5.8 kDa, $Đ_M$ = 1.08. ^{1}H NMR (400 MHz, CDCl$_3$, 298 K): δ 0.85–1.30 (br, 3H, CH$_2$–C–CH$_3$), 1.90–2.20 (br, 2H, –CH$_2$–C–CH$_3$), 1.94 (s, 6H, end group), 3.6–3.75 (br, 2H, –CH$_2$–CH(OH)–CH$_2$–OH), 3.95–4.20 (br, 3H, –CH$_2$–CH(OH)–CH$_2$–OH) 7.85 (d, J = 8.2 Hz, 1H, end group), 7.55 (t, J = 7.4 Hz, 2H end group), 7.41 (t, J = 8.1 Hz, 2H, end group).

Simulation methodology. In order to generate titration profiles of pGMA, CpHMD simulations were performed using the AMBER package. Due to its speed and the small size of polymer, the discrete sampling CpHMD approach was utilised.[58] The protonation state of an ionisable unit during the MD simulation was periodically changed. The acceptance criterion of the new protonation state was determined by a Metropolis Monte Carlo step using the free energy formula given in eqn (1).[66] This is based on the pH of the system, the pK_a of the monomeric unit (p$K_{a,mon}$) and the difference in the Gibbs deprotonation free energies of the system of interest ($\Delta G_{deprot,sys}$) and the monomeric unit in isolation ($\Delta G_{deprot,mon}$). The statistics of the proportion of deprotonated species are recorded during the simulation as a function of pH, which are then subsequently fitted using the Henderson–Hasselbalch function.

$$\Delta G = k_B T(\mathrm{pH} - \mathrm{p}K_{a,mon})\ln 10 + \Delta G_{deprot,sys} - \Delta G_{deprot,mon} \quad (1)$$

Monomer deprotonation free energies for common ionisable amino acids come pre-calculated in AMBER; for novel ionisable units these need to be calculated manually. Consequently the first stage of the simulation was to calculate the deprotonation free energy, $\Delta G_{deprot,mon}$, of the GMA monomer using classical Thermodynamic Integration (TI). TI is a method for calculating the free-energy difference between two end states, termed λ_0 and λ_1. In these simulations, a value of λ_0 indicates that the monomer is protonated, while λ_1 indicates that the monomer is deprotonated. Intermediate values of λ represent hybrids between the two end states and are used to improve the accuracy of the TI simulations. Eleven equally spaced λ-windows between 0 and 1 [0–0.09, 0.09–0.18 …1] were used to calculate the free energy, with each window simulated for 5 ns. Data was collected for the last 4 ns of the simulation. The value of $\Delta U/\Delta\lambda$ was calculated for each window (where U is the potential energy of the system) and integrated with respect to λ to find the overall free energy change between λ_0 and λ_1. The deprotonation free energies for the two alcohol sites on the monomer at different salt concentrations are shown in Table 1.

Implicit solvent constant-pH simulations were performed with AMBER 14 (ref. 67) using the same conditions as the TI simulations. The pGMA molecule was made by combining 30 titratable monomer units in the leap module of AMBER and capping it with AIBN and benzodithioate at the termini, with the small excess charge evenly distributed on one of the backbone atoms of each polymeric unit to ensure neutrality. The polymer was initially minimised in Sander, and then equilibrated for 100 ps prior to the simulation. The simulations were set up to ensure that each residue had its titration state

Table 1 Deprotonation free energies for the two alcohol sites at two salt concentrations. The energies are in units of kcal mol^{-1}

Alcohol site	0 M	0.5 M
Primary	−64.1	−44.5
Secondary	−64.4	−44.4

sampled, on average, every 100 fs. Simulations were performed at every pH integer between 3 and 10, with more performed around the equivalence point of the titration. A time-step of 2 fs was used. Each constant-pH simulation was carried out for 10 ns, with the probability of one of the alcohol sites being deprotonated monitored throughout the simulation. At least four repeats were performed for each pH value to calculate the statistical uncertainty in the value of the average deprotonation probability across the polymer, α. The value of α as a function of pH was used to fit the data using the Henderson–Hasselbalch equation, shown in eqn (2), to find the pK_a.

$$\alpha(pH) = \frac{1}{10^{n(pK_a - pH)} + 1} \qquad (2)$$

In eqn (2), n is the Hill coefficient which indicates whether the deprotonation of two sites are correlated ($n > 1$) or anti-correlated ($n < 1$). Both n and pK_a are fitted parameters in eqn (2).

