

Electronic supplemental material for the manuscript

Local pH and effective pK: Two names for one quantity?

Anastasiia Murmiliuk¹, Peter Košovan^{1,*}, Miroslav Janata², Karel Procházka¹, Filip Uhlík¹, and Miroslav Štěpánek^{*,†}

1. Department of Physical and Macromolecular Chemistry,
Charles University, Hlavova 8, 128 43 Prague 2, Czech Republic and
2. Institute of Macromolecular Chemistry, Czech Academy of Sciences,
Heyrovský square 2, 162 06 Prague 6, Czech Republic

(Dated: September 20, 2018)

CONTENTS

I. Synthesis and characterization of umbelliferone-labeled PMAA	1
A. Materials	1
B. Synthesis of the 4-chloromethyl umbelliferone (4-chloromethyl-7-hydroxycoumarin) initiator	2
C. ATRP polymerization of <i>tert</i> -butyl methacrylate initiated with 4-chloromethylumbelliferone	2
D. Transformation of umbelliferone-labeled poly(<i>tert</i> -butyl methacrylate) to umbelliferone-labeled poly(methacrylic acid)	3
II. Solution properties and ionization of PMAA-Umb	4
A. Materials	4
B. Electrophoretic Light Scattering (ELS)	4
C. Dynamic Light Scattering (DLS)	4
D. Small-angle X-ray Scattering (SAXS)	5
E. Fluorometry	6
III. Derivation of the relation between "local pH" and effective pK_A	7
A. Relation between Δ and μ^{ex}	7
B. Determination of μ^{ex} from $g(r)$	8
C. Determination of $\langle \psi(r=0) \rangle$ using a test charge	9
IV. Simulation model and method	9
V. Additional simulation results	10
A. Alternative representation of the pK_A shifts	10
B. Sensitivity of the simulated pK_A shift to model parameters	13
References	15

I. SYNTHESIS AND CHARACTERIZATION OF UMBELLIFERONE-LABELED PMAA

A. Materials

1,3-Dihydroxybenzene (resorcinol; 99%, Sigma-Aldrich), ethyl 4-chloroacetoacetate (95%, Sigma-Aldrich), sulphuric acid (95%, Lach-Ner, Czech Republic), copper (I) chloride (97%, Sigma-Aldrich), copper (II) chloride (97%, Sigma-Aldrich), 1,1,4,7,10,10-hexamethyltriethylenetetramine (HMTETA; 97%, Sigma-Aldrich), tetrahydrofuran (THF; Lach-Ner, Czech Republic), dichloromethane (Lach-Ner), trifluoroacetic acid (99.5%, Acros) were used as received. Toluene was distilled with lithium aluminum hydride. *Tert*-butyl methacrylate was distilled with calcium hydride and with triisobutyl aluminum before use.

* peter.kosovan@natur.cuni.cz

† stepanek@natur.cuni.cz

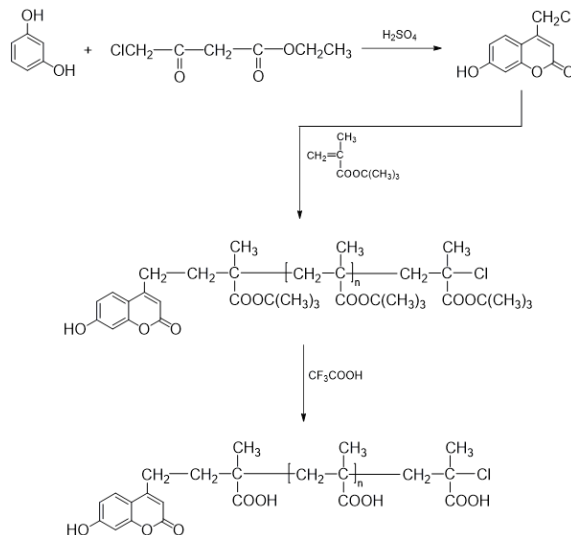


FIG. S1. Reaction scheme of the synthesis of umbelliferone-labeled PMAA.

B. Synthesis of the 4-chloromethyl umbelliferone (4-chloromethyl-7-hydroxycoumarin) initiator

The initiator was synthesized following the procedure described in Ref. [1]. Crushed 1,3-dihydroxybenzene (5.6 g, 50.86 mmol) was added portion-wise with stirring to sulphuric acid (43 mL) in an ice-water bath. Ethyl 4-chloroacetoacetate (5.7 mL, 42.5 mmol) was added drop-wise to the resulting suspension under cooling. The reaction mixture was intensively stirred for 2 h under cooling, subsequently for 24 h at room temperature, and then it was poured portion-wise to ice-cold water (400 mL). The resulting mixture with ochreous precipitate was stirred for ca 1 h at room temperature, filtered and the solid product was washed with ice-cold water to neutral reaction. The isolated product was then recrystallized from ethanol and dried at 40°C in vacuum. The yield was 6.7 g. The product was characterized by ^1H NMR (Fig. S2).

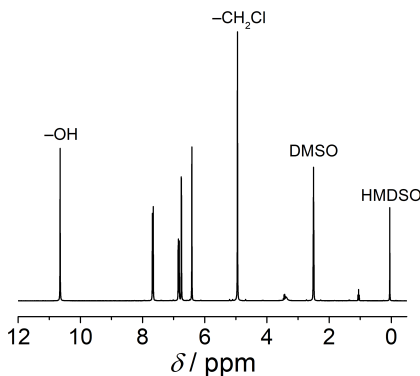


FIG. S2. ^1H NMR spectrum of 4-chloromethylumbelliferone in DMSO: 4.95 ppm (CH_2Cl), 6.41 – 7.68 ppm (protons at carbons C3, C5, C6, and C8), 10.65 ppm (OH).

C. ATRP polymerization of *tert*-butyl methacrylate initiated with 4-chloromethylumbelliferone

Copper(I) chloride (8.46 mg, 0.085 mmol), copper(II) chloride (5.75 mg, 0.043 mmol), and 4-chloromethylumbelliferone (3.6 mg, 0.017 mmol) were placed in the reaction flask. After several vacuum/argon cycles toluene (4.2 mL), *tert*-butyl methacrylate (4.2 mL, 25.6 mmol) and HMTETA (0.023 mL, 0.085 mmol) were added sequentially. The polymeriza-

tion was carried out at 90°C for 24 h. The polymerization mixture was then diluted with THF and precipitated in 70% methanol. The isolated polymer, umbelliferone-labeled poly(*tert*-butyl methacrylate), was dried at 40°C in vacuum. The yield was 1.75 g. SEC chromatogram for the polymer is shown in Fig. S3.

