



UNIVERSITAT DE
BARCELONA

Coupling of binding and conformational equilibria in weak polyelectrolytes. Dynamics and charge regulation of biopolymers in crowded media.

Pablo Miguel Blanco

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

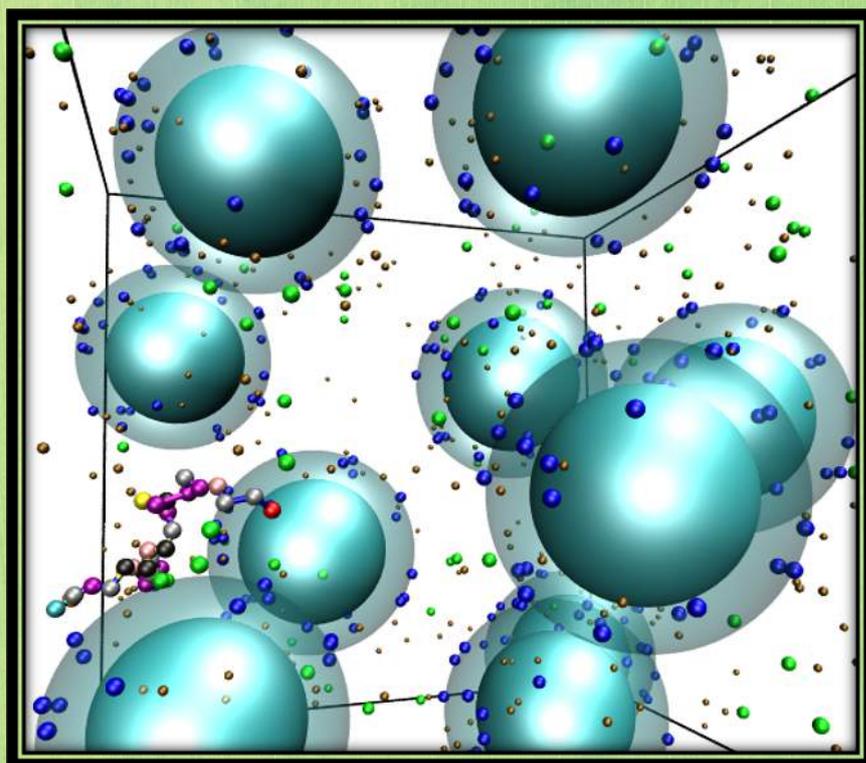
ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

Coupling of binding and conformational equilibria in weak polyelectrolytes. Dynamics and charge regulation of biopolymers in crowded media.

Doctoral Thesis by

Pablo M. Blanco



UNIVERSITAT DE
BARCELONA

UNIVERSITY OF BARCELONA

A thesis submitted in partial fulfillment for the degree of
Doctor of Philosophy in
Theoretical Chemistry and Computational Modelling
entitled

**Coupling of binding and conformational equilibria in weak polyelectrolytes.
Dynamics and charge regulation of biopolymers in crowded media.**

October 2020

Signed:

The Author:

MsC. Pablo Miguel Blanco
BioPhysChem Research Group
Theoretical and Computational Chemistry Inst.
Material Science and Physical Chemistry Dpt.
Faculty of Chemistry
University of Barcelona

The Director:

Dr. Sergio Madurga
BioPhysChem Research Group
Theoretical and Computational Chemistry Inst.
Material Science and Physical Chemistry Dpt.
Faculty of Chemistry
University of Barcelona

The Director:

Dr. Josep Lluís Garcés
BioPhysChem Research Group
AGROTECNIO Research Institute
Chemistry Department
Agricultural Engineering Technical School
Universitat de Lleida

The Tutor:

Prof. Francesc Mas
BioPhysChem Research Group
Theoretical and Computational Chemistry Inst.
Material Science and Physical Chemistry Dpt.
Faculty of Chemistry
University of Barcelona

"We march to victory, or we march to defeat, but we march forward, only forward."

King Stannis Baratheon, character in *A Dance with Dragons* (George R. R. Martin).

Declaration of Authorship

I, Pablo M. Blanco, declare that this thesis titled, "Coupling of binding and conformational equilibria in weak polyelectrolytes. Dynamics and charge regulation of biopolymers in crowded media" and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

Pablo M. Blanco

Abstract

Polymers are ubiquitous in our everyday lives, presenting many industrial applications and playing essential biological roles. For instance, DNA, proteins or polysaccharides (among many others) can be considered polymers, usually referred to as biopolymers. Those polymers containing a large number of charged groups are known as polyelectrolytes and they are frequently classified in terms of their ionization properties strong (completely ionized) and weak (partially ionized). In this thesis, I investigate different theoretical aspects of the conformational, ionization, elastic and diffusive properties of weak polyelectrolytes and biopolymers using theoretical and computational simulation techniques.

On the one hand, I focus in the complex interplay of the ionization and conformational degrees of freedom of weak polyelectrolytes using the recently developed Site Binding Rotational Isomeric State (SBRIS) model. A new analytical technique, the Local Effective Interaction Parameters (LEIP) method, is presented to solve the SBRIS model including long range intramolecular electrostatic interactions. The model is validated against constant pH Monte Carlo (MC) simulations. The SBRIS model is used to study the elastic response of a model weak polyelectrolyte. A new force regime is found for which (i) the force-extension curves are significantly dependent on the pH and the ionic strength values (ii) the polyelectrolyte charge is modified by the stretching force. The conformational and elastic properties of the polyelectrolyte are observed to be significantly affected by the presence of charge regulation.

On the other hand, I study the diffusive and ionization properties of biopolymers in solution with a high concentration of macromolecules. These conditions, known as macromolecular crowding, are characteristic of biological media where up to the 40% of the volume is occupied by macromolecules. The diffusion of two globular proteins, α -chymotrypsin and streptavidin, is analyzed in different crowding conditions with Brownian Dynamics simulations. A new coarse grained model is proposed, named as Chain Entanglement Softened Potential (CESP), which is found to reproduce quantitatively the experimental data. I have investigated the effect of macromolecular crowding in the binding and conformational properties of two Intrinsically Disordered Proteins (IDPs), histatin-5 and β -amyloid 42, which are modelled using a bead and spring model. The crowders are modelled using the CESP model mimicking Bovin Serum Albumin (BSA). The IDPs global charge is found to exhibit significant variations when neutral and charged crowders are added to the system.

Resumen

Los polímeros son omnipresentes en nuestras vidas diarias, presentes en múltiples aplicaciones industriales e involucrados en procesos biológicos esenciales. Por ejemplo, el ADN, las proteínas y los azúcares pueden ser considerados polímeros, normalmente denominados biopolímeros. Los polímeros con un gran número de grupos cargados se denominan polielectrolitos y son clasificados según su ionización como fuertes (totalmente ionizados) y débiles (parcialmente ionizados). En esta tesis, investigo diferentes aspectos teóricos de las propiedades conformacionales, elásticas, difusivas y de ionización de polielectrolitos débiles y biopolímeros.

Por una parte, me centro en la compleja interacción entre los grados de libertad de ionización y conformacionales de polielectrolitos débiles usando el modelo *Site Binding Rotational Isomeric State* (SBRIS), recientemente desarrollado. Se introduce una nueva técnica analítica, el método *Local Effective Interaction Parameters* (LEIP), para resolver el modelo SBRIS incluyendo las interacciones electrostáticas intramoleculares de largo alcance. El modelo se valida comparándolo con simulaciones Monte Carlo a pH constante. El modelo SBRIS se utiliza para estudiar la respuesta elástica de un polielectrolito modelo. Se encuentra un nuevo régimen de fuerza en el que (i) las curvas fuerza-extensión son significativamente dependientes del valor del pH y la fuerza iónica (ii) la carga del polielectrolito es modificada por la fuerza de estiramiento. Se observa que las propiedades conformacionales y elásticas del polielectrolito se ven significativamente afectadas por la regulación de la carga.

Por otra parte, evalúo las propiedades difusivas y de ionización de biopolímeros en solución con una elevada concentración de macromoléculas. Estas condiciones, conocidas como *crowding* macromolecular, son características de los medios biológicos donde hasta el 40% del volumen está ocupado por macromoléculas. Se analiza la difusión de dos proteínas globulares, α -chymotripsina y estreptavidina, en diferentes condiciones de *crowding* macromolecular con simulaciones de Dinámica Browniana. Un nuevo modelo de grano grueso es propuesto, nombrado *Chain Entanglement Softened Potential* (CESP), el cual reproduce cuantitativamente los datos experimentales. Se ha investigado el efecto de la *crowding* macromolecular en las propiedades conformacionales y de ionización de dos proteínas intrínsecamente desordenadas (IDPs), histidina-5 y β -amiloide 42, las cuales son modelizadas usando un modelo de cuentas y muelles. Se agrega una alta concentración de albumina de suero bovino (BSA), modelizada con el modelo CESP. Se encuentra que la carga global de las IDPs exhibe variaciones significativas cuando macromoléculas neutras y cargadas son añadidas al sistema.

Resum

Els polímers són omnipresents a la nostra vida diària, presents en múltiples aplicacions industrials i involucrats en processos biològics essencials. Per exemple, l'ADN, les proteïnes i els sucres poden ser considerats polímers, normalment denominats biopolímers. Aquells polímers amb un gran nombre de grups carregats són coneguts com a polielectròlits i freqüentment classificats en termes de la seva ionització com a forts (completament ionitzats) o febles (parcialment ionitzats). En aquesta tesi, investigo diferents aspectes teòrics de les propietats conformacionals, elàstiques, difusives i d'ionització de polielectròlits febles i biopolímers.

Per un costat, em centro en la complexa interacció entre els graus de llibertat d'ionització i conformacionals dels polielectròlits febles utilitzant el model *Site Binding Rotational Isomeric State* (SBRIS), desenvolupat recentment. S'introdueix una nova tècnica analítica, el mètode *Local Effective Interaction Parameters* (LEIP), per resoldre el model SBRIS incloent les interaccions electroestàtiques intramoleculares de llarg abast. El model es valida comparant els resultats amb la solució obtinguda per simulacions Monte Carlo a pH constant. El model SBRIS s'utilitza per estudiar la resposta elàstica d'un polielectròlit model. Es troba un nou règim de força en el que (i) les corbes força-extensió són significativament dependents del valor del pH i la força iònica (ii) la càrrega del polielectròlit es modificada per la acció de la força d'estirament. S'observa que les propietats conformacionals i elàstiques del polielectròlit es veuen significativament afectades per la presència de fluctuacions en la càrrega.

Per l'altre costat, avaluo les propietats difusives i d'ionització de biopolímers en solució amb una elevada concentració de macromolècules. Aquestes condicions, conegudes com *crowding* macromolecular, son característiques dels medis biològics on fins el 40% del volum es ocupat per macromolècules. S'analitza la difusió de dues proteïnes globulars, α -chymotripsina i estrep-tavidina, en diferents condicions de *crowding* macromolecular amb simulacions de Dinàmica Browniana. Es proposa un nou model de gra gruixut, anomenat *Chain Entanglement Softened Potential* (CESP), el qual es trobat que reproduïx quantitativament les dades experimentals. S'ha investigat l'efecte del *crowding* macromolecular en les propietats conformacionals i de protonació de dues proteïnes intrínsecament desordenades (IDPs), histidina-5 i β -amiloides 42, les quals són representades amb un model de perles i motlles. S'agrega una alta concentració d'albumina de sèrum boví (BSA), que es modelitzada utilitzant el model CESP. Es troba que la càrrega global de les IDPs exhibeix variacions significatives en afegir macromolècules neutres i carregades al sistema.

Acknowledgements

Throughout the development of this thesis, I have been fortunate enough to gather the advice, support and affection of many people. I will acknowledge them in the following lines as a small token of my appreciation and gratitude.

First, I would like to thank the people that supported me in the day-by-day work at the University of Barcelona. This, of course, includes all the actual and former members of the Bio-PhysChem research group. I specially thank my advisors, Drs. Sergio Madurga and Josep L. Garcés, and my tutor, Prof. Francesc Mas, for their guidance, for training me as a scientist and for everything you have taught me. Thank you Sergio for teaching me everything I know about programming. Thank you Josep for your patience and dedication to improve my scientific writing and language. Thank you Francesc for training me in academic teaching, your constant support and thoughtful discussions. I also want to thank my fellow PhD students Cristina Roncero, Juanjo Piñero, Luis Coronas, Raul Morales and the others for our enjoyable meals together and all the good memories that we have gathered this years. Specially thank you Cristina for all your constant emotional support and affection since we first met during our master studies.

My PhD studies have given me the opportunity of traveling abroad to several countries and even stay there for some time. I want to thank Dr. Rita S. Dias for welcoming me to her research group in Trondheim. Thank you Rita for your guidance in the final stages of my thesis and for the (still lasting) very nice coffee-break discussions. I also want to thank Bart van Blokland and Nataša Žuržul for their hospitality and affection to me during my stay in Norway. I also want to thank Dr. Claudio F. Narambuena for welcoming me to his research group in San Rafael. Thank you Claudio for your hospitality, for showing and joining me in traveling to many wonderful locations in Argentina, for a lot of thoughtful discussions and for your personal support. I also want to thank Emmanuel Sanchez, Marco Galat, Luciana Dauverne and Claudina Franchetti for their hospitality and friendship during my time there. Both have been wonderful experiences that have make grow up both as a scientist and as a person.

Por último, pero nunca menos importante, quiero agradecer todo el apoyo de mis familiares y amigos más cercanos. Agradezco a mis padres, Carlos Fco. Blanco y M^a Pilar Andrés, y a mi hermana, Sara P. Blanco, por respaldarme cada día y por todo lo que habéis hecho y hacéis por mí. También agradezco a mis tíos, Javier Andrés y Sara Pol, por todo su afecto y apoyo durante todos estos años. Especialmente le agradezco a mi abuela, Rosario Martino, la fe absoluta que siempre ha tenido en mí y su invaluable cariño. Quiero agradecerles a Víctor Fernández y a Jordi Tauste por nuestra amistad, por todos los buenos momentos que hemos pasado juntos en estos trece años y por todo el apoyo y cariño que me habéis dado siempre. También quiero

agradecer a Jordi Albiac y a Eudald Martínez que vinieran a visitarme a Trondheim, y el viaje que hicimos juntos por Noruega. Quiero también agradecerle a Anna Caparrós su apoyo durante el tiempo que he estado fuera y por todos los consejos que me ha dado estos últimos años. Finalmente, quiero agradecer a todos los "Crussaders", aunque no os nombre a todos explícitamente, por todo vuestro cariño y por todos los buenos momentos que hemos pasado juntos.

Abbreviations

AFM	A tomic F orce M icroscopy
BD	B rownian D ynamics
BSA	B ovin S erum A lbumin
ccMC	constant charge M onte C arlo
CESP	C hain E ntanglement S oftened P otential
CR	C harge R egulation
CF	C harge F luctuation
DNA	D eoxyribo N ucleic A cid
FJC	F reely J ointed C hain
FCS	F luorescence C orrelation S pectroscopy
FRAP	F luorescence R ecovery A fter P hotobleaching
FRC	F reely R otating C hain
HI	H ydrodynamic I nteractions
IDP	I ntrinsically D isordered P rotein
LEIP	L ocal E ffective I nteraction P arameters
LR	L ong- R ange
LPEI	L inear P oly E thylene I mine
MC	M onte C arlo
PEG	P oly(E thylene G lycol)
RIS	R otational I someric S tate
RPY	R otne- P raeger Y amakawa
SB	S ite B inding
SBRIS	S ite B inding R otational I someric S tate
SEV	S teric E xcluded V olume

SGCMC	Semi-Grand Canonical Monte Carlo
SR	Short-Range
WLC	Worm-Like Chain

Dedicated to the ninety anniversary of my grandmother Rosario Martino and to the memory of my grandfather Teófilo Andrés (1919-2008).

Contents

Declaration of Authorship	i
Abstract	iii
Resumen	v
Resum	vii
Acknowledgements	ix
Abbreviations	xi
1 Introduction	1
1.1 Polymer structure	1
1.1.1 Fundamental concepts	1
1.1.2 Freely Jointed Chain model	6
1.1.3 Freely Rotating Chain model	7
1.1.4 Worm-like Chain model	9
1.1.5 Rotational Isomeric State model	10
1.2 Polyelectrolyte ionization	14
1.2.1 Fundamental concepts	14
1.2.2 Site Binding model	18
1.3 Single-chain polymer elasticity	20
1.3.1 Fundamental concepts	20
1.3.2 Force regimes in the force-extension response	21
1.4 Macromolecular crowding	24
1.4.1 Fundamental concepts	24
1.4.2 Polymer diffusion in crowded media	25
1.5 Thesis outline	29
2 Methodology	35
2.1 The Site Binding Rotational Isomeric State model	35
2.2 Polymer models at different levels of description	38
2.3 Semi-Grand Canonical Monte Carlo (SGCMC) simulation	40
2.4 Brownian dynamics simulation	42

2.5	Software developed	44
2.5.1	Semi-Grand Canonical Monte Carlo software	44
2.5.2	Brownian Dynamics software	45
3	Coupling of Charge Regulation and Conformational Equilibria in Linear Weak Polyelectrolytes: Treatment of Long Range Interactions <i>via</i> Effective Short-Ranged and pH-Dependent Interaction Parameters	48
3.1	Summary	49
3.1.1	Introduction	49
3.1.2	Results	51
3.1.3	Conclusions	54
3.2	Publication	55
4	Effect of charge regulation and conformational equilibria in the stretching properties of weak polyelectrolytes	77
4.1	Summary	78
4.1.1	Introduction	78
4.1.2	Results	80
4.1.3	Conclusions	83
4.2	Publication	86
5	Role of Charge Regulation and Fluctuations in the Conformational and Mechanical Properties of Weak Flexible Polyelectrolytes	103
5.1	Summary	104
5.1.1	Introduction	104
5.1.2	Results	105
5.1.3	Conclusions	109
5.2	Publication	112
6	Brownian dynamics computational model of protein diffusion in crowded media with dextran macromolecules as obstacles	133
6.1	Summary	134
6.1.1	Introduction	134
6.1.2	Results	135
6.1.3	Conclusions	139
6.2	Publication	140
7	Macromolecular diffusion in crowded media beyond the hard-sphere model	157
7.1	Summary	158
7.1.1	Introduction	158
7.1.2	Results	160
7.1.3	Conclusions	163
7.2	Publication	165
8	Influence of Macromolecular Crowding in the Charge Regulation of Intrinsically Disordered Proteins	178
8.1	Summary	179

8.1.1	Introduction	179
8.1.2	Results	182
8.1.3	Conclusions	186
8.2	Publication	189
Conclusions		208
List of Publications		212
Appendix		213
A	Supporting information to "Coupling of charge regulation and conformational equilibria in linear weak polyelectrolytes: treatment of long range interactions <i>via</i> effective short-ranged and pH-dependent interaction parameters"	216
A.1	Matricial expression for the SBRIS partition function for a linear polyelectrolyte	216
A.2	Calculation of the mean square distance between two nodes of the chain	218
A.3	Minimal symmetric polyelectrolyte model	219
B	Supporting information to "Effect of charge regulation and conformational equilibria in the stretching properties of weak polyelectrolytes"	222
C	Supporting information to "Role of charge regulation and fluctuations in the conformational and mechanical properties of weak flexible polyelectrolytes"	232
D	Supporting information to "Macromolecular diffusion in crowded media beyond the hard-sphere model"	236
E	Supporting information to "Charge Regulation in Intrinsically Disordered Proteins due to Macromolecular Crowding"	240
E.1	Aminoacid titration curves at $c_s = 0.01$ M without crowding agents	241
E.2	Average charge variation of aspartic acid, histidine and lysine aminoacids for varying salt concentration	242
E.3	Charge regulation in a medium with neutral crowders and salt concentration $c_s = 0.01$ M	243
E.4	IDP global charge variation <i>vs.</i> excluded volume curves	244
E.5	Binding capacitance titrations curves per aminoacid type in a medium crowded with neutral crowders at $c_s = 0.1$ M	246
E.6	Best fit parameters to the scaling law (Eq. 13 in the main text) for the R_g <i>vs.</i> ϕ curves	247
E.7	Binding capacitance per aminoacid titration curves in a media with charged crowders at $c_s = 0.01$ M	248
E.8	Calculation of the effective ionic strength I' for the case with charged crowders	249
E.9	Comparison of $\Delta Q(I' = 0.05\text{M})$ and $\Delta Q'(\phi)$ <i>vs.</i> pH for the case with charged crowders	250

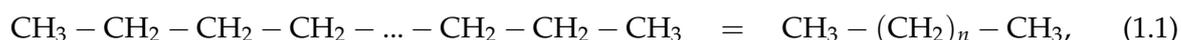
Chapter 1

Introduction

1.1 Polymer structure

1.1.1 Fundamental concepts

Polymers are molecules formed by a large number of atoms, *i.e.* they are macromolecules. The distinctive feature of polymers is their chemical structure, which is formed by repeating units known as monomers. For instance, the polymer with the simplest chemical structure is polyethylene,



which is formed by n repeating units of its monomer CH_2 . Polymers are named after this concept and the word polymer is combination of the greek terms *polus* πολύζ (many) and *meros* μέροςζ (part). Polymers are ubiquitous in our daily life due to their unique physical properties such as their toughness, high elasticity and viscoelasticity. Synthetic polymers (*e.g.* polyethylene, polystyrene, polyvinyl chloride) are commonly used as manufacturing materials. Natural polymers (*i.e.* biopolymers) such as DNA, proteins or polysaccharides are essential for living organisms and play key biological roles¹.

In solution, polymers frequently exhibit a characteristic dynamic structure². Polymer conformational properties are highly dependent on the medium conditions (solvent, salt concentration, temperature...) and chemical composition³. A particular spatial distribution of the polymer atoms is known as a conformation. In solution, polymers usually are constantly rearranging the relative position of its atoms which yields to a huge number of different conformations.

When the polymer does not have any preferred conformation is usually regarded as being in a random coil conformation.

Possible conformations are, however, restricted by the polymer inner chemical geometrical restrictions. For instance, one can consider the simple case of polyethylene chain, illustrated in Fig. 1.1. The monomer j has a covalent bond with its neighboring monomer $j + 1$. As a result, the distance between j and $j + 1$ will be almost constant and, on average, equal to the bond length l , with only very small thermal oscillations. This is also true for the angle formed by the monomers $j - 1$, j and $j + 1$ which also fluctuates around the bond angle θ . The relative position between groups separated by three covalent bonds (e.g. the monomers $j - 1$ and $j + 2$ in Fig. 1.1) is usually measured by its dihedral angle ϕ . ϕ is the angle between the plane formed by the first group and the central ones (yellow plane enclosing monomers $j - 1$, j and $j + 1$ in Fig. 1.1) and the plane formed by the last group and the central ones (blue plane enclosing monomers j , $j + 1$ and $j + 2$ in Fig. 1.1).

ϕ is usually dynamic due to thermal fluctuations and the polymer readily changes it by performing rotations around the axis formed by each bond of its chain. As can be observed in Fig. 1.2 (up), the polymer can change the relative position of the monomers $j - 1$ and $j + 2$ by performing a rotation around the axis defined by the central bond i . In the case depicted in Fig. 1.2 (up), this rotation changes the phi-value from $\phi = 0^\circ$ (*cis* conformation) to $\phi = 180^\circ$ (*trans* conformation).

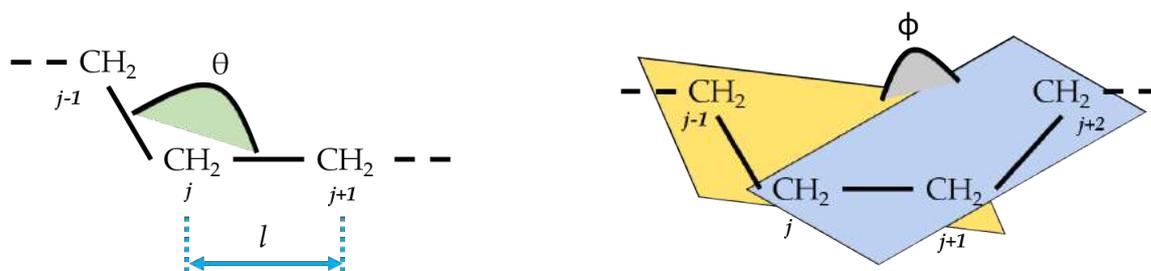


FIGURE 1.1: Scheme illustrating the concepts of bond length l , bond angle θ and dihedral angle ϕ for the particular case of a polyethylene chain. For a given monomer j , l is the equilibrium distance between j and its first neighbours $j \pm 1$. θ is the equilibrium value of the angle formed by $j - 1$, j and $j + 1$. ϕ is the angle between the planes formed by the central monomers and their first neighboring monomers $j - 1$ (orange plane) and $j + 2$ (blue plane).

The rotation of the dihedral angle is not free in general and it has an associated energetic cost (torsional energy) due to factors such as excluded volume, electrostatic interaction, Van der Waals forces or hydrogen bonds. An illustrative example of a torsional energy curve is shown in Fig. 1.2 (down) for a bond i of polyethylene. Typically, the conformational states with $\phi = 60, 180, 300$ are of minimal energy since the groups are alternated. These states are known as *trans* ($\phi = 180$), *gauche+* ($\phi = 60$) and *gauche-* ($\phi = 300$). Conversely, the conformations where the groups are eclipsed (*i.e.* $\phi = 0, 120, 240$) are usually states of maximum energy. Although the bond conformation strongly determines its torsional energy, it also depends on the conformation of the next (first, second,...) neighbouring bonds. In other words, bond conformations are not independent. This is specially dramatic in polymers, where the rotation of the dihedral angle can also be affected by long range intramolecular interactions (*e.g.* electrostatic interactions or hydrogen bonds) involving distant groups of the chain. As a result, the conformation of a given bond is not independent but highly coupled with the conformation of the other bonds in the chain.

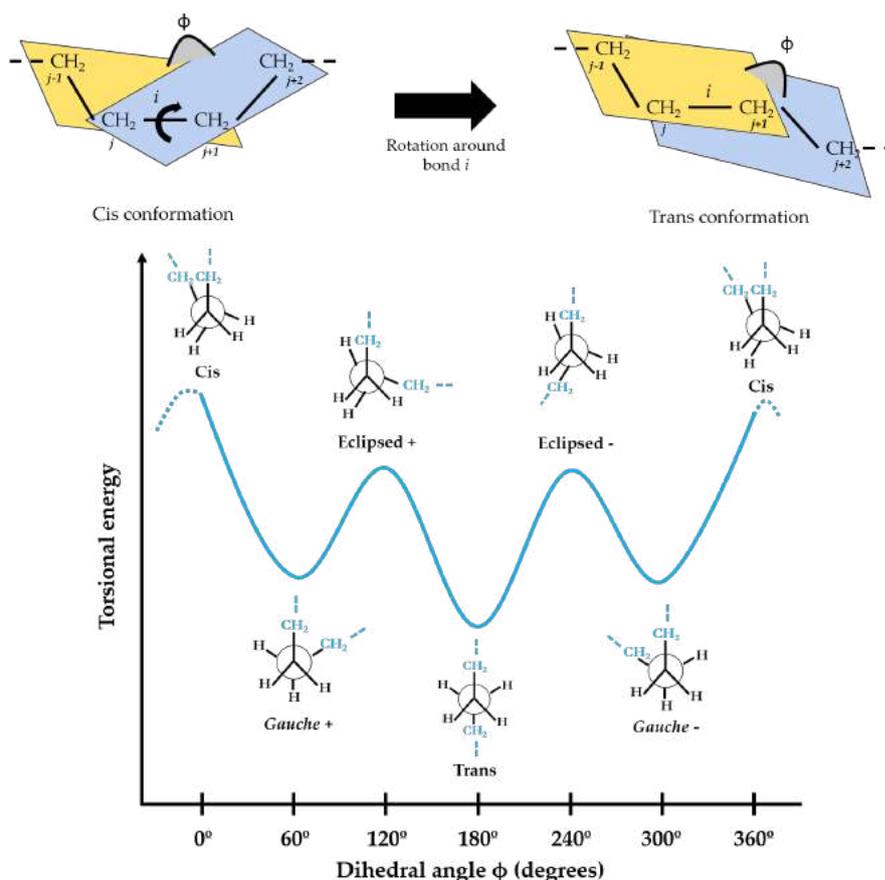


FIGURE 1.2: (Layout of polyethylene showing the torsional energy associated to the rotation around the axis defined by bond i . The rotation changes the dihedral angle ϕ value from 0 (*cis* conformation) to 180 (*trans* conformation). Typically, alternated conformations (*gauche* \pm and *trans*) are states of minimum energy whereas eclipsed conformations are maximum energy states.

There are several magnitudes which measure the conformational properties of polymers. The ones used throughout this thesis are illustrated in Fig. 1.3, again considering the simple example of polyethylene. The magnitudes of study are the following : the end-to-end distance \vec{r} , the radius of gyration R_g , the persistence length l_p and the contour length L_C .

The end-to-end distance \vec{r} measures the distance between the two ends of the polymer chain (See Fig. 1.3). Considering the general case of a polymer with M bonds in its principal chain, \vec{r} can be calculated as

$$\vec{r} = \sum_{i=1}^M \vec{l}_i. \quad (1.2)$$

where \vec{l}_i is a vector with the same direction of the bond i and magnitude equal to the bond length of i . However, the key magnitude that allows to quantify the polymer size is not \vec{r} but the thermal average of its second moment $\langle r^2 \rangle$. Considering Eq. 1.2, $\langle r^2 \rangle$ can be expressed as

$$\langle r^2 \rangle = \langle \vec{r} \cdot \vec{r} \rangle = \sum_{i,j}^M \langle \vec{l}_i \cdot \vec{l}_j \rangle = \sum_{i=1}^M \langle \vec{l}_i^2 \rangle + 2 \sum_{i=1}^M \sum_{j=i+1}^M \langle \vec{l}_i \cdot \vec{l}_j \rangle \quad (1.3)$$

The polymer size can also be quantified with its radius of gyration, R_g , defined as the root mean square distance of the monomers with its center-of-mass (C.M.)

$$R_g^2 = \frac{1}{M+1} \sum_{i=1}^M \langle s_i^2 \rangle \quad (1.4)$$

where s_i is the distance between the group i and the C.M. (See Fig. 1.3). According to a theorem due to Lagrange⁴, Eq. 1.4 can be rewritten as

$$R_g^2 = \frac{1}{(M+1)^2} \sum_{i < j}^M \langle r_{ij}^2 \rangle \quad (1.5)$$

where r_{ij} is the distance between the monomers i and j . Using Eq. 1.5, it can be proved that, for a long enough chain ($M \rightarrow \infty$) and identical bonds, $\langle r^2 \rangle$ and R_g^2 are directly linked by the relation

$$R_g^2 = \frac{\langle r^2 \rangle}{6}. \quad (1.6)$$

The latter equation is valid for any polymer chain only subject to local constrains related to the geometrical features of bond structure, and hindrances of rotation about bonds⁴. However, Eq. 1.6 is no longer applicable when long range intramolecular interactions are non-negligible (*e.g.* in presence of electrostatic interactions between charged groups in the polymer chain).

A useful quantity to measure the chain stiffness is the persistence length l_p , defined as the average sum of the projections of all the bonds $j \geq i$ on bond i in an infinitely long chain (See

Fig. 1.3). Considering the simple case of a polymer with bonds with an equal length l , l_p reads

$$l_p = \frac{1}{l} \sum_{i \leq j}^M \langle l_i \cdot l_j \rangle = l + \frac{1}{l} \sum_{j=i+1}^M \langle l_i \cdot l_j \rangle. \quad (1.7)$$

Comparing the last terms of Eqs. 1.3 and 1.7, one can relate $\langle r^2 \rangle$ and l_p . For a long enough chain ($M \rightarrow \infty$), both equations can be combined, which yields to

$$\langle r^2 \rangle = Ml^2 - 2Ml_p l - 2Ml^2 = Ml(2l_p - l) \quad (1.8)$$

or

$$l_p = \frac{\langle r^2 \rangle}{2Ml} + \frac{l}{2}. \quad (1.9)$$

For short polymer chains, end effects can be significant. As a result, the value of l_p becomes dependent of the first bond chosen in Eq. 1.7, thereby different for each bond of the chain.

The contour length L_C is defined as the polymer maximum length, *i.e.* the polymer length when its chain is totally extended (See Fig. 1.3). L_C and $\langle r^2 \rangle$ can be related using the expression

$$\langle r^2 \rangle = Mll_k = L_C l_k \quad (1.10)$$

where l_k defines the so-called Kuhn length of the polymer. l_k can be understood as the length of a "super bond" linking different nodes of the polymer chain. The resulting chain composed of the super bonds with length l_k behaves, in an effective way, as the Freely Jointed Chain model, which will be introduced in the next section. Eqs. 1.3 and 1.10 can be combined to prove that, for a long enough chain, l_k and l_p are linked by

$$l_k = 2l_p - l. \quad (1.11)$$

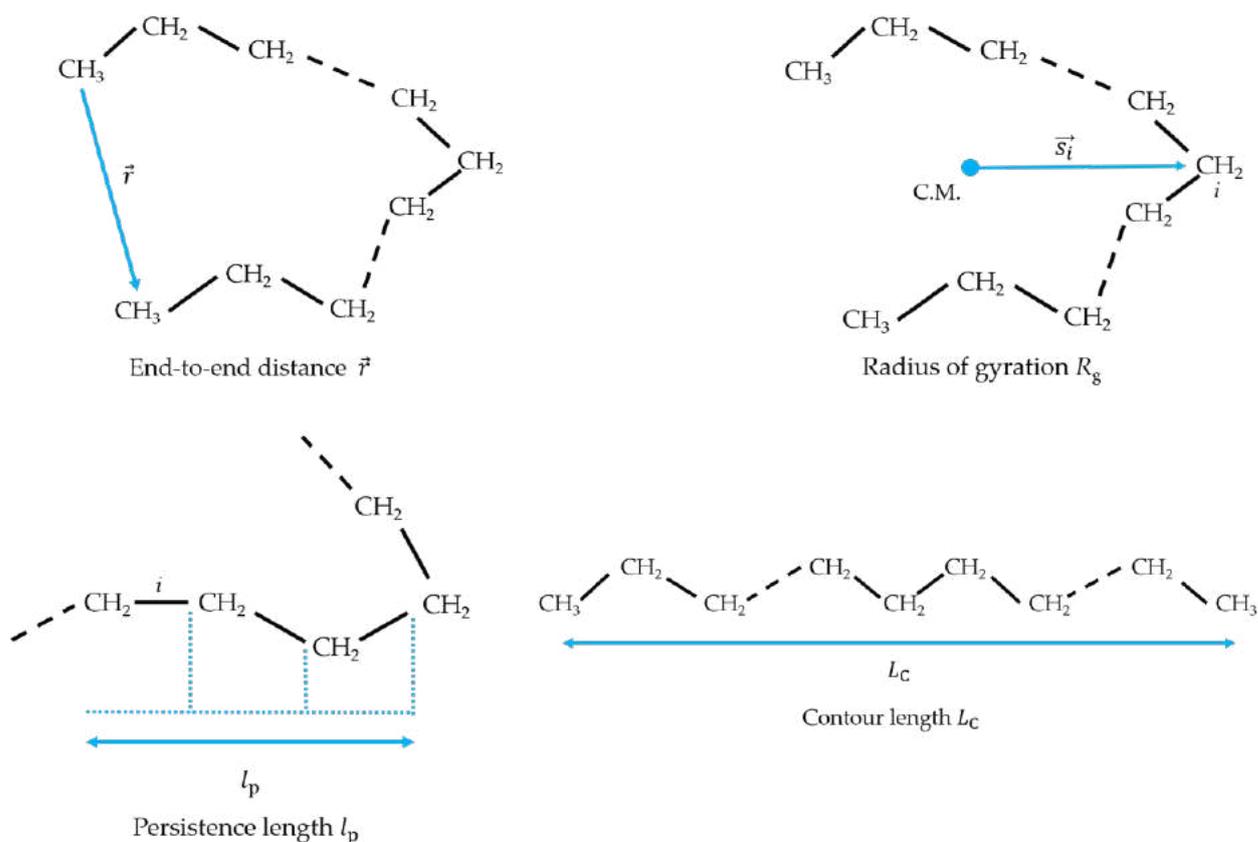


FIGURE 1.3: Outline of the conformational magnitudes used and/or calculated throughout this thesis: end-to-end distance \vec{r} , radius of gyration R_g , persistence length l_p and contour length L_C . \vec{r} measures the distance between the two chain ends. R_g measures the average distance between the polymer center-of-mass (C.M.) and the polymer monomers. For a given bond i , l_p is the length necessary to find a bond whose conformational state is uncorrelated with that of i . L_C is the polymer maximum length, *i.e.* the polymer length when its chain is totally extended.

1.1.2 Freely Jointed Chain model

The Freely Jointed Chain (FJC) model represents the polymer chain as a set of bonds with an equal length l_k , usually referred as Kuhn length. Each bond is represented as a vector \vec{l} which joins two consecutive nodes of the polymer (*i.e.* monomers or atoms), as outlined in Fig. 1.4. The direction of a given bond i , \vec{l}_i , is considered to be independent of the position of any other bond \vec{l}_j , *i.e.* all bonds are uncorrelated and

$$\langle \vec{l}_i \cdot \vec{l}_j \rangle = 0 \quad \forall \quad i \neq j. \quad (1.12)$$

This assumption implies that the directions of the bonds are fully uncorrelated. Moreover, the bonds can inter-cross/overlap each other (for this reason, the FJC chain is also usually referred to be a "phantom" chain).

As a result of considering that the bond directions are independent (Eq. 1.12), it follows from Eq. 1.8 that the end-to-end distance $\langle r^2 \rangle_{\text{FJC}}$ for a FJC chain of M bonds can be straightforwardly calculated as

$$\langle r^2 \rangle_{\text{FJC}} = Ml_k^2 = L_C l_k. \quad (1.13)$$

where $L_{C,\text{FJC}} = Ml_k$ is the contour length of a FJC. A natural consequence of Eq. 1.12 is that, for the FJC model, the persistence length l_p (Eq. 1.9) is equal to the bond length (Kuhn length)

$$l_{p,\text{FJC}} = l_k. \quad (1.14)$$

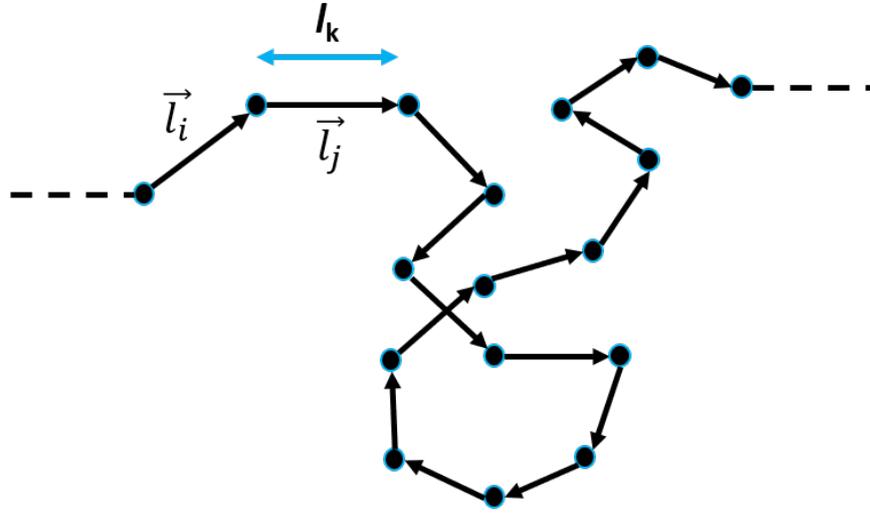


FIGURE 1.4: Scheme illustrating the Freely Jointed Chain (FJC) model. In this model, all the bonds are considered to have the same length l_k , named Kuhn length. The direction of bond i , \vec{l}_i , is considered to be uncorrelated from the directions of the other bonds, which yields to possible bond inter-crossing.

1.1.3 Freely Rotating Chain model

The Freely Rotating Chain (FRC) model improves the FJC and considers a chain that not only has bonds with a fixed length l but also with a fixed bond angle θ (See Fig. 1.5). In the simple case, the values of l and θ are considered to be constant and equal for all bonds. Moreover, the torsional energy of each bond i is constant and all the dihedral angle ϕ values are equally-probable (*i.e.* the rotation about i axis is free). However, unlike the FJC model, the directions between the bonds are correlated due to the geometrical restriction of fixing θ . As can be observed in 1.5, the rotation of the last bond is limited to the positions delimited by the blue circumference.

For a given bond $i + 1$, its projection on the preceding bond i is $l \cos(\pi - \theta)$, as can be observed in Fig. 1.5. In the same way, it is straightforward to find that the L_C of FRC with M bonds is given by

$$L_{C,\text{FRC}} = Ml \cos((\pi - \theta)/2). \quad (1.15)$$

Under the assumption of free rotation around i axis, the projections in directions perpendicular to the bond direction average to zero. It follows that

$$\langle \vec{l}_i \cdot \vec{l}_{i+k} \rangle = l^2 \alpha^k \quad (1.16)$$

where $\alpha = \cos(\pi - \theta)$. For a long enough chain of M bonds, from Eqs. 1.8 and 1.16 it follows that⁴

$$\langle r^2 \rangle_{\text{FRC}} = Ml^2 \frac{(1 + \alpha)}{(1 - \alpha)}. \quad (1.17)$$

Direct comparison of Eqs. 1.10 and 1.17 shows that the Kuhn length for a FRC $l_{k,\text{FRC}}$ is equal to

$$l_{k,\text{FRC}} = l \frac{(1 + \alpha)}{(1 - \alpha)}. \quad (1.18)$$

The persistence length of a FRC $l_{p,\text{FRC}}$ can be calculated as^{4,5}

$$l_{p,\text{FRC}} = \frac{l \cos((\pi - \theta)/2)}{|\ln(\alpha)|} \quad (1.19)$$

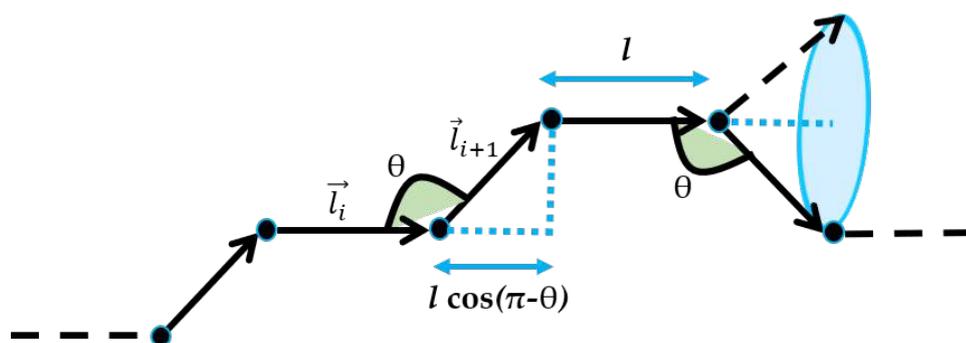


FIGURE 1.5: Outline of the Freely Rotating Chain (FRC) model. Analogously to the FJC chain, all bonds are considered to have the same bond length l . However, the FRC model goes one step further to the FJC chain and considers a fixed bond angle θ between the bond vector \vec{l}_i and its first neighbour \vec{l}_{i+1} .

1.1.4 Worm-like Chain model

The Worm-like Chain (WLC) model was firstly proposed by Porod and Kratky in 1949^{6,7}. In this model, the polymer chain is represented as a string with a continuous curvature, the direction at any point of the trajectory being random. It is particularly suited for describing the conformational properties of semi-flexible polymers (stiff but with a certain torsion modulus), although its use has not been restricted to them alone⁸.

The WLC model can be understood as the continuous curvature limit of the FRC model (See Fig. 1.6). In this limit, the polymer chain is indefinitely divided in M smaller fragments ($M \rightarrow \infty$), reducing the bond length $l' \rightarrow 0$ (and consequently increasing the bond angle value to $\theta' \rightarrow \pi$) but retaining the polymer maximum extension L_C

$$l_{C,WLC} = \lim_{l' \rightarrow 0, M \rightarrow \infty} Ml', \quad (1.20)$$

and persistence length l_p ,

$$l_{p,WLC} = \lim_{l' \rightarrow 0, \theta' \rightarrow \pi} \frac{l \cos((\pi - \theta)/2)}{|\ln(\alpha)|}, \quad (1.21)$$

as finite values, where $\alpha = \cos(\pi - \theta')$. In this limit, the $\langle r^2 \rangle$ for the FRC (Eq. 1.17) converges to

$$\langle r^2 \rangle_{WLC} = 2L_C l_p \left[1 - (l_p/L_C)(1 - e^{-L_C/l_p}) \right]. \quad (1.22)$$

which for an infinitely long chain ($M \rightarrow \infty$) yields to⁴

$$\langle r^2 \rangle_{WLC} = 2l_p L_C. \quad (1.23)$$

Direct comparison of Eqs. 1.13 and 1.23 shows that an infinitely long WLC has a Kuhn length $l_{k,WLC} = 2l_p$.

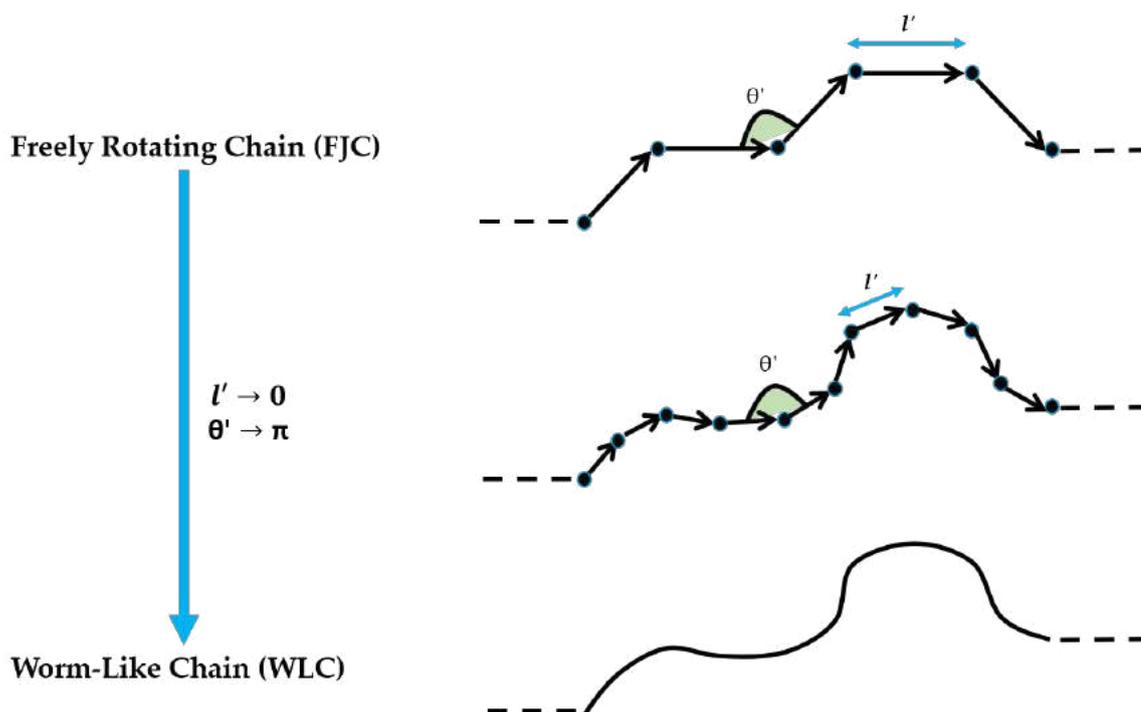


FIGURE 1.6: Layout illustrating the derivation of the Worm-Like Chain (WLC) model from the FRC model. The WLC can be understood as the continuous curvature limit of the FRC model, in which the bond length $l' \rightarrow 0$ and the bond angle $\theta' \rightarrow \pi$, but keeping L_C and l_p constant.

1.1.5 Rotational Isomeric State model

The Rotational Isomeric State (RIS) model was firstly proposed by Flory⁴ as a natural improvement of the FRC model. In the RIS model, the free rotation assumed in the FRC model is replaced by the torsional energy of the bonds. The basic assumption of RIS model is that the torsional energy can be introduced in a discrete way, considering only the more representative rotational states. In the simplest case, only the states of minimum energy are considered, which typically are *trans*, *gauche+* and *gauche-* states (See Fig.1.2). A particular conformational state can thus be defined by a set of variables $\{c_j\}$, $j = 3, \dots, M - 1$, where M is the number of the bonds and its torsional free energy reads

$$\mathcal{F}_{\text{RIS}}(\{c_j\}) = \sum_{j=3}^{M-1} \mathcal{F}(c_j) \quad (1.24)$$

where $\mathcal{F}(c_j)$ is the torsional energy of bond j in the rotational state c_j . Then, the conformational properties can be calculated applying the proper statistical weight to each conformational state and averaging over all the possible conformational states.

The RIS model can be illustrated using the simple examples of butane and pentane molecules. Butane molecule only has one dihedral angle ϕ , whose value changes when the central bond of

the molecule is rotated, as can be observed in Fig. 1.7. The statistically predominant rotational states of butane are those of minima energy: *trans*, *gauche+* and *gauche-* states. Following the RIS model the canonical partition function Z of butane reads⁴

$$Z(\text{butane}) = e^{-\beta\mathcal{F}(t)} + e^{-\beta\mathcal{F}(g^+)} + e^{-\beta\mathcal{F}(g^-)} = 1 + \sigma + \sigma \quad (1.25)$$

where $\sigma = \exp(-\beta\mathcal{F}(g))$. The latter equality in Eq. 1.25 considers the *trans* conformation as the energy origin. The energy of the two possible *gauche* conformations is equal due to the symmetry of the rotation.

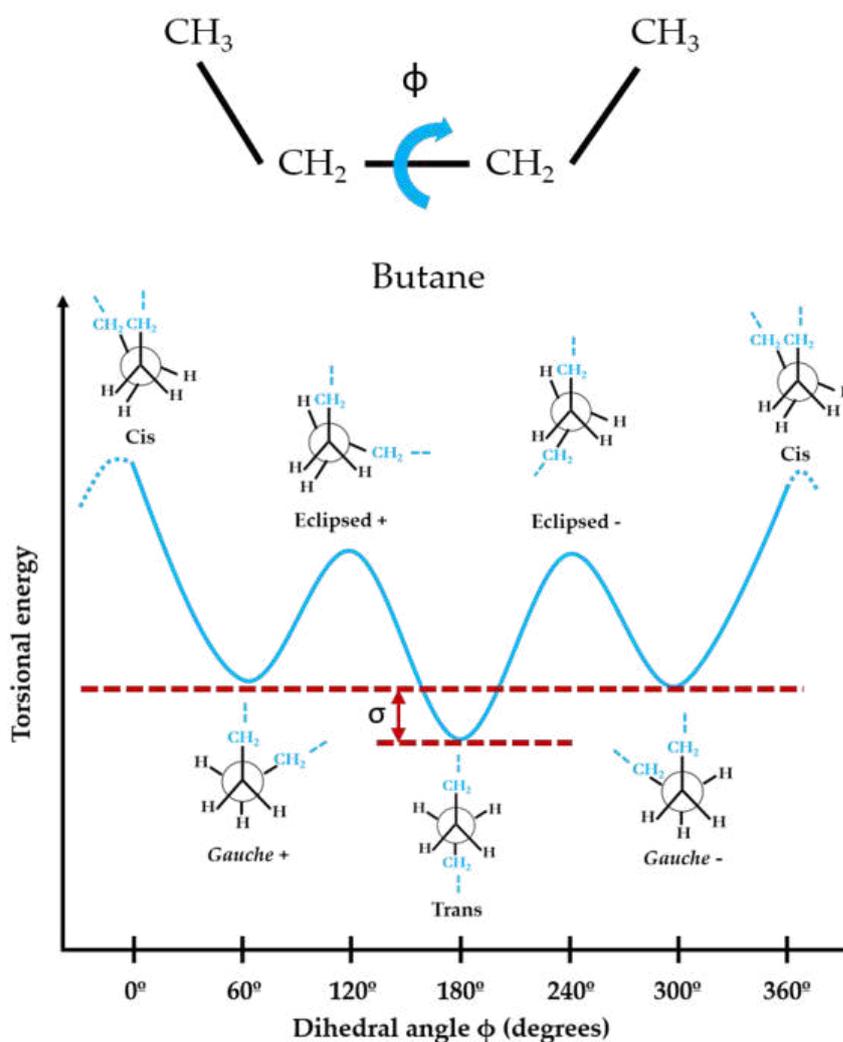


FIGURE 1.7: Qualitative layout of the butane torsional energy as a function of its single dihedral angle ϕ , highlighting the more characteristic rotational states. σ denotes the Boltzmann factor corresponding to the energy difference between the *gauche* states $\sigma = \exp(-\beta\mathcal{F}(g))$.

The pentane molecule has one methylene group more than butane, and consequently two dihedral angles are necessary, as can be observed in Fig. 1.8. Again following the RIS approximation, one can consider that each ϕ can only adopt *trans* and *gauche* \pm states. Then, the pentane Z can be calculated as

$$Z(\text{pentane}) = e^{-\beta\mathcal{F}(t,t)} + 2e^{-\beta\mathcal{F}(t,g\pm)} + 2e^{-\beta\mathcal{F}(g\pm,t)} + e^{-\beta\mathcal{F}(g+,g+)} + e^{-\beta\mathcal{F}(g-,g-)} + e^{-\beta\mathcal{F}(g-,g+)} + e^{-\beta\mathcal{F}(g+,g-)} = 1 + 4\sigma + 2\sigma^2\psi + 2\sigma^2\omega \quad (1.26)$$

where $\mathcal{F}(\phi_1, \phi_2)$ is the torsional energy of the conformational state corresponding to the dihedral angles ϕ_1 and ϕ_2 . The rotation around one bond is hindered by the excluded volume of the neighbouring groups, thereby the conformational states of ϕ_1 and ϕ_2 are no longer independent. The later equality in Eq. 1.26 is obtained by splitting the energy two contributions: (i) the torsional energy of a single bond $\mathcal{F}(\phi)$ in a ϕ conformation and (ii) the interaction energy $\mathcal{F}_I(\phi_1, \phi_2)$ between bonds 1 and 2 in the conformational states ϕ_1 and ϕ_2 , respectively. Then, for molecular symmetry, $\mathcal{F}(g+, g+) = \mathcal{F}(g-, g-) = 2\mathcal{F}(g) + \mathcal{F}_I(g+, g+)$ and $\mathcal{F}(g+, g-) = \mathcal{F}(g-, g+) = 2\mathcal{F}(g) + \mathcal{F}_I(g+, g-)$. The corresponding statistical weights of the interaction energies are $\psi = \exp(-\beta\mathcal{F}_I(g+, g+))$ and $\omega = \exp(-\beta\mathcal{F}_I(g+, g-))$. The chosen origin of energy is the conformation with the two rotating bonds in *trans*, $\mathcal{F}(t,t) = 0$.

In principle, one could compute Z for molecules with an increasing number of methylene groups following the same approach. In practice, the exponentially increasing number of possible conformational states requires a more efficient method than listing all the possible rotational states. The transfer matrix method, borrowed from classical statistical mechanics^{9,10}, allows to

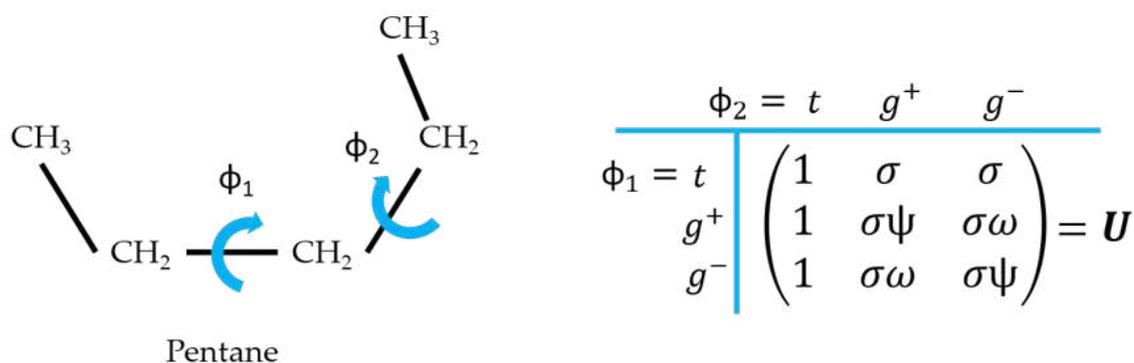


FIGURE 1.8: Example illustrating the transfer matrix \mathbf{U} for the pentane molecule, which has two dihedral angles ϕ_1 and ϕ_2 . Each element U_{ij} contains the statistical weight corresponding to the energy increment of the system in adding a bond in the state j given that previous bond is in the state i where $i, j = \text{trans}, \text{gauche+}$ and gauche- . $\sigma = \exp(-\beta\mathcal{F}(g))$, $\psi = \exp(-\beta\mathcal{F}_I(g+, g+))$ and $\omega = \exp(-\beta\mathcal{F}_I(g+, g-))$ are the statistical weight corresponding to *gauche* state torsional energy $\mathcal{F}(g)$ and the interaction energies $\mathcal{F}_I(g+, g+)$ and $\mathcal{F}_I(g+, g-)$ of having two neighboring bonds in the same/opposite *gauche* state, respectively.

readily obtain the Z of polymer linear chain with M groups if Z for the same polymer but with $M - 1$ groups is known.

For instance, following with the previous examples, $Z(\text{pentane})$ can be directly calculated from $Z(\text{butane})$ using the transfer matrix framework as

$$Z(\text{pentane}) = \vec{Z}(\text{butane})\mathbf{U}\vec{E} = 1 + 4\sigma + 2\sigma^2\psi + 2\sigma^2\omega \quad (1.27)$$

where $\vec{E} = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^T$ is the ending vector, which sums over all the possible final terms, and $\vec{Z}(\text{butane}) = \begin{pmatrix} 1 & \sigma & \sigma \end{pmatrix}$ is a vector whose elements are each of the terms of the butane Z (Eq. 1.25). \mathbf{U} is the transfer matrix, which in this case, reads

$$\mathbf{U} = \begin{pmatrix} 1 & \sigma & \sigma \\ 1 & \sigma\psi & \sigma\omega \\ 1 & \sigma\omega & \sigma\psi \end{pmatrix}. \quad (1.28)$$

Each element \mathbf{U}_{ij} contains the statistical weight corresponding to the energy increment when a new bond in the state j is "added" to the system given that the previous bond is in the state i , as outlined in Fig. 1.8. For instance, the element $\mathbf{U}_{23} = \sigma\omega$ corresponds to the energy variation associated to add a new bond in *gauche-*, which has a torsional energy of $\sigma = \exp(-\beta\mathcal{F}(\text{g}))$, given that the previous bond is in *gauche+*. The Boltzmann factor associated to the interaction energy is $\omega = \exp(-\beta\mathcal{F}_1(\text{g+}, \text{g-}))$.

The transfer matrix method allows to compute Z of linear molecules with any number of methylene groups (butane, pentane, hexane, heptane...). Considering only first neighbour correlations, the Z of a polyethylene chain with M bonds in its backbone reads⁴

$$Z = \vec{I}\mathbf{U}^{M-2}\vec{E} \quad (1.29)$$

where $\vec{I} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}$ and the origin of energy has been settled to the conformation of maximum extension, *i.e.* all bonds in trans conformation. By replacing $M = 3$ and $M = 4$ in Eq. 1.29, one can easily calculate $Z(\text{butane})$ (Eq.1.25) and $Z(\text{pentane})$ (Eq. 1.26), respectively. For a sufficiently large number of methylene groups, one obtains the polyethylene polymer chain (Eq. 1.1).

Despite here introduced for the case of polyethylene, the definition of \mathbf{U} in the form of Eq. 1.28 is valid for any symmetric aliphatic carbon chain with rotational states equal to trans, *gauche+*, and *gauche-*. For asymmetric chains, proper transfer matrix can also be defined. In fact, RIS model can be applied to a wide variety of linear polymers such as polyoxyethylene, polyethyleneimine or polypropylene^{4,11}. The inclusion of polymers with monomers composed

of more than one chemical group or with tacticity (in the case of asymmetric chains) needs of to the definition of more than one transfer matrix per monomer. In these cases, the product in Eq. 1.27 must be generalized, so that

$$Z = \vec{l} \prod_{i=1}^{M-2} \mathbf{U}_i \vec{E}. \quad (1.30)$$

Moreover, the RIS model is not restricted to only consider three rotational states but it can take into account any number of rotational states.

In general, the expression for the RIS model for the $\langle r^2 \rangle$ of a polymer chain with M bonds reads⁴

$$\langle r^2 \rangle (\text{RIS}) = \sum_{i=1}^M l_i^2 + 2Z^{-1} \vec{l} \sum_{i<j}^M \mathbf{U}_1^{i-1} (\mathbf{E}_\nu \times \vec{l}_i^T) \times [(\mathbf{U} \times \mathbf{E}_3) \|\mathbf{T}\|_i^{j-i} (\mathbf{E}_\nu \times \vec{l}_j) \mathbf{U}_j^{M-j} \vec{E} \quad (1.31)$$

where \mathbf{E}_ν is the unit matrix of size ν , \vec{l}_i is the vector pointing in the same direction than the bond i with magnitude equal to the bond length l_i and \times is the direct product of the matrices. \mathbf{T}_i is a transformation matrix which which transforms the coordinates of \vec{l}_{i+1} to the reference frame of bond \vec{l}_i and it is given by

$$\mathbf{T}_i = \begin{pmatrix} \cos(\pi - \theta_i) & \sin(\pi - \theta_i) & 0 \\ \sin(\pi - \theta_i) \cos(\phi_i) & -\cos(\theta_i) \cos(\phi_i) & \sin(\phi_i) \\ \sin(\pi - \theta_i) \sin(\phi_i) & -\cos(\pi - \theta_i) \sin(\phi_i) & -\cos(\phi_i) \end{pmatrix}, \quad (1.32)$$

where θ_i is the bond angle between bonds i and $i + 1$ and ϕ_i is the dihedral angle governing the rotation around bond i . The radius of gyration and the persistence length can also be calculated using the RIS model using expressions similar to Eq. 1.31 that can be found in the Chapter IV of Ref.⁴.

1.2 Polyelectrolyte ionization

1.2.1 Fundamental concepts

Within polymer macromolecules, the ones with monomers holding charged groups, commonly known as polyelectrolytes, exhibit unique technological applications and biological roles very different from their neutral counterparts¹². The unique features of polyelectrolytes make them excellent as glues, binders, thickeners, additives to ceramic glaze and as stabilizing agents in inks, photography, and lithographic processes¹³. Polyelectrolyte conformational properties are greatly influenced by the long-ranged electrostatic generated by the polyelectrolyte charged

groups^{14–17}. It is important to note that these electrostatic interactions can be intramolecular (between the charged groups in the polyelectrolyte chain) and intermolecular (between the charged polyelectrolyte groups and the other charged species in solution). Moreover, the large number of charged groups in polyelectrolytes produces unique electrostatic-based phenomena not observed in small ions. For instance, polyelectrolytes attract a significant fraction of small counter-ions in their nearby, phenomenon known as counterion condensation¹⁸.

In turn, polyelectrolytes can be classified in terms of the dissociation of their acid/base groups in solution. In an analogous way to strong acids/bases, strong polyelectrolytes completely dissociate in solution. An illustrative example is the ionization of a polyanion composed of monomers with a strong acid groups, as depicted in Fig. 1.9. When dissolved, each acid group of the polyanion dissociates following the chemical equation



where HA/A⁻ are the protonated/unprotonated forms of the acid group and H⁺ is a proton. For a strong polyelectrolyte, this chemical reaction is fully shifted to its right side; so the polyanion is always completely dissociated, at least at reasonable pH values. Several examples can be found of polymers that are strong polyelectrolytes, both from synthetic (*e.g.* polyDADMAC) and natural (*e.g.* DNA) sources.

In a similar way as weak acids/bases, weak polyelectrolytes are only partially dissociated in solution¹⁹. Their ionization is dynamic and highly depending on factors such as the temperature, the concentration of binding species (*e.g.* protons or metals), salt concentration or external fields^{20–26}. For instance, in the case of a weak polyanion in aqueous solution, its weak acid

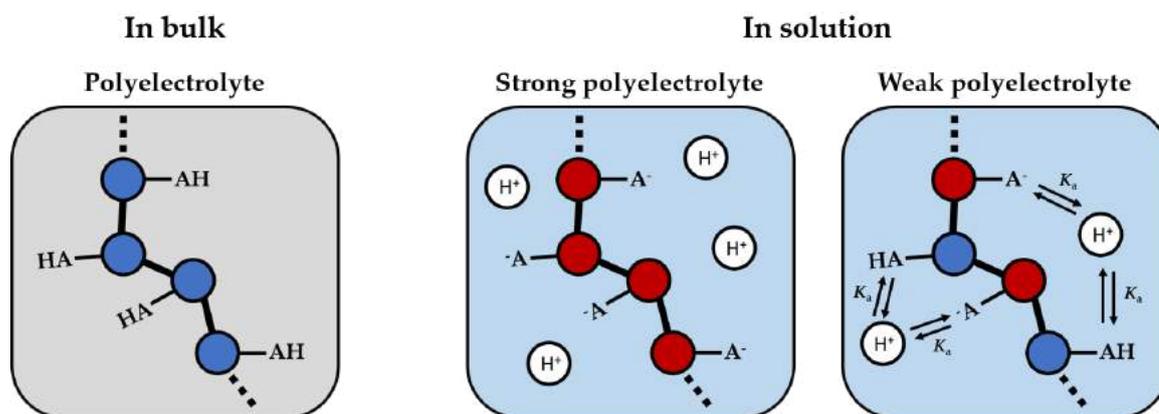
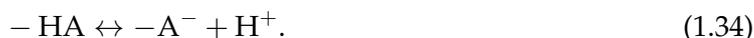


FIGURE 1.9: Outline of the two different ionization behaviours of polyelectrolytes in solution, illustrated for the case of a polyacid. Strong polyelectrolytes completely dissociate in solution while weak polyelectrolyte only have a partial dissociation, holding both uncharged (blue) and charged (red) acid/base groups. The acid/base groups of weak polyelectrolytes have a dynamic binding equilibrium, which is governed by the acid equilibrium constant K_a .

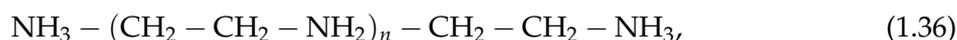
groups dynamically bind/unbind protons (See Fig. 1.9) following the chemical equilibrium



In the absence of interactions, the equilibrium is governed by its intrinsic equilibrium acid constant, K_a , which can be calculated as

$$K_a = \frac{a_{\text{H}} a_{\text{A}}}{a_{\text{HA}}} = K^{-1}. \quad (1.35)$$

where a_{HA} , a_{H} and a_{A} are the the chemical activities of the species HA, H^+ and A^- , respectively. Note that K_a is the inverse of the proton binding constant, K . Both equilibrium constants are usually expressed in logarithmic units $\log(K) = -\log(K_a) = \text{p}K_a$ to be compared with the pH-value, $\text{pH} = -\log(a_{\text{H}})$. An example of a synthetic weak polyelectrolyte is linear poly(ethyleneimine),



and many examples can be also found in biopolymers belonging to proteins or polysaccharide families.

The dynamic nature of the binding equilibrium implies that weak polyelectrolytes can modulate its charge in response to external physicochemical perturbations such as pH variations, changes in the salt concentration, interaction with other macromolecules in solution or external fields. This phenomenon is known as Charge Regulation^{12,27}, and it is ubiquitous in a wide range of processes such as protein-protein²⁸ and protein-surface interactions²⁹, the stability of colloidal systems and nanoparticle coating^{30,31} and receptor-ligand interactions in biochemical systems³² among many others.

For weak polyelectrolytes conformational and binding degrees of freedom are highly coupled: a change in the pH-value alters the polyelectrolyte charge which in turn can induce dramatic structural changes in the macromolecule^{16,17}. A classical example of this phenomenon is the swelling of poly(methacrylic) acid in a very narrow range of pH³³. The coupling between ionization and conformation is fundamental in biopolymers. For instance, it can be observed in the helix-coil transitions of poly(peptides)³⁴ or in the strong influence of ionization in the folding of proteins³⁵. Recently, the ionization and configuration coupling have been found to be specially important in proteins with disordered regions or entirely disordered, known as Intrinsically Disordered Proteins (IDPs)^{36,37}.

The ionization of a weak polyelectrolyte is typically characterized by measuring the polyelectrolyte degree of ionization θ . For a polyelectrolyte with N ionizable groups, θ quantifies the

ratio of charged groups of the polyelectrolyte and it reads

$$\theta = \frac{\langle N_C \rangle}{N}, \quad (1.37)$$

where $\langle N_C \rangle$ is the average number of charged groups. In the ideal case where electrostatic interactions are absent, the corresponding ideal degree of ionization θ^{ID} can be calculated from the well-known Henderson-Hasselbalch equation. For the simple case of a polyacid composed by identical binding sites, θ^{ID} reads

$$\theta^{\text{ID}} = \frac{c_{\text{A}^-}}{c_{\text{HA}} + c_{\text{A}^-}} = \frac{10^{\text{p}K_{\text{a}} - \text{pH}}}{1 + 10^{\text{p}K_{\text{a}} - \text{pH}}}. \quad (1.38)$$

where c_i ($i = \text{A}^-, \text{HA}$) is the "concentration" of i . For the case of a polyacid with N different independent binding sites, Eq. 1.38 can be extended summing the contribution of each site as

$$\theta^{\text{ID}} = \frac{1}{N} \sum_{i=1}^N \frac{10^{\text{p}K_{\text{a},i} - \text{pH}}}{1 + 10^{\text{p}K_{\text{a},i} - \text{pH}}}. \quad (1.39)$$

where $\text{p}K_{\text{a},i}$ is the acid equilibrium constant of the site of type i . In general, electrostatic interactions, both intermolecular and specially intramolecular, can not be neglected and the value of θ is different from the ideal value θ^{ID} .

Those deviations from ideality can be quantified in terms of the effective protonation constant K_c , which estimates the average affinity of the macromolecular sites for the protons. K_c can be calculated generalizing the Henderson-Hasselbalch as

$$\log K_c(\text{pH}, I, \dots) = \text{pH} + \log \left(\frac{\theta}{1 - \theta} \right), \quad (1.40)$$

where $\log(K) = \text{p}K_{\text{a}}$ has been used. For the ideal case where electrostatic interactions are negligible, $\log K_c$ is constant with a value equal to the intrinsic acid constant $\text{p}K_{\text{a}}$. However, in general, $\log K_c$ depends on the pH-value, the ionic strength of the medium I and the interaction of the polyelectrolyte with other charged species present

Finally, the propensity of a species to bind/unbind to the polyelectrolyte can be quantified in terms of its binding capacity C . C is defined as the variance of the probability distribution of N_C and reads^{28,38}

$$C = \frac{d\theta}{d \ln a_{\text{H}}} = \langle (N_C - N\theta)^2 \rangle. \quad (1.41)$$

As can be observed in Eq. 1.41, C is directly proportional to the fluctuation in the charge of the weak polyelectrolyte. The Fluctuation-Dissipation theorem dictates that, given that a system obeys detailed balance, the fluctuations of a physical variable predict the response quantified

by the impedance of the same physical variable. Thus, C can be used as an indicator to pinpoint the conditions where it is more feasible to observe charge regulation.

1.2.2 Site Binding model

Ionization processes have been mainly rationalized using the so-called Site Binding (SB) model. In this model, the polyelectrolyte is defined as a set of sites, whose ionization state can change by binding/unbinding chemical species.²³ No delocalization of proton binding is thus considered. In a similar way to the classical Ising model for ferromagnets, the ionization state of each site i is defined by the state variable s_i , so the complete set $\{s_i\}$, $i = 1, \dots, N$, where N is the total number of binding sites, defines all the possible ionization microstates of the polyelectrolyte. When the SB method is applied to proton binding, s_i take two possible values: $s_i = 0$ (site i unprotonated) or $s_i = 1$ (site i protonated). In the SB model, the electrostatic interactions are divided in "clusters" grouping pairwise interactions, triplet interactions, quadruplet interactions etc. In this way, the free energy of the system \mathcal{F}_{SB} is commonly written as the cluster expansion^{14,16,26}

$$\frac{\beta \mathcal{F}_{\text{SB}}(\{s_i\})}{\ln 10} = \sum_{i=1}^N \mu_i s_i + \sum_{i,j}^{N-1} \phi_{ij} s_i s_j + \sum_{i,j,k}^{N-2} \lambda_{ijk} s_i s_j s_k \dots, \quad (1.42)$$

where $\mu_i = \text{pH} - \text{p}K_{a,i} = -\log(K_a a_{\text{H}})$ is the reduced chemical potential of the ionizable site i , which depends on the proton activity, a_{H} , and the intrinsic acidic equilibrium constant of the site i , $\text{p}K_i$. The term "reduced" refers to the fact that the chemical potential incorporates both the pH-value and $\text{p}K_i$. In using the cluster expansion, the electrostatic interactions between charged sites are calculated by means of their corresponding cluster parameters, for instance: ϕ_{ij} is the pairwise interaction energy between sites i and j and λ_{ijk} is the triplet interaction energy between sites i , j and k . In principle, one could extend the cluster expansion until the exact limit of considering the N th-interactions in the polyelectrolyte chain. However, the main advantage of the cluster expansion is that it converges to the exact value of the free energy very quickly with only a few terms in Eq. 1.42.^{14,23,26} The interaction (or cluster) parameters are usually expressed in thermal units, i.e., $\beta = 1/k_{\text{B}}T = 1$, and divided by a factor $\ln 10$ in order to be compared in the pH scale. In the SB model, the conformational degrees of freedom are disregarded so these cluster parameters should be understood as averages over the possible bond rotations.¹⁶

With the free energy (Eq. 1.42) so defined, the machinery of statistical mechanics can be used to quantify mesoscopic quantities such as the polyelectrolyte titration curves, macroscopic protonation constants, site-site binding correlations and many others^{20,21,23,25,26,39,40}. The proper

statistical ensemble to be used is the semi-grand canonical ensemble, where the system is considered to be in equilibrium with a reservoir that keeps a_{H} constant. The corresponding semi-grand canonical function reads²⁶

$$\Xi = \sum_{\{s_i\}} a_{\text{H}}^n e^{\beta \mathcal{F}(\text{SB})} \quad (1.43)$$

Once Ξ is computed, the thermal averages can be readily computed. For instance, the average degree of ionization θ (Eq. 1.37) can be calculated as

$$\theta = \frac{a_{\text{H}}}{N} \left(\frac{\partial \ln \Xi}{\partial a_{\text{H}}} \right). \quad (1.44)$$

In order to illustrate the SB model, one can consider the paradigmatic case of the proton binding/unbinding to a linear weak polyelectrolyte, such as polyethyleneimine (Eq. 1.36). For simplicity, one can consider that all the binding sites identical and only first neighboring electrostatic interactions are present, as depicted in Fig. 1.10. In this case, the free energy of the system is then given by^{16,17,26}

$$\frac{\beta \mathcal{F}_{\text{SB}}(\{s_i\})}{\ln 10} = \sum_{i=1}^N \mu s_i + \sum_{i=1}^{N-1} \phi s_i s_{i+1}, \quad (1.45)$$

where μ is the reduced chemical potential of the sites and ϕ is the first neighbour interaction parameter. In an analogous way of the one presented for RIS model in section 1.1.5, in this simple case the transfer method can be used to compute the semi-grand canonical function Ξ of this system. Firstly, one defines a partition (file) vector $\vec{\Xi}_N$, whose elements are the partition functions of the chain with N sites with the last state variable s_N constrained to a particular value. The partition function of a polyelectrolyte chain with $N + 1$ sites Ξ_{N+1} can be calculated

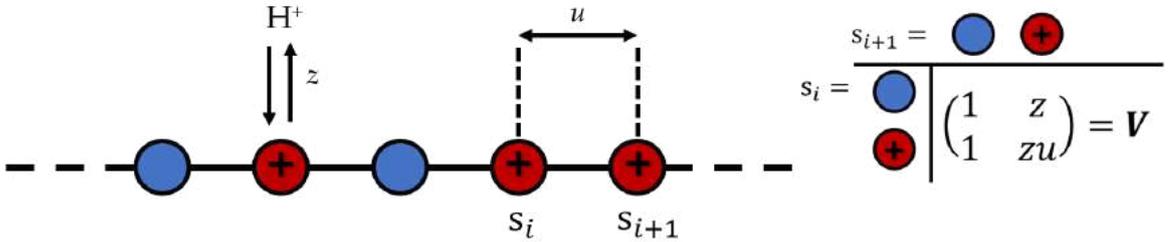


FIGURE 1.10: Outline of the Site Binding (SB) model corresponding to a linear weak polyelectrolyte chain with identical sites, when only neighboring interactions are considered. The system resembles polyethyleneimine (Eq. 1.36), a weak polycation. The polyelectrolyte is modelled as a set of binding sites, each site i having with two possible binding states s_i : protonated (red) or unprotonated (blue). The corresponding transfer matrix of the system U depends on the protonation energy z and the first neighbor interaction u .

using the transfer matrix method^{9,10}, as follows

$$\Xi_{N+1} = \vec{\Xi}_N \mathbf{V}. \quad (1.46)$$

The corresponding transfer matrix \mathbf{V} can be found by inspection of Eq. 1.45

$$\mathbf{V} = \begin{pmatrix} 1 & z \\ 1 & zu \end{pmatrix}, \quad (1.47)$$

where $z = \exp(-\beta\mu)$ and $u = \exp(-\beta\phi)$ are the statistical weights corresponding to the protonation and first neighbour interaction energies, respectively. Each element in \mathbf{V} stands for the energy increment in the system in adding a site $i + 1$ in a particular state s_{i+1} if the previous site i is in the state s_i , as outlined in Fig. 1.10. Using the recursive relation presented in Eq. 1.46, one can compute the semi-grand canonical of a polyelectrolyte with N sites as

$$\Xi = \vec{I} \mathbf{V}^N \vec{E} \quad (1.48)$$

where $\vec{I} = \begin{pmatrix} 1 & 0 \end{pmatrix}$ and $\vec{E} = \begin{pmatrix} 1 & 1 \end{pmatrix}^T$.

The SB model is not restricted to proton binding but it can be adapted to include a wide range of phenomena such as chelate complexation of metal ions²³, proton binding to polyampholytes^{20,41}, super-capacitor charging⁴² or protein-DNA binding⁴³. The model can also be extended to take into account long-range electrostatic interactions in an approximate way¹⁷.

1.3 Single-chain polymer elasticity

1.3.1 Fundamental concepts

The elastic response of a single polymer chain has been object of interest for decades due to its implications in the mechanical properties of polymeric materials. For many time, direct measurement of the elastic response of a single chain was unfeasible, and the field was mainly restricted to theoretical considerations⁸. However, the recent development of methods to manipulate single-molecules has allowed to obtain direct measurements of the mechanical response of single polymer chains⁵⁴. These techniques require two basic elements: (i) a probe that can generate and detect forces and displacements, and (ii) a method for spatially locating the molecules. The typical scenario in the stretching of a single-chain is displayed in Fig. 1.11 for a polyethylene chain. Firstly, one end of the polymer chain is chemically attached to a surface. Then, the mechanical force \vec{F} necessary to maintain a given polymer extension in the direction of the

force L_z is measured. By convention, \vec{F} is typically considered to have the same direction that the laboratory z -axis and perpendicular to the surface to which the chain is attached. It is important not to confuse L_z with the polymer contour length L_C (See Fig. 1.3). L_C is defined as the polymer maximum length and it is strictly only equal to L_z at the infinite force limit, when the polymer chain is completely extended.

Various techniques have been used for the study of single-molecule stretching, with detectable forces ranging from femtonewtons to a few nanonewtons.⁵⁵ The more common tools are mechanical transducers and external field manipulators. On the one hand, mechanical transducers, for instance Atomic Force Microscopy^{56,57} (AFM), apply or sense forces through the displacement of a bendable beam. On the other hand, external field manipulators, for example optical tweezers^{58,59}, can be used to exert forces on macromolecules either by acting on the molecules themselves or by exerting forces through “handles” attached to the molecules. These experiments have proved that some chemical reactions can be induced by stretching a molecule, leading to the creation of the field of mechano-chemistry. For instance, AFM have proven that single-molecule stretching can affect enzyme catalysis⁶⁰, induce *cis*-to-*trans* isomerization of carbon-carbon double bonds⁶¹ or produce surface desorption of polypeptides^{62,63}.