The GMA monomer, AIBN and benzodithioate groups were parameterised using the antechamber module in AMBER,[67] with the partial charges assigned using the AM1-BCC model.[68] Prior to simulation, minimisation was performed using the Sander module of AMBER to remove bad contacts, using an igb keyword of 1. The GAFF force-field was used to model the monomer and pGMA,[69] with the Generalised-Born model of Onufriev[70] (igb = 2) used to describe the implicit GB solvent model. A non-bonded cut-off of 3.0 nm and salt concentrations (NaCl) of 0 M and 0.5 M were used. The Berendsen thermostat was used with a time constant of 2.0 ps to maintain the temperature at 300 K. Bond distances were maintained constant using the SHAKE algorithm.[71]

Results and discussion

The monomer glycerol methacrylate (GMA) was chosen for study because, as previously mentioned, the polymer is water soluble and has a wide range of potential applications. Reversible addition-fragmentation chain transfer (RAFT) polymerisation techniques were used because this monomer has been shown previously to be polymerised with good control by RAFT.[64] The precursor monomer, glycidyl methacrylate (GM), was first polymerised in 1,4-dioxane using a dithiobenzoate chain transfer agent (Fig. 1).[65] Hydrolysis of the resulting polymer with a 10% w/w solution of H_2SO_4, followed by exhaustive dialysis, yielded the pGMA polymer, $M_{n\ NMR} = 5.0$ kDa, $M_{n\ SEC} = 5.8$ kDa, $Đ_M = 1.08$ (see ESI†). ^1H NMR comparison of the end group signals of the CTA with the side chain signals of the GMA group gave a degree of polymerisation (DP) of 30. The controlled character of the polymerisation was demonstrated by the narrow dispersity of the SEC trace and the good agreement between the DP obtained from ^1H NMR analysis of the purified polymer and that predicted from conversion NMR.

The pK_a of the GMA monomer was determined experimentally to be 6.3 (see ESI†). As shown in eqn (1), one of the prerequisites to perform a CpHMD simulation is the monomer pK_a ($pK_{a,mon}$). Despite advances in pK_a calculations of small molecules, the best accuracy for pK_a estimation is typically ± 0.5 pK_a units. This error can be larger for molecules which can have multiple titration sites or have unusual chemical functionalities. Experimental values of the $pK_{a,mon}$ were used in the CpHMD simulations to give a more accurate titration profile for the polymer due to the presence of the diol. This value was used as an input in the CpHMD simulations. The pK_a of the pGMA homopolymer was determined by potentiometric titration in salt-free water and in a 0.5 M NaCl solution.

CpHMD simulations performed on the pGMA shown in Fig. 1 used the AMBER simulation package, with implicit solvent (see experimental for detailed simulation procedure). Simulations were performed at unit intervals in pH between 3 and 10 for salt concentrations of 0 and 0.5 M. The average value of the degree of ionisation, α, recorded during the simulation as a function of pH was fitted to the Henderson–Hasselbalch (HH) equation to determine the Hill coefficient and the pK_a of the polymer (i.e. where $\alpha = 0.5$).

