Supporting Information for

pH dependence of the assembly mechanism and properties of poly(L-lysine) and poly(L-glutamic acid) complexes

Tuuva Kastinen[a,b,c], Dawid Lupa[d,e], Piotr Bonarek[f], Dmitrii Fedorov[c,g], Maria Morga[d], Markus B. Linder[c,g], Jodie L. Lutkenhaus[h,i], Piotr Batys[*d], Maria Sammalkorpi[*a,c,g]

[a.] *Department of Chemistry and Materials Science, Aalto University, P.O. Box 16100, 00076 Aalto, Finland*
[b.] *Faculty of Engineering and Natural Sciences, Chemistry & Advanced Materials, Tampere University, P.O. Box 541, 33014 Tampere University, Finland*
[c.] *Academy of Finland Center of Excellence in Life-Inspired Hybrid Materials (LIBER), Aalto University, P.O. Box 16100, 00076 Aalto, Finland*
[d.] *Jerzy Haber Institute of Catalysis and Surface Chemistry, Polish Academy of Sciences, Niezapominajek 8, PL-30239 Krakow, Poland*
[e.] *M. Smoluchowski Institute of Physics, Jagiellonian University, Łojasiewicza 11, 30-348 Kraków, Poland*
[f.] *Department of Physical Biochemistry, Faculty of Biochemistry, Biophysics and Biotechnology, Jagiellonian University, Gronostajowa 7, 30-387 Krakow, Poland*
[g.] *Department of Bioproducts and Biosystems, Aalto University, P.O. Box 16100, 00076 Aalto, Finland*
[h.] *Artie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, Texas 77840, USA*
[i.] *Department of Materials Science and Engineering, Texas A&M University, College Station, Texas 77840, USA.*

* Corresponding authors: piotr.batys@ikifp.edu.pl; maria.sammalkorpi@aalto.fi

Hydrodynamic diameters of the polypeptides

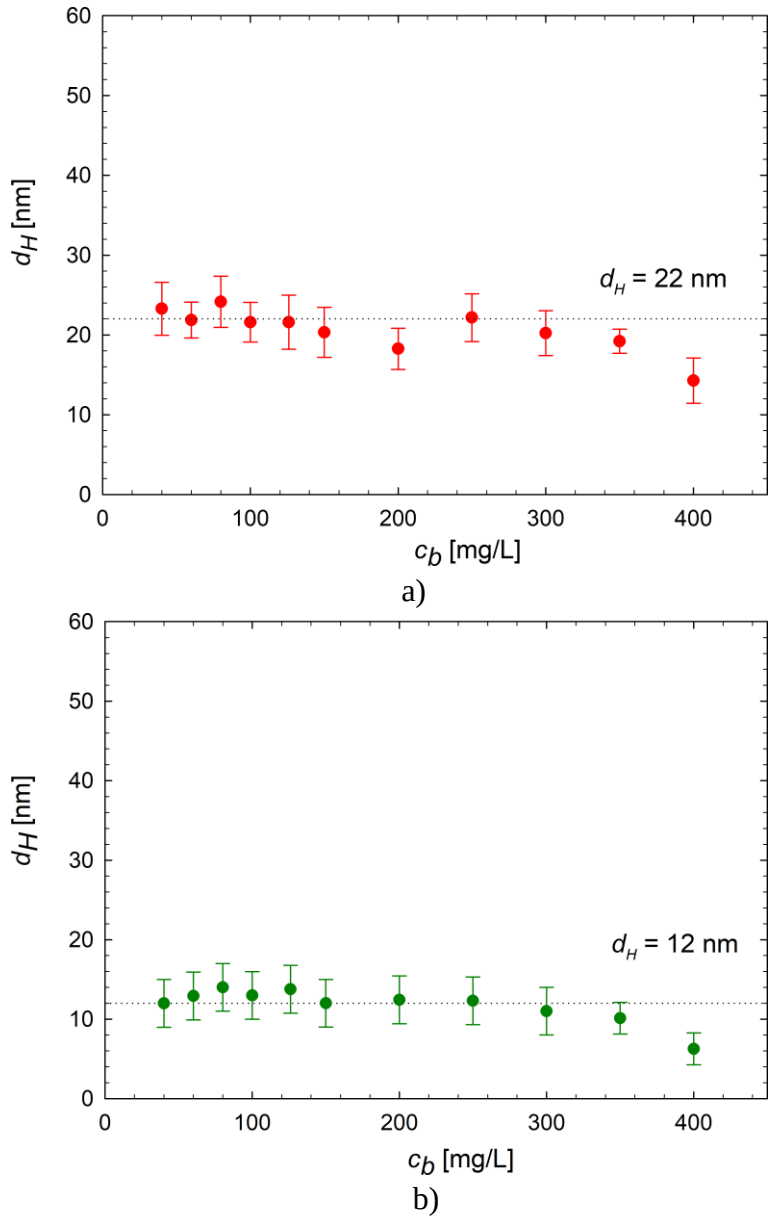


Figure S1. Hydrodynamic diameter d_H dependence on polypeptide bulk concentration c_b of a) PLL and b) PGA, measured at neutral pH.

Sedimentation velocity and sedimentation equilibrium experiments via analytical ultracentrifugation

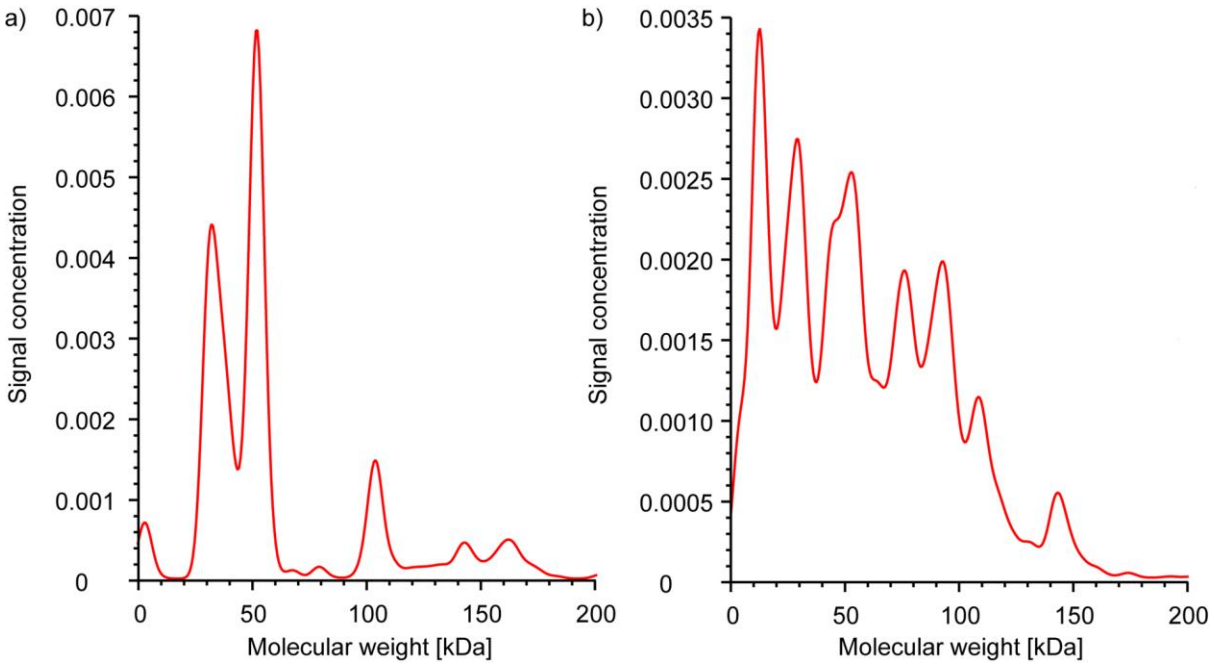


Figure S2. Molecular weight distribution of a) PLL and b) PGA molecules determined by the analytical ultracentrifuge technique (AUC).