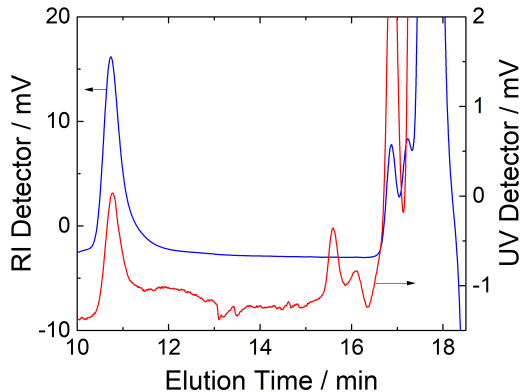


FIG. S3. SEC chromatogram of umbelliferone-labeled poly(*tert*-butyl methacrylate): $M_n = 179000$, $M_w/M_n = 1.14$

D. Transformation of umbelliferone-labeled poly(*tert*-butyl methacrylate) to umbelliferone-labeled poly(methacrylic acid)

The umbelliferone-labeled poly(*tert*-butyl methacrylate) (1.39 g) was placed in the reaction flask. Dichloromethane (40 mL) and trifluoroacetic acid (3.6 mL) were added after several vacuum/argon cycles. The reaction solution was stirred at room temperature for 24 h. The reaction mixture was then evaporated to dryness, the solid product was solubilized in absolute ethanol and the solution was precipitated in hexane. The isolated product, umbelliferone-labeled poly(methacrylic acid), was dried at 40°C in vacuum for 48 h. The yield was 0.88 g. Fig.S4 shows ^1H NMR spectrum of umbelliferone-labeled PMAA.

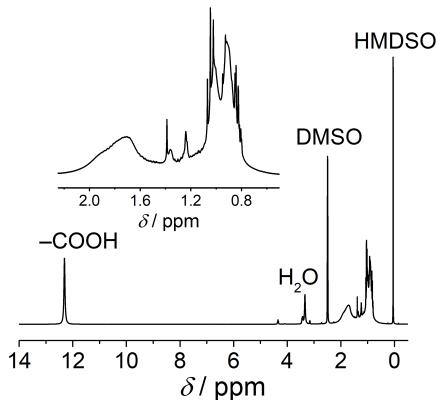


FIG. S4. ^1H NMR spectrum of umbelliferone-labeled PMAA in DMSO: 0.8 – 1.1 ppm (α -CH₃), 1.7 ppm (backbone CH₂), 6 – 8 ppm (protons at carbons C3, C5, C6, and C8 of umbelliferone label), 10.5 ppm (OH proton of umbelliferone label), 12.3 ppm (COOH). Insert: Detail of the spectrum for 2.25 – 0.5 ppm.

Size exclusion chromatography (SEC) of umbelliferone-labeled poly(*tert*-butyl methacrylate) precursor was performed at 25°C with two PLgel MIXED-C columns (300 × 7.5 mm, SDV gel with particle size 5 μm ; Polymer Laboratories, USA) and with UV (UVD 250; Watrex, Czech Republic, operated at 260 nm) and RI (RI-101; Shodex, Japan) detectors. Tetrahydrofuran was used as a mobile phase at a flow rate of 1 mL/min. The molecular weight values were calculated using Clarity software (Dataapex, Czech Republic). Calibration with poly(methyl methacrylate)

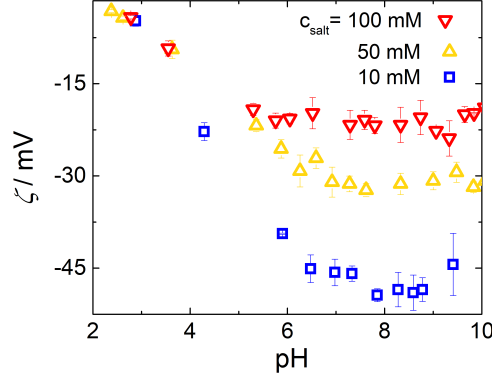


FIG. S5. ζ -potential of Umb-PMAA as a function of pH at various salt concentrations. Concentrations of $\text{Na}_2\text{B}_4\text{O}_7$ are indicated above.

standards (PSS, Germany) was used. ^1H NMR spectra of 4-chloromethylumbelliferone initiator and umbelliferone-labeled poly(methacrylic acid) were measured in dimethyl sulfoxide (DMSO-d_6) at 22°C using a Bruker DPX 300 spectrometer at 300.1 MHz. Hexamethyldisiloxane (HMDSO) was used as an internal standard.

II. SOLUTION PROPERTIES AND IONIZATION OF PMAA-UMB

A. Materials

Umbelliferone (Umb) was obtained from Fluka. The Umb-PMAA concentrations for both fluorescence and scattering measurements were 1 g/L. The Umb concentration for fluorescence measurements was ca. 1.5 mM, which corresponds to the molar concentration of Umb-PMAA. Both Umb and Umb-PMAA were dissolved in aqueous sodium tetraborate (10, 25, 50 or 100 mM) and the solution pH was adjusted by adding either sodium hydroxide or phosphoric acid.

B. Electrophoretic Light Scattering (ELS)

The ζ -potential measurements were carried out with a Nano-ZS Zetasizer (Malvern Instruments, UK). The ζ -potential values were calculated from electrophoretic mobilities (average of three subsequent measurements, each of which consisted of 15 runs) using the Henry equation in the Smoluchowski approximation,

$$\mu_i = \frac{\epsilon \zeta}{\eta_0} \quad (\text{S1})$$

where μ_i is the electrophoretic mobility and ϵ is the dielectric constant of the solvent.