The average degree of deprotonation of pGMA, as a function of pH at 0 M salt concentration determined by both experiment and simulation are plotted in Fig. 2. The predicted pK_a from the CpHMD simulation is 7.4 ± 0.25, which compares favourably with the experimental titration value of 7.5 ± 0.2. The calculated Hill coefficient from the simulation, 0.67 also agrees well with the experimental titration value of 0.72. It is significant to note that the simulated curve begins to deviate from the experimental titration curve at low pH values. Experimentally, at low pH the ionic strength associated with $[H]^+$ is increased, which contributes to an electrostatic screening effect. Landsgesell and co-workers have shown that one of the caveats of the CpHMD method is that it does not capture this pH-induced screening, which leads to a subtle deviation from the experimental curve.[72]

In the absence of salt the pK_a of the pGMA is consistently about 1 pK_a unit higher than its monomer. We attribute this to the increasing difficulty of ionising neighbouring repeat units along the pGMA chain. As the nth group along the pGMA chain becomes negatively charged it requires more energy for the $n + 1$th (neighbouring) group to also be ionised because of the electrostatic repulsion between the groups. The exception to this phenomenon, however, is the terminus of the polymer. As a result, the average pK_a of the polymerised chain is higher than that of the monomer.

Evidence for electrostatic repulsion along the chain can be seen in Fig. 3, which shows the state of deprotonation of each titratable group along the molecule (as a colour coded map) captured at an instant in the simulation when $pH = pK_a$, i.e. at the equivalence point of the polymer. Fig. 3 shows that a protonated site is typically found adjacent to two
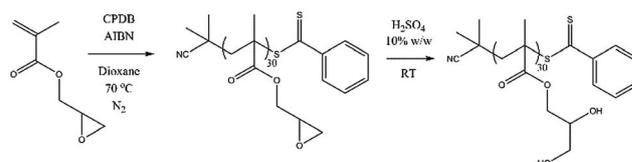

Fig. 1 The synthetic route to the pGMA polyelectrolyte studied in this work.
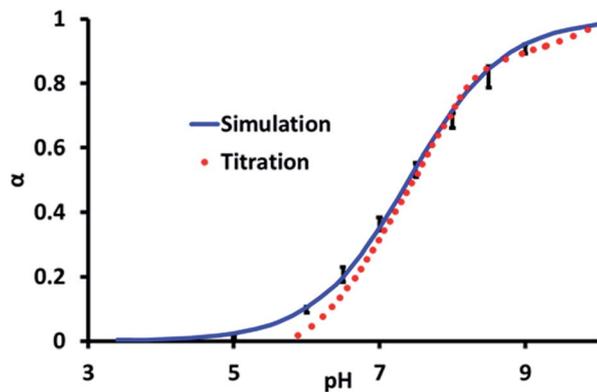
Fig. 2 The average degree of ionisation, α, for all the titratable sites (1 = deprotonated, 0 = protonated) of pGMA as a function of pH at 0 M NaCl concentration from experiment and simulation. The error bars on the simulation data are the standard error in α taken across all of the ionisable groups in the molecule.
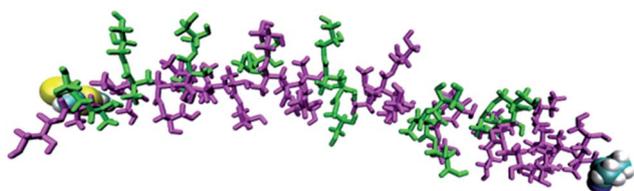

Fig. 3 Simulated protonation state (magenta = deprotonated, green = protonated) for each site in pGMA in a 0 M salt solution at pH = 7.4. The terminal groups are shown as van der Waals spheres.

deprotonated sites, highlighting the synergy between adjacent units to reduce electrostatic repulsion. Indeed, no more than three adjacent deprotonated sites are ever observed in the simulation at pH = pK_a. As the pH of the simulation deviates further from the polymer pK_a, more protonated or deprotonated sites in proximity to each other will occur. A key advantage of the CpHMD method is that the level of detail shown in Fig. 3 cannot be derived currently from experiment. These results can therefore be used to give insights into experimental observations and trends.

One such observation is the formation of micro-domains within polyelectrolytes. Micro-domains are partly folded regions of the polymer conformation which display increased order compared to a fully-extended conformation. Their occurrence is of interest in the fields of drug-delivery and biotechnology, as they can be exploited to trap small molecules.[73–75] The small molecules can be enclosed within the polymer at a certain pH value, and then released within the body under physiological pH.