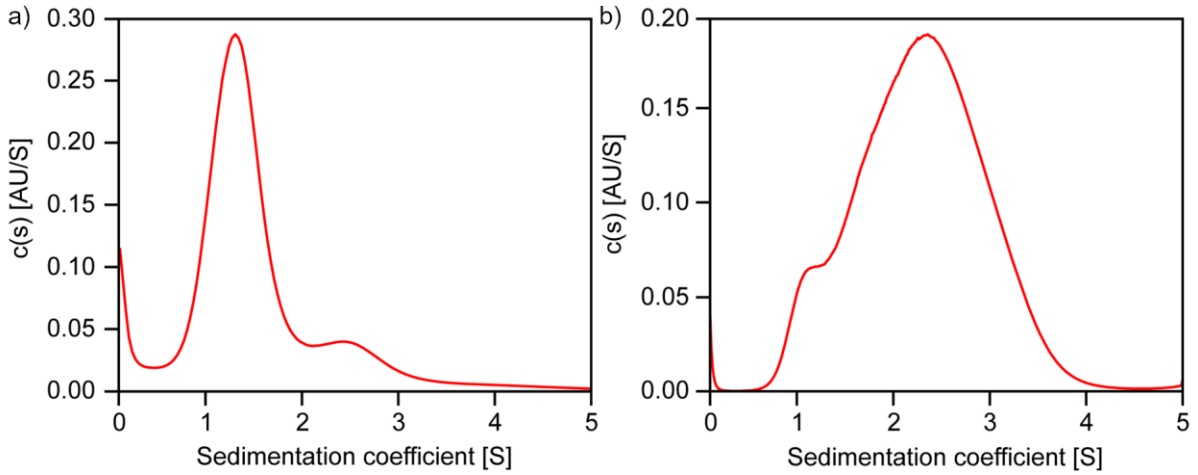


Figure S3. Calculated distributions *c*(*s*) of the sedimentation coefficient *s* obtained by Sedfit for the AUC data. In panel a), data for PLL is presented. The peaks correspond approximately to 44 and 145 kDa molecular weight. In panel b), data for PGA is presented. The peak corresponds approximately to 48.5 kDa molecular weights. The deviation in molecular weight appears due to specificity of analysis. Sedfit uses an averaged frictional ratio in determining molecular weight. In case of polydisperse sample, this can cause deviation in molecular weight.

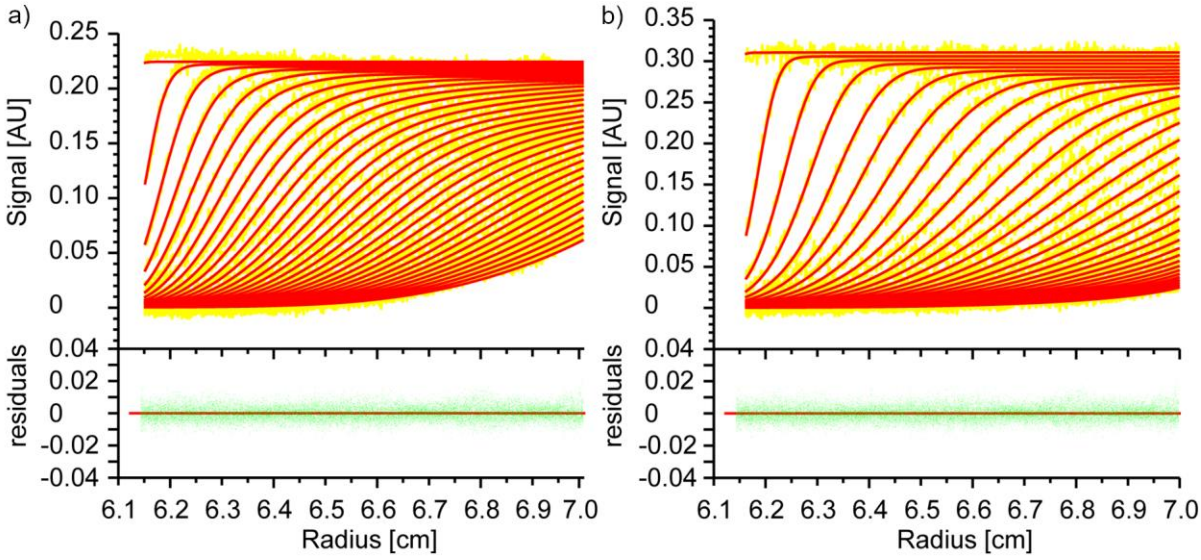


Figure S4. Sedimentation velocity experiments data (yellow lines) and simulated data based on the 2DSA Monte Carlo analysis (red lines) of the AUC data. Panel a) presents PLL and panel b) PGA. Residuals are shown in green in the panel below.

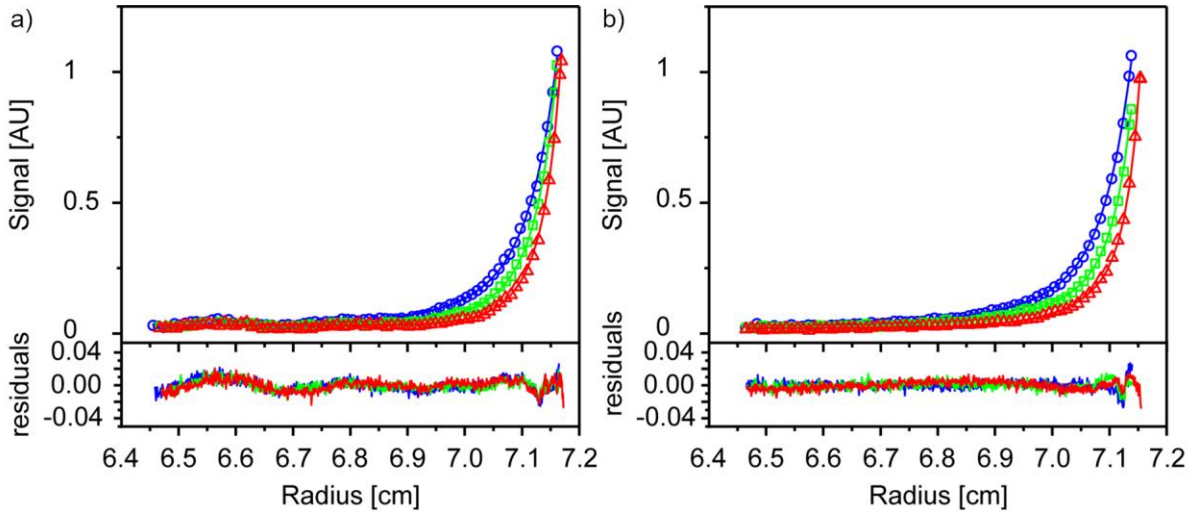


Figure S5. Sedimentation equilibrium distributions of a) PLL (220 mg/L) and b) PGA (220 mg/L) at three different AUC rotation speeds: 20 000 rpm (blue circles), 25 000 rpm (green squares), and 30 000 rpm (red triangles). Fitting curves obtained with the two-species model are shown as continuous line in matching color.

pH-metric titration of PGA by PLL

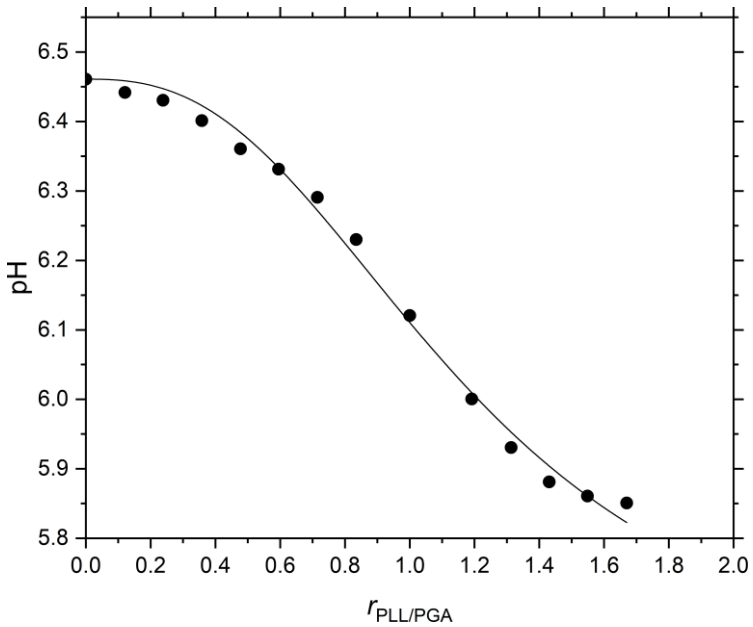


Figure S6. pH vs molar ratio of PLL/PGA $r_{PLL/PGA}$ for PGA solution (100 mg/L) after addition of different volumes of PLL solution (1000 mg/L, pH = 6.5). Solid line denotes non-linear fit to experimental data using a logistic curve. Ionic strength I = 0.01 M NaCl and temperature T = 298 K.