C. Dynamic Light Scattering (DLS)

The light scattering setup (ALV, Langen, Germany) consisted of a 22 mW He-Ne laser ($\lambda = 632.8 \text{ nm}$), an ALV CGS/8F goniometer, an ALV High QE APD detector, and an ALV 5004 multi-bit, multi-tau autocorrelator. The measurements were carried out at 23°C for scattering angle $\theta = 90^\circ$. The polymer mass concentration in the solution was $c_{\text{pol}} = 1 \text{ g/L}$. Prior to measurements the samples were filtered using $0.45 \mu\text{m}$ Acrodisc PVDF membrane filters. DLS measurements were evaluated by fitting the normalized time autocorrelation function of the scattered light intensity using the constrained regularization algorithm (CONTIN), which provides the distribution of relaxation times τ . The distributions were converted to the distributions of apparent hydrodynamic radii, R_H , using the relationship

$$R_H = \tau \frac{8\pi n_0^2 k_B T}{3\eta_0 \lambda^2} \sin^2\left(\frac{\theta}{2}\right) \quad (\text{S2})$$

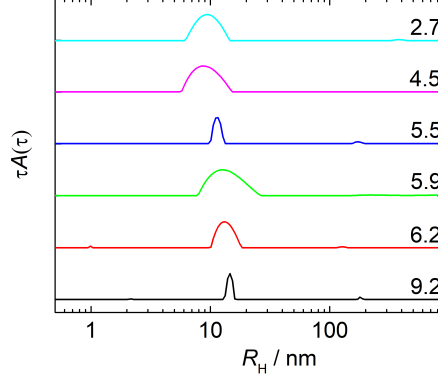


FIG. S6. DLS distributions of hydrodynamic radii of micelles prepared in 1 g/L Umb-PMAA in 50 mM $\text{Na}_2\text{B}_4\text{O}_7$ buffer as function of solution pH (at scattering angle $\theta = 90^\circ$).

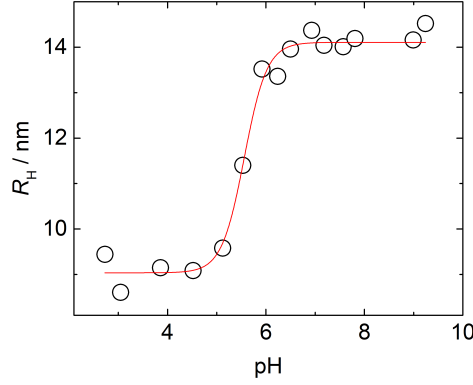


FIG. S7. Mean hydrodynamic radii, R_H , of micelles prepared in 1 g/L Umb-PMAA in 50 mM $\text{Na}_2\text{B}_4\text{O}_7$ as function of solution pH (at scattering angle $\theta = 90^\circ$).

where k_B is the Boltzmann constant, T is the temperature, n_0 is the refractive index of the solvent, and η_0 is the viscosity of the solvent.

D. Small-angle X-ray Scattering (SAXS)

SAXS experiments were performed on the high brilliance beamline ID02 [2] at ESRF (Grenoble, France) in 0.2 cm flow-through quartz capillaries at sample-to-detector distances 2.5 and 8 m, using a monochromatic incident X-ray beam with the energy, $E = 12.46$ keV, corresponding to scattering q vectors from 0.05 to 2.76 nm^{-1} . The SAXS setup utilized a pinhole camera with a beam stop placed in front of a two-dimensional Rayonix MX170-HS high-sensitivity low noise CCD detector having an active area of 100×100 mm, which was divided into 512×512 pixels with 4×4 binning for the sample-to-detector distance 2 m and into 1024×1024 pixels with 2×2 binning for 8 m. Online corrections were applied for the detector, and the sample-to-detector distance, center, transmission, and incident intensity were calibrated. 10 frames for buffer solution and 20 frames for Umb-PMAA solution with the accumulation time of 10 ms were taken and inspected for radiation damage. The frames without presence of radiation damage were azimuthally averaged to determine the dependence of the scattered intensity $I(q)$ on the scattering vector q . The calibration to absolute units (cm^{-1}) was performed by dividing the scattered intensity by the thickness of flow-through capillary (0.2 cm). The scattering from a capillary filled with 0.05 M sodium tetraborate was measured as a background and subtracted from the scattering intensity of the samples. The data were fitted by the cylinder form factor, using the

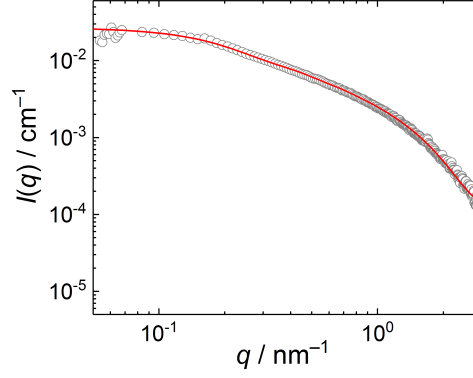


FIG. S8. Scattering curve for PMAA-Umb (1 g/L in 50 mM $\text{Na}_2\text{B}_4\text{O}_7$): SAXS data (open squares) and fit of the experimental data by the cylinder form factor (Eq. S3) (red line). Parameters of the model from the fit are: cylinder length, $L = 24$ nm; cylinder radius, $R = 1.2$ nm.

SASfit software, version 0.94.7 [3]

$$I(q) = 16I_0 \int_0^1 \left(\frac{J_1(qR\sqrt{1-x^2}) \sin(qLx/2)}{q^2 R \sqrt{1-x^2} L x} \right)^2 dx \quad (\text{S3})$$

where I_0 is the forward scattering intensity, L and R are the length and the radius of the cylinder, respectively, and $J_1(y)$ is the regular cylindrical Bessel function of the first order.

E. Fluorometry

All fluorescence emission measurements were carried out in 1 cm quartz cuvettes using a Fluorolog FL 3-22 fluorometer (Horiba Jobin Yvon, France) equipped with double-grating excitation and emission monochromators, Glan-Thompson polarizers, a FluoroHub time-correlated single photon counting module and a TBX single photon counting detector.