It is thought that typical micro-domains result from hydrogen-bonds between triplets of titratable sites.[76] In order to investigate whether pGMA is capable of forming micro-domains, the MD trajectories at pH = pK_a were examined. Despite being a small polymer, the MD simulations show conformational features which are suggestive of incipient micro-domain formation. Bent or 'hairpin' conformations appear frequently in the simulation, and a MD representation of them can be found in Fig. 4. Radial distribution functions

(RDFs) between atoms were calculated which provide another perspective on this conformational trend. The RDF between two atoms gives the probability of finding them at a set distance from each other. The hairpin conformation presented in Fig. 4 could be a consequence of favourable electrostatic attractions between triplets of ionisable sites along the polymer chain, highlighting the benefits of using CpHMD to identify protonation-conformation relationships. This should show up in the RDF between the primary O-atom of one site (n) and the primary O-atom of the n − 3rd and n + 3rd site.

Fig. 5 shows the RDF between site 14, which is at the apex of the bending motion shown in Fig. 3, and near-adjacent sites 17 and 20 in the polymer chain. The RDFs indicate that the greatest probability of finding the O-atoms of sites 17 and 20 closest to the O-atom of site 14 is approximately 4.5 Å for the hairpin conformation, while for the extended conformation it is approximately 7 Å. Fig. 5 therefore indicates the formation of microdomain-like features, even within this relatively small pGMA 30-mer. This behaviour is reproduced for other hairpin conformations, which suggests that ionisable sites spaced 3 units apart are able to interact favourably to form small local-ised microdomains in pGMA.

In the simulations we also explored the effect of increasing the concentration of monovalent salt on the ionisation behaviour of the pGMA 30-mer. The predicted pK_a from the CpHMD simulation at a salt concentration of 0.5 M NaCl is 7.1 ± 0.25 which again agrees well with the experimental titration value of 7.1 ± 0.2 (cf. the pK_a of the pGMA at 0 M was 7.4 ± 0.25). The decrease in pK_a on increasing the salt concentration from 0 to 0.5 M is consistent with an enhanced screening of the electro-static repulsion between adjacent deprotonated units, which causes the sites to be more acidic by approximately 0.4 pK_a units. Classical Debye–Hückel (DH) theory,[77] can account at a mean-field level of approximation for changes from an ideal solution to an electrolyte. DH also predicts a pK_a shift of ca. 0.3 units which is in excellent agreement with the CpHMD shift noted above. Of course DH theory does not provide the level of atomic detail obtained by CpHMD, which is important for resolving other issues relating to this system.

A strength of simulations that include reactions, such as the CpHMD methods, is the ability to follow specific single sites along the polymer chain. The site specific pK_a is plotted in Fig. 6
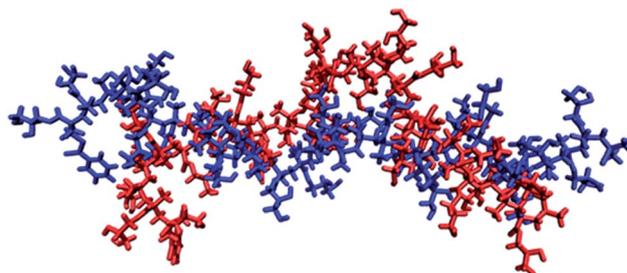

Fig. 4 Two snapshots of the pGMA generated by simulation when it is stretched out (blue) and has a bend in the middle (red), with a solution salt concentration of 0 M.
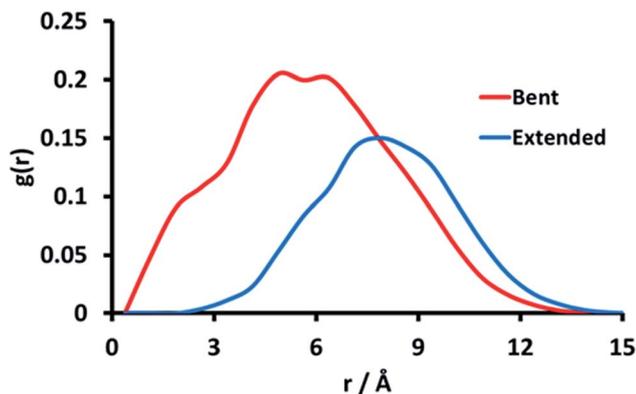
Fig. 5 Radial distribution function between the primary alcohol oxygen atom of site 14 and the primary alcohol oxygen of sites 17 and 20. The RDF for the bent conformation is shown in red, with the extended conformation shown in blue.