Curve fitting procedures for DLS and LDV experiments.

The observed dependencies of hydrodynamic diameter, d_H on the molar ratio of components (r_m), are different for pH conditions promoting and conditions preventing complex formation. This is reflected by the type of curve used to fit the experimental data. For pH conditions preventing complex formation, a single exponential formula was sufficient to fit the data:

$$d_H = a + be^{\frac{r_{PLL/PGA}}{c}} \quad \text{(S1)}$$

In this, a, b, and c are fitting parameters and e is the Euler's number. The fitting parameters obtained in this way were collected to Table S1.

Both for the DLS measurements performed under pH conditions promoting complexation and zeta potential measurements (regardless of pH value), the logistic (sigmoid) curve given in general form by

$$y = a + \frac{b-a}{1+\left(\frac{x}{x_0}\right)^p} \quad \text{(S2)}$$

was used as the fitting curve to the experimental results. Here a and *b* are constants corresponding to the maximum and minimum values, respectively. x_0 is the midpoint and *p* is the slope parameter describing the steepness of the curve.

Table S1. Fitting parameters for experimental results acquired in DLS experiments.

Logistic fit					
pH	***a***	***b***	x_0	***p***	R^2
6.2	734 ± 37	13 ± 1	3.56 ± 0.46	2.62 ± 0.16	0.99
10	280 ± 22	13 ± 2	2.0 ± 0.2	2.64 ± 0.31	0.99

Exponential fit				
pH	***a***	***b***	***c***	R^2
3	30.5 ± 1.2	19.1 ± 1.6	-0.16 ± 0.05	0.95
4	53.7 ± 2.0	-45.7 ± 2.4	-0.15 ± 0.05	0.99
12	56.1 ± 12.2	-42 ± 12.2	-1.18 ± 0.07	0.83

In the fitting procedure applied to the experimental data acquired from LDV measurements, the dependent and independent value is the measured zeta potential and the PLL/PGA molar ratio $r_{PLL/PGA}$, respectively. The parameters *a* and *b* were fixed at values equal to ζ_{PLL} and ζ_{PGA} respectively. Therefore, only two external parameters were fitted to the experimental data. With above, Eq. (S2) can be rewritten in the following form:

$$\zeta = \zeta_{PLL} + \frac{\zeta_{PGA} - \zeta_{PLL}}{1 + \left(\frac{r_{PLL/PGA}}{r_0} \right)^p} \tag{S3}$$

The fitting parameters and coefficients of determination (R^2) for each investigated pH value were collected in Table S2.

Table S2. Fixed parameters and fitting parameters for experimental data presented in Figs. 3 and S7 obtained using Eq. S2.

pH	ζ_{PGA} **[mV]**	ζ_{PLL} **[mV]**	r_0	***p***	R^2
3	-11	64	0.78 ± 0.01	4.98 ± 0.47	0.99
4	-39	60	0.19 ± 0.01	5.15 ± 0.22	0.99
6.2	-53	63	1.19 ± 0.07	1.65 ± 0.22	0.95
10	-73	34	1.13 ± 0.06	1.89 ± 0.21	0.98
12	-79	1	2.87 ± 0.16	0.69 ± 0.04	0.98

Assuming 1:1 complex stoichiometry, the zeta potential of the complex ζ_c was calculated from the following condition:

$$\zeta_c = \zeta\left(r_{PLL/PGA} = 1 \right) \tag{S4}$$

The values of ζ_c obtained in this manner were used to calculate the theoretical values of zeta potential both in the "two stage" and "no interaction" models.

It is worth pointing out that the distribution of electrophoretic mobility obtained via the LDV measurements is Gaussian with a single peak. Therefore, each measurement point in Figs. 3d-e in the main manuscript can be considered to represent an averaged value of zeta potential. Assuming that the molecular contributions are proportional to the molar fractions in the mixture, the following two stage model was proposed to quantitatively describe the observed changes in determined values of zeta potential:

$$\zeta_{R1} = x_c \zeta_c + x_{PGA} \zeta_{PGA} \tag{S5}$$

$$\zeta_{R2} = x_c \zeta_c + x_{PLL} \zeta_{PLL} \tag{S6}$$

Here ζ_c, ζ_{PGA}. and ζ_{PLL} are the zeta potentials of the PLL/PGA complex, PGA, and PLL, respectively. The corresponding molar fractions are x_c, x_{PGA}, and x_{PLL}.

Equation (S5) describes the first stage of experiments R1 in which PGA is in stoichiometric excess with respect to PLL. Ideally, in such systems only PLL/PGA complexes and unreacted PGA are present. Consequently, Eq. (S6) describes the second stage of experiments R2 which corresponds to molar excess of PLL. These equations correspond therefore to the first (R1) and second experimental regimes (R2), respectively. The curves obtained in this way were plotted in Figs. 3d-e in the main manuscript and in Fig. S7 using dashed (R1) and dotted line (R2). Additionally, similar, simple molar fraction weighted additive model for the zeta potential is proposed for non-interacting polypeptides:

$$\zeta = x_{\mathrm{PGA}}\zeta_{\mathrm{PGA}} + x_{\mathrm{PLL}}\zeta_{\mathrm{PLL}} \qquad \text{(S7)}$$

The results obtained using Eq. (S7) were also plotted in Figs. 3d-e in the main manuscript and in Fig. S7, as the "no interaction" model.