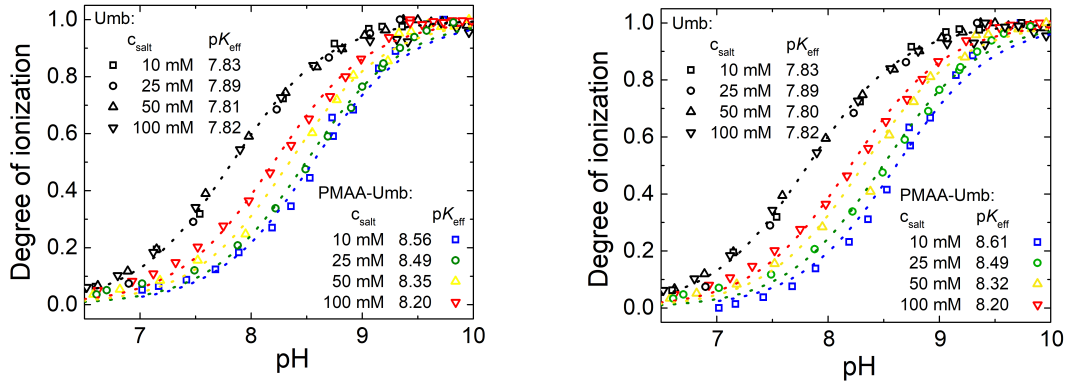


FIG. S9. Experimental titration curves of Umb-PMAA at various salt concentrations (points), fitted by the idea titration curve (dash lines) obtained from the integral (left) or maxima of the fluorescence emission (right).

Steady-state fluorescence measurements were performed using a 450 W high-pressure xenon arc lamp as a light source with the excitation wavelength 378 nm. We determined the Umb ionization using either intensity at the emission

maximum, or the integral of the whole emission spectrum. Both methods yielded very similar titration curves, shown in Fig. II E.

Anisotropy decays were measured using a 378 nm NanoLed diode laser source with the pulse FWHM of ca. 200 ps, operated at the repetition frequency of 1 MHz. Emission decays were measured at the emission wavelength of 450 nm in a time window of 57 ns with a resolution of 55.8 ps per channel. The polarized components $I_{\parallel}(t)$ and $I_{\perp}(t)$ were accumulated quasi-simultaneously with a switching frequency of 30 s. The polarized decays were measured with the emission polarizer set in the fixed vertical position, and with the excitation polarization plane rotated 0° and 90° . The value of the G -factor was determined by an independent experiment. The anisotropy decays, $r(t)$, were fitted by the single exponential model:

$$r(t) = \frac{GI_{\parallel}(t) - I_{\perp}(t)}{GI_{\parallel}(t) + 2I_{\perp}(t)} = (r_0 - r_{\infty})\exp(-t/\varphi) + r_{\infty} \quad (\text{S4})$$

where φ is the rotational correlation time and r_0 and r_{∞} , respectively, are the initial and residual anisotropy.

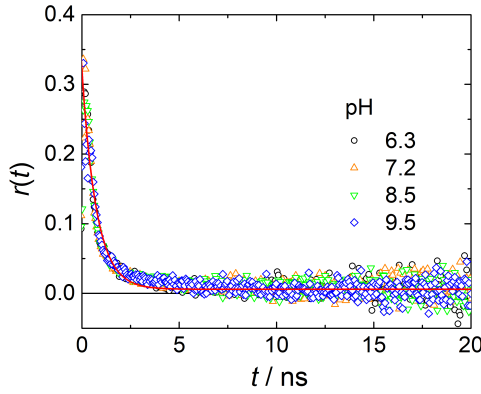


FIG. S10. Fluorescence anisotropy decays for 1 g/L Umb-PMAA in 50 mM $\text{Na}_2\text{B}_4\text{O}_7$ buffer as a function of pH (here excitation wavelength was $\lambda_{ex} = 378$ nm and emission wavelength was $\lambda_{em} = 450$ nm).

III. DERIVATION OF THE RELATION BETWEEN "LOCAL pH" AND EFFECTIVE $\text{p}K_A$

A. Relation between Δ and μ^{ex}

In this section we re-derive the original result of Hartley and Roe [4], starting from the chemical potentials. This will serve to introduce the notation, which is different from the seminal work, and to distinguish where our work deviates from the original work of Hartley and Roe. Consider the schematic chemical reaction of the fluorophore acid group A, attached to the end of the polymer chain:



Chemical equilibrium requires that

$$0 = \sum_i \nu_i \frac{\mu_i}{k_B T} = \frac{1}{k_B T} \sum_i \nu_i (\mu_i^{\ominus} + \mu_i^{\text{id}} + \mu_i^{\text{ex}}), \quad (\text{S6})$$

where $i = \{\text{HA}, \text{A}^-, \text{H}^+\}$, $\mu^{\ominus}$ is the reference chemical potential μ^{id} is the ideal gas chemical potential, and μ^{ex} is the excess chemical potential due to intermolecular interactions. We define the bare acidity constant, K_A , and μ_i^{id} as

$$\ln K_A \equiv -\frac{1}{k_B T} \sum_i \nu_i \mu_i^{\ominus}, \quad \frac{\mu_i^{\text{id}}}{k_B T} \equiv -\ln \frac{c_i}{c^{\ominus}}, \quad (\text{S7})$$

where c_i is the concentration of species i , and $c^{\ominus}$ is the reference concentration. Note that in a nano-heterogeneous system, such as a polyelectrolyte solution, the chemical potential is constant independent of position., whereas the

local concentration can vary with position [5]. We define the degree of ionization as

$$\alpha = \frac{c_{A^-}}{c_{A^-} + c_{HA}} \quad (\text{S8})$$

and rewrite Eq. S6 as

$$\ln K_A = \frac{\mu_{H^+}}{k_B T} + \ln \frac{\alpha}{1 - \alpha} + \frac{1}{k_B T} \sum_{i \neq H^+} \nu_i \mu_i^{\text{ex}} \quad (\text{S9})$$

We now introduce the following symbols

$$\text{pH} = -\log_{10} a_{H^+} = -\frac{\mu_{H^+} - \mu_{H^+}^\ominus}{k_B T \ln(10)}, \quad \text{p}K_A = -\log_{10} K_A, \quad \Delta = \frac{1}{k_B T \ln(10)} \left(\sum_{i \neq H^+} \nu_i \mu_i^{\text{ex}} \right). \quad (\text{S10})$$

Note that the excess chemical potential of H^+ is included in the definition of pH, consistent with the IUPAC definition [6]. Therefore, Δ includes all other excess chemical potentials, except H^+ . We define the effective pH and effective $\text{p}K_A$ as

$$\text{pH}_{\text{eff}} \equiv \text{pH} - \Delta, \quad \text{p}K_{\text{eff}} \equiv \text{p}K_A + \Delta, \quad (\text{S11})$$

Using Δ , $\text{p}K_{\text{eff}}$ and pH_{eff} , we can cast Eq. S9 in the form of the ideal titration curve

$$\text{pH} - \text{p}K_A - \Delta = \text{pH} - \text{p}K_{\text{eff}} = \text{pH}_{\text{eff}} - \text{p}K_A = \log_{10} \frac{\alpha}{1 - \alpha}. \quad (\text{S12})$$

The last equation demonstrates that we can equivalently describe the observed changes in ionization behavior as the effective ionization constant, $\text{p}K_{\text{eff}}$, or as the effective pH, pH_{eff} .