and shows that the ω and α ends of the polymer (numbered sites 1 and 30 here) are more acidic than those near the centre of the molecule. This can be attributed to a reduced electrostatic repulsion from neighbouring sites in that region; the termini of the polymer are neutral, and as such it is easier to ionise units near these regions. Similar behaviour has been noted previously by Castelnovo and co-workers.[51] Of note is that sites 6 to 26 have, within statistics, the same pK_a value at the two salt concentrations with these pK_a values slightly higher than the recorded average pK_a, particularly at 0 M.

Thus it can be concluded that the average, macroscopic, pK_a of a polymer recorded by experiment is not necessarily indicative of the microscopic pK_a of its constituent units. The consistent microscopic pK_a recorded across the polymer suggests that, at $pH = pK_a$, the synergy seen between triplets of ionisable units is likely to be reversible and dynamic. It is foreseeable that, as the pH is varied further from the pK_a, the synergy between triplets
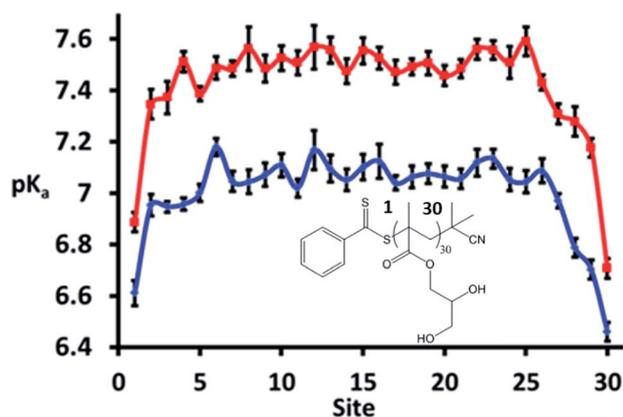
will become less reversible and more long-lived. Under these circumstances, CpHMD is ideally suited to identify these triplets; both through the MD snapshots and also the emergence of pronounced peaks and troughs in the microscopic pK_as.

Conclusions

This work has used both experimental and computational procedures to study the ionisation and titration behaviour of a pGMA. One of the input requirements for a CpHMD simulation is the pK_a of the ionisable monomer, which is usually known for common amino acids but less so for more unusual monomers. By using a monomer pK_a which has been measured experimentally we show that CpHMD can reproduce both the experimental pK_a and the titration behaviour of pGMA as a function of salt concentration, the latter being found to agree well with Debye–Hückel theory for this benchmark system.

One of the strengths of CpHMD lies in its ability to calculate routinely, with modest computational resources, the pK_a's of all the sites in the molecule in a single simulation in a way which incorporates the natural synergy between adjacent sites at various states of protonation. Guided by experimental input, the conformational states likely to be adopted by the real polymer at different pH values can be generated and linked to the protonation states of the units along the polymer backbone. Such detail cannot be determined by classical MD or experiment and, as briefly explored here, can be used to design pH-responsive polymers which can trap and release small molecules within micro-domains depending on the system acidity. Furthermore, this work lays the groundwork for using CpHMD to help assist in the design of new polymers[78] as a guide to experimental synthesis which may enable the preparation of responsive and functional nanomaterials for applications as diverse as electronics or healthcare.

Acknowledgements

Fig. 6 Site specific pK_a values for the model pGMA at 0 (red) and 0.5 M (blue) NaCl solution. The sites are ordered from 1 (site closest to ω end) to 30 (site closest to α end). Error bars indicate the standard error over 8 independent runs (note that 8 repeats were used to study the site specific values when the polymer was simulated at its equivalence point rather than 4 repeats, which was the minimum number of runs performed when studying average values across the entire polymer).