DLS and LDV results for pH 3.0 and 12.0

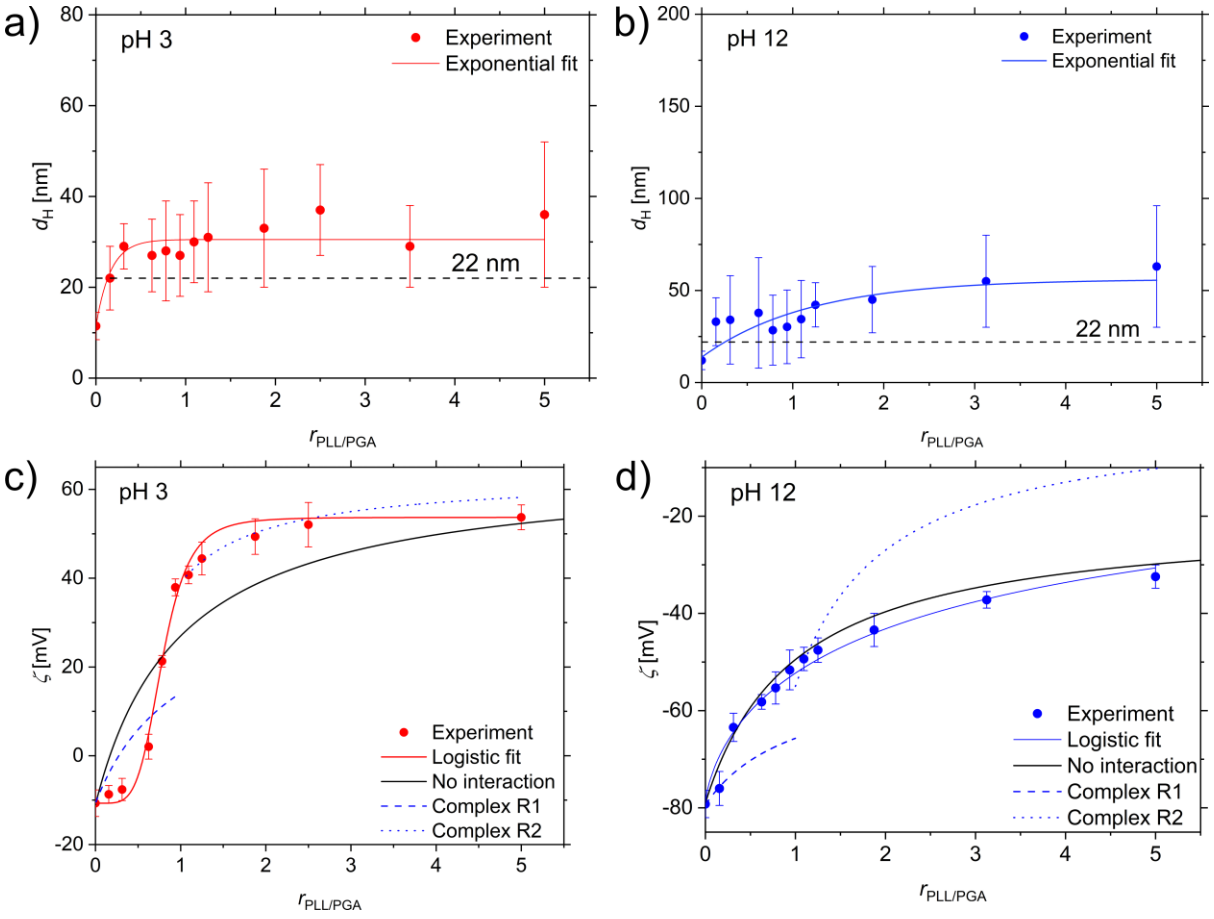


Figure S7. At top, the dependence of the mean hydrodynamic diameter d_H of the PLL/PGA system on the molar ratio of PLL to PGA $r_{PLL/PGA}$ determined for different pH values. Panel a) represents pH 3 and panel b) pH 12. The solid lines show the non-linear fits to experimental data. The dashed line is the theoretical estimate of hydrodynamic diameter of PLL/PGA complex, calculated according to Eq. 2 in the main manuscript. At bottom, the dependence of the mean zeta potential of PLL/PGA system on $r_{PLL/PGA}$ determined for different pH values. Panel c) corresponds to pH 3 and panel d) to pH 12. The dashed and dotted line are the theoretical values of zeta potential calculated for the first (R1) and second (R2) experimental regimes,

respectively. Black solid lines correspond to the "No interaction" model. Ionic strength I = 0.01 M NaCl and temperature T = 298 K.

CD spectra for pH 3.0 and 12.0

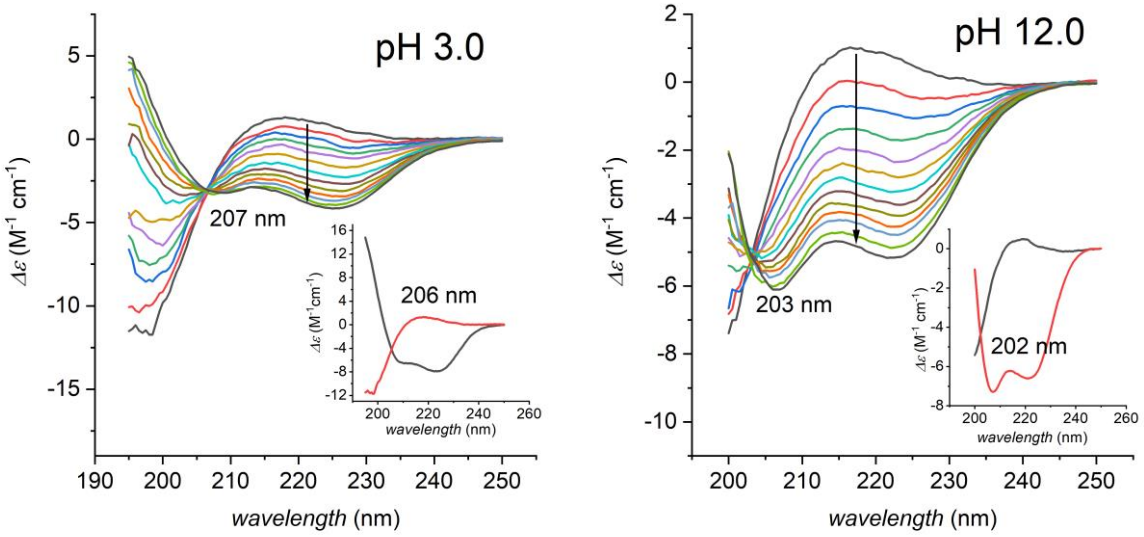


Figure S8. Changes in CD spectra during titrations of 0.1 mg/g PGA into 0.02 mg/g PLL solution at pH 3 and 1 mg/g PLL into 0.1 mg/g PGA solution at pH 12. Insets show the spectra of single component solutions (PGA – black line, PLL – red line). Ionic strength I = 0.01 M NaCl and temperature T = 298 K.

ITC data

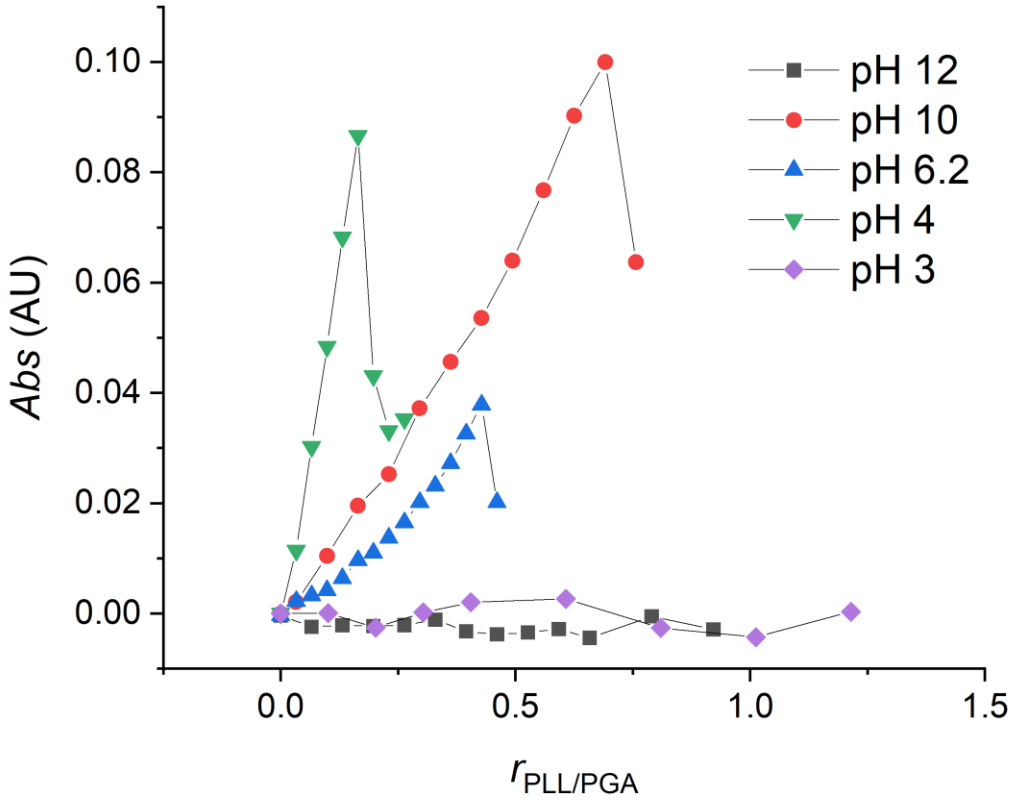


Figure S9. The changes in sample absorbance at 250 nm measured during titration of 1 mg/g PLL solution into 0.1 mg/g PGA at different pH. The exception was pH 3, where 0.2 mg/g PGA was titrated into 0.01 mg/g PLL solution. As PLL and PGA do not absorb at this wavelength, the increase of signal at this band results from light scattering. Ionic strength I = 0.01 M NaCl and temperature T = 298 K.