B. Determination of μ^{ex} from $g(r)$

To determine Δ from the radial distribution functions, we need to estimate the excess chemical potentials, μ^{ex} , from the local concentrations of H^+ . For convenience, we choose the coordinate system such that the dissociating fluorophore is at the origin, and express μ_i^{ex} as[7]

$$\mu_i^{\text{ex}} = \frac{1}{V} \int_0^1 \int_0^\infty w_i(r) g_i(r) 4\pi r^2 dr d\xi \approx \frac{1}{V} \int_0^1 \int_0^{R_{\text{max}}} w_i(r) g_i(r) 4\pi r^2 dr d\xi, \quad (\text{S13})$$

where ξ is the coupling parameter, V is the system volume, $w_i(r) = k_B T \ln g_i(r)$ is the potential of mean force due to all other particles felt by particle i at distance r from the origin and $g_i(r) = c_i(r)/\langle c_i \rangle$ is the radial distribution function of particles of type i around the origin. The upper bound of the integral in Eq. S13 should be ∞ in the thermodynamic limit. When using $g(r)$ from the simulations, it is truncated at R_{max} , such that $g(r > R_{\text{max}}) \approx 1$, and $V = 4\pi R_{\text{max}}^3/3$ is the corresponding volume. Note that Eq. S13 was averaged over the angular coordinates, assuming spherically symmetric interactions. If the potential of mean force is not spherically symmetric, then Eq. S13 is only approximate because it does not explicitly consider the angular dependence.

To establish the link between the local concentration of H^+ ions and the chemical potential, we approximated the potential of mean force potential by the mean electrostatic potential to obtain

$$z_i e \langle \psi(r) \rangle \approx w_i(r) \equiv -k_B T \ln g_{H^+}(r) = -k_B T \ln g_{Na^+}(r), \quad (\text{S14})$$

where $\langle \psi(r) \rangle$ is the mean electrostatic potential at r , z_i is the valency of particle i and e is the elementary charge. In the case of the end-labeled polyelectrolyte chain, $\langle \psi(r) \rangle$ is predominantly determined by charges on the chain; therefore, we can assume that $\langle \psi(r) \rangle$ is independent of ξ . Note that the last equality in Eq. S14 holds only in our simulation because H^+ and Na^+ have the same interaction parameters, therefore, $g_{H^+}(r) = g_{Na^+}(r)$. In reality, the distributions of H^+ and Na^+ will differ on short distances comparable to the ion size, where $w(r)$ is dominated by steric interactions instead of electrostatics. The higher number of Na^+ ions in the simulation box provides $g(r)$ with much better statistical quality (less noise). Therefore, we used profiles of Na^+ ions to determine the "local pH".

Using Eq. S14 we can estimate the mean electrostatic potential at a given point from the simulated radial distribution functions of counterions. The value of $\langle \psi(r) \rangle$ can be then used in Eq. S13 to determine the chemical potentials. Because

$z_{\text{HA}} = 0$, from Eq. S14 we immediately obtain $\mu_{\text{HA}}^{\text{ex}} = 0$. In the simulation described below, the ionized acid group of the fluorophore A^- is located at the origin, and H^+ ion is generated anywhere in the system; therefore: $g_{\text{A}^-}(r) = \delta(r)$, and

$$\mu_{\text{A}^-}^{\text{ex}} = -k_{\text{B}}T \lim_{r \rightarrow 0} \langle e\psi(r) \rangle, \quad (\text{S15})$$

For distances smaller than the particle size, $r \lesssim \sigma$, the $g(r)$ is dominated by short-ranged excluded volume interactions. Therefore, we estimated $\langle \psi(0) \rangle$ for A^- from $g(r)$ beyond the cutoff radius, $r_{\text{cut}} = 1.9\sigma$, where the effect of the excluded volume is negligible $\langle \psi(0) \rangle \approx \langle \psi(r_{\text{cut}}) \rangle \approx -k_{\text{B}}T \ln g_{\text{H}^+}(r_{\text{cut}})$. The value of $r_{\text{cut}} = 1.9\sigma$ was chosen by comparing $g(r)$ from simulations of the same systems at various pH (see Fig. S13). At shorter distances, $g(r)$ exhibited a sharp peak characteristic of the short-range excluded volume repulsion, and the height of this peak slightly depended on pH. Starting from $r_{\text{cut}} = 1.9\sigma$, the $g(r)$ at different pH coincide within the statistical error (see also Fig. S12). Combining Eq. S15 and Eq. S14, we substitute for $\mu_{\text{A}^-}^{\text{ex}}$ into Eq. S10 to obtain

$$\Delta \approx \frac{\mu_{\text{A}^-}^{\text{ex}}}{k_{\text{B}}T} \approx -\frac{e \langle \psi(r=0) \rangle}{k_{\text{B}}T} \approx \ln g_{\text{H}^+}(r_{\text{cut}}) = \ln \frac{c_{\text{H}^+}(r_{\text{cut}})}{\langle c_{\text{H}^+} \rangle} = \ln \frac{c_{\text{Na}^+}(r_{\text{cut}})}{\langle c_{\text{Na}^+} \rangle} \quad (\text{S16})$$

The last equation provides a formula to compute Δ using $\langle \psi(r=0) \rangle$ or $g(r_{\text{cut}})$ obtained from the simulation.