References

1 M. I. Gibson and R. K. O'Reilly, *Chem. Soc. Rev.*, 2013, **42**, 7204–7213.
2 R. Liu, M. Fraylich and B. Saunders, *Colloid Polym. Sci.*, 2009, **287**, 627–643.

3 D. Roy, W. L. A. Brooks and B. S. Sumerlin, *Chem. Soc. Rev.*, 2013, **42**, 7214–7243.

4 F. D. Jochum and P. Theato, *Chem. Soc. Rev.*, 2013, **42**, 7468–7483.

5 A. O. Moughton, J. P. Patterson and R. K. O'Reilly, *Chem. Commun.*, 2011, **47**, 355–357.

6 S. Dai, P. Ravi and K. C. Tam, *Soft Matter*, 2008, **4**, 435–449.

7 D. Schmaljohann, *Adv. Drug Delivery Rev.*, 2006, **58**, 1655–1670.

8 K. E. B. Doncom, C. F. Hansell, P. Theato and R. K. O'Reilly, *Polym. Chem.*, 2012, **3**, 3007–3015.

9 S. Dai, P. Ravi and K. C. Tam, *Soft Matter*, 2009, **5**, 2513–2533.

10 T. Yoshida, T. C. Lai, G. S. Kwon and K. Sako, *Expert Opin. Drug Delivery*, 2013, **10**, 1497–1513.

11 K. M. Huh, H. C. Kang, Y. J. Lee and Y. H. Bae, *Macromol. Res.*, 2012, **20**, 224–233.

12 S. Mura, J. Nicolas and P. Couvreur, *Nat. Mater.*, 2013, **12**, 991–1003.

13 O. Onaca, R. Enea, D. W. Hughes and W. Meier, *Macromol. Biosci.*, 2009, **9**, 129–139.

14 L. Zhu, Y. Shi, C. Tu, R. Wang, Y. Pang, F. Qiu, X. Zhu, D. Yan, L. He, C. Jin and B. Zhu, *Langmuir*, 2010, **26**, 8875–8881.

15 T. Einfalt, R. Goers, I. A. Dinu, A. Najer, M. Spulber, O. Onaca-Fischer and C. G. Palivan, *Nano Lett.*, 2015, **15**, 7596–7603.

16 D. Grafe, J. Gaitzsch, D. Appelhans and B. Voit, *Nanoscale*, 2014, **6**, 10752–10761.

17 H. A. Zayas, A. Lu, D. Valade, F. Amir, Z. Jia, R. K. O'Reilly and M. J. Monteiro, *ACS Macro Lett.*, 2013, **2**, 327–331.

18 A. Lu, D. Moatsou, D. A. Longbottom and R. K. O'Reilly, *Chem. Sci.*, 2013, **4**, 965–969.

19 Y. Xie, M. Wang, X. Wu, C. Chen, W. Ma, Q. Dong, M. Yuan and Z. Hou, *ChemPlusChem*, 2016, **81**, 541–549.

20 J. Zhang, M. Zhang, K. Tang, F. Verpoort and T. Sun, *Small*, 2014, **10**, 32–46.

21 A. Ievins, X. Wang, A. O. Moughton, J. Skey and R. K. O'Reilly, *Macromolecules*, 2008, **41**, 2998–3006.

22 Y. Kotsuchibashi, Y. Wang, Y.-J. Kim, M. Ebara, T. Aoyagi and R. Narain, *ACS Appl. Mater. Interfaces*, 2013, **5**, 10004–10010.

23 C. de las Heras Alarcón, B. Twaites, D. Cunliffe, J. R. Smith and C. Alexander, *Int. J. Pharm.*, 2005, **295**, 77–91.