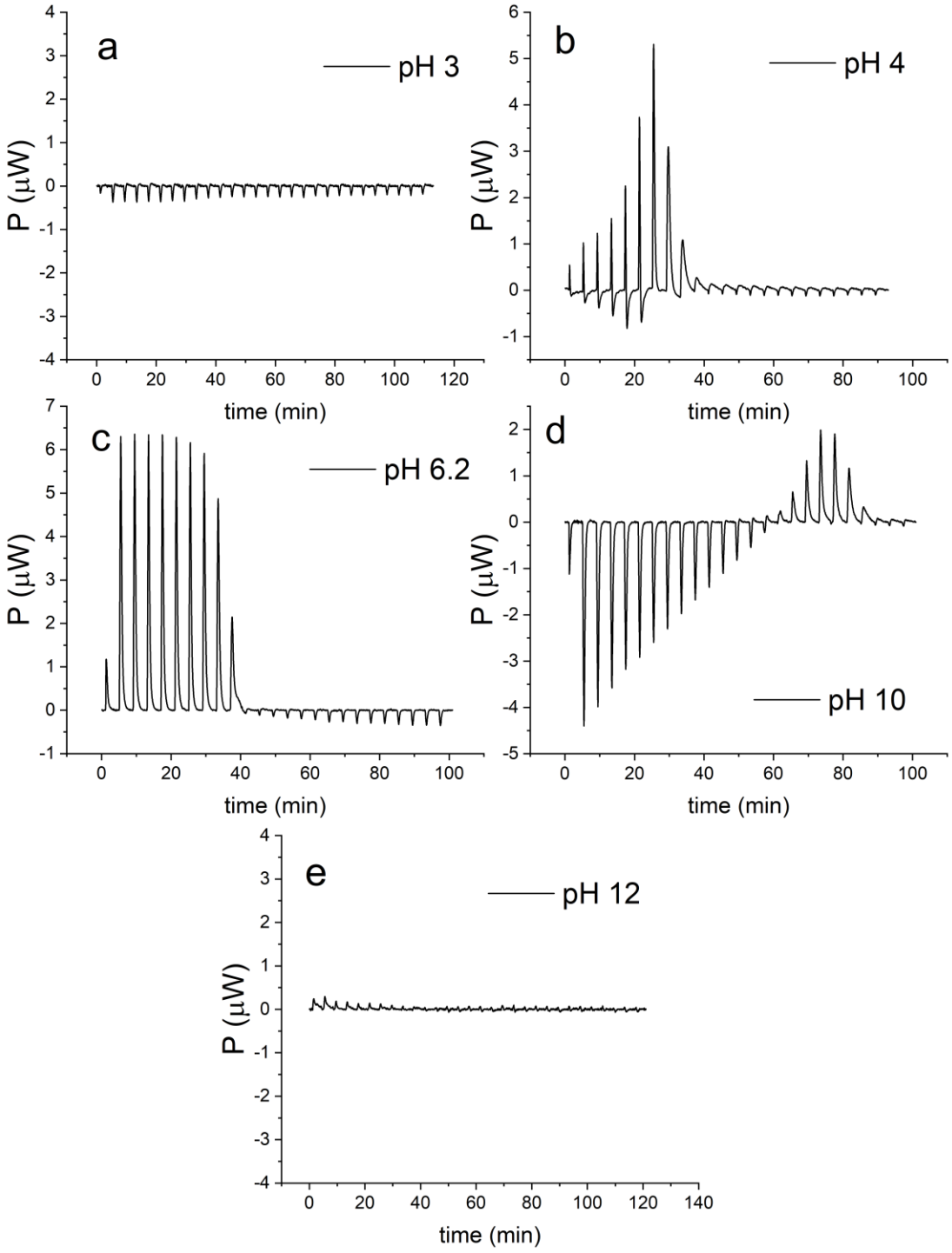


Figure S10. ITC titrations raw power signal P recorded during titrations representing the complexation of PGA with PLL in different pH. Ionic strength I = 0.01 M NaCl and temperature T = 298 K.

Molecular dynamics simulations

Table S3. Evolution of secondary structure as a function of the simulation time for PGA/PLL systems consisting of the fully charged PGA with PLL of different ionization degrees (IDs) in the molecular dynamics simulations. The secondary structures corresponding to α-helices, β-structure, and other are colored blue, red, and white, respectively. The x-axis is the simulation time (a single pixel is 2 ns), while the y-axis is the residue index.

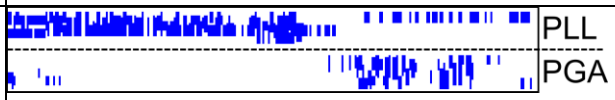
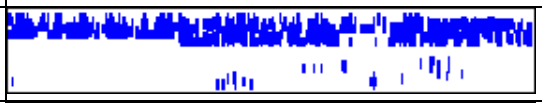
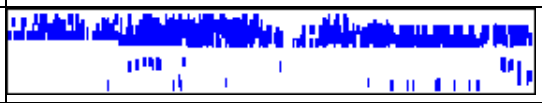
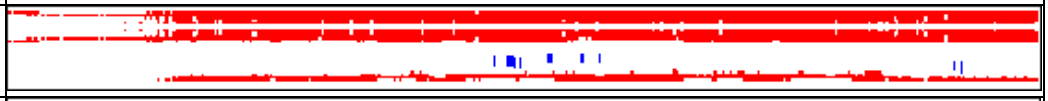
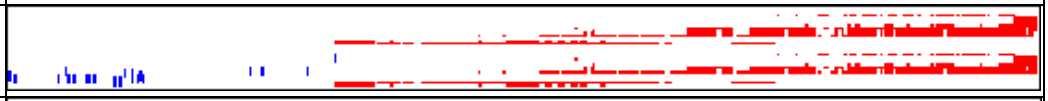
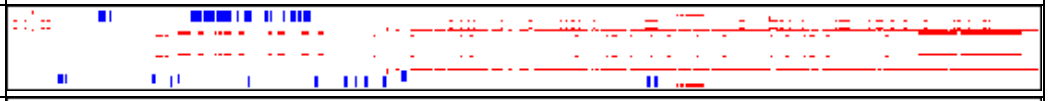
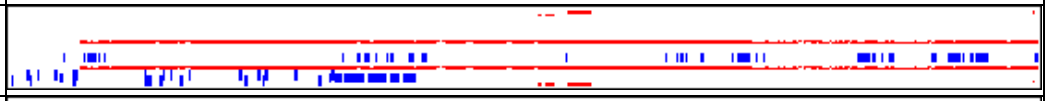
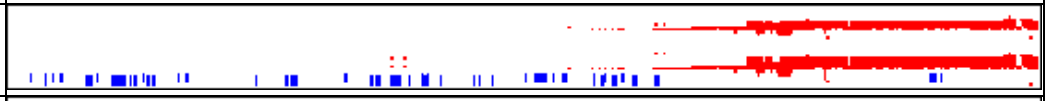
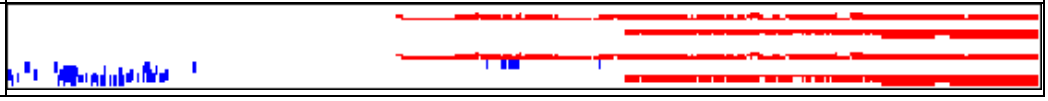
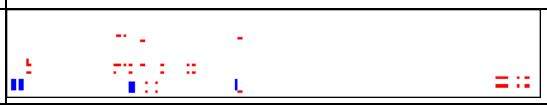
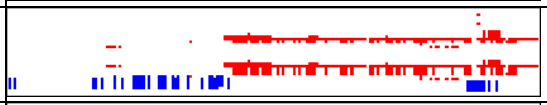
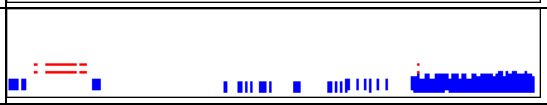
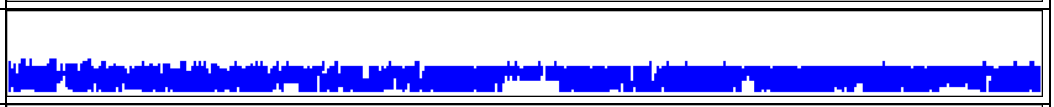
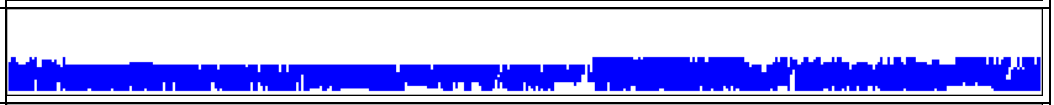
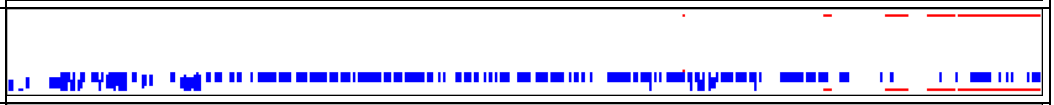
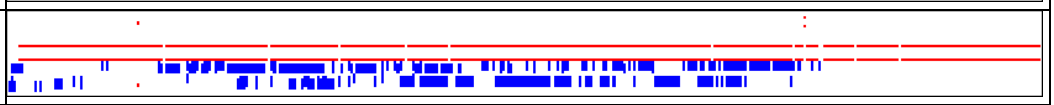
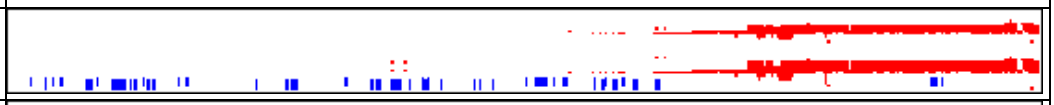
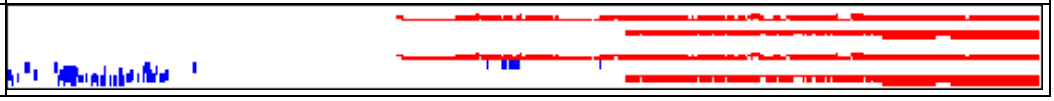
ID_{PLL}	Conf.	Evolution of secondary structure
0	1	
	2	
	3	
0.33	1	
	2	
	3	
0.66	1	
	2	
	3	
1	1	
	2	
	3	