C. Determination of $\langle \psi(r=0) \rangle$ using a test charge

We determined $\psi(r=0)$ by placing a test charge at the fluorophore in the simulation. We ran the standard Langevin dynamics simulation (see next section for details) with a polyelectrolyte chain and uncharged end segments representing the fluorophores. At regular intervals, we placed the test charge at the centre of the fluorophore, and measured the associated change in the electrostatic energy of the whole system which provided the instantaneous value of the electrostatic potential $\psi(r=0, t)$. Then we removed the test charge and continued the original simulation. The average value of $\langle \psi(r=0) \rangle$ was determined by averaging over many measurements of its instantaneous value, same as for any other quantity in the simulation.

IV. SIMULATION MODEL AND METHOD

The simulations were performed using the Kremer-Grest bead-spring model of a polyelectrolyte in the Langevin thermostat [8]. The model is similar to the one described in our previous publication [9]. The excluded volume interactions are modeled using the Weeks-Chandler-Andersen (WCA) potential between all particle pairs:

$$U(r) = \begin{cases} \infty & \text{for } r < r_{\text{off}} \\ 4\epsilon \left(\left(\frac{\sigma}{r-r_{\text{off}}} \right)^{12} - \left(\frac{\sigma}{r-r_{\text{off}}} \right)^6 + \frac{1}{4} \right) & \text{for } r - r_{\text{off}} \leq 2^{1/6}\sigma \\ 0 & \text{for } r - r_{\text{off}} \geq 2^{1/6}\sigma \end{cases} \quad (\text{S17})$$

where we choose $\epsilon = 1 k_{\text{B}}T$ and $\sigma = 0.40 \text{ nm}$, which defines the effective particle size. Unless stated otherwise, we use the offset $r_{\text{off}} = 0$. Bonds between the polymer beads are modeled using the finite-extensible non-linear elastic (FENE) potential:

$$U(r) = -\frac{1}{2}k_{\text{F}}r_{\text{F}}^2 \ln \left[1 - \left(\frac{r}{r_{\text{F}}} \right)^2 \right] \quad \text{for } r < r_{\text{F}} \quad (\text{S18})$$

with $k_{\text{F}} = 10\epsilon/\sigma^2 \approx 62 k_{\text{B}}T/\text{nm}^2$ and $r_{\text{F}} = 1.5\sigma \approx 0.60 \text{ nm}$, which in combination with the WCA potential provides the mean bond length of $b = 1.03\sigma$. Electrostatic interactions between particles i and j are represented by the Coulomb potential,

$$U_{i,j}(r) = z_i z_j k_{\text{B}}T \frac{l_{\text{B}}}{r} = \frac{1}{4\pi\epsilon_0\epsilon_r} \frac{z_i z_j e^2}{r} \quad (\text{S19})$$

where l_{B} is the Bjerrum length, z is the valency, e is the elementary charge, ϵ_0 and ϵ_r are the permittivity of free space and the relative permittivity of the medium, respectively.

Ionization reaction, in the constant pH simulation is implemented as a Monte Carlo move, changing the state from non-ionized to ionized (forward direction) or vice versa (reverse direction). In the forward direction, the chemical identity of the acid group HA is changed to A^- , and the proton H^+ is inserted at a random position in the system. In the reverse direction, one H^+ is deleted, and the A^- group changes to HA. The new state is accepted with the probability [10]

$$P_{\text{acc}}^{\pm} = \min \left[1, \exp \left(-\frac{U_n - U_o}{k_B T} \pm (\text{pH} - \text{p}K_A) \ln(10) \right) \right] \quad (\text{S20})$$

where U_n and U_o are the energy of the new and of the old state. The signs $+$ and $-$ refer to the forward and reverse direction. Our range of pH and ionic strength is safe with respect to artifacts of the constant pH ensemble due to ionic screening effects [11].

Assuming that one monomer in our simulation corresponds to one unit of methacrylic acid with molar mass $M = 86.0 \text{ g/mol}$, and taking $\sigma = 0.40 \text{ nm}$ as the monomer size, we prepared the simulation setup with $n = 3$ chains in the box, and the box size chosen to represent the experimental polymer concentration $c_{\text{pol}} = 1.0 \text{ g/L}$. We set the Bjerrum length $l_B = 1.75 \sigma \approx 0.7 \text{ nm}$ which corresponds to aqueous solution at room temperature. With chain length $N = 51$ segments and 3 chains per box, our simulation box length becomes $L \approx 69 \sigma$. The box contains 263 salt ion pairs at $c_{\text{salt}} = 10 \text{ mmol/L}$ and 2630 salt ion pairs at $c_{\text{salt}} = 100 \text{ mmol/L}$. To check that finite size effects are negligible, on selected systems we performed also simulations with 5 chains per box and 1 chain per box. Because the dissociation of sodium tetraborate in solution yields a complex mixture of monovalent and divalent anions, we also verified that replacing the monovalent anions with the corresponding number of divalent anions has negligible effect on the obtained result. We used the *ESPResSo* simulation software [12, 13] to perform our simulations (development code, git commit b5c1f17389ef54899c). The electrostatic interactions were calculated using the P3M method implemented in *ESPResSo*. Our time step in the Langevin dynamics was $\delta t = 0.01 \tau$, where $\tau = \sigma \sqrt{m/\epsilon}$ is the reduced unit of time. The typical simulation time was $10^7 \delta t = 10^5 \tau$, yielding about 10^3 uncorrelated samples of polymer conformations, measured by the autocorrelation time of the radius of gyration. The reaction steps were performed in the intervals of 35τ , with 5 reaction attempts per each ionizable group in the system, yielding about twice the number of uncorrelated samples of the degree of ionization, as compared to the radius of gyration.

V. ADDITIONAL SIMULATION RESULTS

A. Alternative representation of the $\text{p}K_A$ shifts

Instead of plotting the shift $\Delta = \text{p}K_{\text{eff}} - \text{p}K_A$ in the form of the titration curve, we may use Eq. S12 to determine $\text{p}K_{\text{eff}}$ from each simulation at a given value of pH. The values of $\text{p}K_{\text{eff}}$ obtained using Eq. S12 are shown in Fig. S11, showing that $\text{p}K_{\text{eff}}$ is independent of pH. In the case of $\text{p}K_{\text{eff}}$ obtained from the $g(r_{\text{cut}})$, shown in Fig. S12, there is no dependence on pH either. Fig. S13 shows that some dependence of $g(r)$ can be observed at shorter distances, where the steric interactions dominate over electrostatics. However, even in this case the difference is comparable to the estimated statistical error. This corroborates our assumption that the shift in $\text{p}K_{\text{eff}}$ is predominantly caused by the interaction with charges on the polymer chain, and the influence of Umb ionization plays a minor role. Fig. S14 shows the difference between the $\text{p}K_{\text{eff}}$ obtained from the titration curves, and from the analysis of the radial distribution functions. Within the estimated statistical error, this difference attains a constant value of about 0.2.