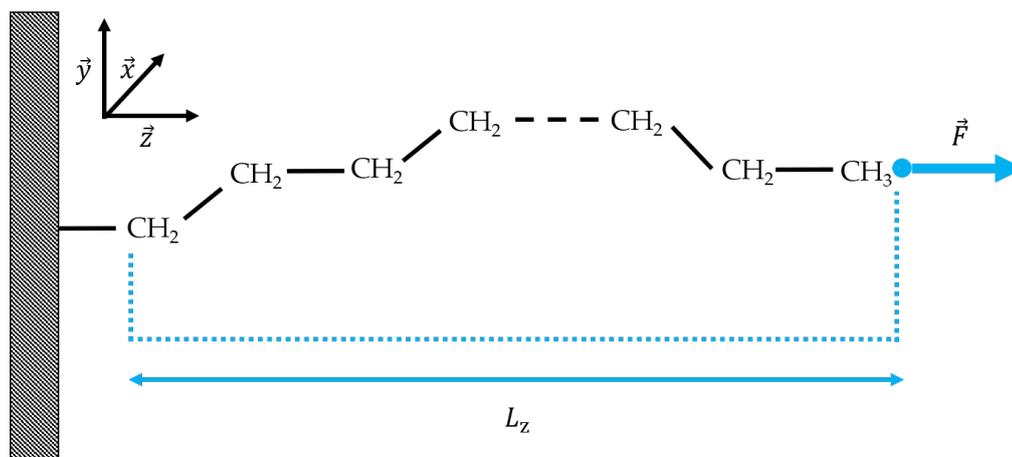


FIGURE 1.11: Outline of a single-chain stretching experiment, illustrated for the case of a polyethylene chain. One end of the chain is chemically attached to a surface, and the mechanical force \vec{F} necessary to fix a given polymer extension in the direction of the force, L_z , is measured. By convention, \vec{F} is typically considered to act along the laboratory z -axis and perpendicular to the surface.

1.3.2 Force regimes in the force-extension response

The polymer force-extension response has been found to exhibit a rich scaling behaviour which depends on the force regime and, in general, it cannot be fitted to a single model in the full force

range. Instead, in each force regime the polymer has a preferred model description, as outlined in Fig. 1.12.

At very small forces, the elastic response of the polymer chain follows the well-known classical entropic spring model. This model basis relies in the idea that the mechanical force is acting against the thermal fluctuations of the polymer. The force/extension curve, which can be derived from the fluctuation-dissipation theorem, reads⁸

$$L_z = \langle r^2 \rangle \beta F \approx \frac{Ml_k^2}{3} \beta F \quad (1.49)$$

where $\beta = 1/(k_B T)$ is the inverse of the thermal energy, with k_B as the Boltzmann constant and T as the absolute temperature. This force regime is known as the linear response regime of the polymer chain, and it is usually valid for $F < \beta^{-1}/\langle r \rangle$. De Gennes⁶⁷ proved that Eq. 1.49 can be extended to consider "swollen" polymer chains, *i.e.* chains whose self-crossing is not

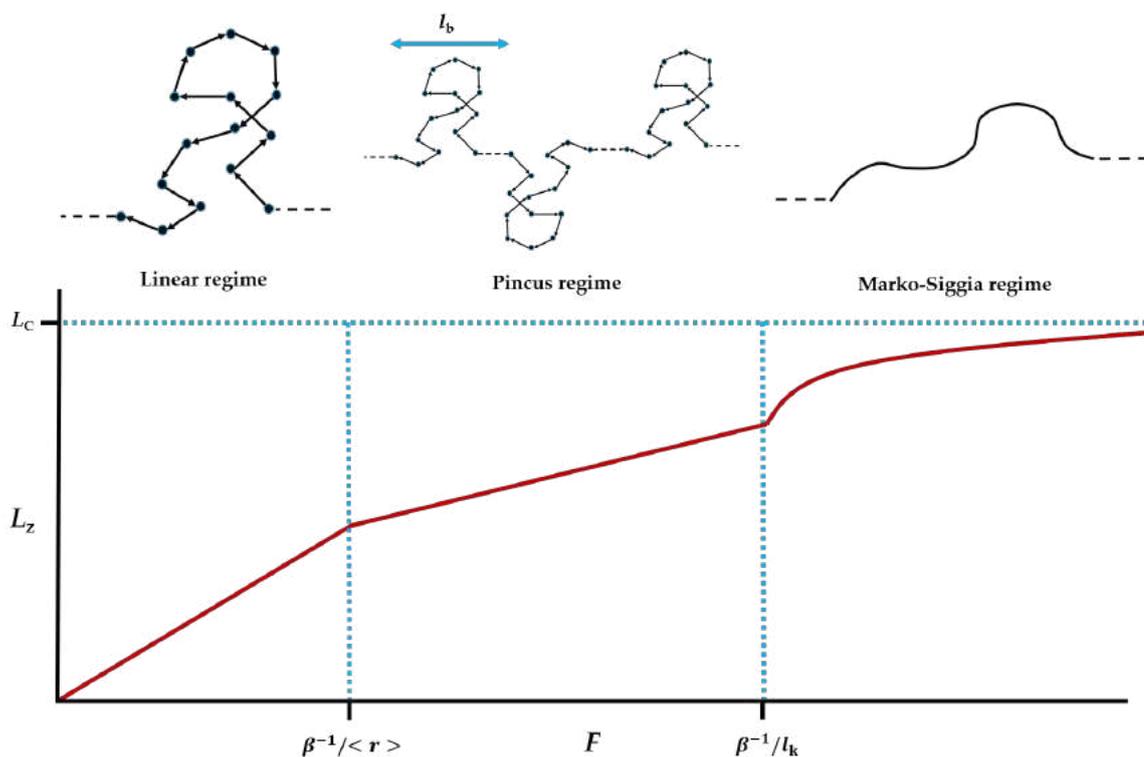


FIGURE 1.12: Outline of the typical force regimes observed in the force/extension curve of a single polymer chain. At low forces ($F < \beta^{-1}/\langle r \rangle$) the polymer exhibits a linear stretching response (Eq. 1.49)⁸. At intermediate forces ($\beta^{-1}/\langle r \rangle < F < \beta^{-1}/l_k$), it behaves as a chain of blobs, with a characteristic size l_b , which acts as a swollen random walk but keeping the blob chain in the direction of force following a Pincus scaling (Eq. 1.51)^{64,65}. At high forces ($F > \beta^{-1}/l_k$), the polymer is well represented by the WLC model and it follows the Marko-Siggia scaling (Eq. 1.52)⁶⁶

negligible, leading to the prediction

$$L_z = \langle r^2 \rangle \beta F \approx \frac{M^{6/5} l_k^2}{3} \beta F \quad (1.50)$$

Pincus^{64,65} found an intermediate force regime ($\beta^{-1}/\langle r \rangle < F < \beta^{-1}/l_k$), where the force/extension curve deviates from the linear regime. In this regime, the polymer acts as a chain of blobs with a characteristic size l_B . At length scales shorter than l_B , each blob acts as a swollen random walk but the blob chain is pulled in the direction of force. This model predicts the force-extension scaling law

$$L_z \sim F^{1/\nu-1}. \quad (1.51)$$

At higher forces ($F > \beta^{-1}/l_k$), the Marko-Siggia elastic model⁶⁶ fits the stretching response of a variety of polymers in which electrostatic interactions are either missing or negligible.⁸ The Marko-Siggia model relies in a WLC description of the polymer, with exponential backbone correlations. In the high-force limit, the force-extension expression for the Marko-Siggia model reads

$$L_z(\text{WLC}) \approx L_C \left(1 - \sqrt{\frac{1}{\beta 4 F l_p}} \right). \quad (1.52)$$

It also worth to mention that an expression, in principle valid in the full force regime, can also be derived for the FJC model using basic statistical mechanics⁶⁸

$$L_z(\text{FJC}) = L_C \left(\coth(\beta F l_k) - \frac{1}{\beta F l_k} \right). \quad (1.53)$$

One can note that at the low force limit ($F \rightarrow 0$), the linear regime is recovered (Eq. 1.49). At high forces ($F \rightarrow \infty$), Eq. 1.53 can be approximated as

$$L_z(\text{FJC}) \approx L_C \left(1 - \frac{1}{\beta F l_k} \right). \quad (1.54)$$

It can be noted that FJC (Eq. 1.54) and WLC (Eq. 1.52) have different predictions for the force/extension curve. It is also worth noting that in FJC and WLC models one does not consider the torsional energy involved in the rotation of the dihedral angles of the bonds. For flexible polymer chains, the elastic response of the chain is dependent on its conformational properties, as it will be assessed in forthcoming chapters of this thesis.

The models above mentioned only account for short range interactions (chain self-crossings) and can not be applied to polyelectrolytes, where the long-ranged electrostatic forces can not be neglected. The presence of electrostatic forces produces new elastic regimes for polyelectrolytes, which strongly deviate from the FJC and WLC models.⁷¹ Further, the polyelectrolyte

elastic response is affected by different factors which are not relevant in the stretching of neutral chains. For instance, experiments with single-stranded nucleic acids (ss-DNA and ss-RNA) have shown that polyelectrolyte force/extension curves are highly dependent on the valence and concentration of the counterions^{58,59,72–76}. Recent Monte Carlo simulations^{77,78}, have shown two elastic regimes for polyelectrolytes. At low forces, the polyelectrolyte behaves as a set of swollen electrostatic blobs on a long length scale, while at high forces a short length, ion-stabilized, crumpling structure is detected^{77–80}. These findings have motivated the development of the “snake chain model”^{79,81}, which seemingly explains well both elastic regimes.

1.4 Macromolecular crowding

1.4.1 Fundamental concepts

In biological media, although the concentration of a single macromolecule is usually low, there is a high total concentration of macromolecules (*e.g.* proteins, polysaccharides, etc.) surrounding the biopolymer. The term “macromolecular crowding” has been coined to refer to these heterogeneous and concentrated solutions of macromolecules⁸². For instance, the weight fraction of protein is around 5% in lymph, 9% in blood plasma and 35% in hemolysate^{83,84}. In general, macromolecular crowding occupies around 30–40% of the volume cellular cytosol⁸⁵. In the recent years, the study of biopolymer reactivity, dynamics and structural properties in crowded media has attracted an increasing interest^{84,85}. The relevance of such studies was firstly highlighted by Laurent⁸⁶: the natural medium where biopolymers carry out their biological functions is very different of the classic *in vitro* experiment. The general scenario is depicted in Fig. 1.13 where a hypothetical biopolymer (grey circles) in solution with small ions (red and green circles) can be found in the left-hand panel. This “ideal” situation is only realistic in *in vitro* studies and it is far from the real situation when it is in biological media, where there is a high macromolecular concentration (right panel in Fig. 1.13).

Macromolecular crowding substantially alters the reactivity^{82,84–90}, conformational properties^{91–96} and diffusion processes^{97–105} of biopolymers. These changes are produced by non-specific interactions among macromolecular species due to excluded volume, Van der Waals, electrostatic and hydrodynamic interactions.

For instance, the impact of macromolecular crowding in enzymatic reactions has been mainly rationalized with thermodynamic models relying on the excluded volume effect^{82,84,86}. For instance, the velocity of some enzymatic reactions is increased when carried out in macromolecular crowding conditions. This is explained as a cage effect due to an effective increase in the

reactants concentration caused by the excluded volume of the crowders. Moreover, recent studies highlight that the effect of macromolecular crowding in enzymatic reactivity also depends on the relative enzyme-crowder size ratio^{87,89}.

Macromolecular crowding is also known to present significant influence on the conformational stability of biopolymers⁹⁵. In general, macromolecular crowding stabilizes compact conformations of the macromolecules due to volume exclusion. However, in the case of proteins without a well-defined structure, the so-called Intrinsically Disordered Proteins (IDPs), crowding affects their structure in a highly heterogeneous manner^{91–96}. In fact, regarding their response to crowding, IDPs have been classified as foldable (fold upon crowding), unfoldable (extend upon crowding) and non-foldable (mostly unaffected upon crowding)⁹⁵. This variability suggests that other IDP-crowder interactions in addition to the exclusion volume may be necessary to rationalize these systems.

1.4.2 Polymer diffusion in crowded media

Polymer diffusion in crowded media is known to be different than that observed in dilute conditions^{100,101,106}. Crowder particles slow down the tracer polymer diffusion due to volume exclusion and hydrodynamic interaction leading to different diffusive regimes. In this section, the fundamental notions of diffusion in dilute and crowded media are briefly outlined.

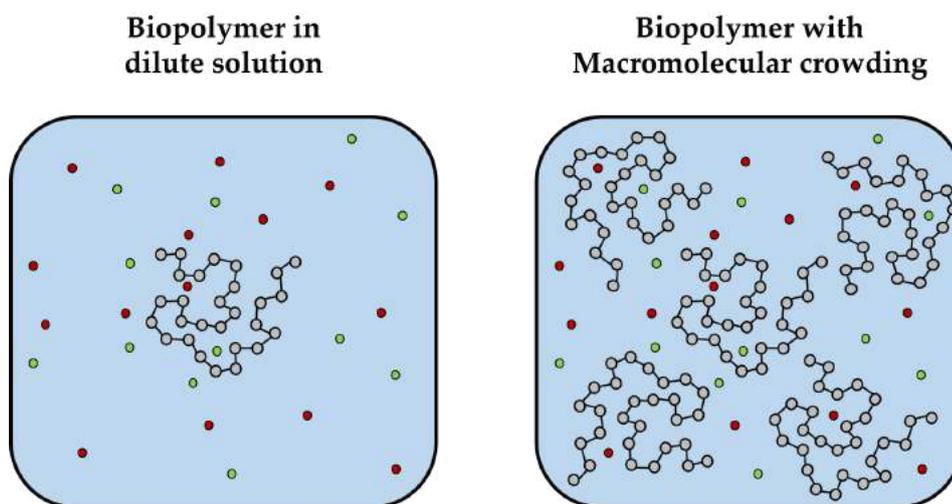


FIGURE 1.13: Scheme of macromolecular crowding. A hypothetical biopolymer is depicted as a set of grey circles together with small ions (red and green circles). Two different situations are shown: (left) the biopolymer is in a "dilute" solution, typical of *in vitro* experiments and (right) the biopolymer is surrounded other biopolymers, referred to as macromolecular crowding, which is the usual scenario in biological media.

Polymer motion in dilute solution follows the well-known random walk or Brownian dynamics¹⁰⁷, which is illustrated in Fig. 1.14. This motion results from the isotropic collisions of the polymer with the solvent molecules, which produce the polymer movements to be erratic and random. The classical magnitude to measure the polymer motion is the mean square displacement $\langle r_{\text{MSD}}^2(t) \rangle$, which is the second moment of the difference between the polymer initial position and its position at the measure time t . In dilute solution, $\langle r_{\text{MSD}}^2(t) \rangle$ follows the well-known Einstein-Smoluchowski relation

$$\langle r_{\text{MSD}}^2(t) \rangle = 2dD_0t, \quad (1.55)$$

where d is the topological dimension of the system and D_0 is the diffusion coefficient of the polymer in dilute solution. Considering the polymer a spherical particle in a Newtonian fluid, D_0 can be calculated by means of the Stokes-Einstein equation

$$D_0 = \frac{k_B T}{6\pi\eta R_H}, \quad (1.56)$$

where η is the solvent viscosity and R_H is the hydrodynamic radius of the particle. Note that D_0 only depend on the medium conditions (η and T) and the particle dimensions (R_H) but it is independent of the time of measure.

When macromolecular crowding is present, the situation changes and the polymer diffusion coefficient D is no longer constant over time^{100,101,106}. As outlined in Fig. 1.15, in this situation the diffusion coefficient has three different temporal regimes: (i) at short times, the diffusion coefficient has a constant value D^{short} ; (ii) at intermediate times, D decays over time and (iii)

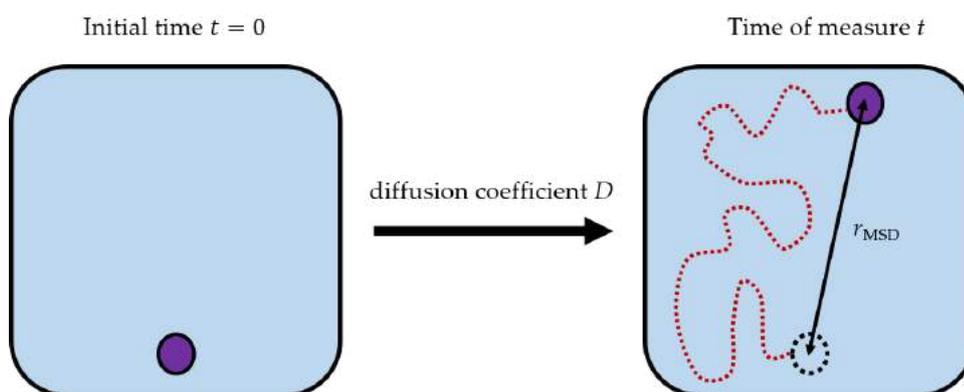


FIGURE 1.14: Graphical scheme illustrating the motion of a generic polymer (purple circle) in solution. The polymer follows a random (brownian) walk due to the isotropic collisions with the solvent. Typically, the polymer diffusion is quantified measuring the displacement r_{MSD} from the polymer position at the measure time t to its position at the initial time $t = 0$. r_{MSD} is directly proportional to the polymer diffusion coefficient D , and follows the well known Einstein-Smoluchowski relation (Eq. 1.55).

at long times, the diffusion coefficient reaches a stationary value D^{long} . The physical reason for such transition can be explained as follows.

The situation at short times is depicted in the panel 1 in Fig. 1.15. The polymer tracer particle (purple circle) is still close to its original position (dashed circle) and it has still not collided with any crowder particle (purple circles). As a first intuition, one would naively expect the diffusion coefficient at these times to be equal to its dilute (ideal) value $D^{\text{short}} = D_0$. However, this is not generally true since the crowders influence the polymer motion through Hydrodynamic Interactions (HI) ^{108–111}. HI are solvent-mediated tracer-crowder interactions due to the solvent fluxes originated by the motion of the crowders. D^{short} can be estimated using the analytical expression derived by Tokuyama ¹⁰⁸. It assumes a solution of equal-sized spherical brownian particles interacting through mean field HI:

$$D^{\text{short}}(\phi) = \frac{D_0}{[1 + H(\phi)]}; \quad H(\phi) = \frac{2b^2}{1-b} - \frac{c}{1+2c} - \frac{bc(2+c)}{(1+c)(1-b+c)}, \quad (1.57)$$

where ϕ is the volume fraction occupied by the crowders, $H(\phi)$ is the contribution due to the HI, only valid in the short-time regime, $b = \sqrt{\frac{9}{8}\phi}$ and $c = \frac{11}{16}\phi$.

At intermediate times, the tracer particles start to collide with the crowders in the nearby (See panel 2 in Fig. 1.15). In this transition regime, the polymer motion deviates from the Einstein-Smoluchowski equation and its diffusion is called anomalous diffusion. Under these conditions, diffusion is described by a generalization of the Einstein-Smoluchowski equation (Eq.

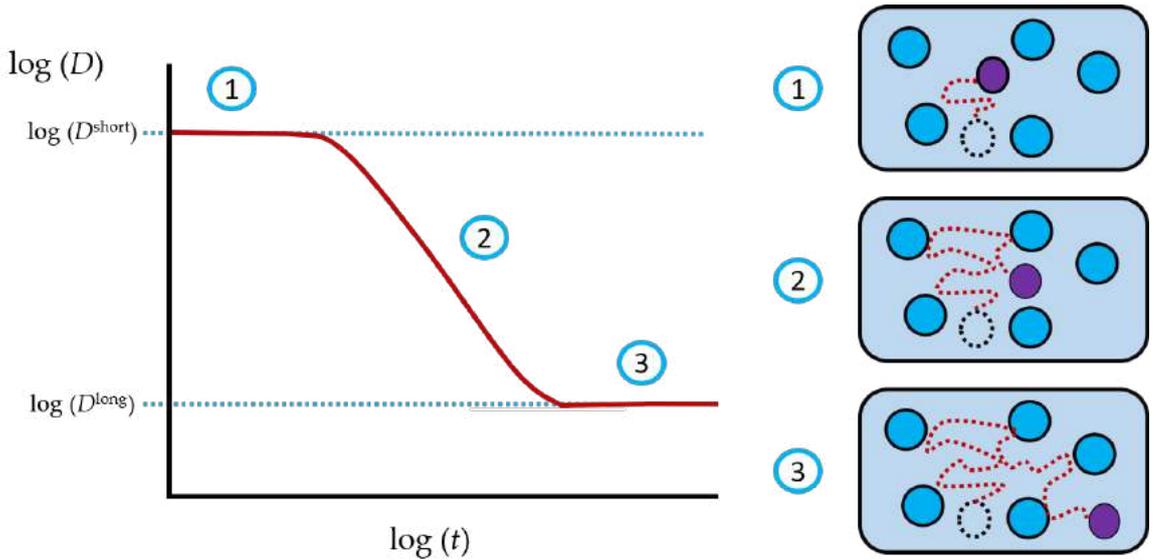


FIGURE 1.15: Scheme outlining the different temporal regimes of the diffusion coefficient D of a polymer in a crowded media. D undergoes a transition from a constant value at short times, D^{short} , to a different constant value at long times D^{long}

1.55), which reads^{100,101,106}

$$\langle r^2 \rangle = (2d)\Gamma_\alpha t^\alpha. \quad (1.58)$$

This is formally equivalent to having a diffusion coefficient that is not constant on time¹¹²:

$$D(t) = \frac{\Gamma t^{\alpha-1}}{2d}, \quad (1.59)$$

where α is the anomalous diffusion exponent and Γ_α is a generalized transport coefficient. The value of α allows to differentiate between the two different types of anomalous diffusion: super-diffusive ($\alpha > 1$) and sub-diffusive ($\alpha < 1$). Macromolecular crowding is known to produce only sub-diffusive behaviour^{97,98,100,101,104}. For $\alpha = 1$, the Einstein-Smoluchowski equation (Eq. 1.55) is recovered.

Finally, at long times, the polymer motion is stationary and the diffusion coefficient reaches an asymptotic value D^{long} (See panel 3 in Fig. 1.15). At long times, one observes the "average" motion of the particle, impeded by the crowded particles. A mean field analytical expression for D^{long} has also been derived by Tokuyama^{109,110}, under the same assumptions used to derive Eq. 1.57, which reads

$$D^{\text{long}}(\phi) = \frac{D^{\text{short}}(\phi)}{\left[1 + \kappa \frac{D^{\text{short}}(\phi)}{D_0} \left(\frac{\phi}{\phi_c}\right) \left(1 - \frac{\phi}{\phi_c}\right)^{-2}\right]}, \quad (1.60)$$

where κ and ϕ_c are parameters with values $\kappa = 2.0$ and $\phi_c = 1.09$.

The diffusion coefficient is usually experimentally measured using spectroscopic techniques, mainly Fluorescence Correlation Spectroscopy¹¹³ (FCS) and Fluorescence Recovery After Photobleaching^{98,99} (FRAP). An scheme of both techniques can be found in Fig. 1.16, where the tracer macromolecules are represented as green circles. In FRAP and FCS experiments, fluorescent markers are specifically bound to tracer particles and the intensity of the fluorescence (I) is measured over time. In FCS, the fluctuations in fluorescence intensity are recorded and used to calculate the temporal auto-correlation function, which allows determining the diffusion coefficient. In FRAP, a small volume of the sample is lighted with a laser beam. The molecules in the lighted region become bleached (grey circles in Fig. 1.16) and no fluorescence is exhibited. Then, the diffusion coefficient of the fluorescent molecule can be estimated by means of the velocity of fluorescence recovery in the bleached region. Both techniques usually measure a characteristic residence time around 1–3 ms^{97–99}, which is defined as the time which the tracer particle needs to traverse the characteristic length of the detection volume. Since this value is much larger than the characteristic times where D^{long} is measured in computational simulation^{97–105}, the D measured in these experiments should reasonably correspond to D^{long} .

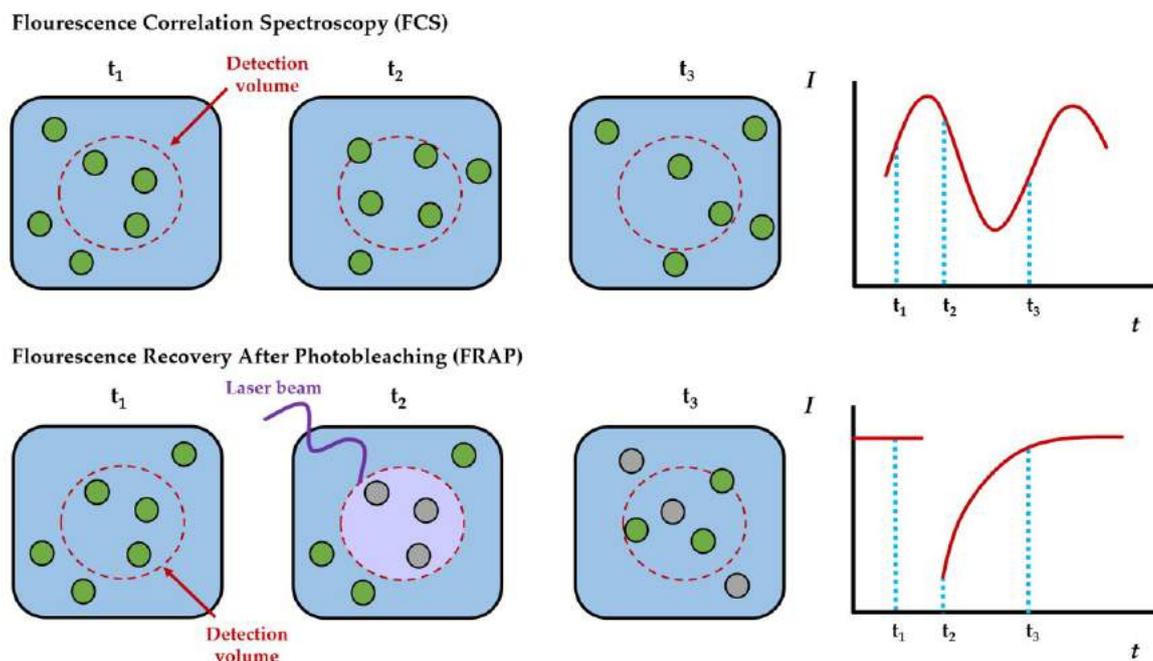


FIGURE 1.16: Outline of two spectroscopic techniques used to measure macromolecular diffusion: Fluorescence Correlation Spectroscopy (FCS) and Fluorescence Recovery After Photobleaching (FRAP). In both techniques, a fluorescent marker is bound to the tracer macromolecule (green circles) and the intensity of fluorescence (I) is measured at different times t . In FCS, the fluctuations in I are measured and used to calculate the tracer diffusion coefficient. In FRAP, a small region of the sample is bleached with a laser beam, making the macromolecules to lose their fluorescence (grey circles). Then, the velocity of the fluorescence recovery in the bleached region is measured, which allows to estimate the diffusion coefficient.

1.5 Thesis outline

I have carried out this thesis in the Biophysical Chemistry of Macromolecules and Colloids research group of the University of Barcelona. In the recent years, the research group have made significant contributions to the study of the conformational, binding, diffusion and reactivity properties of weak polyelectrolytes and biopolymers^{14,16,17,52,53,87,89,98–101,104,130}.

On the one hand, the group has contributed to the design and development of the Site Binding Rotational Isomeric State (SBRIS) model for weak polyelectrolytes (Sec. 2.1)¹⁴. They have also successfully used the SBRIS model to reproduce the complex experimental titration curves of polyethylenimine¹⁶. At high ionic strength vales, the SBRIS analytical prediction of polyethylenimine radius of gyration was also in good agreement with constant pH simulations, where long range electrostatic interactions where explicitly included. Moreover, they have recently extended the Site Binding model (Sec. 1.2.2) to include long range electrostatic interactions using the Gibbs-Bogoliubov variational principle¹⁷.

On the other hand, the research group has done significant contributions in the study of biopolymer diffusion and reactivity in crowded media^{87,89,98-101}. They tracked the diffusion of α -chymiotrypsin in solutions crowded by different-sized dextran macromolecules^{98,99}. They rationalized the experimental results comparing them with *on-lattice* Monte Carlo (MC) simulations, observing a qualitative agreement between both sets of data^{100,101}. Moreover, the group reported that the enzyme kinetics of L-Lactate Dehydrogenase are altered by macromolecular crowding, dependent on factors such as excluded volume and crowder size^{87,89}.

The present thesis follows the efforts of the research group in the development of theoretical and computational models for weak polyelectrolytes and biopolymers. In Chapter 2, the different theoretical and computational methodologies employed in this thesis are presented. Chapters 3 to 7 include articles published in peer-reviewed scientific journals. Chapter ?? encloses a manuscript that has been also submitted to a scientific journal and it has been just accepted after peer-review. A brief summary of each publication is included at the beginning of each chapter, prior to the corresponding original article. The thesis is organized as follows:

- In Chapter 2, the theoretical and computational methods used in this thesis are introduced. The Site Binding Rotational Isomeric State (SBRIS) model is explained, which is later used in chapters 3 to 5. Moreover, a general outline of the polymer models more relevant for the scope of this thesis is given, highlighting their different levels of the description of the macromolecules. The different computational methods used in this thesis, Semi Grand Canonical Monte Carlo (Chapters 3 to 5 and Chapter 8) and Brownian Dynamics (Chapters 6 and 7) simulations, are explained. Finally, computational details are given explaining the two computational codes that I developed throughout this thesis.
- In Chapter 3, the analytical resolution of the SBRIS model is extended to include intramolecular long range electrostatic interactions. The long range interactions are taken into account by defining effective short range and pH-dependent parameters, such as effective microscopic protonation constants and rotational bond energies. The technique is tested with a minimal model of a flexible linear polyelectrolyte, containing only one type of rotating bond. The analytical prediction of structural and binding properties of a model polyelectrolyte is compared to those obtained by Semi-Grand Canonical Monte Carlo (SGCMC) simulation. The chapter corresponds to the publication "P. M. Blanco, S. Madurga, F. Mas, and J. L. Garcés. Coupling of charge regulation and conformational equilibria in linear weak polyelectrolytes: Treatment of long-range interactions via effective short-ranged and pH-dependent interaction parameters. *Polymers*, 10(8):811, 2018"¹³⁴. I programmed the SGCMC code and I have performed all the simulations exhibited in the chapter. I wrote the main text of the article, in collaboration with the other authors of the publication.

- In Chapter 4, the model of a flexible linear polyelectrolyte proposed in the previous chapter is extended to investigate the elastic response of a model weak polyelectrolyte chain with SGCMC simulations. The hypothesis that the ionization of weak polyelectrolyte is sensitive to the action of a force stretching the polymer chain is proposed and validated. The force/extension curves of the polyelectrolyte chain are analyzed at different pH and ionic strength conditions and different force regimes are identified and rationalized. Moreover, the structural, binding and elastic properties of the polyelectrolytes are assessed in the different force regimes. The chapter corresponds to the publication "P. M. Blanco, S. Madurga, F. Mas, and J. L. Garcés. Effect of Charge Regulation and Conformational Equilibria in the Stretching Properties of Weak Polyelectrolytes. *Macromolecules*, 52:8017–8031, 2019"¹³⁵. I did the necessary modifications to the SGCMC code developed in the previous chapter to include the action of an stretching force and I have performed all the simulations and I did all the data analysis. I have also planned the simulation framework and I proposed the analysis of the Pincus exponent as a function of the pH-value. I wrote the main text of the article, in collaboration with the other authors of the publication.
- In Chapter 5, the study presented in the previous chapter is continued to investigate the effect of charge fluctuation in the conformational and stretching properties of weak polyelectrolytes. The structural and elastic magnitudes measured by SGCMC simulation are compared with those obtained with MC simulation with the same charge but keeping it constant. The Pincus stretching regime is studied in detail, extending the results obtained in the previous chapter. Moreover, the effect of excluded volume in the stretching and conformational properties of the weak polyelectrolyte is assessed in a wide range of pH and ionic strength conditions. The chapter corresponds to the publication "P. M. Blanco, S. Madurga, C. F. Narambuena, F. Mas, and J. L. Garcés. Role of Charge Regulation and Fluctuations in the Conformational and Mechanical Properties of Weak Flexible Polyelectrolytes. *Polymers*, 11:1962, 2019"¹³⁶. This work was done in collaboration with Dr. Claudio F. Narambuena from the Universidad Tecnológica Nacional (UTN) and CONICET (San Rafael, Argentina). I participated in the conceptual framework of the article and I proposed to perform the constant charge MC simulation to assess the role of charge fluctuation. I programmed the constant charge MC code, performed all the simulations shown in the chapter and I did all the data analysis. I wrote the main text of the article, in collaboration with the other authors of the publication.
- In Chapter 6, the diffusion of two globular proteins, α -chymiotrypsin and streptavidin, in media crowded by dextran macromolecules is investigated with BD simulations. The macromolecules were modelled as hard spheres, using a sing sphere coarse-grained model with an effective radius. Two different treatments of the hydrodynamic interactions, the

Rotne-Prager-Yamakawa and Tokuyama methods, are investigated. Different possible choices for the effective radius of the spheres representing the dextran crowders are analysed. Finally, the diffusive properties measured in the BD simulations are tested against experimental data. The chapter corresponds to the publication "P. M. Blanco, M. Via, J. L. Garcés, S. Madurga, and F. Mas. Brownian dynamics computational model of protein diffusion in crowded media with dextran macromolecules as obstacles. *Entropy*, 19(3):105, 2017"¹³⁷. I programmed the BD code with the different treatments for the hydrodynamic interactions and I have performed all the simulations. I proposed the idea that an effective radii is necessary to describe dextran macromolecules in crowded media, due to their expected compaction and entanglement in such conditions. I wrote the main text of the article, in collaboration with the other authors of the publication.

- In Chapter 7, the study done in the previous chapter is improved by proposing a new coarse-grained model, which is able to reproduce macromolecular entanglement in crowded media in a effective way. The model, named Chain Entanglement Softned Potencial (CESP), describes the macromolecules as double-shelled spheres. The spheres have a soft outer region (corresponding to the less dense, branched region of the macromolecule) and a hard inner region (corresponding to the dense core of the macromolecule). The CESP model is used to study streptavidin diffusion in dextran-crowded media with BD simulations and it is tested against experimental data and compared against the results obtained in the previous chapter. Moreover, a new empirical expression is proposed which describes the full temporal evolution of the diffusion coefficient observed in crowded media. The chapter corresponds to the publication "P. M. Blanco, J. L. Garcés, S. Madurga, and F. Mas. Macromolecular diffusion in crowded media beyond the hard-sphere model. *Soft Matter*, 14(16):3105–3114, 2018"¹³⁸. I proposed the idea of describing the macromolecular flexibility using a double-shelled description and designed the equation used for the CESP model. I have also proposed the empirical expression used to fit the diffusion coefficient temporal evolution obtained in the simulations. I designed the simulation framework of the study, adapted the BD code to the CESP potential and performed all the simulations exhibit in the chapter. I wrote the main text of the article, in collaboration with the other authors of the publication.
- In Chapter 8, the fundamental and technical knowledge gathered in all the previous chapters is applied to study the influence of macromolecular crowding in the structural and binding properties of two Intrinsically Disordered Proteins (IDPs), histatin-5 and β -amyloid. The investigation is motivated by the hypothesis that, in the same way that the conformational and ionization degrees of freedom are coupled in weak polyelectrolytes, the macromolecular crowding should also cause a charge regulation in IDPs whose flexible structure is similar to those of weak polyelectrolytes. Two possible mechanisms

through by macromolecular crowding can alter the IDP ionization are proposed: (i) the increase in the effective ionic strength (same ions in a reduced volume) and (ii) the IDP compaction due to macromolecular crowding. We investigate the binding properties of the IDPs using SGCMC simulations with MOLSIM software¹³⁹. The IDPs are modelled with a bead-and-spring model and the crowders are modelled after Bovin Serum Albumin (BSA) protein using the CESP model developed in the previous chapter. The chapter corresponds to a just accepted manuscript, "P. M. Blanco, S. Madurga, J. L. Garcés, F. Mas, R. S. Dias. Influence of Macromolecular Crowding in the Charge Regulation of Intrinsically Disordered Proteins. *Soft Matter*, 2020, DOI: 10.1039/D0SM01475C ". This work was done in collaboration with Dr. Rita S. Dias from Norges Teknisk-Naturvitenskapelige Universitet NTNU (Trondheim, Norway), during a research stay on her research group. Dr. Dias collaborated with me in the development of the research plan, the discussion of the results and the manuscript writing. I proposed the investigation and the different hypothesis aforementioned. I have implemented the CESP potential as a module in MOLSIM software and I have performed all the simulations. I wrote the main text of the article, in collaboration with the other authors of the manuscript.

Chapter 2

Methodology

2.1 The Site Binding Rotational Isomeric State model

The Site Binding Rotational Isomeric State (SBRIS) model was proposed as a combination of the SB and RIS models¹⁴. The resulting model treats the conformational and binding equilibria on the same foot^{14,16}. In an analogous way to RIS model (sub-section 1.1.5), the conformational state of the polyelectrolyte is defined by the set of variables $\{c_i\}$, $i = 1, \dots, M$, where M is the total number bonds in the polymer backbone, which describe the rotational state of each bond. In the same fashion as in the SB model (sub-section 1.2.2), the ionization state of the binding sites of the polyelectrolyte is given by the state variables $\{s_i\}$, $i = 1, \dots, N$, where N is the total number of binding sites. The SBRIS free energy is the sum of both energy contributions, RIS and SB, and reads^{14,16}

$$\mathcal{F}_{\text{SBRIS}}(\{c_i\}, \{s_i\}) = \mathcal{F}_{\text{p}}(\{c_i\}, \{s_i\}) + \mathcal{F}_{\text{RIS}}(\{c_i\}), \quad (2.1)$$

where $\mathcal{F}_{\text{RIS}}(\{c_i\})$ is the free energy of the reference state, corresponding to the fully unbinded state of each conformer, and follows Eq. 1.24. $\mathcal{F}_{\text{p}}(\{c_i\}, \{s_i\})$ is the protonation free energy, which calculated in a similar way to the SB model (Eq. 1.45) with the cluster expansion^{14,16}

$$\frac{\beta \mathcal{F}_{\text{p}}(\{c_i\}, \{s_i\})}{\ln 10} = \sum_{i=1}^N \mu_i(c_i) s_i + \sum_{i,j}^{N-1} \phi_{ij}(c_i) s_i s_j + \sum_{i,j,k}^{N-2} \lambda_{ijk}(c_i) s_i s_j s_k + \dots, \quad (2.2)$$

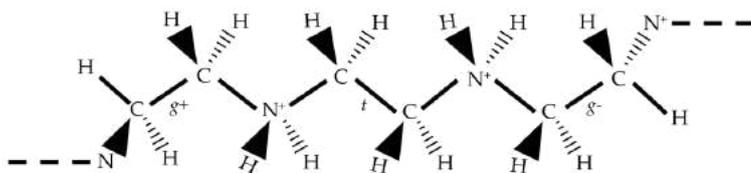
where $\mu_i(c_i)$ is the reduced chemical potential of the site i , $\phi_{ij}(c_i)$ is the pairwise interaction energy between sites i and j and $\lambda_{ijk}(c_i)$ is the triplet interaction energy between sites i , j and k . The parameters depend on the rotational state $\{c_i\}$ of bond i which holds the sites i and $i + 1$. The difference between SBRIS and SB models is precisely that in the SBRIS model the

conformational degrees of freedom are explicitly introduced whereas in the SB model they are only present in an average way through the interaction parameters.¹⁴

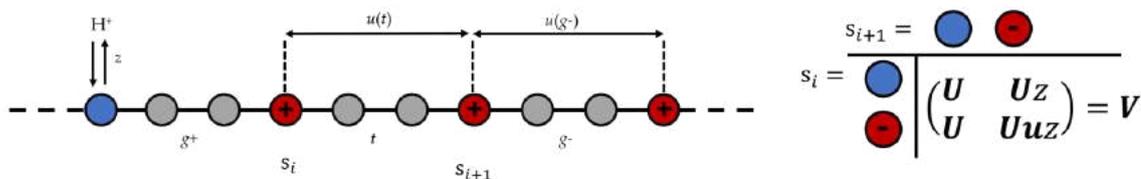
The SBRIS model can be illustrated considering the hypothetical case of a polycarboxylic acid which has one carboxylic group pending of every carbon of its backbone chain as depicted in Fig. 2.1. The monomeric units of the polyelectrolyte can be represented as binding sites joined by rotating bonds. In the simplest case, one can (i) reduce the rotational states to those of minima energy (trans, *gauche+* and *gauche-*), (ii) consider the binding constant to be independent of the conformational state ($\mu_i(c_i) = \mu_i$) and (iii) consider only nearest interactions between bonds and sites. Under these assumptions, the protonation energy reads

$$\frac{\beta \mathcal{F}_p(\{c_i\}, \{s_i\})}{\ln 10} = \sum_{i=1}^N \mu_i s_i + \sum_{i=1}^{N-1} \phi(c_i) s_i s_{i+1}, \quad (2.3)$$

Linear polyethyleneimine



Site Binding Rotational Isomeric State (SBRIS) model



Site Binding (SB) model

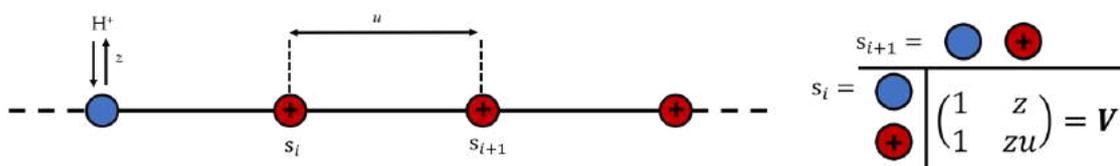


FIGURE 2.1: Scheme of the Site Binding Rotational Isomeric State (SBRIS) model, corresponding to the case of a linear poly(ethyleneimine) chain with only first neighbour interaction. Each monomer is enclosed in a binding site, which is joined with its neighbours with rotating bonds that can be in three conformational states (c): trans (t), *gauche+* ($g+$) and *gauche-* ($g-$). The conformational state of the bond holding two neighboring sites changes the value of the first neighbour interaction $u(c)$. The semi-grand canonical partition function can be calculated as a product of the super matrix V which depends on the RIS model transfer matrix U , the protonation energy z and the first neighbour interaction energy matrix $u(c)$.

The semi-grand canonical partition function of this system, Ξ , can be computed as the product (Eq. 1.43) of the transfer super-matrix \mathbf{V} , which in this case reads¹⁶

$$\mathbf{V} = \begin{pmatrix} \mathbf{U} & \mathbf{Uz} \\ \mathbf{U} & \mathbf{Uzu} \end{pmatrix}; \quad \mathbf{u} = \begin{pmatrix} u(t) & 0 & 0 \\ 0 & u(g+) & 0 \\ 0 & 0 & u(g-) \end{pmatrix}, \quad (2.4)$$

where \mathbf{U} is the RIS model transfer matrix (defined in Eq. 1.28) and $z = \exp(-\beta \ln 10(\text{pH} - \text{pK}_a))$. $u(c)$ is the statistical weight corresponding to interaction energy between site i and $i \pm 1$ through a bond in the conformational state c .

The SBRIS model has been used to rationalize the complex titration curves of poly(acrylic acid) (PAA) and poly(methacrylic acid) (PMAA)¹⁴. In particular, the "jump" in the experimental PMMA titration curve was quantitatively reproduced using the SBRIS model and found to be a consequence of a conformational transition. In turn, the experimental titration curve of poly(ethylenimine) has been found to follow the SBRIS model for polyelectrolyte chains of different length¹⁶. In this work, the SBRIS model was compared with Monte Carlo (MC) simulations, where long range interactions (electrostatic and excluded volume) are explicitly considered. For long poly(ethylenimine) chains, the SBRIS was found to overestimate the radius of gyration value in comparison with the corresponding value obtained by MC simulation. This result suggests that long range interactions are necessary for a proper theoretical description of poly(ethylenimine) conformational properties.

2.2 Polymer models at different levels of description

A wide variety of models, with varying levels of description, have been designed to explain the structural, binding, elastic and dynamic properties of polymers, polyelectrolytes and biopolymers^{2,12,16,19,46,50,128–130}. Regarding this thesis, the more relevant ones are presented in Fig. 2.2, where they are illustrated for the case of polyethylenimine linear chain. The models are ordered from top to bottom in decreasing level of detail.

Atomistically-detailed models aim to accurately describe the specific interactions between all the atoms of the system^{88,128}. These models are typically used in the study small molecules or fragments of macromolecules^{128,131}. A remarkable exception is the study of conformational properties of proteins, where all-atom molecular dynamics simulation are commonly used¹³² (and references quoted therein). These simulations are computationally expensive and they have characteristic running times of several weeks. The computational cost can be reduced removing the solvent atoms and approximating the solvent as a continuum with a dielectric constant ϵ , *i.e.* performing simulations with implicit solvent¹¹³ (and references quoted therein). This approximation implies the loss of specific solvent-solute interactions and solvation effects.

One strategy to avoid the computationally demanding computation associated to atomistically-detailed simulations is to reduce the system resolution, retaining the more relevant physico-chemical aspects for the problem of interest. One example of this procedure is the Site Binding Rotational Isomeric State (SBRIS) model^{14,16}, where only the atoms of the polymer backbone chain are explicitly simulated and the rest of atoms of the chain are included implicitly in the torsional energy of the chain. Furthermore, the conformational degrees of freedom of the polymer chain are reduced to a selected number of torsional states. Another example is the use of bead-and-spring models, which are specially popular in polymer computational simulation^{19,46,47,50,53,129,130,133}. These models reduce the number of particles of the systems by representing the polymer as a set of spheres, usually named beads. Each bead encloses some of the atoms of the polymer, typically one monomeric unit, effectively occupying the same volume.

SBRIS and bead-and-spring models (when studied in the Semi-Grand Canonical ensemble) still retain the coupling between the conformational and binding degrees of freedom, allowing to study them at the same foot. Classical theoretical models usually focus in either the conformational or the binding degrees of freedom and consider the other in an average way. On the one hand, the Freely Jointed Chain (Sec. 1.1.2), Freely Rotating Chain (Sec. 1.1.3), Worm-Like Chain (Sec. 1.1.4) and Rotational Isomeric State (Sec. 1.1.5) models describe the polymer conformational degrees of freedom^{2,11,71,131}. On the other hand, the Site Binding model (Sec. 1.2.2)

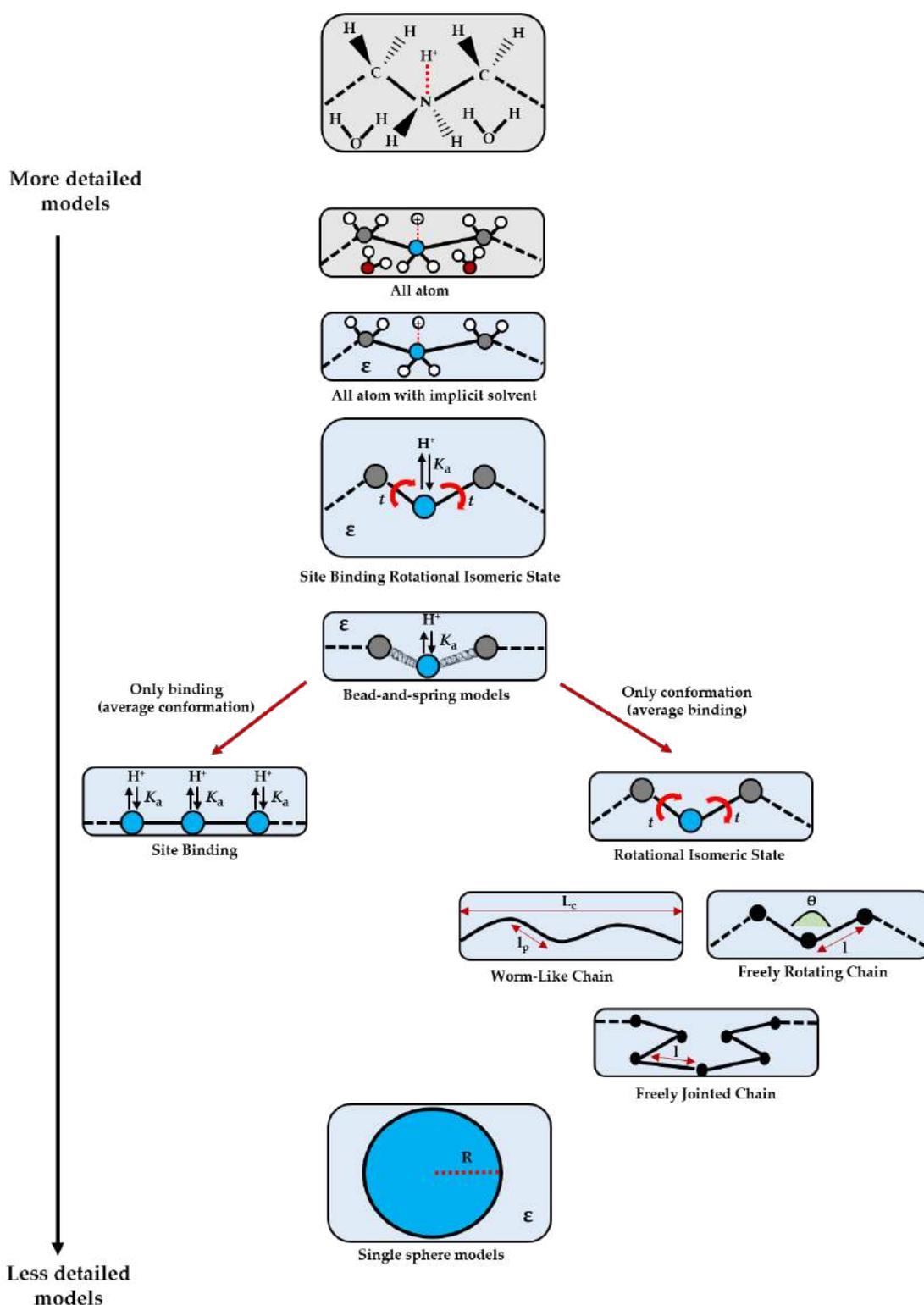


FIGURE 2.2: Schematic outline of the different models used in the theoretical and computational investigation of polymer chains. The models are illustrated using as example a polyethylenimine linear chain. The models are ordered, from top to bottom, in decreasing level of detail.

is designed to study binding properties of weak polyelectrolytes^{20,23,41}. These theoretical models are often used to rationalize the experimental data, for instance polymer force-extension curves⁵⁴, since they usually have analytical solution.

In the limit of lower resolution, the whole macromolecule is modelled as a single sphere with an effective radius R . This single-sphere coarse-grained models are usually employed in the study of highly concentrated solutions of macromolecules, where the quantity of atoms makes the computational cost of all-atom models prohibitive^{105,120}.

2.3 Semi-Grand Canonical Monte Carlo (SGCMC) simulation

Alternatively to the transfer matrix method described for SB and SBRIS models, the ionization properties of weak polyelectrolytes can be calculated using Monte Carlo (MC) simulation techniques. The more characteristic feature of Semi-Grand Canonical Monte Carlo (SGCMC) simulations is that the polyelectrolyte charge is not constant over the simulation but it fluctuates as a natural result of the dynamic proton binding/unbinding. Semi-Grand Canonical Monte Carlo (SGCMC) simulations were originally developed to study phase behaviour of reactive mixtures⁴⁴. In SGCMC, the system is in equilibrium with a reservoir which maintains the chemical potential of one component of the mixture fixed. Its use was firstly extended to weak polyelectrolytes by Reed and Reed⁴⁵, which proposed a linear polyelectrolyte chain in equilibrium with a reservoir which maintains the proton chemical potential constant. A direct consequence of fixing the proton chemical potential is that the pH value of the medium is also kept constant.¹⁹

In MC simulation, the phase space of the system is explored by using the Metropolis algorithm^{22,40}, which is outlined in Fig. 2.3. Starting from an initial state, the algorithm generates random trial states, which are accepted or rejected depending on the energy difference between the initial and trial states $\Delta\mathcal{F}$. If $\Delta\mathcal{F} \leq 0$ the trial state is always accepted. Otherwise, the trial state is accepted with a probability $P = \exp(-\beta\Delta\mathcal{F})$. If the trial state is accepted, the system is updated to that state whereas in the opposite case the system remains in the initial state. This procedure is repeated until the magnitudes of interest converge to their equilibrium value. When MC simulations are applied to simulate polymer chains, these trial movements typically can include translations of a whole polymer chain, pivot rotations of segments of the chain and small translations of a monomer of the chain (among others).

In SGCMC simulations, in addition of these trial movements, the Metropoli algorithm generates new ionization states of the weak polyelectrolyte by randomly changing the ionization state of the ionizable groups, as outlined in Fig. 2.3. For instance, the algorithm can randomly

pick an ionized group of the polyelectrolyte (A^-) and change it to its protonated (HA) counterpart. Naturally, the reverse case can also occur and the algorithm can change the protonation state of protonated group to its unprotonated state. When salt ions are explicitly taken into account, The electroneutrality is usually maintained by coupling this change of identity with the insertion/deletion of a counter/co-ion. If $\Delta\mathcal{F} > 0$, the probability of accepting the trial state is given by^{19,45}

$$P = \min[\exp(-\beta\Delta\mathcal{F} + \zeta(\text{pH} - \text{p}K_a) \ln 10)], \quad (2.5)$$

where $\text{p}K_a$ is the intrinsic acid constant value of the chosen group and ζ is the extent of the protonation reaction, which corresponds to one step in the protonation ($\zeta = 1$) or the deprotonation ($\zeta = -1$) direction.

SGCMC simulations have been used to rationalize a wide spectra of systems including weak polyelectrolytes such as nanoparticle coating and bridging by weak polyelectrolytes^{46,47}, adsorption of weak polyelectrolytes and proteins on charged surfaces^{29,48–51}, ion binding to ionizable surfaces^{52,53}, weak polyelectrolyte conformational properties¹⁶ and many others¹⁹ (and references quoted therein).

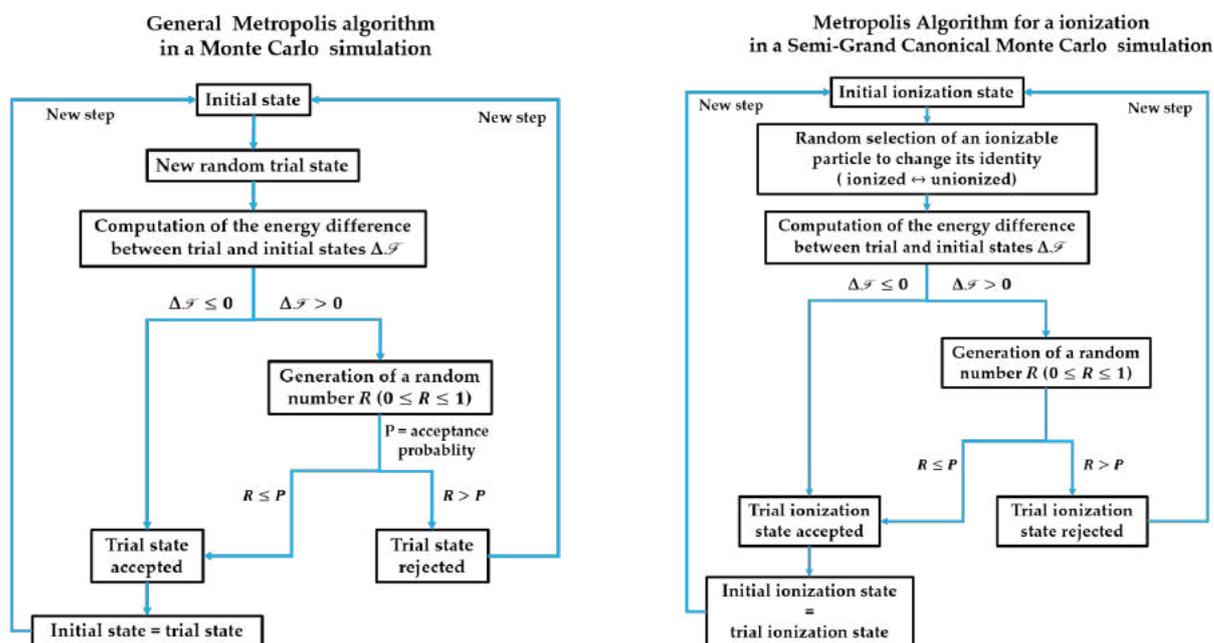


FIGURE 2.3: (left) Scheme illustrating the Metropolis algorithm^{22,40} applied to Monte Carlo. (right) Scheme showing the ionization step in a Semi-Grand Canonical Monte Carlo simulation. In this case, the algorithm proposes a change in the identity of one random ionizable particle (ionized \leftrightarrow unionized) in each realization.

2.4 Brownian dynamics simulation

The diffusion of biopolymers in crowded media has been studied with computational techniques ranging from *on-lattice* Monte Carlo simulations^{100,101} to Molecular Dynamics simulations^{88,114,115}. One popular approach is the use of Brownian Dynamics (BD) simulations^{88,105,116,117}, where the effect of the solvent is implicitly included in the simulation, reducing the computational cost.

BD simulations rely on the classical Langevin equation, which describes the brownian motion of a particle immersed in a fluid¹¹⁸. The Langevin equation reads

$$m \frac{d^2 \vec{r}}{dt^2} = -\vec{\nabla} V(\vec{r}) - k_B T \mathbf{D}^{-1} \frac{d\vec{r}}{dt} + \sqrt{2\mathbf{D}} \vec{\zeta}(t) \quad (2.6)$$

where m is the particle mass, \vec{r} is the particle position, $\vec{\nabla} V(\vec{r})$ is the gradient in potential energy and \mathbf{D} is the diffusion tensor. The last term in Eq. 2.6 mimics the brownian motion of the particle due to the solvent as a gaussian random force. By construction, $\vec{\zeta}(t)$ is a random vector that satisfies $\langle \vec{\zeta}(t) \rangle = 0$ and $\langle \vec{\zeta}(t) \cdot \vec{\zeta}(t') \rangle = \delta(t - t')$, where δ is the Dirac delta function, at any given t time.

In BD simulation, the Langevin equation (Eq. 2.6) is integrated in the so-called overdamped limit, in which one assumes that the particle acceleration is, on average, zero¹⁰⁷. Under this assumption, the integration of the Langevin equation yields¹⁰⁷

$$\vec{r}(t + \Delta t) = \vec{r}(t) - \frac{\Delta t}{k_B T} \mathbf{D} \vec{\nabla} V(\vec{r}) + \sqrt{2\mathbf{D}\Delta t} \vec{\zeta}(t) \quad (2.7)$$

where Δt is the integration time step.

The different BD algorithms can be classified in terms of the treatment of the inter-particle Hydrodynamic Interactions (HI), which are briefly outlined in Fig. 2.4. In the classic BD algorithm proposed by Ermak and McCammon¹¹⁹, HI are neglected and the diffusion tensor \mathbf{D} is considered to be equal to the ideal diffusion coefficient $\mathbf{D} = D_0 \mathbf{I}$ (where \mathbf{I} is the identity matrix). Under this approximation, anomalous diffusion in crowded conditions is still observed due to volume exclusion^{88,117,120}. A natural consequence of neglecting HI is that in these simulations the diffusion coefficient at short times is equal to D_0 , $D^{short} = D_0$.

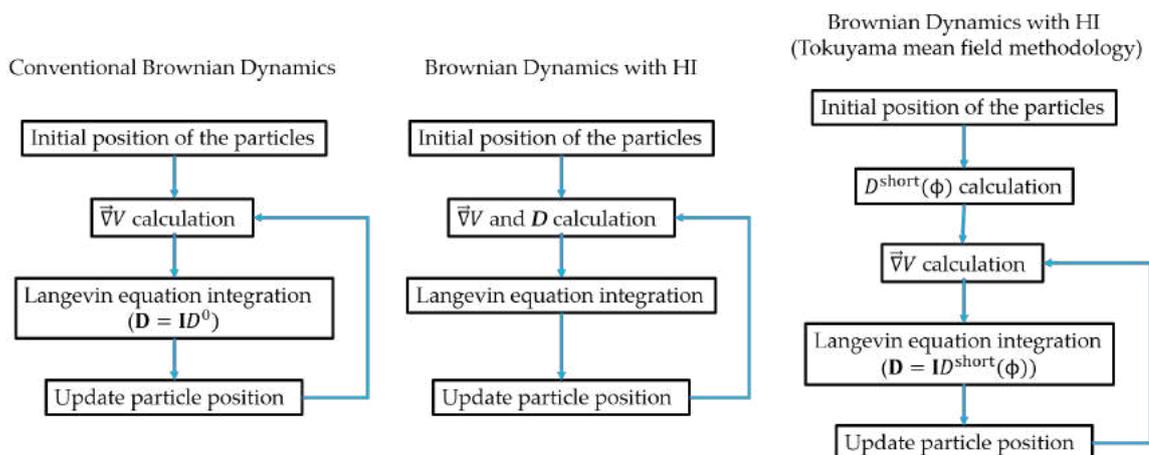


FIGURE 2.4: Outline of the possible algorithms in Brownian Dynamics. In conventional Brownian Dynamics (BD) the diffusion tensor \mathbf{D} calculation is avoided by taking $\mathbf{D} = D_0$, which implies neglecting inter-particle Hydrodynamic Interaction (HI). In BD with HI, \mathbf{D} needs to be computed every time step which is computationally very demanding. Alternatively, HI can be included in BD following Tokuyama mean field approach^{108–110}, which replaces \mathbf{D} by the diffusion coefficient at short times D^{short} calculated with Eq. 1.57 as a function of the volume fraction occupied by the crowders, ϕ .

HI interactions have been reported to play a crucial role in macromolecular diffusion in crowded media^{111,121}. In order to accurately include HI in the BD simulation, one needs to compute \mathbf{D} and its factorization which is a computationally expensive procedure as it scales with the number of particles N as N^3 ¹²². Typically, \mathbf{D} is computed using procedures based on the Rotne-Prager-Yamakawa (RPY) method^{123,124}. It is worth to mention that the RPY method can be extended to include many-body and near-field HI^{125,126} and it has been recently generalized to different-sized particles¹²⁷.

Alternatively, one can include the HI in the BD simulation following the mean-field approach developed by Tokuyama^{108–110}. In this method, HI are included in an effective way by replacing in Eq. 2.7 \mathbf{D} by $D^{\text{short}}\mathbf{I}$, where the diffusion coefficient at short times D^{short} is estimated with Eq. 1.57. Within this procedure, Tokuyama *et al.*¹⁰⁹ have obtained good agreement for the diffusion coefficient at long times (D^{long} , Eq. 1.60) between Brownian Dynamics simulations, Molecular Dynamics with explicit solvent and experimental data.

2.5 Software developed

2.5.1 Semi-Grand Canonical Monte Carlo software

I have programmed a Monte Carlo (MC) code in C++, which allows the simulation of a single linear polymer chain in solution. The polymer can be either a neutral polymer, a strong polyelectrolyte or a weak polyelectrolyte. In the latter case, the program can perform simulations in the Semi-Grand Canonical ensemble, where the pH of the media is fixed (*i.e.* it is an input variable) and the charge of the polyelectrolyte fluctuates (*i.e.* it is an output). The program is prepared, but not restricted, to work with the Site Binding Rotational Isomeric State method. The program allows to specify the geometrical restrictions (bond length, bond angle, torsional energy), charge, hard-sphere radius, pK_a -value and specific first-neighbour electrostatic interaction energy for each unit (site) of the chain. Moreover, the code can deal with any arbitrary distribution of the sites along the chain, which is chosen by the user. Thus, the code is not restricted to periodic polymer chains but can also deal with atactic chains. The program is prepared to simulate not only the behaviour of a polymer in solution but also to perform simulations of single-molecule stretching.

As output, the program returns the polymer average conformational (end-to-end distance, radius of gyration, persistence length, kuhn length and the probability of the bonds of being in a given conformational state), binding (degree of ionization and binding capacitance) and elastic (chain extension) properties. The program allows to perform several MC simulations (with different random seeds) in parallel using OPENMPI libraries, averaging the results over all the simulations. This procedure allows to compute the precision of the measurements, which is estimated as the standard deviation of the results obtained in the different simulations. The program also gives the fluctuation of the measures, calculated as the standard deviation within one MC simulation. All the magnitudes are measured after a certain number of equilibration MC steps, which the user can settle as input. Moreover, a supplementary script (programmed in Python3) can be used to calculation the correlation between conformational states.

The solvent is approximated as a continuum with a dielectrical constant ϵ . The salt ions are implicitly included in the simulations in the ionic strength of the solution I . Both input parameters, ϵ and I , change the electrostatic interaction between the charged sites of the chains, which is calculated with the mean-field Debye-Hückel potential. The bond length and bond angle are simulated as springs with a harmonic potential. The torsional energy of the bonds is calculated with specific energetic parameters for each possible conformation, following the Rotational Isomeric State method. In one MC step, the program attempts to change (i) the ionization state of one binding site, (ii) the rotational state of all bonds, (iii) all bond lengths, (iv)

all bond angles and (v) the spatial orientation of the polymer chain in the laboratory reference frame. The generated trial states are accepted/discarded following the conventional Metropoli algorithm^{22,40}. The program ensures a 20% of acceptance ratio of the trial movements using a self-consistent algorithm. This algorithm measures the acceptance ratio of each trial movement throughout the simulation. If the acceptance ratio is below the 20 %, the algorithm generates trial states more similar to the initial state (the algorithm reduces the differential between states). Otherwise, the algorithm generates states more distant to the initial state (the algorithm increases the differential between states).

2.5.2 Brownian Dynamics software

I have developed a Brownian Dynamics (BD) program in C++, which allows the simulation of macromolecular diffusion in crowded media. The program supports two different models for the macromolecules: the hard-sphere and the Chain Entanglement Soft Potential (CESP) models. The program distinguishes between two kinds of particles: the tracer and the crowders. For each kind of particle, the user can specify their characteristic radii and their ideal diffusion coefficient. As output, the program returns the average mean square displacement (r_{MSD}^2) time profile of each kind of particle. The program allows to perform several Brownian Dynamics simulations in parallel using OPENMPI libraries, averaging the r_{MSD}^2 values over all the simulations.

Several supporting scripts allow to do further analysis of the simulations. I programmed an Octave script that can be used to fit the r_{MSD}^2 temporal profile to the new empirical equation (Eq. ??) proposed in Chapter 7. Another script, programmed in FORTRAN90, allows to compute the radial distribution functions between different kinds of particles of the systems.

The program can perform BD simulations without Hydrodynamics Interactions (HI), with HI using Tokuyama mean-field method¹⁰⁸⁻¹¹⁰ and with HI computing the diffusion tensor \mathbf{D} . When HI are included using Tokuyama mean-field method, the diffusion coefficient at short times D^{short} , necessary as input, can be calculated using a supplementary script provided with the code. The scripts solve Tokuyama mean-field equations (Eqs. 1.57 and 1.60) and compute D^{short} and D^{long} as a function of the excluded volume of the system. When HI are included calculating the diffusion tensor, the conventional Rotne-Prager-Yamakawa (RPY) tensor^{123,124} is calculated at every time step and it is checked to be always definite positive. When different-sized particles are simulated in the system, the RPY tensor is corrected using the procedure explained in Ref.¹²⁷ Periodic boundary conditions are applied in all the directions of space and the particles interactions are calculated using the minimum image convention.

Chapter 3

Coupling of Charge Regulation and Conformational Equilibria in Linear Weak Polyelectrolytes: Treatment of Long Range Interactions *via* Effective Short-Ranged and pH-Dependent Interaction Parameters

3.1 Summary

3.1.1 Introduction

The ionization state of weak polyelectrolytes in solution is a dynamic phenomenon due to the binding of proton, metals and other chemical species to the macromolecule skeleton. The ability of weak polyelectrolytes to modulate its charge in response of medium conditions (pH, ionic strength) or external perturbations, known as Charge Regulation (CR), is fundamental to understand the physico-chemistry of charged macromolecules in multiple systems.^{24,32,140,143,145,148} Although CR can occur in rigid structures²² (e.g. nanoparticles or charged surfaces), weak polyelectrolytes frequently possess a flexible structure with significant thermal fluctuations. In a flexible weak polyelectrolyte a conformational change can modify its ionization state. For instance, if the conformational change brings together two equally-charged groups of the polyelectrolyte chain, the electrostatic repulsion between them can change the ionization state of these groups. In turn, changes in the ionization state of the weak polyelectrolyte can produce a conformational change of the macromolecule. For example, the ionization of two consecutive binding sites of the macromolecule with a charge of equal sign promotes extended conformations, which reduce the electrostatic repulsion between the sites. In summary, ionization and conformational degrees of freedom are highly coupled in flexible weak polyelectrolytes. Several examples of the conformation-ionization interplay in synthetic polyelectrolytes and biopolymers can be found in the literature.^{33–37}

Theoretically, the ionization of weak polyelectrolytes has been mainly studied using the Site Binding (SB) model.^{20,21,23,25,26,39,40} In this model, the macromolecule is described as a set of binding sites, which have two possible states: ionized or neutral. The conformational degrees of freedom are implicitly included in the binding parameters, which should be understood as proper averages over all the possible conformations of the polyelectrolyte. Recently, the SB model has been combined with the Rotational Isomeric State (RIS) model. The resulting method, named Site Binding Rotational Isomeric State (SBRIS) model, allows to study the ionization and conformational degrees of freedom on the same foot.¹⁴ The SBRIS model has been used in a recent work to analyze the conformational and ionization degrees of freedom of polyethylene(imine).¹⁶

For a linear polyelectrolyte chain, the SB, RIS and SBRIS can be solved using the transfer matrix method. This method relies in a recursive relation between the partition function of a system with N ionizable sites with the one of a system with $N - 1$ sites.^{9,10} Once the partition function of the system is calculated, one can obtain the thermal averages of the properties of interest (end-to-end distance, bond state probabilities, average degree of ionization and many

others). As a result, the transfer matrix method has been adapted to study many different phenomena^{20,20,23,42,43}. The principal restriction of the transfer matrix method is that it can only consider Short Range (SR) interactions. This is an important limitation when dealing with polyelectrolytes, since electrostatic interactions are long-ranged. This limits the utility of the transfer matrix to high ionic strength conditions, where electrostatic interactions are screened and Long Range (LR) interactions can be safely neglected.

In a recent work¹⁷, a new method has been proposed to solve the SB model including the LR interactions within the transfer matrix framework. The method, named Local Effective Interaction Parameters (LEIP), modifies the SR free energy to include the LR interactions in an effective way. In this way, the new effective SR parameters (protonation energies, pair interaction energies, etc.) can be calculated using Gibbs–Bogoliubov variational principle¹⁶³, which quickly converges to the exact solution. In this study, the results obtained with LEIP method and with Monte Carlo simulation exhibited an almost exact agreement. However, the LEIP method was only studied in rigid polyelectrolytes, where the conformational degrees of freedom are absent.

The objective of this chapter is to extend the LEIP method to solve the SBRIS model, *i.e.* to the study of flexible weak polyelectrolytes. The new treatment includes not only effective ionization parameters but also effective conformational parameters. The details of the theoretical derivation can be found in the publication enclosed at the end of the chapter. The new method is studied in a model linear polyelectrolyte, which has an ionizable site (blue) every three chain positions as shown in Fig. 3.1. Only the bonds joining two non-binding neutral sites (grey) are allowed to rotate, which can be in the conformational states *trans*, *gauche+* and *gauche-*. When using the LEIP method, the torsional energy of these rotating bonds is replaced by an effective rotational energy, as outlined Fig. 3.1. This effective rotational energy includes the LR interaction energy and it is dependent on the pH and ionic strength of the medium. The analytical resolution of the LEIP method is compared with Semi-Grand Canonical Monte Carlo (SGCMC) simulations, which allows to assess the accuracy of LEIP method when applied to flexible polyelectrolytes.

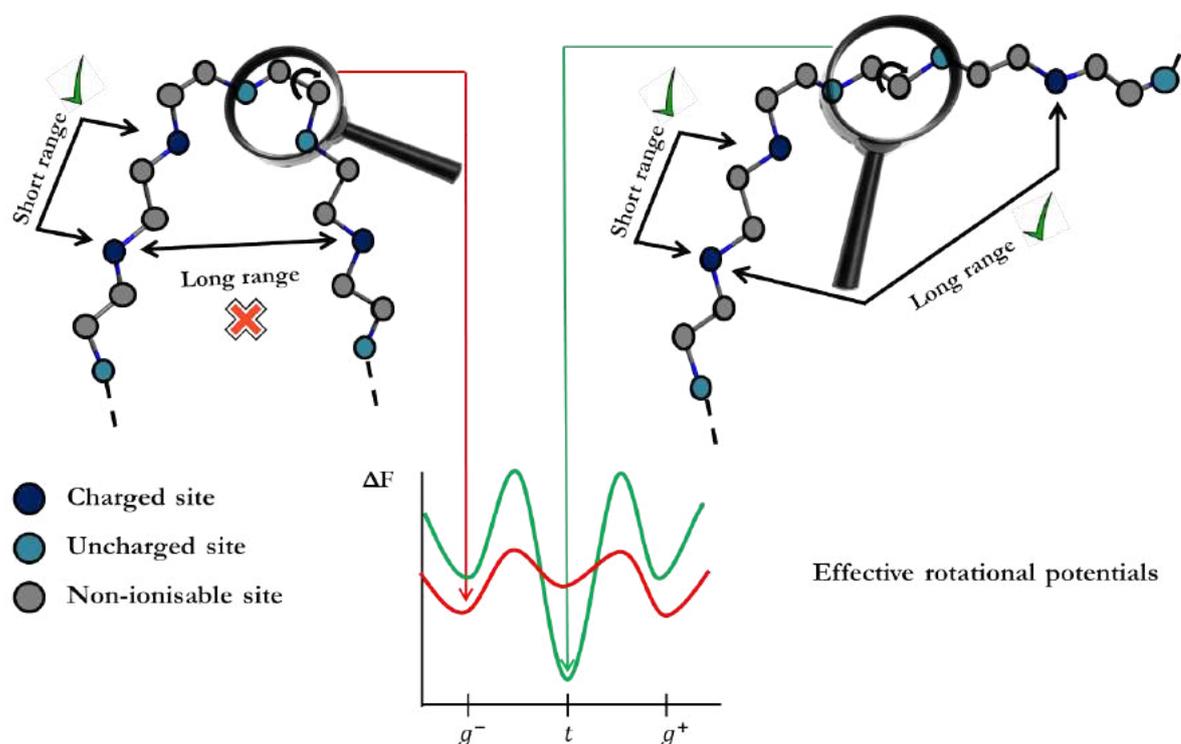


FIGURE 3.1: Outline of the model polyelectrolyte used in the publication. The polyelectrolyte has one ionisable site (blue) every three chain positions. Only the bonds holding two non-ionisable sites (grey) are allowed to rotate. The LEIP method is used to analytically solve the SBRIS model, which accounts for the conformational and ionization degrees of freedom on the same foot. In this way, the rotational energy of the bonds becomes effective and dependent on the long range electrostatic interaction of the charged sites.

3.1.2 Results

The model weak polyelectrolyte used in the publication has the following assumptions:

- Only proton binding is considered. All binding sites are equal (*i.e.* all sites have the same protonation constant $pK_a = 9$). The sites are charged when protonated and neutral otherwise.
- Only the bonds holding two non-ionisable sites (from now on *c* bonds) are allowed to rotate. The bond rotations are independent when the macromolecule is uncharged.
- All *c* bonds are equal and can only be in the conformational *trans*, *gauche+* and *gauche-*. Both *gauche* states are symmetrical and have a torsional energy of the *trans* \longleftrightarrow *gauche* rotation is given by σ .

- The nearest neighbour electrostatic interactions are characterized by the energies ϵ_t (when the c bond is in *trans*) and ϵ_g (when the c bond is in *gauche*).
- co- and counter ions are only present through the Debye-Hückel potential describing the long-range electrostatic interactions. Excluded volume effects are only present through the electrostatic inter-site repulsion.

The model can be regarded as a "minimal" model of a flexible weak polyelectrolyte with only four energetic parameters (pK_a , ϵ_t , ϵ_g and σ). The conformational (probability of a bond of being in a *gauche* state) and ionization (average degree of ionization θ) properties of this model polyelectrolyte are calculated with the LEIP method and SGCMC simulations. Here, the only the case with $\epsilon_t = 1$, $\epsilon_g \rightarrow \text{inf}$ and $\sigma = 10$ are shown, although a more complete analysis can be found in the publication enclosed at the end of the chapter. The case with $\epsilon_g = \text{inf}$ and $\sigma = 10$ is interesting because it implies that, for a given bond **c**, its *gauche* states are promoted when it is holding at least one uncharged site (for instance because of hydrogen bonding) but are they are forbidden otherwise. As a result, in this case the propensity of the conformational change with the ionization of the sites, *i.e.* with the pH-value.

The titration curve obtained at different ionic strength (I) values ranging from 1 M to 0.001 M (from top to bottom) can be observed in Fig. 3.2. The red dot markers depicts the results obtained from SGCMC simulation while the black continuous lines represent the prediction obtained with LEIP method correcting pK_a and σ . It can be observed that the LEIP method reproduces almost exactly the SGCMC results for all the range of pH and I values. The dashed line represents the titration curve at $I = 0.001$ M obtained with the LEIP method when only pK_a is corrected. Comparing the results with the corresponding ones obtained with SGCMC simulation, a clear deviation is observed. This indicates that it is necessary to correct the effect of the LR interaction in the torsional energy in order to properly reproduce the SGCMC results. In this dramatic case, neglecting the coupling between ionization and conformational degrees of freedom results in a phase transition-like artifact at $\text{pH} \simeq 5$.

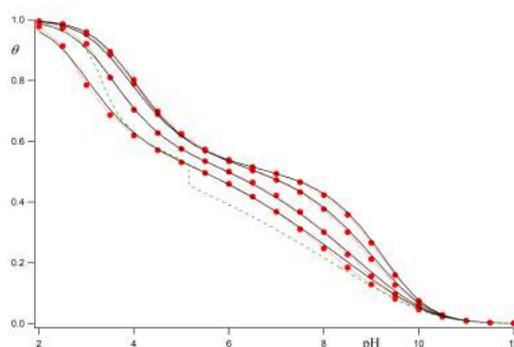


FIGURE 3.2: Average degree of ionization θ as a function of the pH-value for a SBRIS model linear polyelectrolyte with $pK = 9$, $\epsilon_t = 1$, $\epsilon_g \rightarrow \infty$ and $\sigma = 10$. From bottom to top the ionic strength values are 0.001 M, 0.01 M, 0.1 M and 1 M. Red markers represent the results obtained by SGC MC simulation while black continuous lines is the solution of the LEIP method with pK and $p\sigma$ corrected. The green dashed line depicts the LEIP values obtained when only pK is corrected, for the case with $I = 0.001$ M.

The probability of a rotating bond to be in *gauche* state as function of the pH-value is depicted in Fig 3.3 at different I -values: 0.001 M (triangles and dashed line), 0.01 M (squares and dotted line) and 1 M (circles and continuous line). Again, the red markers correspond to the values obtained by SGC MC simulation and the black lines to the LEIP values correcting both pK_a and σ . A very good agreement between the LEIP prediction and the SGC MC results is obtained. Two asymptotic behaviors can be observed. At low pH-values, the *gauche* probability tends to zero while at high pH-values it tends to almost 1. This is because at high pH-values the polyelectrolyte is almost uncharged and the *gauche* states are energetically promoted. When the pH-value is decreased, the polyelectrolyte is, in turn, charged. The electrostatic repulsive interactions cause the polyelectrolyte to be on average more extended, which promotes *trans* conformations where the sites are more separated between them.

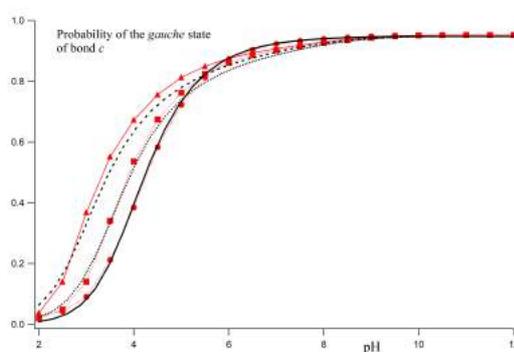


FIGURE 3.3: Probability of the *gauche* state as a function of the pH-value for a SBRIS model linear polyelectrolyte with $pK = 9$, $\epsilon_t = 1$, $\epsilon_g \rightarrow \infty$ and $\sigma = 10$ at different I -values: 0.001 M (triangles and dashed line), 0.01 M (squares and dotted line) and 1 M (circles and continuous line). Again, the red markers correspond to the values obtained by SGC MC simulation and the black lines to the LEIP values correcting both pK_a and σ .

The effect of the long range interactions in the rotational properties of the polyelectrolytes can be better analyzed in terms of the LEIP correction to the σ , x_σ . x_σ can be understood as the average energy contribution to the torsional energy of the \mathbf{c} bonds due to LR electrostatic interaction. In Fig. 3.4, x_σ is plotted as a function of the pH-value at different I values ranging from 0.001 M to 1 M (from top to bottom). It can be appreciated that x_σ increases when I increases. This is because at low I -value the electrostatic repulsive are less screened, thus being the electrostatic LR influence in σ is more important. It can also be observed that at high pH-values x_σ tends to zero, since the polyelectrolyte is almost uncharged.