Table S4. Evolution of secondary structure as a function of the simulation time for PGA/PLL systems consisting of the fully charged PLL with PGA of different ionization degrees (IDs) in the molecular dynamics simulations. The secondary structures corresponding to α-helices, β-structure, and other are colored blue, red, and white, respectively. The x-axis is the simulation time (a single pixel is 2 ns), while the y-axis is the residue index.

ID_{PGA}	Conf.	Evolution of secondary structure
0	1	
	2	
	3	
0.33	1	
	2	
	3	
0.66	1	
	2	
	3	
1	1	
	2	
	3	

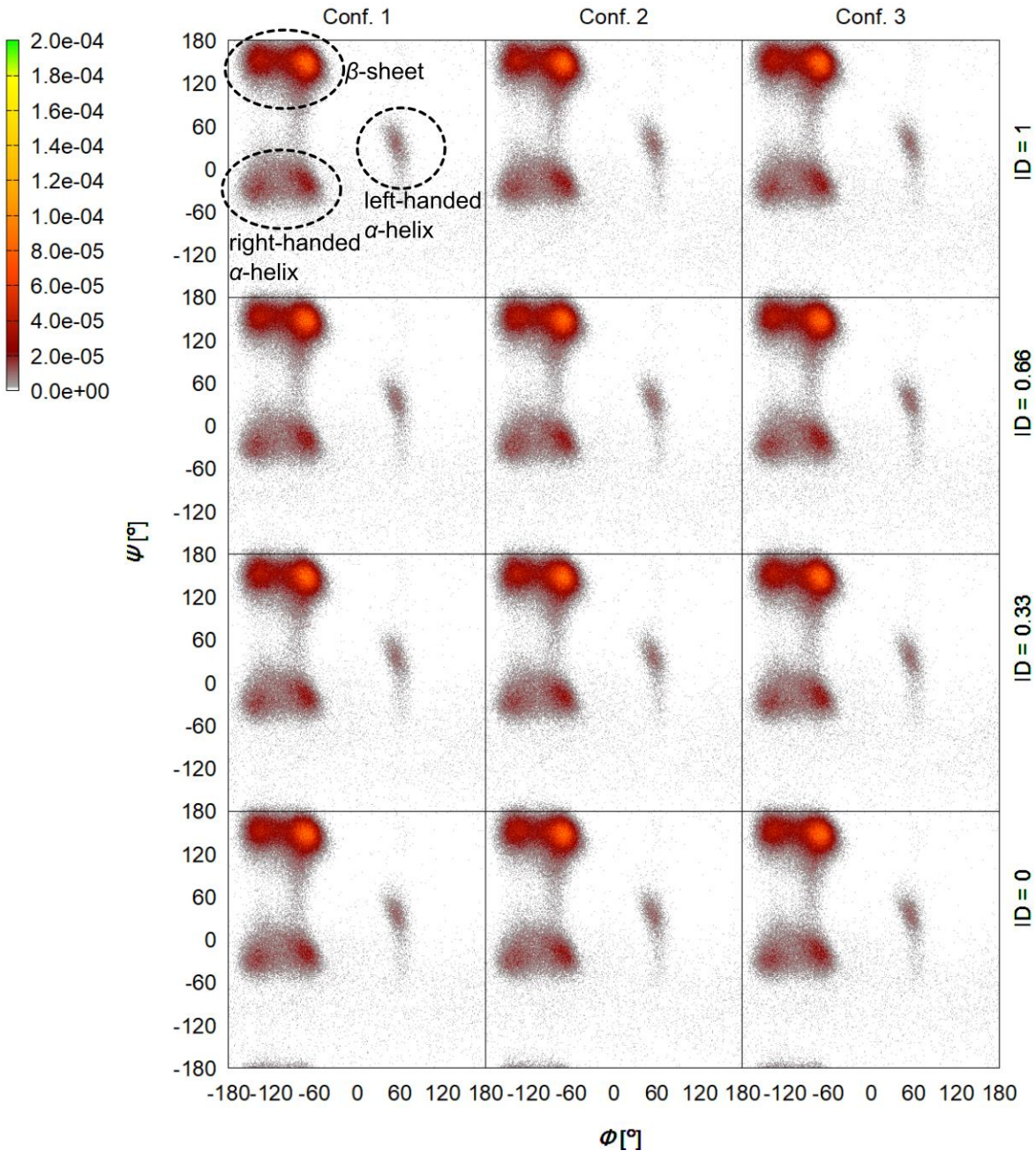


Figure S11. Ramachandran plots for the systems consisting of the fully charged PGA and PLL with different IDs in the molecular dynamics simulations. Data of three different initial configurations are presented. The regions corresponding to different secondary structures are highlighted in the top-left plot.