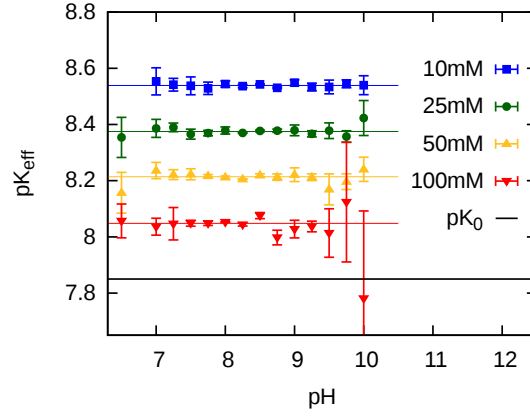


FIG. S11. The value of pK_{eff} calculated from the degree of ionization α using Eq. S12 for each simulation at various pH values. The horizontal lines represent the pK_{eff} values obtained from fitting the titration curves. Systems at low and high pH, which correspond to $\alpha \approx 0$ and $\alpha \approx 1$ respectively, have much higher uncertainties than those at intermediate values of α . All simulations performed with $N = 51$, $c_{\text{pol}} = 1 \text{ g/L}$, $l_B/\sigma = 1.75$, and 1 spacer.

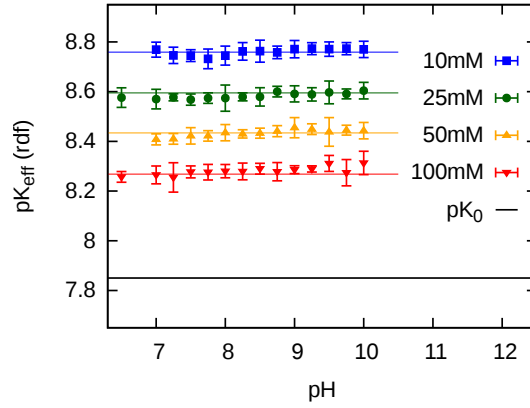


FIG. S12. pK_{eff} calculated from rdf using Eq. S16 for each simulation at various pH values. All simulations performed with $N = 51$, $c_{\text{pol}} = 1 \text{ g/L}$, $l_B/\sigma = 1.75$, and 1 spacer.

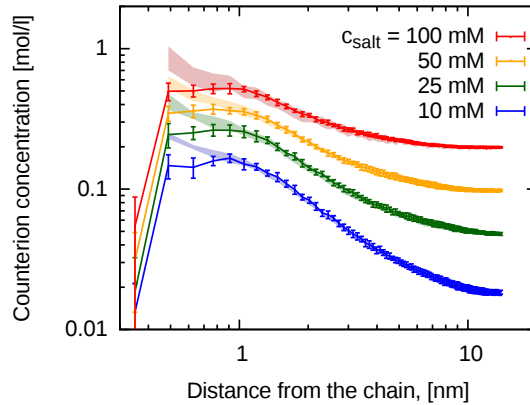


FIG. S13. Radial distribution functions, $g(r)$, of Na^+ ions around the fluorophore at two extreme pH values: pH=7.0 (points connected with lines) and pH=10.0 (shaded areas), and various salt concentrations. Intermediate pH values fall in between the two extremes. At distances $r \lesssim 0.7 \text{ nm}$ the $g(r)$ at different pH and the same salt concentration differ due to steric effects (local crowding). Beyond this distance they coincide within the statistical error. All simulations performed with $N = 51$, $c_{\text{pol}} = 1 \text{ g/L}$, $l_B/\sigma = 1.75$, and 1 spacer.

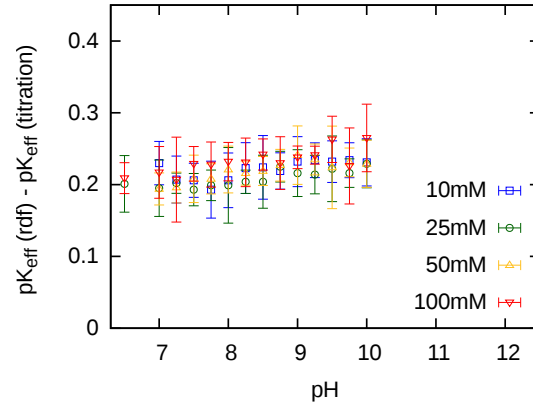


FIG. S14. Difference between pK_{eff} calculated from the degree of ionization α using Eq. S12, and from rdf using Eq. S16 for each simulation at various pH values. This figure confirms that within the estimated uncertainty this difference attains a constant value of ≈ 0.2 . All simulations performed with $N = 51$, $c_{\text{pol}} = 1 \text{ g/L}$, $l_B/\sigma = 1.75$, and 1 spacer.

B. Sensitivity of the simulated pK_A shift to model parameters

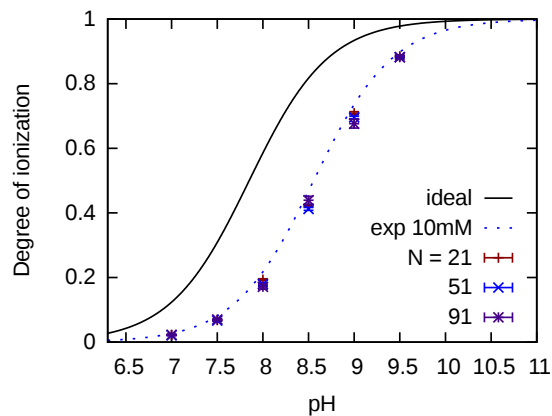


FIG. S15. Influence of chain length on the titration curve of the end segment: an increases or decrease in the chain length roughly by a factor of two has no significant effect on the ionization. Reference simulations performed with $c_{\text{pol}} = 1 \text{ g/L}$, $l_B/\sigma = 2.0$, and 1 spacer.