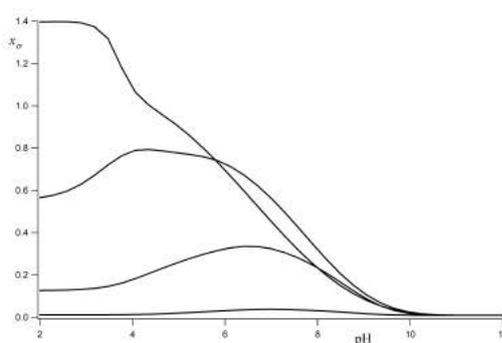


FIGURE 3.4: LEIP correction to torsional energy of bonds \mathbf{c} x_σ as a function of the pH-value. From top to bottom the I values are 0.001 M, 0.01 M, 0.1 M and 1 M.

3.1.3 Conclusions

In this chapter, the Local Effective Interaction Parameters (LEIP) method is extended to calculate the conformational and ionization properties of a flexible weak polyelectrolyte using the Site Binding Rotational Isomeric State (SBRIS) model. This method was recently developed to study the ionization properties of a rigid polyelectrolyte in the full ionic strength range. The LEIP theoretical prediction for a flexible polyelectrolyte is validated against Semi-Grand Canonical Monte Carlo (SGCMC) simulations.

For polyelectrolytes with different energetic parameters, the degree of ionization and the *gauche* state probabilities are computed both with LEIP method and SGCMC simulation. For all the cases of study, a very good agreement is observed in both magnitudes. The computational cost associated is, however, is orders of magnitude lower in LEIP case. As a result, it could be used to fit experimental titration data, obtaining conformational and binding information. This information could be used, for instance, to develop pH-dependent force-fields based on experimental data. Moreover, the LEIP method could also be extended to study the competitive binding of metal ions in weak polyelectrolytes.

3.2 Publication

Article

Coupling of Charge Regulation and Conformational Equilibria in Linear Weak Polyelectrolytes: Treatment of Long-Range Interactions via Effective Short-Ranged and pH-Dependent Interaction Parameters

Pablo M. Blanco ^{1,*} , Sergio Madurga ^{1,*} , Francesc Mas ¹  and Josep L. Garcés ²

¹ Physical Chemistry Unit, Department of Materials Science and Physical Chemistry & Research Institute of Theoretical and Computational Chemistry (IQTCUB) of Barcelona University (UB), 08028 Barcelona, Catalonia, Spain; fmas@ub.edu

² Department of Chemistry, Technical School of Agricultural Engineering & Agrotecnio of Lleida University (UdL), 25003 Lleida, Catalonia, Spain; jlgarces@quimica.udl.cat

* Correspondence: pmblanco@ub.edu (P.M.B.); s.madurga@ub.edu (S.M.)

Received: 15 June 2018; Accepted: 23 July 2018; Published: 24 July 2018



Abstract: The classical Rotational Isomeric State (RIS) model, originally proposed by Flory, has been used to rationalize a wide range of physicochemical properties of neutral polymers. However, many weak polyelectrolytes of interest are able to regulate their charge depending on the conformational state of the bonds. Recently, it has been shown that the RIS model can be coupled with the Site Binding (SB) model, for which the ionizable sites can adopt two states: protonated or deprotonated. The resulting combined scheme, the SBRIS model, allows for analyzing ionization and conformational equilibria on the same foot. In the present work, this approach is extended to include pH-dependent electrostatic Long-Range (LR) interactions, ubiquitous in weak polyelectrolytes at moderate and low ionic strengths. With this aim, the original LR interactions are taken into account by defining effective Short-Range (SR) and pH-dependent parameters, such as effective microscopic protonation constants and rotational bond energies. The new parameters are systematically calculated using variational methods. The machinery of statistical mechanics for SR interactions, including the powerful and fast transfer matrix methods, can then be applied. The resulting technique, which we will refer to as the Local Effective Interaction Parameters (LEIP) method, is illustrated with a minimal model of a flexible linear polyelectrolyte containing only one type of rotating bond. LEIP reproduces very well the pH dependence of the degree of protonation and bond probabilities obtained by semi-grand canonical Monte Carlo simulations, where LR interactions are explicitly taken into account. The reduction in the computational time in several orders of magnitude suggests that the LEIP technique could be useful in a range of areas involving linear weak polyelectrolytes, allowing direct fitting of the relevant physical parameters to the experimental quantities.

Keywords: polyelectrolytes; charge regulation; long-range interactions; Debye–Hückel interactions; transfer matrix; Ising models; semi-grand canonical ensemble; Monte Carlo simulations; conformational equilibria; variational methods

1. Introduction

The ionization state of charged macromolecules in solution is regulated by the binding of small ions (protons, metal ions, etc.) present in the backward medium. In particular, acid-basic equilibria in weak polyelectrolytes represent the paradigmatic mechanism of charge regulation due to the ubiquitous presence of proton ions in aqueous solution. These processes are of paramount importance

to understand the physicochemical behavior of charged macromolecules in a wide range of situations. Just to mention a few examples, charge regulation plays a fundamental role in receptor–ligand equilibria in biochemical systems [1–4], supramolecular chemistry [5–7], the role of natural organic matter in geochemical cycle of metal ions [8], wastewater treatment [9], stability of colloidal systems [10], advanced coating in material science [11–13] or drug delivery [14]. Charge regulation can take place on rigid structures, such as surfaces or nano-particles [15], but in general polyelectrolytes are flexible and conformational and ionization degrees of freedom are strongly coupled. This fact can result in dramatic structural changes in the macromolecule. Classical examples are the helix–coil transitions of poly(peptides) [16], the swelling of poly(methacrylic) acid in a very narrow range of pH [17] or the strong influence of ionization in the folding of proteins [18]. More recently, the importance of the ionization configuration in the conformational properties of intrinsically disordered proteins, whose function–structure relationship still remains a controversial matter, has been recognized [19,20].

The understanding of the ionization processes has been mainly based on the so-called Site Binding (SB) model. In this approach, the ionization configuration of the macromolecule is defined as a set of sites which can be in two possible states, i.e., protonated or deprotonated, as outlined in Figure 1a. The free energy is then parametrized by site-specific microscopic protonation constants and interaction energies between sites. Triplet or higher-order interactions among sites can also be considered. Once the system is parametrized, the machinery of statistical mechanics can be used in order to quantify the relevant physical properties such as titration curves, site-specific binding probabilities, macroscopic protonation constants, microscopic protonation enthalpies, site–site binding correlation, etc. [5,15,21–25]. For systems with a small number of sites ($N \leq 20$), the necessary thermal averages can be performed by direct enumeration, while, for a large number of sites, Monte Carlo (MC) simulations become necessary [26–37].

In the case of linear polyelectrolytes, the transfer matrix method can be used to compute the relevant thermal averages [15]. This powerful and elegant technique was originally designed to solve the classical Ising model of ferromagnets. It is based on the fact that the partition function of a system with $N + 1$ sites can be related in a recursive way to the one of a system with N sites. The resulting recursive relationship can be expressed in terms of the transfer matrix, whose elements represent the contributions of the new site to the partition function, for a given state of the preceding site [38,39]. The technique is very versatile and can be generalized to systems composed with repetitive units (spins, bonds or binding sites), which can take in principle more than two states. When applied to the binding of ions to polyelectrolytes, the method can be adapted to include a wide range of phenomena such as triplet interactions between sites [21], chelate complexation of metal ions [23], proton binding to polyampholytes [40,41], protein–DNA binding [42], super-capacitor charging [43] or coupling between ionization and conformational degrees of freedom [44–46].

Probably the most productive application of transfer matrices was proposed by Flory in the context of the Rotational Isomeric State (RIS) model [47,48], aiming to compute conformational properties of neutral linear molecules. The RIS model relies on the observation that, although a particular bond can adopt in principle any rotation angle, only those of minimum energy (typically *trans*, *gauche+* and *gauche-*) are significantly populated. As a consequence, each bond can be regarded as a ‘unit’ of the system adopting three possible states. The corresponding partition function and the necessary thermal averages (bond probabilities, end-to-end distance, radius of gyration, etc.) can be calculated using a proper product of transfer matrices. In recent works [45,46], it has been shown that SB and RIS models can be combined in a unique scheme so that conformational and ionization equilibria can be analyzed on the same foot. It has been shown that all the matrixial expressions of RIS can be systematically extended to account for the ionization degrees of freedom. The resulting SBRIS model, outlined in Figure 1b, has been recently applied to the detailed characterization of the conformational and ionization properties of linear poly(ethylene)imine [46].

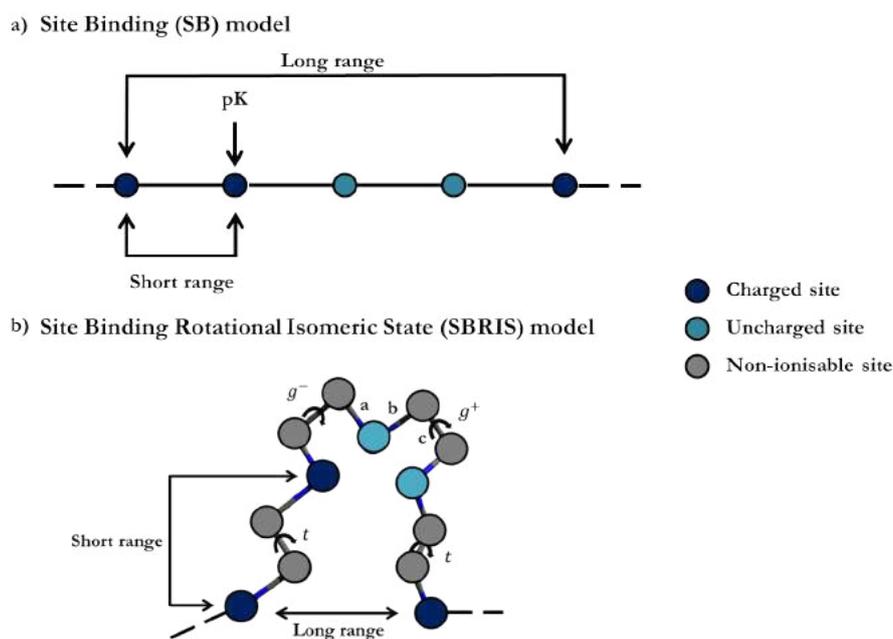


Figure 1. (a) Outline of the Site Binding (SB) model for a linear polyelectrolyte represented as a linear chain of ionizable sites. The ionization state of the macromolecule is characterized by a set of variables $s = \{s_i\}$, $i = 1, \dots, N$, which can adopt two possible values: $s_i = 1$ if the site i is protonated (dark blue circles) and $s_i = 0$ if it is deprotonated (cyan circles); (b) sketch of a linear chain joining ionizable sites by means of rotating bonds as an example of the Site Binding-Rotational Isomeric State (SBRIS) model. Both ionization and conformation degrees of freedom are now taken into account. In the depicted chain, only the bonds type c are able to rotate, which can take three possible states of minimum energy, i.e., *trans*, *gauche+* and *gauche-*.

The main limitation of the transfer matrices used in SB, RIS and SBRIS models is that they can only deal with Short-Range (SR) interactions [49,50]. SR interactions are chemically specific and can produce important correlations between neighbouring sites and bonds. They cannot be modeled by simple continuous force fields (such as van der Waals or Debye–Hückel potentials) [51] but, in exchange, they can be easily implemented in a transfer matrix scheme. For polyelectrolytes, however, this is an important restriction, due to the Long-Range (LR) nature of coulombic interactions, which severely restricts the range of application of the transfer matrix approach. In practice, the possibility of neglecting LR coulombic interactions must be restricted to high ionic strengths, an important limitation specially for polyelectrolytes which become insoluble under such conditions [52–55].

In a recent paper [56], the SB model has been extended to include LR interactions by introducing a modified free energy involving Local Effective Interaction Parameters (LEIP), which account for the LR interactions in an effective way. The LR force field is thus replaced by a short-ranged effective one. The new local effective parameters, i.e., effective protonation free energies, effective pair interactions and so on, can be systematically calculated by using the Gibbs–Bogoliubov variational principle [39]. The resulting modified free energy converges very fast to the exact free energy. It was found that the correction to the site protonation pK (first order correction) is enough to obtain an excellent, exact from a practical point of view, agreement between theory and MC simulations. This previous study, however, was restricted to rigid molecules, and conformational degrees of freedom were not taken into account.

The main goal of the present work is to extend the LEIP method to account for the coupling between charge regulation and conformational equilibria involving LR interactions. In addition to allow much faster computations of ionization/conformational properties (computational times are reduced in orders of magnitude), the methodology here presented adds new physical insight in the

interplay of conformational and ionization degrees of freedom in polymeric structures. For instance, the energy of the *gauche* state of a bond will now depend on the pH and the ionic strength, even if such a bond does not hold any ionizable group. The use of the LEIP technique in the SB model is reviewed in Section 2. In Section 3, the technique is generalized in order to include conformational equilibria coupled to LR coulombic interactions represented by the Debye–Hückel potential. In Section 4, semi-grand canonical Monte Carlo simulations are introduced as a tool to test LEIP accuracy when applied to flexible polyelectrolytes. In Section 5, LEIP theoretical results are compared to MC simulations. The new ideas here introduced are illustrated with a minimal model of a flexible linear weak polyelectrolyte containing only one type of rotating bond.

2. Simultaneous Treatment of Short- and Long-Range Interactions in Rigid Molecules

The ionization state of a macromolecule with N ionizable sites can be characterized by a set of variables $s = \{s_i\}$, $i = 1, \dots, N$, which can adopt two possible values: $s_i = 1$ if the site i is protonated, and $s_i = 0$ if it is deprotonated. The corresponding reduced free energy can be expressed in terms of the variables s_i by means of the so-called cluster expansion [24]

$$\frac{H(s)}{\ln 10} = \sum \mu_i s_i + \sum_{i>j} \phi_{ij} s_i s_j + \sum_{i>j>k} \tau_{ijk} s_i s_j s_k + \dots, \quad (1)$$

where $\mu_i = \text{pH} - \text{p}K_i = -\log(K_i a_H)$ is the reduced chemical potential, which depends on the proton activity, a_H , and the protonation $\text{p}K$ -value of the ionizable site i , $\text{p}K_i$; ϕ_{ij} represents the interaction energy of the sites i and j ; τ_{ijk} accounts for possible triplet interactions among sites i , j and k , and so on. The term “reduced” refers to the fact that the chemical potential incorporates both the pH and the protonation $\text{p}K$, which simplifies the subsequent expressions. The interaction (or cluster) parameters are expressed in thermal units, i.e., $\beta = 1/k_B T = 1$, and divided by a factor $\ln 10$ in order to be compared in the pH scale. Note that the conformation degrees of freedom are omitted in Equation (1), so that the interaction parameters should be understood as proper averages over the conformational states. The mathematical form of these averages is not trivial and expressions for them are given in [45]. Throughout this work, we will assume that a site is charged when it is protonated, i.e., we are dealing with poly-cations. However, the subsequent arguments are also applicable to poly-anions with a suitable change in the protonation variables [15]. The expansion of the free energy (1) usually converges very fast to the exact free energy, and, for most of the cases, the inclusion of triplet interactions is enough to accurately reproduce the measurable quantities, such as the degree of ionization of the individual sites [22]. These can be obtained from $H(s)$ by means of the semi-grand canonical partition function

$$\Xi = \sum_s e^{-H(s)}. \quad (2)$$

The average degree of protonation of a particular site i is related to Ξ as

$$\theta_i = \langle s_i \rangle = -\frac{\partial \log \Xi}{\partial \mu_i} = \frac{1}{\ln 10} \frac{\partial \Omega}{\partial \mu_i}, \quad (3)$$

where $\Omega = -\ln \Xi$ is the thermodynamic potential associated with the semi-grand canonical ensemble. The average number of bound protons is given by

$$\nu = \left\langle \sum_i s_i \right\rangle = \frac{\partial \Omega}{\partial \ln a_H}. \quad (4)$$

The correlation of the protonation degrees of two sites i and j , a quantity which will be used later, can be expressed as

$$h_{ij} = \langle s_i s_j \rangle = -\frac{1}{(\ln 10)^2} \frac{\partial \Omega}{\partial \mu_i \partial \mu_j}. \quad (5)$$

As can be seen, the quantification of all the relevant physical quantities relies in the accurate determination of the partition function Ξ . If the number of sites is small ($N \leq 20$), Ξ can be evaluated by direct enumeration of all the possible ionization states. Otherwise, Monte Carlo (MC) simulations must be performed. In some cases, however, methods borrowed from Statistical Mechanics can be used. Among them, probably the most elegant one is the transfer matrix method, consisting of relating the partition function for a system with $N + 1$ sites with that with N sites in a recursive way. This method was firstly used in the exact solution of the Ising model of ferromagnets [38,39]. The link between both partition functions is the transfer matrix whose elements are the Boltzmann factors corresponding to the increase in the reduced free energy. For instance, for the linear polyelectrolyte sketched in Figure 1a and assuming only nearest neighbour interactions, the partition function can be expressed as [47]

$$\Xi = \mathbf{q}\mathbf{T}^N\mathbf{p}^T, \tag{6}$$

where \mathbf{T} is the transfer matrix

$$\mathbf{T} = \begin{pmatrix} 1 & z \\ 1 & zu \end{pmatrix}. \tag{7}$$

$z = Ka_H$ represents the reduced activity and $\epsilon = -\log u$ is the interaction free energy between neighbouring sites. $\mathbf{q} = (1, 0)$ and $\mathbf{p} = (1, 1)$ are the initiating and terminating vectors. This would be the simplest use of the transfer matrix.

The main limitation of the transfer matrix methods is that they can be only used when Long-Range (LR) interactions are neglected, since the size of the transfer matrices grows exponentially with the range of the interactions [50]. This is an important limitation of the method when dealing with polyelectrolytes, since it can be only used at high enough ionic strengths, for which the screening is enough to avoid the LR interactions. In a recent paper [56], we introduced a method which allows for including the LR interactions in a very accurate way. In this approach, the full free energy Equation (1) is replaced by a new one involving only Short-Range (SR) interaction parameters, accounting for the LR interactions in an effective way. The resulting formalism deals with both SR and LR interactions simultaneously. The method can be used for any kind of molecular or surface geometry, but it is restricted to rigid structures, so that the conformational degrees of freedom are not explicitly taken into account. Since the main goal of this work is to extend this formalism to flexible molecules and polyelectrolytes, we briefly outline the basic ideas of the method. The details of the derivations are given in Reference [56]. Although the following arguments can be readily generalized to the general form of the free energy Equation (1), let us consider the simplest case of a rigid linear chain with identical sites, such as the one sketched in Figure 1a. For this system, $\mu_1 = \mu_2 = \dots = \mu$ and triplet interactions are omitted, i.e., $\tau_{ijk} = 0$. The reduced free energy H can be split into two contributions $H = H_0(x) + \Delta H(x)$ such as

$$\begin{aligned} \frac{H_0}{\ln 10} &= (\mu - x) \sum_i s_i + \epsilon \sum_i s_i s_{i+1}, \\ \frac{\Delta H}{\ln 10} &= \sum_{j>i+1} \phi_{ij} s_i s_j + x \sum_i s_i, \end{aligned} \tag{8}$$

where x is a parameter to be determined. Note that H_0 corresponds to a reduced free energy containing only nearest neighbour interactions of energy $\epsilon = \phi_{i,i+1}$, which can be exactly solved by using the transfer matrix (7). Now, we can use the Gibbs–Bogoliubov variational principle [39,57]

$$\Omega \leq \tilde{\Omega} = \Omega_0(x) + \langle \Delta H(x) \rangle_0 \tag{8}$$

to determine the optimal value of x , where $\Omega_0(x) = -\ln \Xi_0$ and $\langle \dots \rangle_0$ represent the free energy and

the thermal average corresponding to H_0 , respectively. Minimizing $\tilde{\Omega}$ with respect to x , it is found that x fulfills the equation [56]

$$x = \frac{d\varphi_0/dx}{dv_0/dx} = \frac{d\varphi_0}{dv_0}, \quad (9)$$

where

$$\varphi_0 = \left\langle \sum_{j>i+1} \phi_{ij} s_i s_j \right\rangle_0 = \sum_{j>i+1} \phi_{ij} h_{ij}^0 \quad (10)$$

is the LR energy averaged over the unperturbed free energy H_0 , whose correlation function h_{ij}^0 , can be exactly evaluated using (5). If the optimal value for x is used in the computations, the variational principle (8) implies that all the thermal averages (degree of protonation, correlation functions, etc.) can be obtained replacing the average $\langle \dots \rangle$ by $\langle \dots \rangle_0$, which can be exactly determined since only SR interactions are involved. Equation (9) provides a transparent physical interpretation of x : it is the average change in the LR interaction energy when a new proton is bound to the molecule at a given pH-value. As expected, x vanishes in the absence of LR interactions and the nearest neighbour interaction model becomes exact. Therefore, x can be interpreted as the necessary correction to the reduced chemical potential μ in order to account for the LR interactions but in a local effective way. By the definition of $\mu = \text{pH} - \text{pK}$, x can also be understood as the correction to the site pK-value, so that $\text{pK}^{\text{eff}} = \text{pK} - x$ is the effective pK-value, and it represents the extra energetic cost of the site protonation due to the presence of LR interactions. We will refer to this procedure as the Local Effective Interaction Parameters (LEIP) method. It is important to highlight that LEIP, unlike other approaches involving some mean-field approximation (such as the Bragg–Williams approximation in Ising models), includes the correlations via Equation (10), although in an approximate way. This approximation, however, results in being extremely accurate, as can be observed in Figure 2a, where the titration curves corresponding to a rigid linear chain with identical sites are depicted. The chosen parameters are $\text{pK} = 9$ and $\epsilon = 1.5$. In this model, the LR interactions between distant sites are described by the Debye–Hückel potential

$$\phi_{ij} = \frac{1}{\ln 10} \frac{\ell_B e^{-\kappa d_{ij}}}{d_{ij}} ; \quad j > i + 1, \quad (11)$$

where $\ell_B \simeq 0.7$ nm is the Bjerrum length in water at 298 K, d_{ij} is the distance between the sites i and j , and κ^{-1} (nm) = $0.304/\sqrt{I(\text{M})}$ is the Debye length at the ionic strength I . For a rigid linear chain, as the one shown in Figure 1, $d_{ij} = |j - i|b$, where b is the separation between consecutive protonating sites. We have plotted the titration curves obtained by Monte Carlo (MC) simulations in the semi-grand canonical ensemble, i.e., at constant pH (blue circles), together with the ones calculated using Equations (9) and (10) (continuous lines) for all the range of ionic strengths and $b = 0.2$ nm. Surprisingly, simulated and calculated curves overlap, so that, for this model, the LEIP solution can be regarded as exact from the practical point of view. The computational cost of LEIP methods is many orders of magnitude lower than that required in MC simulations, allowing the fitting of parameters to experimental titration curves. The correction to the pK, x , shown in Figure 2b, increases in lowering the pH (i.e., increasing the charge), and in decreasing the ionic strength (lower electrostatic screening), since the energetic cost to protonate a site increases with the macromolecular charge and with the intensity of the LR interactions.

Another advantage of the LEIP method is that it can be systematically improved by selectively correcting other cluster parameters. For instance, one could decide to correct, not only the pK-value

($pK \rightarrow pK - x$), but also the nearest neighbour interaction energy ($\epsilon \rightarrow \epsilon + x_\epsilon$). Proceeding in the same way, it can be shown that x and x_ϵ fulfill the nonlinear system of equations [56]

$$\mathbf{J} \begin{pmatrix} x \\ x_\epsilon \end{pmatrix} = \begin{pmatrix} \left(\frac{\partial \varphi_0}{\partial x}\right)_{x_\epsilon} \\ \left(\frac{\partial \varphi_0}{\partial x_\epsilon}\right)_x \end{pmatrix}; \mathbf{J} = \begin{pmatrix} \left(\frac{\partial v_0}{\partial x}\right)_{x_\epsilon} & \left(\frac{\partial D_0}{\partial x}\right)_{x_\epsilon} \\ \left(\frac{\partial v_0}{\partial x_\epsilon}\right)_x & \left(\frac{\partial D_0}{\partial x_\epsilon}\right)_x \end{pmatrix}, \quad (12)$$

where $v_0(x, x_\epsilon)$ and $D_0 = \langle s_i s_{i+1} \rangle_0 = h_{12}^0(x, x_\epsilon)$ represent the average number of protons and the average number of nearest neighbour interactions, respectively, which can be exactly calculated using H_0 . Solving Equation (12), the correction to the pK and ϵ are obtained as functions of the pH. The physical meaning of x and x_ϵ becomes transparent if Equation (12) are rewritten in terms of v_0 and D_0 as independent variables. After some elementary algebra, x and x_ϵ adopt the much simpler form

$$x = \left(\frac{\partial \varphi_0}{\partial v_0}\right)_{D_0}; x_\epsilon = \left(\frac{\partial \varphi_0}{\partial D_0}\right)_{v_0}. \quad (13)$$

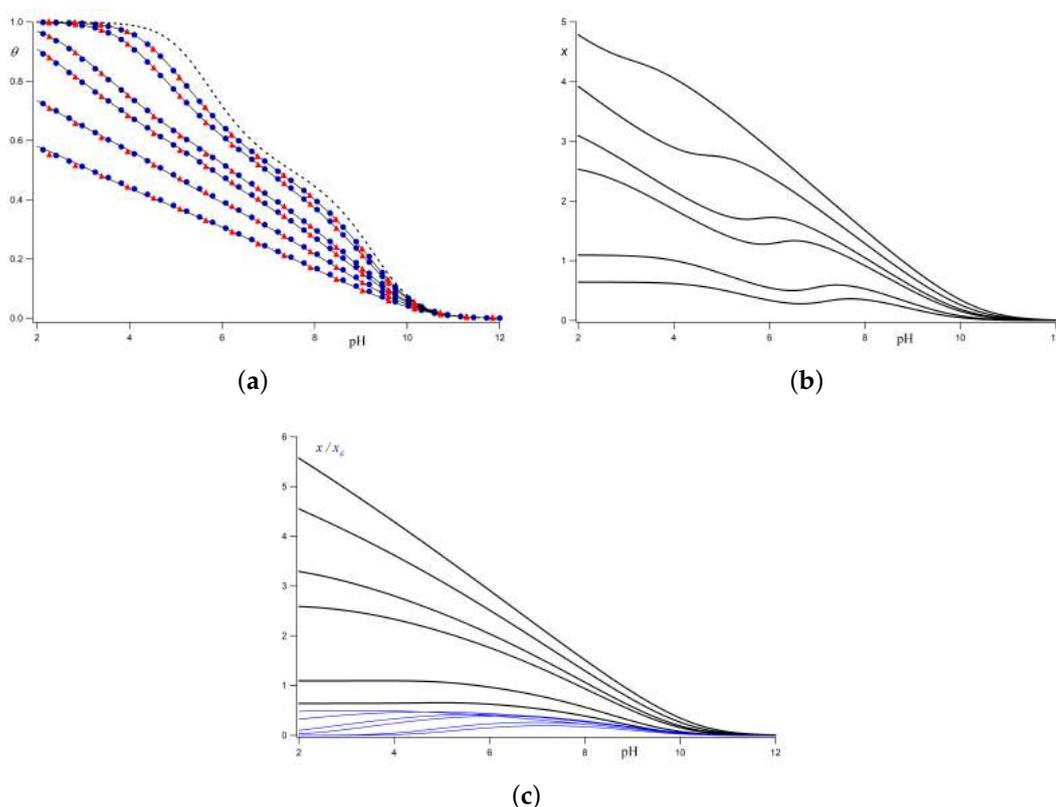


Figure 2. (a) Titration curves corresponding to a rigid linear chain with interacting ionizable sites separated by a distance $b = 0.2$ nm obtained using Monte Carlo (MC) simulations (blue circles), Local Effective Interaction Parameters (LEIP) method correcting only the pK -value (continuous line) and LEIP method correcting the pK -value and the nearest neighbour interaction energy ϵ (red triangles). The chosen parameters are $pK = 9$ and $\epsilon = 1.5$. The Long-Range (LR) interactions are calculated using the Debye–Hückel potential. The dashed line represents the titration curve in the absence of LR interactions. Note that the correction to the pK -value is enough to reproduce almost exactly the MC simulations and no significant improvement is obtained in correcting ϵ ; (b) correction x to the pK -value using the LEIP method; (c) corrections x (black lines) and x_ϵ (blue lines) to the pK -value and the nearest neighbour interaction energy ϵ , respectively. In all the figures, from top to bottom, the ionic strengths are 1 M, 0.5 M, 0.1 M, 0.05 M, 0.01 M and 0.001 M.

Equation (13) states that x represents the increase in φ_0 for a constant number of interactions D_0 , while x_ϵ can be interpreted as the change in φ_0 in creating a nearest neighbour interaction, keeping constant the number of bound protons ν_0 . Intuitively, one can guess that x_ϵ is much smaller than x , so that the correction to the pK is enough to reproduce almost exactly the exact free energy, generating physical properties almost indistinguishable from the MC simulations. In Figure 2a, the titration curves have been recalculated using the correction to ϵ . As expected, no significant improvement is obtained. x and x_ϵ as functions of the pH are shown in Figure 2c, where it is clearly observed that x_ϵ is much lower than x for all the ionic strengths. Note that the wavy behaviour of x in Figure 2b is no longer present in Figure 2c, and seems to be replaced by the contribution x_ϵ . Using the same procedure, corrections to higher order interactions, such as triplet or next-nearest neighbour interactions, can be calculated until the desired accuracy is obtained, and expressions of type (13) can be generalized in a straightforward manner. The same treatment leads to very good results for heterogeneous polyelectrolytes and polyampholytes, by correcting the pK -values of the different kind of sites ($pK_i \rightarrow pK_i - x_i$) [56].

3. Coupling of Ionization and Conformational Equilibria

For a linear macromolecule composed by M bonds, a particular conformational state is denoted by a set of variables $c = \{c_\alpha\}$, $j = 1, \dots, M$. The variables c_α can adopt several values corresponding to the rotational angles of the bond α . The possible states of the bonds are usually chosen as those of minimum energy, three in the simplest situation: *trans*, *gauche+* and *gauche-*. The selection of a finite number of rotational states instead of working with the full continuous rotational potential greatly simplifies the statistical mechanics treatment, and constitutes the basis of the Rotational Isomeric State (RIS) model, mainly developed by Flory [47]. In the case of linear polymers, the transfer method can be used to determine the conformational partition function Ξ_{rot} , which can be expressed as

$$\Xi_{\text{rot}} = \mathbf{q} \mathbf{U}_1 \mathbf{U}_2 \dots \mathbf{U}_{M-1} \mathbf{U}_M \mathbf{p}^T, \quad (14)$$

where \mathbf{U}_α is the transfer matrix corresponding to the bond α . For a symmetric chain, for which the states *gauche+* and *gauche-* have the same energy and identical bonds, the transfer matrices are of the form

$$\mathbf{U} = \begin{pmatrix} 1 & \sigma & \sigma \\ 1 & \sigma\psi & \sigma\omega \\ 1 & \sigma\omega & \sigma\psi \end{pmatrix}, \quad (15)$$

where σ , ψ and ω are the Boltzmann factors associated with the conformational energies of the bonds: $-k_B T \ln \sigma$ is the free energy of the *gauche* states while ψ (ω) are related to the interaction energies between two consecutive *gauche* states of different (same) orientation. ψ and ω equate one if the rotation of the bonds is independent. $\mathbf{q} = (1, 0, 0)$ and $\mathbf{p} = (1, 1, 1)$ are the initiating and terminating vectors. As in the SB model, the necessary thermal averages can be obtained by performing proper derivatives of the partition function. For instance, the average number of bonds in the *gauche* state is given by [47]

$$g = \frac{\partial \ln \Xi_{\text{rot}}}{\partial \ln \sigma}. \quad (16)$$

The RIS model can be generalized in order to take into account the protonation degrees of freedom. If the macromolecule is in a protonation state s , the pair (s, c) defines a *roto-microstate* with reduced free energy

$$F(s, c) = F_{\text{rot}}(c) + F_p(s, c). \quad (17)$$

$F_{\text{rot}}(c)$ is the free energy corresponding to the fully deprotonated state of each conformer, while $F_p(s, c)$ represents the reduced free energy due to the protonation process, which, for a given conformation, can be expressed as

$$\frac{\beta F_p(s, c)}{\ln 10} = \sum_i \mu_i(c) s_i + \sum_{j>i} \phi_{ij}(c) s_i s_j, \tag{18}$$

where triplet interactions have been neglected. Note that the cluster parameters now depend on the conformational state c . The reduced free energy Equations (17) and (18) combines the RIS and the SB model and defines the SBRIS model, which allows for studying conformational and ionization properties on the same foot. In recent publications, the SBRIS model has been used to explain conformational transitions in weak linear polyelectrolytes [45] and in the characterization of ionization/conformational properties of linear poly(ethylenimine) [46]. The probability of a specified roto-microstate is given by

$$p(s, c) = \frac{e^{-\beta F(s, c)}}{\Xi_{\text{SBRIS}}}, \tag{19}$$

where the SBRIS partition function Ξ_{SBRIS} is defined as

$$\Xi_{\text{SBRIS}} = \sum_{s, c} e^{-\beta F(s, c)}. \tag{20}$$

The SBRIS partition function can alternatively be expressed in the fashion

$$\Xi_{\text{SBRIS}} = \sum_s \Xi_{\text{rot}}(s), \tag{21}$$

where $\Xi_{\text{rot}}(s)$ denotes the rotational partition function for the macromolecule in a ‘frozen’ binding configuration $s = \{s_1, s_2, \dots, s_N\}$. $\Xi_{\text{rot}}(s)$ can then be calculated as a RIS partition function as in Equation (14), but now decorating the transfer matrices with the suitable binding parameters. The sum over the protonation states can be performed by using proper matricial methods described elsewhere [46]. They are outlined as supplementary information and here we just comment the final results. The SBRIS partition function is obtained by replacing the conformational RIS transfer matrices \mathbf{U} (Equation (15)) for suitable super-matrices. The rule is that, if a bond is holding at its ends two ionization groups, \mathbf{U} must be replaced by

$$\mathbf{U} \rightarrow \mathbf{B} = \begin{pmatrix} \mathbf{U} & \mathbf{Uz} \\ \mathbf{U} & \mathbf{Uuz} \end{pmatrix}, \tag{22}$$

where \mathbf{u} is a diagonal matrix containing the Boltzmann factors corresponding to the short-range interactions

$$\mathbf{u} = \begin{pmatrix} u_t & 0 & 0 \\ 0 & u_g & 0 \\ 0 & 0 & u_g \end{pmatrix}. \tag{23}$$

In matrix (23), $-k_B T \ln u_t$ and $-k_B T \ln u_g$ represent the short-range interaction energy between two protonated sites separated by a bond in *trans* and *gauche* conformation, respectively. For the bonds which do not hold ionization sites, the substitution is

$$\mathbf{U} \rightarrow \mathbf{B} = \begin{pmatrix} \mathbf{U} & \mathbf{0} \\ \mathbf{0} & \mathbf{U} \end{pmatrix}. \tag{24}$$

The resulting SBRIS partition function reads

$$\Xi_{\text{SBRIS}} = \mathbf{r} \mathbf{B}_1 \mathbf{B}_2 \dots \mathbf{B}_{M-1} \mathbf{B}_M \mathbf{t}^T, \quad (25)$$

where $\mathbf{r} = (\mathbf{q} \ \mathbf{q})$ and $\mathbf{t} = (\mathbf{p} \ \mathbf{p})$ are the initiating and terminating vectors, respectively. The average number of bound protons and the bond state probabilities can be again obtained by proper derivatives of Equation (25). It can also be shown that matricial expressions for other physical quantities derived in the context of the RIS model can also be generalized to ionizable molecules by performing suitable substitutions by super-matrices [45]. For instance, there are available matricial expressions for the average square distance between two sites of the chain. These expressions are used in this work to estimate the average distance between charged sites and the corresponding LR interaction energy.

As commented on in the preceding section, transfer matrix methods can only be applied if only SR interactions are taken into account. In this work, we propose to use the LEIP technique to include the LR interactions via local parameters, as done in the case of rigid molecules. Now, however, not only the ionization parameters, such as $pK \rightarrow pK - x$, but also the conformational parameters must be corrected as outlined in Figure 3. In the simplest case, with only one kind of rotating bonds, the substitution $p\sigma \rightarrow p\sigma + x_\sigma$ where $p\sigma = -\log \sigma$ will be necessary. The treatment is almost identical to the one used for rigid molecules. Now, the “unperturbed” free energy is $\Omega_0 = -\ln \Xi_{\text{SBRIS}}(x, x_\sigma)$. It can be easily shown that the corrections x for the pK and x_σ for $p\sigma$ fulfill equivalent equations to (12)

$$\mathbf{J} \begin{pmatrix} x \\ x_\sigma \end{pmatrix} = \begin{pmatrix} \left(\frac{\partial \varphi_0}{\partial x} \right)_{x_\sigma} \\ \left(\frac{\partial \varphi_0}{\partial x_\sigma} \right)_x \end{pmatrix}; \quad \mathbf{J} = \begin{pmatrix} \left(\frac{\partial v_0}{\partial x} \right)_{x_\sigma} & \left(\frac{\partial g_0}{\partial x} \right)_{x_\sigma} \\ \left(\frac{\partial v_0}{\partial x_\sigma} \right)_x & \left(\frac{\partial g_0}{\partial x_\sigma} \right)_x \end{pmatrix}, \quad (26)$$

where $v_0(x, x_\sigma)$ represents the average number of bound protons (Equation (4)) and $g_0(x, x_\sigma)$ the average number of bonds in the *gauche* state (Equation (16)), calculated using the unperturbed free energy. If we use v_0 and g_0 as independent variables, instead of x and x_σ , Equation (26) can be rewritten in a similar fashion as Equation (13)

$$x = \left(\frac{\partial \varphi_0}{\partial v_0} \right)_{g_0}; \quad x_\sigma = \left(\frac{\partial \varphi_0}{\partial g_0} \right)_{v_0}, \quad (27)$$

which essentially tell us that x represents the average change in φ_0 when a new proton is bound (keeping constant the number of bonds in *gauche*) while x_σ is the the average change in φ_0 when a bond is brought to its *gauche* state (keeping constant the number of bound protons). Note that the LEIP method always leads to expressions for the interaction corrections of the same type of Equations (13) and (27).

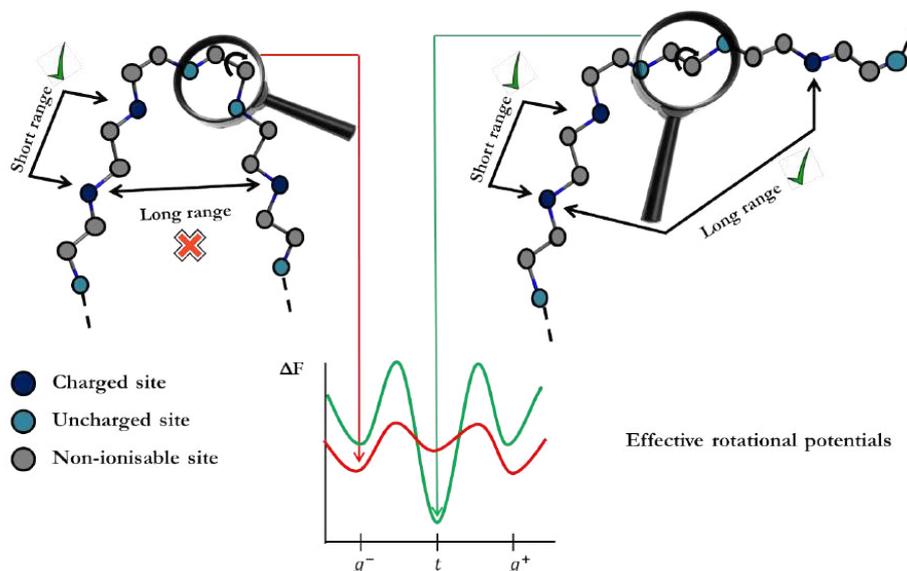


Figure 3. Outline of the Local Effective Interaction Parameters (LEIP) method. LEIP accounts for the Short-Range (SR) interactions exactly, but LR interactions are replaced by effective SR free energies. In the resulting scheme, only SR interactions are present, which considerably simplify the theoretical treatment. The bonds “feel” the presence of the LR interactions in an effective way, which leads to an apparent pH-dependent rotational energy. Other parameters, such as the protonation p*K*-values, also become pH-dependent due to the presence of LR interactions.

As in the previous section, LR interactions are described by the Debye–Hückel potential, although the method could in principle be applied to other kind of interactions such as van der Waals interactions. Moreover, by including convenient “hard core” terms in the interaction potentials, the excluded volume effect could in principle be taken into account. The study of this effect, however, is not trivial and it is out of the scope of this work. Unlike rigid molecules, for flexible molecules, the average LR interaction energy φ_0 for the unperturbed free energy can only be approximately calculated. In this work, as a first approximation, we assume that

$$\varphi_0 = \left\langle \sum_{ij} \phi(d_{ij}) s_i s_j \right\rangle_0 \simeq \sum_{ij} \phi \left(\sqrt{\langle d_{ij}^2 \rangle_0} \right) \langle s_i s_j \rangle_0. \quad (28)$$

This approximation could in principle be improved by using higher order moments of d_{ij} . Matricial expressions for $\langle d_{ij}^2 \rangle_0$ and higher moments were derived by Flory and Jernigan [47,58] for neutral chains. Here, these expressions are modified in order to account for the protonation degrees of freedom. An outline of the derivations is provided as supplementary information.

4. Monte Carlo Simulations

In order to estimate the accuracy of the LEIP method when applied to flexible polyelectrolytes, we compare the theoretical values with those resulting from MC simulations. Two main MC techniques have been previously proposed: the Reaction Ensemble approach, for which the pH is a calculated quantity [59,60], and the constant pH method, corresponding to the semi-grand canonical ensemble [33,61–63]. Since the control variable in the LEIP method is the pH-value, as indicated by the reduced free energies in Equations (1) and (17), the constant pH method has been chosen here. In previous studies about polyelectrolyte ionization properties, both Reaction Ensemble and constant pH methods have been coupled to Molecular Dynamics schemes in order to deal with explicit

ions. In this work, free protons, co- and counter-ions are not explicit in the simulations and the screening effects are taken into account via the Debye length parameter, κ^{-1} . The MC code generalises the one previously used in the computation of conformational and ionization properties of linear poly(ethylenimine) [46]. The polyelectrolyte is modeled as a linear chain with rigid bond lengths and angles. Bonds can adopt one of the three states of minimum energy (*trans*, *gauche+* or *gauche-*). Each change of a bond state implies a 120° rotation of its dihedral angle and the recalculation of distances among the sites situated before and after the rotating bond. The linear chain is composed of interacting nodes which can correspond to inert or protonating groups. In Figure 4, two snapshots of Monte Carlo simulations at ionic strength 0.001 M and two pH-values (four in Figure 4a and eight in Figure 4b) are presented. As in Figures 1 and 3, the ionizable sites are depicted in blue (dark blue if they are protonated and cyan otherwise). It can be observed that a decrease in the pH-value promotes the elongation of the chain, and the consequent reduction of the electrostatic repulsion, by increasing the number of bonds in the *trans* conformations.

In the MC simulations, the free energy of the system is divided into SR and LR terms

$$F(s, c) = F_{\text{SR}} + F_{\text{LR}} + \sum_i (\text{pH} - \text{p}K_i) s_i. \quad (29)$$

The SR term is computed using SBRIS free energy (Equation (17)) which involves the energies present in the transfer matrices (σ , ψ , ω , u_t and u_g), while the LR contribution is calculated using the Debye–Hückel potential (Equation (11)). If F_{LR} is set to zero, the obtained results coincide, within the numerical error, with those obtained using the transfer matrix method. This was one of the tests used to check the reliability of the Monte Carlo code. A Metropolis algorithm [15,27] is used to generate roto-microstates at constant pH in a chain with 50 ionizable sites (i.e., 148 nodes or 147 bonds). In each new MC configuration, the polyelectrolyte can change either (i) the conformational state of a rotating bond or (ii) the ionization state of a binding site, with trial probabilities of 0.999 and 0.001, respectively. These trial probabilities allow us to obtain a good equilibration of the conformational structure for each ionization state and the system does not become trapped in local minima. The probability to accept a new configuration is obtained by computing the free energy difference ($\Delta F(s, c)$) between trial and actual conformations. When the state of the bond α is changed, the following free energy differences must be calculated: (i) the conformational energy of bond α and its interaction with bonds $\alpha \pm 1$ (corresponding to the parameters σ , ψ and ω); (ii) the electrostatic SR interaction between the two sites bound to α when they are charged (corresponding to u_t and u_g , which depend on the new conformation of α); and (iii) the change in the LR Debye–Hückel interaction among sites before and after α , which involves the recalculation of the distances between the charged sites. On the other hand, a change in the ionization state of a site s_i implies to recalculate: (i) the reduced chemical potential of the site i by an amount $\Delta\mu_i = (\text{pH} - \text{p}K_i)\Delta n$, where $\Delta n = \pm 1$ is the variation in the number of protons; (ii) the SR repulsive interaction between s_i and $s_{i\pm 1}$; and (iii) the LR Debye–Hückel interactions between the trial protonating site and the rest of ionized sites. Once $\Delta F(s, c)$ is computed, the new configuration is always accepted if $\Delta F(s, c) < 0$ and accepted with a probability $\exp(-\beta\Delta F(s, c))$ if $\Delta F(s, c) > 0$. The values presented are the average over eight different MC simulations. Each MC simulation has been equilibrated in the first 5×10^7 configurations and the thermal averages have been computed in the following 4.5×10^8 realizations. The simulations were performed using a parallel code developed in C++ on a 126 CPU cluster. For each pH and ionic strength (one point of the curves), typical jobs were run using 8 CPUs during 1 to 2 h.

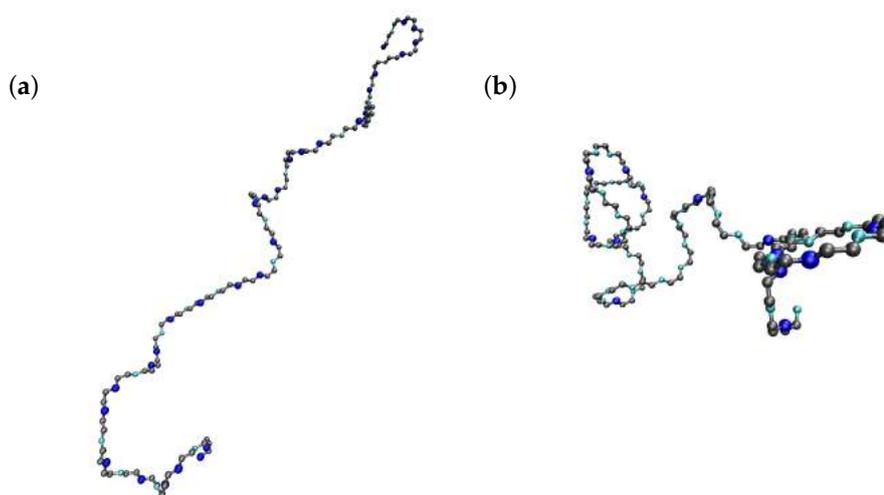


Figure 4. Two snapshots of Monte Carlo simulations with $pK = 9$, $\epsilon_t = 1$, $u_g = 0$, $\sigma = 10$ and pH = 4 (a) and pH = 8 (b). Note that elongated conformations are promoted at low pH as a consequence of polyelectrolyte global charge increase.

5. Results and Discussion

As a model system, we use the linear polyelectrolyte outlined in Figure 1b, with protonating sites situated every three chain positions. Only c bonds are allowed to rotate and they can adopt the three states of minimum energy, i.e., *trans*, *gauche+* or *gauche-*. The conformation of c bonds determines the intensity of the SR interactions between neighbouring protonated sites. The rest of bonds (a and b in the figure) are forced to be in the *trans* state. The energy of the *gauche* state of the c bonds is denoted as $p\sigma = -\log \sigma$. The c bonds are assumed to rotate independently when the macromolecule is uncharged ($\omega = \psi = 1$ in Equation (15)). The protonating sites are considered to be identical with the same protonation pK-value. The interactions between protonated sites are characterized by the energies $\epsilon_t = -\log u_t$ (when the bond c is in *trans* state) and $\epsilon_g = -\log u_g$ (when the bond c is in *gauche* state). In the computations, the used values of the bond length and the bond angle are 0.2 nm and 120° , respectively. This model can be regarded as a minimal model of a flexible polyelectrolyte, with only four energetic parameters involved (σ , ϵ_t , ϵ_g and pK), and it is here used to illustrate the application of the LEIP method to the analysis of the interplay of conformational and protonation degrees of freedom.

Let us firstly consider the case for which c bonds can freely rotate when the adjacent sites are deprotonated (i.e., $\sigma = 1$). When both sites are charged, however, the very strong SR repulsion hinders the *gauche* conformation, so that we take $u_g = 0$ (i.e., $\epsilon_g \rightarrow \infty$). The resulting titration curves are shown in Figure 5a for ionic strengths ranging from 1 M to 0.001 M. The chosen parameters are pK = 9 and $\epsilon_t = 1$. The black continuous lines represent the average protonation degree θ calculated using the LEIP method correcting both the pK-value ($pK \rightarrow pK - x$) and the conformational energy of c bonds ($p\sigma \rightarrow p\sigma + x_\sigma$), while the red circles represent the results of the MC simulations. It is observed that the LEIP method reproduces very accurately the MC simulations for all the range of pH-values and ionic strengths. The dashed line depicts the values provided by LEIP for $I = 0.001$ M if only the pK-value is corrected, while σ remains constant. Although relative good prediction of the MC simulations is obtained, the quality of the titration curve clearly improves if σ is corrected. This means that the rotational energy of the c bonds is affected by LR interactions even if their pendant sites are not charged, as a result of the tendency of the chain to separate the rest of charged groups. Actually, the system behaves as if c bonds “feel” the LR interactions in an effective way. This effect is more remarkable in the subsequent case.

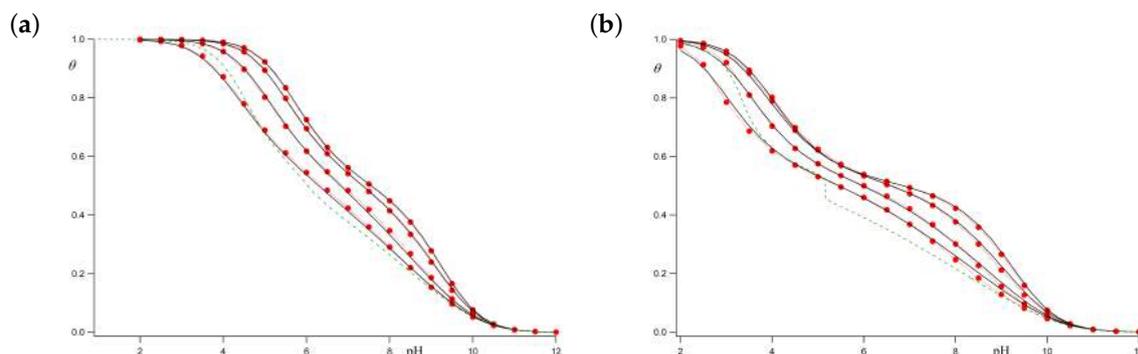


Figure 5. Titrations curve for the model polyelectrolyte depicted in Figure 2 with $pK = 9$, $\epsilon_t = 1$, $u_g = 0$ (i.e., $\epsilon_g \rightarrow \infty$) and $\sigma = 1$ (a); $\sigma = 10$ (b). The chosen ionic strengths are, from top to bottom, 1 M, 0.1 M, 0.01 M and 0.001 M. Black continuous lines represent calculations using the LEIP method in which two effective short-range parameters (pK and $p\sigma$) has been corrected. Red circles depict the MC values. The green dashed line corresponds to the LEIP values for $I = 0.001$ M when only the pK -value is corrected, while $p\sigma$ is kept constant.

Let us check the accuracy of LEIP method when the *gauche* states of the *c* bonds are favored, for instance, because of the existence of hydrogen bond, which means that $\sigma > 1$. When the adjacent sites are both protonated, on the contrary, the electrostatic repulsion is so strong that the *gauche* states are forbidden ($u_g = 0$). In this case, the conformational propensity changes when the ionization state of the sites change. Figure 5b compares the titration curves obtained using LEIP correcting pK and $p\sigma$ and MC simulations for $pK = 9$, $\epsilon_t = 1$ and $\sigma = 10$. As can be observed, LEIP method and MC simulations yield to almost identical titration curves. In this case, however, the correction of $p\sigma$ becomes compulsory. If only pK is corrected, the titration curve obtained by LEIP at 0.001 M (green dashed line) exhibits a phase transition-like behaviour at $pH \approx 5$. This is an artifact resulting from the impossibility to explain the complex interplay of charge regulation and conformational transition without taking into account the influence of LR interactions in the effective energy of the *gauche* state.

For the two cases commented above, the *gauche* state probabilities versus the pH are shown in Figure 6a,b. Markers correspond to MC simulations while black lines represent the theoretical values at ionic strengths 1 M, 0.01 M and 0.001 M, for $pK = 9$, $\epsilon_t = 1$ and $u_g = 0$. Figure 6a corresponds to the case with $\sigma = 1$. A good correspondence between simulated and theoretical profiles is obtained for all the ionic strengths. Since at low pH-values, the polyelectrolyte is almost fully protonated, the *gauche* state probability tends to zero because of the high electrostatic repulsion between the nearest charged sites in the *gauche* position ($u_g = 0$). On the other hand, at high pH-values, the macromolecule is completely uncharged and the *c* bonds are freely to rotate. As a result, the probability of the two *gauche* conformers tends to $2/3$. For $I = 1$ M, the LR interactions can be neglected and total correspondence between simulated and calculated values is found. At higher ionic strengths (0.01 M and 0.001 M), for which the Debye–Hückel potential is not screened enough, some small differences arise. However, still now, good agreement between the LEIP method and MC simulations is observed.

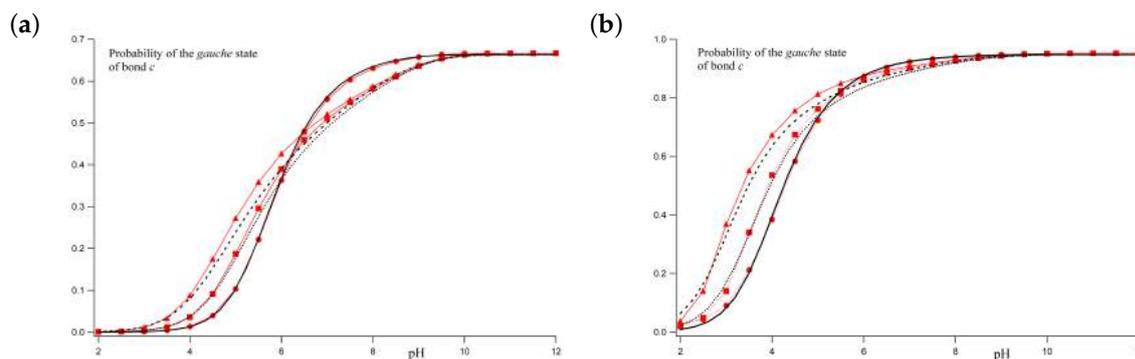


Figure 6. *Gauche* state probabilities of bond c versus the pH-value computed by means of MC simulations (red markers) and using the LEIP method (black lines) for ionic strengths: 1 M (circles and continuous line), 0.01 M (squares and dotted line) and 0.001 M (triangles and dashed lines). The chosen parameters are $pK = 9$, $\epsilon_t = 1$, $u_g = 0$ and $\sigma = 1$ (a); $\sigma = 10$ (b).

Figure 6b corresponds to the case with $\sigma = 10$. Simulated and theoretical probabilities are also in good agreement. Again, the *gauche* state probability tends to zero for low pH-values, while, at high pH-values, the population of the *gauche* conformer is $2\sigma / (2\sigma + 1) = 0.95$. A continuous transition from *gauche* to *trans* conformations as the pH decreases is observed. This transition is sharper than in the previous case. Note that, from the LEIP point of view, this transition occurs because of a double effect. On the one hand, there is the charging process, so that two adjacent sites tend to minimize the repulsion when the bonds adopt the *trans* conformation. This effect is present even when LR interactions are not present. On the other hand, the effective *gauche* state energy is increasing due to the average effect of the LR interactions ($p\sigma \rightarrow p\sigma + x_\sigma$). Both effects are important to correctly reproduce the MC simulations. Otherwise, the lack of flexibility in the conformational energy leads to the spurious phase transition observed in Figure 5b (green dashed line).

Figure 7 shows the LEIP method correction x_σ to the bond conformational energy ($p\sigma \rightarrow p\sigma + x_\sigma$) for $\sigma = 1$ (Figure 7a) and $\sigma = 10$ (Figure 7b). In both cases, it is observed that x_σ tends to zero at high pH-values, since the molecule is uncharged and no LR interactions are present, so no correction is necessary. As a general tendency, x_σ tends to increase as the pH decreases due to the charging process and the corresponding increase in the LR interactions. This effect is larger at low values of the ionic strength since the Debye–Hückel potential is less screened. For the case $\sigma = 10$, a wavy behaviour is observed for ionic strengths 0.1 M and 0.01 M and x_σ exhibits a smooth maximum at $pH \simeq 4$, which coincides with the pH regime where the *trans* to *gauche* transition is sharper. This fact could be due to correlations between the rotation of neighbouring bonds or because part of the correction is effectively included in the pK -correction x . Further analysis would probably be necessary in order to clarify this point.

In the two cases discussed above, we have taken $u_g = 0$, which means that bonds between two protonated sites cannot be in the *gauche* state. Let us now relax this condition and take a finite value for u_g , so that the electrostatic interaction between two charged sites in *gauche* is not forbidden but only penalized. LEIP predictions (black lines) and MC simulations (red markers) are plotted in Figure 8. Figure 8a shows the computed titration curves with $\epsilon_g = 2$ at ionic strengths ranging, from top to bottom, from 1 M to 0.001 M. Excellent agreement between the theoretical predictions and simulations is obtained for all the ionic strengths, so that the relaxation of the condition $u_g = 0$ does not seem to affect the accuracy of the LEIP approach. *Gauche* state probabilities versus pH at three ionic strengths are depicted in Figure 8b: 1 M (circles and continuous line), 0.01 M (squares and dotted line) and 0.001 M (triangles and dashed line). As expected, even at low pH-values, some bonds can remain in the *gauche* state due to the finite value of u_g . Despite the complexity of the obtained profiles for this case, LEIP is able to accurately reproduce the MC simulations.

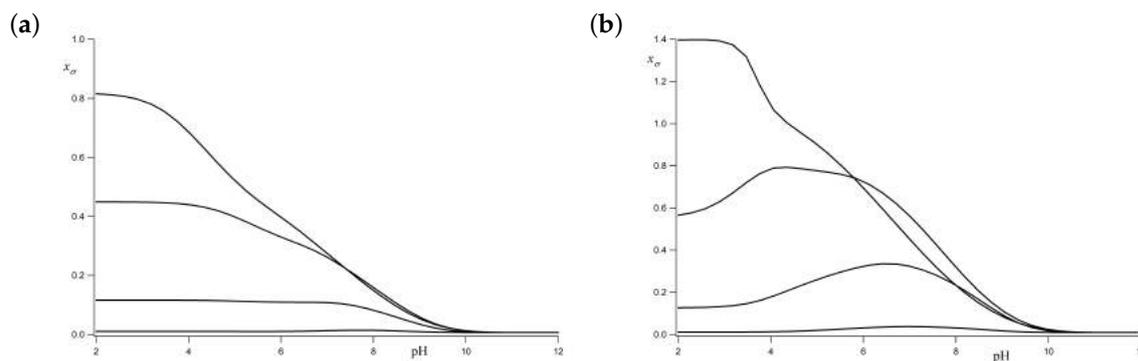


Figure 7. Correction to the *gauche* state energy x_σ versus the pH for $pK = 9$, $\epsilon_t = 1$, $u_g = 0$ and $\sigma = 1$ (a); $\sigma = 10$ (b). From the bottom to the top, the ionic strengths are 1 M, 0.1 M, 0.01 M and 0.001 M. x_σ represents the average effective energy “felt” by c bonds as a result of the LR interactions.

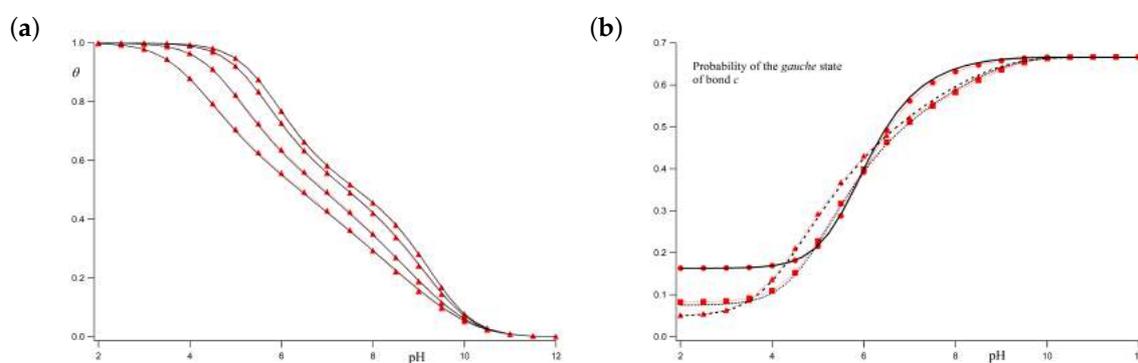


Figure 8. Titration curves (a) and *gauche* state probabilities (b) obtained using the LEIP method (black lines) and MC simulations (red markers). The chosen parameters are $pK = 9$, $\epsilon_t = 1$, $\epsilon_g = 2$ and $\sigma = 1$. (a) the ionic strength are, from top to bottom, 1 M, 0.1 M, 0.01 M and 0.001 M; (b) three different ionic strengths are shown: $I = 1$ M (circles and continuous line), 0.01 M (squares and dotted line) and 0.001 M (triangles and dashed line).

6. Conclusions

The ionization and conformational properties of polyelectrolytes are determined by a combination of Short-Range (SR) and Long-Range (LR) interactions between bonds and ionizable sites. In particular, electrostatic LR interactions can only be neglected at high enough ionic strengths, which is an important limitation for many macromolecular systems of interest. The present work explores the possibility of defining local, short-ranged, interaction parameters which are corrected to account for the LR interactions in an effective way. The new parameters are systematically calculated by using variational methods and equations for them are provided. The resulting approach, the Local Effective Interaction Parameters (LEIP) method, was firstly developed to study the binding properties of rigid polyelectrolytes. In this paper, these ideas are extended to flexible polyelectrolytes, for which conformational and ionization equilibria (charge regulation) are strongly coupled. With this aim, LEIP is combined with the Site Binding Rotational Isomeric State (SBRIS) model in order to deal simultaneously with conformational and protonation degrees of freedom for the full range of ionic strengths.

The LEIP method is illustrated by using a model of a linear symmetric polyelectrolyte containing protonating sites situated regularly along the polymer backbone. The charged sites interact by means of the Debye–Hückel potential, which accounts for the electrostatic screening in an average way, while excluded volume effects are neglected. The bonds linking the ionizable sites can be in three possible states, i.e., *trans*, *gauche+* and *gauche-*. This model with only four relevant parameters

(protonation pK -value, *gauche* state energy and SR electrostatic interactions between neighbouring sites through bonds in *trans* and *gauche* states) can be regarded as a minimal model of a flexible polyelectrolyte where conformational and binding equilibria are strongly coupled. The LEIP method is applied to correct both the protonation pK -values and the *gauche* state energy. As a result, local pH dependent rotational potentials are obtained. The correction to the *gauche* energy represents the contribution of the LR interactions in rotating a bond to its *gauche* state.

The degree of protonation and the *gauche* state probabilities obtained using the LEIP method are compared with those computed using semi-grand canonical Monte Carlo (MC) simulations. In all of the studied cases, the agreement between LEIP and MC simulations is excellent. The computational cost, however, is orders of magnitude lower in the LEIP method. This fact allows using LEIP to directly fit parameters to experimental information. The LEIP method could also represent a complementary tool to the study of other aspects of the polyelectrolyte physical chemistry, such as the dependence of the molecular size on the pH, the influence of excluded volume interactions, the presence of attractive hydrophobic interactions or the competitive binding of metal ions. The clarification of these points, which have not been the subject of the present study, would be desirable in order to extend the applicability of the LEIP method. We think that the ideas presented here could be useful in the design of pH-dependent force fields based on experimental ionization and conformational properties.

Supplementary Materials: General equations for SBRIS partition function and mean square distance between sites, necessary in the computations of this work, are provided as supplementary information (can be available online at <http://www.mdpi.com/2073-4360/10/08/811/s1>). The transfer matrices corresponding to the model polyelectrolyte used here are also reported.

Author Contributions: Conceptualization, F.M. and J.L.G.; Methodology, F.M. and J.L.G.; Software, P.M.B. and S.M.; Validation, P.M.B. and S.M.; Investigation, P.M.B. and J.L.G.; Writing—Original Draft Preparation, P.M.B., S.M., F.M. and J.L.G.; Writing—Review and Editing, P.M.B., S.M., F.M. and J.L.G.; Visualization, P.M.B. and J.L.G.

Funding: We acknowledge the financial support from Generalitat de Catalunya (Grants 2017SGR1033, 2017SGR1329 and XrQTC). Josep Lluís Garcés also acknowledges the Spanish Ministry of Science and Innovation (project CTM2016-78798-C2-1-P). Sergio Madurga and Francesc Mas acknowledge the funding of the EU project 8SEWP-HORIZON 2020 (692146). Pablo M. Blanco also acknowledges the financial support from Grant (FI-2017) of Generalitat de Catalunya.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

LEIP	Local Effective Interaction Parameters
LR	Long-Range
MC	Monte Carlo
RIS	Rotational Isomeric State
SB	Site Binding
SBRIS	Site Binding Rotational Isomeric State
SR	Short-Range

References

1. Wyman, J.; Gill, S.J. *Binding and Linkage*; University Science Books: Mill Valley, CA, USA, 1990.
2. Warshel, A.; Sharma, P.K.; Kato, M.; Parson, W.W. Modeling electrostatic effects in proteins. *BBA Proteins Proteom.* **2006**, *1764*, 1647–1676. [[CrossRef](#)] [[PubMed](#)]
3. Lund, M.; Jönsson, B. Charge regulation in biomolecular solution. *Q. Rev. Biophys.* **2013**, *46*, 265–281. [[CrossRef](#)] [[PubMed](#)]
4. Lipfert, J.; Doniach, S.; Das, R.; Herschlag, D. Understanding Nucleic Acid–Ion Interactions. *Annu. Rev. Biochem.* **2014**, *83*, 813–841. [[CrossRef](#)] [[PubMed](#)]
5. Borkovec, M.; Koper, G.J.M.; Piguët, C. Ion binding to polyelectrolytes. *Curr. Opin. Colloid Interface Sci.* **2006**, *11*, 280–289. [[CrossRef](#)]
6. Hamacek, J.; Borkovec, M.; Piguët, C. Simple thermodynamics for unravelling sophisticated self-assembly processes. *Dalton Trans.* **2006**, *12*, 1473–1490. [[CrossRef](#)] [[PubMed](#)]
7. Li, Y.; Zhao, T.; Wang, C.; Lin, Z.; Huang, G.; Sumer, B.D.; Gao, J. Molecular basis of cooperativity in pH-triggered supramolecular self-assembly. *Nat. Commun.* **2016**, *7*, 13214. [[CrossRef](#)] [[PubMed](#)]
8. Stumm, W.; Morgan, J.J. Aquatic Chemistry. In *Chemical Equilibria and Rates in Natural Waters*, 3rd ed.; Wiley: New York, NY, USA, 1996.
9. Rivas, L.; Pereira, E.D.; Moreno-Villoslada, I. Water-soluble polymer–metal ion interactions. *Prog. Polym. Sci.* **2003**, *28*, 173–208. [[CrossRef](#)]
10. Trefalt, G.; Behrens, S.H.; Borkovec, M. Charge Regulation in the Electrical Double Layer: Ion Adsorption and Surface Interactions. *Langmuir* **2016**, *32*, 380–400. [[CrossRef](#)] [[PubMed](#)]
11. Dobrynin, A.V. Theory and simulations of charged polymers: From solution properties to polymeric nanomaterials. *Curr. Opin. Colloid Interface Sci.* **2008**, *13*, 376–388. [[CrossRef](#)]
12. Chapel, J.P.; Berret, J.F. Versatile electrostatic assembly of nanoparticles and polyelectrolytes: Coating, clustering and layer-by-layer processes. *Curr. Opin. Colloid Interface Sci.* **2012**, *17*, 97–105. [[CrossRef](#)]
13. Carnal, F.; Clavier, A.; Stoll, S. Polypeptide-nanoparticle interactions and corona formation investigated by monte carlo simulations. *Polymers* **2016**, *8*, 203. [[CrossRef](#)]
14. Hartig, S.M.; Greene, R.R.; Dikov, M.M.; Prokop, A.; Davidson, J.M. Multifunctional nanoparticulate polyelectrolyte complexes. *Pharm. Res.* **2007**, *24*, 2353–2369. [[CrossRef](#)] [[PubMed](#)]
15. Borkovec, M.; Jönsson, B.; Koper, G.J.M.J. *Surface and Colloid Science*; Matijevic, E., Ed.; Plenum Press: New York, NY, USA, 2001; Volume 16.
16. Olander, D.S.; Holtzer, A. The Stability of the Polyglutamic Acid alpha Helix. *J. Am. Chem. Soc.* **1968**, *90*, 4549–4560. [[CrossRef](#)] [[PubMed](#)]
17. Uyaver, S.; Seidel, C. First-order conformational transition of annealed polyelectrolytes in a poor solvent. *Europhys. Lett.* **2003**, *64*, 536–542.
18. Whitten, S.T.; García-Moreno, B.; Hilser, V.J. Local conformational fluctuations can modulate the coupling between proton binding and global structural transitions in proteins. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 4282–4287. [[CrossRef](#)] [[PubMed](#)]
19. Das, R.K.; Pappu, R.V. Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 13392–13397. [[CrossRef](#)] [[PubMed](#)]
20. Das, R.K.; Ruff, K.M.; Pappu, R.V. Relating sequence encoded information to form and function of intrinsically disordered proteins. *Curr. Opin. Struct. Biol.* **2015**, *32*, 102–112. [[CrossRef](#)] [[PubMed](#)]
21. Borkovec, M.; Koper, G.J.M. Ising models of polyprotic acids and bases. *J. Phys. Chem.* **1994**, *98*, 6038–6045. [[CrossRef](#)]
22. Borkovec, M.; Koper, G.J.M. A cluster expansion method for the complete resolution of microscopic ionization equilibria from NMR titrations. *Anal. Chem.* **2000**, *72*, 3272–3279. [[CrossRef](#)] [[PubMed](#)]
23. Koper, G.J.M.; Borkovec, M. Binding of metal ions to polyelectrolytes and their oligomeric counterparts: An application of a generalized potts model. *J. Phys. Chem. B* **2001**, *105*, 6666–6674. [[CrossRef](#)]
24. Koper, G.J.M.; Borkovec, M. Proton binding by linear, branched, and hyperbranched polyelectrolytes. *Polymer* **2010**, *51*, 5649–5662. [[CrossRef](#)]
25. Borkovec, M.; Cakara, D.; Koper, G.J.M. Resolution of microscopic protonation enthalpies of polyprotic molecules by means of cluster expansions. *J. Phys. Chem. B* **2012**, *116*, 4300–4309. [[CrossRef](#)] [[PubMed](#)]

26. Cannas, S.A. One-dimensional Ising model with long-range interactions: A renormalization-group treatment. *Phys. Rev. B* **1995**, *52*, 3034–3037. [[CrossRef](#)]
27. Ullner, M.; Jönsson, B. Monte Carlo study of titrating polyelectrolytes in the presence of salt. *Macromolecules* **1996**, *29*, 6645–6655. [[CrossRef](#)]
28. Nishio, T. Monte Carlo studies on potentiometric titration of poly(glutamic acid). *Biophys. Chem.* **1998**, *71*, 173–184. [[CrossRef](#)]
29. Zarzycki, P. Monte Carlo study of the topographic effects on the proton binding at the energetically heterogeneous metal oxide/electrolyte interface. *Langmuir* **2006**, *22*, 11234–11240. [[CrossRef](#)] [[PubMed](#)]
30. Zarzycki, P. Effective adsorption energy distribution function as a new mean-field characteristic of surface heterogeneity in adsorption systems with lateral interactions. *J. Colloid Interface Sci.* **2007**, *311*, 622–627. [[CrossRef](#)] [[PubMed](#)]
31. Ulrich, S.; Seijo, M.; Stoll, S. A Monte Carlo study of weak polyampholytes: Stiffness and primary structure influences on titration curves and chain conformations. *J. Phys. Chem. B* **2007**, *111*, 8459–8467. [[CrossRef](#)] [[PubMed](#)]
32. Companys, E.; Garcés, J.L.; Salvador, J.; Galceran, J.; Puy, J.; Mas, F. Electrostatic and specific binding to macromolecular ligands. A general analytical expression for the Donnan volume. *Colloid Surf. A* **2007**, *306*, 2–13. [[CrossRef](#)]
33. Madurga, S.; Garcés, J.L.; Companys, E.; Rey-Castro, C.; Salvador, J.; Galceran, J.; Puy, J.; Mas, F. Ion binding to polyelectrolytes: Monte Carlo simulations versus classical mean field theories. *Theor. Chem. Acc.* **2009**, *123*, 127–135. [[CrossRef](#)]
34. Garcés, J.L.; Rey-Castro, C.; David, C.; Madurga, S.; Mas, F.; Pastor, I.; Puy, J. Model-independent link between the macroscopic and microscopic descriptions of multidentate macromolecular binding: Relationship between stepwise, intrinsic, and microscopic equilibrium constants. *J. Phys. Chem. B* **2009**, *113*, 15145–15155. [[CrossRef](#)] [[PubMed](#)]
35. Carnal, F.; Ulrich, S.; Stoll, S. Influence of explicit ions on titration curves and conformations of flexible polyelectrolytes: A Monte Carlo study. *Macromolecules* **2010**, *43*, 2544–2553. [[CrossRef](#)]
36. Carnal, F.; Stoll, S. Chain stiffness, salt valency, and concentration influences on titration curves of polyelectrolytes: Monte Carlo simulations. *J. Chem. Phys.* **2011**, *134*, 044909. [[CrossRef](#)] [[PubMed](#)]
37. Aleksandrov, A.; Polydorides, S.; Archontis, G.; Simonson, T. Predicting the Acid/Base Behavior of Proteins: A Constant-pH Monte Carlo Approach with Generalized Born Solvent. *J. Phys. Chem. B* **2010**, *114*, 10634–10648. [[CrossRef](#)] [[PubMed](#)]
38. Baxter, R.J. *Exactly Solved Models in Statistical Mechanics*; Academic Press: London, UK, 1982.
39. Chandler, D. *Introduction to Modern Statistical Mechanics*; Oxford University Press: Oxford, UK, 1987.
40. Koper, G.J.M.; Borkovec, M. Ising models of polyprotic acids and bases. *Langmuir* **1994**, *10*, 2863–2865.
41. Koper, G.J.M.; Borkovec, M. Exact affinity distributions for linear polyampholytes and polyelectrolytes. *J. Chem. Phys.* **1996**, *104*, 4204. [[CrossRef](#)]
42. Vladimir, B.T. General transfer matrix formalism to calculate DNA-protein-drug binding in gene regulation: Application to OR operator of phage λ . *Nucleic Acids Res.* **2007**, *35*, 3519–3867.
43. Lee, A.A.; Kondrat, S.; Kornyshev, A.A. Single-file charge storage in conducting nanopores. *Phys. Rev. Lett.* **2014**, *113*, 048701. [[CrossRef](#)] [[PubMed](#)]
44. Poland, D.; Scheraga, H.A. *Theory of Helix-Coil Transitions in Biopolymers*; Academic Press: New York, NY, USA, 1970.
45. Garcés, J.L.; Koper, G.J.M.; Borkovec, M. Ion binding to polyelectrolytes. *J. Phys. Chem. B* **2006**, *110*, 10937–10950. [[CrossRef](#)] [[PubMed](#)]
46. Garcés, J.L.; Madurga, S.; Borkovec, M. Coupling of conformational and ionization equilibria in linear poly(ethylenimine): A study based on the site binding/rotational isomeric state (SBRIS) model. *Phys. Chem. Chem. Phys.* **2014**, *16*, 4626–4638. [[CrossRef](#)] [[PubMed](#)]
47. Flory, P.L. *Statistical Mechanics of Chain Molecules*; John Wiley: New York, NY, USA, 1969.
48. Flory, P.L. Foundations of Rotational Isomeric State Theory and General Methods for Generating Configurational Averages. *Macromolecules* **1974**, *7*, 381–392. [[CrossRef](#)]
49. Vila, J. Statistical mechanics in homopolypeptides. I. Theoretical approach for titration process in α helix. *J. Chem Phys.* **1986**, *84*, 6421–6425. [[CrossRef](#)]

50. Reed, C.E.; Reed, W.F. Monte Carlo study of titration of linear polyelectrolytes. *J. Chem. Phys.* **1992**, *96*, 1609–1620. [[CrossRef](#)]
51. Ullner, M. Comments on the Scaling Behavior of Flexible Polyelectrolytes within the Debye-Huckel Approximation. *J. Phys. Chem. B* **2003**, *7*, 8097–8110. [[CrossRef](#)]
52. Alfrey, T.; Morawetz, H. Amphoteric Polyelectrolytes. I. 2-Vinylpyridine-Methacrylic Acid Copolymers. *J. Am. Chem. Soc.* **1952**, *74*, 436–438.
53. Mazur, J.; Silberberg, A.; Katchalsky, A. Potentiometric behavior of polyampholytes. *J. Polym. Sci.* **1959**, *35*, 43–70. [[CrossRef](#)]
54. David, C.; Companys, E.; Galceran, J.; Garcés, J.L.; Mas, F.; Rey-Castro, C.; Salvador, J.; Puy, J. Competitive ion complexation to polyelectrolytes: Determination of the stepwise stability constants. The $\text{Ca}^{2+}/\text{H}^{+}$ /polyacrylate system. *J. Phys. Chem. B* **2007**, *111*, 10421–10430. [[CrossRef](#)] [[PubMed](#)]
55. David, C.; Companys, E.; Galceran, J.; Garcés, J.L.; Mas, F.; Rey-Castro, C.; Salvador, J.; Puy, J. Competitive $\text{Cd}^{2+}/\text{H}^{+}$ complexation to polyacrylic acid described by the stepwise and intrinsic stability constants. *J. Phys. Chem. B* **2008**, *112*, 10092–10100. [[CrossRef](#)] [[PubMed](#)]
56. Garcés, J.L.; Madurga, S.; Rey-Castro, C.; Mas, F. Dealing with long-range interactions in the determination of polyelectrolyte ionization properties. Extension of the transfer matrix formalism to the full range of ionic strengths. *J. Polym. Sci. Part B* **2017**, *55*, 275–284. [[CrossRef](#)]
57. Burak, Y.; Netz, R.R. Charge Regulation of Interacting Weak Polyelectrolytes. *J. Phys. Chem. B* **2004**, *108*, 4840–4849. [[CrossRef](#)]
58. Flory, P.J.; Jernigan, R.L. Second and fourth moments of chain molecules. *J. Chem. Phys.* **1965**, *42*, 3509–3519. [[CrossRef](#)]
59. Landsgesell, J.; Holm, C.; Smiatek, J. Wang-Landau Reaction Ensemble Method: Simulation of Weak Polyelectrolytes and General Acid-Base Reactions. *J. Chem. Theory Comput.* **2016**, *13*, 852–862. [[CrossRef](#)] [[PubMed](#)]
60. Landsgesell, J.; Holm, C.; Smiatek, J. Simulation of weak polyelectrolytes: A comparison between the constant pH and the reaction ensemble method. *Eur. Phys. J. Spec. Top.* **2017**, *226*, 725–736. [[CrossRef](#)]
61. Baptista, A.M.; Teixeira, V.H.; Soares, C.M. Constant-pH molecular dynamics using stochastic titration. *J. Chem. Phys.* **2002**, *117*, 4184–4200. [[CrossRef](#)]
62. Mongan, J.; Case, D.A.; McCammon, J.A. Constant pH Molecular Dynamics in Generalized Born Implicit Solvent. *J. Comput. Chem.* **2004**, *25*, 2038–2048.
63. Madurga, S.; Rey-Castro, C.; Pastor, I.; Vilaseca, E.; David, C.; Garcés, J.L.; Puy, J.; Mas, F. A semi-grand canonical Monte Carlo simulation model for ion binding to ionizable surfaces: Proton binding of carboxylated latex particles as a case study. *J. Chem. Phys.* **2011**, *135*, 1–11. [[CrossRef](#)] [[PubMed](#)]



Chapter 4

Effect of charge regulation and conformational equilibria in the stretching properties of weak polyelectrolytes

4.1 Summary

4.1.1 Introduction

The mechanical stretching response of a single polymer chain has been widely studied experimentally, theoretically and computationally. The setup of a single-molecule stretching experiment is the following.^{54,58} First, one chemically attaches the polymer chain to a surface. Then, a controlled force is applied to the macromolecule and the average extension of the chain is measured. The extension of the chain is usually measured using spectroscopic techniques, mainly Atomic Force Microscopy^{56,57,61,166,167} (AFM) and optical tweezers.^{58,59} Single-chain stretching experiments have proved that many interesting physico-chemical events can be triggered by the action of a mechanical force. For instance, one can induce isomerization reactions^{61,166,167}, break hydrogen bonding^{56,57} or even cause ring opening reactions, forbidden in normal conditions.^{168,169}

Several theoretical models have been proposed to explain the single-molecule stretching experiments.^{54,58,172} Many of them rely on the Freely Jointed Chain (FJC) and Worm-Like Chain (WLC) models, which propose a coarse-grained description of the polymer chain. Both models have been used to study a wide variety of synthetic polymers and biopolymers,^{5,54,66,173–175} obtaining good agreement with the experimental stretching response of neutral polymers and polyelectrolytes with negligible electrostatic interactions. However, strong deviations from the ideal FJC and WLC models have been found for polyelectrolytes where long-range coulombic interactions are important.⁷¹ In particular, the stretching response of single stranded nucleic acids have been found to exhibit two unique elastic regimes.^{77,78} Moreover, the elastic response of these strong polyelectrolytes was found to be dependent on the concentration and valence of the counterions.^{77,78} All these observations have motivated the development of the snake chain model,^{79,81} which seemingly is able to explain well the elastic response of strong polyelectrolytes.