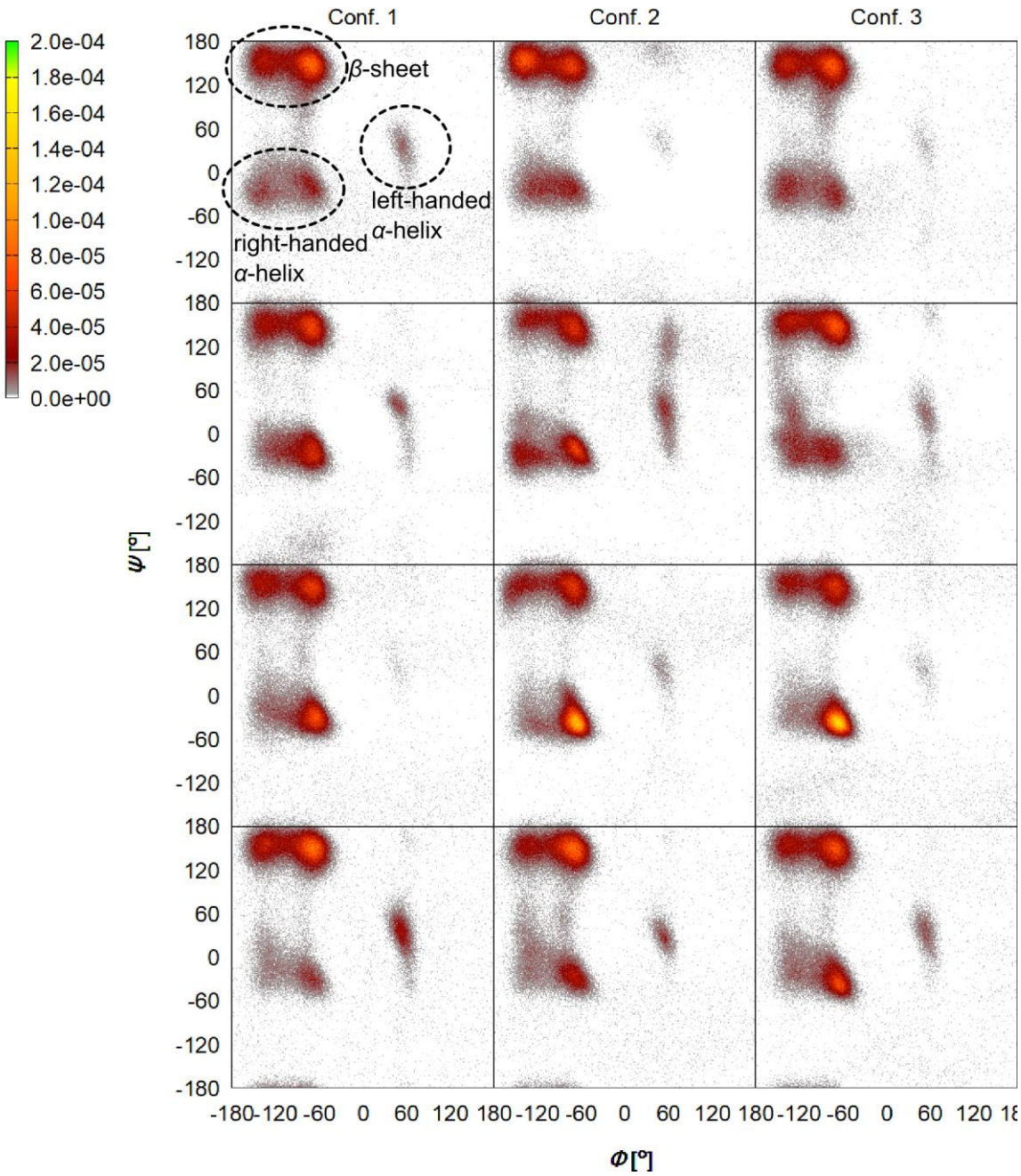


Figure S12. Ramachandran plots for the systems consisting of the fully charged PLL and PGA with different IDs in the molecular dynamics simulations. Data of three different initial configurations are presented. The regions corresponding to different secondary structures are highlighted in the top-left plot.

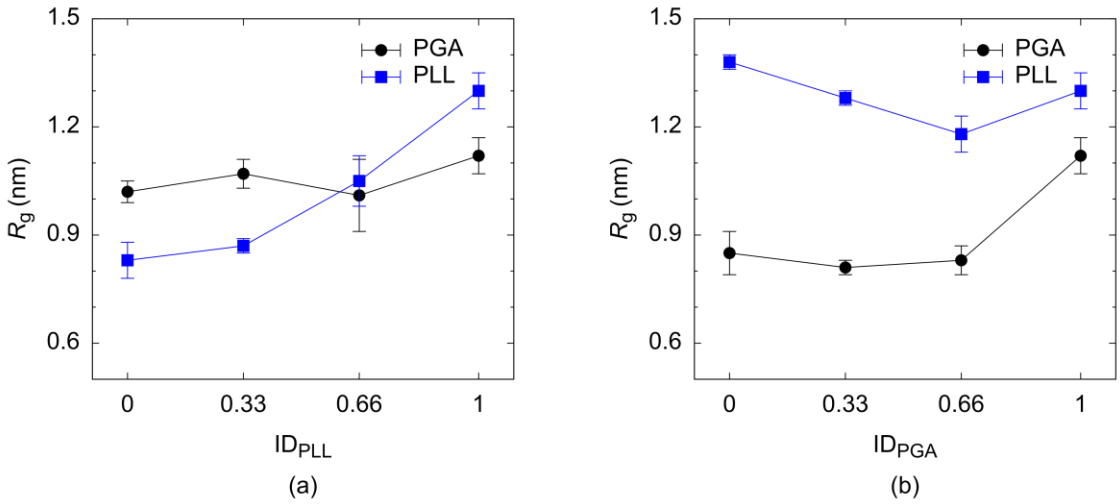


Figure S13. Radius of gyration as a function of the ID of a) PLL in the systems with the fully charged PGA and of b) PGA in the systems with the fully charged PLL in the molecular dynamics simulations.

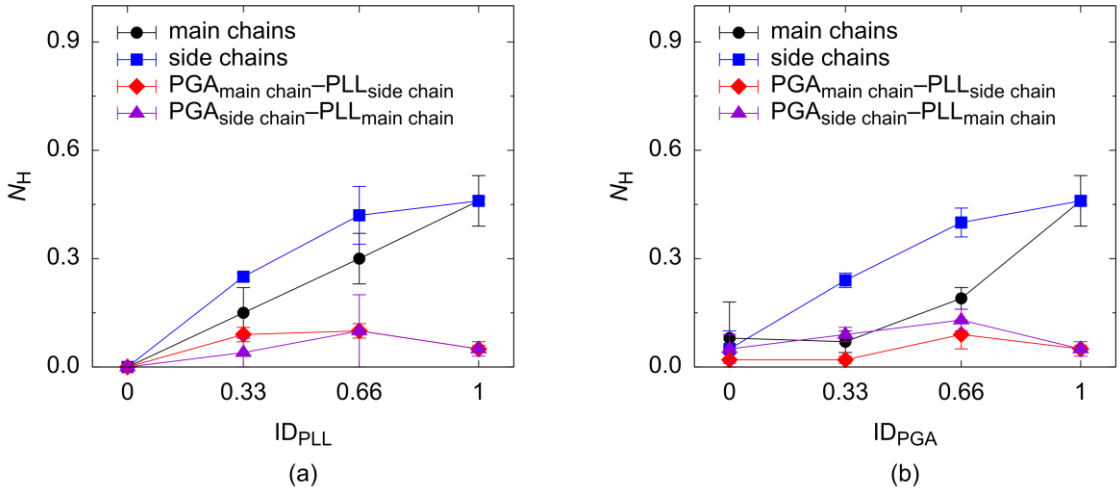


Figure S14. Number of different intermolecular hydrogen bonds (N_H) per single amino acid as a function of the ID of a) PLL in systems with fully charged PGA and (b) PGA in systems with fully charged PLL in the molecular dynamics simulations.

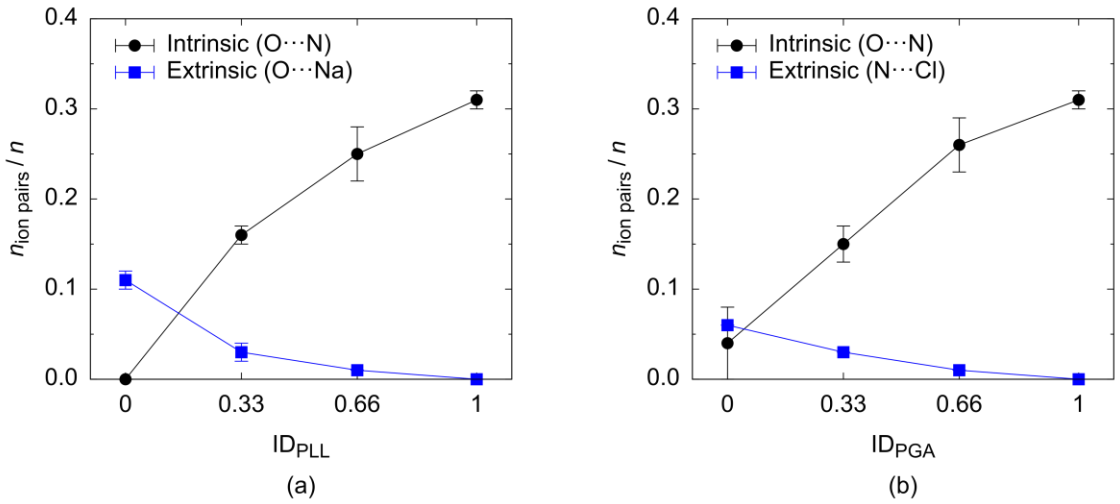


Figure S15. Number of the intrinsic and extrinsic ion pairs ($n_{\text{ion pairs}}$) per the number of repeating units in the peptide chain (n) as a function of the ionization degree ID of a) PLL in systems with fully charged PGA and b) PGA in systems with fully charged PLL in the molecular dynamics simulations.