24 N. P. Gerard, R. L. Eddy, T. B. Shows and C. Gerard, *J. Biol. Chem.*, 1990, **265**, 20455–20462.

25 J.-M. Y. Carrillo and A. V. Dobrynin, *Macromolecules*, 2011, **44**, 5798–5816.

26 Q. Liao, A. V. Dobrynin and M. Rubinstein, *Macromolecules*, 2003, **36**, 3386–3398.

27 D. J. Sandberg, J.-M. Y. Carrillo and A. V. Dobrynin, *Langmuir*, 2007, **23**, 12716–12728.

28 H. Dong, H. Du, S. R. Wickramasinghe and X. Qian, *J. Phys. Chem. B*, 2009, **113**, 14094–14101.

29 U. P. Strauss and M. S. Schlesinger, *J. Phys. Chem.*, 1978, **82**, 571–574.

30 I. Borukhov, D. Andelman, R. Borrega, M. Cloitre, L. Leibler and H. Orland, *J. Phys. Chem. B*, 2000, **104**, 11027–11034.

31 J. D. Ziebarth and Y. Wang, *Biomacromolecules*, 2010, **11**, 29–38.

32 H. Lee, S. H. Son, R. Sharma and Y.-Y. Won, *J. Phys. Chem. B*, 2011, **115**, 844–860.

33 O. Colombani, E. Lejeune, C. Charbonneau, C. Chassenieux and T. Nicolai, *J. Phys. Chem. B*, 2012, **116**, 7560–7565.

34 E. L. Mehler, M. Fuxreiter, I. Simon and B. E. Garcia-Moreno, *Proteins: Struct., Funct., Bioinf.*, 2002, **48**, 283–292.

35 F. B. Sheinerman, R. Norel and B. Honig, *Curr. Opin. Struct. Biol.*, 2000, **10**, 153–159.

36 M. Schaefer, M. Sommer and M. Karplus, *J. Phys. Chem. B*, 1997, **101**, 1663–1683.

37 J. Kim, J. Mao and M. R. Gunner, *J. Mol. Biol.*, 2005, **348**, 1283–1298.

38 A. H. Elcock and J. A. McCammon, *J. Mol. Biol.*, 1998, **280**, 731–748.

39 Z. S. Hendsch, T. Jonsson, R. T. Sauer and B. Tidor, *Biochemistry*, 1996, **35**, 7621–7625.

40 K. P. Tan, T. B. Nguyen, S. Patel, R. Varadarajan and M. S. Madhusudhan, *Nucleic Acids Res.*, 2013, **41**, W314–W321.