The aforementioned experiments were conducted with strong polyelectrolytes (single stranded nucleic acids), where the charge is constant and not fluctuating. There are few experiments on the elastic response of weak polyelectrolytes, which have mainly focused in the effect of temperature or have been carried out in pH-conditions where charge regulation is neglectable.^{57,183} Thus, the impact on the elastic properties of weak polyelectrolytes of the conditions such as the pH or the ionic strength is still unknown.

In this chapter, it is investigated the physicochemistry involved in the mechanical stretching of a single weak polyelectrolyte. The SBRIS model polyelectrolyte introduced in the previous

chapter, which is outlined in Fig. 4.1a, is extended to study its elastic response. This is done performing SGCMC simulations at varying conditions of pH, ionic strength and stretching force. A snapshot of one of the simulations can be found in 4.1b. In the publication presented in this chapter, we answer the following questions:

- In the previous chapter, special focus have been placed in the coupling between the ionization and conformational degrees of freedom of weak polyelectrolytes. Is it possible to change the ionization state of a weak polyelectrolyte by mechanically stretching its chain?
- In affirmative case, until which extent is it possible to modify the polyelectrolyte charge and in which conditions?
- Are the force/extension curves of flexible weak polyelectrolyte dependent of the pH and ionic strength conditions?
- Which are the different force regimes observed in weak polyelectrolyte stretching?

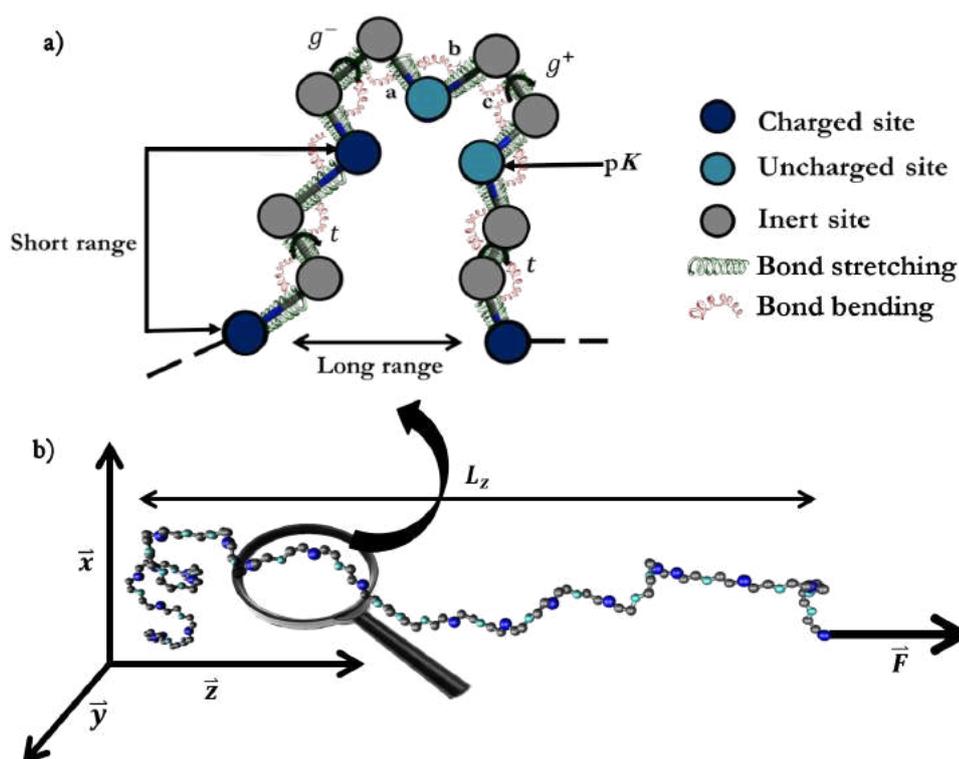


FIGURE 4.1: (a) Model weak polyelectrolyte used in the publication presented in this chapter. The polyelectrolyte has an ionizable site (blue circles) every three chain positions. Only the bonds joining two non-ionizable sites (grey circles) can change its conformation, which is restricted to *trans* (*t*), *gauche+* (g^+) and *gauche-* (*gauche-*) states. The ionizable sites can dynamically change its charge via protonation/deprotonation processes. (b) Snapshot taken from one of the SGCMC simulations exhibited in the publication presented in this chapter.

4.1.2 Results

The model polyelectrolyte used in the publication exhibited in this chapter aims to retain the more fundamental physicochemical aspects relevant in the stretching of a weak polyelectrolyte: proton binding, dihedral angle rotation, bond stretching and bending. The model can be considered an extension of the SBRIS model presented in the previous chapter and it follows the same assumptions listed in Section 3.1.2. In addition to those, two new assumptions are considered:

- The bond stretching and bond bending are included in the model with harmonic potential, which allow small thermal fluctuations of the bond length and bond angle around their equilibrium value.
- All the bonds are considered to have the same equilibrium bond length, equilibrium bond angle and harmonic constant values for the bending and stretching potentials.

The results presented in this section correspond to the case with $pK_a = 9$, $\sigma = 10$, $\epsilon_{u,t} = 1$, $\epsilon_{u,g} = \infty$, equilibrium bond length $l_0 = 1.5 \text{ \AA}$ and equilibrium bond angle $\alpha_0 = 120^\circ$. In the publication presented in this chapter, the case with $\sigma = 1$ is also studied and a more complete casuistry, including wider pH and ionic strength conditions is studied.

The elastic response of the model weak polyelectrolyte can be observed in Fig. 4.2, where the polyelectrolyte extension L_z is shown as a function of the stretching force applied F . In Fig. 4.2a, the force/extension are obtained at a constant ionic strength $I = 0.001 \text{ M}$ and pH-values

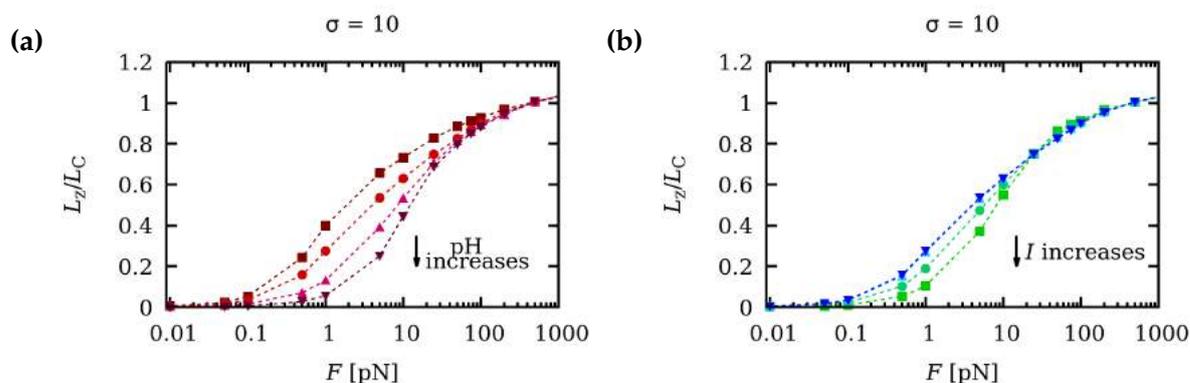


FIGURE 4.2: Chain extension L_z normalised to the contour length $L_C = Nl_0 \cos((\pi - \alpha_0)/2)$, where N is the number of bonds of the polyelectrolyte. (a) constant ionic strength $I = 0.001 \text{ M}$ and pH-values 4 (squares), 6 (circles), 8 (upwards triangles) and 10 (downwards triangles). (b) constant pH-value $\text{pH} = 6$ and ionic strengths: 1 M (green squares), 0.1 M (turquoise circles), 0.01 M (cyan upwards triangles) and 0.001 M (blue downwards triangles).

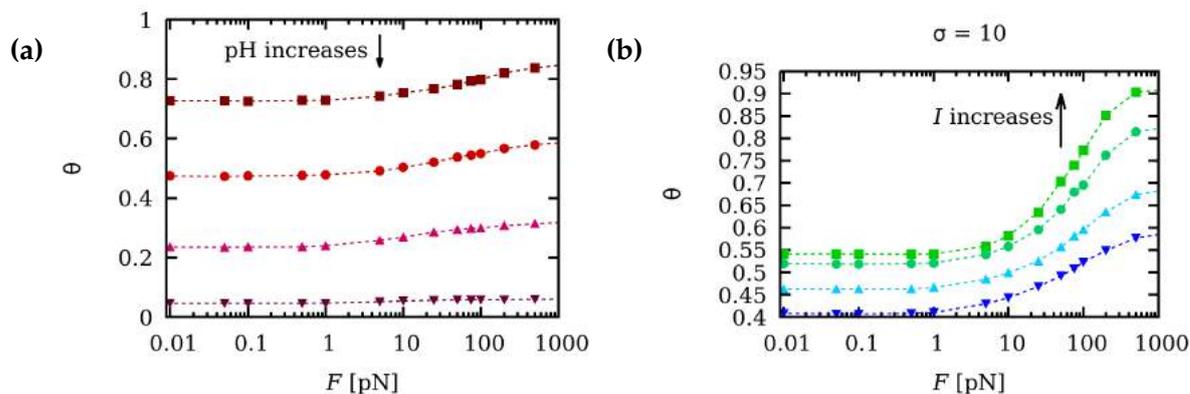


FIGURE 4.3: Average ionization degree θ as a function of the stretching force F for simulations with (a) constant ionic strength $I = 0.001$ M and pH-values ranging from 4 to 10 and (b) constant pH-value $\text{pH} = 6$ and I -values ranging from 1 M to 0.001 M. The markers follow the same symbols as the ones depicted in Fig. 4.2.

ranging from 2 to 10, from top to bottom. It can be observed that, for a given F -value, the polyelectrolyte is easier stretched at small pH-values. This can be understood because at low pH values most of the polyelectrolyte binding sites are protonated, *i.e.* charged. The intramolecular electrostatic repulsion between charged sites contributes to stretching the polymer chain. The same fundamental reasoning can be applied to explain the force/extension curves in Fig. 4.2b, which are obtained at constant $\text{pH} = 6$ and I -values ranging from 0.001 M to 1 M, from top to bottom. In this case, for a given F -value, the chain is easier stretched at low I -values because the intramolecular electrostatic interaction is less screened. It is important to note that those effects only occur for $0.5 \leq F \leq 100$ pN.

The effect of the mechanical stretching of the polyelectrolyte on its ionization is investigated analyzing the degree of ionization θ as function of F , which is depicted in Fig. 4.3. The SGC MC results shown in Fig. 4.3a are obtained at a constant ionic strength $I = 0.001$ M and pH-values ranging from 4 to 10 and the ones in Fig 4.3b at a constant pH-value $\text{pH} = 6$ and I -values ranging from 1 M to 0.001 M. It can be observed that θ increases in increasing F for all the cases of study. This is because the charged sites are separated when the polyelectrolyte chain is elongated. As a result, the electrostatic repulsion among charged sites is reduced and the system admits more charged sites. Note that, again, the force regime where the θ increase is observed is also $0.5 \leq F \leq 100$ pN. It is also relevant to highlight that this effect is more important at intermediate pH-values and high I -values. The reason of this latter observation will be explained in the following chapter.

The conformational properties of the polyelectrolyte chain can be assessed in Fig. 4.4, where the polymer persistence length l_p is shown as a function of F . Again, the left-side panel (Fig. 4.4a) corresponds to SGC MC simulations at constant ionic strength $I = 0.001$ M while those in

the right-side panel (Fig. 4.4b) correspond to results at constant pH-value $\text{pH} = 6$. It can be observed that the chain is stiffer (larger l_p values) at low pH and I values. This can be justified since at these conditions the intramolecular repulsive electrostatic interactions are more important, which cause the chain to be more extended (at low pH values the polyelectrolyte is more charged and at low I values the electrostatic interactions are less screened). Note that, while l_p is constant for $F < 1$ pN, it increases in increasing F otherwise. This is because the mechanical stretching elongates the chain which makes the chain stiffer (in turn increasing l_p). A deeper analysis of the conformational properties of the polyelectrolyte, including the response of the rotational states to the mechanical stretching, can be found in the publication included in this chapter.

Three different force regimes have been identified. At low forces $F < F_E = k_B T / l_p \simeq 1$ pN, the chain behaves as a set of segments with characteristic length l_p .⁵ The existence of this "entropic" force regime can be assessed in Fig. 4.4, where it can be observed that l_p remains constant in this force regime. In Fig. 4.5, the low force regime is analyzed in detail at pH-values ranging from 2 to 10, from top to bottom. The simulated elongation results (markers) are fitted to the linear prediction of the force/extension response Eq. 1.49 (dashed lines) and to the Pincus scaling law Eq. 1.51 (continuous lines). It can be observed that at very low forces, $F < 0.1$ pN the elongation follow a linear behavior. For higher forces, $0.1 \text{ pN} < F < 1 \text{ pN}$, the simulations results deviate from the linear regime and follow the Pincus scaling law. The Pincus scaling exponent, ν , exhibits two asymptotic behaviors: at low pH values (where the chain is charged) it tends to $\nu = 3/5$ while at high pH values (where the chain is uncharged) it tends to $\nu = 1/2$. The first value was predicted by Pincus⁶⁴ for strong polyelectrolytes and it is result of the swelling of the polyelectrolyte due to the intramolecular electrostatic repulsion of its charged groups. The latter limiting value recovers the linear regime, expected for a neutral polymer following ideal chain statistics. Interestingly, at intermediate pH-values a transition between both limiting

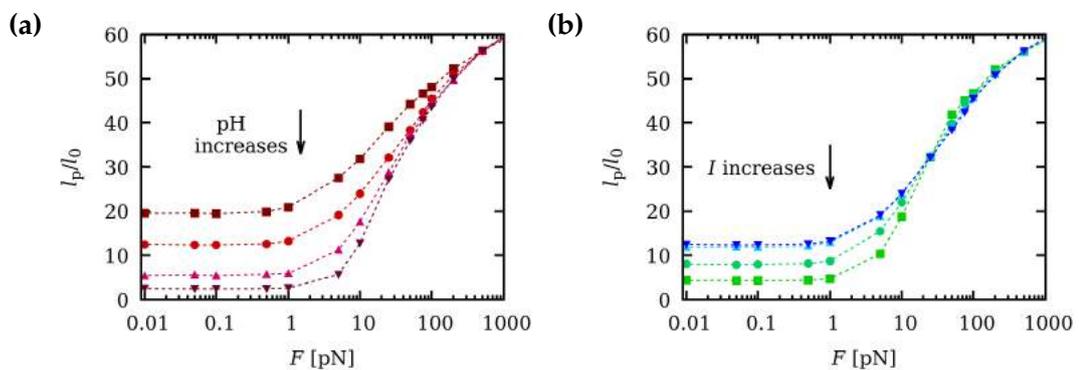


FIGURE 4.4: Persistence length l_p normalized to the equilibrium bond length l_0 as a function of the stretching force F for simulations with (a) constant ionic strength $I = 0.001$ M and pH-values ranging from 4 to 10 and (b) constant pH-value $\text{pH} = 6$ and I -values ranging from 1 M to 0.001 M. The meaning of the markers used is the same as in Fig. 4.2.

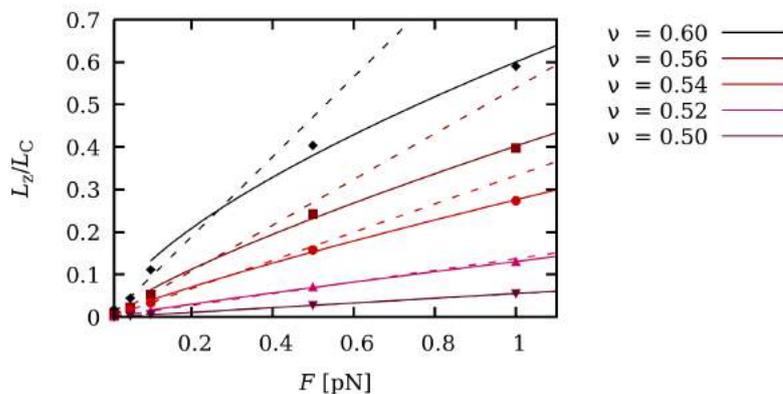


FIGURE 4.5: Force/extension curves obtained in the low-force regime at constant $I = 0.001$ M and pH values from top to bottom of 2, 4, 6, 8 and 10. Markers corresponds to the results obtained with SGCMC. Lines show the best of the simulation results to the linear Eq. 1.49 (dashed lines) and to the Pincus scaling law 1.51 (continuous lines).

situations is observed, suggesting that weak polyelectrolytes can be understood as intermediate situation between strong polyelectrolytes and neutral chains. This transition is analyzed in more detail in the following chapter.

At intermediate forces ($1 \text{ pN} < F < 100 \text{ pN}$), the mechanical stretching force is directly affecting the conformational properties of the polymer. This fact can also be observed in Fig. 4.4, where l_p increases with F for $F > 1 \text{ pN}$. Further prove of this affirmation is provided in the publication at the end of this chapter, where the dynamics of the conformational bond states is analyzed in the full force range. It is in this regime, in which the stretching force changes the conformation of the polyelectrolyte, when the ionization is also modified, as can be observed in Fig. 4.3. Finally, at high forces $F > 100 \text{ pN}$, the chain is almost fully elongated, as can be observed in Fig. 4.2. In this high force regime, the mechanical force starts to elongate the bond angle and bond length, in this order. In the publication presented in this chapter, this latter affirmation is assessed studying the average values of the bond length and the bond angle. The results indicate that their values only increase with the stretching force in this high force regime.

4.1.3 Conclusions

In this chapter, the physico-chemistry involved in the mechanical stretching of a weak polyelectrolyte chain is investigated with Semi-Grand Canonical Monte Carlo (SGCMC) simulations at different pH, ionic strength (I) and stretching force F conditions. The chosen model for the weak polyelectrolyte is an extension of the SBRIS model presented in the previous chapter, with an ionizable site very three chain positions. The conformational (persistence length l_p and the probabilities of the *gauche* states), ionization (average degree of ionization θ and effective

protonation constant) and elastic (chain extension L_z are analyzed in the full range of stretching forces at varying conditions of pH and I .

Three different force regimes are identified. At low forces $F < 1$ pN, the polymer behaves as a structureless chain of segments of constant length equal to l_p . In this "entropic" force regime, l_p is constant and independent of the stretching force. The Pincus scaling exponent ν is found to exhibit two limiting behaviors with the pH. At low pH-values (highly charged chain) ν tends to $3/5$, which is the value theoretically predicted for a strong polyelectrolyte. At high pH-values (neutral chain) ν tends to $1/2$ and the linear regime is found, which is the theoretical prediction for ideal neutral chain. At intermediate pH-values, ν is found to have a transition between these two limiting values, indicating that weak polyelectrolytes have an intermediate behavior between a strong polyelectrolyte and a neutral ideal chain.

At intermediate forces $1 \text{ pN} < F < 100$ pN, the stretching force acts in the rotational degrees of freedom. The stretching force promotes *trans* conformational state, where the polymer is more elongated, and the *gauche* state probabilities are observed to decay. In turn, l_p becomes force-dependent and it starts to increase in increasing F . In this force regime, θ is observed to increase as F increases. This is because in elongating the chain, one separates the distance between charged sites, diminishing the intramolecular electrostatic repulsion and allowing the creation of new charged sites. Moreover, the force/stretching curves are found to be pH and I dependent on this force regime. As a general trend, the polyelectrolyte is found to be easier stretched in conditions of significant electrostatic interaction (low pH and I values). In summary, it is in the intermediate force regime where the elastic, conformational and ionization properties are highly coupled, exhibiting an interesting and complex interplay.

At high forces $F > 100$ pN, the force/stretching curves are found to be independent of the I and pH values. The chain is fully stretched and the bonds are found to be all in *trans* states. In this force regime, the stretching force is found to first increase the average value of the bond angle and then increase the average bond length.

4.2 Publication

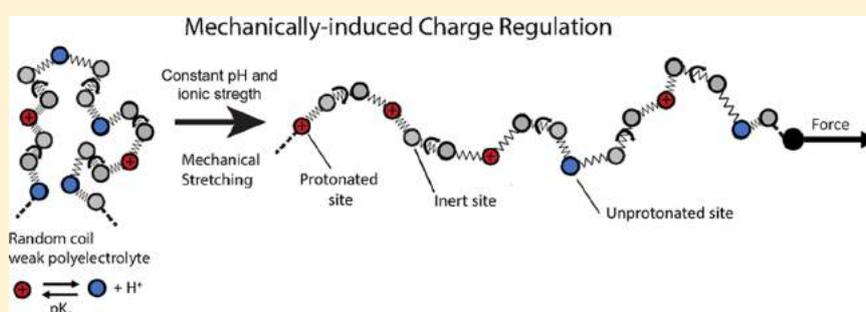
Effect of Charge Regulation and Conformational Equilibria in the Stretching Properties of Weak Polyelectrolytes

Pablo M. Blanco,^{*,†} Sergio Madurga,^{*,†} Francesc Mas,[†] and Josep L. Garcés[‡]

[†]Physical Chemistry Unit, Materials Science and Physical Chemistry Department & Research Institute of Theoretical and Computational Chemistry (IQTCUB) of Barcelona University (UB), Barcelona 08028, Catalonia, Spain

[‡]Chemistry Department, Technical School of Agricultural Engineering & AGROTECNIO of Lleida University (UdL), Lleida 25198, Catalonia, Spain

Supporting Information



ABSTRACT: Weak polyelectrolytes can modulate their charge in response to external perturbations, such as changes in the pH, ionic strength (I), or electrostatic interactions with other charged species, a phenomenon known as charge regulation (CR). On the other hand, it is well established that CR is highly coupled with the conformational degrees of freedom. In this paper, the influence of CR in the stretching properties of weak polyelectrolytes is analyzed, and the possibility of CR induced by mechanical stretching is explored. With this aim, we make use of a minimal model, which captures the fundamental aspects present in the stretching of a flexible weak linear polyelectrolyte: internal angle rotation, bond stretching, bond bending, and proton binding, which is the paradigmatic mechanism of CR. The angle rotation is described by using the rotational isomeric state approximation, while for protonation, the site binding model is assumed. Mechanical stretching is studied by performing semi-grand canonical Monte Carlo simulations at different pH and ionic strength conditions. The simulations simultaneously provide both conformational (bond state probabilities, persistence length l_p , and chain elongation) and protonation properties (degree of protonation θ and the effective protonation constant K_c). The obtained force–extension curves suggest that the pH value and the ionic strength I have a significant effect on polyelectrolyte stretching. Three different force regimes can be observed. For large forces ($F > 100$ pN for typical force constants), the force–extension curve is almost independent of the pH and I . For low forces, the persistence length l_p is force-independent, although it strongly increases with the pH value. Under this regime, linear and Pincus scaling behaviors are observed. Finally, in the intermediate-force regime, both rotational and protonation degrees of freedom are mechanically activated, and the picture becomes more complicated. It is found that l_p increases with F and, under certain conditions, a significant increase of θ with F is observed, indicating that CR could in principle be induced by means of mechanical stretching. This fact can be explained by analyzing the coupling between θ and the probability of a bond to be in the *gauche* state $P(g)$. $P(g)$ decreases with F as the bonds adopt the *trans* conformation so that the electrostatic repulsion is reduced and θ increases. Finally, the intricate interplay between short-range and long-range interactions is analyzed, leading to apparently contradictory behaviors ($P(g)$ and l_p simultaneously decrease with I), which can only be explained by CR and the presence of complex spatial correlations.

INTRODUCTION

In the last two decades, the development of single-molecule force spectroscopy has led to an extraordinary expansion of the field of mechanochemistry.^{1,2} By applying a controlled external force to a molecule that is chemically attached to a surface, a wide range of mechanically induced physicochemical events can take place. Just to mention a few examples, AFM has been used to mechanically induce *cis*-to-*trans* isomerization of carbon–carbon double bonds,³ prolyl *cis*-*trans* isomeriza-

tion,^{4,5} or conformational chair–boat transitions or hydrogen bond breaking in polysaccharides.^{6,7} Some ring-opening reactions, normally forbidden by orbital symmetry, become possible if a tensile force is applied to the polymer chain.^{8,9} Single-molecule AFM experiments have been recently used in

Received: June 12, 2019

Revised: September 16, 2019

Published: October 16, 2019

monitoring force-dependent enzyme catalysis¹⁰ and surface desorption of polypeptides;^{11,12} the characterization of new supramolecular polymers based on host-enhanced π - π interaction;¹³ or the design of mechanophores embedded in macrocycles, which allows pinpointing of the mechanochemical bond rupture.¹⁴ Optical tweezers have also been used in the study of the elastic properties of biomacromolecules such as single-stranded DNA (ss-DNA).^{2,15}

In parallel to the experimental work, several theoretical approaches have been developed, differing in the detail of description of the macromolecular structure.^{1,2,16} In the freely jointed chain (FJC) model, the polymer chain is represented at the coarse-grain level by a set of rigid links joined with fully random orientations. Although able to account for the stretching properties of a wide variety of synthetic polymers with different structures and solvents,¹ this model was shown to present clear deviations from the elastic response of many other macromolecules of interest, such as double-stranded DNA (ds-DNA). Aiming at overcoming these limitations, Marko and Siggia modeled the polymer as a worm-like chain (WLC), which, assuming exponential decaying correlations between chain segments, accounted for the capability of the chain to deform on short-length scales.¹⁷ The resulting high-force regime matched very well with a variety of polymers for which electrostatic interactions can be neglected, such as some polypeptides¹⁸ or ds-DNA.¹⁹ Models including freely rotating bonds,²⁰ bond elasticity,²¹ or ligand-receptor equilibria²² have also been proposed, leading to theoretical predictions of new force-extension regimes.

As an alternative to these coarse-graining approaches, theoretical methods based on first principles, which account for the detailed atomistic structure of the macromolecular backbone, have also been proposed.²³ In most of these studies, ab initio calculations are first performed in order to detect the more stable conformational states of the interacting monomers at different elongations of the bonds.²⁴ Once the structural microscopic information is available, the necessary thermal averages are performed by using Monte Carlo (MC) or transfer matrix techniques.^{25,26} The resulting scheme has been successful in reproducing the experimental force-extension curves of several polymers.^{20,27-29} In particular, the stretching behavior of poly(ethylene glycol) (PEG) has been analyzed in detail.^{27,30,31} In essence, this methodology can be regarded as a generalization of the rotational isomeric state (RIS) model developed mainly by Flory to study the conformational properties of linear chains^{32,33} in which only the rotational states of minimum energy (commonly *trans*, *gauche+*, and *gauche-*) are taken into account in the computation of the thermal averages.

The methods mentioned above only account for short-range interactions and cannot thus be applied to charged macromolecules for which the long-range Coulombic forces cannot be neglected. The presence of self-avoiding electrostatic forces produces, however, new elastic regimes, which strongly deviate from the ideal, non-interacting FJC and WLC models.³⁴ The resulting stretching behavior, mainly studied in single-stranded nucleic acids (ss-DNA and ss-RNA), is extremely dependent on the valence and concentration of the counterions^{2,15,35-39} and seems to be well explained by the recently proposed "snake chain model".^{40,41} This model was motivated by recent MC simulations with explicit ions,^{42,43} which suggested two elastic regimes. At low forces, the polyelectrolyte behaves as a set of swollen electrostatic blobs on a long-length scale, while

at high forces, a short-length, ion-stabilized, crumpling structure is detected.⁴¹⁻⁴⁴ Since ss-DNA and ss-RNA are strong polyelectrolytes, in those studies, the macromolecular charge is considered constant and independent of the degree of stretching.

However, this could not be the case of weak polyelectrolytes for which the charge is in general a dynamical and fluctuating variable. This fact leads to the phenomenon of charge regulation (CR), defined as the capability of weak polyelectrolytes to modulate their ionization state as a response to some physicochemical perturbation.^{45,46} The main aim of the present work is to study the influence of CR in the stretching properties of weak polyelectrolytes. The possibility of mechanically induced CR will also be explored. CR is ubiquitous in a wide range of processes of biological, environmental, and technological interest. A few examples are the stability of colloidal systems and nanoparticle coatings,^{47,48} receptor-ligand interactions in biochemical systems,⁴⁹ and protein-protein⁵⁰ and protein-surface interactions,⁵¹ among many others, which can be found in ref 52 and references quoted therein. The paradigmatic mechanism for CR is the binding of protons and other small ions present in the backward medium. Although CR can take place in rigid structures such as nanoparticles or surfaces, most polyelectrolytes are flexible so that alterations in the ionization state induce changes in the rotational states of the bonds. Sometimes, this can even produce dramatic conformational transitions in the global macromolecular structure, such as the helix-coil transitions of polypeptides⁵³ or the abrupt swelling of poly(methacrylic acid) in a very narrow range of pH values.⁵⁴

The mechanochemistry of weak polyelectrolytes is still a fairly unexplored area from the experimental, theoretical, and computational point of view. Although some AFM experiments^{7,55} have been performed on weak polyelectrolytes (such as hyaluronic acid), they have been either focused on the temperature effect⁷ or carried out at pH conditions where CR is negligible.⁵⁵ As a result, the effect of other environmental variables such as pH or salt concentration is still unknown. In this work, we introduce a minimal model that captures the fundamental aspects present in the stretching of a flexible weak linear polyelectrolyte: internal angle rotation, bond stretching, bond bending, and proton binding. The model presented is based on the site binding rotational isomeric state (SBRIS) model.^{52,56-58} The model is implemented in a Monte Carlo simulation scheme in the semi-grand canonical ensemble (SGCMC) widely used in computational modeling of CR phenomena.^{51,59-66} An outline of the used methodology is reported in [Model and Simulations](#), while the [Results and Discussion](#) section is devoted to analyzing the behavior of both the conformational (bond state probabilities, persistence length, and chain extension) and protonation properties (degree of protonation θ).

MODEL AND SIMULATIONS

Minimal Site Binding Rotational Isomeric State (SBRIS) Model of Stretched Weak Polyelectrolytes. In this work, we will make use of a model, which, containing a minimum number of parameters, still captures the fundamental aspects present in the stretching of a flexible weak linear polyelectrolyte: internal angle rotation, bond stretching, bond bending, and proton binding. The polyelectrolyte, outlined in [Figure 1a](#), can be considered a simplification of a previously

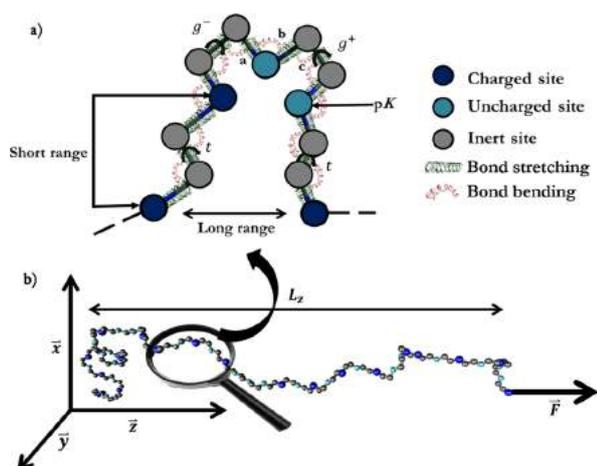


Figure 1. (a) Outline of the model of a weak polyelectrolyte proposed in the present work. The monomers are represented as sites joined by flexible harmonic bonds. Only the rotational states corresponding to minimum energy (rotational isomeric state approximation), that is, *trans* (t), *gauche+* (g^+), and *gauche-* (g^-), are taken into account. In order to minimize the number of parameters, only the bonds holding protonable sites (c bonds) are allowed to rotate, while the rest of the bonds (a and b) are forced to be in the *trans* state. The macromolecular chain is considered symmetric so that g^+ and g^- have the same energy ε_σ . Two kinds of sites are considered: inert sites (gray) and protonating sites. The latter can adopt two possible states: protonated (dark blue) or deprotonated (cyan), with protonation constant K (site binding model). Long-range (LR) and short-range (SR) interactions are treated in a different way. LR interactions (mediated by the solvent) are described by the Debye–Hückel potential. Conversely, neighboring protonated sites linked by a c bond in the *trans* state interact with energy $\varepsilon_{u,t}$ since SR interactions are mainly mediated by the macromolecular skeleton. Two neighboring sites linked by a c bond in the *gauche* state cannot be simultaneously protonated due to the huge electrostatic repulsion ($\varepsilon_{u,g} \rightarrow \infty$). (b) Snapshot from a semi-grand canonical Monte Carlo (SGCMC) simulation of the stretching of a weak polyelectrolyte with $pK = 9$, $pH = 6$, $I = 10^{-3}$ M, $F = 10$ pN, $\varepsilon_\sigma = -1$, $\varepsilon_{u,t} = 1$, $l_0 = 1.5$ Å, and $\alpha_0 = 120^\circ$. Both the extension and the degree of protonation depend on the applied force in the z axis direction.

proposed model for linear poly(ethylenimine) (LPEI).⁵⁷ A similar model has been recently proposed to study the role of long-range interactions in the conformational/protonation coupling.⁵² Let us assume that the chain is symmetric (i.e., the chain has a plane of symmetry when it is completely elongated), thus avoiding the question of tacticity, and contains a protonating site situated every three chain positions. In Figure 1, inert and protonating sites are depicted in gray and blue, respectively. A macromolecule with N protonating sites thus contains $M = 3N - 3$ bonds.

The protonation equilibria are treated using the site binding (SB) model for which the protonating sites can adopt two possible states: protonated (dark blue) and deprotonated (cyan). Within the SB approach, the ionization state of the macromolecule can be characterized by a set of variables $s = \{s_i\}$, $i = 1, \dots, N$, with values 0 (deprotonated) or 1 (protonated). On the other hand, the conformational degrees of freedom are treated assuming the rotational isomeric state (RIS) approximation;^{32,33} that is, only the rotational states corresponding to local energy minima are taken into account, typically *trans* (t), *gauche+* (g^+), and *gauche-* (g^-). A conformational state of the macromolecule can now be defined

by a set of variables $c = \{c_j\}$, $j = 1, \dots, M$ where c_j is the row vector with as many components as the number of states can adopt the bond j . In our case, each c_j can only take three different values: $c_j = (1, 0, 0)$ if bond j is in *trans*, $c_j = (0, 1, 0)$ if it is in *gauche+*, and $c_j = (0, 0, 1)$ if it is in *gauche-*.

Since the chain is symmetric, *gauche+* and *gauche-* states have the same energy. For simplicity, let us also assume that only the bonds with adjacent protonating sites (c bonds in Figure 1a) are allowed to rotate, while the rest of the bonds (a and b) are forced to be in the *trans* state. This approximation has been previously used in the modeling of stretching properties of poly(ethylene glycol) (PEG),²⁷ the neutral counterpart of the model proposed here. As a result, in our model, only $N - 1$ bonds from the total M bonds are allowed to rotate. Finally, we introduce the possibility of elastic bond stretching and bending.

Combining the SB and RIS approaches, we obtain the SBRIS model, which deals with ionization and conformational equilibria simultaneously.^{52,56,57} The resulting free energy \mathcal{F} can be expressed as the sum of five contributions

$$\mathcal{F} = W + \mathcal{F}_{\text{length}} + \mathcal{F}_{\text{angle}} + \mathcal{F}_{\text{SR}} + \mathcal{F}_{\text{LR}} \quad (1)$$

The term

$$W = -F r \quad (2)$$

represents the mechanical work exerted by the applied force F , which is considered to act in the z axis direction, as shown in Figure 1, and r is the polyelectrolyte chain end-to-end vector. $\mathcal{F}_{\text{length}}$ and $\mathcal{F}_{\text{angle}}$ quantify the elastic deformation of the length and the angles of the M bonds, respectively, which can be important at large forces. In this work, they are represented by the harmonic potentials

$$\mathcal{F}_{\text{length}} = \sum_{j=1}^M \frac{k_{\text{length},j}}{2} (l_j - l_{j,0})^2 \quad (3)$$

and

$$\mathcal{F}_{\text{ang}} = \sum_{j=1}^{M-1} \frac{k_{\text{angle},j}}{2} (\alpha_j - \alpha_{j,0})^2 \quad (4)$$

where l_j , α_j , $l_{j,0}$ and $\alpha_{j,0}$ represent, respectively, the length, the bending angle, the equilibrium length, and the equilibrium bending angle of bond j . Finally, $k_{\text{length},j}$ and $k_{\text{angle},j}$ denote the bond stretching and bending force constants. Note that the geometrical parameters and the force constants in the potentials (eqs 3 and 4) do not depend on the ionization state of the sites. At this level of description, this is a reasonable approximation, according to quantum-mechanical computations, which show only small variations in the bond lengths (see, for instance, the results for LPEI at different degrees of ionization⁶⁷). On the other hand, as will be shown in the next section, the bond bending and bond stretching will be essentially induced by the mechanical work at high enough forces rather than by electrostatic repulsions.

The electrostatic/conformational interaction free energy has been split into short-range (SR) \mathcal{F}_{SR} and long-range (LR) \mathcal{F}_{LR} contributions, as depicted in Figure 1. This distinction becomes necessary due to the fundamental differences in the physical chemistry of SR and LR interactions. It is well established that LR interactions are chemically unspecific,

mediated by the solvent, and can be reasonably approximated by a simple pair-interaction continuous force field. Conversely, SR interactions between neighboring sites and bonds are mediated by the macromolecular skeleton so that they depend on the detailed chemical environment of the site. As a result, they cannot be described by simple potentials, and specific interaction parameters must be used. These parameters will depend on the particular rotational state of the bond connecting the two protonating sites (in our case, the c bonds). \mathcal{F}_{SR} is the result of two contributions

$$\mathcal{F}_{\text{SR}} = \mathcal{F}_{\text{rot}}(c) + \mathcal{F}_{\text{p}}(s, c), \quad (5)$$

$\mathcal{F}_{\text{rot}}(c)$, corresponding to the classical RIS model,^{32,33} represents the conformational energy of the bonds for a given conformational state $c = \{c_j\}$ when the polyelectrolyte is uncharged. On the other hand, $\mathcal{F}_{\text{p}}(s, c)$ includes the binding free energy and the SR electrostatic interaction between charged sites and accounts for the coupling between the conformational and ionization degrees of freedom. The term \mathcal{F}_{rot} can be expressed as

$$\frac{\beta \mathcal{F}_{\text{rot}}(c)}{\ln 10} = \sum_{j=1}^M \epsilon_j c_j^T + \sum_{j=1}^M \epsilon_j E_j c_j^T + \dots \quad (6)$$

where $\beta = 1/k_{\text{B}}T$ is the inverse of the thermal energy, ϵ_j is a row vector whose elements are the free energies associated with the rotational state of bond j , and E_j is a square matrix containing the interaction energies between the neighboring bonds j and $j + 1$. ϵ and E are expressed in thermal units and divided by a factor $\ln 10$ in order to be compared with the pH scale. The sum in eq 6 could be extended to take into account three-bond interactions, four-bond interactions, and so on. In this work, however, we only considered SR interactions involving the first neighbor bonds. Following Flory,^{32,33} we choose as the ground state the configuration with all the bonds in the *trans* state. In this way, the rotational parameters ϵ_j and E_j can be expressed as

$$\epsilon_j = (0, \epsilon_{\sigma}, \epsilon_{\psi}, \epsilon_{\omega}), \quad E_j = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \epsilon_{\psi} & \epsilon_{\omega} \\ 0 & \epsilon_{\omega} & \epsilon_{\psi} \end{pmatrix}_j \quad (7)$$

where $\epsilon_{\sigma, j}$ is the free energy of the *gauche* states, while $\epsilon_{\psi, j}$ and $\epsilon_{\omega, j}$ are related to the interaction energies between two consecutive *gauche* states with the same and different orientation, respectively. The Boltzmann factor corresponding to ϵ_{α} is precisely α , that is, $\epsilon_{\alpha} = -\log \alpha$ ($\alpha = \sigma, \psi, \omega$). In choosing this notation, the resulting Boltzmann factors are denoted by the same symbols used in previous works,^{33,56–58} which make use of the transfer matrix approach. For instance, the transfer matrix corresponding to the free energy (eq 6) is given by

$$\mathbf{T}_j = \begin{pmatrix} 1 & 1 & 1 \\ 1 & \sigma_{\psi} & \sigma_{\omega} \\ 1 & \sigma_{\omega} & \sigma_{\psi} \end{pmatrix}_j \quad (8)$$

On the other hand, the second term in eq 5, \mathcal{F}_{p} , can be expressed in terms of the ionization state $s = \{s_i\}$ by means of the cluster expansion⁵²

$$\frac{\beta \mathcal{F}_{\text{p}}(s, c)}{\ln 10} = \sum_{i=1}^N \mu_i s_i + \sum_{i=1}^{N-1} \epsilon_{u, i} \epsilon_{3i-1} s_i s_{i+1} + \dots, \quad (9)$$

where $\mu_i = \text{pH} - \text{p}K_i = -\log(Ka_{\text{H}})$ is the reduced chemical potential of the ionizable site i , which depends on the proton activity, a_{H} , and the $\text{p}K$ values of the protonation constant of site i , $\text{p}K_i$. ϵ_u is a row vector whose components correspond to the electrostatic interaction energy between the neighboring charged sites i and $i + 1$. The intensity of the interaction is determined by the rotational state of the c bond between the two sites (bond number $3i - 1$ in the chain)

$$\epsilon_{u, i} = (\epsilon_{u, t}, \epsilon_{u, g}, \epsilon_{u, g})_i \quad (10)$$

In eq 10, we have assumed that only the bonds of type c are able to rotate. Note that in the SBRIS model, the SR electrostatic interactions depend on the conformation of the bond linking of the sites, which couples the ionization and the conformational degrees of freedom. As in eq 6, the sum in eq 9 could be extended to take into account triplet interactions, quadruplet interactions, and so on. In this work, however, only neighboring pair SR interactions will be taken into account.

Finally, as in most of the previous literature,^{57,68} LR electrostatic interactions will be described by the Debye–Hückel (DH) potential

$$\beta \mathcal{F}_{\text{LR}} = \sum_{i=1}^N \sum_{j=i+2}^N \frac{l_{\text{B}}}{d_{ij}} e^{-\kappa d_{ij}} s_i s_j \quad (11)$$

where $l_{\text{B}} \approx 0.7$ nm is the Bjerrum length in water at 298.15 K, d_{ij} is the distance between sites i and j , and $\kappa^{-1}(\text{nm}) = 0.304/\sqrt{I(M)}$ is the Debye length for water 298.15 K at ionic strength I .

Since we are interested in a model with the minimum number of parameters in order to analyze the fundamental aspects of the stretching, in this work, we will restrict ourselves to the special situation in which all the bonds have the same length, bond angle ($l_{j,0} = l_0$ and $\alpha_{j,0} = \alpha_0$), and force constants ($k_{\text{length}, j} = k_{\text{length}}$ and $k_{\text{angle}, j} = k_{\text{angle}}$). Moreover, we consider that all the protonating sites are identical ($\text{p}K_i = \text{p}K$) and the possible end effects of the chain are neglected so that $\epsilon_i = \epsilon$ and $\epsilon_{u, i} = \epsilon_u$. It is also assumed that when two neighboring sites are charged, the very strong SR repulsion hinders the *gauche* conformation of the c bond so that $u_{\text{g}} = 0$ or $\epsilon_{u, \text{g}} \rightarrow \infty$. Since one of every two consecutive bonds is always in the *trans* state, the interaction terms ϵ_{ψ} and ϵ_{ω} become irrelevant, and they can be taken as zero without loss of generality. To summarize, the model presented involves the following assumptions:

1. The SBRIS model is used to describe the conformational and protonation equilibria on the same foot.
2. The molecule contains one protonating site every two inert groups, as shown in Figure 1.
3. The ionizable sites are identical, with the same $\text{p}K$ value, and the bonds have the same length, bending angle, and constant forces.
4. Only the bonds of type c are allowed to rotate, while bonds of types a and b are constrained to be in the *trans* state. In practice, this implies that the rotation of the bonds is independent when the macromolecule is fully uncharged.
5. LR interactions are described by the DH potential, which accounts for screening effects so that co- and

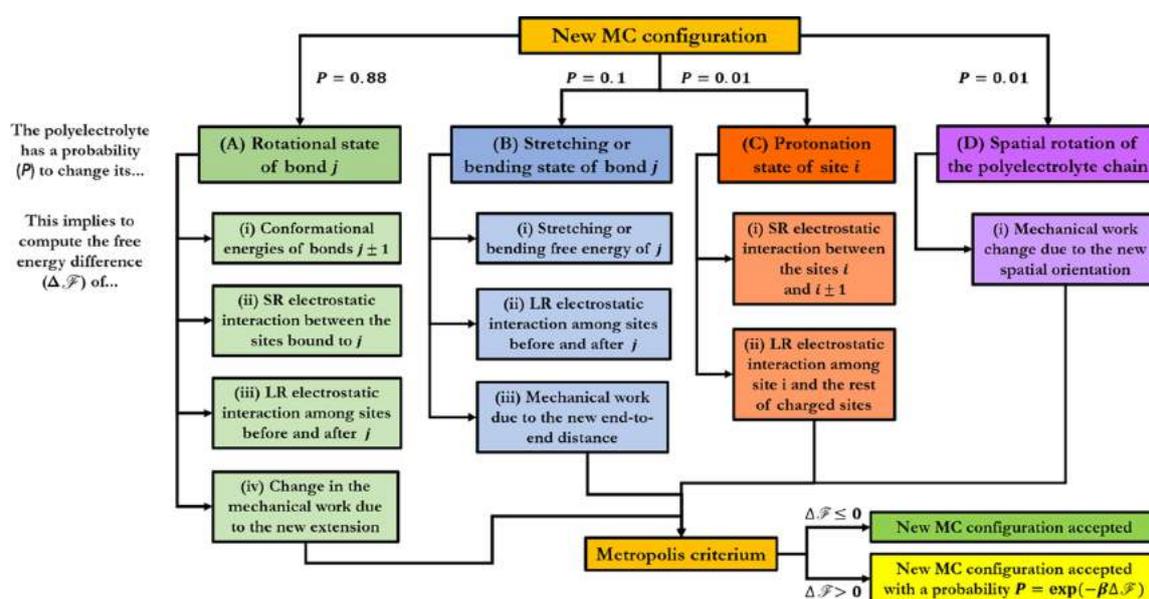


Figure 2. Metropolis algorithm of the SMGMC simulation code. In each new MC configuration, the polyelectrolyte can change either (A) the rotational state of a bond, (B) the length or angle of a bond, (C) the ionization state of a binding site, or (D) the spatial orientation of the polyelectrolyte chain in a laboratory coordinate frame with trial probabilities of 0.88, 0.1, 0.01, and 0.01, respectively.

counterions intervene only in an effective way. Only excluded volume effects induced by electrostatics are taken into account.

- Specific parameters are used to describe interactions between neighboring sites. Moreover, when two neighboring sites are simultaneously protonated, the c bond linking the sites cannot adopt the *gauche* state.
- As a result, the parameters involved in the model are ε_σ (free energy difference between *gauche* and *trans* states), interaction energy between neighboring sites ($\varepsilon_{u,i}$) when the c bonds are in the *gauche* state, equilibrium length and equilibrium angle of the bonds (l_0 and α_0), and constant forces for the bending and bond stretching (k_{length} and k_{angle}). The control variables are the reduced chemical potential of the protonating sites ($\mu = \text{pH} - \text{pK}$) and the ionic strength (I).

Finally, note that in the absence of LR interactions ($\mathcal{F}_{\text{LR}} = 0$), that is, at high enough ionic strengths, the model can be exactly solved by using the transfer matrix (TF) method.^{52,57,58} When applied to calculate stretching properties, the resulting TF combine conformational energies, by means of TF of the type (eq 8), and geometrical restrictions imposed by the macromolecular skeleton. In the absence of ionization processes, this method has been used to study the stretching of chains with freely rotating bonds²⁰ and the stretching properties of POE.^{20,27} However, in this work, we are especially interested in the effect of electrostatic interactions that are long ranged, and the LR term (eq 11) cannot be neglected. As a consequence, the transfer matrix approach would be too restrictive since only the cases of high ionic strength could be analyzed. For this reason, a Monte Carlo computational code has been developed in order to implement the model, which is described in the next subsection.

Monte Carlo Simulations at Constant pH Value. The proposed SBRIS model is analyzed by means of simulations in the semi-grand canonical Monte Carlo (SGCMC); that is, the pH value is the control variable, and it is kept constant along

the computation. The SGCMC code is a modification of the one previously developed by our group to compute conformational and ionization properties of linear polyelectrolytes.^{52,57}

In particular, it has been extended in order to include the effect of mechanical work. As a result, bending and stretching of the bonds have also been implemented. The resulting program is rather general since it allows working with sites and bonds of different pK values, interaction energies, conformational energies, and so on. Excluded volume effects can also be taken into account. Moreover, the code can deal with any arbitrary distribution of the sites along the chain, which is chosen by the user. However, in this paper, we restrict its use to the assumptions detailed previously. A snapshot of one of the SGCMC realizations is shown in Figure 1.

The Metropolis algorithm^{68,69} generates new states at constant pH in a chain with $N = 50$ ionizable sites (i.e., 148 nodes or $M = 147$ bonds), a number which is large enough to avoid end effects and ensures the reproducibility of the intensive properties of the polymer, such as bond state probabilities or degree of protonation. An outline of the algorithm is depicted in Figure 2. In each new MC configuration, the polyelectrolyte can change either (A) the rotational state of a bond, (B) the length or bending angle of a bond, (C) the ionization state of a binding site, or (D) the spatial orientation of the polyelectrolyte chain in the laboratory reference frame, with trial probabilities of 0.88, 0.1, 0.01, and 0.01, respectively. These values allow us to obtain a good equilibration of the conformational structure for a given ionization state so that the system does not get trapped in local minima. Each change in the rotational state of a bond j implies a $\pm 120^\circ$ rotation of its dihedral angle and the recalculation of distances among the sites situated before and after the rotating bond. The changes in the stretching and bending states of bond j are $\Delta l_j = \pm 0.01 \text{ \AA}$ and $\Delta \alpha_j = \pm 0.5^\circ$, respectively. These variations provide an average acceptance ratio of $\sim 20\%$, which is an acceptable value to make proper statistics. The global spatial orientation of the polymeric chain is altered by

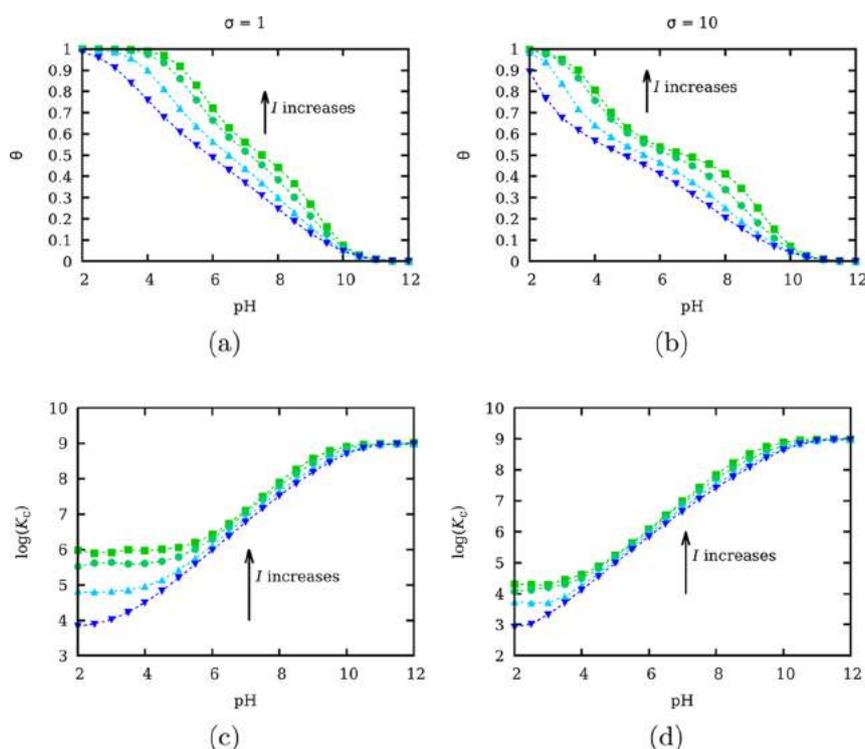


Figure 3. (a, b) Titration curves and (c, d) effective pK value for the polyelectrolyte depicted in Figure 1 in the absence of the pulling force. The images on the left side correspond to $\sigma = 1$, and those on the right side refer to $\sigma = 10$. The chosen ionic strengths are 1 M (green squares), 0.1 M (turquoise circles), 0.01 M (cyan upward triangles), and 0.001 M (blue downward triangles). The rest of the parameters have the same values as those in Figure 1b.

changing the polar angle θ and the azimuth angle ϕ of the first and second bonds of the polyelectrolyte with respect to the laboratory coordinate frame by amounts $\Delta\theta = \pm 2^\circ$ and $\Delta\cos(\phi) = \pm 0.1$, respectively. The latter change is important to avoid preferred orientations in the space at zero force. Once the free energy difference ($\Delta\mathcal{F}$ of eq 1) between trial and current conformations is calculated, the new configuration is always accepted if $\Delta\mathcal{F} < 0$ and accepted with a probability $\exp(-\beta\Delta\mathcal{F})$ if $\Delta\mathcal{F} > 0$.

The results presented represent the average over eight different SGCMC simulations. Each simulation has been equilibrated in the first 5×10^7 configurations, and the thermal averages have been computed in the following 4.5×10^8 realizations. The simulations were performed using a parallel code developed in C++ on a 126 CPU cluster. For each pH, ionic strength, and force, typical jobs were run using eight CPUs for 1 to 2 h. The chosen parameters are $pK = 9$ and $u_t = 10$, similar to those corresponding to LPEI. Note, however, that the reduced free energy only depends on the reduced chemical potential $\mu = \text{pH} - pK$. This means that the results and conclusions taken from the simulations are the same for any pK value by choosing a suitable pH value for which the difference $\text{pH} - pK$ is the same. The simulations are performed at room temperature $T = 298.15$ K. Free protons, co-ions, and counterions are not explicit in the simulations, and the screening effects are taken into account via the Debye length parameter, κ^{-1} , in eq 11. The chosen values for the parameters in the stretching and bending potentials are $l_0 = 1.5$ Å, $\alpha_0 = 120^\circ$, $k_{\text{length}} = 300$ kcal mol $^{-1}$ Å $^{-2}$, and $k_{\text{angle}} = 0.01$ kcal mol $^{-1}$ deg $^{-2}$, which are typical values used in molecular dynamics force fields for C–C bonds.⁷⁰

The average degree of protonation (θ) is computed as

$$\theta = \frac{\langle N_+ \rangle}{N} \quad (12)$$

where $\langle N_+ \rangle$ is the thermal average number of protonated sites. Note that since the simulations are performed at constant pH, N_+ is a fluctuating quantity, different in each new accepted configuration. Another interesting quantity is the effective protonation constant (K_c), which provides information about the average affinity of the macromolecular sites for the protons.^{62,71,72} In general, K_c depends on the charge of the macromolecule, different at each pH value. It is defined as

$$\log K_c = \text{pH} + \log\left(\frac{\theta}{1 - \theta}\right) \quad (13)$$

The probability of a rotating c bond to be in the *gauche* state, $P(g)$, is calculated as

$$P(g) = \frac{\langle N_g \rangle}{N - 1} \quad (14)$$

where $\langle N_g \rangle$ is the thermal average number of rotating bonds in a *gauche* state. Other quantities, such as the probability of having two neighboring c bonds in given conformations, (e.g., tt , tg^+ , g^+g^+ , etc.) can be calculated in the same way. The extension of the polyelectrolyte chain (L_z) in the direction of the mechanical force, that is, the z axis, is obtained as

$$L_z = \langle z_{M+1} - z_1 \rangle \quad (15)$$

where z_i is the z coordinate of site i in the laboratory coordinate frame. Finally, a very useful quantity to understand

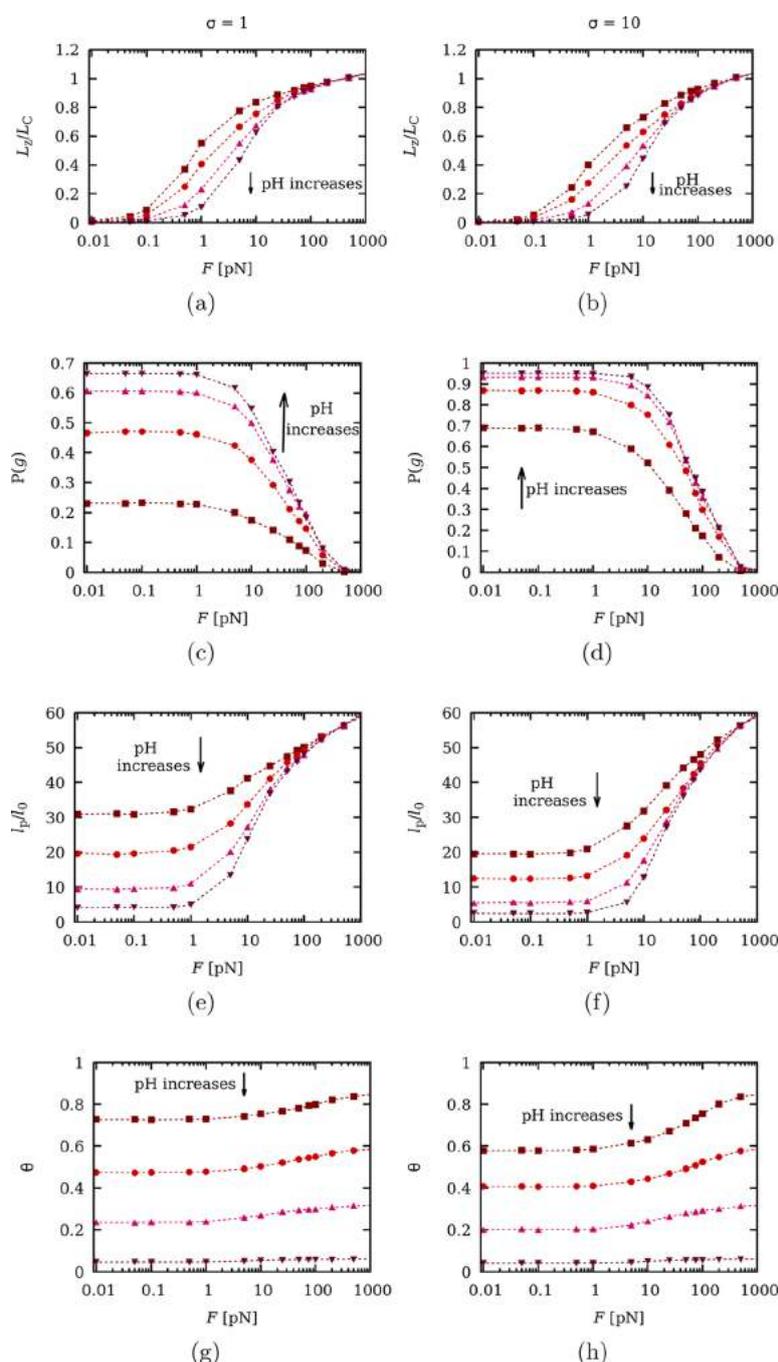


Figure 4. (a, b) Chain elongation L_z normalized to the contour length L_C , (c, d) *gauche* state probability, (e, f) persistence length, and (g, h) degree of protonation versus the applied force F at a constant ionic strength of 0.001 M and pH values of 4 (squares), 6 (circles), 8 (upward triangles), and 10 (downward triangles). The images on the left side correspond to $\sigma = 1$, and those on the right side refer to $\sigma = 10$. The rest of the parameters have the same values as those in Figure 1b.

the mechanism of macromolecular stretching is the persistence length l_p , defined as the average sum of the projections of all the bonds $j \geq i$ on bond i in an indefinitely long chain

$$l_p/l_0 = \sum_{j \geq i} \langle \mathbf{b}_i \cdot \mathbf{b}_j \rangle \quad (16)$$

where \mathbf{b}_i are unitary vectors pointing to the direction of the bonds. It is straightforward to show that, for a long enough

chain ($M \rightarrow \infty$), l_p is related to the average square end-to-end distance $\langle \mathbf{r}^2 \rangle$ by the relationship³²

$$l_p = \frac{\langle \mathbf{r}^2 \rangle}{2Ml_0} + \frac{l_0}{2} = \frac{\langle (\mathbf{r}_{M+1} - \mathbf{r}_1)^2 \rangle}{2Ml_0} + \frac{l_0}{2} \quad (17)$$

where r_i is the position of site i . As a consequence of eq 17, the Kuhn length

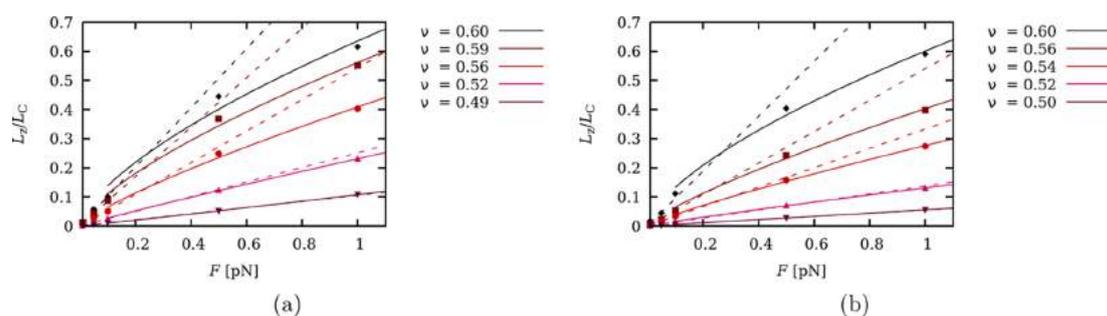


Figure 5. Normalized chain extension versus force in the low-force regime for (a) $\sigma = 1$ and (b) $\sigma = 10$ at constant ionic strength $I = 0.001$ M but with pH values (from top to bottom) of 2, 4, 6, 8, and 10. The simulation results (markers) are compared to the linear eq 20 (dashed lines) and to the Pincus scaling law, eq 21 (continuous lines). The rest of the parameters have the same values as those in Figure 1b.

$$l_K = \frac{\langle \mathbf{r}^2 \rangle}{Ml_0} \quad (18)$$

can be directly related to the persistence length by $l_K = 2l_p - l_0$.

RESULTS AND DISCUSSION

In this section, we will discuss the effect of the pH value and the ionic strength on the force–extension curves by simultaneously analyzing the dependence of conformational (chain elongation, bond state probabilities, and persistence length) and protonation properties (degree of protonation and effective protonation constant). As commented above, the microscopic pK value of the protonating sites will be fixed at $pK = 9$ throughout this section without loss of generality. Concerning the energy of the *gauche* state in the absence of charge, two cases physically relevant are considered. In the first case, the polyelectrolyte exhibits free rotation (i.e., *gauche* and *trans* states have the same energy, $\epsilon_\sigma = 0$, $\sigma = 1$). In the second case, the *gauche* states of the c bonds are favored, for instance, because of the existence of hydrogen bonding between two consecutive protonating sites, which means that $\sigma > 1$ (we take $\epsilon_\sigma = 1$, $\sigma = 10$). This phenomenon has been observed in LPEI and POE.⁵⁷ The case $\sigma < 1$ is not much interesting since most of the bonds are in the *trans* state so the chain is basically extended even in the absence of force. Finally, the interaction energy between two charged neighboring sites through a *trans* c bond is fixed at $\epsilon_{u,t} = 1$ ($u_t = 0.1$) for all the studied cases, which is the order of magnitude found in a number of weak polyelectrolytes by using potentiometry. These works indicate that, for the same molecule, neighboring interactions are rather independent of the ionic strength.^{56,57,89}

Since charge regulation is a key ingredient of the model, let us first analyze the behavior of the degree of protonation θ when no mechanical force is applied, which will be useful in the foregoing discussion. In Figure 3, the titration curves for the cases $\sigma = 1$ (a) and $\sigma = 10$ (b) are shown at four different ionic strengths (I): 1, 0.1, 0.01, and 0.001 M, from top to bottom. It can be observed that in both cases, lowering the ionic strength results in a decrease of the degree of protonation for all of the pH values, which is explained by the increase in the LR electrostatic repulsion. Note that this effect is larger in the case $\sigma = 10$, which can be explained by the fact that *gauche* states are energetically favored, which hinders the possibility of having two neighboring sites charged (since $u_g = 0$). The effective protonation constant K_c is depicted in Figure 3c,d for $\sigma = 1$ and $\sigma = 10$, respectively. Clearly, K_c presents two asymptotic values. At high pH, the macromolecule is not

charged, electrostatic interactions are absent, and the K_c value corresponds to the microscopic pK value of the ionizable sites ($pK = 9$). However, as pH decreases, sites get ionized and the work needed to protonate an empty site increases due to electrostatic repulsion. This results in a decrease in K_c , which, at low enough pH values, reaches a new asymptotic value. This decrease in the affinity for the proton is especially relevant at low ionic strengths for which the LR interactions are stronger since screening is weaker. Finally, note that, for the same pH and I values, the decrease in pK_c in the case $\sigma = 10$ is more pronounced than in the case $\sigma = 1$. For $\sigma = 10$, the *gauche* states are energetically promoted, the chain is more folded, and the distance between charge sites is shorter, which leads to larger LR interactions.

Effect of the pH Value on the Force–Extension Curves. The force–extension curves are shown in Figure 4a,b for the cases $\sigma = 1$ and $\sigma = 10$, respectively, for pH values ranging from 4 to 10 (top to bottom). The chain extension is normalized to the contour length

$$L_C = Nl_0 \cos((\pi - \alpha_0)/2) \quad (19)$$

defined as the length of the fully extended chain (i.e., all the bonds are in the *trans* state) with bond lengths and angles in their equilibrium position. The chosen value for the ionic strength is 0.001 M, a small value for which the LR electrostatic interactions are maximized. In order to better understand the extension curves, the *gauche* state probability (Figure 4c,d) and persistence length (Figure 4e,f) are also represented. On the other hand, the degree of protonation is depicted in Figure 4g,h. The images in the left side always correspond to the case $\sigma = 1$, while those in the right side refer to the case $\sigma = 10$. It is important to stress that stretching, conformational, and ionization properties are highly coupled, so we will discuss them all at once. Owing to the lack of space, in the main document of this work, we just present and discuss the force–extension curves, which are more relevant for the purpose of this study. However, for the lecturers interested, the full casuistry, covering the complete range of pH values and ionic strengths, is reported in the Supporting Information.

Let us first analyze the different force regimes in terms of the progressive activation of the degrees of freedom of the bonds, that is, rotational, bending, and stretching. First of all, note that, for low enough forces and for all the pH values, both the *gauche* state probability $P(g)$ and the normalized persistence length l_p/l_0 remain constant. Despite this fact, the chain is considerably extended, around 10% for pH = 10 and a remarkable 50% for pH = 4. This indicates the existence of a

low-force regime (corresponding to $F < F_E = k_B T / l_p \approx 1$ pN)²⁰ under which the chain behaves as a structureless set of segments with characteristic length l_p .

The dependence of $P(g)$ and l_p on the pH value is shown in Figure 4. As observed in Figure 4c,d, $P(g)$ is strongly affected by the pH value at low forces. This fact can be better explained by observing the behavior of the average degree of protonation θ (Figure 4g,h) for $F < F_E$. For $\sigma = 1$, θ increases from $\theta \approx 0.05$ at pH = 10 to $\theta \approx 0.7$ at pH = 4, while for $\sigma = 10$, θ increases from $\theta \approx 0.05$ at pH = 10 to $\theta \approx 0.6$ at pH = 4. In both cases, when the adjacent sites are simultaneously protonated, the electrostatic repulsion is so strong that the *gauche* states are forbidden ($u_g = 0$). As a result, the rotational properties change with the pH value, resulting in two limit behaviors. At high pH, when the polymer is discharged, the bonds present free rotation (when $\sigma = 1$) or preference for the *gauche* state (when $\sigma = 10$). On the contrary, at low enough pH values, when the macromolecule is almost completely charged, the restriction $u_g = 0$ forces all the bonds to adopt the *trans* conformation. The increase of the number of bonds in the *trans* state due to lowering the pH value can be clearly observed in Figure 4c,d. As a result, the polymer chain gets stiffer, and the persistence length increases, as can be observed in Figure 4e,f. In turn, this fact leads to the increase in the chain elongation observed in Figure 4a,b, which is more marked for $\sigma = 1$ than for $\sigma = 10$. In the latter case, the *gauche* state is energetically favored so that a larger charge (i.e., a lower pH) is required to obtain the same number of *trans* bonds and to increase the persistence length.

The elongation versus force curves in the low force regime are shown in Figure 5 where the simulation values are depicted as markers. As can be observed, the low-force regime can be divided into two subregimes. For very low forces ($F < 0.1$ pN), the chain behaves as an entropic spring, and the elongation responds linearly to the applied force (dashed lines in the figure). The relation between elongation and force is given by

$$L_z / Ml_0 = \beta \frac{l_K F}{3} = \beta \frac{2l_p - l_0}{3} F \quad (20)$$

This expression is independent of the structure of the chain since it comes directly from the fluctuation–dissipation theorem.⁷³ Under this subregime, the mechanical work is much smaller than the thermal energy. However, for larger forces ($0.1 < F < 1$ pN), the situation becomes more complex. It is found that the extension follows the Pincus scaling law^{73,74} (continuous lines)

$$L_z \approx F^{1/\nu-1} \quad (21)$$

Note that ν is found to range from $\nu = 1/2$ at high pH (uncharged chain and corresponding to the linear regime) to $\nu = 3/5$ at low pH (at which the chain is almost fully charged). The latter value was first predicted by Pincus⁷⁴ for strong polyelectrolytes, and it can be explained as the effect of electrostatic excluded volume interactions and the corresponding swelling of the macromolecule. Interestingly, both limiting values are obtained with great accuracy from the simulations, which nicely confirms Pincus theory. For pH values ranging from 6 to 8, intermediate values of ν are obtained, indicating that a partially charged weak polyelectrolyte can be seen as an intermediate situation between the neutral chain and a strong polyelectrolyte. For the case $\sigma = 10$, the transition between the two limiting cases is more gradual than for $\sigma = 1$ due to the fact

that lower pH values are necessary to fully charge the chain (see Figure 3).

So far, we have analyzed the low-force regime for which the persistence length remains constant with the applied force. For larger forces, however, the rotational degrees of freedom are activated. This fact makes $P(g)$ decrease with the force. In this new situation, the bonds, which were initially in the *gauche* state, gradually adopt the *trans* state by effect of the pulling force. The stretching mechanism is no longer entropic, but it depends on the free energy of the *trans/gauche* transition, and the persistence length becomes force-dependent. This fact is in contrast with DNA, with a much more rigid structure and for which charge regulation can be neglected.³⁵ The new characteristic force F_R for which the rotational degrees of freedom are activated can be roughly estimated by equating the mechanical work per monomer to the free energy of the bond state transition $\Delta F_{t \rightarrow g}$ so that

$$F_R l_0 \approx \Delta F_{t \rightarrow g} \approx k_B T \ln(2\sigma) \quad (22)$$

The F_R resulting values are 30 and 70 pN for $\sigma = 1$ and $\sigma = 10$, respectively, in agreement with the observed range of forces (1–100 pN) for which the variation of $P(g)$ is more pronounced. Moreover, the conformational degrees of freedom are coupled with proton binding due to the term (eq 9) in the reduced free energy. It is observed that, in increasing the force value, the change in the rotational states from *gauche* to *trans* is simultaneous with the increase in the macromolecular charge. Two asymptotic behaviors are found again. At low forces, the protonation state is the same as the one of the non-stretched molecule. As commented above, θ is lower for $\sigma = 10$ than for $\sigma = 1$. At large enough forces, however, a new plateau arises, and θ is the same value for both σ values. The gap between the θ value at low and high forces is thus larger for $\sigma = 10$ and depends on the pH value. For instance, for $\sigma = 10$ and pH = 4, θ ranges from less than 0.58 in the linear regime to 0.83 for the larger forces. At large pH values, however, the change in θ is smaller in absolute terms although it can be significant in relative terms. For instance, for $\sigma = 10$ and pH = 8, θ ranges from 0.2 to 0.32, which means an increase of more than 50%. Using the definition (eq 13), the effective protonation constant K_c can be calculated as a function of the force. The obtained effective pK value, which is not shown here but can be found in the Supporting Information, slightly increases with F for both σ values. However, this effect is very weak since the pK value at large forces exceeds the low force limit by, at most, 0.5 pK units. However, this variation seems to be enough to cause significant changes in the macromolecular charge in applying an external force. This point will be discussed in more detail in the next subsection for the full range of ionic strengths.

Finally, when the force is large enough to deform bond angles and lengths, a third situation arises. The average bond length $\langle l \rangle$ (green squares) and bond angle $\langle \alpha \rangle$ (black dots), normalized by their respective equilibrium values, are shown in Figure 6 as functions of the applied force. It can be observed that the bond length and angle only start to be significantly elongated for values larger than the characteristic force $F_S > 300$ pN. A rough estimation of F_S using the Hooke law also confirms this value. Note that the bond angle is slightly easier to deform than the bond length. Although it is not shown in the figure, bond stretching and bending have been found to be independent of the pH value and ionic strength. This fact explains why the force–extension curves in Figure 4a,b also become almost independent of the pH for forces larger than

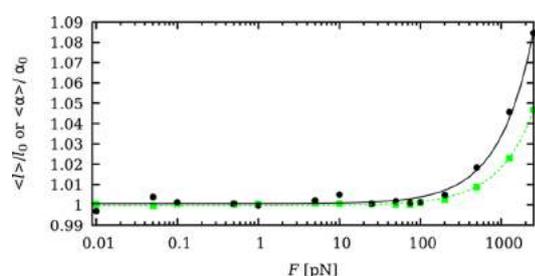


Figure 6. Average bond length $\langle l \rangle$ (green squares) and bond angle $\langle \alpha \rangle$ (black dots) versus applied force F , normalized to their equilibrium values, l_0 and α_0 . Lines are only to guide the eyes. The simulations have been performed at $\text{pH} = 6$, $I = 0.01$ M, and $\sigma = 1$. The rest of the parameters have the same values as those in Figure 1b.

F_S . At this point, most of the bonds are in the *trans* state. Since the force constants are independent of the ionization state, the response to the applied force is also the same for any degree of protonation. Finally, for forces around and larger than F_S , the

bending potential becomes anharmonic, finally leading to bond breaking, as reported in several AFM single-molecule stretching experiments.¹

We conclude that, for intermediate forces and suitable pH and ionic strength values, CR can be induced by an applied force when the mechanism of CR is proton binding. For other types of binding mechanisms, such as metal binding, the ionic charge and binding constants are much larger, and the binding mechanism strongly depends on the conformational state (for instance, because of the presence of chelate complexes). As a result, CR could be significantly enhanced. In those cases, which are out of the scope of the present work, the interplay between stretching and CR could be of technological interest.

Effect of the Ionic Strength. Let us investigate the effect of the ionic strength in the force–extension curves, which are represented in Figure 7a,b for ionic strengths (from bottom to top) of 1, 0.1, 0.01, and 0.001 M. Now, the pH value is fixed to $\text{pH} = 6$. For this pH value, the macromolecule is approximately half charged, so it is a suitable value to discuss the influence of

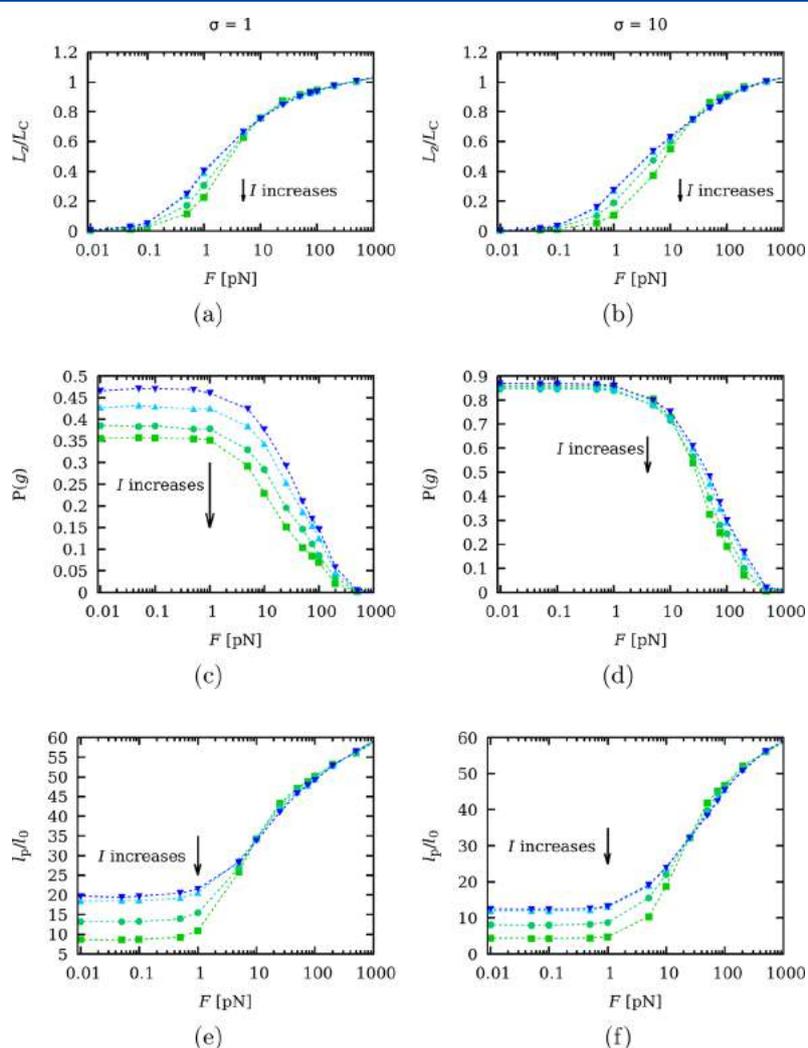


Figure 7. (a, b) Chain elongation L_z normalized to the contour length L_C , (c, d) *gauche* state probability, and (e, f) persistence length versus applied force F at a constant pH value of $\text{pH} = 6$ and ionic strengths of 1 M (green squares), 0.1 M (turquoise circles), 0.01 M (cyan upward triangles), and 0.001 M (blue downward triangles). The images on the left side correspond to $\sigma = 1$, and those on the right side refer to $\sigma = 10$. The rest of the parameters have the same values as those in Figure 1b.

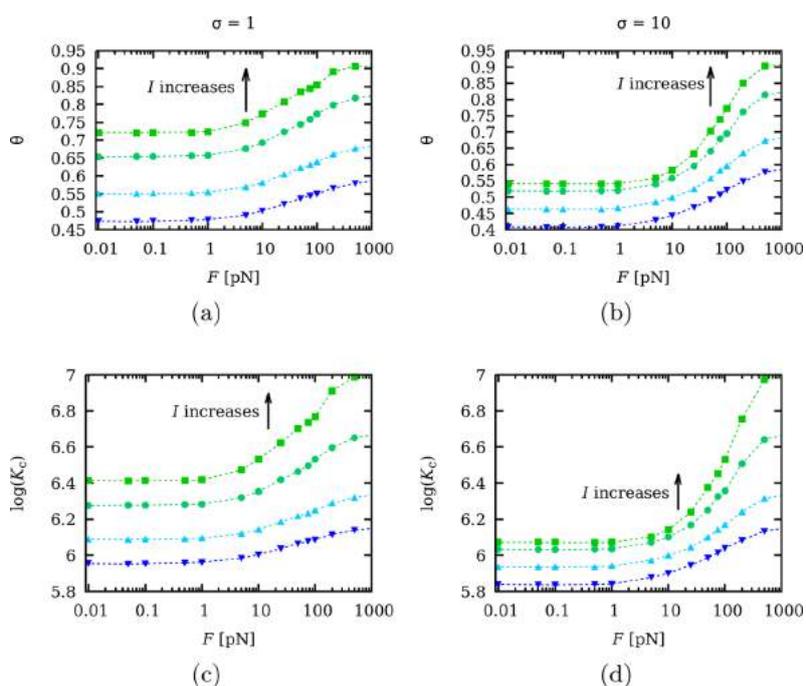


Figure 8. (a, b) Degree of protonation and (c, d) effective pK value versus applied force F at a constant pH value of pH = 6 and ionic strengths of 1 M (green squares), 0.1 M (turquoise circles), 0.01 M (cyan upward triangles), and 0.001 M (blue downward triangles). The images on the left side correspond to $\sigma = 1$, and those on the right side refer to $\sigma = 10$. The rest of the parameters have the same values as those in Figure 1b.

charge regulation. The full results for the rest of the pH values (4, 8, and 10) are delivered in the Supporting Information. The *gauche* state probability (Figure 7c,d), the persistence length (Figure 7e,f), degree of protonation (Figure 8a,b), and effective pK value (Figure 8c,d) are also presented. Again, two cases are considered, $\sigma = 1$ and $\sigma = 10$, which correspond to the images on the left side and on the right side, respectively.