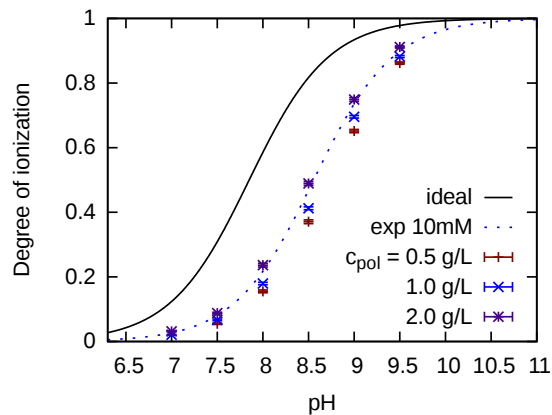


FIG. S16. Influence of the polymer concentration on the titration curve of the end segment: dilution of the polymer suppresses ionization. Reference simulations performed with $N = 51$, $l_B/\sigma = 2.0$, and 1 spacer.

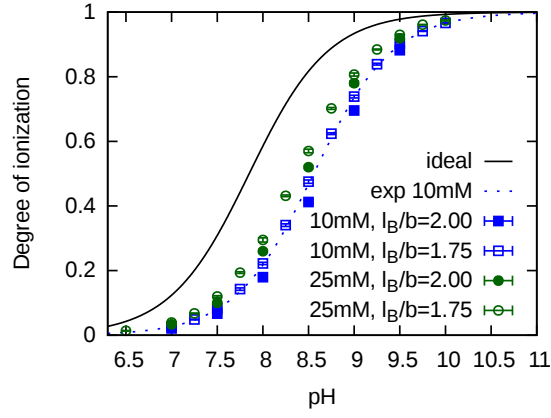


FIG. S17. Influence of the electrostatic coupling strength, l_B/σ on the titration curve of the end segment: an increase in the coupling strength suppresses the ionization. Reference simulations performed with $N = 51$, $c_{\text{pol}} = 1 \text{ g/L}$, and 1 spacer.

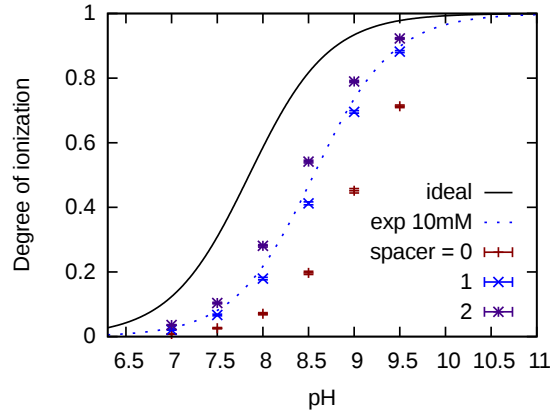


FIG. S18. Influence of the number of spacers between the polyelectrolyte chain and the end segment: addition of a spacer leads to higher ionization. Reference simulations performed with $N = 51$, $c_{\text{pol}} = 1 \text{ g/L}$ and $l_B/\sigma = 2.0$.

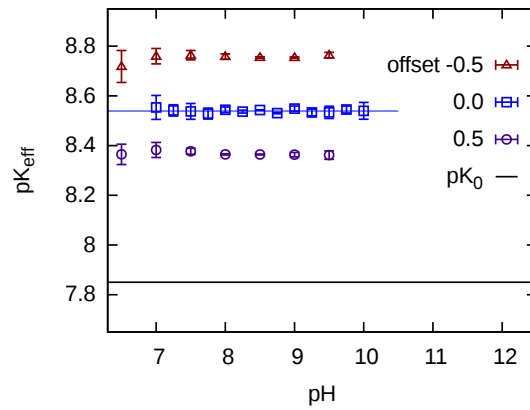


FIG. S19. Influence of the fluorophore size on the titration curve of the end segment: an increase in the fluorophore size increases its ionization. The offset reported in the legend is the relative increase in the fluorophore size. Reference simulations performed with $N = 51$, $c_{\text{pol}} = 1 \text{ g/L}$, $l_B/\sigma = 1.75$, and 1 spacer.

-
- [1] W. Lin, L. Long, J. Feng, B. Wang, and C. Guo, *European Journal of Organic Chemistry* **2007**, 4301.
 - [2] P. Van Vaerenbergh, J. Léonardon, M. Sztucki, P. Boesecke, J. Gorini, L. Claustre, F. Sever, J. Morse, and T. Narayanan, in *AIP Conference Proceedings*, Vol. 1741 (2016) p. 030034.
 - [3] I. Breßler, J. Kohlbrecher, and A. F. Thünemann, *Journal of Applied Crystallography* **48**, 1587 (2015).
 - [4] G. S. Hartley and J. W. Roe, *Trans. Faraday Soc.* **35**, 101 (1940).
 - [5] L. Nová, F. Uhlík, and P. Košovan, *Phys. Chem. Chem. Phys.* **19**, 14376 (2017).
 - [6] IUPAC. *Compendium of Chemical Terminology*, 2nd ed. (the "Gold Book"). Compiled by A. D. McNaught and A. Wilkinson. Blackwell Scientific Publications, Oxford (1997). doi:10.1351/goldbook.
 - [7] D. A. McQuarrie, *Statistical Mechanics* (Harper Collins, New York, 1976).
 - [8] G. S. Grest and K. Kremer, *Physical Review A* **33**, 3628 (1986).
 - [9] P. Košovan, T. Richter, and C. Holm, *Macromolecules* **48**, 7698 (2015).
 - [10] C. E. Reed and W. F. Reed, *The Journal of Chemical Physics* **96**, 1609 (1992).
 - [11] J. Landsgesell, C. Holm, and J. Smiatek, *The European Physical Journal Special Topics* **226**, 725 (2017).
 - [12] A. Arnold, O. Lenz, S. Kesselheim, R. Weeber, F. Fahrenberger, D. Röhm, P. Košovan, and C. Holm, in *Meshfree Methods for Partial Differential Equations VI*, Lecture Notes in Computational Science and Engineering, Vol. 89, edited by M. Griebel and M. A. Schweitzer (Springer, 2013) pp. 1–23.
 - [13] "ESPResSo homepage," <http://espressomd.org>.