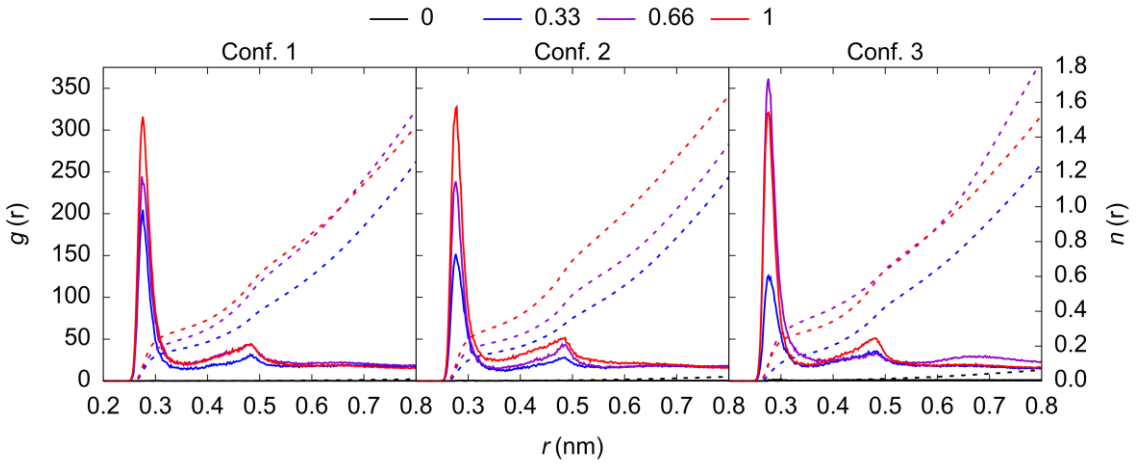


Figure S16. Radial distribution function g(r) (solid lines) and the corresponding cumulative number count n(r) (dashed lines) calculated between the charge groups of PGA and PLL in the systems consisting of the fully charged PGA and PLL with different IDs in the molecular dynamics simulations. The data corresponds to intrinsic ion pair O· · · N formation in the system.

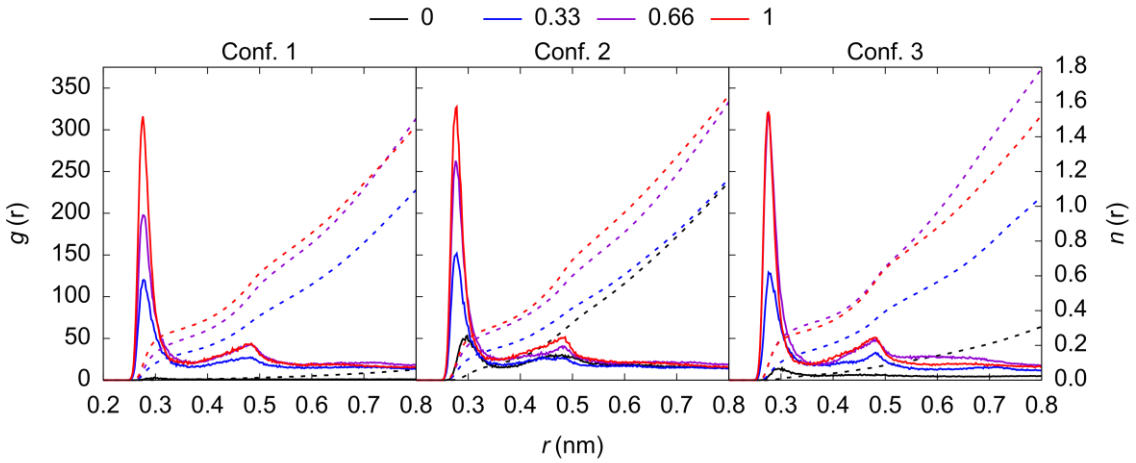


Figure S17. Radial distribution function *g*(r) (solid lines) and the corresponding cumulative number count *n*(r) (dashed lines) calculated between the charge groups of PGA and PLL in the systems consisting of the fully charged PLL and PGA with different IDs in the molecular dynamics simulations. The data corresponds to intrinsic ion pair O…N formation in the system.

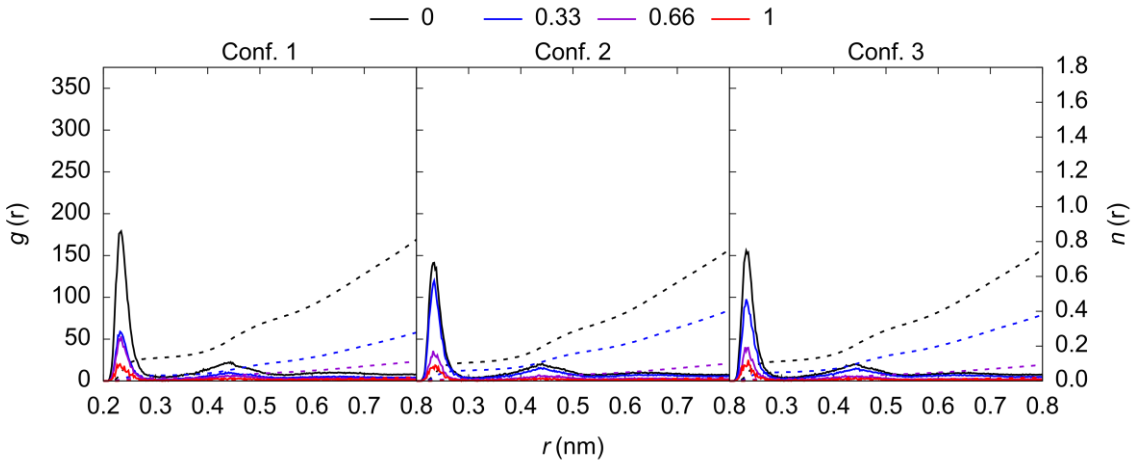


Figure S18. Radial distribution function *g*(r) (solid lines) and the corresponding cumulative number count *n*(r) (dashed lines) calculated between the oxygen in the charge groups of PGA (O^-) and the Na^+ counter ions in the systems consisting of the fully charged PGA and PLL with different IDs in the molecular dynamics simulations. The data corresponds to extrinsic ion pair formation in the system.

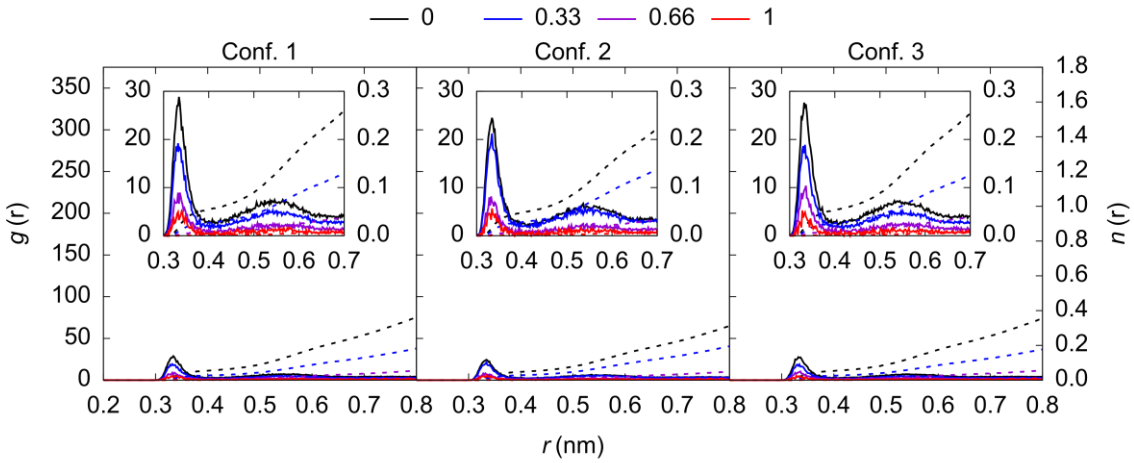


Figure S19. Radial distribution function *g*(r) (solid lines) and the corresponding cumulative number count *n*(r) (dashed lines) calculated between the nitrogen in the charge groups of PLL (N^+) and Cl^- counter ions in the systems consisting of the fully charged PLL and PGA with different IDs in the molecular dynamics simulations. The data corresponds to extrinsic ion pair formation in the system.