First, note that the effect of the ionic strength, for the interval of accessible experimental values, is overall weaker than the effect of the pH value. For instance, at $F = 1$ pN and $\sigma = 1$, the normalized chain extension varies from 0.18 at $I = 1$ M to 0.35 at $I = 0.001$ M (a difference of 0.17 units), and in the case $\sigma = 10$, the extension ranges from 0.8 at $I = 1$ M to 0.24 at $I = 0.001$ M (a difference of 0.16 units). The three force regimes are again observed for all the ionic strengths: the entropic regime; the intermediate regime, for which the rotational and ionization degrees of freedom are activated; and finally the large-force regime, corresponding to deformations in the bond angle and length. However, unlike the effect of the pH value, the effect of ionic strength on the conformational properties is more complicated due to protonation and the complex interplay between SR and LR interactions.

Let us first analyze the dependence of the binding properties on the applied force, depicted in Figure 8. In all the cases, θ increases with F for forces larger than the low-force regime $F > F_E \approx 1$ pN. As a general trend, the polyelectrolyte chain is on average more elongated as F increases so that the mean distance between sites increases and the LR electrostatic repulsion decreases, allowing more sites to be protonated. Concerning the SR interactions, note that $P(g)$ experiences an important decrease in the interval $F = 10$ – 100 pN (see Figure 7c,d), which is especially dramatic in the case $\sigma = 10$. This fact implies a drastic change in the chemical environment of the ionizable sites, which become separated by *trans* bonds,

through which the repulsion is much weaker. Charge regulation is clearly induced by the mechanical force. Interestingly, the larger the ionic strength, the more intense charge regulation is. Especially remarkable is the case $I = 1$ M and $\sigma = 10$ for which the charge is almost doubled at high forces. This indicates that the charging process is basically a local phenomenon, which is essentially driven by SR interactions and the conformational state of the c bonds and is rather independent of the ionic strength. Conversely, LR interactions, which increase on lowering the ionic strength, weaken charge regulation because they discharge the molecule in all the force regimes. In the same way, the effective pK value also increases with the stretching process for $F > F_E \approx 1$ pN, as can be observed in Figure 8c,d, so that a larger affinity for the protons is induced by the applied force. Again, this effect is especially relevant for $\sigma = 10$ and at high ionic strengths.

Concerning the dependence of the *gauche* probabilities on the ionic strength, note that, as observed in Figure 7c,d, $P(g)$ seems to be in contradiction with the behavior of the persistence length. On increasing the ionic strength, $P(g)$ decreases and so does the number of *gauche* bonds, and apparently, the chain should be stiffer. However, for forces below $F = F_R \approx 20$ pN, the persistence length also decreases, so actually the chain gets more folded. This effect can be observed for the two σ values although it is especially relevant for $\sigma = 10$. This apparent paradox can be explained by taking a look at the degree of protonation. As commented above, θ increases with the ionic strength. As a result, the probability of having two charged neighboring sites increases. Since they cannot be both protonated and linked by a bond in the *gauche* state at the same time ($u_g = 0$), the number of *gauche* states decreases. Note that this effect is due to action of the SR interactions. However, unlike the *gauche* probability, the persistence length is a “global” property, the result of the

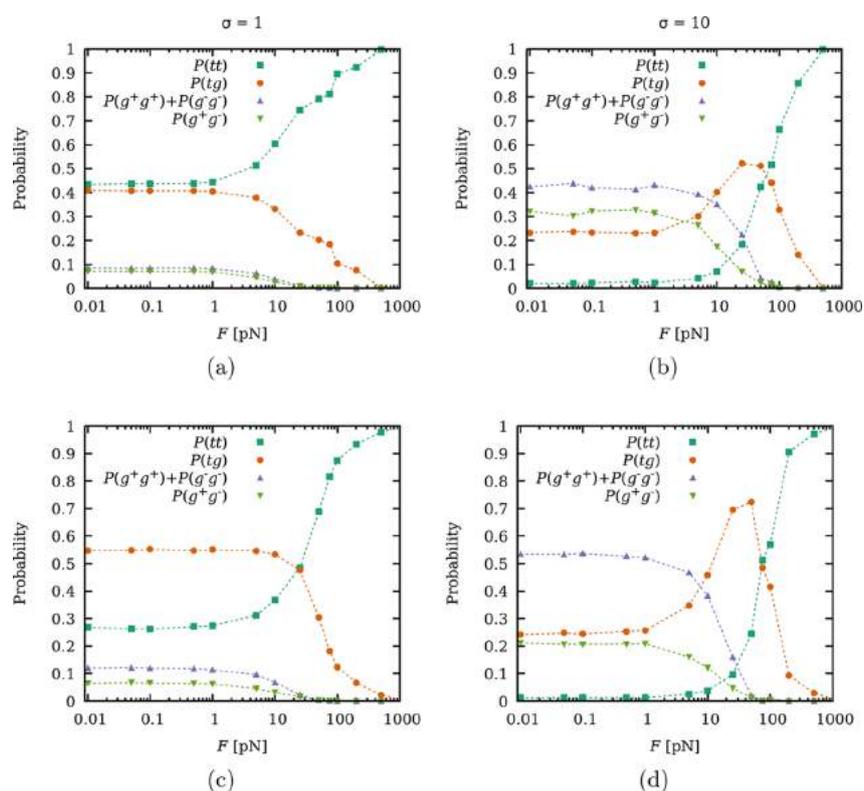


Figure 9. Probability of having two neighbor bonds in given conformations $P(c_i c_{i+1})$ versus F at two different ionic strengths of (a, b) $I = 1$ M and (c, d) $I = 0.001$ M at constant $\text{pH} = 6$. The combinations shown are both bonds in *trans* $P(tt)$ (blue squares), one bond in *trans* and the neighbor in *gauche* $P(tg)$ (orange circles), both bonds in *gauche* with the same orientation $P(g^+g^+) + P(g^-g^-)$ (purple upward triangles), and both bonds in *gauche* with the opposite orientation $P(g^+g^-)$ (green downward triangles). The images on the left side (a and c) correspond to $\sigma = 1$, and those on the right side (b and d) refer to $\sigma = 10$. The rest of the parameters have the same values as those in Figure 1b.

simultaneous action of many bonds and sites. As a consequence, LR interactions play a more important role in the behavior of l_p . The same argument can be used for the chain elongation, which decreases with the ionic strength in all the curves below for $F < 20$ pN.

For $F > 20$ pN, the elongation becomes almost independent of I , but on examining in detail the curves, we can observe an unexpected result: the extension slightly decreases with I . This intriguing trend of the elongation is consistent with the behavior of the persistence length: for $F < F_R$ the chain gets stiffer on lowering the ionic strength, but, for larger forces, it gets slightly more flexible. For lower pH values, this effect is more visible (as in the curves for $\text{pH} = 4$ shown in the Supporting Information). This certainly small effect seems to be apparently irrelevant but again points out the non-trivial connection between SR and LR interactions. We suspect that this result is related to spatial correlations of the states of neighboring bonds although probably a deeper insight is necessary.

As an example of the formation of spatial patterns, the probabilities of having two consecutive c bonds in a given conformation $P(c_i c_{i+1})$ as a function of F have been calculated at two different ionic strengths $I = 1$ M (Figure 9a,b) and $I = 0.001$ M (Figure 9c,d). Again, the images on the left side and on the right side correspond to the cases $\sigma = 1$ and $\sigma = 10$, respectively. Due to the polyelectrolyte symmetry, there are only four different combinations of bond states: both bonds in *trans* (with a probability $P(tt)$), one bond in *trans* and its neighbor in *gauche* ($P(tg)$), both bonds in *gauche* with the

same orientation ($P(g^+g^+) + P(g^-g^-)$), and both bonds with different orientations ($P(g^+g^-)$). As expected, $P(tt)$ monotonically increases as the mechanical force increases and, for large enough forces, $P(tt)$ tends to 1 independently of the value of σ . For $\sigma = 1$ and $I = 0.001$ M, the preferred combination is *trans-gauche* at low forces, but there is a crossover, which can be observed at $F \approx 50$ pN, curiously at the same force interval at which the change of tendency in l_p and in the elongation is observed. For $\sigma = 10$, the preferred combination at low forces is g^+g^+ . However, the most intriguing point is the observed maximum in $P(tg)$ at $F \approx 50$ pN. This maximum implies the existence of an intermediate situation where the mechanical work contribution, which promotes the *trans* state, competes with the energetic stabilization of the *gauche* states due to the fact that $\sigma > 1$. In this force regime, the polyelectrolyte adopts a highly ordered structure, which alternates bonds in the *trans* state with bonds in the *gauche* state. Again, the presence of this maximum coincides with the force interval for which l_p and the elongation switch their dependence on the ionic strength. We would like to highlight that, even for the very simple model of polyelectrolyte presented here, one finds a rather rich physical-chemical behavior, which includes charge regulation, complex conformational transitions, and highly correlated spatial structures.

CONCLUSIONS

In this work, the influence of charge regulation, highly coupled with the conformational degrees of freedom, in the stretching

properties of weak polyelectrolytes is studied. With this aim, we propose a model, which captures the fundamental aspects present in a flexible weak linear polyelectrolyte (internal angle rotation, bond stretching, bond bending, and proton binding) with a minimum number of parameters. The model was inspired by the structure of linear poly(ethylenimine) (LPEI), a symmetric weak polyelectrolyte with an ionizable site every three chain positions. It is based on the site binding rotational isomeric state (SBRIS) model, which allows studying conformational and ionization properties on the same foot. Short-range (SR) and long-range (LR) electrostatic interactions are treated separately. LR interactions are chemically unspecific and can be reasonably implemented using the mean-field Debye–Hückel potential. Conversely, SR interactions between neighboring sites and bonds are mediated by the macromolecular skeleton so that they depend on the detailed chemical environment of the site. As a result, specific energetic parameters are used to describe them. Bond stretching and bending are included by means of harmonic potentials. The resulting scheme is used to perform semi-grand canonical Monte Carlo (SGCMC) simulations at constant pH and applied force. Concerning the energy of the *gauche* state, two situations are studied, controlled by the energy of the *gauche* state (corresponding to the Boltzmann factor σ): when *trans* and *gauche* states have the same energy ($\sigma = 1$) and when the *gauche* state is energetically stabilized, for instance, by hydrogen bonding (we take $\sigma = 10$). The influence of the pH value and the ionic strength in the force–extension curves is analyzed. In order to understand the different mechanisms of chain stretching, the degree of protonation θ , bond state probabilities of the *gauche* state $P(g)$, and persistence length l_p as functions of the force are also analyzed.

As a general trend, three force regimes are found. In the low-force regime, the persistence length is force-independent, and two subregimes are identified. Up to 0.1 pN linear behavior is found, as demanded by the fluctuation–dissipation theorem, for all the pH values. From 0.1 to 1 pN, however, the chain presents Pincus scaling behavior depending on the pH value. For high pH values (i.e., neutral chain), the elongation is still linear with Pincus scaling exponent $\nu = 1/2$, while for low pH values (fully charged chain), $\nu = 3/5$, the theoretically predicted value for strong polyelectrolytes. For intermediate pH values, ν has been found to present a gradual transition between the two limiting values. In the large-force regime, most of the bonds are in the *trans* state so that the stretching becomes approximately independent of the pH and the ionic strength. Finally, there is an intermediate regime, between 1 and 100 pN, for which the rotational and protonation degrees of freedom, which are highly coupled, are activated. This force regime is the most interesting one since conformational transitions, charge regulation, and spatial correlations are observed. It is in this force regime that the pH value and the ionic strength maximally influence the chain elongation. Mechanically induced charge regulation is mainly driven by SR interactions. When the macromolecule is elongated, the *trans* states are promoted so that the electrostatic interaction between neighboring sites decreases, favoring the affinity for the protons. This effect seems to be larger at large ionic strengths and pH values for which the molecule is partially charged.

The role of the pH value is relatively straightforward to understand. On lowering the pH value, the macromolecule gets charged, promoting *trans* states and larger distances

between sites, thus reducing the electrostatic repulsion. This results in an increase in the persistence length, a reduction in the number of *gauche* states, and an easier extension of the chain. Therefore, a significant influence of the pH value on the curve–extension curves is found both for $\sigma = 1$ and $\sigma = 10$.

On the other hand, the effect of the ionic strength for a fixed pH value is more complicated since it depends on the complex interplay between SR phenomena (bond conformations and protonation) and the LR interactions. It is found that the exhibited tendency of $P(g)$ seems to be in contradiction with the behavior of the persistence length. On increasing the ionic strength, $P(g)$ decreases and so does the number of *gauche* bonds, and apparently, the chain should be stiffer. However, for forces below $F = F_R \approx 20$ pN, the persistence length also decreases, so the chain gets more folded. This apparent paradox can be solved by observing that the charge decreases on increasing the ionic strength while the intensity of LR interactions is enhanced. Ionization equilibria therefore play a fundamental role in the stretching properties of weak polyelectrolytes. Finally, it is found that in the intermediate-force regime, spatial correlations are formed, which also determine some subtle aspects of the stretching process.

We would like to highlight that, even for the very simple model of polyelectrolyte presented here, one finds a rather rich physical–chemical behavior, which includes charge regulation, complex conformational transitions, and highly correlated spatial structures. To our knowledge, this work is the first attempt to study, at least by means of computational simulation, the mechanical stretching of a weak polyelectrolyte including the coupling of charge regulation and conformational equilibria.

■ ASSOCIATED CONTENT

📄 Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.macromol.9b01160](https://doi.org/10.1021/acs.macromol.9b01160).

Full casuistry of chain extension, *gauche* state probability, persistence length, degree of protonation, and binding equilibrium constant as a function of the mechanical force F , covering the complete range of pH values and ionic strengths (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: pmbianco@ub.edu (P.M.B.).

*E-mail: s.madurga@ub.edu (S.M.).

ORCID

Sergio Madurga: [0000-0002-8135-7057](https://orcid.org/0000-0002-8135-7057)

Francesc Mas: [0000-0002-1362-4002](https://orcid.org/0000-0002-1362-4002)

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We acknowledge the financial support from Generalitat de Catalunya (grants 2017SGR1033, 2017SGR1329, and XrQTC), and Spanish Structures of Excellence Maria de Maeztu program through (grant MDM-2017-0767). J.L.G. also acknowledges the Spanish Ministry of Science and Innovation (project CTM2016-78798-C2-1-P). F.M. and S.M. acknowledge the funding of the EU project 8SEWP-

HORIZON 2020 (692146). P.M.B. also acknowledges the financial support from a grant (FI-2017) of Generalitat de Catalunya. We thank Professor Michal Borkovec for introducing us to the topic, helpful discussions, and suggestions.

REFERENCES

- (1) Giannotti, M. I.; Vancso, G. J. Interrogation of single synthetic polymer chains and polysaccharides by AFM-based force spectroscopy. *ChemPhysChem* **2007**, *8*, 2290–2307.
- (2) Camunas-Soler, J.; Ribezzi-Crivellari, M.; Ritort, F. Elastic Properties of Nucleic Acids by Single-Molecule Force Spectroscopy. *Annu. Rev. Biophys.* **2016**, *45*, 65–84.
- (3) Radiom, M.; Kong, P.; Maroni, P.; Schäfer, M.; Kilbinger, A. F. M.; Borkovec, M. Mechanically induced cis-to-trans isomerization of carbon-carbon double bonds using atomic force microscopy. *Phys. Chem. Chem. Phys.* **2016**, *18*, 31202–31210.
- (4) Valiaev, A.; Lim, D. W.; Oas, T. G.; Chilkoti, A.; Zauscher, S. Force-induced prolyl cis-trans isomerization in elastin-like polypeptides. *J. Am. Chem. Soc.* **2007**, *129*, 6491–6497.
- (5) Rognoni, L.; Most, T.; Žoldák, G.; Rief, M. Force-dependent isomerization kinetics of a highly conserved proline switch modulates the mechanosensing region of filamin. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 5568–5573.
- (6) Marszalek, P. E.; Oberhauser, A. F.; Pang, Y.-P.; Fernandez, J. M. Polysaccharide elasticity governed by chair-boat transitions of the glucopyranose ring. *Nature* **1998**, *396*, 661–664.
- (7) Giannotti, M. I.; Rinaudo, M.; Vancso, G. J. Force spectroscopy of hyaluronan by atomic force microscopy: From hydrogen-bonded networks toward single-chain behavior. *Biomacromolecules* **2007**, *8*, 2648–2652.
- (8) Klukovich, H. M.; Kouznetsova, T. B.; Kean, Z. S.; Lenhardt, J. M.; Craig, S. L. A backbone lever-arm effect enhances polymer mechanochemistry. *Nat. Chem.* **2013**, *5*, 110–114.
- (9) Wang, J.; Kouznetsova, T. B.; Niu, Z.; Ong, M. T.; Klukovich, H. M.; Rheingold, A. L.; Martinez, T. J.; Craig, S. L. Inducing and quantifying forbidden reactivity with single-molecule polymer mechanochemistry. *Nat. Chem.* **2015**, *7*, 323–327.
- (10) Wiita, A. P.; Perez-Jimenez, R.; Walther, K. A.; Gräter, F.; Berne, B. J.; Holmgren, A.; Sanchez-Ruiz, J. M.; Fernandez, J. M. Probing the chemistry of thioredoxin catalysis with force. *Nature* **2007**, *450*, 124–127.
- (11) Friedsam, C.; Gaub, H. E.; Netz, R. R. Probing surfaces with single-polymer atomic force microscope experiments. *Biointerphases* **2006**, *1*, MR1–MR21.
- (12) Krysiak, S.; Liese, S.; Netz, R. R.; Hugel, T. Peptide desorption kinetics from single molecule force spectroscopy studies. *J. Am. Chem. Soc.* **2014**, *136*, 688–697.
- (13) Liu, Y.; Liu, K.; Wang, Z.; Zhang, X. Host-enhanced $\pi - \pi$ Interaction for water-soluble supramolecular polymerization. *Chem. – Eur. J.* **2011**, *17*, 9930–9935.
- (14) Schütze, D.; Holz, K.; Müller, J.; Beyer, M. K.; Lüning, U.; Hartke, B. Pinpointing mechanochemical bond rupture by embedding the mechanophore into a macrocycle. *Angew. Chem., Int. Ed.* **2015**, *54*, 2556–2559.
- (15) Bosco, A.; Camunas-Soler, J.; Ritort, F. Elastic properties and secondary structure formation of single-stranded DNA at monovalent and divalent salt conditions. *Nucleic Acids Res.* **2014**, *42*, 2064–2074.
- (16) Strick, T. R.; Dessinges, M.-N.; Charvin, G.; Dekker, N. H.; Allemand, J.-F.; Bensimon, D.; Croquette, V. Stretching of macromolecules and proteins. *Rep. Prog. Phys.* **2003**, *66*, 1–45.
- (17) Marko, J. F.; Siggia, E. D. Stretching DNA. *Macromolecules* **1995**, *28*, 8759–8770.
- (18) Rief, M.; Gautel, M.; Oesterhelt, F.; Fernandez, J. M.; Gaub, H. E. Reversible Unfolding of Individual Titin Immunoglobulin Domains by AFM. *1997*, *276*, 1109–1112, DOI: [DOI: 10.1126/science.276.5315.1109](https://doi.org/10.1126/science.276.5315.1109).
- (19) Bustamante, C.; Marko, J. F.; Siggia, E. D.; Smith, S. Entropic elasticity of lambda-phage DNA. *Science* **1994**, *1599*.
- (20) Livadaru, L.; Netz, R. R.; Kreuzer, H. J. Stretching response of discrete semiflexible polymers. *Macromolecules* **2003**, *36*, 3732–3744.
- (21) Kierfeld, J.; Niamploy, O.; Sa-Yakanit, V.; Lipowsky, R. Stretching of semiflexible polymers with elastic bonds. *Eur. Phys. J. E* **2004**, *14*, 17–34.
- (22) Radiom, M.; Borkovec, M. Influence of ligand-receptor interactions on force-extension behavior within the freely jointed chain model. *Phys. Rev. E* **2017**, *96*, No. 062501.
- (23) Radiom, M.; Maroni, P.; Wesolowski, T. A. Size extensivity of elastic properties of alkane fragments. *J. Mol. Model.* **2018**, *24*, 36.
- (24) Hugel, T.; Rief, M.; Seitz, M.; Gaub, H. E.; Netz, R. R. Highly stretched single polymers: Atomic-force-microscope experiments versus ab-initio theory. *Phys. Rev. Lett.* **2005**, *94*, No. 048301.
- (25) Chandler, D. *Introduction to Modern Statistical Mechanics*; Oxford University Press, 1987.
- (26) Koper, G. J. M.; Borkovec, M. Proton binding by linear, branched, and hyperbranched polyelectrolytes. *Polymer* **2010**, *51*, 5649–5662.
- (27) Kreuzer, H. J.; Grunze, M. Stretching of single polymer strands: A first-principles theory. *Europhys. Lett* **2001**, *55*, 640–646.
- (28) Hanke, F.; Serr, A.; Kreuzer, H. J.; Netz, R. R. Stretching single polypeptides: The effect of rotational constraints in the backbone. *Europhys. Lett.* **2010**, *92*, 53001.
- (29) Neuert, G.; Hugel, T.; Netz, R. R.; Gaub, H. E. Elasticity of poly(azobenzene-peptides). *Macromolecules* **2006**, *39*, 789–797.
- (30) Oesterhelt, F.; Rief, M.; Gaub, H. E. Single molecule force spectroscopy by AFM indicates helical structure of poly(ethylene-glycol) in water. *New J. Phys.* **1999**, *1*, 6.
- (31) Liese, S.; Gensler, M.; Krysiak, S.; Schwarzl, R.; Achazi, A.; Paulus, B.; Hugel, T.; Rabe, J. P.; Netz, R. R. Hydration Effects Turn a Highly Stretched Polymer from an Entropic into an Energetic Spring. *ACS Nano* **2017**, *11*, 702–712.
- (32) Flory, P. *Statistical Mechanics of Chain Molecules*; John Wiley & Sons, Inc., 1967.
- (33) Flory, P. J. Foundations of Rotational Isomeric State Theory and General Methods for Generating Configurational Averages. *Macromolecules* **1974**, *7*, 381–392.
- (34) Jacobson, D. R.; McIntosh, D. B.; Stevens, M. J.; Rubinstein, M.; Saleh, O. A. Single-stranded nucleic acid elasticity arises from internal electrostatic tension. *PNAS* **2017**, *114*, 5095–5100.
- (35) Netz, R. R. Strongly stretched semiflexible extensible polyelectrolytes and DNA. *Macromolecules* **2001**, *34*, 7522–7529.
- (36) Dessinges, M.-N.; Maier, B.; Zhang, Y.; Peliti, M.; Bensimon, D.; Croquette, V. Stretching Single Stranded DNA, a Model Polyelectrolyte. *Phys. Rev. Lett.* **2002**, *89*, 248102.
- (37) Seol, Y.; Skinner, G. M.; Visscher, K. Elastic properties of a single-stranded charged homopolymeric ribonucleotide. *Phys. Rev. Lett.* **2004**, *93*, 118102.
- (38) Saleh, O. A.; McIntosh, D. B.; Pincus, P.; Ribbeck, N. Nonlinear low-force elasticity of single-stranded DNA molecules. *Phys. Rev. Lett.* **2009**, *102*, No. 068301.
- (39) McIntosh, D.; Saleh, O. A. Electrostatic effects of multivalent salts on ssDNA elasticity. *Biophys. J.* **2011**, *100*, 483a.
- (40) Ullner, M. Comments on the Scaling Behavior of Flexible Polyelectrolytes within the Debye - Hückel Approximation. *J. Phys. Chem. B* **2003**, *107*, 8097–8110.
- (41) Jacobson, D. R.; McIntosh, D. B.; Saleh, O. A. The snakelike chain character of unstructured RNA. *Biophys. J.* **2013**, *105*, 2569–2576.
- (42) Stevens, M. J.; McIntosh, D. B.; Saleh, O. A. Simulations of stretching a strong, flexible polyelectrolyte: Using Long chains to access the pincus scaling regime. *Macromolecules* **2013**, *46*, 6369–6373.
- (43) Stevens, M. J.; Berezney, J. P.; Saleh, O. A. The effect of chain stiffness and salt on the elastic response of a polyelectrolyte. *J. Chem. Phys.* **2018**, *149*, 163328.

- (44) Stevens, M. J.; McIntosh, D. B.; Saleh, O. A. Simulations of stretching a strong, flexible polyelectrolyte. *Macromolecules* **2012**, *45*, 5757–5765.
- (45) Muthukumar, M. 50th Anniversary Perspective: A Perspective on Polyelectrolyte Solutions. *Macromolecules* **2017**, *50*, 9528–9560.
- (46) Boroudjerdi, H.; Kim, Y.-W.; Naji, A.; Netz, R. R.; Schlagberger, X.; Serr, A. Statics and dynamics of strongly charged soft matter. *Phys. Rep.* **2005**, *416*, 129–199.
- (47) Carnal, F.; Clavier, A.; Stoll, S. Polypeptide-nanoparticle interactions and corona formation investigated by monte carlo simulations. *Polymer* **2016**, *8*, 203.
- (48) Trefalt, G.; Behrens, S. H.; Borkovec, M. Charge Regulation in the Electrical Double Layer: Ion Adsorption and Surface Interactions. *Langmuir* **2016**, *32*, 380–400.
- (49) Lipfert, J.; Doniach, S.; Das, R.; Herschlag, D. Understanding Nucleic Acid–Ion Interactions. *Annu. Rev. Biochem.* **2014**, *83*, 813–841.
- (50) Lund, M.; Jönsson, B. Charge regulation in biomolecular solution. *Q. Rev. Biophys.* **2013**, *46*, 265–281.
- (51) Narambuena, C. F.; Longo, G. S.; Szleifer, I. Lysozyme adsorption in pH-responsive hydrogel thin-films: the non-trivial role of acid-base equilibrium. *Soft Matter* **2015**, *11*, 6669–6679.
- (52) Blanco, P. M.; Madurga, S.; Mas, F.; Garcés, J. L. Coupling of charge regulation and conformational equilibria in linear weak polyelectrolytes: Treatment of long-range interactions via effective short-ranged and pH-dependent interaction parameters. *Polymer* **2018**, *10*, 811.
- (53) Olander, D. S.; Holtzer, A. The Stability of the Polyglutamic Acid alpha Helix. *J. Am. Chem. Soc.* **1968**, *90*, 4549–4560.
- (54) Uyaver, S.; Seidel, C. First-order conformational transition of annealed polyelectrolytes in a poor solvent. *Europhys. Lett.* **2003**, *64*, 536–542.
- (55) Berezney, J. P.; Saleh, O. A. Electrostatic Effects on the Conformation and Elasticity of Hyaluronic Acid, a Moderately Flexible Polyelectrolyte. *Macromolecules* **2017**, *50*, 1085–1089.
- (56) Garcés, J. L.; Koper, G. J. M.; Borkovec, M. Ionization Equilibria and Conformational Transitions in Polyprotic Molecules and Polyelectrolytes. *J. Phys. Chem. B* **2006**, *110*, 10937–10950.
- (57) Garcés, J. L.; Madurga, S.; Borkovec, M. Coupling of conformational and ionization equilibria in linear poly(ethylenimine): A study based on the site binding/rotational isomeric state (SBRIS) model. *Phys. Chem. Chem. Phys.* **2014**, *16*, 4626–4638.
- (58) Garcés, J. L.; Madurga, S.; Rey-Castro, C.; Mas, F. Dealing with long-range interactions in the determination of polyelectrolyte ionization properties. Extension of the transfer matrix formalism to the full range of ionic strengths. *J. Polym. Sci., Part B: Polym. Phys.* **2017**, *55*, 275–284.
- (59) Baptista, A. M.; Teixeira, V. H.; Soares, C. M. Constant-pH molecular dynamics using stochastic titration. *J. Chem. Phys.* **2002**, *117*, 4184–4200.
- (60) Mongan, J.; Case, D. A.; McCammon, J. A. Constant pH molecular dynamics in generalized Born implicit solvent. *J. Comput. Chem.* **2004**, *25*, 2038–2048.
- (61) Narambuena, C. F.; Beltramo, D. M.; Leiva, E. P. M.; Narambuena, C. F.; Beltramo, D. M.; Leiva, E. P. M. Polyelectrolyte Adsorption on a Charged Surface. Free Energy Calculation from Monte Carlo Simulations Using Jarzynski Equality. 2008, 8267–8274, DOI: [10.1021/ma800325e](https://doi.org/10.1021/ma800325e).
- (62) Madurga, S.; Garcés, J. L.; Companys, E.; Rey-Castro, C.; Salvador, J.; Galceran, J.; Vilaseca, E.; Puy, J.; Mas, F. Ion binding to polyelectrolytes: Monte Carlo simulations versus classical mean field theories. *Theor. Chem. Acc.* **2009**, *123*, 127–135.
- (63) Madurga, S.; Rey-Castro, C.; Pastor, I.; Vilaseca, E.; David, C.; Garcés, J. L.; Puy, J.; Mas, F. A semi-grand canonical Monte Carlo simulation model for ion binding to ionizable surfaces: Proton binding of carboxylated latex particles as a case study. *J. Chem. Phys.* **2011**, *135*, 184103.
- (64) Torres, P.; Bojanich, L.; Sanchez-Varretti, F.; Ramirez-Pastor, A. J.; Quiroga, E.; Boeris, V.; Narambuena, C. F. Protonation of β -lactoglobulin in the presence of strong polyelectrolyte chains: a study using Monte Carlo simulation. *Colloids Surf., B* **2017**, *160*, 161–168.
- (65) Stormes, M.; Linse, P.; Dias, R. S. Monte Carlo Simulations of Complexation between Weak Polyelectrolytes and a Charged Nanoparticle. Influence of Polyelectrolyte Chain Length and Concentration. *Macromolecules* **2017**, *50*, 5978–5988.
- (66) Narambuena, C. F. On the reasons for α -lactalbumin adsorption on a charged surface: a study by Monte Carlo simulation. *Colloids Surf., B* **2019**, *174*, 511–520.
- (67) Boudon, S.; Wipff, G. Conformational analysis of protonated ethylenediamine in the gas phase and in water. *J. Mol. Struct.: THEOCHEM* **1991**, *228*, 61–70.
- (68) Ullner, M.; Jönsson, B. A Monte Carlo study of titrating polyelectrolytes in the presence of salt. *Macromolecules* **1996**, *29*, 6645–6655.
- (69) Borkovec, M.; Jönsson, B.; Koper, G. J. M. J. In *Surface and Colloid Science*; Matigerc, E., Ed.; Springer: New York, USA, 2001; Vol. 10; Chapter 2, pp 99–339.
- (70) Sliozberg, Y. R.; Kröger, M.; Chantawansri, T. L. Fast equilibration protocol for million atom systems of highly entangled linear polyethylene chains. *J. Chem. Phys.* **2016**, *144*, 154901.
- (71) Garcés, J. L.; Mas, F.; Puy, J.; Galceran, J.; Salvador, J. Use of activity coefficients for bound and free sites to describe metal–macromolecule complexation. *J. Chem. Soc., Faraday Trans.* **1998**, *94*, 2783–2794.
- (72) Garcés, J. L.; Mas, F.; Cecilia, J.; Companys, E.; Galceran, J.; Salvador, J.; Puy, J. Complexation isotherms in metal speciation studies at trace concentration levels. Voltammetric techniques in environmental samples. *Phys. Chem. Chem. Phys.* **2002**, *4*, 3764–3773.
- (73) Saleh, O. A. Perspective: Single polymer mechanics across the force regimes. *J. Chem. Phys.* **2015**, *142*, 194902.
- (74) Pincus, P. Excluded Volume Effects and Stretched Polymer Chains. *Macromolecules* **1976**, *9*, 386–388.

Chapter 5

Role of Charge Regulation and Fluctuations in the Conformational and Mechanical Properties of Weak Flexible Polyelectrolytes

5.1 Summary

5.1.1 Introduction

In the previous chapters, special focus have been paid to the coupling in the conformational and ionization degrees of weak polyelectrolytes. Unlike strong polyelectrolytes, weak polyelectrolyte can change its charge in response of conformational changes and *vice versa*. In the previous chapter, it has been proven that, as a consequence of the ionization-conformation interplay, it is possible to induce a Charge Regulation (CR) by mechanically stretching the weak polyelectrolyte chain.

CR is fundamental to understand the physico-chemistry of many different charged polymeric systems.^{35,143,148,191–193} A direct implication of CR is that the charge is not a fixed quantity but it fluctuates. The fluctuations in the charge can lead to counter-intuitive phenomena. For instance, attraction between equally-charged macromolecules has been observed due to Charge Fluctuations (CF).^{28,196} In many theoretical works aiming to study the conformational properties of weak polyelectrolytes, the charge is assumed to be constant for instance considering a constant charge density.^{197,198} In other studies, CF is explicitly considered but the bonds are rigid or described as an harmonic potential and the rotational degrees of freedom are not well described.^{33,40,81,152} The influence of CF in the conformational and elastic properties of a flexible weak polyelectrolytes is still never been assessed.

In this chapter, the effect of CR and CF in the conformational and elastic properties of a model weak polyelectrolyte are analyzed. The effect of CF can be directly studied with computer simulations, where CF can be easily switched on/off.¹⁹⁰ In the work presented here, the exact average degree of ionization θ is computed at different pH and ionic strength (I) conditions using Semi-Grand Canonical Monte Carlo (SGCMC) simulations. Then, we assign to each binding site a constant charge equal to θ and perform constant charge Monte Carlo (ccMC) simulations, as outlined in Fig. 5.1. The conformational (persistence length and gauche state probability) and elastic (Pincus exponent and chain extension) obtained with both simulations, SGCMC and ccMC, are compared. In addition, we quantify the effect of excluded volume in the conformational and elastic properties.

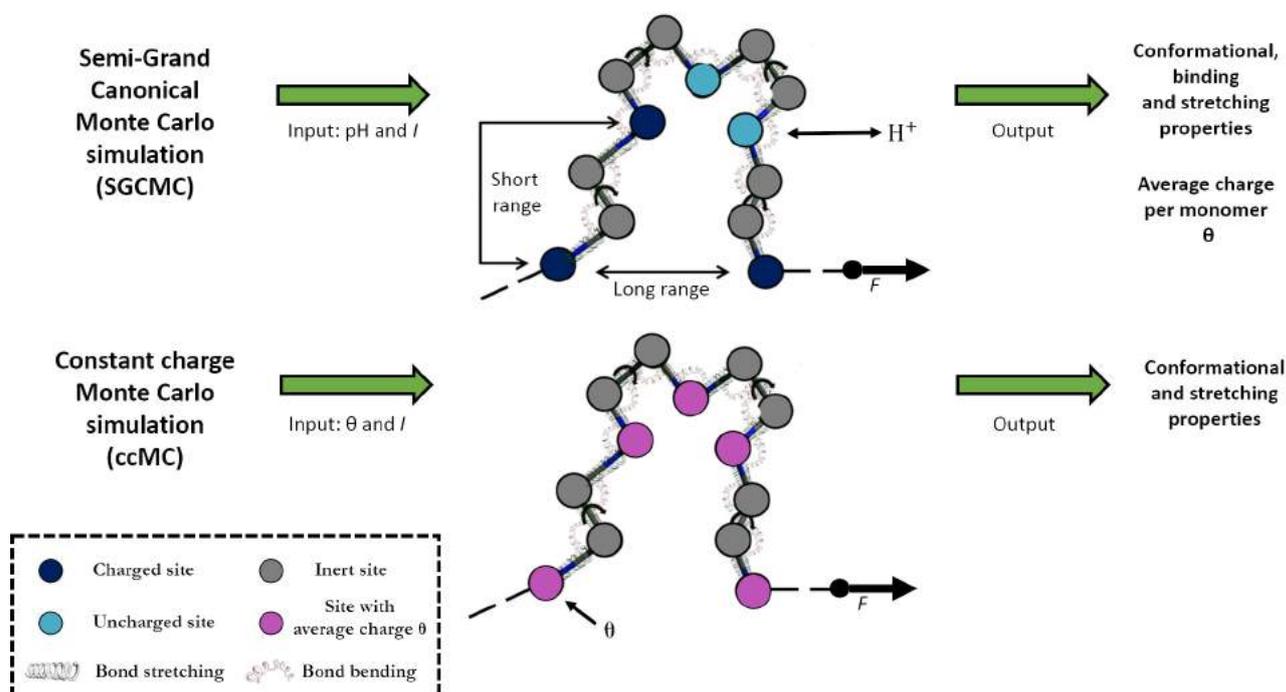


FIGURE 5.1: Scheme of the different simulations performed in the publication presented in this chapter. (up) Semi-Grand Canonical Monte Carlo (SGCMC) simulations, where the charge of the binding sites (blue circles) fluctuates and the average degree of ionization, θ , can be exactly calculated. (down) constant charge Monte Carlo (ccMC), where the charge of the sites is assumed to be constant and equal to θ (purple circles). The model polyelectrolyte considered has a binding site every three chain positions. The rest of the sites (grey circles) are neutral and non-ionizable.

5.1.2 Results

The weak polyelectrolyte is modeled using the SBRIS model and it has a ionizable site every three chain positions. Concerning the structural properties of the chain, the model incorporates the assumptions already presented in Sections 3.1.2 and 4.1.2. However, two new assumptions are involved in some of the simulations performed in this work:

- When constant charge Monte Carlo (ccMC) simulations are performed, the charge of all the binding sites is considered equal to the average degree of ionization θ . θ is obtained from Semi-Grand Canonical Monte Carlo (SGCMC) simulations at same pH and ionic strength (I) conditions. In keeping θ constant, the fluctuations in the charge are neglected. As a consequence, the differences in the average charge of sites located in different positions on the chain are neglected (for instance, the central sites of the chain are in average less charged than the ones in the ends).

- When excluded volume is included in the model, all the sites are modeled as hard sphere with radii $R = 1.7$ and $R = 1.55$, for inert and ionizable sites, respectively.

The results presented in this section correspond to the case with $pK_a = 9$, $\sigma = 1$, $\epsilon_{u,t} = 1$, $\epsilon_{u,g} = \infty$, equilibrium bond length $l_0 = 1.5 \text{ \AA}$ and equilibrium bond angle $\alpha_0 = 120^\circ$.

The intensity of the charge fluctuation can be quantified by the binding capacitance, C , whose dependence on the pH-value is shown in Fig. 5.2 at I -values ranging from 1 M to 0.001 M. It can be observed that at high ionic strengths, C has two maxima at pH values around 6 and 9 units. Due to the strong first neighbour electrostatic interaction, the system behaves as it had a second effective pK_a value (roughly equal to $pK_a - 2\epsilon_{u,t}$) at low pH-values, when it is highly charged. The minima observed at $pH \sim 7.5$ is due to the formation of highly ordered structure with alternated charged and uncharged sites.¹⁶³ For low values of I , the long range electrostatic interactions are strengthened, which destroy the binding correlation between neighboring sites. As a consequence, C curves are flattened at low I -values.

The role of the CF in the conformational properties of a weak polyelectrolyte is investigated by analyzing the probability of the *gauche* state $P(g)$ as a function of θ , which is depicted in Fig. 5.3a at I -values ranging from 1 M to 0.001 M (from top to bottom). Results of different kinds of simulations are shown: SGCMC without excluded volume (filled markers), SGCMC with excluded volume (stars) and ccMC (empty markers). It can be observed that excluded volume is only relevant at small θ values, since at large θ values the electrostatic repulsion is enough to produce chain self-avoiding. At small θ -values, excluded volume slightly diminishes $P(g)$ because some folded *gauche* conformations are forbidden. Comparing SGCMC and ccMC results, it can be observed important deviations at intermediate θ -values. ccMC, where CF is absent, underestimate $P(g)$ and thus have more extended chains. In other words, CF allow the

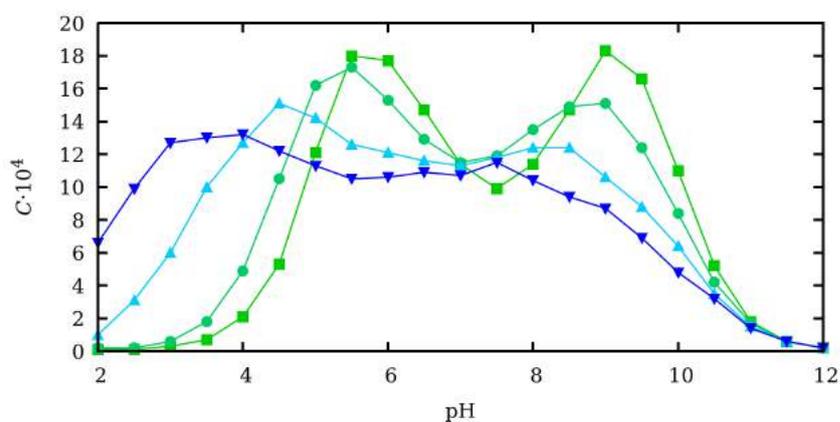


FIGURE 5.2: Binding capacitance C titration curve obtained from SGCMC simulations at ionic strengths of 0.001 M (blue downwards triangles), 0.01 M (cyan upwards triangles), 0.1 M (turquoise circles) and 1 M (green squares)

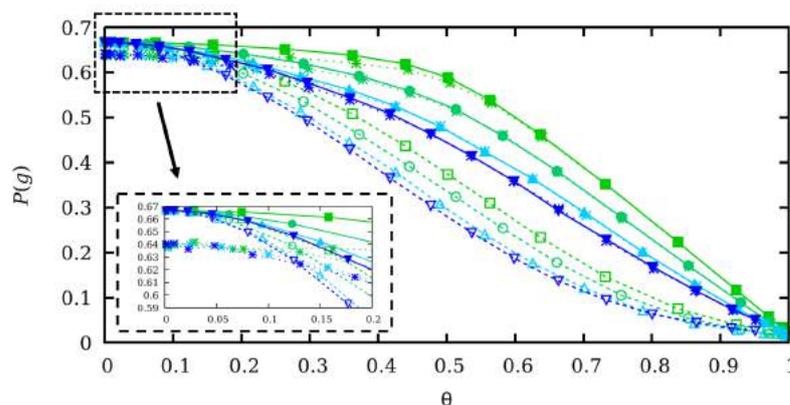


FIGURE 5.3: Probability of a bond of being in *gauche* conformational state $P(g)$ as a function of the average degree of ionization θ at I -values ranging from 1 M to 0.001 M following the same markers as in Fig. 5.2. The results of different kinds of simulations are presented: SGCMC without excluded volume (filled markers), SGCMC with excluded volume (stars) and ccMC (empty markers).

polyelectrolyte to exhibit a more folded, compact form. In the publication presented in this chapter, the same trend is observed for the persistence length, which is larger when CF are absent (ccMC simulations). This is in good agreement with previous theoretical and experimental investigations on the role of CF in the intermolecular interaction of weak polyelectrolyte, which predicted an attractive potential due to CF.^{28,190,196} In the same way, CF produce, in an effective way, an attractive contribution to the interaction energy between charged sites in the same chain.

In the previous chapter, the low force elastic regime of weak polyelectrolytes is analyzed using SGCMC simulations. The simulation results are well-fitted to the Pincus scaling law (Eq. 1.51).⁶⁴ The Pincus scaling exponent ν is found to have a transition from $\nu = 3/5$ at low pH-values to $\nu = 1/2$ at high pH-values. Both limiting values correspond to the theoretical predictions for a strong polyelectrolyte and a neutral ideal polymer, respectively.⁶⁴ In this chapter, it is investigated if the ν intermediate values at intermediate pH-values are a consequence of CR and CF or directly caused by the change in the charge of the polyelectrolyte. In Fig. 5.4, the ν -values obtained from SGCMC without excluded volume (filled markers), SGCMC with excluded volume (stars) and ccMC (empty markers) are shown. Two different ionic strengths are presented $I = 0.001$ M (blue circles) and $I = 1$ M (green squares). It can be observed that the transition of ν is independent of the presence of CR and CF. In general, CR and CF does not have a significant impact on ν except for θ around 0.8 at $I = 1$ M. Excluded volume does not show any significant effect on the ν value. Interestingly, the transition between the two limiting values of ν is found to be linear with θ . The dashed line in Fig. 5.4 is the best fit of the simulation results at $I = 0.001$ M to a linear equation of the form $\nu = m\theta + n$, with two fitting parameters m and n .

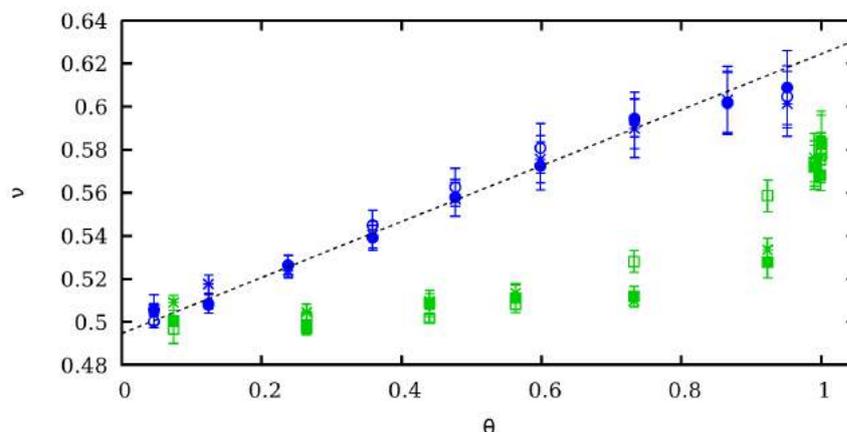


FIGURE 5.4: Pincus scaling exponent ν as function of theta. Three different kinds of simulations are shown: SGCMC without excluded volume (filled markers), SGCMC with excluded volume (stars) and ccMC (empty markers). Two different cases are shown with $I = 1$ M (green squares) and $I = 0.001$ M (blue circles). The dashed line represents the best fit of the simulations to a linear equation of the form $\nu = m\theta + n$.

In the previous chapter, the increase of θ with the stretching force was studied. A similar study can be observed in Fig. 5.5a at constant $I = 1$ M and pH values ranging from 4 to 10 from

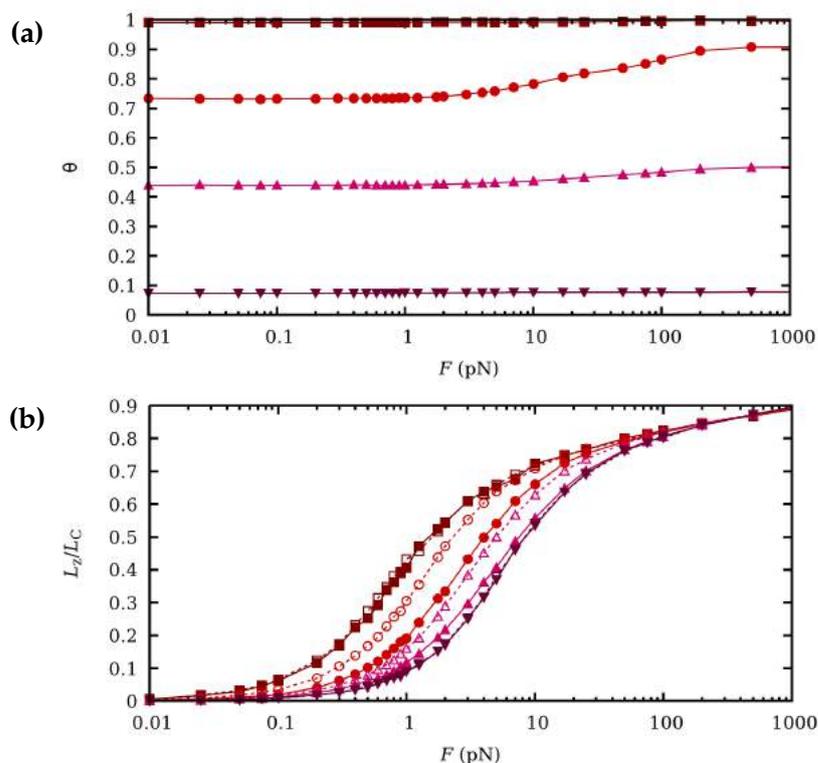


FIGURE 5.5: (a) θ and (b) chain extension L_z as a function of the stretching force F at constant $I = 1$ M and pH-values: 10 (purple downwards triangles), 8 (pink upwards triangles), 6 (red circles) and 4 (wine squares). Filled markers correspond to SGCMC simulations while empty markers refer to ccMC simulations.

top to bottom. Again, the stretching force is only observed to affect θ in the intermediate force regime. Note, that the impact in θ is larger at pH values of 6 and 8. These values are also where the binding capacitance and, therefore, the CF are larger. In other words, the ideal conditions to induce charge regulation by mechanical stretching are those where CF is more important. The relevance of CF in the force/extension curve in the full force range can be assessed in Fig. 5.5b, where the results of SGCMC (filled markers) and ccMC (empty markers) are shown. Clear deviations between both sets of data can be observed, indicating that CF play a relevant role in weak polyelectrolyte stretching. Again, these deviations are stronger when charge fluctuation is more important, *i.e.* at pH and I values of large binding capacitance. This effect can be quantified by comparing these results with the ones obtained at $I = 0.001$ M, where CF is weaker. One can observe that for $I = 0.001$ M the deviations between SGCMC and ccMC are smaller. Although not shown here, excluded volume was not found to have any significant contribution to the elastic response of the polyelectrolyte. The same behavior is observed in the l_p vs. F curves, which can be found in the publication at the end of the chapter.

5.1.3 Conclusions

In this chapter, the effect of charge regulation and charge fluctuations in the conformational and elastic properties of a flexible weak polyelectrolyte is investigated. First, Semi-Grand Canonical Monte Carlo (SGCMC) simulations are performed where the charge fluctuates and the average degree of ionization θ is exactly computed. Then, constant charge Monte Carlo (ccMC) simulations are performed where the charge of the binding sites is assumed to be constant and equal to θ . The conformational (persistence length l_p and *gauche* state probabilities $P(g)$) and elastic (chain extension L_z and Pincus exponent ν) properties obtained with both, SGCMC and ccMC

When charge fluctuations are switched off, larger values of l_p and smaller values of $P(g)$ are found. Therefore, when charge fluctuation is absent the polymer is more extended, which indicates that charge fluctuation contributes to polymer folding. This effect is found to be larger in conditions where charge fluctuations is more relevant, *i.e.* at conditions of large binding capacitance.

The effect of charge fluctuation on the elastic response of the polyelectrolyte has been investigated in the full force regime. In the low force regime, it is found that ν presents a transition with the pH from 3/5 (low pH values) to 1/2 (high pH values). This transition is found to happen independently of the presence of charge fluctuation, which suggests that is mainly caused by the change in the average charge of the polyelectrolyte. In the intermediate force regime, charge fluctuation have been found to have a significant effect on the L_z and l_p (which becomes force-dependent) stretching curves. Again, this effect is found to be larger at pH and ionic strength

conditions of large binding capacitance values, where the fluctuations in the charge are more intense. In general, the steric excluded volume is found to not be significant compared when the electrostatic repulsion for all the cases of studies.

5.2 Publication

Article

Role of Charge Regulation and Fluctuations in the Conformational and Mechanical Properties of Weak Flexible Polyelectrolytes

Pablo M. Blanco ^{1,*}, Sergio Madurga ¹, Claudio F. Narambuena ², Francesc Mas ^{1,*} and Josep L. Garcés ³

- ¹ Physical Chemistry Unit, Materials Science and Physical Chemistry Department & Research Institute of Theoretical and Computational Chemistry (IQTCUB) of Barcelona University (UB), 08028 Barcelona, Catalonia, Spain; s.madurga@ub.edu
 - ² Facultad Regional San Rafael, Universidad Tecnológica Nacional & Instituto de Física Aplicada (INFAP), Universidad Nacional de San Luis-CONICET, 5600 San Rafael, Argentina; claudionarambuena@hotmail.com
 - ³ Chemistry Department, Technical School of Agricultural Engineering & AGROTECNIO of Lleida University (UdL), 25003 Lleida, Catalonia, Spain; garcesjl22@yahoo.es
- * Correspondence: pmbianco@ub.edu (P.M.B.); fmas@ub.edu (F.M.)

Received: 25 October 2019; Accepted: 25 November 2019; Published: 29 November 2019



Abstract: This work addresses the role of charge regulation (CR) and the associated fluctuations in the conformational and mechanical properties of weak polyelectrolytes. Due to CR, changes in the pH-value modifies the average macromolecular charge and conformational equilibria. A second effect is that, for a given average charge per site, fluctuations can alter the intensity of the interactions by means of correlation between binding sites. We investigate both effects by means of Monte Carlo simulations at constant pH-value, so that the charge is a fluctuating quantity. Once the average charge *per* site is available, we turn off the fluctuations by assigning the same average charge to every site. A constant charge MC simulation is then performed. We make use of a model which accounts for the main fundamental aspects of a linear flexible polyelectrolyte that is, proton binding, angle internal rotation, bond stretching and bending. Steric excluded volume and differentiated treatment for short-range and long-range interactions are also included. This model can be regarded as a kind of “minimal” in the sense that it contains a minimum number of parameters but still preserving the atomistic detail. It is shown that, if fluctuations are activated, *gauche* state bond probabilities increase and the persistence length decreases, so that the polymer becomes more folded. Macromolecular stretching is also analyzed in presence of CR (the charge depends on the applied force) and without CR (the charge is fixed to the value at zero force). The analysis of the low force scaling behavior concludes that Pincus exponent becomes pH-dependent. Both, with and without CR, a transition from 1/2 at high pH-values (phantom chain) to 3/5 at low pH-values (Pincus regime) is observed. Finally, the intermediate force stretching regime is investigated. It is found that CR induces a moderate influence in the force-extension curves and persistence length (which in this force regime becomes force-dependent). It is thus concluded that the effect of CR on the stretching curves is mainly due to the changes in the average charge at zero force. It is also found that, for the cases studied, the effect of steric excluded volume is almost irrelevant compared to electrostatic interactions.

Keywords: polyelectrolytes; charge regulation; charge fluctuations; weak polyelectrolyte; annealed polyelectrolyte; Monte Carlo simulation; semi-grand canonical ensemble; binding equilibria; conformational equilibria; constant pH ensemble; stretching; scaling law

1. Introduction

Charge regulation (CR) is defined as the capability of ionizable macromolecules, nano-particles and surfaces, to modify their ionization state as a response to external physico-chemical perturbations. In contrast to strong polyelectrolytes, such as DNA or RNA, whose phosphate groups are charged in very different environmental conditions, weak polyelectrolytes are specially sensitive to changes in the pH-value or the ionic strength, solvent composition, interactions with metal ions or other charged molecules. The paradigmatic mechanism of CR is the binding of small ions present in solution, and, in particular, acid-base equilibria, due to the ubiquitous presence of protons in aqueous solutions [1,2]. In a wide range of situations, the physicochemical behaviour of charged polymers cannot be understood without the presence of CR. A few examples would be the stability of colloidal systems [3], protein-surface [4,5], protein-protein [6,7] and protein-polyelectrolyte [8,9] interactions, nano-particle coating [10], supramolecular chemistry [11,12], ligand-receptor binding in biochemistry [13–15], drug delivery [16], protein folding [17], among many others.

Besides the ability to modulate the electric charge to external changes, there are two relevant aspects of CR which make the difference compared to systems with fixed charge. Firstly, although CR can also take place in rigid structures, weak polyelectrolytes usually are very flexible. That means that, due to the presence of electrostatic interactions between charged groups, the system tries to minimize the electrostatic repulsion (or maximize the attraction) by means of changes in the conformational structure. In the same way, modifications in the conformational structure affect the interactions between charged sites and thus their ionization state. Conformational and ionization degrees of freedom are thus highly coupled [18,19]. The natural mechanism for this fact is the rotation of the chemical bonds. It has been recently shown that one can even build effective rotational potentials which explicitly depend on the pH-value [20,21]. A natural conclusion of this fact is that the stretching properties of weak-polyelectrolytes should depend up to some extent on the ionization state of the macromolecule (i.e., the pH-value and the ionic strength) since the application of an external force modifies its conformational structure.

In the last two-decades, the development of single-molecule force spectroscopy has lead to a huge expansion of the field of mechano-chemistry [22]. Mechanically induced chemical reactions or conformational transitions have been recently described both in neutral and charged macromolecules [23–29]. The stretching of strong polyelectrolytes such as DNA and RNA, two strong polyelectrolytes, have also been the subject of a number of studies, which results to be extremely dependent on the valence and concentration of the counterions [30–32]. It has been also shown that self-avoiding electrostatic repulsion forces produce new elastic regimes and scaling behaviors [33–35]. In a recent paper, our group explored the possibility of induced charge regulation by means of the application of an external force to a weak polyelectrolyte [36]. Mechanical stretching leads to an increase in the distance between charged groups and significant changes in the degree of protonation are observed at certain pH and ionic-strength conditions. In the same way, changes in the pH-value affect the extension/force curves.

A second aspect involved in CR is that the charge is no longer a fixed quantity but a fluctuating one, which can lead to surprising effects. For instance, it is well known that, under certain conditions, charge fluctuations (CF) can produce attraction between two macromolecules with the same average charge. This phenomenon was firstly predicted in a classical work by Kirwood and Shumaker [37], who used statistical mechanics perturbation methods and since then it has been verified in a number of simulation and experimental studies on protein-ligand, protein-protein and protein-membrane interactions [2]. Simulation methods have proved to be specially useful in the quantification of this effect by turning on and off the CF [6]. However, the influence of CF in the conformational and mechanical properties of flexible weak polyelectrolytes has been so far hardly addressed. In most the studies aimed determining the conformational properties (end-to-end distance, radius of gyration, persistence length, etc.) the charge is taken as a constant, sometimes by considering the polymer as a string with constant charge density [38,39].

In a number of works, CF are taken explicitly into account by regarding the polymer as a set of punctual protonating sites linked by rigid or harmonic bonds, so that rotational conformations are not taken into account [40–43].

The present work addresses the role of charge regulation and the associated fluctuations in the conformational and mechanical properties of weak polyelectrolytes by means of Monte Carlo simulations at constant pH-value, so that the charge is a fluctuating quantity. Once the average charge *per* site is available, we turn off the fluctuations by assigning the same average charge to every site. A constant charge MC (ccMC) simulation is then performed. With this aim, we make use of a model, described in Section 2, which accounts for the main fundamental aspects of a linear flexible polyelectrolyte that is, proton binding, dihedral angle rotation, bond stretching and bending. Steric excluded volume and differentiated treatment of short range (SR) and long-range (LR) interactions are also included. This model can be regarded as a kind of “minimal” in the sense that it contains a minimum number of parameters but it still preserves the atomistic detail. In Section 3, the effect of CF in the conformational properties of a weak polyelectrolyte at zero force is analyzed. Section 4 focuses on the influence of CR and CF in the scaling behavior at low forces, that is, linear and Pincus regime. Finally, Section 5 is devoted to the effect of CR and CF in the extension-force curves.

2. Charge Regulation and Stretching of Weak Polyelectrolytes

2.1. Minimal Model of a Weak Flexible Polyelectrolyte

Let us consider a model, depicted in Figure 1, which captures the most relevant aspects of a weak flexible polyelectrolyte (proton binding, dihedral angle rotation, bond stretching and bending) but it involves a minimum number of parameters. Bond rotation represents the main mechanism of stretching at moderate and even relatively high forces. Bond stretching and bond angle bending are only relevant at very high forces (typically larger than 500 pN) for which the stretching regime is fully independent of the ionization state, as shown in a recent publication [36]. Assuming that the polymer is symmetric, (it presents planar symmetry when it is fully extended) complications due to tacticity can be avoided. As usual, the ionization state of the macromolecule is described by a set of variables $s = \{s_i\}$ with $i = 1 \dots N$, where $s_i = 1$ indicates that the site is protonated and $s_i = 0$ otherwise. This is the basis of the Site Binding (SB) model [1]. The conformational state is determined by the set $c = \{\phi_j\}$, $j = 1 \dots M$, where ϕ_j is the rotational angle of the bond j . It will be assumed that only the rotational states corresponding to energy minima (typically *trans* (t), *gauche+* (g^+) and *gauche-* (g^-)) are significantly populated. This is the central assumption of the Rotational Isomeric State (RIS) model, firstly proposed by Flory in order to calculate conformational properties of neutral chain molecules [44]. Combining RIS and SB models, we obtain the Site Binding Rotational Isomeric State (SBRIS) model, previously proposed in order to deal with ionization and conformational degrees of freedom on the same foot [18,19,21]. The SBRIS free energy can be expressed as

$$F(s, c) = \mathcal{F}_{\text{rot}} + \mathcal{F}_{\text{p}} + \mathcal{F}_{\text{E}} + \mathcal{F}_{\text{lenght}} + \mathcal{F}_{\text{angle}} + \mathcal{F}_{\text{SEV}} + W \quad (1)$$

and

$$\beta \cdot \mathcal{F}_{\text{rot}} = \sum_{j=1}^M \epsilon_{\text{rot},j}(\phi_j) \quad (2)$$

represents the RIS free energy contribution which is the sum of the torsional energies $\epsilon_{\text{rot},j}(\phi_j)$ of the M rotating bonds and $\beta = 1/k_B T$. Under the RIS approximation, we choose for ϕ_j the possible values $\phi_j = 0$ (t), $\phi_j = +2\pi/3$ (g^+) and $\phi_j = -2\pi/3$ (g^-). Since the chain is symmetric, $\epsilon_{\text{rot}}(+2\pi/3) = \epsilon_{\text{rot}}(-2\pi/3)$. In this work, we will assume that the three rotational states have the same energy so that $\epsilon_{\text{rot},j}(\phi_j) = 0$.

The term \mathcal{F}_p represents the contribution of proton binding to the free energy

$$\beta \mathcal{F}_p = \sum_{i=1}^N \ln(10) \mu_i s_i = \sum_{i=1}^N \ln(10) (\text{pH} - \text{p}K_i) s_i \quad (3)$$

where $\mu_i = \log(K_i a_H)$ is the reduced chemical potential corresponding to the proton activity a_H and the protonation constant of site i , K_i . A site is positively charged when it is protonated. The chain contains a protonating site every three chain positions so that it can be regarded as a simplified version of linear poly-ethylene imine (LPEI), for which a complete conformational study is reported in Reference [19]. In the particular case of identical sites and bonds $\text{p}K_i = \text{p}K$.

The electrostatic interaction energy \mathcal{F}_E between charged sites is split into short range \mathcal{F}_{SR} and long range \mathcal{F}_{LR} contributions

$$\mathcal{F}_E = \mathcal{F}_{\text{SR}} + \mathcal{F}_{\text{LR}}.$$

This distinction is necessary due to the fundamental differences in the physical mechanism of SR and LR interactions. LR interactions are chemically unspecific and mediated by the solvent and they can be reasonably described by simple pair-potentials. In this work, the Debye-Hückel (DH) potential has been chosen

$$\beta \mathcal{F}_{\text{LR}} = \sum_{i=1}^N \sum_{j=i+2}^N \frac{\ell_B}{d_{ij}} e^{-\kappa d_{ij}} s_i s_j \quad (4)$$

where d_{ij} is the distance between the sites i and j , $\ell_B \simeq 0.7$ nm and κ^{-1} (nm) = $0.304/\sqrt{I(\text{M})}$ represent, respectively, the Bjerrum and the Debye lengths in water at 298.15 K and ionic strength I . SR interactions between neighboring sites, however, are mediated by the macromolecular skeleton and they strongly depend on the chemical environment of the interacting sites [1,19]. As a consequence, they need specific parameters to be described. \mathcal{F}_{SR} reads

$$\beta \mathcal{F}_{\text{SR}} = \ln(10) \sum_{i=1}^{N-1} \epsilon_{\text{int},3i-1}(\phi_{3i-1}) s_i s_{i+1} \quad (5)$$

where $\epsilon_{\text{int},3i-1}(\phi_{3i-1})$ corresponds to the interaction energy between two neighboring sites linked by a bond in a given rotational state ϕ_{3i-1} . This term clearly couples the ionization and the conformational degrees of freedom, a characteristic feature of the SBRIS model.

The elasticity of the bond length l_j and bond angle α_j are included *via* the harmonic potentials

$$\mathcal{F}_{\text{length}} = \sum_{j=1}^M \frac{k_{\text{length},j}}{2} (l_j - l_{j,o})^2, \quad (6)$$

and

$$\mathcal{F}_{\text{angle}} = \sum_{j=1}^{M-1} \frac{k_{\text{angle},j}}{2} (\alpha_j - \alpha_{j,o})^2, \quad (7)$$

where $k_{\text{length},j}$ and $k_{\text{angle},j}$ are the bond stretching and bending force constants and $l_{j,0}$ and $\alpha_{j,0}$ denote the equilibrium length and the equilibrium bending angle of bond j , respectively. The stretching of the chain exerted by the applied force F is quantified by the mechanical work

$$W = -F r \quad (8)$$

where r is the end-to-end vector. Finally, \mathcal{F}_{SEV} accounts for the steric excluded volume (SEV) effects, due to the finite size of the sites and chemical groups composing the chain. They are implemented by means of a hard-sphere potential

$$\mathcal{F}_{\text{SEV}} = \sum_{i,j=i+4}^{M+1} \begin{cases} \infty & d_{ij} \leq R_i + R_j \\ 0 & d_{ij} > R_i + R_j \end{cases} \quad (9)$$

where R_i is the hard-sphere radius of site i . They will be turned on and off in the computations throughout this work in order to quantify their importance in the conformational and stretching properties. When SEV effects are switched off, all sites are considered to be a point charge, that is, $R_i = 0 \forall i$. More details about the model and its approximations can be found in Reference [36].

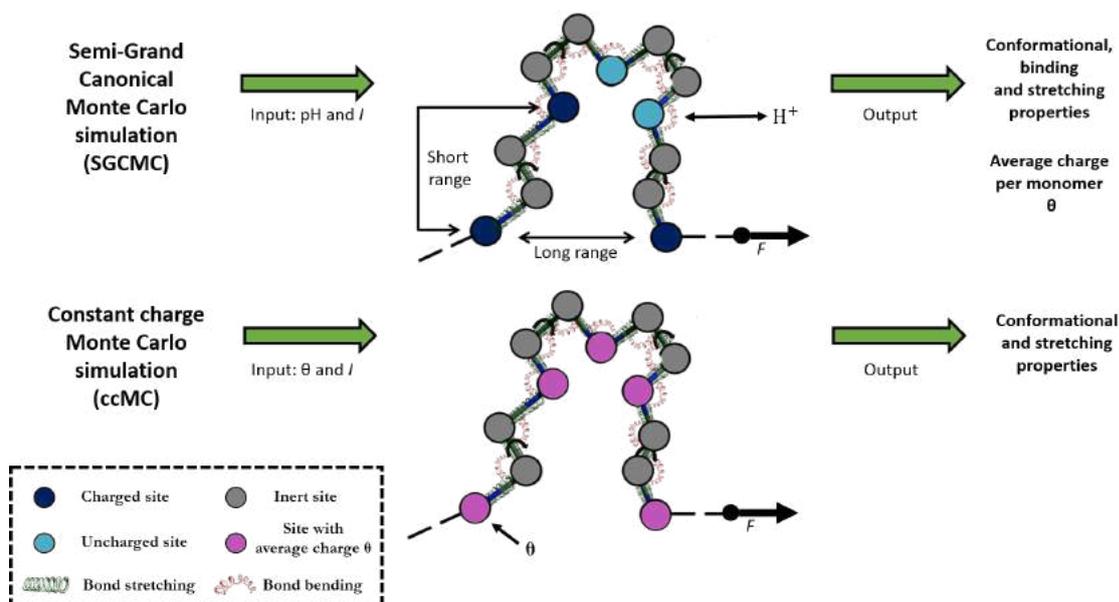


Figure 1. Outline of the two different kind of simulations of a weak flexible polyelectrolyte performed in this study: Semi-Grand Canonical Monte Carlo (SGCMC) and constant charge Monte Carlo (ccMC) simulations. In SGCMC, the pH-value is kept constant and the charge is free to fluctuate by means of proton equilibria (blue and cyan circles depict protonated and deprotonated sites, respectively). Conversely, in ccMC simulations, the charge of a site is fixed to its average value (purple circles). Grey circles represent inert sites. The bonds holding two ionizable sites are allowed to rotate. Bond stretching and angle bending and the mechanical stretching due to the action of an external force are also included in the model.

2.2. Constant Charge versus Constant-pH Simulations

In order to explore the effect of CR and CF on the structural and stretching properties of a single linear weak polyelectrolyte, we make use of a Semi-Grand Canonical Monte Carlo (SGCMC) code previously

developed by our group [19,21]. This program has been recently used to study the possibility of inducing CR by means of mechanical stretching [36]. Since in the present work we are particularly interested in the behavior of the system when CR and CF are turned on and off, the code has been modified in order to perform both constant-pH (SGCMC) and constant charge Monte Carlo (ccMC) simulations, as outlined in Figure 1. In SGCMC simulations, the pH is the control variable and CR is explicitly taken into account. Statistical averages of the conformational, stretching and binding properties are computed and they are pH-dependent. The average charge per monomer θ is defined as

$$\theta = \frac{\langle N_+ \rangle}{N} = \langle s_i \rangle, \quad (10)$$

where $\langle N_+ \rangle$ is the average number of protonated sites. Note that since the simulations are performed at constant pH, N_+ is a fluctuating quantity. The macroscopic quantity which measures the intensity of the fluctuations is the Binding Capacitance C [2,13] which is defined as the variance of the probability distribution of N_+

$$C = \langle (N_+ - N\theta)^2 \rangle = N \left(\frac{\partial \theta}{\partial \mu} \right)_F \quad (11)$$

The second identity can be proved by means of elementary statistical mechanics. As commented in the introduction, CF are the responsible of interesting phenomena such as the effective attraction of macromolecules with the same charge sign, under certain conditions [37].

In order to assess the relevance of CF, we will compare the same conformational properties of interest obtained from ccMC and SGCMC simulations for the same average charge. Unlike SGCMC, in ccMC simulations CR and CF are switched off. The charge state of each site is imposed beforehand and kept constant at the value θ obtained from the equivalent SGCMC simulation, which means that θ is an output for SGCMC and an input for ccMC. Since correlations and CF are absent, the ionization and conformational degrees of freedom are now decoupled. The ionization states of two different sites become desconected. As a result, the average electrostatic interaction between sites i and j energy $\epsilon_{\text{int},i,j}$ reads

$$\langle \epsilon_{\text{int},i,j}(c) s_i s_j \rangle \simeq \epsilon_{\text{int},i,j}(c) \langle s_i s_j \rangle \simeq \epsilon_{\text{int},i,j}(c) \theta^2 \quad (12)$$

which is exact only in the limit of long enough distances between sites. This is equivalent to take $s_i = \theta$ in Equations (4) and (5). As a consequence, only conformational degrees of freedom are free to change and \mathcal{F}_p does not contribute to the free energy.

Concerning the conformational properties, special attention will be paid to the *gauche* state probability

$$P(g) = \frac{\langle M_g \rangle}{M} \quad (13)$$

where $\langle M_g \rangle$ is the average number of bonds in the *gauche* state and to the persistence length l_p , defined as the average sum of the projections of all the bonds $j \geq i$ on a given bond i in an infinitely long chain

$$l_p/l_0 = \sum_{j \geq i} \langle \mathbf{b}_i \cdot \mathbf{b}_j \rangle \quad (14)$$

where \mathbf{b}_i denote unitary vectors pointing to the direction of the bonds. l_p is related to the average square end-to-end distance $\langle \mathbf{r}^2 \rangle$ by the relationship [44]

$$l_p = \frac{\langle \mathbf{r}^2 \rangle}{2Ml_0} + \frac{l_0}{2} = \frac{\langle (\mathbf{r}_{M+1} - \mathbf{r}_1)^2 \rangle}{2Ml_0} + \frac{l_0}{2} \quad (15)$$

where $\langle \mathbf{r}_i \rangle$ is the position of site i .

2.3. Parameters Used in the Simulations

We consider a chain with $N = 50$ identical ionizable sites with $pK = 9$, a large enough number to avoid end effects. The interaction parameters are $\epsilon_{\text{int}}(t) = 1$ and $\epsilon_{\text{int}}(g^+) = \epsilon_{\text{int}}(g^-) = 3$, which means that a bond holding two neighboring charged sites has a very small probability of being in the *gauche* state. As in LPEI, the chain consists of an ionisable site every two inert sites (i.e., 148 nodes or $M = 147$ bonds), as shown in Figure 1. When excluded volume interactions are included, the sites are treated as hard spheres with radii $R = 1.7 \text{ \AA}$ and $R = 1.55 \text{ \AA}$ for inert and ionisable sites, respectively. These values were used in a previous study on the conformational and binding properties of LPEI [19]. For simplicity, all ccMC simulations have been performed without SEV effects, that is, with $R = 0$. All bonds are considered to have equal bond stretching and bending parameters $l_0 = 1.5 \text{ \AA}$, $\alpha_0 = 120^\circ$, $k_{\text{length}} = 300 \text{ kcal mol}^{-1} \text{ \AA}^{-2}$ and $k_{\text{angle}} = 0.01 \text{ kcal mol}^{-1} \text{ deg}^{-2}$, which are the typical values used in Molecular Dynamics for C-C bonds [45]. Only the bonds with pending ionisable sites are allowed to exhibit free internal angle rotation ($\epsilon_\sigma(t) = \epsilon_\sigma(g^+) = \epsilon_\sigma(g^-) = 0$) whereas the rest of the bonds are forced to remain in *trans* conformation. The simulations are performed at room temperature $T = 298.15 \text{ K}$. The reported results represent the average over 8 to 16 different SGCMC simulations, which have been equilibrated in the first 5×10^7 configurations. The thermal averages have been computed in the following 1×10^9 realizations. More computational details about the used algorithm can be found in the Supplementary Materials and explained in detail in Reference [36].

3. Effect of Charge Regulation in the Binding and Conformational Properties at Zero Force

Let us first analyze the influence of charge regulation on the conformational properties of the polyelectrolyte when no mechanical force is applied. Since the average degree of protonation θ is at once output from the SGCMC and input for ccMC simulations, we firstly discuss the dependence of this quantity on the pH-value. The resulting titration curves are shown in Figure 2a for four different ionic strengths ranging from 1 to 0.001 M (from top to bottom). As a general trend, it is observed that θ decreases when lowering the ionic strength, since electrostatic interactions become stronger and more energy is needed to protonate a site.

It is important to recall that in our model nearest neighbor (SR) interactions and LR interactions are treated in a different way. SR are described by chemically specific parameters (ϵ_{int}) which, in accordance with experiments [1,19], are taken as independent of the ionic strength. At $\theta \simeq 1/2$, which corresponds to the plateau observed in the titration curve, an ordered structure consisting of alternated protonated and deprotonated sites is formed [46]. Due to the interaction with the two neighbors, the empty sites now bind protons with a different effective pK -value (roughly $pK - 2\epsilon_{\text{int}}$). As a result, the titration curve resembles that of a system with two different pK -values.

Conversely to SR, LR interactions are described by the DH potential, which is strongly dependent on the ionic strength. LR interactions tend to destroy binding correlation between neighboring sites and it is more amenable to mean-field treatments. This is reflected in the shape of the titration curves when the ionic strength decreases, which becomes flatter and the plateau at $\theta \simeq 1/2$ progressively disappears. The differences between SR and LR interaction can be better understood by computing the binding capacitance C as a function of the pH-value, which is shown in Figure 2b. As expressed by Equation (11), C allows to directly quantify the intensity of charge fluctuations. It can be observed that, for large ionic

strength, C presents two maxima at $\text{pH} \simeq 6$ and $\text{pH} \simeq 9$, which correspond to the inflection points in the titration curve and a maximum in the charge fluctuation. Furthermore, a minimum value is observed at $\text{pH} \simeq 7.5$, related to the presence of the “ordered” alternating state at $\theta \simeq 1/2$. When the ionic strength decreases, C becomes progressively wider and flatter until the maxima and the minimum disappear as a result of the decorrelation introduced by LR interactions.

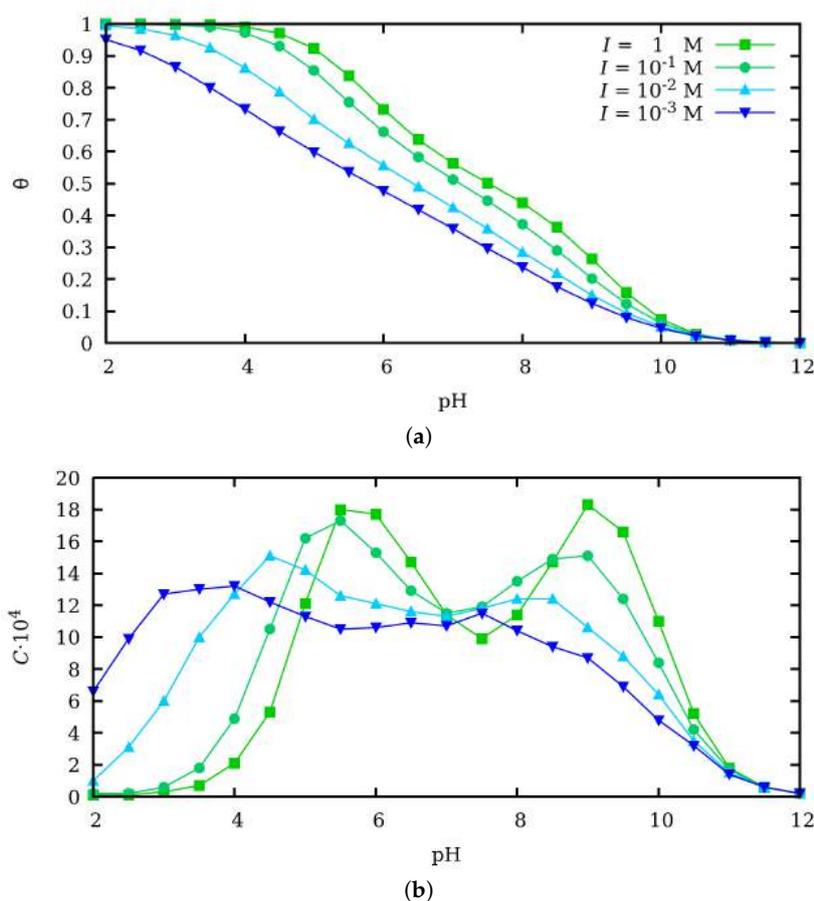


Figure 2. (a) Average degree of protonation θ and (b) binding capacitance C at ionic strengths (from top to bottom) 1 M (green squares), 0.1 M (turquoise circles), 0.01 M (cyan upwards triangles) and 0.001 M (blue downwards triangles) obtained by means of pH-constant, SGCMC simulations. Steric excluded volume contribution is not shown because no effect is observed in the obtained θ -values.

Let us turn CR and CF off and discuss the effect in the conformational properties. The θ -value obtained from SGCMC is now used as an input in the ccMC simulations, in which all the sites have a fixed charge equal to θ . Unfortunately, up to our knowledge, there is no established theory for the role of CF in the intra-molecular interactions and its consequences on the conformational structure of flexible weak polyelectrolytes. However, the contribution of CF in the force between two polyelectrolytes has been the object of a number of previous works, both from the theoretical and experimental point of view [2,6,37]. The interaction energy $U(R)$ between two identical macromolecules with average charge Q separated by a distance R depend not only on their average charge but also on their binding capacitances. In absence of counterions, it is given by [2,37]

$$\beta U(R) = \frac{\ell_B Q^2}{R} - \frac{\ell_B^2}{R^2} QC - \frac{\ell_B^2}{2R^2} C^2. \quad (16)$$

Probably, the most striking and counter-intuitive consequence of this theory is the presence of negative attractive terms in (16), despite the charge of both interacting molecules has the same sign. The first term corresponds to the usual coulombic repulsion. The second and third term in the r.h.s. of (16) are a direct consequence of CR and CF.