41 N. Uddin, T. H. Choi and C. H. Choi, *J. Phys. Chem. B*, 2013, **117**, 6269–6275.

42 W. L. Jorgensen, J. M. Briggs and J. Gao, *J. Am. Chem. Soc.*, 1987, **109**, 6857–6858.

43 W. L. Jorgensen and J. M. Briggs, *J. Am. Chem. Soc.*, 1989, **111**, 4190–4197.

44 C. Lim, D. Bashford and M. Karplus, *J. Phys. Chem.*, 1991, **95**, 5610–5620.

45 J. Ho and M. L. Coote, *Theor. Chem. Acc.*, 2009, **125**, 3–21.

46 H. Dong, H. Du and X. Qian, *J. Phys. Chem. B*, 2009, **113**, 12857–12859.

47 J. Landsgesell, C. Holm and J. Smiatek, *J. Chem. Theory Comput.*, 2017, **13**, 852–862.

48 M. Ullner and B. Jönsson, *Macromolecules*, 1996, **29**, 6645–6655.

49 M. Ullner, B. Jönsson, B. Söderberg and C. Peterson, *J. Chem. Phys.*, 1996, **104**, 3048–3057.

50 M. Ullner and C. E. Woodward, *Macromolecules*, 2000, **33**, 7144–7156.

51 M. Castelnovo, P. Sens and J.-F. Joanny, *Eur. Phys. J. E: Soft Matter Biol. Phys.*, 2000, **1**, 115–125.

52 J. Baschnagel, J. P. Wittmer and H. Meyer, arXiv:cond-mat/0407717 [cond-mat.soft], 2014.

53 J. Landsgesell, C. Holm and J. Smiatek, *J. Chem. Theory Comput.*, 2017, **13**, 852–862.

54 G. B. Goh, J. L. Knight and C. L. Brooks, *J. Chem. Theory Comput.*, 2013, **9**, 935–943.

55 R. Bürgi, P. A. Kollman and W. F. van Gunsteren, *Proteins: Struct., Funct., Bioinf.*, 2002, **47**, 469–480.

56 A. Aleksandrov, S. Polydorides, G. Archontis and T. Simonson, *J. Phys. Chem. B*, 2010, **114**, 10634–10648.

57 J. Mongan, D. A. Case and J. A. McCammon, *J. Comput. Chem.*, 2004, **25**, 2038–2048.

58 J. Mongan and D. A. Case, *Curr. Opin. Struct. Biol.*, 2005, **15**, 157–163.

59 M. S. Lee, F. R. Salsbury and C. L. Brooks, *Proteins: Struct., Funct., Bioinf.*, 2004, **56**, 738–752.

60 M. S. Bodnarchuk, D. M. Heyes, D. Dini, S. Chahine and S. Edwards, *J. Chem. Theory Comput.*, 2014, **10**, 2537–2545.

61 J. L. Knight and C. L. Brooks, *J. Comput. Chem.*, 2009, **30**, 1692–1700.

62 S. Donnini, F. Tegeler, G. Groenhof and H. Grubmüller, *J. Chem. Theory Comput.*, 2011, **7**, 1962–1978.

63 H. Lee, S. H. Son, R. Sharma and Y. Y. Won, *J. Phys. Chem. B*, 2011, **115**, 844–860.

64 L. P. D. Ratcliffe, A. J. Ryan and S. P. Armes, *Macromolecules*, 2013, **46**, 769–777.

65 C. S. Gudipati, M. B. H. Tan, H. Hussain, Y. Liu, C. He and T. P. Davis, *Macromol. Rapid Commun.*, 2008, **29**, 1902–1907.

66 N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller and E. Teller, *J. Chem. Phys.*, 1953, **21**, 1087–1092.

67 D. A. Case, V. Babin, J. T. Berryman, R. M. Betz, Q. Cai, D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, H. Gohlke, A. W. Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, A. Kovalenko, T. S. Lee, S. LeGrand, T. Luchko, R. Luo, B. Madej, K. M. Merz, F. Paesani, D. R. Roe, A. Roitberg, C. Sagui, R. Salomon-Ferrer, G. Seabra, C. L. Simmerling, W. Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu and P. A. Kollman, *AMBER2014*, University of California, San Francisco, 2014.

68 A. Jakalian, D. B. Jack and C. I. Bayly, *J. Comput. Chem.*, 2002, **23**, 1623–1641.

69 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.

70 A. Onufriev, D. A. Case and D. Bashford, *J. Comput. Chem.*, 2002, **23**, 1297–1304.

71 J.-P. Ryckaert, G. Ciccotti and H. J. C. Berendsen, *J. Comput. Phys.*, 1977, **23**, 327–341.

72 J. Landsgesell, J. Smiatek and C. Holm, 2017, arXiv:1702.04911 [cond-mat.soft].

73 M. Qi, S. Duan, B. Yu, H. Yao, W. Tian and F.-J. Xu, *Polym. Chem.*, 2016, **7**, 4334–4341.

74 G. Fundueanu, M. Constantin, C. Stanciu, G. Theodoridis and P. Ascenzi, *J. Mater. Sci.: Mater. Med.*, 2009, **20**, 2465.

75 S. Yamaguchi, R. Kawabata, E. Yamamoto and T. Nagamune, *Polym. J.*, 2014, **46**, 880–886.

76 J. L. Garcés, G. J. M. Koper and M. Borkovec, *J. Phys. Chem. B*, 2006, **110**, 10937–10950.

77 R. A. Cox, *Angew. Chem., Int. Ed.*, 2013, **52**, 7638.

78 P. Cotanda, D. B. Wright, M. Tyler and R. K. O'Reilly, *J. Polym. Sci., Part A: Polym. Chem.*, 2013, **51**, 3333–3338.