One of the objectives of this work is to clarify whether a similar effect can play a role, not only for inter-molecular interactions between charged macromolecules but in the intra-molecular interactions between different regions of a polyelectrolyte. If this was the case, one could expect CR to facilitate polymer folding due to the attractive contribution of CF. In order to put some light on this point, we have investigated the influence of CR in two quantities: the probability of a bond to be in *gauche* state, $P(g)$, and the persistence length l_p . $P(g)$ is a “local” conformational property, in the sense that it refers to the behavior of a single bond. The persistence length, l_p , however, is a “global” quantity, closely connected with the end-to-end distance given by Equation (15) and its value is the result of the behavior of many coupled bonds and sites.

In Figure 3, the *gauche* state probability $P(g)$ (a) and the persistence length l_p (b) are plotted as a function of θ with CR (filled markers) and without CR (empty markers) for ionic strengths ranging from 1 to 0.001 M. We have also included the results of SGCMC simulations if the SEV effects are also present (star-shaped markers). Clearly, in all the studied cases, SEV effects are very weak and almost irrelevant compared to the self-avoiding electrostatic repulsions. The corresponding capacitance versus θ is also reported in Figure 3c. First of all, in Figure 3a it is clear that in both cases, with CR and without CR, $P(g)$ decreases when the ionic strength decreases, since electrostatic screening is weaker, LR interactions are stronger and the macromolecule swells by forming more *trans* states. Moreover, two limiting behaviors for $P(g)$ can be observed, which are also common to SMGMC and ccMC simulations. For low θ -values, the polyelectrolyte is uncharged so that $P(g) \rightarrow 2/3$ since the three conformational states have the same energy. When excluded volume is included, the asymptotic value is slightly smaller since some combinations involving *gauche* conformations are forbidden (See Figure 3a inset). For θ close to unity, the chain is fully extended in order to minimize the electrostatic repulsion, which implies that all the bonds are in *trans* state and $P(g) \rightarrow 0$. Clearly, there are no fluctuations in these two limiting situations and the binding capacitance is very small. However, at intermediate θ -values, CR and CF are important as the values of the binding capacitance indicate and clear differences arise between ccMC and SGCMC simulations. $P(g)$ is significantly smaller without CR than with CR, independently of the θ -value and ionic strength. This fact can be explained because fluctuations allow to create uncharged regions in the chain which allow the chain to fold (through forming *gauche* states) while the extended regions preserve the total average charge. By assuming that all the sites have the same average charge, the possibility of the interplay between conformation and charge equilibria is lost and, as a result, the polyelectrolyte chain gets stiffer. This effect is confirmed by the behavior of the persistence length in Figure 3c. For intermediate θ -values ($0.3 < \theta < 0.8$), ccMC clearly overestimates l_p with specially significant deviations at high ionic strengths, that is, when CF are larger. As a conclusion, CR allows the chain to get more folded while the absence of CR makes it stiffer. Probably, a mechano-statistical theory based on first principles, similar to Kirkwood theory for interacting macromolecules [37], would be desirable in order to understand this point better.

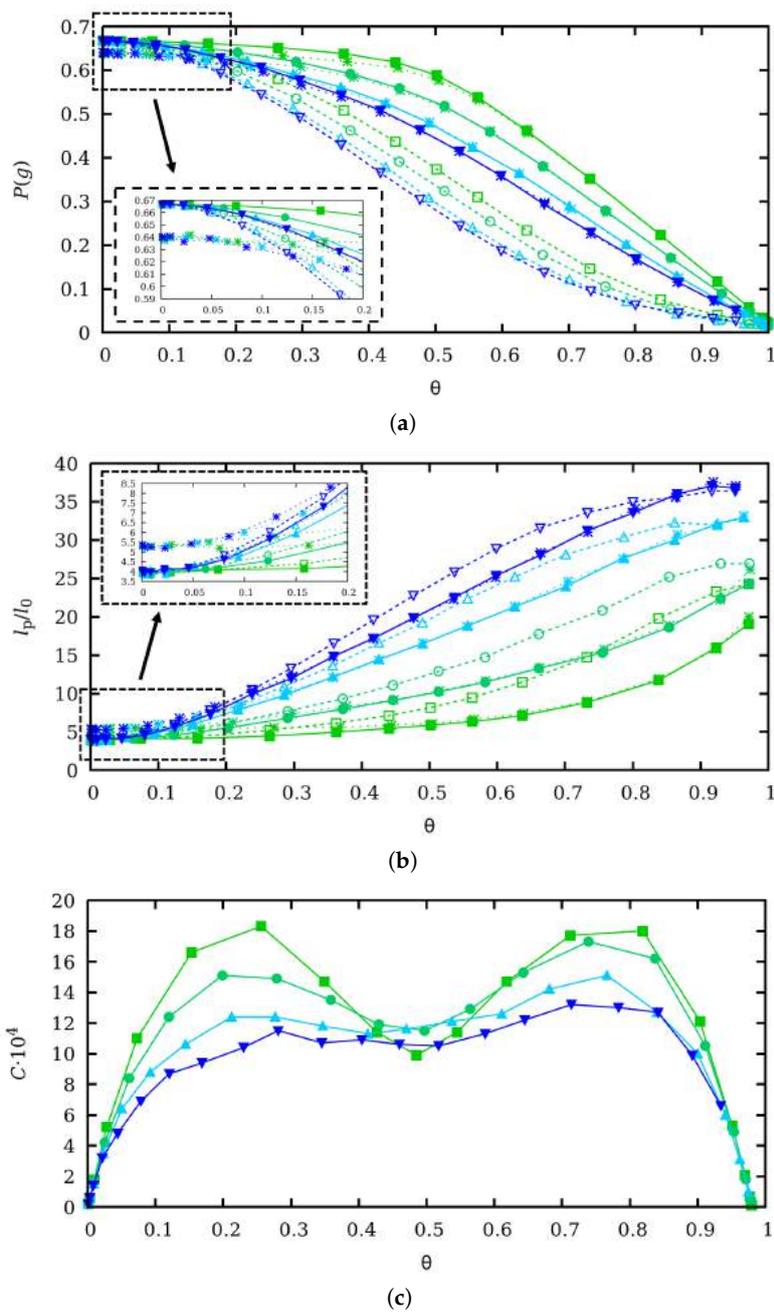


Figure 3. (a) *gauche* state probability, $P(g)$, (b) persistence length l_p and (c) binding capacitance C vs. the average degree of protonation θ at ionic strengths (from top to bottom) 1 M (green squares), 0.1 M (turquoise circles), 0.01 M (upwards cyan triangles) and 0.001 M (downwards blue triangles). l_p is normalized to the equilibrium bond length $l_0 = 1.5 \text{ \AA}$. Filled markers correspond to SGCMC without excluded volume, while empty markers correspond to ccMC simulations. In Figure 3a,b, star-shaped markers denote results obtained with SGCMC but they include steric excluded volume effects. The low charge regime is amplified in the insets of Figure 3a,b.

4. Scaling Properties of Mechanical Stretching in the Low Force Regime

Let us discuss the role of charge fluctuations in the stretching properties at the low force regime, that is, when $F < k_B T/l_p \simeq 1$ pN, under which the chain can be seen as a set of freely joined fragments with characteristic length equal to the Kuhn length l_K [47]. This regime can be in turn divided into two different sub-regimes. For very low forces ($F < 0.3$ pN), the extension L_z responds linearly to force

$$L_z/Ml_0 = \beta F \frac{l_K}{3} = \beta F \frac{2l_p - l_0}{3} \quad (17)$$

which is a direct consequence of the fluctuation-dissipation theorem [48]. The Kuhn length l_K is related to the persistence length as $l_K = 2l_p - l_0$ [44]. It is important to note that l_p , as it will be shown in the next section, can only be considered constant in the low force regime. This fact is due to the activation of the rotational degrees of freedom at intermediate forces [36]. For forces ranging $0.3 < F < 1$ pN, the action of electrostatic self-avoiding interactions makes the force/extension curve to follow the Pincus scaling law [33,48,49]

$$L_z \propto F^{1/\nu-1} \quad (18)$$

which indicates the existence of a second low force sub-regime. The limiting value $\nu = 3/5$ corresponds to strong polyelectrolytes such as DNA whereas for $\nu = 1/2$ the linear behavior of a phantom chain is recovered.

The extension/force curves resulting from SGCMC simulations (markers) are shown in Figure 4 for pH-values ranging from 2 to 10 (from top to bottom). Dashed lines are obtained solving Equation (17) using the l_K value from simulations at zero force, at the corresponding pH and I values. The best fit of Pincus scaling law Equation (18) (continuous lines), in its range of validity $0.3 < F < 1$ pN, is also plotted. Two ionic strengths are considered: $I = 1$ M (Figure 4a) and $I = 0.001$ M (Figure 4b). For large pH-values that is, when the polyelectrolyte is neutral, $\nu = 1/2$ and linear behavior is found for all the force-values. In the other limiting situation, for very low pH-values, when the polyelectrolyte is fully charged, simulations deviate from the linear behavior and they follow the Pincus scaling law (Equation (18)) with $\nu \simeq 3/5$, which nicely matches with the theoretical predictions [49]. The presence of SEV does not affect this conclusion, as shown in the Supplementary Materials.

For intermediate pH-values and charges, a preliminary study [36] suggested that weak polyelectrolytes exhibit intermediate ν -values between the two limiting cases. However, whether such a transition is consequence of the CF or it is a direct consequence of the change in the average charge induced by modifying the pH-value is not clear. With this aim, we perform ccMC simulations using as input the average degree of protonation θ obtained from SGCMC at $F = 0$ (Figure 2a). The scaling exponents ν resulting from SGCMC without SEV (filled markers), SGCMC with SEV (star-shaped markers) and ccMC simulations (empty markers) versus θ are reported in Figure 5 for two ionic strengths $I = 1$ M (green squares) and $I = 0.001$ M (blue circles).

Comparing the results obtained from SGCMC and ccMC simulations, it is observed that, in effect, ν presents a transition between linear and Pincus behavior both in presence or in absence of CR. Curiously, for $I = 0.001$ M the transition between the two limiting ν -values is found to be linear with θ

$$\nu = m\theta + n \quad (19)$$

where $m = 0.129 \pm 0.004$ and $n = 0.494 \pm 0.001 \simeq 0.5$ are fitted parameters to the SGCMC points. Best fitted Equation (19) is depicted as a black dashed line in Figure 5. The situation, however, becomes more complex for $I = 1$ M, for which no significant deviation from the limiting value $\nu = 1/2$ is observed for $\theta < 0.5$. For $\theta > 0.5$. A drastic increase in the ν -value is formed when CR is taken into account. It is also

worth mentioning that the limiting value $\nu = 3/5$ is not observed even at $\theta = 1$, when the polyelectrolyte is fully charged, independently of the presence of CF. This fact suggests that this could be also the case for strong polyelectrolytes at high ionic strengths. Again, SEV does not produce any significant effect in ν , even when the polyelectrolyte is almost uncharged ($\theta \simeq 0$).

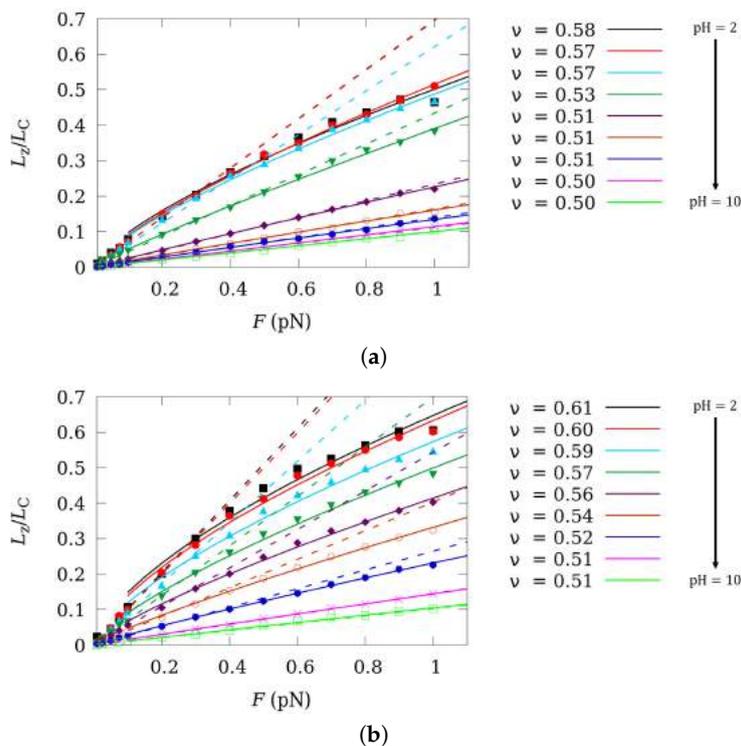


Figure 4. Normalized chain extension versus force curves in the low force regime for pH-values ranging from 2 to 10 (from top to bottom) obtained from SGC MC simulations without SEV at two ionic strengths 1 M (a) and 0.001 M (b). Dashed and continuous lines represent the best fit linear and Pincus scaling law, respectively. Chain extension is normalized to the contour length $L_C = Nl_0 \cos((\pi - \alpha_0)/2)$.

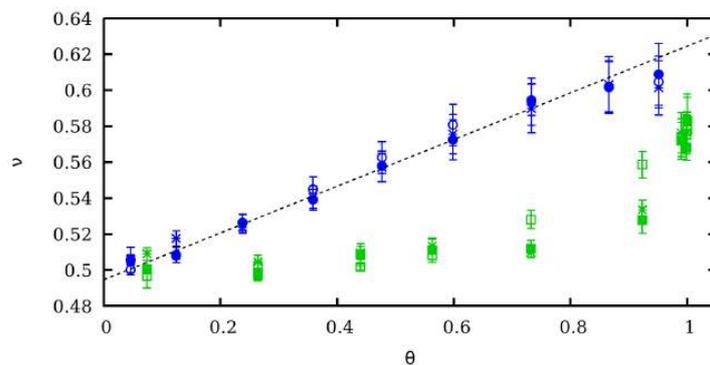


Figure 5. Pincus scaling exponent ν versus θ at two ionic strengths: 1 M (green markers) and 0.001 M (blue markers). The results of three different kind of simulations are plotted: SGC MC without SEV (filled markers), SGC MC with SEV (star-shaped markers) and ccMC (empty markers). Dashed and continuous lines represent the best fit linear and Pincus scaling law, respectively.

5. Influence of Charge Fluctuations in the Intermediate Force Regime

It is well-established that conformational and binding degrees of freedom in weak polyelectrolytes are highly correlated because of charge regulation. Since in flexible polyelectrolytes, mechanical stretching dramatically changes the distance within the macromolecule, changes in θ are expected when an external force is applied. This effect has been extensively discussed in a recent publication [36]. θ versus force is plotted in Figure 6 at different pH-values (4, 6, 8 and 10) and for two ionic strengths $I = 1$ M (Figure 6a) and $I = 0.001$ M (Figure 6b). For intermediate pH-values and low ionic strengths, moderate CR is observed. This effect is significantly enhanced if the *gauche* bond state is favored, for instance, because of the formation of a hydrogen bond, although this issue is out of the scope of this work. Although this point would probably require a more intensive study, we present here some preliminary results.

The influence of CR and CF in the extension-force curves is evaluated by again comparing the results obtained by SGCMC and ccMC. They are shown in Figure 7 for the same pH-values and ionic strengths as in Figure 6. It is observed that for $I = 1$ M clear differences can be appreciated for pH = 8 but specially for pH = 6. It is interesting to note that those are the conditions under which the binding capacitance and thus the binding fluctuations are larger (see Figure 2b). As expected, no effect is observed when the polymer is fully charged (pH = 4) or fully uncharged (pH = 10). For $I = 0.001$ M fluctuations are weaker and the observed differences are smaller than for $I = 1$ M. Although not shown in the figure, we find that SEV does not affect significantly the extension/force curves. Similar trends are found in the persistence length, depicted in Figure 8. Note that, again, the persistence length is a function of the pH-value for intermediate forces due to the activation of the rotational degrees of freedom. The main differences between SGCMC and ccMC also take place for pH = 6 and $I = 1$ M, when correlation effects are larger.

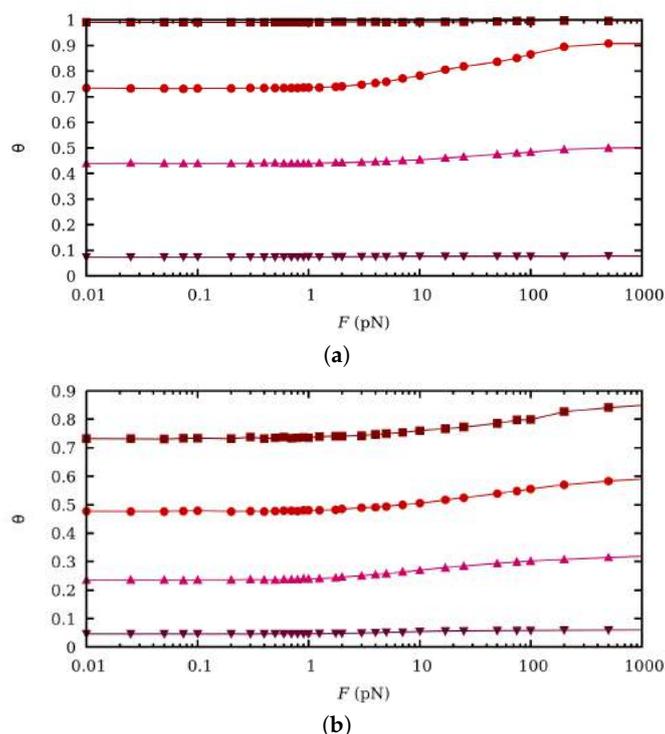


Figure 6. θ versus force curves at fixed pH-values 4 (wine squares), 6 (red circles), 8 (pink upwards triangles) and 10 (purple downwards triangles), from top to bottom, and ionic strengths 1 M (a) and 0.001 M (b), the simulation results correspond to SGCMC without excluded volume.

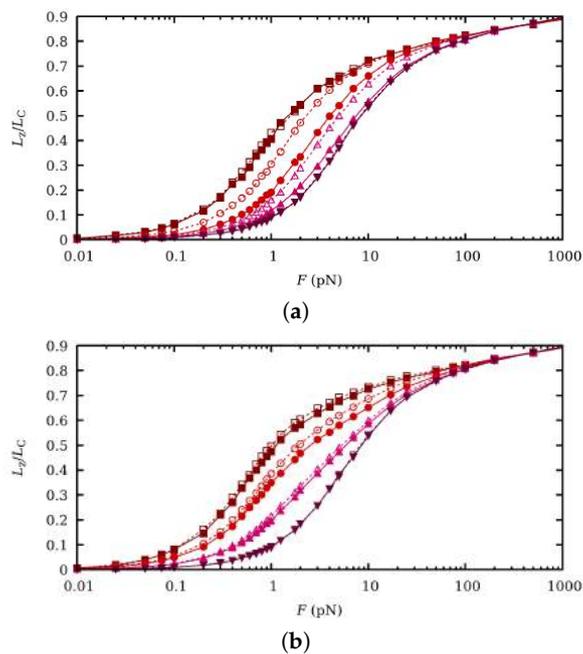


Figure 7. Normalized chain extension versus force curves at fixed pH-values 4 (wine squares), 6 (red circles), 8 (upwards pink triangles) and 10 (downwards purple triangles), from top to bottom, and ionic strengths 1 M (a) and 0.001 M (b). Filled markers correspond to SGCMC while empty markers refer to ccMC. The chain extension L_z is normalized to the contour length $L_C = Nl_0 \cos((\pi - \alpha_0)/2)$.

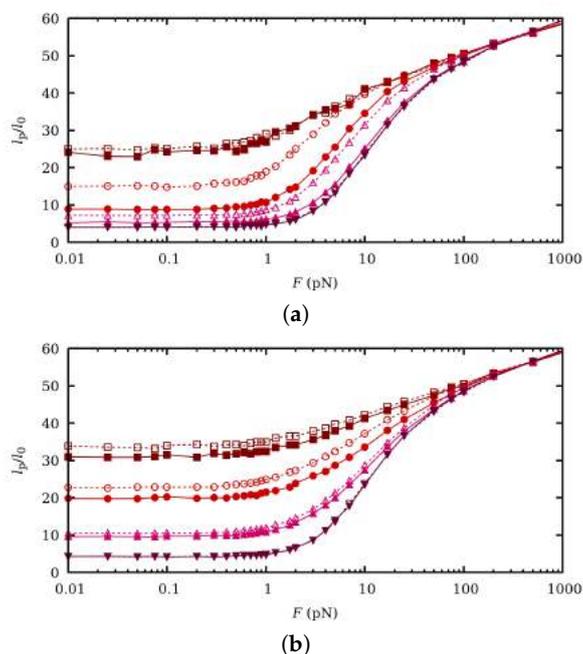


Figure 8. Normalized persistence length l_p versus force at pH-values (from top to bottom) 4 (wine squares), 6 (red circles), 8 (upwards pink triangles) and 10 (downwards purple triangles) and ionic strengths 1 M (a) and 0.001 M (b). Filled markers correspond to SGCMC while empty markers refer to ccMC. l_p is normalized to the equilibrium bond length $l_0 = 1.5 \text{ \AA}$.

6. Conclusions

This work addresses the role of charge regulation (CR) and the resulting fluctuations in the conformational and mechanical properties of flexible weak polyelectrolytes, moreover it is motivated by recent findings which suggest that charge regulation can be induced by mechanical stretching [36]. Due to CR, changes in the pH-value modify the average macromolecular charge and conformational equilibria. A second effect is that, for a given average charge per site, fluctuations can alter the intensity of the interactions by means of the correlation between binding sites. We investigate both effects by means of Semi-Grand Canonical Monte Carlo (SGCMC) simulations at constant pH-value, so that the charge is a fluctuating quantity. Once the average charge *per site* is available, we turn off the fluctuations by assigning the same average charge to every site. The molecule is now in a “frozen” ionization state and a MC simulation at constant charge (ccMC) is performed. The main conformational and stretching properties with and without CR are then compared.

We make use of a model which accounts for the main fundamental aspects of a linear flexible polyelectrolyte that is, proton binding, dihedral angle rotation, bond stretching and bending. Steric excluded volume and specific treatment of short range and long-range interactions are also included in the model. This model can be regarded as a kind of “minimal” in the sense that it contains a minimum number of parameters but it still preserves the atomistic detail. We firstly study the case when no external force is applied. It is shown that, if fluctuations are activated, *gauche* state probabilities become larger, and the persistence length smaller, so that the polymer becomes more folded. Electrostatic repulsion is thus enhanced if the charge is fixed and weakened when charge fluctuations, which are quantified by means of the binding capacitance, are present.

In the presence of an applied force, macromolecular stretching is also analyzed with CR (the charge depends on the applied force) and without CR (the charge is fixed to the value at zero force). The analysis of the scaling behavior at the low force regime concludes that Pincus exponent becomes pH-dependent. Both with and without CR, a transition from 1/2 at high pH-values (phantom chain) to 3/5 to low pH-values (Pincus regime), is observed. This fact suggests that Pincus regime is essentially driven by the average charge and that CR plays a minor role. Finally, the intermediate force stretching regime is investigated. It is found that CR induces a moderate influence in the force-extension curves and in the persistence length (which in this force regime becomes force-dependent). It is thus concluded that the effect of CR on the stretching curves is mainly due to changes in the average charge at zero force, although some differences arise at intermediate pH-values. It is also found that the effect of steric excluded volume is almost irrelevant compared to electrostatic self-avoiding interactions for all the cases studied.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2073-4360/11/12/1962/s1>, Figure S1: Outline of the metropolis algorithm of the simulation code, Figure S2: Normalized chain extension versus force curves in the low force regime for pH-values ranging from 2 to 10 obtained from ccMC and SGCMC simulations with SEV.

Author Contributions: Conceptualization, F.M. and J.L.G.; Methodology, F.M. and J.L.G.; Software, P.M.B. and S.M.; Validation, P.M.B. and C.F.N.; Investigation, P.M.B. and J.L.G.; Writing—Original Draft Preparation, P.M.B., S.M., F.M. and J.L.G.; Writing—Review & Editing, P.M.B., S.M., C.F.N., F.M. and J.L.G.; Visualization, P.M.B. and J.L.G.

Funding: P.M.B., S.M., J.L.G. and F.M. acknowledge the financial support from Generalitat de Catalunya (Grants 2017SGR1033, 2017SGR1329 and XrQTC), and Spanish Structures of Excellence María de Maeztu program through (Grant MDM-2017-0767). J.L.G. also acknowledges the Spanish Ministry of Science and Innovation (project CTM2016-78798-C2-1-P). F.M. and S.M. acknowledge the funding of the EU project 8SEWP-HORIZON 2020 (692146). P.M.B. also acknowledges the financial support from grant (FI-2017) of Generalitat de Catalunya. C.F.N. acknowledges the support from Universidad Tecnológica Nacional of Argentina under project PID 5173.

Acknowledgments: Language assistance by Andrea Muñoz is gratefully acknowledged.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ccMC	constant charge Monte Carlo
CF	Charge Fluctuation
CR	Charge Regulation
LPEI	Linear PolyEthylenImine
LR	Long Range
MC	Monte Carlo
RIS	Rotational Isomeric State
SB	Site Binding
SBRIS	Site Binding Rotational Isomeric State
SEV	Steric Excluded Volume
SGCMC	Semi-Grand Canonical Monte Carlo
SR	Short Range

References

- Borkovec, M.; Jönsson, B.; Koper, G.J.M.J. *Surface and Colloid Science*; Matigeric, E., Ed.; Plenum Press: New York, NY, USA, 2001; Volume 10, Chapter 2, pp. 99–339.
- Lund, M.; Jo, B. Charge regulation in biomolecular solution. *Q. Rev. Biophys.* **2013**, *46*, 265–281. [[CrossRef](#)] [[PubMed](#)]
- Trefalt, G.; Behrens, S.H.; Borkovec, M. Charge Regulation in the Electrical Double Layer: Ion Adsorption and Surface Interactions. *Langmuir* **2016**, *32*, 380–400. [[CrossRef](#)] [[PubMed](#)]
- Hanakam, F.; Günther, G.; Lotz, S.; Alt, T.; Seelig, A. Binding of Hisactophilin I and II to Lipid Membranes Is Controlled by a pH-dependent myristoyl–histidine switch. *Biochemistry* **1996**, *35*, 11036–11044. [[CrossRef](#)] [[PubMed](#)]
- Narambuena, C.F.; Longo, G.S.; Szleifer, I. Lysozyme adsorption in pH-responsive hydrogel thin-films: The non-trivial role of acid-base equilibrium. *Soft Matter* **2015**, *11*, 6669–6679. [[CrossRef](#)]
- Lund, M.; Jönsson, B. On the Charge Regulation of Proteins. *Biochemistry* **2005**, *44*, 5722–5727. [[CrossRef](#)]
- Mason, A.C.; Jensen, J.H. Protein-protein binding is often associated with changes in protonation state. *Proteins Struct. Funct. Bioinform.* **2008**, *71*, 81–91. [[CrossRef](#)]
- Torres, P.; Bojanich, L.; Sanchez-Varretti, F.; Ramirez-Pastor, A.J.; Quiroga, E.; Boeris, V.; Narambuena, C.F. Protonation of β -lactoglobulin in the presence of strong polyelectrolyte chains: A study using Monte Carlo simulation. *Colloids Surf. B Biointerfaces* **2017**, *160*, 161–168. [[CrossRef](#)]
- Torres, P.B.; Quiroga, E.; Ramirez-Pastor, A.J.; Boeris, V.; Narambuena, C.F. Interaction between β -Lactoglobuline and Weak Polyelectrolyte Chains: A Study Using Monte Carlo Simulation. *J. Phys. Chem. B* **2019**, *123*, 8617–8627. [[CrossRef](#)]
- Carnal, F.; Clavier, A.; Stoll, S. Polypeptide-nanoparticle interactions and corona formation investigated by monte carlo simulations. *Polymers* **2016**, *8*, 203. [[CrossRef](#)]
- Hamacek, J.; Borkovec, M.; Piguët, C. Simple thermodynamics for unravelling sophisticated self-assembly processes. *Dalton Trans.* **2006**, 1473–1490. [[CrossRef](#)]
- Li, Y.; Zhao, T.; Wang, C.; Lin, Z.; Huang, G.; Sumer, B.D.; Gao, J. Molecular basis of cooperativity in pH-triggered supramolecular self-assembly. *Nat. Commun.* **2016**, *7*, 13214. [[CrossRef](#)] [[PubMed](#)]
- Cera, E.D. Stochastic linkage: Effect of random fluctuations on a twostate process Stochastic linkage: Effect of random fluctuations on a two-state process. *J. Phys. Chem.* **1991**, *95*, 5082–5086. [[CrossRef](#)]
- Svensson, B.; Jonsson, B.; Thulin, E. Binding of Calcium to Calmodulin and Its Tryptic Fragments: Theory and Experiment. *Biochemistry* **1993**, *32*, 2828–2834. [[CrossRef](#)] [[PubMed](#)]
- Aguilar, B.; Anandakrishnan, R.; Ruscio, J.Z.; Onufriev, A.V. Statistics and physical origins of pK and ionization state changes upon protein-ligand binding. *Biophys. J.* **2010**, *98*, 872–880. [[CrossRef](#)] [[PubMed](#)]

16. Hartig, S.M.; Greene, R.R.; Dikov, M.M.; Prokop, A.; Davidson, J.M. Multifunctional nanoparticulate polyelectrolyte complexes. *Pharm. Res.* **2007**, *24*, 2353–2369. [[CrossRef](#)] [[PubMed](#)]
17. Whitten, S.T.; Garcia-Moreno E.B.; Hilser, V.J. Local conformational fluctuations can modulate the coupling between proton binding and global structural transitions in proteins. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 4282–4287. [[CrossRef](#)]
18. Garcés, J.L.; Koper, G.J.M.; Borkovec, M. Ionization Equilibria and Conformational Transitions in Polyprotic Molecules and Polyelectrolytes. *J. Phys. Chem. B* **2006**, *110*, 10937–10950. [[CrossRef](#)]
19. Garcés, J.L.; Madurga, S.; Borkovec, M. Coupling of conformational and ionization equilibria in linear poly (ethylenimine): A study based on the site binding/rotational isomeric state. *Phys. Chem. Chem. Phys.* **2014**, *16*, 4626–4638. [[CrossRef](#)]
20. Garcés, J.L.; Madurga, S.; Rey-Castro, C.; Mas, F. Dealing with long-range interactions in the determination of polyelectrolyte ionization properties. Extension of the transfer matrix formalism to the full range of ionic strengths. *J. Polym. Sci. Part B Polym. Phys.* **2017**, *55*, 275–284. [[CrossRef](#)]
21. Blanco, P.M.; Madurga, S.; Mas, F.; Garcés, J.L. Coupling of charge regulation and conformational equilibria in linear weak polyelectrolytes: Treatment of long-range interactions via effective short-ranged and pH-dependent interaction parameters. *Polymers* **2018**, *10*, 811. [[CrossRef](#)]
22. Giannotti, M.I.; Vancso, G.J. Interrogation of single synthetic polymer chains and polysaccharides by AFM-based force spectroscopy. *ChemPhysChem* **2007**, *8*, 2290–2307. [[CrossRef](#)] [[PubMed](#)]
23. Rief, M.; Gautel, M.; Oesterhelt, F.; Fernandez, J.M.; Gaub, H.E. Immunoglobulin Domains by AFM Reversible Unfolding of Individual Titin Immunoglobulin Domains by AFM. *Science* **1997**, *276*, 1109–1112. [[CrossRef](#)] [[PubMed](#)]
24. Valiaev, A.; Lim, D.W.; Oas, T.G.; Chilkoti, A.; Zauscher, S. Force-Induced Prolyl Cis - Trans Isomerization in Elastin-like Polypeptides. *J. Am. Chem. Soc.* **2007**, *129*, 6491–6497. [[CrossRef](#)] [[PubMed](#)]
25. Liu, Y.; Liu, K.; Wang, Z.; Zhang, X. Host-enhanced π - π Interaction for water-soluble supramolecular polymerization. *Chem. Eur. J.* **2011**, *17*, 9930–9935. [[CrossRef](#)] [[PubMed](#)]
26. Wang, J.; Kouznetsova, T.B.; Niu, Z.; Ong, M.T.; Klukovich, H.M.; Rheingold, A.L.; Martinez, T.J.; Craig, S.L. Inducing and quantifying forbidden reactivity with single-molecule polymer mechanochemistry. *Nat. Chem.* **2015**, *7*, 323–327. [[CrossRef](#)] [[PubMed](#)]
27. Radiom, M.; Kong, P.; Maroni, P.; Schäfer, M.; Kilbinger, A.F.M.; Borkovec, M. Mechanically induced cis-to-trans isomerization of carbon-carbon double bonds using atomic force microscopy. *Phys. Chem. Chem. Phys.* **2016**, *18*, 31202–31210. [[CrossRef](#)]
28. Jacobson, D.R.; Mcintosh, D.B.; Stevens, M.J.; Rubinstein, M.; Saleh, O.A. Single-stranded nucleic acid elasticity arises from internal electrostatic tension. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 5095–5100. [[CrossRef](#)]
29. Radiom, M.; Borkovec, M. Influence of ligand-receptor interactions on force-extension behavior within the freely jointed chain model. *Phys. Rev. E* **2017**, *96*, 062501. [[CrossRef](#)]
30. Stevens, M.J.; Saleh, O.A. Simulations of Stretching a Strong, Flexible Polyelectrolyte. *Macromolecules* **2012**, *45*, 5757–5765. [[CrossRef](#)]
31. Jacobson, D.R.; Mcintosh, D.B.; Saleh, O.A. The Snakelike Chain Character of Unstructured RNA. *Biophys. J.* **2013**, *105*, 2569–2576. [[CrossRef](#)]
32. Stevens, M.J.; Berezneya, J.P.; Saleh, O.A. The Effect of Chain Stiffness and Salt on the Elastic Response of a Polyelectrolyte. *J. Chem. Phys.* **2018**, *149*, 163328. [[CrossRef](#)] [[PubMed](#)]
33. Saleh, O.A.; McIntosh, D.B.; Pincus, P.; Ribbeck, N. Nonlinear low-force elasticity of single-stranded DNA molecules. *Phys. Rev. Lett.* **2009**, *102*, 068301. [[CrossRef](#)] [[PubMed](#)]
34. Mcintosh, D.B.; Saleh, O.A. Salt Species-Dependent Electrostatic Effects on ssDNA Elasticity. *Macromolecules* **2011**, *44*, 2328–2333. [[CrossRef](#)]
35. Stevens, M.J.; Saleh, O.A. Simulations of Stretching a Strong, Flexible Polyelectrolyte: Using Long Chains To Access the Pincus Scaling Regime. *Macromolecules* **2013**, *46*, 6369–6373. [[CrossRef](#)]
36. Blanco, P.M.; Madurga, S.; Mas, F.; Garcés, J.L. Effect of Charge Regulation and Conformational Equilibria in the Stretching Properties of Weak Polyelectrolytes. *Macromolecules* **2019**, *52*, 8017–8031. [[CrossRef](#)]

37. Kirkwood, J.G.; Shumaker, J.B. Forces between Protein Molecules in Solution Arising from Fluctuations in Proton Charge and Configuration. *Proc. Natl. Acad. Sci. USA* **1952**, *38*, 863–871. [[CrossRef](#)]
38. Odijk, T. Polyelectrolytes Near the Rod Limit. *J. Polym. Sci. Polym. Phys. Ed.* **1977**, *15*, 477–483. [[CrossRef](#)]
39. Barrat, J.L.; Joanny, J.F. Persistence Length of Polyelectrolyte Chains. *Europhys. Lett.* **1993**, *24*, 333–338. [[CrossRef](#)]
40. Ullner, M.; Jonsson, B. Monte Carlo study of titrating polyelectrolytes in the presence of salt. *Macromolecules* **1996**, *29*, 6645–6655. [[CrossRef](#)]
41. Ullner, M. Comments on the Scaling Behavior of Flexible Polyelectrolytes within the Debye-Hückel Approximation. *J. Phys. Chem. B* **2003**, *107*, 8097–8110. [[CrossRef](#)]
42. Ulrich, S.; Seijo, M.; Stoll, S. A Monte Carlo study of weak polyampholytes: Stiffness and primary structure influences on titration curves and chain conformations. *J. Phys. Chem. B* **2007**, *111*, 8459–8467. [[CrossRef](#)] [[PubMed](#)]
43. Uyaver, S.; Seidel, C. First-order conformational transition of annealed polyelectrolytes in a poor solvent. *Europhys. Lett.* **2003**, *64*, 536–542. [[CrossRef](#)]
44. Flory, P. *Statistical Mechanics of Chain Molecules*; John Wiley: New York, NY, USA, 1967.
45. Sliozberg, Y.R.; Kröger, M.; Chantawansri, T.L. Fast equilibration protocol for million atom systems of highly entangled linear polyethylene chains. *J. Chem. Phys.* **2016**, *144*, 154901. [[CrossRef](#)] [[PubMed](#)]
46. Burak, Y.; Netz, R.R. Charge Regulation of Interacting Weak Polyelectrolytes. *J. Phys. Chem. B* **2004**, *108*, 4840–4849. [[CrossRef](#)]
47. Livadaru, L.; Netz, R.R.; Kreuzer, H.J. Stretching response of discrete semiflexible polymers. *Macromolecules* **2003**, *36*, 3732–3744. [[CrossRef](#)]
48. Saleh, O.A. Perspective: Single polymer mechanics across the force regimes. *J. Phys. Chem.* **2015**, *142*, 194902. [[CrossRef](#)] [[PubMed](#)]
49. Pincus, P. Excluded Volume Effects and Stretched Polymer Chains. *Macromolecules* **1976**, *9*, 386–388. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

Chapter 6

Brownian dynamics computational model of protein diffusion in crowded media with dextran macromolecules as obstacles

6.1 Summary

6.1.1 Introduction

In the previous chapters, the conformational, ionization and elastic properties of flexible weak polyelectrolytes have been studied. As a consequence of the coupling between conformational and ionization degrees of freedom, it has been observed that these properties are highly correlated with each other. However, those studies have been done focused on a single polyelectrolyte chain, *i.e.* in infinite dilution conditions, and only the salt ions were considered to have a significant concentration.

Dilute conditions are, however, far from the real environment in which biopolymers are present in their natural media.⁸² In biological media, macromolecules in solution can occupy a 20-40 % of the total volume.⁸⁵ Unlike in dilute conditions, the non-specific interactions (excluded volume, electrostatic, hydrodynamic, among others) of macromolecules with each other can not be neglected. These conditions, named as macromolecular crowding, have been found to significantly influence the reactivity,^{82,84-90} structure⁹¹⁻⁹⁶ and diffusion⁹⁷⁻¹⁰⁵ of biopolymers.

The effect of macromolecular crowding in biopolymer diffusion have been investigated experimentally and theoretically. Experimentally, the diffusion of the macromolecule is usually measured using spectroscopic techniques, mainly Fluorescence Correlation Spectroscopy (FCS) and Fluorescence Recovery After Photobleaching (FRAP).⁹⁷⁻⁹⁹ When the experiments are performed *in vitro*, the macromolecular crowding is mimicked adding high concentrations of "inert" macromolecules (usually poly(ethylene glycol), dextran or ficoll). Theoretically, the diffusion in crowded media has been studied using computational simulation techniques, namely: Monte Carlo¹⁰⁰, Molecular Dynamics⁸⁸ and Brownian Dynamics (BD) simulations^{200,201}.

Macromolecular crowding have been found to alter significantly the diffusion of tracer macromolecules. The tracer particles exhibit three different time regimes,^{100,106} as shown in Fig. 6.1. At short times, the diffusion coefficient of the tracer is found to have a constant value D^{short} . At intermediate times, the diffusion of the tracer is anomalous, due to the collision of the tracer with the crowding agents. The diffusion coefficient becomes time-dependent and it presents a power law dependence with time, characterized by the anomalous exponent α . At long times, the diffusion coefficient returns to a constant value D^{long} since one observes the tracer macromolecule average motion in crowded media.

The crowders hinders the tracer macromolecule diffusion mainly due to excluded volume and hydrodynamic interactions. The crowder motion modifies the tracer particle motion *via* solvent-mediated interaction, *i.e.* Hydrodynamic Interactions (HI). The role of HI in diffusion

in crowded media have been studied mainly with Brownian Dynamics simulations.^{108–111,121} These studies proved that HI have an important contribution to macromolecular diffusion and they need to be included in BD to be accurate.

There are two main different approaches to include HI in BD simulations. On the one hand, one can compute the diffusion tensor using the Rotne-Prager-Yamakawa (RPY) method,^{123,124} which aims to accurately calculate the HI between particles. However, the conventional RPY diffusion tensor neglects the lubrication forces (near-field HI),^{125,126} which have been found crucial to describe diffusion in crowded media.^{111,121} The computation of the diffusion tensor is a computationally expensive procedure and several efforts have been carried out to optimize the BD calculations.¹²² On the other hand, one can use the Tokuyama method^{108–110} which introduces the HI in the BD simulations in the mean-field way avoiding the computation of the diffusion tensor.

In this chapter, the diffusion of two proteins, streptavidin and α -chymotripsin, in a media crowded with differently sized dextran macromolecules is studied by means of BD simulations. The simulations resemble previous *in vitro* experiments carried out in the same conditions^{97–99,101}. The macromolecules are modeled as hard spheres with an effective radius, as can be observed in the snapshot depicted in Fig. 6.1. The results of BD simulations using different treatments of the HI are compared with each other and with previous experimental measurements.^{97–99,101} The long-time diffusion coefficient and the anomalous exponent of both proteins are calculated at different concentrations and crowding agents.

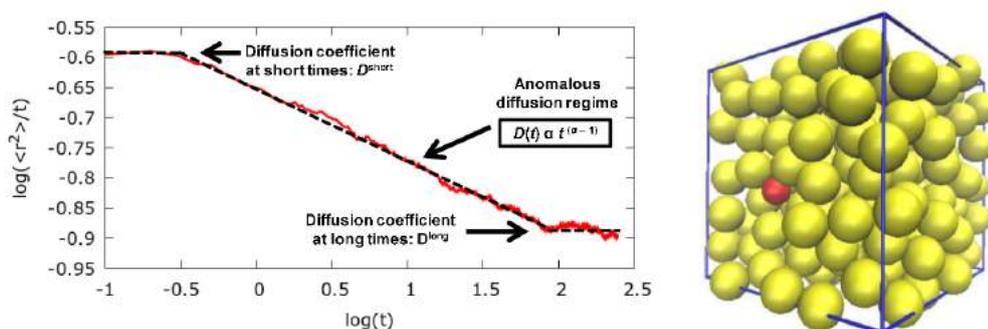


FIGURE 6.1: (left) Average diffusion coefficient $\langle r^2 \rangle / t$ vs. time t obtained by means of Brownian Dynamics with crowder volume fraction of $\phi = 0.17$. (right) Snapshot of the system studied in this chapter. A protein (red sphere) diffuses in a media crowded by dextran macromolecules (yellow spheres).

6.1.2 Results

The macromolecules are modeled, using a coarse-grained approach, as a single hard sphere. The radii of the tracer proteins, streptavidin and α -chymotripsin, is considered to be equal to its

hydrodynamic radii.^{97–99} For dextran crowding macromolecules, it has been found that using their experimental hydrodynamic radius²¹³ R_H or radius of gyration²¹⁴ R_g leads to very high excluded volume fractions at experimental values of the crowder concentration (50 - 300 g L⁻¹). This is because the experimental values of R_H and R_g are measured in dilute conditions.^{213,214} In crowded conditions, the dextran size can be reduced by the steric compression exerted by the crowder macromolecules. Alternatively, the dextran size can be estimated from its specific volume, approximating the dextran macromolecules as a compact sphere. The resulting radius, named as compact radius R_C , was found to be too small. In particular, diffusion coefficients of the proteins obtained with the BD simulations performed with crowders with radius equal to R_C were significantly larger than the experimental diffusion coefficients obtained in FCS and FRAP experiments.^{97–99} This is probably because R_C neglects the solvation effects, resulting in excluded volume fractions too small. In this work, we propose an effective radius R_{eff} for the dextran, whose value is obtained fitting the diffusion coefficient obtained by BD dynamics with the corresponding diffusion coefficient obtained from experimental measurement. The resulting R_{eff} values are observed to have intermediate values between R_H and R_C . A more detailed discussion can be found in the publication enclosed in this chapter, together with more details about the parametrization of the system and the simulations.

First, the case of a system composed with spheres of equal size is considered. This simple case is used to compare the prediction of D^{long} obtained with different treatments of the HI. In Fig. 6.2, D^{long} at different excluded volume fractions ϕ computed with BD without HI (red triangles), BD with HI computed with the conventional RPY method (blue squares) and BD with HI computed with Tokuyama method (cyan circles) is shown. D^{long} is normalized to the dilute solution diffusion coefficient of the particles D^0 . It can be observed that the BD simulations

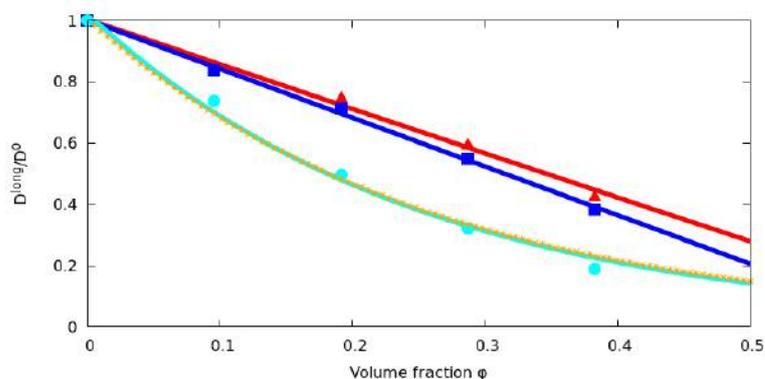


FIGURE 6.2: Long time diffusion coefficient D^{long} as a function of the crowder excluded volume ϕ obtained with: (red triangles) Brownian Dynamics (BD) without hydrodynamic interactions (HI), (cyan circles) BD with HI using Tokuyama method and (blue squares) BD with HI using the conventional RPY method. The orange dashed lines stands for the analytical prediction of D^{long} given by 1.60. The D^{long} is normalized to the ideal diffusion coefficient D^0 . The lines are only to guide the lecturer.

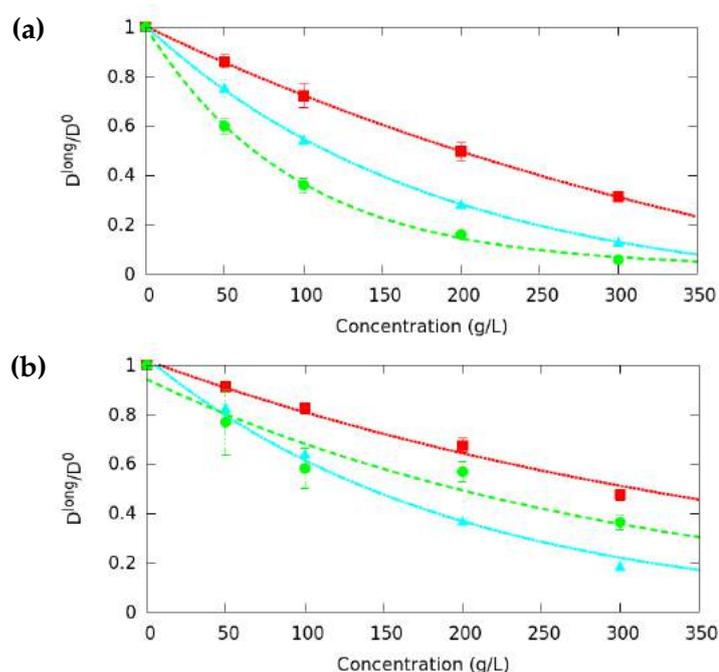


FIGURE 6.3: D^{long} as a function of concentration of the D50 crowding agent for (a) streptavidin and (b) α -chymotripsin. D^{long} is obtained from BD simulations where HI are either not included (red squares) or included with Tokuyama method (cyan triangles). The results obtained are compared with experimental measurements of the diffusion coefficient of both proteins (green circles), obtained from FCS⁹⁷ (streptavidin) and FRAP^{98,99} (α -chymotripsin) experiments. Lines are only to guide the lecturer.

with HI using the conventional RPY method are very close to those obtained in BD without HI. This result agrees with previous studies^{111,121}, where long-range HI have been found to be screened in crowded media. Conversely, near-HI or lubrication forces are expected to have a high impact on crowded media. These near-HI are better captured by the Tokuyama method, which clearly predicts a larger decay of D^{long} in increasing ϕ than that predicted by the BD simulations without HI and with HI using the RPY method. The D^{long} prediction obtained with BD using Tokuyama method can be compared with the analytical prediction of D^{long} given by Eq. 1.60 (orange dashed line). Very good agreement can be observed between the simulation results and the analytical prediction. However, when differently-sized tracer and crowders particle are considered, the analytical prediction of D^{long} is no longer accurate and the BD results show relevant deviations, as can be found in the publication presented in this chapter.

The case of streptavidin and α -chymotripsin diffusing in crowded media is analyzed now, where the tracer proteins and the crowders have different sizes. In this summary, only dextran molecules with molecular weight of 48.6 kg mol^{-1} are considered (from now on, named as D50). In the publication presented in this chapter, other dextran sizes are considered, and the results agree with the discussion that follows. In Fig. 6.3, the D^{long} of streptavidin (Fig. 6.3a) and α -chymotripsin (Fig. 6.3a) as a function of D50 concentration is shown. The simulation results

are obtained from BD without HI (red squares) and BD with HI using Tokuyama method (cyan triangles) and they are compared with the D^{long} -values reported from FCS⁹⁷ (streptavidin) and FRAP^{98,99} (α -chymotripsin) experiments (green circles). As a general trend, it can be observed that D^{long} decreases when the concentration of crowding agent increases, as a result of the protein collisions with the crowders. It can also be noted that the D^{long} obtained with BD with HI are in better agreement with the experimental data than the results obtained without HI. These results agree with previous studies highlighting the impact of HI in crowded media.¹¹¹

The anomalous exponent is also analyzed for the same cases studied in Fig. 6.3 (same markers are used), which is shown Fig. 6.4. It can be noted that α decreases when the concentration of the crowders increases. Since α is related to the transition rate between D^{short} and D^{long} , this indicates that the transition occurs faster in the more crowded systems. As a general trend, a good agreement is observed between the BD simulation results and the experimental data. It can be noted that the inclusion of the HI is not significant for the determination of α , indicating that transition between D^{short} and D^{long} is mainly driven by the collisions of the tracer particles with the crowders.

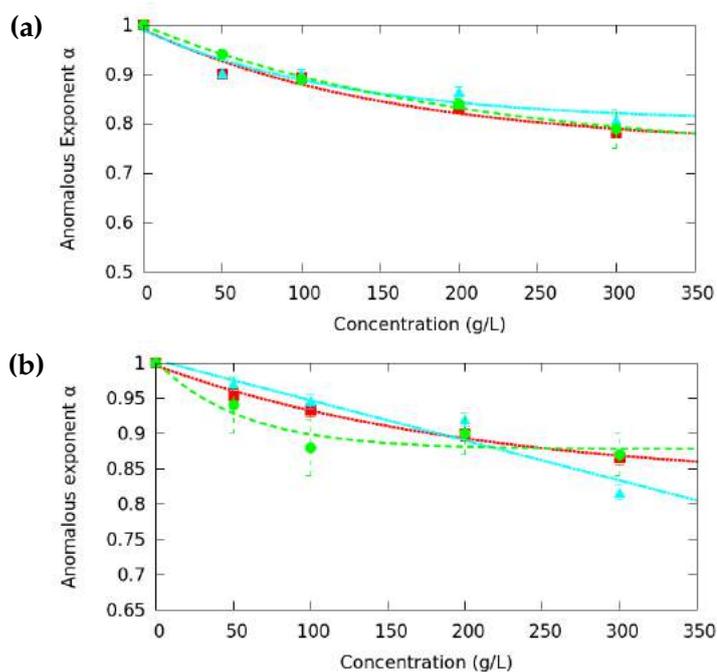


FIGURE 6.4: Anomalous diffusion exponent α as a function of D50 concentration for (a) streptavidin and (b) α -chymotripsin. D^{long} is obtained from BD simulations where HI are either not included (red squares) or included with Tokuyama method (cyan triangles). The results obtained are compared with the experimental results obtained from FCS⁹⁷ (streptavidin) and FRAP^{98,99} (α -chymotripsin) experiments (green circles). Lines are only to guide the lecturer.

6.1.3 Conclusions

The diffusion of two proteins, streptavidin and α -chymotrypsin, in crowded media is studied by means of Brownian Dynamics (BD) simulations. As crowding agents, different sizes of dextran molecules are considered, reproducing previous *in vitro* experiments.^{97–99} The macromolecules, following a coarse-grained approach, are modeled as single spheres with an effective radius, which accounts for macromolecular compaction in crowded media. Conventional BD are compared with BD with different treatments of the Hydrodynamic Interactions (HI), in particular the conventional RPY and Tokuyama methods. The long time diffusion coefficient D^{long} and the anomalous exponent α obtained from the different BD simulations is compared with each other and with the previously reported experimental results.^{97–99}

As a general trend, the D^{long} values obtained with BD with HI using Tokuyama method are smaller than the ones obtained without HI or with HI using the conventional RPY method. Moreover, the D^{long} values computed with BD with HI using Tokuyama method are in good agreement with the ones experimentally reported for most of the cases of study. The results indicate that both the steric volume and the hydrodynamic interactions are the main responsible for the D^{long} decay with macromolecular crowding. Conversely, HI are not found to have significant contribution in the determination of α .

Although, in general, qualitative agreement is observed between the BD simulations and the experimental data, some quantitative discrepancies can be observed. These discrepancies can be caused by the use of either (i) the Tokuyama method for describing the HI or (ii) the coarse-grained model employed. On the one hand, the equations proposed by Tokuyama are conceived for systems of equally-sized spheres.^{108–110} The differences of size are corrected in our simulations in an effective way by explicitly considering the size of the protein and the crowder particles. On the other hand, the hard sphere model here proposed is not able to fully reproduce the macromolecular compaction and entanglement expected in crowded media.²²⁴ In the next chapter, a new coarse-grained model will be proposed, able to account for these effects.

6.2 Publication

Article

Brownian Dynamics Computational Model of Protein Diffusion in Crowded Media with Dextran Macromolecules as Obstacles

Pablo M. Blanco ^{1,2,*}, Mireia Via ^{1,2}, Josep Lluís Garcés ³, Sergio Madurga ^{1,2,*}
and Francesc Mas ^{1,2}

¹ Department of Material Science and Physical Chemistry, Barcelona University, 08028 Barcelona, Spain; mireia.via@ehu.eus (M.V.); fmas@ub.edu (F.M.)

² Institute of Theoretical and Computational Chemistry (IQTC), Barcelona University, 08028 Barcelona, Spain

³ Department of Chemistry, University of Lleida (UdL), 25003 Lleida, Spain; jlgarcés@quimica.udl.cat

* Correspondence: pmlblanco@ub.edu (P.M.B.); s.madurga@ub.edu (S.M.)

Academic Editor: Antonio M. Scarfone

Received: 22 December 2016; Accepted: 3 March 2017; Published: 9 March 2017

Abstract: The high concentration of macromolecules (i.e., macromolecular crowding) in cellular environments leads to large quantitative effects on the dynamic and equilibrium biological properties. These effects have been experimentally studied using inert macromolecules to mimic a realistic cellular medium. In this paper, two different experimental in vitro systems of diffusing proteins which use dextran macromolecules as obstacles are computationally analyzed. A new model for dextran macromolecules based on effective radii accounting for macromolecular compression induced by crowding is proposed. The obtained results for the diffusion coefficient and the anomalous diffusion exponent exhibit good qualitative and generally good quantitative agreement with experiments. Volume fraction and hydrodynamic interactions are found to be crucial to describe the diffusion coefficient decrease in crowded media. However, no significant influence of the hydrodynamic interactions in the anomalous diffusion exponent is found.

Keywords: macromolecular crowding; Brownian dynamics; dextran modelling; macromolecule diffusion; hydrodynamic interactions

1. Introduction

The study of reaction and diffusion processes in biological media has been a challenging topic of recent research. Although single macromolecules are present in low concentrations, there is a high total concentration of macromolecules (such as proteins, polysaccharides, etc.) in cellular environments. In this context, macromolecular crowding can be defined as “*macromolecular cosolutes that are nominally inert with respect to the reaction of interest*” [1]. In general, cell cytosol presents an occupied volume fraction of 20%–30%, which means an approximate macromolecule concentration of 200–300 g/L. Moreover, macromolecular crowding is also relevant outside the cells [2] (e.g., blood plasma has a non-negligible 80 g/L protein concentration).

Since in vitro experiments are usually carried out at low concentrations (1–10 g/L), alternative approaches are necessary in order to evaluate the effect of macromolecular crowding in the thermodynamic and kinetic properties of the system. At the experimental level, high concentrations of crowding agents (usually dextran or ficoll macromolecules), which are considered to interact only by means of steric non-specific interactions, have been used to mimic the in vivo environment (usually called in vivo-like media) [3,4]. The experimental studies of macromolecule diffusion are mainly based on two experimental techniques: Fluorescence Correlation Spectroscopy (FCS) [5] and Fluorescence Recovery After Photobleaching (FRAP) [6,7].

FCS relies on the detection and temporal analysis of the fluorescence signal emitted from a small confocal detection volume. In these studies, fluorescent markers are specifically bound to the tracer molecule allowing a wide range of possible tracers (e.g., proteins, polymers, metal-complexes, etc.). The fluctuations in fluorescence intensity in a small region of the sample are recorded and used to calculate the temporal auto-correlation function, which allows determining the diffusion coefficient and the anomalous diffusion exponent of the tracer.

On the other hand, in FRAP, a small volume of the sample is lighted with a laser beam. As a result, the molecules in the lighted region become bleached and no fluorescence is exhibited. The diffusion coefficient and the anomalous diffusion exponent of the fluorescent molecule can be estimated by means of the velocity of fluorescence recovery in the bleached region.

In addition, computational studies have been performed in order to understand the effect of macromolecular crowding [8]. These studies use different approaches such as *on-lattice* Monte Carlo simulations [9] or *off-lattice* Brownian Dynamics (BD) simulations [10,11]. Recently, macromolecular crowding effect on diffusion has been modelled by means of atomistic models of cytoplasm using Molecular Dynamics [12].

Macromolecular crowding is crucial to describe macromolecular diffusion in biological media. Under these conditions, the well-known Einstein–Smoluchowski equation is no longer valid and three different diffusion regimes are observed, as shown in Figure 1 [9,13]. At short times, the particles of the system have not collided yet and the diffusion coefficient (D^{short}) remains constant. Generally, D^{short} is not equal to the dilute solution diffusion coefficient because the motion of the particles is slowed down due to Hydrodynamic Interactions (HI) with the other particles. As a result of the inter-particle collisions, the diffusion coefficient decays until it reaches a new stationary state at long times (D^{long}). In the intermediate regime, also known as anomalous regime, the mean square displacement of the particles ($\langle r^2 \rangle$) becomes non linear over time and can be expressed as:

$$\langle r^2 \rangle = (2d)\Gamma t^\alpha. \quad (1)$$

This is formally equivalent to having a diffusion coefficient that is not constant over time [14]:

$$D(t) = \frac{\Gamma t^{\alpha-1}}{2d}, \quad (2)$$

where α is the so-called anomalous diffusion exponent, d is the topological dimension of the system and Γ is a generalized transport coefficient (also known as anomalous diffusion coefficient).

The effect of HI is also a key factor in macromolecular dynamics [15]. These interactions emerge from the fact that the motion of different particles becomes correlated by means of solvent interactions. Since in BD the solvent is not explicitly included in the simulations, HI need to be included in the algorithm in an effective way. In general, two different approaches have been used to take into account the HI: those which rely on calculating the diffusion tensor [16] using the Rotne–Prager–Yamakaya (RPY) method [17,18], and those which benefit from the Tokuyama model [19–21].

The RPY method assumes a far-field approximation involving pairs of equal sized particles. The calculation of the diffusion tensor and its factorization is a computationally expensive procedure as it scales with the number of particles N as N^3 . Several approximations can be implemented in order to speed up the results of BD simulations [22]. These procedures are mainly based on avoiding the Cholesky decomposition of the diffusion tensor since it is the bottleneck of the simulation [23,24]. Recently, several efforts have been made to improve the RPY approach such as generalization to different-sized particles [25] or inclusion of many-body and near-field HI by means of the Durlofsky–Brady–Bossis method [26,27]. Moreover, several studies have shown that the inclusion of the short ranged HI or lubrication forces are crucial for describing macromolecular diffusion in crowded media since long ranged HI become screened [28,29].

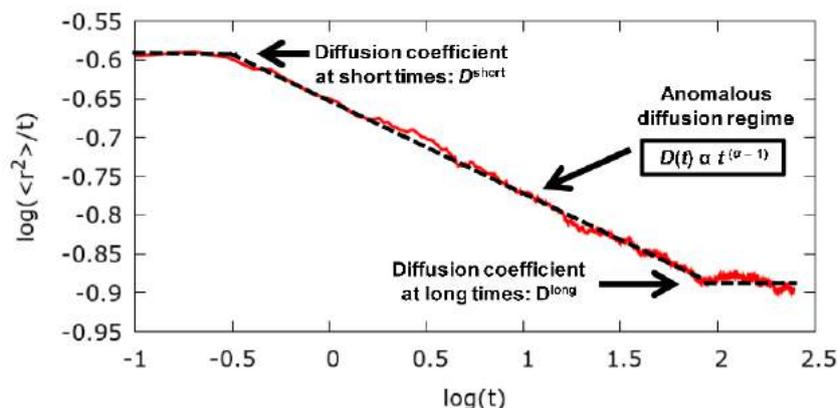


Figure 1. Time evolution of the diffusion coefficient obtained in a system with volume fraction of $\phi = 0.17$. The red line is the result of averaging 400 BD simulations of a tracer particle diffusing among 68 obstacles with radius 4.9 and 1.2 nm respectively. The simulation time is 250 ns and the length of the simulation box is 18 nm.

Unlike the RPY approach, the Tokuyama model is a mean-field approximation for equal-sized soft core spheres that mimic the self-diffusion of biomolecules in solution. This model provides a better description of short range HI than the conventional RPY diffusion tensor. This procedure has started to be used in BD simulations in crowded media [30,31] because it is computationally cheaper than the RPY approach since it allows for introducing the HI contributions without calculating the diffusion tensor.

In the present paper, macromolecule diffusion in crowded media is studied by using BD simulations. Two different experimental systems have been modelled using a new dextran model involving an effective radius which allows accounting for their steric compression in crowded media. The HI are included using the Tokuyama model and their role in diffusion is discussed. Finally, the effect of non-specific interactions, obstacle size and HI on the diffusion coefficient and the anomalous diffusion exponent is analyzed.

2. Methodology and System Parametrization

2.1. BD Simulations with HI

The large amount of solvent molecules in the macromolecule solution makes all-atom Molecular Dynamics simulations computationally very expensive. In this context, Langevin equation [32] provides a suitable procedure since the solvent is implicitly included by adding a stochastic force in the classic Newton equations of motion which accounts for the collisions of the Brownian particles with the solvent. In BD time scales, it is more convenient to apply the over-damped limit [33] for which the equation of motion reads [34]:

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) - \frac{D\Delta t}{k_B T} \nabla V(\mathbf{r}, t) + \sqrt{2D\Delta t} \boldsymbol{\zeta}(t), \tag{3}$$

where $\boldsymbol{\zeta}$ is a vector of $3N$ Gaussian random numbers with zero mean and unit variance.

Since we wish to model highly concentrated macromolecular solutions, we have chosen a coarse-grained approach where each macromolecule is modelled as a single sphere using a proper effective radius. In order to avoid overlapping, a harmonic pairwise repulsion is applied which acts when the distance between two macromolecules is smaller than the sum of their radii:

$$\mathbf{V}_{ij}(\mathbf{r}_i, \mathbf{r}_j) = \begin{cases} \frac{1}{2}k(d_{ij} - R_{ij})^2, & d_{ij} < R_{ij}, \\ 0, & d_{ij} \geq R_{ij}, \end{cases} \tag{4}$$

where R_{ij} is the sum of the radius of the interacting particles i and j , d_{ij} is the distance between i and j particles and k is a parameter accounting for the stiffness of the potential [35]. The cubic simulation box is used as outlined in Figure 2. The length of the simulation box is adapted to the dextran size, which is taken as 18 nm, 38 nm and 77 nm for D5, D50 and D400 dextran molecules, respectively (see the Table in Section 3). This allowed the number of particles to range between 50 and 200 for all the dextran sizes. Periodic Boundary Conditions (PBC) are applied in all directions. The Mean Square Displacement of the particles is calculated after a thermalisation time of 10 ns to avoid the effect of unrealistic particle overlapping due to the random initial configuration. All the performed simulations are 1000 ns long with a time step of 0.1 ns. The system temperature is 298.15 K. For each studied system, a minimum of 400 different BD realizations are averaged over. D^{long} values have been calculated by fitting the Einstein–Smoluchowski equation to the simulated mean square displacement ($\langle r^2 \rangle$) in the long-time regime.

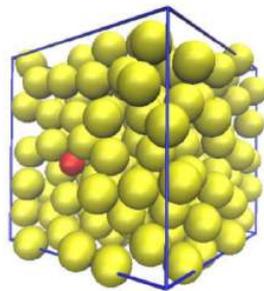


Figure 2. Snapshot of one of the performed dynamics. A protein (in red) diffuses among dextran molecules (in yellow) which act as crowding agents.

HI are included in the BD equation of motion using the Tokuyama model [19]. In this approach, a Fokker–Planck equation for the single-particle distribution function is proposed which is coupled with the Navier–Stokes equation. This equation is analytically solved at short times considering a stationary configuration of equal-sized Brownian particles. Correlation effects and direct interactions are neglected. As a result, the diffusion coefficient of the particle at short times (D^{short}) is found to be a function of the volume fraction ϕ , which has been calculated using [5,36–38]

$$\phi = \frac{4\pi(R_T^3 + N_O R_O^3)}{3L^3}, \quad (5)$$

where R_T is the radius of the tracer protein, and R_O is the radius of the dextran obstacles, N_O is the number of dextran obstacles and L is the side of the simulation box. Following the Tokuyama approach, the short time diffusion coefficient is related to the volume fraction by the equation

$$D^{\text{short}}(\phi) = \frac{D_0}{[1 + H(\phi)]}, \quad (6)$$

where D_0 is the diffusion coefficient at dilute solution and $H(\phi)$ is the contribution due to the HI valid in the short-time regime. $H(\phi)$ is given by

$$H(\phi) = \frac{2b^2}{1-b} - \frac{c}{1+2c} - \frac{bc(2+c)}{(1+c)(1-b+c)}, \quad (7)$$

where $b = \sqrt{\frac{9}{8}\phi}$ and $c = \frac{11}{16}\phi$. Recently, Tokuyama et al. [20] have obtained good agreement for the diffusion coefficient at long times (D^{long}) between Equation (6) and Molecular Dynamics with explicit

solvent. It is also worth mentioning that Tokuyama [21] has proposed an analytical expression of the D^{long} in terms of the volume fraction:

$$D^{\text{long}}(\phi) = \frac{D^{\text{short}}(\phi)}{\left[1 + \kappa \frac{D^{\text{short}}(\phi)}{D_0} \left(\frac{\phi}{\phi_c}\right) \left(1 - \frac{\phi}{\phi_c}\right)^{-2}\right]}, \quad (8)$$

where κ and ϕ_c are parameters settled to $\kappa = 2.0$ and $\phi_c = 1.09$. In this work, simulations with and without HI are compared. In the simulations without HI, the diffusion coefficient D in Equation (3) corresponds to the diffusion coefficient in dilute solution D_0 . HI are included by means of the Tokuyama method [20] replacing D in Equation (3) by D^{short} calculated using Equations (6) and (7).

Tokuyama equations for D^{short} and D^{long} are derived assuming equal-sized spheres. Since our systems contain two different-sized particles, the tracer protein and the dextran obstacles, the use of Equations (6) and (7) for D^{short} involves an approximation. D^{long} , however, is calculated using direct simulated mean square displacements at long times, so that the difference size effect is included in a straightforward way via the inter-particle maximum approach R_{ij} in Equation (4). The resulting D^{long} values are then compared to those predicted by Equation (8) in Section 2.3.

2.2. Effect of the Interaction Potential

The effect of the interaction potential stiffness (Equation (4)) has been also analysed to ensure that the selected constant ($k = 50,000 \text{ J}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}$) was enough to prevent particles to overlap. With this aim, preliminary computations were performed which are shown in Figure 3. No significant difference was observed between the three proposed values for k ($k = 1 \times 10^5$, $k = 5 \times 10^4$ and $k = 2.5 \times 10^4 \text{ J}\cdot\text{mol}^{-1}\cdot\text{nm}^{-2}$), which means that the potential was rigid enough to avoid overlapping.

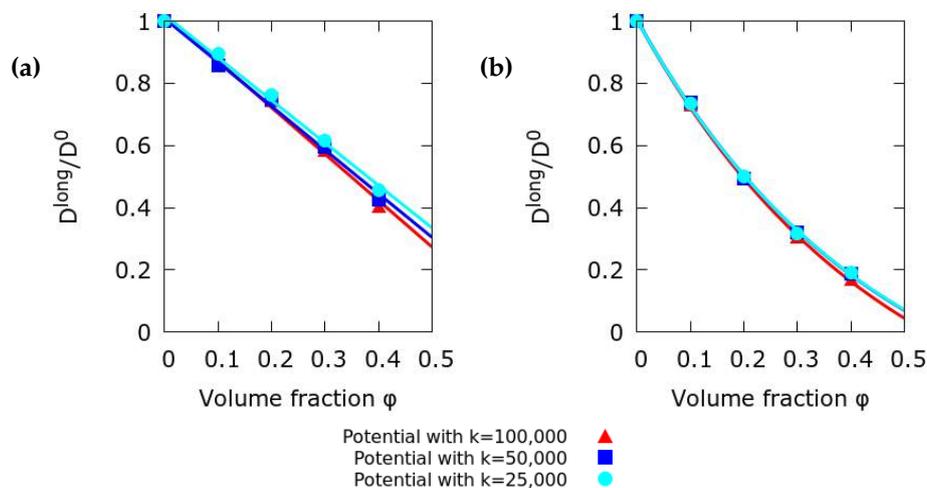


Figure 3. Decrease of the normalised diffusion coefficient at long times (D^{long}) with the volume fraction at three different values of the stiffness constant k of the interaction potential. BD simulation performed: (a) without HI; (b) with HI using the Tokuyama model. The results show that the potential is rigid enough to avoid overlapping. Continuous lines are only to guide the lecturer.

2.3. Effect of Long Range Hydrodynamic Interactions

The effect of the HI in macromolecular diffusion in crowded media has been evaluated by selecting a simple system of equal sized spheres at different volume fractions. We have performed calculations comparing the conventional RPY tensor [17,18] and Tokuyama method [19–21].

Long range HI experience a slow decay with inter-particle distances, a well-known effect in electrostatic interactions. This fact could lead to size-dependent results and to unphysical non-definite positive diffusion tensors [39]. Therefore, in principle, procedures like Ewald sums [40] should be implemented. In order to avoid this problem, we performed simulations at different system sizes. We found that for $N > 50$, the finite size box effect becomes rather small [41]. It was also checked that the diffusion tensor was always definite positive at every time step. This agrees with previous studies [28,29] indicating that long range HI become considerably screened at high volume fractions.

The obtained values for D^{long} without HI, with HI described by RPY diffusion tensor, and with HI using the Tokuyama method are shown in Figure 4. We can observe that the simulations without HI and with HI using the RPY diffusion tensor provide almost identical D^{long} values. This fact supports previously reported results [28,29] highlighting the long range HI screening in crowded media. Conversely, significant differences arise in implementing the Tokuyama method, which account for short range HI (usually called lubrication forces). A marked decay of the diffusion coefficient is observed, in agreement with the relevance of lubrication forces reported in previous works. It is also important to highlight the almost perfect agreement, for a system of equal-sized particles, between the Tokuyama D^{long} analytical prediction (Equation (8)) and BD simulations using the Tokuyama model, as Tokuyama previously reported [20].

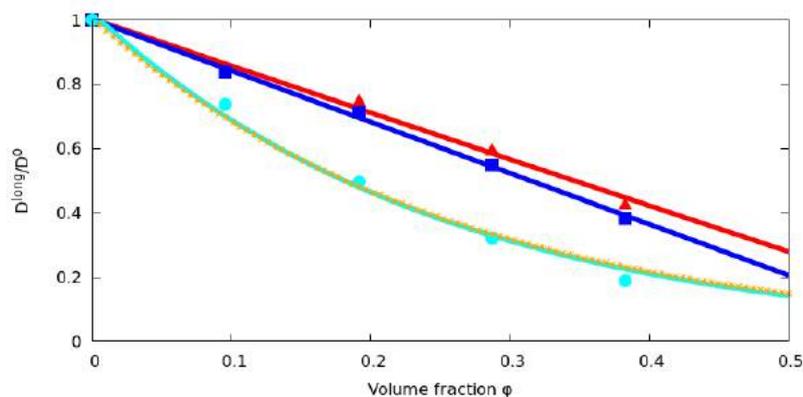


Figure 4. Comparison of the diffusion coefficient calculated using BD without HI (red triangles), BD with HI using the conventional RPY diffusion tensor (blue squares), BD with HI using the Tokuyama model (cyan circles) and Tokuyama analytical expression (Equation (8), orange discontinuous line). The results show the important role of short range HI (lubrication forces) in the diffusion reduction of the diffusion coefficient, while the effect of the of long range HI is very low. It is also worth noting the good agreement between the Tokuyama analytical prediction of D^{long} and the BD simulation with HI using Tokuyama model for D^{short} . The lines are to guide the lecturer.

2.4. Effect of the Difference in Size between Tracer and Obstacles

Tokuyama analytical expressions were derived assuming the same size for all of the particles. However, in our experimental systems, the tracer particle (a protein) and the obstacles (dextran macromolecules) present different sizes. In order to check the accuracy of the Tokuyama prediction for D^{long} , we have performed simulations that implement the difference in size by a suitable modification of the minimum inter-particle distance R_{ij} in the interaction potential (Equation (4)). The Tokuyama effective diffusion coefficient D^{short} is still used to describe the short range HI forces since, to our knowledge, no expression for D^{short} is available yet for different sized particles.

D^{long} resulting from BD simulations and the ones calculated using the Tokuyama analytical expression (Equation (8)) versus volume fraction are depicted in Figure 5. It can be observed that when the tracer (with radius $R = 2.33$ nm) and obstacle ($R = 2.9$ nm) present similar sizes, very good agreement between Tokuyama equations and simulations is found, as shown in Figure 5a. In contrast, if the tracer particle is far bigger ($R = 4.90$ nm) than the obstacles ($R = 1.2$ nm), important

differences arise. As shown in Figure 5b, D^{long} values predicted using the Tokuyama analytic method clearly overestimate the ones obtained by BD computations, with differences up to 60% between both approaches. In conclusion, Tokuyama analytic equations provide a reliable prediction for D^{long} only for systems of similarly-sized spheres so that BD computations are unavoidable otherwise.

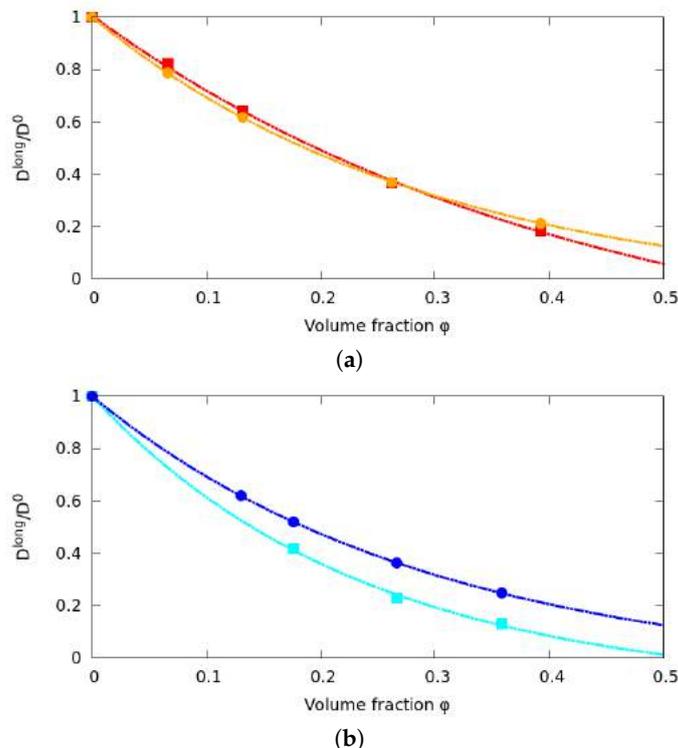


Figure 5. Comparison between the normalized D^{long} obtained using Tokuyama analytic equations (Equations (6)–(8)) and the ones resulting from BD simulations at different volume fractions. In (a), the tracer particle (with radius $R = 2.33$ nm) and the obstacles ($R = 2.9$ nm) present a similar size. D^{long} values obtained by Tokuyama method (orange circles) and the ones obtained by BD simulations (red squares) clearly agree. In contrast, in (b), the tracer particle ($R = 4.90$ nm) is far bigger than the obstacles ($R = 1.2$ nm). In this case, the Tokuyama method (blue circles) clearly overestimates D^{long} values obtained by BD simulations (cyan squares).

2.5. Dextran Model

Here, we assume that each macromolecule is a single sphere. In similar coarse-grained models [12,31], the radius of the particles is calculated using their dilute solution diffusion coefficient via the Stokes–Einstein equation. In these approaches, the dilute solution diffusion coefficient is estimated using HYDROPRO software (Version 10) [42]. HYDROPRO needs as input the atomic coordinates of the macromolecule, which is usually obtained from crystallographic data or NMR spectra. HYDROPRO uses this atomistic structure to compute its hydrodynamic properties such as the diffusion coefficient in dilute solution. Unfortunately, this information is not available for dextran macromolecules yet.

However, experimental hydrodynamic radius (R_H) obtained using quasi-elastic light scattering [43], and radius of gyration (R_G) obtained from capillary viscometry [44] are available. We have fitted R_H and R_G versus molecular weight data to the power law:

$$R_x = K \cdot M_w^\gamma, \quad (9)$$

where R_x can be either R_G or R_H . K and γ are the best fitted parameters reported in Table 1. Unexpectedly, we have found that these experimental radii lead to high volume fractions at experimental concentrations (50–300 g/L). This behaviour increases dramatically with dextran

size, resulting in unphysical volume fractions larger than one for the biggest sizes. This can be explained taking into account the very complex branched structure of dextran macromolecules [45,46]. As dextran macromolecular size increases, the molecular structure becomes more branched. Moreover, the structure of dextran is flexible allowing steric compression in concentrated solution.

Table 1. Best fitted parameters of power law (9) to the experimental radius of gyration and the hydrodynamic radius versus molecular weight.

	R_G	R_H	R_c	R_{eff}
K (nm·Da $^{-\gamma}$)	0.018 ± 0.001	0.043 ± 0.002	0.063	0.045
γ	0.544 ± 0.005	0.445 ± 0.004	$\frac{1}{3}$	0.387

An alternative to calculating the volume fraction at high dextran concentrations is the use of an average dextran specific volume ($\nu = 0.625$ cm 3 /g) [47]. Considering the dextran molecules as compact spheres with a radius R_c :

$$R_c = \sqrt[3]{\frac{3\nu}{4\pi N_A} M_W^{\frac{1}{3}}} = K \cdot M_w^{\gamma}, \quad (10)$$

where N_A is the Avogadro number, $\gamma = \frac{1}{3}$ and $K = \sqrt[3]{\frac{3\nu}{4\pi N_A}} = 0.063$ nm·Da $^{-\frac{1}{3}}$. However, previous studies using Equation (10) showed bad agreement with the experimental diffusion coefficients [5–7] obtained using FRAP and FCS. More specifically, the obtained results exhibit a slower decay of the diffusion coefficient with dextran concentration than that experimentally observed. This was caused by the neglect of solvation effects resulting in volume fractions that were too low.

In summary, a suitable effective radius (R_{eff}) lying in between the hydrodynamic and the compact radius is necessary for dextran macromolecules at high concentration. Such an effective radius has been determined by performing BD simulations scanning the physically meaningful values for K and γ parameters. In these simulations, K ranged from 0.043 to 0.063 nm·Da $^{-\gamma}$ while γ ranged from $\frac{1}{3}$ to 0.445. Among the performed calculations, the best agreement between experimental [5–7] and computed D^{long} was obtained with $\gamma = 0.387$ and $K = 0.045$ nm·Da $^{-\gamma}$.

It is also worth noting that the chosen parameters for R_{eff} are closer to those of R_c than to those of R_H (Figure 6). This means that steric compression is crucial to properly describe dextran size in crowded media. Moreover, the large difference between R_H and R_c at high dextran molecular weight reveals the importance of the steric compression as dextran size increases.

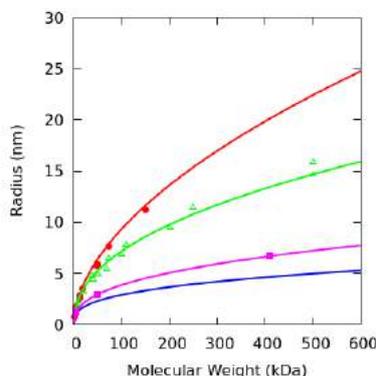


Figure 6. Radius of gyration (R_G , red circles) taken from [44] and hydrodynamic radius (R_H , green triangles) taken from [43] versus molecular weight. The fittings corresponding to the power law (Equation (9)) are represented with red and green lines. The compact radius (R_c , blue) and the effective radius (R_{eff} , purple line) versus molecular weight are also depicted. The effective radius for the chosen dextran obstacles (purple squares) in our computations is also plotted. The lines are plotted only to guide the lecturer.

3. Results and Discussion

Two different experimental systems have been selected and computationally modeled in the present work. The first one is the FCS study of Streptavidin protein ($R_H = 4.90$ nm) diffusion at different concentrations of six distinct dextran macromolecules performed by Banks et al. [5]. The second one uses FRAP to record α -Chymotrypsin protein ($R_H = 2.33$ nm) diffusion [6,7] at different concentrations of three different-sized dextran macromolecules acting as crowding agents. We have chosen the dextran molecules D5, D50 and D400 as obstacles since they have been used in both experimental studies, allowing better comparison between them. The characteristic parameters of the used dextran macromolecules are shown in Table 2. In FCS, the fluorescence fluctuations of a sample of tracer proteins are associated with the tracer auto-correlation function [5]. This allows for experimentally measuring the anomalous exponent α and a characteristic residence time τ_D . τ_D is defined as the time which the tracer particle needs to traverse the characteristic length of the detection volume. A similar approach is followed in the analysis of the fluorescence recovery measured by FRAP [6,7]. α and τ_D are used to calculate the effective diffusion coefficient (D^{eff}) at a time equal to τ_D . The experimentally found τ_D , around 1–3 ms [5,6], is much longer than the computational times at which D^{long} is calculated. As a consequence, the computed values of D^{long} should reasonably correspond to the experimentally obtained D^{eff} .

Table 2. Characteristic parameters of the dextran molecules chosen as obstacles. M_W : averaged molecular mass in weight; R_G : radius of gyration [5–7]; R_H : hydrodynamic radius [43]; R_c : compact spheres radius; R_{eff} : the chosen effective radius. They have been obtained using power law (9) with the parameters shown in Table 1.

Dextran	M_W (kDa)	R_G (nm)	R_H (nm)	R_c (nm)	R_{eff} (nm)
D5	5.2	1.7	1.9	1.1	1.2
D50	48.6	5.8	5.2	2.3	2.9
D400	409.8	17	13.5	4.7	6.7

3.1. Long Time Diffusion Coefficient

As mentioned above, diffusion in crowded media has three different temporal regimes over time (Figure 1). In order to compare computations and experiments, it is necessary to calculate the diffusion coefficient in the long time regime (D^{long}). In the selected experimental systems, the diffusion of a tracer protein is studied at different concentrations of three different-sized dextran macromolecules (D5, D50 and D400), which act as inert obstacles. Two different approaches have been applied in the simulations. In the first one, the HI are not considered and thus only a pairwise repulsive harmonic potential is acting. In the second one, HI are taken into account using the Tokuyama method.

Figures 7 and 8 show the effect of macromolecular crowding in the diffusion of Streptavidin and α -Chymotrypsin, respectively. D^{long} is normalized using the diffusion coefficient at infinite dilution (D^0) of the particle. The simulations including HI using the Tokuyama model exhibit, in general, better qualitative and quantitative agreement with the experimental data for all the dextran sizes studied. As a consequence, the inclusion of HI is crucial to describe the reduction of D^{long} experimentally observed. This is in concordance with the importance of the short range HI in crowded media previously reported in the literature [29]. It is also worth mentioning that D^{long} always decays as obstacle concentration increases, as a result of the increase of particle collisions, which hinder protein diffusion.

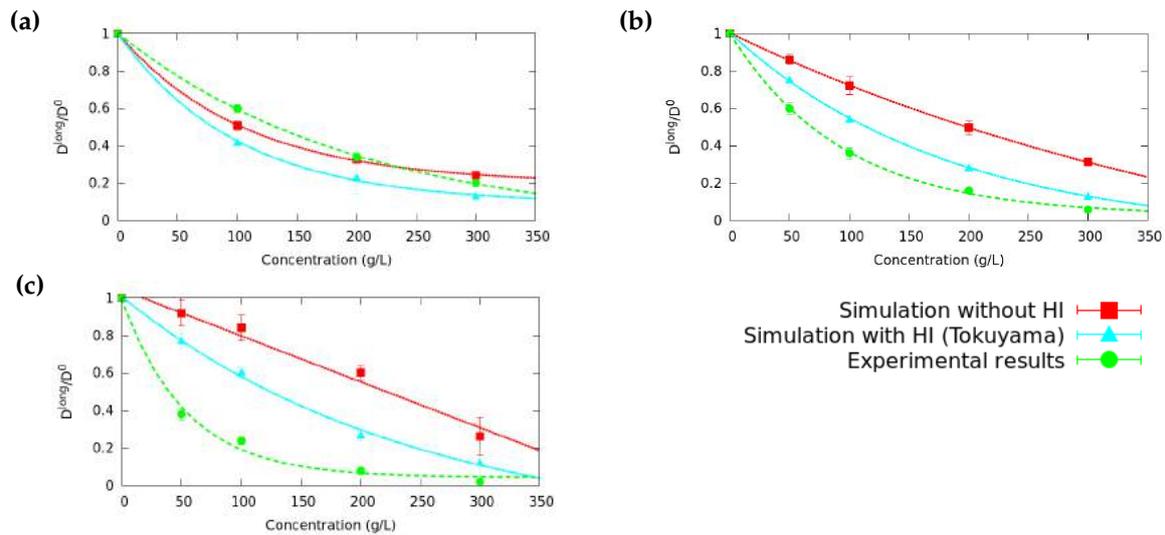


Figure 7. Decay of D^{long} corresponding to Streptavidin protein versus dextran concentration. Three different sizes of dextran obstacles: (a) D5; (b) D50; and (c) D400 have been used. The obtained results including the HI with Tokuyama model exhibit better agreement in general with the experimental data showing the relevance of HI in the D^{long} values in crowded media. The lines are only for guiding the lecturer.

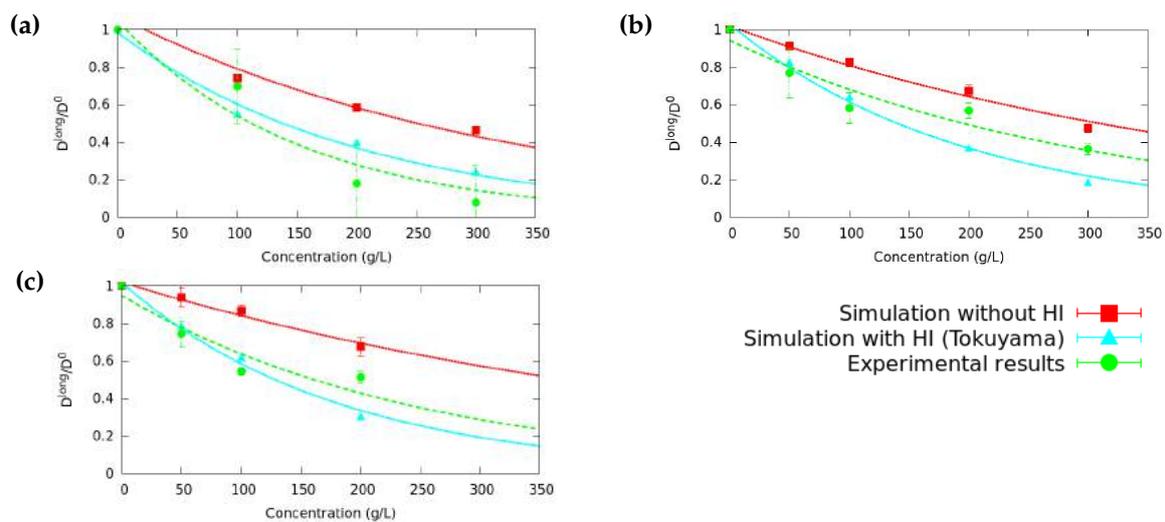


Figure 8. α -Chymiotrypsin protein D^{long} decay with dextran concentration. Three different sizes of dextran obstacles: D5 (a); D50 (b); and D400 (c) have been used. The obtained results including HI with Tokuyama model exhibit in general better agreement with the experimental data showing the relevance of HI in the D^{long} values in crowded media. Continuous lines are only to guide the lecturer.

BD simulations show a clear effect of the tracer particle size, which is different in both systems. Streptavidin shows a faster decay with dextran concentration than α -Chymiotrypsin. This is logical since Streptavidin (with $R = 4.9$ nm) is quite larger than α -Chymiotrypsin ($R = 2.33$ nm). As a result, the loss of mobility due to an increase in the number of collisions of Streptavidin is larger than in the case of α -Chymiotrypsin. On the other hand, BD simulations show a small contribution of the dextran size in D^{long} for the same dextran concentration. This is probably due to the limitations of Tokuyama approach, which assumes equal sizes for the particles in the derivation of the equations. As a consequence, Tokuyama equations only account for the dependence on the volume fraction, but not on the obstacle size. This is the case of α -Chymiotrypsin, but not that of Streptavidin, for which

D^{long} clearly decays faster as the obstacle size increases. This fact points out the interest of future extensions of the Tokuyama approach to systems of different-sized particles.

3.2. Anomalous Diffusion Exponent

The anomalous diffusion exponents α for the two studied diffusing proteins are depicted in Figures 9 and 10. In concordance with the obtained results for D^{long} , the α exponent also decays with dextran concentration. A good quantitative prediction of the experimental data is obtained for both proteins except for the smallest dextran size.

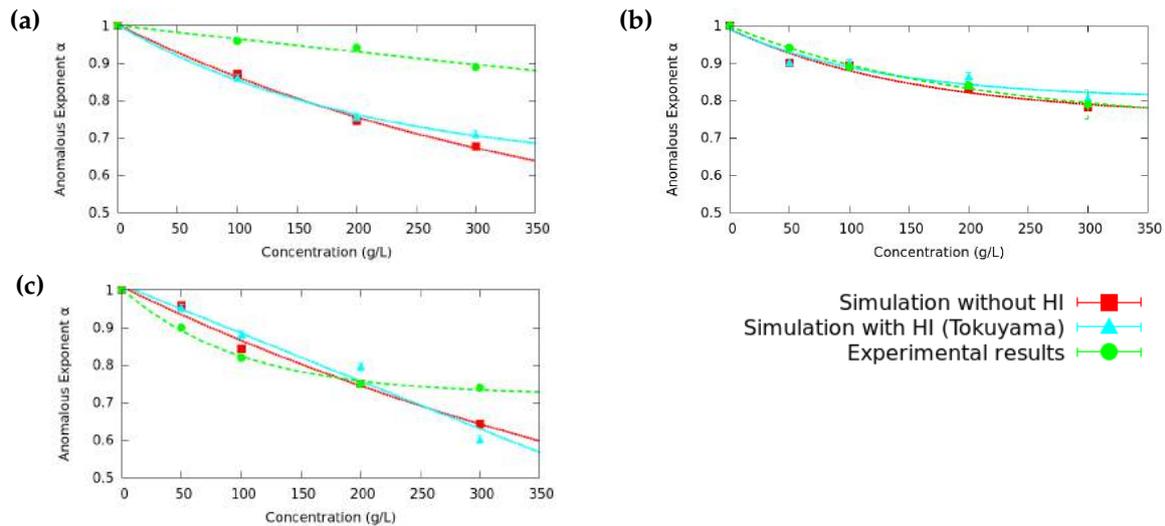


Figure 9. Anomalous diffusion exponent α of Streptavidin versus dextran concentration for three different sizes of dextran obstacles: (a) D5; (b) D50; and (c) D400. The obtained results exhibit good quantitative agreement with the experimental data except for the smallest dextran size D5. In general, no significant influence of HI in the α exponent is obtained. The lines are to guide the lecturer.

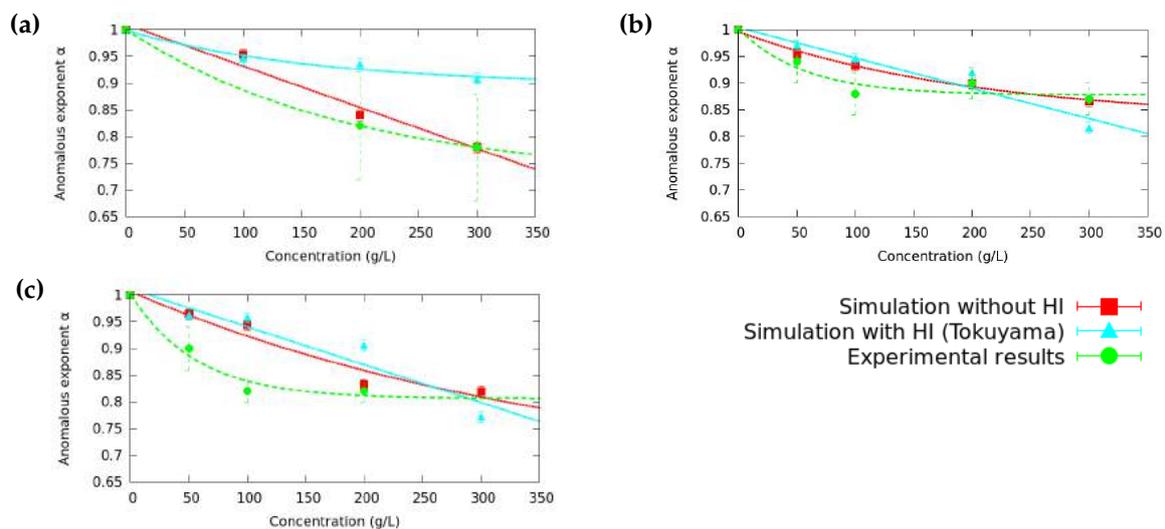


Figure 10. α -Chymotrypsin anomalous diffusion exponent α versus dextran concentration for three different sizes of dextran obstacles: (a) D5; (b) D50; and (c) D400. Computations show good quantitative agreement with experiments except for the smallest dextran size D5. The small difference between the results obtained with and without HI reveals HI are not important in the α exponent value. The lines are only for guiding the lecturer.

In general, the obtained results with and without HI are very similar indicating that, unlike D^{long} , the α exponent is not significantly influenced by HI. Since the anomalous diffusion exponent is related to the transition rate between D^{short} and D^{long} regimes, this means that although HI slow down D^{short} and D^{long} , they do not affect the rate of the transition. One possible explanation is that the collision rate of particles, the driving mechanism of the transition, does not change substantially with the inclusion of HI in the BD simulations of crowded media. For Streptavidin, it is also observed that the anomalous exponent decreases with dextran size. However, the non-monotonous dependence of the anomalous exponent for α -Chymotrypsin is not clear, given the high experimental error for this parameter. Probably, more experiments should be necessary in order to fully clarify this point.

4. Conclusions

A Brownian Dynamics computational model for two experimental systems of diffusing proteins (Streptavidin and α -Chymotrypsin) in crowded media has been proposed. In both cases, dextran macromolecules have been used as crowding agents. Dextran macromolecules have been modeled as spheres with effective radii accounting for macromolecular compression. The obtained long time diffusion coefficients (D^{long}) are, in general, close to the ones experimentally observed. In the analysis of D^{long} , the inclusion of the HI ends up being fundamental to explain the experimental decay. This means that steric effects and HI are the two main contributions to the decrease of D^{long} in crowded media. However, no significant influence of HI in the anomalous diffusion exponent is detected. Although the qualitative behaviour is properly modelled using the Tokuyama approach, in some cases, the obtained results by simulation are still far from the experimental ones. This could be mainly due to two different reasons.

The first one is that the Tokuyama model assumes equal-sized spheres and our systems are heterogeneous. Therefore, this method is less accurate when the difference in size between the tracer protein and obstacles particles increases. A generalization of the Tokuyama model for different spheres is not available yet and it would be a promising procedure to improve our results.

The second factor that can be responsible for the difference between simulation and experiments is the procedure used to model dextran molecules. This method does not take into account that the degree of compression of the dextran macromolecules increases as their concentration increases. This suggests the possibility to improve the dextran description by changing the harmonic potential for a new potential able to allow dextran radius to evolve from its hydrodynamic radius to its compact radius as the concentration increases [48]. This new potential could also account for the entanglement of the polymer branches, allowing some overlap between the macromolecules. This effect could be also relevant in the diffusion properties of macromolecules in crowded media.

Acknowledgments: We acknowledge the financial support from: the Spanish Ministry of Science and Innovation (project CTM2012-39183 and CTM2016-78798-C2-1-P) and Generalitat de Catalunya (Grants 2014SGR1017, 2014SGR1132 and XrQTC). Sergio Madurga and Francesc Mas acknowledge the funding of the project 8SEWP-HORIZON 2020 (692146).

Author Contributions: Pablo M. Blanco, Josep Lluís Garcés, Sergio Madurga and Francesc Mas designed the BD code and the BD simulations. Pablo M. Blanco and Mireia Via developed and tested the computational code and performed the BD simulations. Pablo M. Blanco, Sergio Madurga and Francesc Mas discussed and proposed the new dextran macromolecular model. Pablo M. Blanco, Josep Lluís Garcés, Sergio Madurga and Francesc Mas collaborated to write the present paper. All authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

BD	Brownian Dynamics
FCS	Fluorescence Correlation Spectroscopy
FRAP	Fluorescence Recovery After Photobleaching
HI	Hydrodynamic Interactions
RPY	Rotne–Prager–Yamakawa

References

- Zhou, H.X.; Rivas, G.; Minton, A.P. Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences. *Annu. Rev. Biophys.* **2008**, *37*, 375–397.
- Ellis, R.J. Macromolecular crowding: Obvious but underappreciated. *TRENDS BioChem. Sci.* **2001**, *26*, 597–604.
- Pastor, I.; Pitulice, L.; Balcells, C.; Vilaseca, E.; Madurga, S.; Isvoran, A.; Mas, F. Effect of crowding by Dextrans in enzymatic reactions. *Biophys. Chem.* **2014**, *185*, 8–13.
- Balcells, C.; Pastor, I.; Pitulice, L.; Hernández, C.; Via, M.; Garcés, J.L.; Madurga, S.; Vilaseca, E.; Isvoran, A.; Cascante, M.; et al. Macromolecular Crowding upon in-vivo-Like Enzyme-Kinetics: Effect of Enzyme-Obstacle Size Ratio. *New J. Chem.* **2015**, *24*, 3–16.
- Banks, D.S.; Fradin, C. Anomalous Diffusion of Proteins Due to Molecular Crowding. *Biophys. J.* **2005**, *89*, 2960–2971.
- Pastor, I.; Vilaseca, E.; Madurga, S.; Garcés, J.L.; Cascante, M.; Mas, F. Diffusion of α -Chymotrypsin in Solution-Crowded Media. A Fluorescence Recovery After Photobleaching Study. *J. Phys. Chem.* **2010**, *114*, 4028–4034.
- Vilaseca, E.; Pastor, I.; Isvoran, A.; Madurga, S.; Garcés, J.L.; Mas, F. Diffusion in macromolecular crowded media: Monte Carlo simulation of obstructed diffusion vs. FRAP experiments. *Theor. Chem. Acc.* **2011**, *128*, 795–805.
- Elcock, A.H. Models of macromolecular crowding effects and the need of quantitative comparisons with experiment. *Curr. Opin. Struct. Biol.* **2010**, *20*, 196–206.
- Vilaseca, E.; Isvoran, A.; Madurga, S.; Pastor, I.; Garcés, J.L.; Mas, F. New insights into diffusion in 3D crowded media by Monte Carlo simulations: Effect of size, mobility and spatial distribution of obstacles. *Phys. Chem. Chem. Phys.* **2011**, *13*, 7396–7407.
- Hasnain, S.; McClendon, C.L.; Hsu, M.T.; Jacobson, M.P.; Bandyopadhyay, P. A New Coarse-Grained Model for E. coli Cytoplasm: Accurate Calculation of the Diffusion Coefficient of Proteins and Observation of Anomalous Diffusion. *PLoS ONE* **2014**, *9*, doi:10.1371/journal.pone.0106466.
- Gomez, D.; Klumpp, S. Biochemical reactions in crowded environments: Revisiting the effects of volume exclusion with simulations. *Front. Phys.* **2015**, doi:10.3389/fphy.2015.00045.
- Yu, I.; Mori, T.; Ando, T.; Harada, R.; Jung, J.; Sugita, Y.; Feig, M. Biomolecular interactions modulate macromolecular structure and dynamics in atomistic model of a bacterial cytoplasm. *eLife* **2016**, *5*, e19274.
- Saxton, M.J. Anomalous Diffusion Due to Obstacles: A Monte Carlo Study. *Biophys. J.* **1994**, *66*, 394–401.
- Bauchaud, J.P.; Georges, A. Anomalous diffusion in disordered media: Statistical mechanics, model and physical application. *Phys. Rep.* **1990**, *185*, 127–293.
- Elcock, A.H. Molecule-Centered Method for Accelerating the Calculation of Hydrodynamic Interactions in Brownian Dynamics Simulations Containing Many Flexible Biomolecules. *J. Chem. Theory Comput.* **2013**, *9*, 3224–3239.
- Ermak, D.L.; McCammon, J.A. Brownian dynamics with hydrodynamic interactions. *J. Chem. Phys.* **1978**, *69*, 1352–1360.
- Rotne, J.; Prager, S. Variational treatment of hydrodynamic interaction in polymers. *J. Chem. Phys.* **1969**, *50*, 4831–4837.
- Yamakawa, H.J. Transport properties of polymer chains in dilute solution: Hydrodynamic interaction. *Chem. Phys.* **1970**, *53*, 436–443.
- Tokuyama, M.; Oppenheim, I. Dynamics of hard-sphere suspensions. *Phys. Rev. E* **1994**, *50*, R16, doi:10.1103/PhysRevE.50.R16.
- Tokuyama, M.; Moriki, T.; Kimura, Y. Self-diffusion of biomolecules in solution. *Phys. Rev. E* **2011**, *83*, 081402.

21. Tokuyama, M. Mean-field theory of glass transitions. *Physica A* **2006**, *364*, 23–62.
22. Rodríguez, R.; Hernández, J.G.; García, J. Comparison of Brownian dynamics algorithms with hydrodynamic interactions. *J. Chem. Phys.* **2011**, *135*, 084116.
23. Fixman, M. Implicit algorithm for Brownian Dynamics of Polymers. *Macromolecules* **1986**, *19*, 1195–1204.
24. Geyer, T.; Winter, U. An $O(N^2)$ approximation for hydrodynamic interactions in Brownian dynamics simulations. *J. Chem. Phys.* **2009**, *130*, 114905.
25. Zuk, P.J.; Wajnryb, E.; Mizerski, K.A.; Szymczak, P. Rotne–Prager–Yamakawa approximation for different-sized particles in application to macromolecular bead model. *J. Fluid. Mech.* **2014**, *741*, R5, doi:10.1017/jfm.2013.668.
26. Durlofsky, L.; Brady, J.F.; Bossis, G. Dynamic simulation of hydrodynamically interacting particles. *J. Fluid Mech.* **1987**, *180*, 21–49.
27. Brady, J.F.; Bossis, G. Stokesian Dynamics. *Annu. Rev. Fluid Mech.* **1988**, *20*, 111–157.
28. Phillips, R.J.; Brady, J.F.; Bossis, G. Hydrodynamic transport properties of hard-sphere dispersions. I. Suspensions of freely mobile particles. *Phys. Fluids* **1988**, *31*, 3462–3472.
29. Ando, T.; Chow, E.; Skolnick, J. Dynamic simulation of concentrated macromolecular solutions with screened long-range hydrodynamic interactions: Algorithm and limitations. *J. Chem. Phys.* **2013**, *139*, 121922.
30. Sun, J.; Weinstein, H. Toward realistic modeling of dynamic processes in cell signaling: Quantification of macromolecular crowding effects. *J. Chem. Phys.* **2007**, *127*, 155105.
31. Mereghetti, P.; Wade, R.C. Atomic Detail Brownian Dynamics Simulations of Concentrated Protein Solutions with a Mean Field Treatment of Hydrodynamic Interactions. *J. Chem. Phys.* **2012**, *116*, 8523–8533.
32. Gillespie, D.T. The multivariate Langevin and Fokker–Planck equations. *Am. J. Phys.* **1996**, *64*, 1246–1257.
33. Mori, H. Transport, Collective Motion and Brownian Motion. *Prog. Theor. Phys.* **1965**, *33*, 423–455.
34. Dhont, J.K.G. Brownian motion of non-interacting particles. In *An Introduction to Dynamics of Colloids*; Elsevier: Amsterdam, The Netherlands, 1996; Chapter 2.
35. Schöneberg, J.; Noé, F. ReaDDy—A Software for Particle-Based Reaction-Diffusion Dynamics in Crowded Cellular Environments. *PLoS ONE* **2013**, *8*, e74261.
36. Roosen-Runge, F.; Hennig, M.; Zhang, F.; Jacobs, R.M.; Sztucki, M.; Schober, H.; Seydel, T.; Schreiber, F. Protein self-diffusion in crowded solutions. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 11815–11820.
37. Qin, S.; Zhou, H.X. Atomistic modeling of macromolecular crowding predicts modest increases in protein folding and binding stability. *Biophys. J.* **2009**, *97*, 12–19.
38. Sastry, S.; Truskett, T.M.; Debenedetti, P.G.; Torquato, S.; Stillinger, F.H. Free volume in the hard sphere liquid. *Mol. Phys.* **1998**, *95*, 289–297.
39. Brady, J.F.; Phillips, R.J.; Lester, J.C.; Bossis, G. Dynamic simulation of hydrodynamically interacting suspensions. *J. Fluid Mech.* **1988**, *195*, 257–280.
40. Beenakker, C.W.J. Ewald sum of the Rotne–Prager tensor. *J. Chem. Phys.* **1986**, *85*, 1581–1582.
41. Ladd, A.J. Hydrodynamic transport coefficients of random dispersions of hard spheres. *J. Chem. Phys.* **1990**, *93*, 3484–3494.
42. Ortega, A.; Amorós, D.; de la Torre, J.G. Prediction of hydrodynamic and other solution properties of rigid proteins from atomic-and residue-level models. *Biophys. J.* **2011**, *101*, 892–898.
43. Armstrong, J.K.; Wenby, B.; Meiselman, H.J.; Fisher, T.C. The Hydrodynamic Radii of Macromolecules and Their Effect on Red Blood Cell Aggregation. *Biophys. J.* **2004**, *87*, 4259–4270.
44. Fundueanu, G.; Nastruzzi, C.; Carpov, A.; Desbrieres, J.; Rinaudo, M. Physico-chemical characterization of Ca-alginate microparticles produced with different methods. *Biomaterials* **1999**, *20*, 1427–1435.
45. Sloan, J.W.; Alexander, B.H.; Lohmar, R.L.; Wolff, I.A.; Rist, C.E. Determination of dextran structure by periodate oxidation techniques. *J. Am. Chem. Soc.* **1954**, *76*, 4429–4434.
46. Sabatie, J.; Choplin, L.; Moan, M.; Doublie, J.L.; Paul, F.; Monsan, P. The effect of synthesis temperature on the structure of dextran NRRL B 512F. *Carbohydr. Polym.* **1988**, *9*, 87–101.
47. Schaefer, D.W. A unified model for the structure of polymers in semidilute solution. *Polymer* **1984**, *25*, 387–394.
48. Vilaseca, P.; Franzese, G. Softness dependence of the anomalies for the continuous shouldered well potential. *J. Chem. Phys.* **2010**, *133*, 084507.



Chapter 7

Macromolecular diffusion in crowded media beyond the hard-sphere model

7.1 Summary

7.1.1 Introduction

In the previous chapter, protein diffusion in a media crowded by dextran macromolecules have been studied by means of Brownian Dynamics (BD) simulations. The macromolecules were modeled as single hard spheres with an effective radius. The use of an effective radius was necessary since the hydrodynamic radius and the radius of gyration obtained in dilute solution were found to yield to unphysical excluded volume fractions at high concentrations. In using the hard sphere model, the long-time diffusion coefficient of the proteins obtained with BD simulations have been found in qualitative agreement with the corresponding experimental measures.⁹⁷⁻⁹⁹ However, in some cases the results were not in quantitative agreement. This discrepancy suggests that coarse-grained models beyond the hard-sphere, able to describe macromolecular flexibility, could be necessary to describe macromolecular diffusion in crowded media.²²⁴

The dynamic conformational behavior of macromolecules is well reproduced with atomistically-detailed models. However, solutions with a high number of macromolecules are computationally very demanding. Moreover, it is necessary to reach the microsecond time scale to measure the long-time diffusion coefficient. For these reasons, the computational study of macromolecular diffusion in crowding conditions have been mainly restricted to coarse-grained models, which often rely on the single hard-sphere description.^{88,105,116,120,200}

Effective softened potentials, accounting for macromolecular flexibility, have been previously proposed for colloidal systems.²²⁵⁻²³¹ These systems, mainly star-polymers and polymer micelles, exhibit a structure composed of a dense core and a disperse corona. Coarse-grained simulations of pluronic polymeric micelles have proved that systems exhibiting both repulsive and attractive interactions have an inter-particle potential of mean force which is purely repulsive.²²⁶ Star polymers have been modeled employing the Derjaguin approximation, yielding to an effective potential with a soft repulsion between the macromolecular branches and hard repulsion between the macromolecules dense cores.²³²⁻²³⁴

In this chapter, a new coarse-grained model is presented, named Chain Entanglement Softened Potential (CESP), relying in a softened inter-particle potential. The CESP model proposes to describe the macromolecule as a double-shelled sphere, as outlined in Ref. 7.1. The outer shell is soft and it represents the outer, less dense region of the macromolecule. The CESP model allows overlapping in this soft region, which represents the macromolecular entanglement. The inner shell is hard-core and it represents the dense core of the macromolecules. The CESP model has only one parameter to be fitted, U_r , which is associated with the entanglement energetic

cost. The CESP model is implemented in Brownian Dynamics (BD) simulations. As a reference system, the simulations are parametrized to resemble the *in vitro* experiments of streptavidin on systems crowded with dextran macromolecules. A new empiric equation is proposed, able to describe the temporal evolution of the diffusion coefficient, which is used to measure the short-time diffusion coefficient (D^{short}), the long-time diffusion coefficient (D^{long}) and the anomalous exponent α .

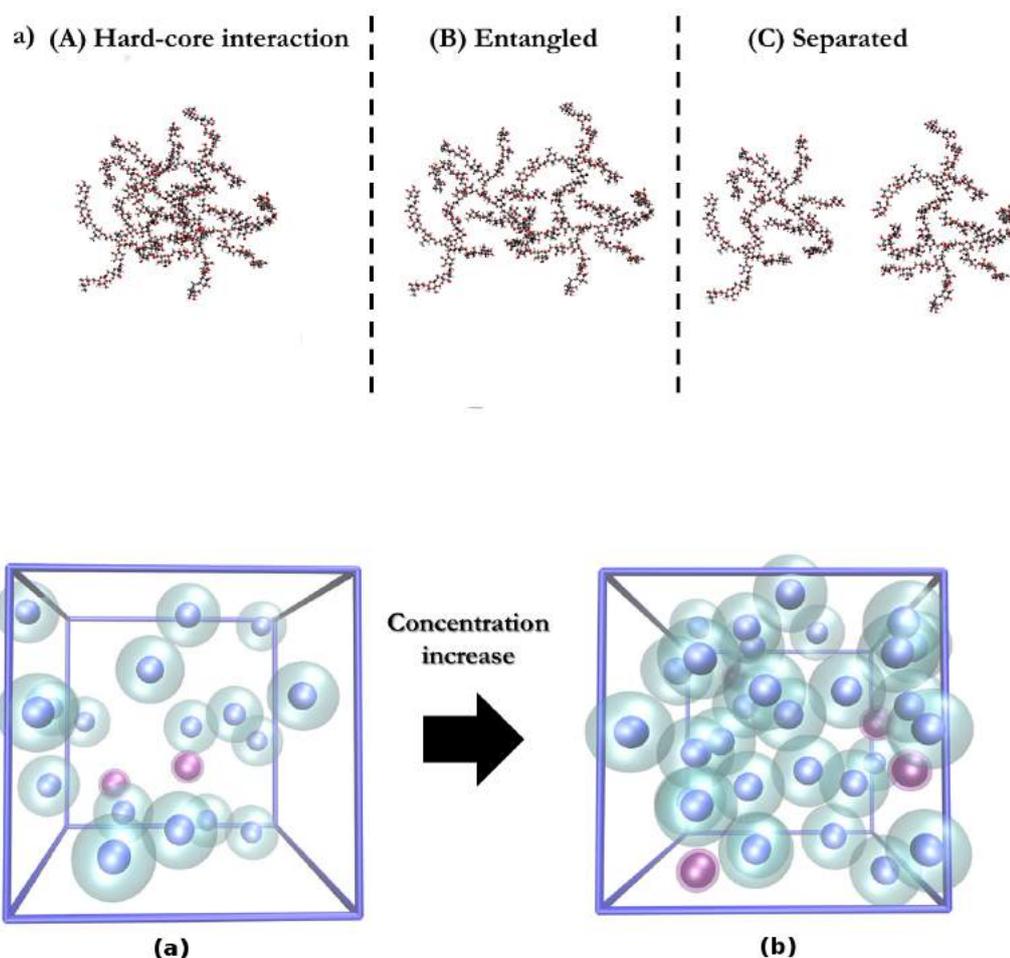


FIGURE 7.1: (up) Scheme showing the three situations contemplated by the CESP model: (A) Hard-core interaction between the dense core of the macromolecules, (B) soft interactions between the macromolecular branches and (C) absence of interaction when the macromolecular are far enough. (down) Two snapshots of Brownian Dynamics (BD) simulations of streptavidin (purple spheres) diffusing between dextran obstacles (blue spheres). The macromolecules are represented as double-shelled spheres with an outer soft region (transparent) and an inner hard region. Dextran concentration is of (left) 25 g/L and (right) 100 g/L.

7.1.2 Results

In this chapter, the principal objective is to develop a model able to reproduce the experiments of streptavidin diffusion in a media crowded with dextran macromolecules.⁹⁷ Here, the diffusion of streptavidin in a media with four different sizes of dextran crowding agents is analyzed. The dextran of study have molecular weights of 11.6 kg mol^{-1} , 48.6 kg mol^{-1} , $409.8 \text{ kg mol}^{-1}$, $667.8 \text{ kg mol}^{-1}$. From now on, these dextrans will be referred as D10, D50, D400 and D700, respectively. Aiming to improve the hard sphere model used in the previous chapter, a new inter-particle softened potential, named as Chain Entanglement Softened Potential (CESP). The CESP model describes the macromolecules as double-shelled spheres with an outer soft shell, with a radius equal to the macromolecule hydrodynamic radius R_H , and an inner hard core shell with a radius R_C estimated from the macromolecule specific volume. The toughness of the soft shell can be tuned varying the entanglement energetic cost U_r , which is the only parameter to be fit in CESP model. The U_r value have been settled fitting the D^{long} obtained from BD simulations to diffusion coefficient obtained from Fluorescence Correlation Spectroscopy (FCS) experiments of streptavidin diffusion using D50 dextran as crowding agent.⁹⁷ A more detailed description of the CESP model, its parametrization and the BD simulations can be found in the publication enclosed at the end of the chapter.

In the previous chapter, the three different temporal regimes of the diffusion coefficient D were presented. These regimes can be appreciated in Fig. 7.2, where the temporal evolution of the D of streptavidin diffusing in a media with a D50 concentration of 200 g L^{-1} is shown (red continuous line). It can be observed that D undergoes a transition from a constant value at short times, D^{short} , to a smaller constant value at long times D^{long} . The rate of the transition

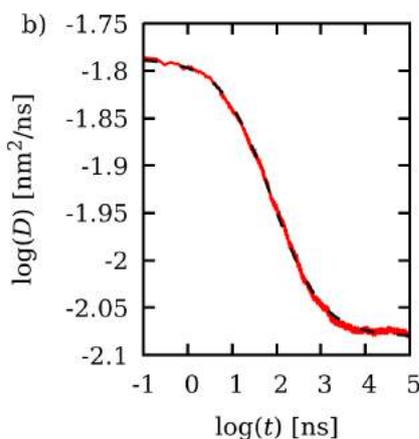


FIGURE 7.2: Streptavidin diffusion coefficient D vs. time t curve obtained from BD simulations of streptavidin diffusion in media with a concentration of D50 crowders of 200 g L^{-1} (red continuous line). The best fit of the computational results to Eq. 10 (See publication enclosed to this chapter) is shown as a black dashed line.

between these regimes is given by the anomalous exponent α . In this chapter, a new empiric equation is proposed, which describes D in the full time regime (Eq. 10 in the publication at the end of the chapter). This equation allows to readily obtain D^{short} , α and D^{long} by fitting the D vs. time curves. The equation have been found to fit very well to the temporal evolution of D obtained from all the cases under study, which cover a wide range of dextran sizes and concentrations. An example can be found in Fig. 7.2, where the best fit of the BD computational data to the empiric equation is shown as a dashed black line.

The D^{long} and α values obtained with BD simulations can be found in Figs. 7.3a and 7.3b, respectively, as a function of the dextran concentration. The results obtained from BD simulations (Filled markers following the dashed lines) are compared with the experimental diffusion coefficient measured from FCS taken from Ref. ⁹⁷ (empty markers). Four different dextran are used as obstacles D10 (green squares), D50 (red circles), D400 (cyan upwards triangles) and D700

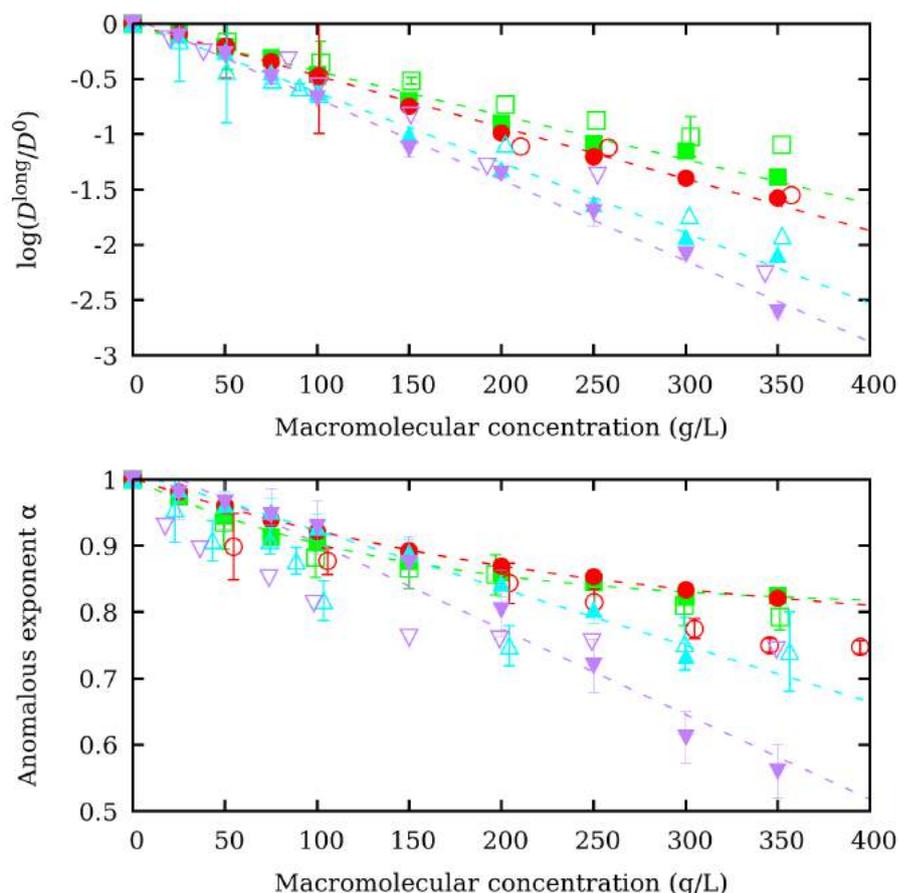


FIGURE 7.3: (up) D^{long}/D^0 and (down) anomalous exponent α as a function of the dextran concentration obtained with BD simulations (filled markers following the dashed lines) and with FCS experiments ⁹⁷ (empty markers). For different dextran sizes are shown: D700 (downwards purple triangles), D400 (upwards blue triangles), D50 (red circles) and D10 (green squares). Lines follow the simulation results and are only to guide the eye.

(purple downwards triangles). The D^{long} values are again normalized to D^0 . Regarding D^{long} , it can be observed a very good quantitative agreement between the BD and the experimental results. Regarding α , a qualitative agreement is observed between the results obtained from BD simulations and FCS experiments. This discrepancies could be due to the different ways of obtaining α , since experimentally it is obtained fitting the autocorrelation function of the protein fluorescence to Eq. 1.58. It can be also observed that the D^{long} and α values, at a given macromolecular concentration, decrease when the dextran size is increased. This is because, at a given crowder concentration, bigger crowders occupy larger volumes than smaller ones thus impeding more the tracer diffusion.

The degree of entanglement of the macromolecules is assessed by computing the radial distribution function (rdf) of the dextran crowders. The rdf have been computed from the BD simulations averaging the center-to-center distance r of a dextran with respect the other dextrans in solution. In Fig. 7.4, the rdf obtained from systems with D50 crowders is shown for D50 concentration ranging from 350 g L^{-1} (top) to 25 g L^{-1} (bottom). Two characteristic distances are highlighted, which correspond to the sum of the hydrodynamic and compact radii of two D50 dextran, $d_e = 2R_H(\text{D50})$ and $d_c = 2R_c(\text{D50})$, respectively. The formation of two coordination shells of entangled macromolecules is observed, as indicated by the presence of two peaks at distances slightly smaller than d_e and d_c . The intensity of the peaks increase in increasing the macromolecular concentration, indicating an increase of the number of macromolecules entangled. The number of entangled macromolecules exponentially increases with the dextran concentration, as can be assessed in the publication at the end of the chapter. Note that the macromolecules become entangled even if the only potential acting is CESP potential, which in this case is purely repulsive. This result indicates that entanglement is necessary to

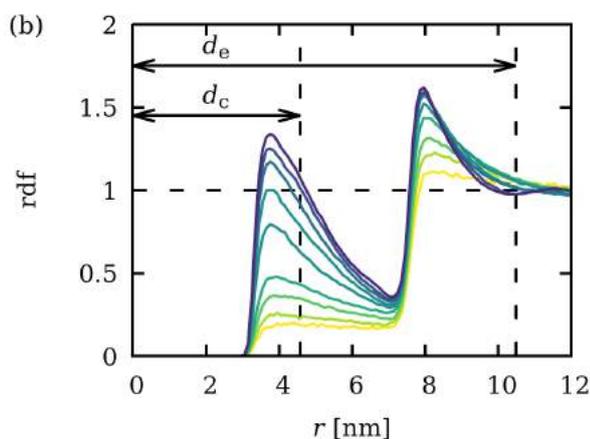


FIGURE 7.4: D50-D50 dextran radial distribution function as a function of the dextran center-to-center distance r obtained from BD simulations using the CESP model. Eight different concentrations are shown, ranging from 350 g L^{-1} (purple) to 25 g L^{-1} (yellow), ordered from top to bottom.

minimize volume exclusion effect, even if at the cost of a energetic penalty. In the publication enclosed in this chapter, the rfd computer for the other dextran sizes can be found.

7.1.3 Conclusions

In this chapter, a new coarse grained model is proposed, named as Chain Entanglement Softened Potential (CESP) model. The CESP model goes one step beyond the hard-sphere model, aiming to effectively describe macromolecular flexibility. The macromolecules are described as double-shelled spheres with an outer soft shell and an inner hard-core shell, representing the branched and dense regions of the macromolecules, respectively. The model only has one energetic parameter, U_r , which determined the outer shell toughness. The CESP model is integrated in Brownian Dynamics (BD) simulations. As a system of study, it is studied the diffusion of streptavidin in media crowded with dextran macromolecules of different concentrations and sized. A new empiric equation is proposed, which allows to readily compute the short-time diffusion coefficient D^{short} , long-time diffusion coefficient D^{long} and anomalous exponent α from the full temporal evolution of the diffusion coefficient obtained from BD simulations. This equation have been found to fit very well the BD data for all the cases studied in this chapter.

The U_r value have been settled fitting the D^{long} obtained from BD simulations to diffusion coefficient obtained from Fluorescence Correlation Spectroscopy (FCS) experiments of streptavidin diffusion using D50 dextran as crowding agent.⁹⁷ Without any further parametrization, a very good agreement is observed between the D^{long} and α values obtained from BD simulations and those reported experimentally, for all the cases with differently-sized dextran obstacles studied here. The results obtained with CESP model clearly improve the ones obtained with the hard-sphere description used in the previous chapter. Radial distribution functions between dextran macromolecules are calculated from the Brownian Dynamics, which show that a significant number of crowder molecules are entangled. This effect becomes more important in increasing the dextran concentration, which exponentially increases the number of entangled molecules. These results indicate that macromolecular entanglement plays an important role on molecular diffusion, which can not be captured by the standard hard-sphere description.

7.2 Publication


 Cite this: *Soft Matter*, 2018,
14, 3105

Macromolecular diffusion in crowded media beyond the hard-sphere model†

 Pablo M. Blanco,^a Josep Lluís Garcés,^c Sergio Madurga^{*ab} and Francesc Mas^{ab}

The effect of macromolecular crowding on diffusion beyond the hard-core sphere model is studied. A new coarse-grained model is presented, the Chain Entanglement Softened Potential (CESP) model, which takes into account the macromolecular flexibility and chain entanglement. The CESP model uses a shoulder-shaped interaction potential that is implemented in the Brownian Dynamics (BD) computations. The interaction potential contains only one parameter associated with the chain entanglement energetic cost (U_e). The hydrodynamic interactions are included in the BD computations *via* Tokuyama mean-field equations. The model is used to analyze the diffusion of a streptavidin protein among different sized dextran obstacles. For this system, U_e is obtained by fitting the streptavidin experimental long-time diffusion coefficient D^{long} versus the macromolecular concentration for D50 (indicating their molecular weight in kg mol^{-1}) dextran obstacles. The obtained D^{long} values show better quantitative agreement with experiments than those obtained with hard-core spheres. Moreover, once parametrized, the CESP model is also able to quantitatively predict D^{long} and the anomalous exponent (α) for streptavidin diffusion among D10, D400 and D700 dextran obstacles. D^{long} , the short-time diffusion coefficient (D^{short}) and α are obtained from the BD simulations by using a new empirical expression, able to describe the full temporal evolution of the diffusion coefficient.

 Received 27th January 2018,
Accepted 20th March 2018

DOI: 10.1039/c8sm00201k

rsc.li/soft-matter-journal

1 Introduction

Biological media are known to contain a high concentration of a wide variety of macromolecular species such as proteins, polysaccharides or nucleic acids. For instance, the weight fraction of protein is around 5% in lymph, 9% in blood plasma and 35% in hemolysate.^{1,2} In the cellular cytosol, three-dimensional analysis of electron micrographs revealed a 20% volume fraction of fibrous supramolecular structures (*i.e.* F-actin, microtubules and intermediate filaments).³ These conditions, known as “macromolecular crowding” involve non-specific interactions among macromolecular species due to the excluded volume, van der Waals, electrostatic and hydrodynamic interactions. Macromolecular crowding substantially alters the diffusion processes, conformational properties and reactivity of the macromolecular species.^{4–8}

Several efforts are underway to properly understand the effect of crowding in macromolecular diffusion.⁹ Recently, *in vivo* and

in vitro experimental studies have been carried out to study macromolecular diffusion in crowded media. In *in vivo* experiments, fluorescent tracer proteins are introduced into the cell by means of transfection, microinjection or recombinant expression.¹⁰ Although this strategy allows the direct study of the macromolecular crowding, the intrinsic differences between the different intracellular microenvironments make the interpretation of the results difficult. *In vitro* experiments attempt to overcome this problem by reducing the complexity and heterogeneity of the media. The crowded environment is recreated using highly concentrated polymer solutions (usually dextran or ficoll).^{11–13} The motion of a fluorescent protein is then studied using spectroscopic techniques,¹⁴ mainly Fluorescence Correlation Spectroscopy (FCS)¹¹ and Fluorescence Recovery After Photobleaching (FRAP).^{12,13}

In order to interpret the experimental results, an increasing number of computational studies have been performed¹⁵ that provide a highly controlled environment. The comparison of the computational results with the experimentally observed results facilitates the quantification of the factors governing macromolecular diffusion in crowded media. Different computational approaches have been applied ranging from on-lattice Monte Carlo simulations^{16,17} and off-lattice Brownian Dynamics (BD)^{18–25} to Molecular Dynamics simulations.^{23,26,27} In implicit solvent simulations like BD, the Hydrodynamic Interactions (HIs) of the macromolecules have been found to be crucial to properly describe their motion in crowded media.²⁸

^a Department of Material Science and Physical Chemistry, Barcelona University, 08028 Barcelona, Spain. E-mail: fmas@ub.edu, pmlblanco@ub.edu, s.madurga@ub.edu

^b Institute of Theoretical and Computational Chemistry (IQTC), Barcelona University, 08028 Barcelona, Spain

^c Department of Chemistry, University of Lleida (UdL), 25003 Lleida, Spain. E-mail: jlgarcés@quimica.udl.cat

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c8sm00201k



Macromolecular diffusion is known to become sub-diffusive in crowded media. The time-dependent diffusion coefficient (D) of the macromolecule is defined as $D \equiv \langle r^2 \rangle / (2dt)$, where d is the topological dimension of the system and $\langle r^2 \rangle$ is the mean square displacement of the particles. In crowded media, D undergoes a transition from an initial value at short times (D^{short}) to an asymptotic final value at long times (D^{long}). During the transition between both regimes, macromolecular motion shows a non-linear dependence on time (usually called anomalous diffusion):^{16,29}

$$\langle r^2 \rangle = (2d)\Gamma_\alpha t^\alpha \quad (1)$$

where α is the anomalous diffusion exponent and Γ_α is a generalized transport coefficient.

In order to study macromolecular diffusion in the long-time regime, it is necessary to reach μs time scales. However, the large number of atoms present in a highly concentrated macromolecular solution causes atomistic-detailed computations to be limited to smaller time scales. As a consequence, in order to reach larger time scales, coarse grained approaches become necessary. Several approximations have been developed to take into account the HIs. The more accurate descriptions are those that involve the calculation of the diffusion tensor,³⁰ mainly based on the Rotne–Prager–Yamakawa approach.^{31–33} However, the calculation of the diffusion tensor at every time step is computationally very expensive. An alternative method is based on the mean field equations proposed by Tokuyama,^{34–36} which introduce an effective diffusion coefficient accounting for the HIs. This effective diffusion coefficient is computed as a function of the volume fraction and the dilute solution diffusion coefficient. However, the Tokuyama method is limited since it was deduced for systems of equal sized hard-core spheres. In previous studies that use Tokuyama equations, macromolecules are considered as hard spheres and implicit solvent *via* Brownian Dynamics (BD) is used.^{18,20–23}

In general, macromolecules have a flexible structure. As a result, when two macromolecules approach each other, their branches can become entangled and non-specific attractive and/or repulsive interactions between their chains are expected. This fact can play an important role in macromolecular diffusion in crowded media. Therefore, models going beyond the hard-core spheres, in which the conformational dynamic behaviour of macromolecules is lost, are necessary.³⁷ Softened interaction potentials have been previously proposed in the study of colloidal systems such as polymeric micelles^{38,39} and star-polymers.^{40–44} These systems consist of a dense core surrounded by a rather sparse corona. For pluronic polymeric micelles, studies based on implicit solvent and coarse grained computations have shown that both attractive and repulsive forces are present. However, the resulting potential of mean force between two particles is found to be purely repulsive.³⁹ Similar behaviour is found in the effective interparticle potentials used in star-polymer models. Star polymers can be modelled as surfaces coated with grafted polymers interacting by excluded volume effects. The particles are large enough to employ the Derjaguin approximation.^{45–47} The proposed potential is the result of the soft repulsion

between the polymeric branches and a hard-core repulsion between the rigid cores.

In this paper, diffusion of the streptavidin protein among dextran molecules will be analysed. Dextran has a branched, highly hydrated, polymeric structure. Dextran molecules are significantly smaller than micelles and star-polymers, so that some approximations previously used in these systems, such as the Derjaguin approximation, are no longer valid. On the other hand, their very sparse structure and chemical composition, including polar alcoholic groups, can compensate in part the steric repulsions and favour macromolecular entanglement, so that a singular treatment becomes necessary.

In Section 2, a new coarse-grained approach is proposed: the Chain Entanglement Softened Potential model (CESP). The CESP model is based on an inter-particle interaction potential depicted in Fig. 1. It includes both hard-core and soft interaction regions at proper characteristic lengths. In the soft region, where the potential is shoulder-shaped, particle overlapping is possible, so that macromolecular branches are allowed to become entangled. A single parameter controls the entanglement energy. This potential is introduced in the BD scheme in which the HIs are implemented using the mean-field Tokuyama equations. D^{long} and α are obtained from the simulated temporal evolution of D by using a new empirical expression. The new equation, unlike the mostly used power law (eqn (1)), is able to describe the evolution of D for the three temporal regimes.

In Section 3, the CESP model is used to study streptavidin diffusion in crowded media among dextrans of different sizes and concentrations. Firstly, the shoulder shaped potential is

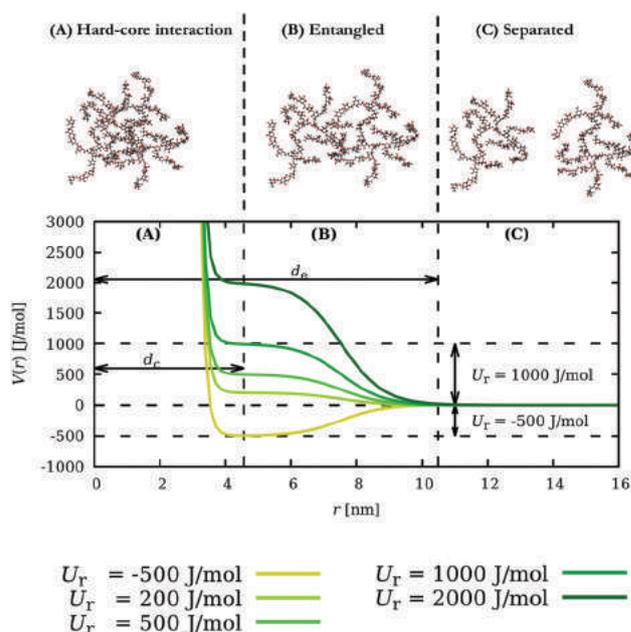


Fig. 1 Outline of the Chain Entanglement Softened Potential (CESP) model and the potential energy $V(r)$ vs. interparticle distance r (eqn (2)). Three interaction regions can be observed: no interaction (C), soft interaction where entanglement is allowed (B), and hard-core interaction (A). The chosen parameters are $d_c = 4.58$ nm and $d_e = 10.48$ nm, for U_T ranging from -500 to 2000 J mol⁻¹.



parametrized by fitting the computed D^{long} to the experimental values¹¹ for D50 (indicating its molecular weight in kg mol⁻¹) dextran obstacles. The BD simulations performed with the CESP model show a quantitative agreement between the calculated and experimental long time diffusion coefficient (D^{long}). This agreement is much better than the one obtained when the hard sphere model is used.²⁴ Once parametrized, we show that the CESP model is able to quantitatively predict D^{long} and the anomalous exponent (α) for the rest of the dextran sizes (D10, D400 and D700) in a wide range of macromolecular concentrations. Analysis of the radial distribution functions shows a significant increase in the population of entangled macromolecules with macromolecular concentration. This result highlights the importance of including the conformational dynamics of the macromolecules in the diffusion processes in crowded media.

2 Methodology

2.1 Chain Entanglement Softened Potential (CESP) model

In BD simulations of protein diffusion in crowded media, it is usual to reduce the resolution of the macromolecules to effective hard spheres.^{18,20–23} This approximation significantly reduces the computational cost allowing larger time scales to be reached. However, the loss of the macromolecular structural properties could become unrealistic in many cases.³⁷ As the volume fraction increases, the macromolecules are expected to start to become entangled, a fact that cannot be properly described within the hard-core sphere model.²⁴

In the present work, macromolecular entanglement is introduced *via* a coarse-grained model, the Chain Entanglement Softened Potential model (CESP). The model relies on an empiric interparticle interaction potential that resembles the continuous shouldered well potential used in computational models of coarse-grained fluids.^{48,49} In these studies, density anomalies have been reported when a softened inter-particle potential is applied. The obtained radial distribution functions showed that an inner coordination shell is formed when the fluid density increases. Moreover, the population of the inner coordination shell was found to be proportional to the fluid density. Here, a redefinition of this idea is proposed. The coarse grained particles are no longer fluid molecules but macromolecules in solution. As the density of macromolecules in solution increases, they start to become entangled. Similar potentials have been used in the description of protein–surface interactions.⁵⁰

The interaction potential corresponding to the CESP model is shown in Fig. 1. Two clear interaction regions can be observed that define two characteristic distances: the entanglement distance (d_e) and the core distance (d_c). If the macromolecules are at a distance r larger than d_e , they are considered to be separated, and they interact weakly (Fig. 1C). However, when the macromolecules are at $d_c \leq r \leq d_e$, they are considered to become entangled (Fig. 1B). In the latter situation, steric repulsions start to arise between the macromolecular chains. Such repulsions

can be compensated in part by the presence of van der Waals and hydrogen bonding attraction forces. As a result, a shoulder is developed in the proposed potential (Fig. 1). If the macromolecules get closer ($r < d_c$), the steric repulsion between the macromolecular skeletons dramatically increases and the hard-core region emerges (Fig. 1A). The resulting picture resembles models previously proposed for polymer micelles and star-polymers, composed of a dense core and a sparse corona. Unlike these systems, dextran molecules are polysaccharides with a complex branched structure where no structurally differentiated parts can be identified.^{51,52} Therefore (d_e) and (d_c) should be regarded as the characteristic lengths of the different interaction regimes rather than chemically defined parts of the macromolecules. The proposed potential $V(r)$ acting between two particles at a distance r reads

$$V(r) = \varepsilon_0 \left(\frac{d_e}{r} \right)^n - \frac{U_r}{2} \tanh \left(\frac{d_c/d_0}{d_e - d_c} \left(r - \frac{d_e + d_c}{2} \right) \right) + \frac{U_r}{2} \quad (2)$$

where U_r is a parameter that quantifies the entanglement energetic cost. The sign of U_r indicates which forces, repulsive (positive sign) or attractive (negative sign), dominate in the soft interaction region. The first term in eqn (2) accounts for the hard-core repulsion region. Subsequently,^{48–50} the exponent n is set to $n = 24$. $\varepsilon_0 = 1 \text{ J mol}^{-1}$ and $d_0 = 1 \text{ nm}$ are just to set the units. $V(r)$ is depicted in Fig. 1 for U_r values ranging from -500 to 2000 J mol^{-1} . Although the real situation is most probably more complex, eqn (2) can be regarded as a minimal model, compatible with the physicochemical properties of macromolecules considered, with only one parameter to be fitted.

The proposed interaction potential will be used to analyze the diffusion of streptavidin among different sized dextrans acting as inert obstacles,¹¹ tracked down using FCS. Both the tracer protein and the dextran obstacles are considered to be able to become entangled. The differences between protein/dextran and dextran/dextran (protein/protein interactions are not relevant due to the large excess of dextran) are then included in an effective way by the specific values of d_e and d_c for the protein and dextran, while the U_r value should be understood as an average interaction parameter. In order to estimate d_e and d_c , the interaction regions in a macromolecule are considered to be concentric spheres with two characteristic radii: the entanglement radius (R_e) and the hard-core radius (R_c), so that for two interacting particles i and j , $d_{e,ij} = R_{e,i} + R_{e,j}$ and $d_{c,ij} = R_{c,i} + R_{c,j}$.

In this work, R_e is associated with the hydrodynamic radius of the macromolecule in dilute solution. For dextran molecules, it has been computed by fitting experimental hydrodynamic radii *vs.* molecular weight obtained by quasi-elastic light scattering to an empiric power law function.^{24,53} The dilute solution diffusion coefficient (D_0) for dextrans is calculated by incorporating the resulting R_e values into the Stokes–Einstein equation. On the other hand, the HYDROPRO software (Version 10)⁵⁴ has been used to compute D_0 corresponding to streptavidin. HYDROPRO considers each atom in the protein surface as a hydrodynamic frictional sphere. Protein atomic coordinates corresponding to three different crystallographic structures^{55–57} were used in the calculations. D_0 was taken as the average of the



values resulting from the three structures. R_c of streptavidin is then obtained *via* the Stokes–Einstein equation.

Here, we propose to estimate R_c as the maximum approach distance between macromolecules, and it has been estimated using the specific volume ν , considering them as compact spheres²⁴

$$R_c = \sqrt[3]{\frac{3\nu}{4\pi N_A} M_w} \quad (3)$$

where N_A is the Avogadro number and M_w is the molecular weight. We have used $\nu = 0.625 \text{ cm}^3 \text{ g}^{-1}$ for dextrans⁵⁸ and $\nu = 0.71 \text{ cm}^3 \text{ g}^{-1}$ for streptavidin.⁵⁹

M_w , D_0 , R_c and R_e for streptavidin and different sized dextrans are reported in Table 1. Interestingly, the ratio R_e/R_c increases with dextran size, which could be caused by an increase in dextran branching with molecular weight. Note also that D10 and streptavidin have similar D_0 and R_c values. However, D10 has a higher ratio R_e/R_c than streptavidin, in accordance with the very different chemical structures of both macromolecules, much more compact in the last case.

2.2 Brownian Dynamics algorithm

Brownian Dynamics (BD) simulations make use of the Langevin equation in the over-damped limit,^{60,61} so that the particles are considered to be under the effect of a stochastic force that accounts for the collisions with the solvent. The stochastic force is mathematically represented by Gaussian random noise,^{19,24,25,62} and the equation of motion for the N particles reads

$$\mathbf{x}(t + \Delta t) = \mathbf{x}(t) - \frac{\Delta t}{RT} \mathbf{D} \nabla V(r) + \sqrt{2\mathbf{D}\Delta t} \boldsymbol{\xi}(t) \quad (4)$$

where $\boldsymbol{\xi}$ is a vector of $3N$ Gaussian random numbers with zero mean and unit variance, T is the system temperature, \mathbf{x} is a vector with the $3N$ Cartesian coordinates, V is the potential of the CESP model (eqn (2)) and \mathbf{D} is the diffusion tensor.

Since the solvent is simulated implicitly, the HIs must be included in eqn (4). HIs are solvent mediated correlations of the particle motions and they have been found to play an important role in macromolecular diffusion in crowded media.^{24,28,63} HIs can be included in eqn (4) by calculating \mathbf{D} following the Rotne–Prager–Yamakawa (RPY) method at each time step,³⁰ which is computationally very demanding.

Tokuyama equations^{34–36} offer an alternative to the RPY method based on a mean field approximation. In this approach, an effective diffusion coefficient that accounts for the particle

mobility reduction due to the HIs is analytically calculated. The expression for the diffusion coefficient is deduced starting from first principles, *i.e.*, the Fokker–Planck equation for the single-particle distribution function coupled with the Navier–Stokes equations. This equation is analytically solved at short times considering a stationary configuration of equal-sized Brownian particles. In doing so, an approximation of the effective diffusion coefficient at short times (D^{short}) is obtained. For a particle diffusing in a system with a volume fraction ϕ , D^{short} reads

$$D^{\text{short}} = \frac{D_0}{1 + H(\phi)} \quad (5)$$

where $H(\phi)$ describes the HI contributions in the short-time regime

$$H(\phi) = \frac{2b^2}{1-b} - \frac{c}{1+2c} - \frac{bc(2+c)}{(1+c)(1-b+c)} \quad (6)$$

with $b = \sqrt{\frac{9}{8}}\phi$ and $c = \frac{11}{16}\phi$.

In this approach, \mathbf{D} in eqn (4) is replaced by the effective D^{short} , which is a function of ϕ and D_0 . However, the choice of ϕ presents two main difficulties. On the one hand, for a number of obstacles N_o , one could naively calculate ϕ as

$$\phi = \frac{4\pi N_o R_o^3}{3V} \quad (7)$$

by identifying the obstacle radius R_o as the hydrodynamic radius. However, this leads to unphysical ϕ values larger than unity for the higher dextran concentrations, suggesting the existence of interpenetration between dextran molecules. On the other hand, the soft nature of the particles and the possibility of entanglement cause ϕ to be undetermined. A possible solution to this problem could be to interpret ϕ as an effective volume fraction ϕ^{eff} , depending on the macromolecular concentration and molecular size *via* some parameters to be determined. Then, the problem would be to find the value for these parameters that better predicts the experimental D^{long} using BD dynamics and Tokuyama equations. However, finding the best-fitted value of these parameters would require a repeated iteration of the BD dynamics, which is computationally very expensive. Instead, a heuristic approach has been chosen, consisting of estimating ϕ^{eff} using an analytic expression for D^{long} deduced by Tokuyama for equal sized hard-core spheres in the absence of the interaction force

$$D^{\text{long}} = \frac{D^{\text{short}}}{1 + \kappa \frac{D^{\text{short}} \phi^{\text{eff}}}{D_0 \phi_c} \left(1 - \frac{\phi^{\text{eff}}}{\phi_c}\right)^{-2}} \quad (8)$$

where κ and ϕ_c are parameters set to $\kappa = 2.0$ and $\phi_c = 1.09$.³⁵ By equating eqn (8) to the experimental D^{long} , we can obtain a first approximation to ϕ^{eff} . The resulting ϕ^{eff} values have been found to depend not only on dextran concentration but on the dextran size. For the same dextran mass concentration, larger dextran sizes lead to larger ϕ^{eff} values, so that dextrans seem to effectively occupy more volume as the molecular size increases. In fact, the computed ϕ^{eff} values have been found to follow the

Table 1 Molecular weight (M_w) in kg mol^{-1} (kDa),¹¹ entanglement radius (R_e) and hard core radius (R_c) for streptavidin and different sized dextrans (D10, D50, D400 and D700)

	M_w [kDa]	R_e [nm]	R_c [nm]	R_e/R_c	D_0 [$\text{nm}^2 \text{ ns}^{-1}$]
Streptavidin	52.8	3.04	2.45	1.24	0.085
D10	11.6	2.77	1.43	1.94	0.089
D50	48.6	5.24	2.29	2.29	0.047
D400	409.8	13.52	4.67	2.90	0.018
D700	667.8	16.81	5.51	3.14	0.015



simple empiric equation (which resembles the Langmuir-Freundlich isotherm⁶⁴)

$$\phi^{\text{eff}} = \frac{z^m}{1 + z^m} \quad (9)$$

with $z = k\nu c \frac{R_c}{R_e}$, where c is the dextran concentration, ν represents the specific value, and R_e/R_c is the characteristic radii ratio. k and m are fitted parameters whose values are $k = 3.7 \pm 0.1$ and $m = 1.43 \pm 0.08$. The best-fitted ϕ^{eff} as a function of z for D10 (squares), D50 (circles), D400 (upward-pointing triangles) and D700 (downward-pointing triangles) is shown in Fig. 2. The dashed line represents the best-fitted curve eqn (9). This procedure is actually expected to provide an approximation to ϕ^{eff} , whose suitability can only be justified a posteriori, if the BD simulations coupled to the CESP model properly reproduce the experimental D^{long} values. This is the case, as will be shown in the next section.

BD simulations are performed in a cubic box under periodic boundary conditions in all directions. The time step is set to 0.1 ns and the simulation time ranges from 10 to 150 μs depending on the time necessary to reach the stationary value of the diffusion coefficient at long times (D^{long}). The system temperature is 298.15 K. D is computed via eqn (5) and (6) using the ϕ^{eff} predicted by eqn (9) for each type of particle (dextran and protein). All the studied systems have 100 obstacle dextran particles and only one to five tracer protein particles to ensure that the volume occupied by the tracer is less than 1% of the volume occupied by the obstacles. The length of the simulation box is properly tuned to match the experimental dextran concentration and ranges from 17.56 nm (for the smaller dextran) to 164.32 nm (for the larger dextran). $\langle r^2 \rangle$ is computed after a thermalisation time of 0.5 μs in order to avoid possible initial unrealistic motion due to the random starting configuration. For each studied system, $\langle r^2 \rangle$ is obtained by averaging from 1000 (higher macromolecular concentration) to 15 000

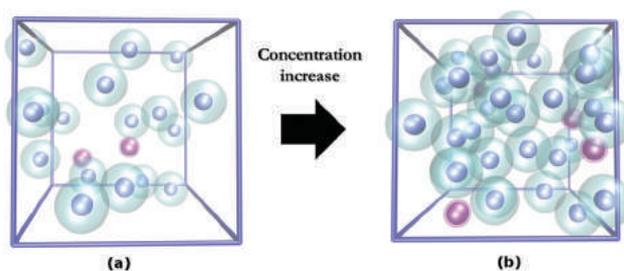


Fig. 3 Illustration of the simulation box of two BD simulations of the streptavidin protein (purple) diffusing among D50 dextran obstacles (blue). Dextran concentrations are (a) 25 g L⁻¹ and (b) 100 g L⁻¹. Each macromolecule is represented as two concentric spheres representing the soft (transparent) and hard-core (opaque) interaction regions. As concentration increases, steric compression promotes macromolecules to become entangled.

(lower macromolecular concentration) BD realizations in order to ensure appropriate statistics.

In Fig. 3, two BD simulation boxes are depicted to illustrate the CESP model. Tracer streptavidin protein molecules (purple) diffuse among D50 dextran obstacles (blue). Each macromolecule is represented as two concentric spheres with radii of R_c and R_e , which characterize the soft (transparent) and the hard-core (opaque) interaction regions, respectively. The macromolecular concentrations of the solutions are 25 g L⁻¹ (Fig. 3a) and 100 g L⁻¹ (Fig. 3b). In the less concentrated solution, it can be observed that the soft interaction is enough to avoid macromolecule overlapping. In the more concentrated solutions, steric compression allows the macromolecules to overcome the initial steric repulsion and become entangled.

2.3 Determination of D^{long} , D^{short} and α from BD simulations: a new empiric equation for the time evolution of D

As mentioned before, the diffusion coefficient D of a particle is known to have a transition between two steady states: D^{short} and D^{long} . $\log(D)$ vs. $\log(t)$ profiles computed from BD simulations can be very well fitted to the function

$$\log(D) = B \tanh(\beta(\log(t) - \gamma)) + A \quad (10)$$

where A , B , β and γ are parameters related to the relevant dynamic properties of the system (the mathematical details are given in the appendix): $A = \log(D^{\text{short}}D^{\text{long}})/2$, $B = \log(D^{\text{long}}/D^{\text{short}})/2$, $\beta = (\alpha - 1)/B$ and γ is the inflection time between D^{short} and D^{long} . This new empiric equation is able to describe the diffusion coefficient in the three temporal regimes. This fact represents an advantage compared to the scaling law (eqn (1)), which is unable to describe the asymptotic behaviour at short and long times. Moreover, it provides a simple method to compute D^{short} , D^{long} and α from the trajectory of a tracer particle.

In Fig. 4, the time evolution of the simulated diffusion coefficient of streptavidin in a D50 crowded solution, at a macromolecular concentration of 350 g L⁻¹, is shown (blue line). As expected, the diffusion coefficient undergoes a transition from D^{short} to D^{long} as macromolecular crowding slows down the motion of the tracer particle. It can be observed that eqn (10), depicted in red, fits very well to the computed values

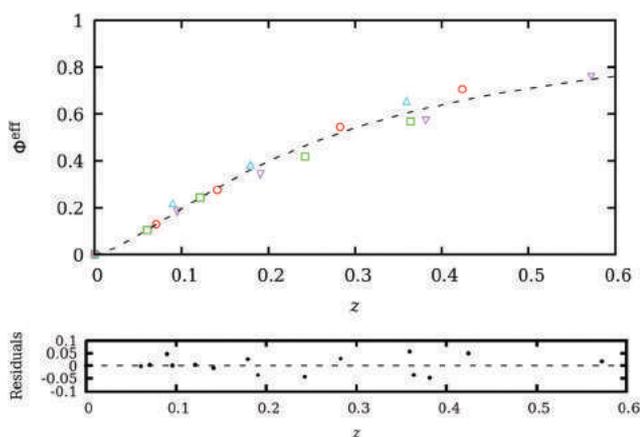


Fig. 2 Effective volume fraction (ϕ^{eff}) as a function of $z = k\nu c R_e/R_c$ estimated using eqn (5)–(8) for D10 (squares), D50 (circles), D400 (upward-pointing triangles) and D700 (downward-pointing triangles) dextran molecules. The dashed line represents the fitted eqn (9). The residuals of the fitting procedure are also depicted and show a very good fitting.



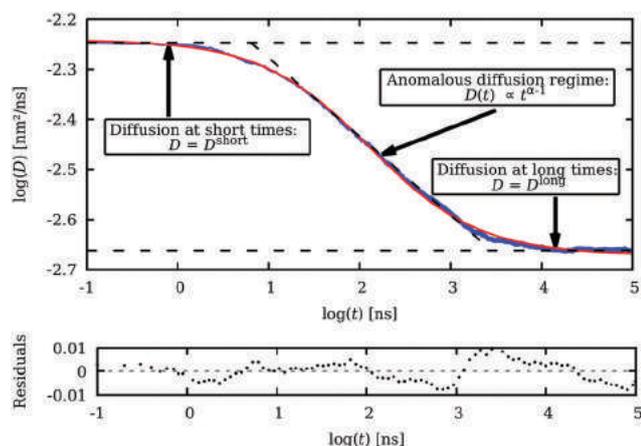


Fig. 4 Time evolution of the diffusion coefficient (D) in logarithmic scale. The blue line results from averaging 3000 BD simulations of 100 μs long in a cubic simulation box with a side of 28.466 nm. The simulation box contains 5 tracer particles (streptavidin) diffusing among 100 obstacle particles (D50 dextran). The macromolecular concentration is 350 g L^{-1} . Macromolecular crowding slows down the tracer diffusion coefficient from D^{short} to D^{long} . The red line is the result of fitting eqn (10) to the simulated results, which allows D^{short} , D^{long} and α to be easily determined. The resulting fitting is very good, as shown by the residuals.

over all the temporal regimes, allowing rigorous and easy determination of D^{short} , D^{long} and α .

3 Results and discussion

3.1 Model parametrization

The proposed shoulder-shaped potential used in the CESP model includes only one parameter to be fitted, *i.e.*, the entanglement energetic cost U_r (eqn (2)). In order to properly set U_r , BD simulations have been performed with U_r values ranging from 200 to 2000 J mol^{-1} . In these simulations, D^{long} of streptavidin (normalised to D_0) has been computed at seven macromolecular concentrations of D50 dextran: 50, 100, 200, 210, 260, 300 and 360 g L^{-1} . D^{long} is obtained by fitting the computed $\log(D)$ vs. $\log(t)$ profile to eqn (10).

The obtained values are shown in Fig. 5 in green colour gradation. As can be observed, D^{long} decreases with U_r because of the increase in the impenetrability of the particles and the consequent reduction in the available space for the particle to diffuse. Experimental D^{long} values reported in ref. 11 are also depicted in the same figure (red markers). For $U_r = 500 \text{ J mol}^{-1}$, very good quantitative agreement between simulations and experiments is obtained. In all the BD computations performed in the rest of this work, U_r has been set to this value, which indicates that, for dextran, the repulsion forces arising from chain entanglement slightly dominate over the attractive ones. D^{long} values obtained using the hard sphere model, reported in ref. 24, are also depicted in Fig. 5 (black markers). In this previous computational work, dextran macromolecules were modeled considering the dextran molecules as hard-core spheres with an effective radius, which was chosen as an intermediate value lying between the hydrodynamic radius⁵³ and the compact radius.

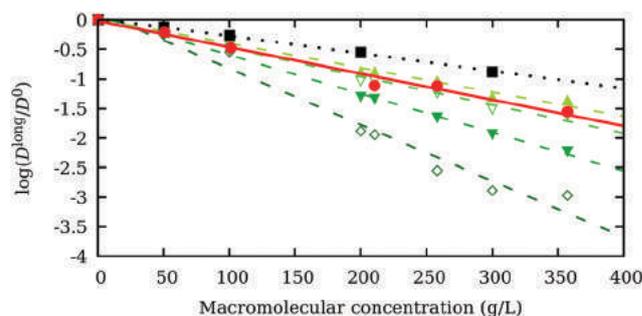


Fig. 5 Calculated D^{long}/D_0 as a function of the concentration corresponding to D50 dextran for U_r values: 200 J mol^{-1} (green triangles), 500 J mol^{-1} (empty green inverted triangles), 1000 J mol^{-1} (green inverted triangles) and 2000 J mol^{-1} (green diamonds). Black square markers represent the values obtained using the hard sphere model (taken from ref. 24) while red circle markers denote the experimental D^{long} value (taken from ref. 11). For $U_r = 500 \text{ J mol}^{-1}$, good agreement between the CESP model and the experimental values is observed. Lines are only to guide the eye.

This approach accounted for the possibility of macromolecular overlapping in an average way, but represented a static description of the conformational state, being unable to properly mimic macromolecular entanglement. As is clearly observed in Fig. 5, D^{long} values obtained using the CESP model are much closer to the experimental ones than those coming from the hard-core sphere model.²⁴ For the rest of the dextran sizes (D10, D400 and D700), direct comparison of D^{long} predicted by the hard-sphere model²⁴ and the CESP model can be found in the ESI.† Again, it is found that the CESP model shows a significant improvement in the prediction of the obtained experimental D^{long} over that using the hard-core sphere model. Therefore, the inclusion of coupling of the dynamic and conformational properties seems to be necessary to accurately describe diffusion in crowded media.

3.2 Model prediction for D^{long} and the anomalous exponent α

In the previous section, we have found that for $U_r = 500 \text{ J mol}^{-1}$, the CESP model is able to describe D^{long} of streptavidin for D50 dextran obstacles. Let us now see that this model can also predict, without any further parametrization, D^{long} and α of streptavidin for different dextran sizes (D10, D400 and D700) and concentrations (from 25 to 350 g L^{-1}). D^{long} and α are again obtained from the BD simulations by fitting eqn (10) to the computed D time evolution. The curve $\log(D)$ versus $\log(t)$ has been plotted in Fig. 6, together with the best fitted eqn (10), for four dextran sizes and concentrations: (Fig. 6a) D10 at 300 g L^{-1} , (Fig. 6b) D50 at 200 g L^{-1} , (Fig. 6c) D400 at 100 g L^{-1} and (Fig. 6d) D700 at 50 g L^{-1} . As can be observed, for all the systems presented, eqn (10) fits very well to the computational results.

The computed D^{long} versus macromolecular concentration plots are depicted in Fig. 7 for four dextran sizes: D10, D50, D400 and D700. The results are shown together with the experimental values reported in ref. 11 for the same systems. Dashed lines are only to guide the eye and they follow the simulation results. In general, the computed results are in very good quantitative



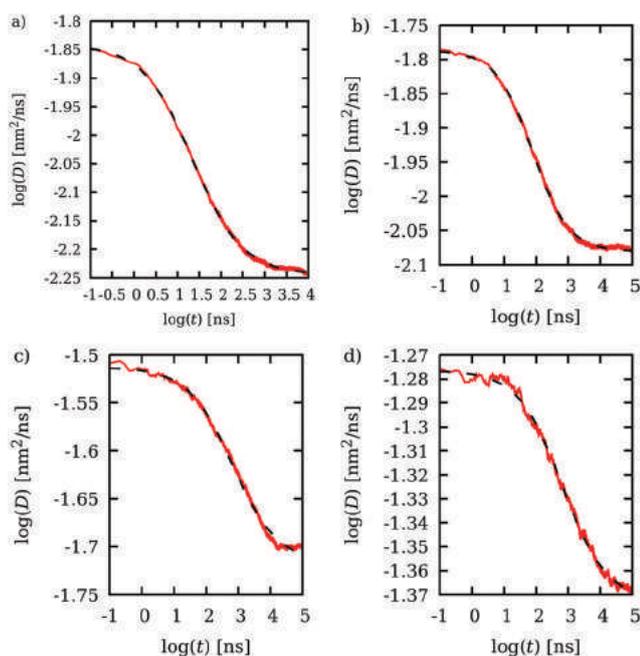


Fig. 6 Time evolution of the diffusion coefficient D versus time in the logarithmic scale obtained from BD computations (red line). The best-fitted eqn (10) is plotted as a black dashed line. Four cases are depicted: (a) D10 at 300 g L^{-1} , (b) D50 at 200 g L^{-1} , (c) D400 at 100 g L^{-1} and (d) D700 at 50 g L^{-1} . For all the systems presented, eqn (10) fits very well to the computed values, allowing easy determination of D^{long} , D^{short} and α from the temporal evolution of D .

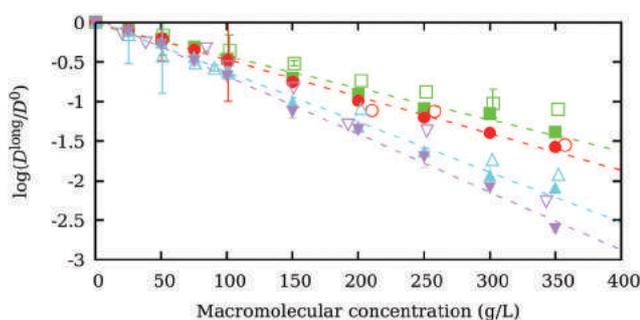


Fig. 7 $\log(D^{\text{long}}/D^0)$ for streptavidin versus macromolecular concentration. The filled markers represent BD simulations and the empty ones correspond to experimental values¹¹ for four dextran sizes: D10 (green squares), D50 (red circles), D400 (upward pointing blue triangles) and D700 (downward pointing purple triangles). Dashed lines are only to guide the eye and they follow the simulation results.

agreement with the experimental values. However, the values corresponding to D10 exhibit some systematic error, which could be explained taking into account that the potential energy barrier U_r has been set, regardless of the dextran size, to the best fit D^{long} for D50. This suggests that U_r should be smaller for D10 than for D50, D400 and D700. Since U_r accounts for the energy cost of the macromolecular entanglement, this fact could suggest a lower energy cost for D10 entanglement than the one for D50, D400 and D700. In agreement with the experiments reported in ref. 11, BD simulations predict a

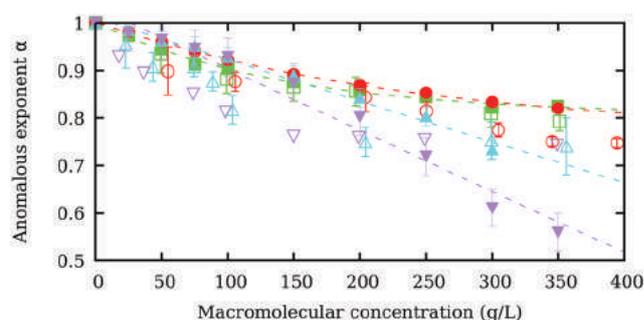


Fig. 8 Anomalous exponent α versus macromolecular concentration. Filled markers represent BD simulations and empty markers correspond to the experimental values¹¹ for four dextran sizes: D10 (green squares), D50 (red circles), D400 (upward pointing blue triangles) and D700 (downward pointing purple triangles). Dashed lines are only to guide the eye and they follow the simulation results.

decrease in D^{long} with macromolecular concentration, as a result of volume exclusion and hydrodynamic interactions. Alternative potentials based on a self-consistent field method appropriate for weak excluded interactions⁴⁵ have been implemented in BD simulations of star-polymer micelles.⁴² As in dextran molecules, the computational results exhibit similar D^{long} decay with the particle density. However, at the quantitative level, the calculated values are much larger than the experimental ones. The discrepancies could be explained by the fact that HIs were not included in the computations of ref. 42. In the present work, HIs are included in the BD scheme, and it has been found that they are crucial for the quantitative reproduction of the experimental D^{long} values.^{24,28}

The anomalous exponent α provides information about the transition between the short and the long-time regimes. α values close to unity indicate a smooth transition. As α deviates from unity, the transition becomes more pronounced. The simulated α versus macromolecular concentration plots are depicted in Fig. 8 together with the experimental values obtained using FCS.¹¹ As can be observed, the CESP model qualitatively captures the experimentally observed behaviour. α decays with macromolecular concentration as a result of an increase in particle collisions. Moreover, α also decreases with dextran size, which could be a consequence of the increase of ϕ^{eff} , so that the volume fraction of the system seems to play a crucial role in α and therefore in the transition between the D^{short} and D^{long} steady states. However, quantitative discrepancies are observed, which could be explained not only by the limitations of the model, but by the way in which α has been calculated. The experimental α values have been obtained by fitting the auto-correlation function of the tracer protein fluorescence using the power law described in eqn (1) while here, α has been determined as the slope in the inflection point of the curve $\log(D)$ versus $\log(t)$ using eqn (10).

3.3 Radial distribution functions

The simulated radial distribution functions (rdfs) for the cases analyzed in the previous section (D10 (a), D50 (b), D400 (c) and



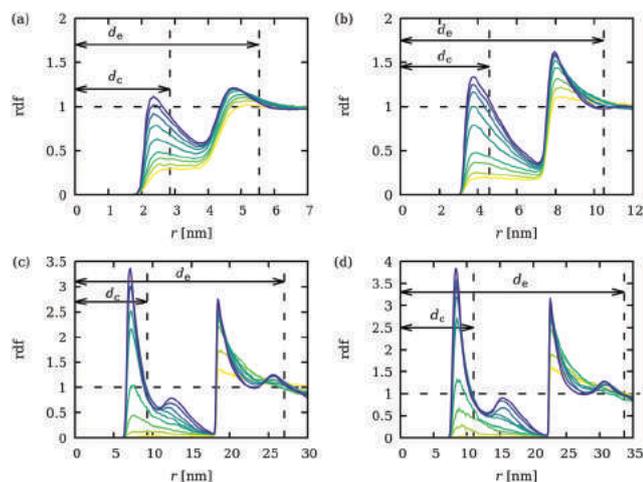


Fig. 9 Simulated radial distribution functions (rdfs) versus inter-particle distance using the CESP model for four dextran sizes: (a) D10, (b) D50, (c) D400 and (d) D700. For each size, the rdf has been computed at eight concentrations ranging from 25 to 350 g L⁻¹, depicted in color gradation from yellow (25 g L⁻¹) to purple (350 g L⁻¹).

D700 (d)) are shown in Fig. 9 for macromolecular concentrations ranging from 25 g L⁻¹ to 350 g L⁻¹. The rdfs have been computed from the BD computations by calculating the center-to-center distance of every dextran obstacle to the rest of the dextran particles. Averages have been taken over 10 000 configurations every 2 ns of simulation. As can be observed, the CESP model predicts an increase in the number of entangled macromolecules with macromolecular concentration, as revealed by the gradual increase in the rdf at $r = d_c$. This fact is observed for all the studied dextran sizes. Two sharp peaks at distances slightly lower than d_c and d_e define two coordination shells of entangled macromolecules, corresponding to highly entangled macromolecules (close to d_c) and weakly entangled macromolecules (close to d_e), respectively. It is also interesting to note the presence of two secondary peaks for D400 and D700 at macromolecular concentrations of 250, 300 and 350 g L⁻¹, corresponding to additional coordination shells and suggesting other possible degrees of entanglement for the biggest dextran molecules. Note that the two main peaks in the entanglement region are developed even in the absence of explicit attractive terms in the proposed potential. Entanglement is thus the way the system can minimize the excluded volume effects, even paying a small energetic cost. It is reasonable to conclude that the presence of attractive forces would enhance this effect, although it is not necessary to produce it.

The first rdf peak has been integrated in order to determine the number of entangled particles (N_E) close to the core of the reference particle. N_E is shown in Fig. 10 for the four analyzed dextran sizes. It can be observed that N_E increases exponentially with macromolecular concentration, so that entanglement is promoted by steric compression. This effect is enhanced by the increase in dextran size, as a consequence of the increase in ϕ^{eff} , so that bigger dextran obstacles are able to effectively occupy higher volumes than the smaller ones for the same macromolecular concentration.

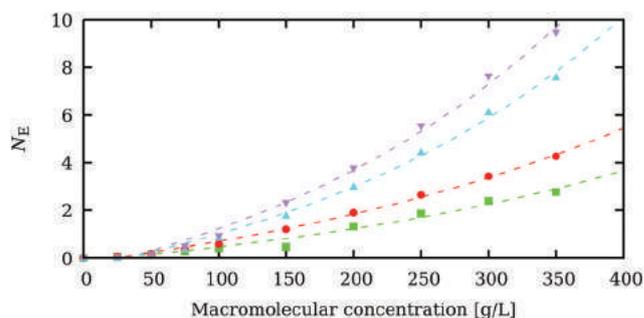


Fig. 10 Number of entangled macromolecules around a reference particle (N_E) versus macromolecular concentration for four dextran sizes: D10 (green squares), D50 (red circles), D400 (blue upward pointing triangles) and D700 (purple downward pointing triangles). The number of entangled particles increases with macromolecular concentration as a result of steric compression.

4 Conclusions

A new coarse grained approach to macromolecular diffusion in crowded media that goes beyond the hard sphere model is presented: the Chain Entanglement Softened Potential (CESP) model. Two interaction regions between macromolecules are considered: a soft interaction region caused by chain entanglement, a combination of repulsive and attractive effects, and a hard-core interaction region caused by the dramatic increase in steric repulsions between macromolecular skeletons when they approach below a certain distance. The resulting picture is quantified by means of a shoulder-shaped potential that allows macromolecular conformational flexibility and chain entanglement to be taken into account. The model can be parametrized by a unique parameter (U_r) that is associated with the entanglement energetic cost. For streptavidin protein diffusion in D50 crowded solution, U_r has been quantified by fitting the long time diffusion coefficient (D^{long}) resulting from BD simulations to the experimental values obtained using Fluorescence Correlation Spectroscopy (FCS).¹¹

In order to determine D^{long} and the anomalous exponent α from BD simulations, a new empiric expression for the temporal evolution of the diffusion coefficient, valid for all the temporal regimes, has been proposed. The new function goes beyond the standard power law and depends on the four relevant dynamic properties: D^{short} , D^{long} , α and the inflection time γ . This function fits very well to the BD computations for all the obstacle concentrations and sizes under study, and it could be useful in forthcoming computational studies focused on anomalous diffusion.

The computed D^{long} using the CESP model clearly improves on the results obtained in previous studies based on hard-core spheres.²⁴ Without any further parametrization of $V(r)$, the CESP model is able to provide quantitative predictions of D^{long} and α of streptavidin in media crowded by D10, D400 and D700 at several dextran concentrations (25 to 350 g L⁻¹). Radial distribution functions exhibit a significant increase in the population of entangled macromolecules with the macromolecular concentration and dextran size. Our results indicate that an adequate



description of macromolecular entanglement is essential to understand macromolecular diffusion in crowded media.

Conflicts of interest

The authors declare no conflicts of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Appendix

The new empirical equation proposed here for the time evolution of D (eqn (10)) contains four parameters that are related to the relevant dynamic properties of the system, *i.e.*, D^{short} , D^{long} , α and the inflection time γ . At short times ($\log(t) \rightarrow -\infty$), eqn (10) must fulfill $\log(D) \rightarrow \log(D^{\text{short}})$, while at long times ($\log(t) \rightarrow +\infty$), we have $\log(D) \rightarrow \log(D^{\text{long}})$. Taking limits in eqn (10)

$$\lim_{t \rightarrow -\infty} \log(D) = -B + A = \log(D^{\text{short}}) \quad (11)$$

$$\lim_{t \rightarrow +\infty} \log(D) = B + A = \log(D^{\text{long}}) \quad (12)$$

from which it follows that $A = \frac{1}{2} \log(D^{\text{short}} D^{\text{long}})$ and $B = \frac{1}{2} \log(D^{\text{long}}/D^{\text{short}})$. By construction, γ corresponds to the inflection point of the time evolution curve. Some elementary algebra shows that

$$\left(\frac{d^2 \log(D)}{d \log(t)^2} \right)_{\log(t)=\gamma} = 0 \quad (13)$$

Finally, β can be related to the anomalous diffusion exponent α , used when the transition regime is analyzed in terms of the standard power law (eqn (1))

$$D = \Gamma_x t^{\alpha-1} \Leftrightarrow \log(D) = (\alpha - 1) \log(t) + \log(\Gamma_x) \quad (14)$$

By taking the first derivative in eqn (10) and comparing to eqn (14), one obtains

$$\left(\frac{d \log(D)}{d \log(t)} \right)_{\log(t)=\gamma} = \beta \quad B = \alpha - 1 \quad (15)$$

which leads to $\beta = (\alpha - 1)/B$.

Acknowledgements

We acknowledge the financial support from the Spanish Ministry of Science and Innovation (project CTM2016-78798-C2-1-P) and Generalitat de Catalunya (Grants 2014SGR1017, 2014SGR1132 and XrQTC). Sergio Madurga and Francesc Mas acknowledge the funding of the EU project 8SEWP-HORIZON 2020 (692146). Pablo M. Blanco also acknowledges the financial support from the grant (FI-2017) of Generalitat de Catalunya. The authors also want to thank Prof. Giancarlo Franzese (University of Barcelona) for his suggestion to use a shouldered interaction potential to model macromolecular flexibility.

References

- 1 A. P. Minton, *Biopolymers*, 1981, **20**, 2093–2120.
- 2 S. B. Zimmerman and S. O. Trach, *J. Mol. Biol.*, 1991, **222**, 599–620.
- 3 A. Minton, *Biophys. J.*, 1992, **63**, 1090–1100.
- 4 R. J. Ellis, *Trends Biochem. Sci.*, 2001, **26**, 597–604.
- 5 H.-X. Zhou, G. Rivas and A. P. Minton, *Annu. Rev. Biophys.*, 2008, **37**, 375–397.
- 6 C. Balcells, I. Pastor, E. Vilaseca, S. Madurga, M. Cascante and F. Mas, *J. Phys. Chem. B*, 2014, **118**, 4062–4068.
- 7 I. Pastor, L. Pitulice, C. Balcells, E. Vilaseca, S. Madurga, A. Isvoran, M. Cascante and F. Mas, *Biophys. Chem.*, 2014, **185**, 8–13.
- 8 C. Balcells, I. Pastor, L. Pitulice, M. Via, S. Madurga, E. Vilaseca, A. Isvoran, M. Cascante and F. Mas, *New Frontiers in Chemistry*, 2015, **24**, 3–16.
- 9 G. Rivas and A. P. Minton, *Trends Biochem. Sci.*, 2016, **41**, 970–981.
- 10 M. C. Konopka, I. A. Shkel, S. Cayley, M. T. Record and J. C. Weisshaar, *J. Bacteriol.*, 2006, **188**, 6115–6123.
- 11 D. S. Banks and C. Fradin, *Biophys. J.*, 2005, **89**, 2960–2971.
- 12 I. Pastor, E. Vilaseca, S. Madurga, M. Cascante and F. Mas, *J. Phys. Chem. B*, 2010, **114**, 4028–4034.
- 13 I. Pastor, E. Vilaseca, S. Madurga, M. Cascante and F. Mas, *J. Phys. Chem. B*, 2010, **114**, 12182.
- 14 J. A. Dix and A. Verkman, *Annu. Rev. Biophys.*, 2008, **37**, 247–263.
- 15 M. Feig, I. Yu, P.-h. Wang, G. Nawrocki and Y. Sugita, *J. Phys. Chem. B*, 2017, 1–17.
- 16 E. Vilaseca, A. Isvoran, S. Madurga, I. Pastor, J. L. Garcés and F. Mas, *Phys. Chem. Chem. Phys.*, 2011, **13**, 7396–7407.
- 17 E. Vilaseca, I. Pastor, A. Isvoran, S. Madurga, J. L. Garcés and F. Mas, *Theor. Chem. Acc.*, 2011, **128**, 795–805.
- 18 J. Sun and H. Weinstein, *J. Chem. Phys.*, 2007, **127**, 155105.
- 19 P. Mereghetti and R. C. Wade, *J. Phys. Chem. B*, 2012, **116**, 8523–8533.
- 20 S. Hasnain, C. L. McClendon, M. T. Hsu, M. P. Jacobson and P. Bandyopadhyay, *PLoS One*, 2014, **9**, 1–12.
- 21 S. Kondrat, O. Zimmermann, W. Wiechert and E. V. Lieres, *Phys. Biol.*, 2015, **12**, 046003.
- 22 T. Sentjabrskaja, E. Zaccarelli, C. De Michele, F. Sciortino, P. Tartaglia, T. Voigtmann, S. U. Egelhaaf and M. Laurati, *Nat. Commun.*, 2016, **7**, 11133.
- 23 I. Yu, T. Mori, T. Ando, R. Harada, J. Jung, Y. Sugita and M. Feig, *eLife*, 2016, **5**, 1–22.
- 24 P. M. Blanco, M. Via, J. L. Garcés, S. Madurga and F. Mas, *Entropy*, 2017, **19**, 105.
- 25 S. Smith and R. Grima, *J. Chem. Phys.*, 2017, **146**, 024105.
- 26 S. Bucciarelli, J. S. Myung, B. Farago, S. Das, G. A. Vliegthart, O. Holderer, R. G. Winkler, P. Schurtenberger, G. Gompper and A. Stradner, *Sci. Adv.*, 2016, **2**, e1601432.
- 27 P. H. Wang, I. Yu, M. Feig and Y. Sugita, *Chem. Phys. Lett.*, 2017, **671**, 63–70.
- 28 T. Ando, E. Chow and J. Skolnick, *J. Chem. Phys.*, 2013, **139**, 121922.



- 29 M. Saxton, *Biophys. J.*, 1994, **66**, 394–401.
- 30 D. L. Ermak and J. A. McCammon, *J. Chem. Phys.*, 1978, **69**, 1352–1360.
- 31 J. Rotne and S. Prager, *J. Chem. Phys.*, 1969, **50**, 4831–4837.
- 32 H. Yamakawa, *J. Chem. Phys.*, 1970, **53**, 436–443.
- 33 J. Brady, *Annu. Rev. Fluid Mech.*, 1988, **20**, 111–157.
- 34 M. Tokuyama and I. Oppenheim, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1994, **50**, 6–9.
- 35 M. Tokuyama, *Phys. A*, 2006, **364**, 23–62.
- 36 M. Tokuyama, T. Moriki and Y. Kimura, *Phys. Rev. E: Stat., Nonlinear, Soft Matter Phys.*, 2011, **83**, 1–8.
- 37 A. Irbäck and S. Mohanty, *Eur. Phys. J.-Spec. Top.*, 2017, **226**, 627–638.
- 38 D. Bedrov, C. Ayyagari and G. D. Smith, *J. Chem. Theory Comput.*, 2006, **2**, 598–606.
- 39 D. Bedrov, G. D. Smith and J. Yoon, *Langmuir*, 2007, **23**, 12032–12041.
- 40 C. N. Likos, H. Löwen, A. Poppe, L. Willner, J. Roovers, B. Cubitt and D. Richter, *Phys. Rev. E: Stat. Phys., Plasmas, Fluids, Relat. Interdiscip. Top.*, 1998, **58**, 6299–6307.
- 41 C. N. Likos, H. Löwen, M. Watzlawek, B. Abbas, O. Jucknischke, J. Allgaier and D. Richter, *Phys. Rev. Lett.*, 1998, **80**, 4450–4453.
- 42 B. Loppinet, G. Fytas, D. Vlassopoulos, C. N. Likos, G. Meier and J. L. Guo, *Macromol. Chem. Phys.*, 2005, **206**, 163–172.
- 43 L. Liu, W. K. den Otter and W. J. Briels, *Soft Matter*, 2014, **10**, 7874–7886.
- 44 E. Locatelli, B. Capone and C. N. Likos, *J. Chem. Phys.*, 2016, **145**, 174901.
- 45 S. T. Milner, T. A. Witten and M. E. Cates, *Macromolecules*, 1988, **21**, 2610–2619.
- 46 J. Mewis, W. J. Frith, T. A. Strivens and W. B. Russel, *AIChE J.*, 1989, **35**, 415–422.
- 47 U. Genz, B. D'Aguanno, J. Mewis and R. Klein, *Langmuir*, 1994, **10**, 2206–2212.
- 48 P. Vilaseca and G. Franzese, *J. Chem. Phys.*, 2010, **133**, 1–11.
- 49 F. Leoni and G. Franzese, *J. Chem. Phys.*, 2014, **141**, 174501.
- 50 P. Vilaseca, K. A. Dawson and G. Franzese, *Soft Matter*, 2012, 6978–6985.
- 51 J. W. Sloan, B. H. Alexander, R. L. Lohmar, I. A. Wolff and C. E. Rist, *J. Am. Chem. Soc.*, 1954, **76**, 4429–4434.
- 52 J. Sabatie, L. Choplin, M. Moan, J. L. Doublier, F. Paul and P. Monsan, *Carbohydr. Polym.*, 1988, **9**, 87–101.
- 53 J. Armstrong, R. Wenby, H. Meiselman and T. Fisher, *Biophys. J.*, 2004, **87**, 4259–4270.
- 54 A. Ortega, D. Amorós and J. G. De La Torre, *Biophys. J.*, 2011, **101**, 892–898.
- 55 M. Fairhead, D. Krndija, E. D. Lowe and M. Howarth, *J. Mol. Biol.*, 2014, **426**, 199–214.
- 56 T. Terai, M. Kohno, G. Boncompain, S. Sugiyama, N. Saito, R. Fujikake, T. Ueno, T. Komatsu, K. Hanaoka, T. Okabe, Y. Urano, F. Perez and T. Nagano, *J. Am. Chem. Soc.*, 2015, **137**, 10464–10467.
- 57 L. Baugh, I. Le Trong, P. S. Stayton, R. E. Stenkamp and T. P. Lybrand, *Biochemistry*, 2016, **55**, 5201–5203.
- 58 D. W. Schaefer, *Polymer*, 1984, **25**, 387–394.
- 59 A. Pähler, W. A. Hendrickson, M. A. G. Kolks, C. E. Argaña and C. R. Cantor, *J. Biol. Chem.*, 1987, **262**, 13933–13937.
- 60 H. Mori, *Prog. Theor. Phys.*, 1965, **33**, 423–455.
- 61 D. T. Gillespie, *Am. J. Phys.*, 1996, **64**, 1246–1257.
- 62 J. K. Dhont, *An Introduction to Dynamics of Colloids*, Elsevier, Amsterdam, The Netherlands, 1996, ch. 2.
- 63 R. J. Phillips, J. F. Brady and G. Bossis, *Phys. Fluids*, 1988, **31**, 3462–3472.
- 64 J. L. Garcés, F. Mas, J. Cecilia, E. Companys, J. Galceran, J. Salvador, J. Puy and J. Galceran, *Phys. Chem. Chem. Phys.*, 2002, **4**, 3764–3773.



Chapter 8

Influence of Macromolecular Crowding in the Charge Regulation of Intrinsically Disordered Proteins

8.1 Summary

8.1.1 Introduction

In chapters 3 to 5, the coupling of conformational and ionization degrees of freedom of flexible weak polyelectrolytes have been analyzed. In particular, it has been proved that a perturbation in the conformational properties of the polyelectrolyte produced by a mechanical force causes in turn a perturbation in its ionization properties. Later on this thesis, in chapters 6 and 7, it has been found that an adequate description of the flexible structure of macromolecules is necessary to describe the dynamic properties of biopolymers on crowded media. This result indicated that the dynamic and conformational of biopolymers in crowded media are coupled, due to effects such as macromolecular entanglement and macromolecular compaction. All these observations come together when considering the case of Intrinsically Disordered Proteins (IDPs) in crowded media.

IDPs are an abundant family of proteins whose principle feature is being highly flexible, therefore lacking of a well-defined stable structure. For this reason, owing to the paradigm of the "lock-and-key" principle,²⁴¹ they were considered to not present any specific biological function.^{242,243} However, recently many biological processes are being discovered in which IDPs play a fundamental role.²⁴⁵ The research of macromolecular crowding impact on the IDPs properties have been focused in their conformational properties, which have been found to be in general dependent on each IDP-crowder combination.^{91-94,96}

IDPs and flexible weak polyelectrolytes share a common feature of having a dynamic structure in solution. Moreover, the conformational and ionization degrees of freedom of IDPs have been found to be also coupled^{36,37}. For this reason, bead-and-spring models,^{16,29,46,47,50,51,134-136,248,249} commonly used for studying weak polyelectrolytes, are starting to be also applied to IDPs. These models have been integrated in Semi-Grand Canonical Monte Carlo (SGCMC) simulations, which have been able to reproduce ellipsometry^{48,49} and X-ray scattering^{250,251} experiments.

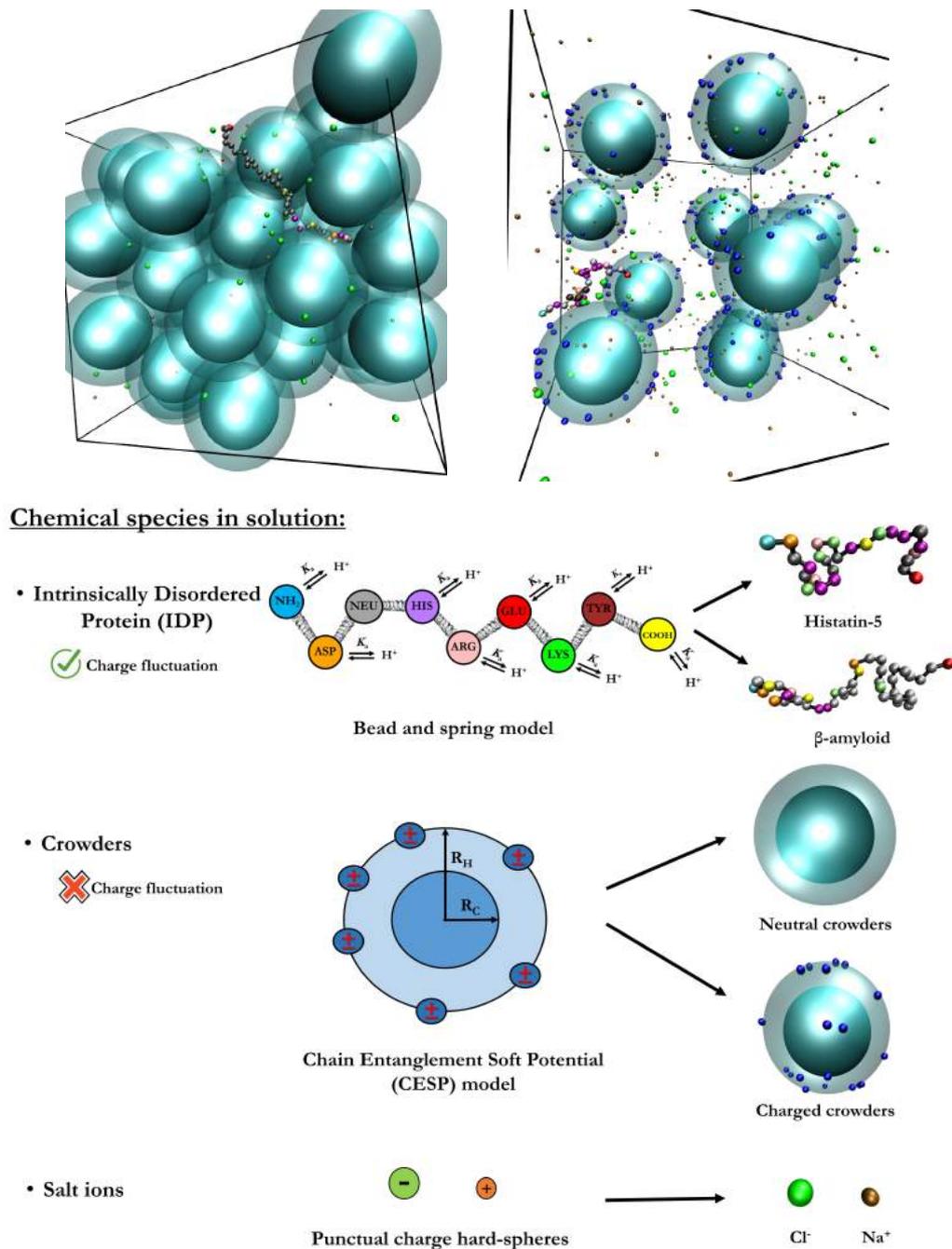


FIGURE 8.1: (Top) Two snapshots taken from Semi-Grand Canonical Monte Carlo (SGCMC) simulations of β -amyloid 42 in solution with neutral crowders at excluded volume fraction $\phi = 0.39$ and (right) histatin-5 in solution with charged crowders at $\phi = 0.08$. (Bottom) General outline of the cases contemplated in this chapter. Two IDPs are studied, histatin-5 and β -amyloid 42, which are modeled with a bead-and-spring model (coloured chain). The charge of their ionizable aminoacid fluctuates, and it is exactly calculated within the SGCMC simulation. The crowders (double-shelled cyan spheres) are modeled with the Chain Entanglement Softened Potential (CESP), developed in the previous chapter, and parametrized to resemble Bovine Serum Albumin (BSA). When charged crowders are considered, hard beads containing one elementary charge are placed in the surface, whose charge is assumed to be fixed. The added salt ions (green and orange spheres) are approximated as hard sphere with an elemental charge in the center, mimicking Na^+ and Cl^- ions.

In this chapter, a possible charge regulation in IDPs produced by macromolecular crowding is investigated. This effect is expected to happen owing to two hypothetical reasons:

- As aforementioned, macromolecular crowding can change the conformational properties of IDPs. Generally, macromolecular crowding produces protein compactation. Since the conformational and ionization degrees of freedom are coupled,^{36,37} the conformational change could induce, in turn, a change in the ionization of the IDPs.
- Adding macromolecular crowding reduces the available volume in solution. For a system with a constant salt concentration, this implies that the same number of ions are in a reduced space, which implies an increase in the effective ionic strength. In case of adding charged crowders, this effect is coupled with adding more charged species in solution, which should cause a more significant increase in the effective ionic increase. These changes in the effective ionic strength could induce a charge regulation of the IDPs, in a similar way that changes in the ionic strength can regulate the charge of weak polyelectrolyte as shown in the previous chapters of this thesis.

These hypothesis are investigated for the particular case of two IDPs: histatin-5 and β -amyloid 42. Histatin-5 is a protein present in human saliva, which anti-fungal and anti-bacterial activity.²⁵²⁻²⁵⁴ β -amyloid 42 to be the main component of the fibers found in the brain of patients with Alzheimer disease.²⁵⁷⁻²⁵⁹ Both IDPs are modeled using a bead-and-spring model, similar to the ones used in previous works.^{48,49,250} The crowders are modeled using the Chain Entanglement Softened Potential (CESP), developed in the previous chapter, which is parametrized for Bovine Serum Albumin (BSA) protein. The salt ions are explicitly included in the system and modeled as hard spheres with a elementary charge in its center. The ionization (average charge, binding capacitance and average charge per aminoacid) and conformational (radius of gyration) of both proteins are computed with SGCMC using MOLSIM software¹³⁹ in (i) absence of crowders, (ii) presence of neutral crowders and (iii) presence of charged crowders. A complete outline of the casuistry studied and the models employed can be found in Fig. 8.1 together with two representative snapshots of the simulated systems.

8.1.2 Results

A multi-layered modeling description is used in this chapter. On the one hand, the IDPs are modeled as a set of beads joined by harmonic springs. Each bead encloses an amino acid of the protein chain and an extra bead is added at each chain end to account for the terminal amino/-carboxylic groups. The beads containing ionizable amino acids have a fluctuating charge, mimicking the proton binding/unbinding. The crowders are modeled as double-shelled spheres, following the Chain Entanglement Softened Potential (CESP) model, and they are parametrized to resemble Bovin Serum Albumin (BSA). When charged crowders are considered, extra beads are randomly distributed in the surface of the crowders, containing one elementary charge per bead. The number of beads on the crowder surface is adapted to match the BSA global charge at a given pH-value, obtained from capillary electrophoresis experiments.²⁶² Charge regulation is only considered in the IDP chains, and the charge of the crowders is assumed to be constant. Na^+ and Cl^- ions are included as added salt, which are modeled as hard spheres with an elementary charge in its center. In this summary, the main results obtained will be discussed showing mainly the results for the systems with histatin-5. The results obtained with β -amyloid 42 are in qualitative agreement with those obtained for histatin-5 and they support the following discussion. More details about the model parametrization, the SGCMC simulations and the graphics corresponding to β -amyloid 42 can be found in the publication enclosed in this chapter.

First, the titration properties (average charge Q and binding capacitance C) of histatin-5 are analyzed in absence of crowding agents, which can be observed in Fig. 8.2 for added salt concentrations c_s ranging, from top to bottom, from 0.5 M to 0 M (no added salt). In Fig. 8.2a, it

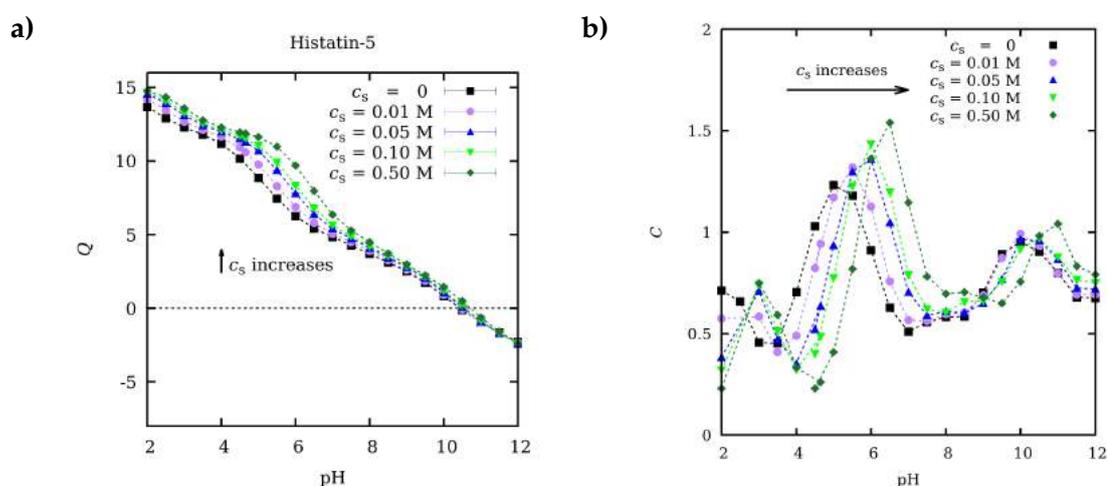


FIGURE 8.2: (a) average charge Q and (b) binding capacitance C titration curves for histatin-5. The concentration of added salt are: no added salt (squares), $c_s = 0.01$ M (circles), 0.05 M (triangles), 0.1 M (inverted triangles) and 0.5 M (diamonds).

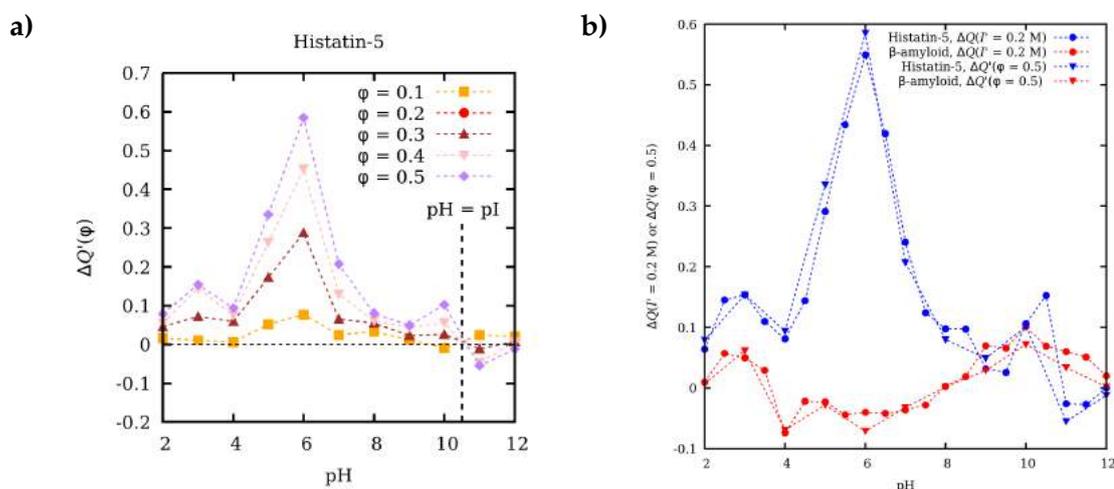


FIGURE 8.3: (a) Charge variation of histatin-5 due to the presence of excluded volume ϕ of neutral crowders, $\Delta Q'(\phi) = |Q(\phi)| - |Q(0)|$. The results are obtained from SGCMC simulations at an added salt concentration $c_s = 0.1 \text{ M}$ of NaCl, at ϕ values ranging from 0.1 to 0.5. (b) Comparison between: i) charge variation produced when adding neutral crowders with resulting $\phi = 0.5$, $\Delta Q'(\phi = 0.5)$; and ii) charge variation produced when the added salt concentration is incremented from $c_s = 0.1 \text{ M}$ to 0.2 M without crowders $\Delta Q'(I' = 0.2 \text{ M})$. The meaning of the markers and color of the lines used can be read inside the panels.

can be observed that histatin-5 behaves similarly to a polycation with Q ranging from +15 to -2 elementary charges and an isoelectric point of $pI \sim 10.5$. The pI obtained for histatin-5 is in good agreement with previous theoretical and computational calculations^{276–279}. For a given pH-value, increasing the c_s -value produces an increase in Q due to the salt screening of the intramolecular repulsion. This effect is more significant near $\text{pH} \sim 5.5$, which coincides with the pH region of which C exhibits maximum values, as can be observed in Fig. 8.2b. It can be also noted that increasing c_s results in a shift to the right in the titration curve of C . Both observations are in good agreement with the titration behavior of weak polyelectrolytes, already discussed in chapter 5. Moreover, the titration curves for Q and C are in good agreement with previous computational works.⁴⁸

Second, the effect of adding neutral crowders in the Q of histatin-5 is assessed for a constant added salt concentration of 0.1 M . This is done computing the difference between the absolute Q -value with and without crowding $\Delta Q'(\phi) = |Q(\phi)| - |Q(0)|$ where $Q(\phi)$ is the average charge of the protein at a given crowder excluded volume ϕ value. $\Delta Q'(\phi)$ as function of the pH-value is presented in Fig. 8.3a, for systems with ϕ values ranging from 0.1 to 0.5, from bottom to top. It can be observed that, for a given pH-value, increasing ϕ provokes an increase in the charge of the protein. The pH-values where this effect is more significant coincide with those where the titration curve of C exhibits maxima values, as can be observed in 8.2b. This is in good agreement with the previous finding obtained in chapter 5 for weak polyelectrolytes: the polyelectrolyte ionization is more affected by the mechanical stretching at the pH and ionic

conditions were the binding capacitance is high, *i.e.* charge fluctuations are more important. In order to validate/refuse the first hypothesis presented in the previous section, the case with $\phi = 0.5$ is analyzed. For this case, the effective ionic strength is $I' = 0.2$ M, the double than the initial added salt concentration $c_s = 0.1$ M. If the charge variation caused by macromolecular crowding is due to the increase of the effective ionic increase, then the same charge variation should be achieved changing the added salt in the system without crowding. This latter variation is measured as $\Delta Q(I' = 0.2\text{M}) = |Q(\phi = 0, c_s = 0.2\text{M})| - |Q(\phi = 0, c_s = 0.1\text{M})|$. Both magnitudes, $\Delta Q(I' = 0.2\text{M})$ (circles) and $\Delta Q'(0.5)$ (triangles), are compared in Fig. 8.3b in the full pH-range for histatin-5 (red) and β -amyloid (blue). Interestingly, it can be observed that both magnitudes are approximately equal in the full pH-range. This result indicates that the increase in the effective ionic strength is the main responsible of the charge regulation in the IDPs due to macromolecular crowding. Conversely, the structural changes due to macromolecular crowding is not observed to have significant influence in the charge regulation of the IDPs. This is because macromolecular crowding is not significantly changing the size of the IDPs, as it will be discussed later.

Finally, the charge regulation of the IDPs produced by the addition of charged crowders is studied. The $\Delta Q'(\phi)$ values obtained as a function of the pH-value for systems with ϕ values ranging from 0.05 to 0.2, from bottom to top, are shown in Fig. 8.4a. The values here shown

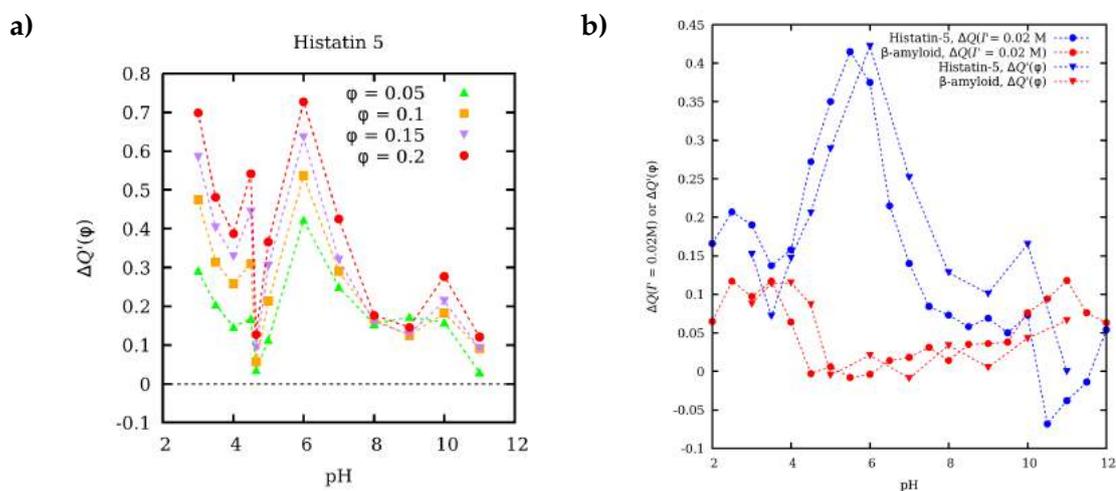


FIGURE 8.4: (a) Charge variation of histatin-5 due to the presence of excluded volume ϕ of charged crowders, $\Delta Q'(\phi) = |Q(\phi)| - |Q(0)|$. The results are obtained from SGCMC simulations at an added salt concentration $c_s = 0.01$ M of NaCl, at ϕ values ranging from 0.05 to 0.2. (b) Comparison between: i) charge variation produced when adding charged crowders, $\Delta Q'(\phi)$, where ϕ is chosen to ensure $I' = 0.02$ M; and (ii) charge variation produced when the added salt concentration is incremented from $c_s = 0.01$ M to 0.02 M without crowders $\Delta Q(I' = 0.02\text{M})$. The meaning of the markers and color of the lines used can be read inside the panels.

correspond to those obtained from SGCMC simulation of histatin-5 with an added salt concentration $c_s = 0.01$ M of NaCl. It can be observed that, again, for a given pH-value, the charge of histatin-5 increases when ϕ increases. For a given ϕ value, this effect is observed to be larger with charged crowder than in the neutral crowder case. For instance, for $\phi = 0.2$ and pH = 6, $\Delta Q'(0.1) \sim 0.1$ for the neutral crowder and $\Delta Q'(0.1) \sim 0.5$ for the charged crowder. This is because in adding charged crowders, one is adding more charged species to the system (the crowders and their counter-ions), which increases the effective ionic strength more than in the neutral crowder case. In order to analyze this latter point, $\Delta Q(I' = 0.02\text{M})$ (circles) and $\Delta Q'(\phi)$ (triangles), are compared in Fig. 8.4b in the full pH-range for histatin-5 (red) and β -amyloid (blue). Note that for the charged crowders, the ϕ value which produces $I' = 0.02\text{M}$ is different for each pH value, since the charge of the crowder varies with the pH. More details about how to compute I' in this case can be found in the publication presented at the end of the chapter, where an equation to calculate it is provided (Eq. 18). It can be observed that $\Delta Q(I' = 0.02\text{M})$ and $\Delta Q'(\phi)$ agree reasonably well for all the pH-range, indicating that in this case the increment in the effective ionic strength is also the main contribution to the IDP charge regulation. However, a shift can be observed between both sets of data, which suggests a possible contribution of the electrostatic interaction of the IDP with the charged crowders.

Finally, the impact of macromolecular crowding on the radius of gyration (R_g) of histatin-5 are presented in Figs. 8.5a and 8.5b as a function of ϕ for neutral and charged crowders, respectively. Different curves are shown obtained from SGCMC simulations at pH values ranging

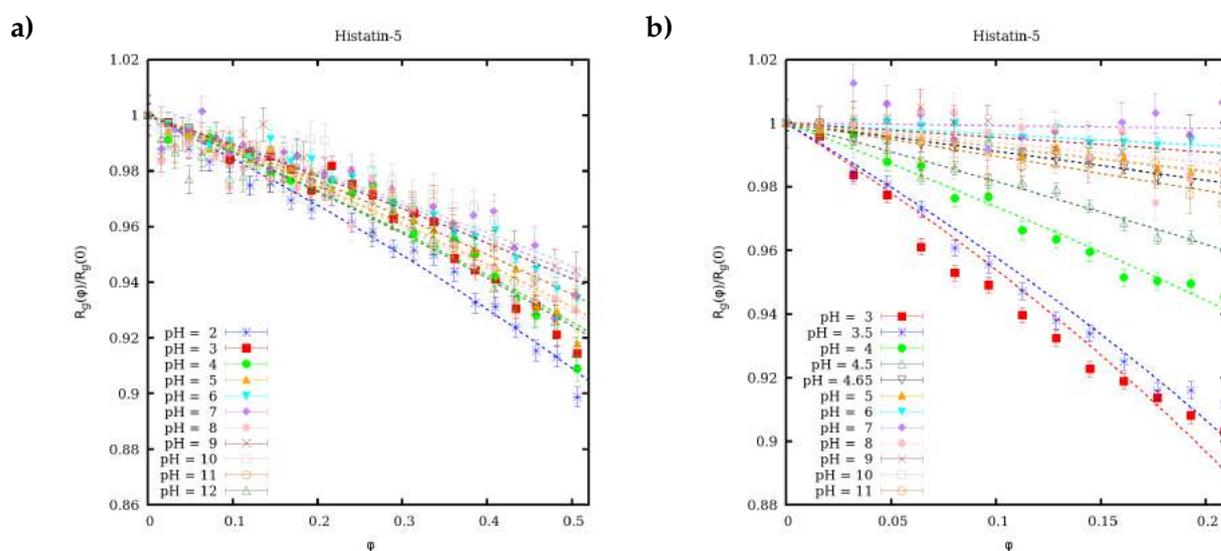


FIGURE 8.5: Histatin-5 radius of gyration R_g (normalized to its value in absence of crowding agents $R_g(0)$) as a function of ϕ . The panels show the cases with (a) neutral crowders and (b) charged crowders. Dashed lines are the best fit of the scaling law^{282,283} to the SGCMC data (Eq. 17 in the publication at the end of the chapter).

from 2 to 12. R_g is always normalized to the radius of gyration at the corresponding pH without crowding agents $R_g(0)$. As a general trend, it can be observed that R_g decays in increasing the ϕ value. However, the reduction of the IDP size is rather small, reaching a maximum reduction of about a 10 % at most. This is probably the reason why the impact of macromolecular crowding on the IDP structure is not affecting its charge. The R_g vs. ϕ curves are observed to be well-fitted to a scaling law proposed in earlier works^{282,283} (Eq. 17 in the publication at the end of the chapter). The best fits of this scaling law to the simulation data are displayed as dashed lines in in Figs. 8.5a and 8.5b.

8.1.3 Conclusions

The possibility of charge regulation induced by macromolecular crowding is investigated in two IDPs, histatin-5 and β -amyloid 42, using Semi-Grand Canonical Monte Carlo simulations. The IDPs are modeled using a bead-and-spring model whereas the crowders are modeled using the Chain Entanglement Softened Potential (CESP), developed in the previous chapter. The ionization (average global charge, average charge per aminoacid and binding capacitance) and structural (radius of gyration) of both proteins are analyzed in different conditions: (i) in absence of crowding, (ii) with neutral crowders and (iii) with charged crowders.

In absence of crowders, the titration curve of histatin-5 is found to be highly dependent on the added salt concentration while the β -amyloid titration curve is found to nearly not affected by the salt. The pH-values where the salt effect is more significant in both proteins coincide with those where their binding capacitance exhibits maxima values. This is in good agreement with the behavior observed for weak polyelectrolytes in chapter 5. Both proteins are observed to fold when increasing the salt concentration (smaller radius of gyration) at pH-values far from their isoelectric point. This is because at those pH-values the proteins are highly charged and the salt screens the intra-molecular repulsive interactions.

In presence of crowders, significant variations of the charge of the IDP respect its value in absence of crowders are observed. These charge variations are driven by the increase in the effective ionic strength (same ions in a reduced space) caused by the crowders. In the case of charged crowders, a possible contribution of the interaction between the charged crowder and the IDP is also observed. Macromolecular crowding is observed to produce larger charge variations in the IDP at pH-values close to its binding capacitance maxima. In chapter 5, we observed that charge regulation induced by mechanical stretching is also stronger at conditions of maximum binding capacitance. Both findings highlight the utility of the binding capacitance to assess the conditions where the charge of the system is more susceptible to be altered. The analysis of the average charge per aminoacid reveals that the charge regulation of the IDPs can

be mediated in some cases by neutralizing aminoacid with opposite charge than the IDP global charge.

The radius of gyration of the proteins is found to slightly decay when the excluded volume of crowders is increased, for both neutral and charged crowders. The results are found to be well fitted to a scaling law proposed in previous works.^{282,283} The compactation of the IDP due to macromolecular crowding is not found to have any significant effect in the charge of the IDP.

8.2 Publication

Journal Name

ARTICLE TYPE

Cite this: DOI: 00.0000/xxxxxxxxxx

Influence of Macromolecular Crowding in the Charge Regulation of Intrinsically Disordered Proteins[†]Pablo M. Blanco,^{a,*} Sergio Madurga,^a Josep L. Garcés,^b Francesc Mas^{a,*} and Rita S. Dias^{c,*}

Received Date

Accepted Date

DOI: 00.0000/xxxxxxxxxx

In this work we study the coupling between ionization and conformational properties of two IDPs, histatin-5 and β -amyloid 42, in the presence of neutral and charged crowders. The latter are modeled to resemble Bovin Serum Albumin (BSA). With this aim, Semi-Grand Canonical Monte Carlo simulations are performed, so that the IDPs charge is a dynamic property, undergoing protonation/deprotonation processes. Both ionization properties (global and specific aminoacid charge, binding capacitance) and radius of gyration are analyzed in a large range of pH values and salt concentrations. Without crowder agents, the titration curve of histatin-5, a polycation, is salt-dependent while that of β -amyloid 42, a polyampholyte, is almost unaffected. The salt concentration is found to be specially relevant at pH values where the protein binding capacitance (directly linked with charge fluctuation) is larger. Upon addition of neutral crowders, charge regulation is observed in histatin-5, while for β -amyloid 42 this effect is very small. The main mechanism for charge regulation is found to be the effective increase in the ionic strength due to excluded volume. In presence of charged crowders, significant increase in the charge of both IDPs is observed at almost all the pH range. In this case, the IDP charge is altered not only by the increase in the effective ionic strength but also by its direct electrostatic interaction with the charged crowders.

1 Introduction

For many years, the basis of enzymatic biochemistry were laid following the classic “lock-and-key” principle, which relates the specific conformation of an enzyme with its biological function.¹ Un-

der this paradigm, and despite their abundance in all proteomes, proteins with partially or entirely disordered sequences, the so-called Intrinsically Disordered Proteins (IDPs), were considered to have no specific biologic role or simply that of connecting ordered domains^{2,3}. Nowadays, however, the characteristic structural flexibility of IDPs has been found to exhibit many biological advantages such as multi-binding, the capability to avoid non-desired interactions through functional misfolding,⁴ or overcoming steric restrictions due to macromolecular crowding⁵. Due to these advantages, IDPs are involved in a wide spectra of biological processes⁶, such as signalling processes, supramolecular assembly or allosteric regulation, so that they are appealing targets for drugs.

On the other side, biological media are highly concentrated in macromolecular species, the so-called macromolecular crowding^{7,8}. For instance, the weight fraction of protein in blood plasma is around 9%, with serum albumin being the most abundant protein with concentration around 35-55 g/L⁹. In the cell cytosol, macromolecules occupy a volume fraction of 20%-40%, with approximate macromolecule concentrations 200–400 g/L. The impact of macromolecular crowding in proteins with a defined 3D structure has been extensively studied; it has been found to substantially affect the reactivity, conformational properties,

^a Physical Chemistry Unit, Materials Science and Physical Chemistry Department & Research Institute of Theoretical and Computational Chemistry (IQTCUB) of Barcelona University (UB), Barcelona (Catalonia, Spain) E-mail: pmbianco23@gmail.com; fmas@ub.edu

^b Chemistry Department, Technical School of Agricultural Engineering & AGROTECNIO of Lleida University (UdL), Lleida (Catalonia, Spain).

^c Department of Physics, Norwegian University of Science and Technology (NTNU), Trondheim (Norway) E-mail: rita.dias@ntnu.no

[†] Electronic Supplementary Information (ESI) available: Additional graphics that can be found as part of the supporting information: (S1) Charge and binding capacitance per aminoacid vs. pH in dilute solution; (S2) Variation of the charge of aspartic acid, histidine and lysine due to the added salt vs. pH in dilute solution; (S3) Charge variation vs. pH for the case with neutral crowders with salt concentration 0.01 M; (S4) Variation of the IDP global charge vs. excluded volume (S5) Binding capacitance per aminoacid vs. pH for the case with neutral crowders with salt concentration 0.1 M; (S6) Best fit parameters to the scaling law (Eq. 13) for the R_g vs. ϕ curves (S7) Binding capacitance per aminoacid vs. pH for the case with charged crowders with salt concentration 0.01 M; (S8) Calculations with effective ionic strength I' for the case with charged crowders; and (S9) Comparison of $\Delta Q(c_s)$ and $\Delta Q'(\phi)$ vs. pH for systems with $I' = 0.05$ M.

diffusion processes and self-assembly^{7,8,10–22}. In the case of IDPs, the latter point is specially relevant to unravel their physiological role and to prevent the pathways that lead to unwanted self-association or association of these proteins with other macromolecules.

Research on the response of IDPs to macromolecular crowding has been mainly focused on their conformational properties. Macromolecular crowding effect in the conformational properties of IDPs has been found to be rather heterogeneous and, in general, it is dependent on each IDP-crowder combination^{23–27}. Therefore, IDPs have been classified as foldable (fold upon crowding), unfoldable (extended upon crowding) and non-foldable (mostly unaffected)⁵. This variability suggests that the impact of molecular crowding on IDPs is not only due to the excluded volume effect, but that other variables, such as the macromolecular charge, could also play a key role. For example, the presence of crowders may alter the conformation of the protein, which in turn affects the ionization state of the weak acid/base groups of the protein.

Coupling between conformational and ionization properties has been reported to be significant in IDPs^{28,29} and it is a rather general trend of weak polyelectrolytes^{30–36}. Some examples are the swelling of poly(methacrylic) acid in a narrow range of pH³⁷, the important influence of ionization in protein folding³⁸ or the helix–coil transitions of poly(peptides)³⁹. The ability of a macromolecule to modulate its charge as a response to an external stimulus (e.g. pH variations, salt concentration changes, interactions with other macromolecules or external fields) is known as Charge Regulation^{40,41}. This capability has led to the use of weak polyelectrolytes in pH-responsive smart systems with multiple technological applications (see, for instance, Ref.⁴² and references quoted therein).

Moreover, both IDPs and weak polyelectrolytes possess a very flexible structure so that simulation techniques borrowed from weak polyelectrolytes can be straightforwardly used in modeling IDPs. In particular, “bead-and-spring” models^{32,34–36,43–49} have been successfully used to simulate IDPs^{50–52}. Using Monte Carlo simulations of this nature, Skepö *et al.* have been able to reproduce small-angle X-ray scattering^{52,53} and ellipsometry^{50,51} experiments. Furthermore, in both cases ionization equilibria can be explicitly taken into account by using Monte Carlo simulations in the Semi-Grand Canonical Monte Carlo ensemble^{32,34–36,54,55}.

In this work we will focus on the effect of macromolecular crowding on charge regulation and conformational structure of two IDPs with very different biological functions: histatin-5 and β -amyloid 42. Histatin-5 is present in human saliva^{50–53} and presents antibacterial⁵⁶, antifungal^{57,58}, and polyphenolic-binding^{59,60} properties. β -amyloid 42 is known for its ability to self-assemble, forming fibres in the brain of patients with Alzheimer disease^{61–63}. This article is organised as follows. In section two, the theoretical model, its parameterization and the computational details are introduced. Following a step-wise approach, we analyse in section three the ionization and conformational properties of the IDPs in the (i) absence of crowding agents; (ii) presence of neutral crowders, and (iii) presence of charged crowders.

2 Methodology

Histatin-5 and β -amyloid 42 present a rather different ionization properties: Histatin-5 is a polycation while β -amyloid 42 is a polyampholyte. In this work they are modelled using a simple bead-and-spring model. The conformational and ionization properties are computed by means of Semi-Grand Canonical Monte Carlo simulations using MOLSIM software⁶⁴. The crowders are described by means of the recently developed Chain Entanglement Softened Potential (CESP) model²¹. Within this model, they are regarded as double-shelled spheres with an inner hard-core and an outer soft shell. In order to evaluate the effect of the crowder agents, we perform simulations with charged and uncharged crowders. When considered, the crowders charge is fixed at a given pH, and modelled by placing charged beads at the sphere surface. Therefore, only the IDP charge is allowed to fluctuate. In Fig. 1, two examples of the performed simulations are shown. In the left upper panel we have depicted β -amyloid 42 in a solution of neutral crowders (cyan double-shelled spheres) occupying a large fraction of the volume. In the right-hand side upper panel the solution contains histatin-5 and charged crowders (cyan double-shelled spheres with blue beads in its surface representing the charges). The chemical species in solution, as well as the different models used, are outlined in the bottom panel of Fig. 1. Let us discuss in detail the methods used in this work.

2.1 Theoretical model

For simplicity, let us consider the case in which only one IDP chain is present in a medium populated by crowder macromolecules and salt ions. The atomistic modeling of this medium is computationally challenging due to the massive quantity of atoms in solution and the physicochemical processes involved. Our model aims to reduce the system complexity while retaining the most relevant physicochemical aspects (electrostatic interactions and steric hinderance) involved in the acid/base equilibria of the IDP. We make use of a multi-layered coarse-grained description with different degrees of detail depending on the chemical identity of each component, as outlined in Fig. 1.

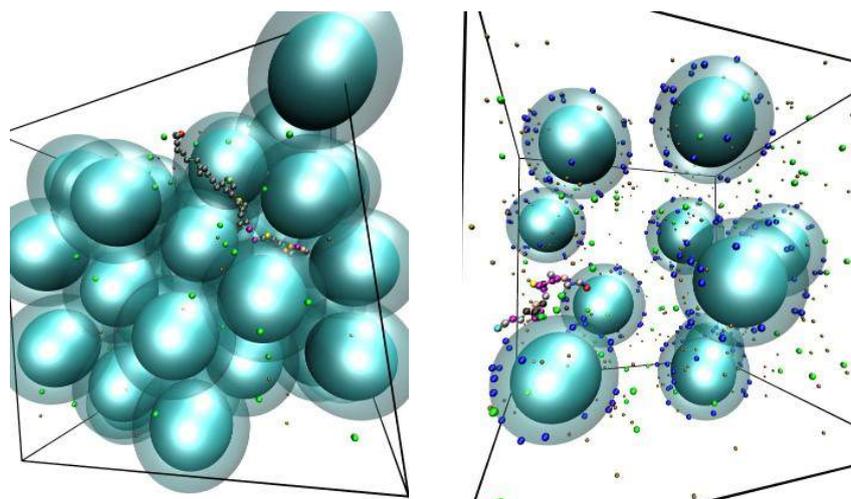
The IDP chain is described as a set of M beads joined by elastic bonds. The bonds are approximated as springs with a harmonic potential

$$F_B = \frac{k_{\text{Bond}}}{2} \sum_{j=1}^{M-1} (l_j - l_0)^2, \quad (1)$$

where l_j is the length of bond j , l_0 is the bond equilibrium length and k_{Bond} is the harmonic force constant. Each bead represents an aminoacid residue and two extra beads are used to describe the terminal amino/carboxylic groups, as depicted in Fig. 1. The beads holding an acid/base group are allowed to change their ionization state, mimicking proton binding/unbinding. For a protein chain with S acid/basic beads, the protonation free energy F_p reads^{32,34–36,67}

$$\frac{\beta F_p}{\ln(10)} = \sum_{i=1}^S \mu_i s_i = \sum_{i=1}^S (\text{pH} - \text{p}K_{a,i}) s_i, \quad (2)$$

where $\text{pH} = -\log(a_{\text{H}})$ depends on the proton activity a_{H} , $\beta =$



Chemical species in solution:

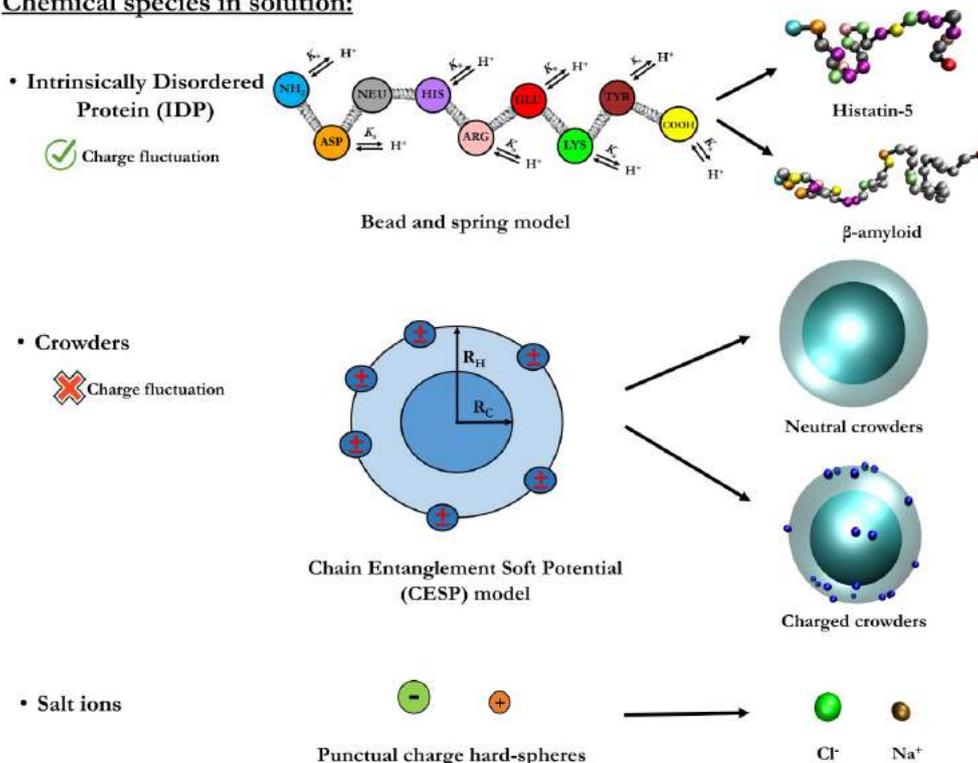


Fig. 1 Upper panel: two snapshots of Semi-Grand Canonical Monte Carlo simulations of: (left) β -amyloid 42 in a medium containing neutral crowders with excluded volume fraction $\phi = 0.39$ and (right) histatin-5 surrounded by charged crowders with $\phi = 0.08$. Bottom panel: Outline of the chemical species in solution. Both IDPs (coloured chain) are modeled by means of a bead and spring model^{50–52}, with acid/base equilibria explicitly taken into account. The crowders (cyan double-shelled spheres) are modeled using the Chain Entanglement Softened Potential (CESP) model²¹, for which the crowders have an inner dense core (opaque) and an outer softer region (transparent). In the case of charged crowders, small charged beads (blue) are added mimicking the Bovin Serum Albumin surface charge. Na⁺ (orange spheres) and Cl⁻ (green spheres) ions are also present with concentration of 0.01 M.

Particle	-ARG-	-LYS-	-TYR-	-HIS-	-GLU-	-ASP-	-NH ₂	-COOH
$pK_{a,i}$	12.0	10.4	9.6	6.3	4.4	4.0	7.5	3.8

Table 1 pK_a -values for the titrable aminoacid residues and for the terminal amino (-NH₂) and carboxylic (-COOH) groups, obtained from Ref.⁶⁶

Histatin-5 sequence: (24 aminoacids: 15 basic and 5 acid groups)

NH₂ – ASP – SER – HIS – ALA – LYS – ARG – HIS – HIS – GLY – TYR – LYS – ARG – LYS –
– PHE – HIS – GLU – LYS – HIS – HIS – SER – HIS – ARG – GLY – TYR – COOH

 β -amyloid 42 sequence: (42 aminoacids: 7 basic and 8 acid groups)

NH₂ – ASP – ALA – GLU – PHE – ARG – HIS – ASP – SER – GLY – TYR – GLU – VAL – HIS –
– HIS – GLN – LYS – LEU – VAL – PHE – PHE – ALA – GLU – ASP – VAL – GLY – SER –
– ASN – LYS – GLY – ALA – ILE – ILE – GLY – LEU – MET – VAL – GLY – GLY – VAL –
– VAL – ILE – ALA – COOH

Fig. 2 Aminoacid sequence of Histatin-5^{50,65} and β -amyloid 42⁶¹. The amino acids with basic and acid groups are coloured in blue and red, respectively.

$1/(k_B T)$ is the inverse of the thermal energy with k_B the Boltzmann constant and T the temperature. For a given ionizable group i , μ_i and $pK_{a,i}$ represent its reduced chemical potential and its acid constant, respectively. s_i is a variable representing its protonation state, whose value can be either 1 (protonated) or 0 (unprotonated).

The crowders are modelled using the recently developed Chain Entanglement Softened Potential (CESP) model²¹. In this model, the particles interact through an empiric inter-particle interaction potential containing a continuous shouldered well, similar to that used in computational models of coarse-grained fluids^{68,69}. The particle is divided into two distinct regions: (i) the core, which is densely populated by atoms and modeled as a hard-sphere of radius R_C , and (ii) the outer part, in contact with the solvent, which is less dense and can be penetrated by other particles with energetic cost U_r . Its radius is estimated as the hydrodynamic radius of the macromolecule R_H , as depicted in Fig. 1. The mean-field pair-wise potential, F_{CESP} , corresponding to this model, is given by the empirical expression

$$F_{CESP} = \sum_{i,j}^N \epsilon_0 \left(\frac{d_C}{d_{ij}} \right)^{24} - \frac{U_r}{2} \tanh \left(\frac{d_C/d_0}{d_E - d_C} \left(d_{ij} - \frac{d_E + d_C}{2} \right) \right) + \frac{U_r}{2} \quad (3)$$

where d_{ij} is the distance between the centers of the interacting particles i and j . $\epsilon_0 = 1 \text{ J mol}^{-1}$ and $d_0 = 1 \text{ nm}$ are added to correct the units²¹ while $d_C = R_{C,i} + R_{C,j}$ and $d_H = R_{H,i} + R_{H,j}$ represent the sum of hard core and hydrodynamic radii of particles i and j , respectively. Conversely, the steric repulsion between protein beads, ions and between beads and ions is calculated only using the first, repulsive, term in Eq. 3.

The electrostatic interaction in the implicit solvent assumption obeys the Coulomb potential

$$\beta F_C = \sum_{i=1}^N \sum_{j=i+1}^N \frac{\ell_B}{d_{ij}} q_i q_j \quad (4)$$

where ℓ_B is the Bjerrum length, which is $\ell_B \simeq 0.71 \text{ nm}$ at 298 K.

The total free energy of the system reads

$$F = F_B + F_P + F_{CESP} + F_C. \quad (5)$$

2.2 Parameterization

The aminoacid sequences of histatin-5 and β -amyloid 42 can be found in Fig. 2. Both IDPs are short peptide chains with a flexible random-coil structure but they exhibit very different acid/base properties. Histatin-5 consists of a large number of basic aminoacids (mainly histidine), so it is positively charged in a wide range of pH-values. β -amyloid 42, on the other hand, presents almost the same number of acid and basic aminoacids, so its charge undergoes a transition from positive at low pH-values to negative at high pH-values. Inspired by the model suggested by Skepö *et al.*, we describe aminoacid residues as beads with hard-core radius $R_C = 0.2 \text{ nm}$ ^{50–53}. There are also two extra beads for the terminal amino/carboxylic groups with the same radius $R_C = 0.2 \text{ nm}$. The aminoacids hard-core and hydrodynamic radii are taken to be equal. The beads are joined by harmonic bonds with equilibrium bond length $l_0 = 0.41 \text{ nm}$ and force constant $k_{\text{Bond}} = 240 \text{ kJ mol}^{-1} \text{ nm}^{-2}$ or 0.4 N m^{-1} . The intrinsic acid constants ($pK_{a,i}$) of the aminoacids are taken from Ref. ⁶⁶ and can be also found in Tab. 1.

Aiming to mimic a biologically realistic scenario, the crowders have been chosen to resemble Bovine Serum Albumin (BSA). We have chosen BSA instead of its human variant because it has been more extensively used in research and it is thus better parameterized. Since we are also interested in the effect of the crowder charge, we firstly study the simpler situation of neutral crowders with the same dimensions as BSA. The hard-core radius R_C can be estimated as²¹

$$R_C(\text{BSA}) = \sqrt[3]{\frac{3\nu M_W}{4\pi N_A}} = 2.73 \text{ nm}, \quad (6)$$

where $\nu = 0.736 \text{ cm}^3 \text{ g}^{-1}$ and $M_W = 70 \text{ kg mol}^{-1}$ are the specific volume⁷¹ and the molecular weight⁷² of BSA, respectively, and N_A is the Avogadro constant. R_H is calculated using the Stokes-

pH	3.00	3.50	4.00	4.50	4.65	5.00	6.00	7.00	8.00	9.00	10.0	11.0
Net charge/ <i>e</i>	+22	+18	+10	+6	0	-4	-9	-15	-19	-22	-26	-27

Table 2 Average net charge of Bovine Serum Albumin (BSA) at different pH values. The values were obtained from capillary electrophoresis by measuring the electrophoretic mobility of BSA⁷⁰.

Einstein equation

$$R_H(\text{BSA}) = \frac{k_B T}{6\pi\eta D} = 3.64 \text{ nm}, \quad (7)$$

where $\eta = 0.932 \text{ mPa}$ is water viscosity at room temperature⁷³. $D = 6.32 \cdot 10^{-2} \text{ nm}^2\text{ns}^{-1}$ represents the BSA translational diffusion coefficient at 293 K, obtained from dynamic light scattering⁷⁴ and quasi-elastic light scattering⁷⁵ experiments at pH = 5, near the isoelectric point of BSA (pH \simeq 4.65)⁷⁰. Although the diffusion coefficient of BSA slightly varies with the pH-value⁷⁴ (which leads to different R_H values ranging from 4.29 nm for pH = 3 to 3.53 nm for pH = 7), we will consider, for simplicity, the size of BSA to be pH-independent. The entanglement energetic cost U_r has been determined in a recent publication²¹ to be $U_r = 5000 \text{ J mol}^{-1}$ by fitting brownian dynamic simulations using CESP potential (Eq. 3) to experimental long-time diffusion coefficients. This value prevents crowders to overlap at low concentrations while it allows crowder-particle inter-crossing at high crowder concentrations.

In the simulations with charged crowders, we have to account for the fact that the net charge of the BSA is sensitive to the pH of the medium. One could in principle treat explicitly the ionization equilibrium of BSA and the consequent charge fluctuations. However, this turns to be computationally very expensive. We have thus decided to assign charge to the charged crowders according to the solution pH and keep it constant during the simulations, i.e., charge regulation is included in the IDPs but not in charged crowders. The net charge of BSA for different pH values was obtained from capillary electrophoresis experimental data⁷⁰ and can be found in Tab. 2. It is distributed in hard-sphere beads with $R_C = 0.2 \text{ nm}$ containing one elementary charge, which are randomly placed at the surface of the crowders (at a distance R_H from the center, see Fig. 1). Previous works have pointed out the relevance of the charge distribution in protein-polyelectrolyte interactions^{76–80}. This is not the case of our system, where preliminary calculations showed no significant effect of the charge distribution of the crowders in the conformational and ionization properties of the IDPs. This difference is possibly due to the fact that polyelectrolytes, in general, have larger charge densities than IDPs. For a given pH-value, all the crowders are considered to have equal charge distribution. In order to avoid anisotropy problems, the crowders are allowed to rotate during the simulation.

Finally, the electroneutrality of the solution is ensured by adding the appropriate amount of sodium Na^+ and chloride Cl^- ions, also modeled as hard spheres with radii equal to its ionic radii: $R_C(\text{Na}^+) = 0.10 \text{ nm}$ and $R_C(\text{Cl}^-) = 0.18 \text{ nm}$ ^{81–83}.

2.3 Computational details

Semi-Grand Canonical Monte Carlo simulations^{32,34–36,54,55} were performed using MOLSIM v6.4.7 software⁶⁴. An additional routine with the CESP contribution to the free energy was added to the original MOLSIM code. The routine was tested by reproducing the results presented in Ref.²¹. The simulations were performed at room temperature $T = 298.15 \text{ K}$ in a cubic simulation box with a side length $L = 22 \text{ nm}$. L was tested and found to be large enough to prevent finite-size artifacts. Periodic boundary conditions were applied in all space directions. The long-range electrostatic interactions are corrected with standard Ewald sums with a cutoff in real space $R_{\text{cut}} = L/2 = 11 \text{ nm}$ and reduced Ewald splitting parameter $\alpha^* = 3.0$ with error in F_C of around $10^{-6} \text{ kJ mol}^{-1}$ ⁸⁴. The simulations consisted of 10^6 to 10^7 MC steps, with thermal equilibration of 10^5 MC steps. Typical runs lasted from 1 to 8 days using the Idun cluster (Norwegian University of Science and Technology, NTNU)⁸⁵, depending on the total particle number, which ranged from 50 to 1300. In each MC step, the following trial moves are performed: (i) translation of a particle (protein beads, crowder particles, ions); (ii) translation of the IDP chain as a whole; (iii) pivot rotation of a chain section of the IDP chain; (iv) slithering of the IDP chain, where a section of the chain is moved to the other end of the chain, mimicking the reptation of a snake; (v) protonation/deprotonation of an acid/basic bead of the IDP chain. The charge of a counter-ion is accordingly changed to maintain the system electroneutrality; and (vi) rotation of the crowders. The increment in the free energy is computed and the standard Metropolis algorithm applied: the trial configuration is always accepted if $\Delta F \leq 0$ and it is accepted with a probability equal to $\exp(-\beta\Delta F)$ otherwise.

The sought thermal averages were the average charge of a specific residue type A , $\langle q_A \rangle$; the binding capacitance of A ^{36,86,87}, $\langle c_A \rangle$, defined as

$$\langle c_A \rangle = \left(\frac{\partial q_A}{\partial \mu_A} \right) = \langle q_A^2 \rangle - \langle q_A \rangle^2, \quad (8)$$

the protein average charge Q

$$Q = \sum_A^G N_A \langle q_A \rangle, \quad (9)$$

where N_A is the number of ionizable groups of type A among the G different types of acid/base groups; the global IDP binding capacitance, given by^{36,86,87}

$$C = \frac{\partial Q}{\partial \mu} = \sum_A^G N_A \left(\frac{\partial \langle q_A \rangle}{\partial \mu} \right) = \sum_A^G N_A c_A. \quad (10)$$

Finally, we also computed the radius of gyration R_g , which reads

$$R_g^2 = \sum_{i,j}^M \frac{1}{2M^2} (r_i - r_j)^2. \quad (11)$$

where r_i is the position of bead i and M is the number of beads of the IDP.

3 Results and discussion

3.1 Histatin-5 and β -amyloid titration in absence of crowders

Let us firstly discuss the ionization and conformational properties of histatin-5 and β -amyloid 42 in absence of crowder agents. The average protein charge Q as a function of the pH-value is shown in Figs. 3a (histatin-5) and 3b (β -amyloid 42) at added salt concentrations ranging from $c_s = 0$ (no added salt) to $c_s = 0.5$ M. Since we only add 1:1 salts, c_s can be directly identified with the ionic strength I . Note that in calculating I in this way the concentration of ions H^+ and OH^- coming from water self-ionization, which could be significant at very low and high pH-values, are neglected. However, as it will be seen, at these pH-values the influence of the ionic strength is very weak, so that we can identify I as the added salt concentration quite accurately.

It can be observed that histatin-5 essentially behaves as a polyelectrolyte with average charge ranging from +15 to -2 and an isoelectric point pI close to 10.5, due to the presence of 15 basic but only 5 acidic groups. The pI -value of Histatin-5 is in good agreement with previous theoretical and computational calculations^{88–91} Conversely, β -amyloid 42 contains eight acidic and seven basic groups, and exhibits an amphoteric behaviour. Its charge ranges from +7 to -8 with $pI \sim 5$, which is consistent with the observed self-aggregation of this protein at pH values close to 5⁹². The obtained histatin-5 titration curves agree reasonably well with previous work⁵⁰, where the average charge was computed using Semi-Grand Canonical Monte Carlo simulations with implicit ions using the Debye-Hückel potential. This mean field approximation is known to fail when increasing the ionic strength^{93,94}, so that moderate deviations from our computations are found at low pH-values for ionic strength larger than 0.1 M.

Clearly, the impact of the added salt on the titration curves is much more pronounced in the case of histatin-5, while it is very weak for β -amyloid 42. This point is better discussed by computing the increment in the charge due to the addition of salt at concentration c_s , *i.e.*, $\Delta Q(c_s) = |Q(c_s)| - |Q(0)|$, shown in Figs. 3c (histatin-5) and 3d (β -amyloid 42). It can be observed that in general $\Delta Q(c_s)$ increases with c_s . This is not surprising since the added salt screens the electrostatic repulsion between equally-charged aminoacids along the chain favoring ionization^{34,35}. Moreover, $\Delta Q(c_s)$ is larger for histatin-5 than for β -amyloid 42. This is possibly a consequence of the more balanced composition of β -amyloid 42, since the increase in c_s screens both the attractive and the repulsive interactions.

The binding capacitance as a function of the pH value is plotted in Figs. 3e and 3f for histatin-5 and β -amyloid 42, respectively. Note that the maxima of C and $\Delta Q(c_s)$ (panels c and d) take place at the same pH values, with $pH \simeq pK_{a,i}$ of the most

abundant aminoacids, that is, histidine for histatin-5 and aspartic and glutamic acid for β -amyloid 42. Therefore, the salt screening effect is maximum for pH-values in which the protein charge regulation and fluctuation are maxima (recall that capacitance and fluctuations are directly linked, see Eqn. 10)^{36,86,87}. It can be also observed that for histatin-5 the maxima of C are higher and shifted towards larger pH values as c_s increases. This behavior is also found in weak polyelectrolytes^{36,50}. The capacitance and charge increment $\Delta q_A(c_s) = |q_A(c_s)| - |q_A(0)|$ for each individual amino acid type as a function of the pH and c_s can be found in the Supporting Information[†] (Figs. S1 and S2).

The radii of gyration R_g as a function of pH and salt concentration are presented in Figs. 4a (histatin-5) and 4b (β -amyloid 42). It can be observed that for histatin-5, R_g decreases as the pH increases (panel a), while, for β -amyloid 42, R_g exhibits a minimum around $pH = 5$ (panel b). This fact can be easily explained: at the isoelectric point (pI) of the IDPs ($pI \sim 10.5$ for histatin-5 and $pI \sim 5$ for β -amyloid 42) the aminoacids of the proteins are either neutral or perfectly balanced, and, as a result, the repulsive electrostatic intramolecular interactions are minimized and the protein is, on average, more compact. The increase in c_s from zero (squares) to 0.5 M (diamonds) generally leads to a decrease in the R_g . Again, the presence of salt screens the intramolecular interactions, leading to less extended conformation. This effect is more pronounced at pH values far from the isoelectric point. Interestingly, for pH-values close to pI , the situation is reversed and R_g slightly increases with the ionic strength of the medium. In this case, the intramolecular attractive electrostatic forces seems to be more important than their repulsive counterpart.

3.2 Charge regulation of IDPs in crowded media with neutral obstacles

Let us now populate the system with neutral crowders with the same geometrical properties as BSA. For simplicity, all results shown in this section refer to systems with constant salt concentration $c_s = 0.1$ M, except otherwise stated. We have also performed computations at $c_s = 0.01$ M, with similar results, which are reported as Supplementary Information[†] (Fig. S3). The titration curves of histatin-5 and β -amyloid 42 were calculated at different excluded volume fractions ϕ , defined as

$$\phi = \frac{4\pi N_c R_C^3}{3L^3}, \quad (12)$$

where $R_C = 2.73$ nm is the hard-core radius of the crowders, $L = 22$ nm is the length of the simulation box and N_c is the number of crowder particles, which ranges from 2 to 63. Note that the outer soft shell of the crowders can be penetrated, specially at high crowder concentrations²¹, which is the reason why only the hard-core radius is considered in Eq. 12.

The effect of macromolecular crowding in the ionization properties of the IDPs is quantified by means of the difference between the absolute charges with and without crowding $\Delta Q'(\phi) = |Q(\phi)| - |Q(0)|$ where $Q(\phi)$ is the average charge of the protein at a given ϕ value. $\Delta Q'(\phi)$ is shown in Figs. 5a (histatin-5) and 5b (β -amyloid 42) as a function of the pH at ϕ -values ranging from:

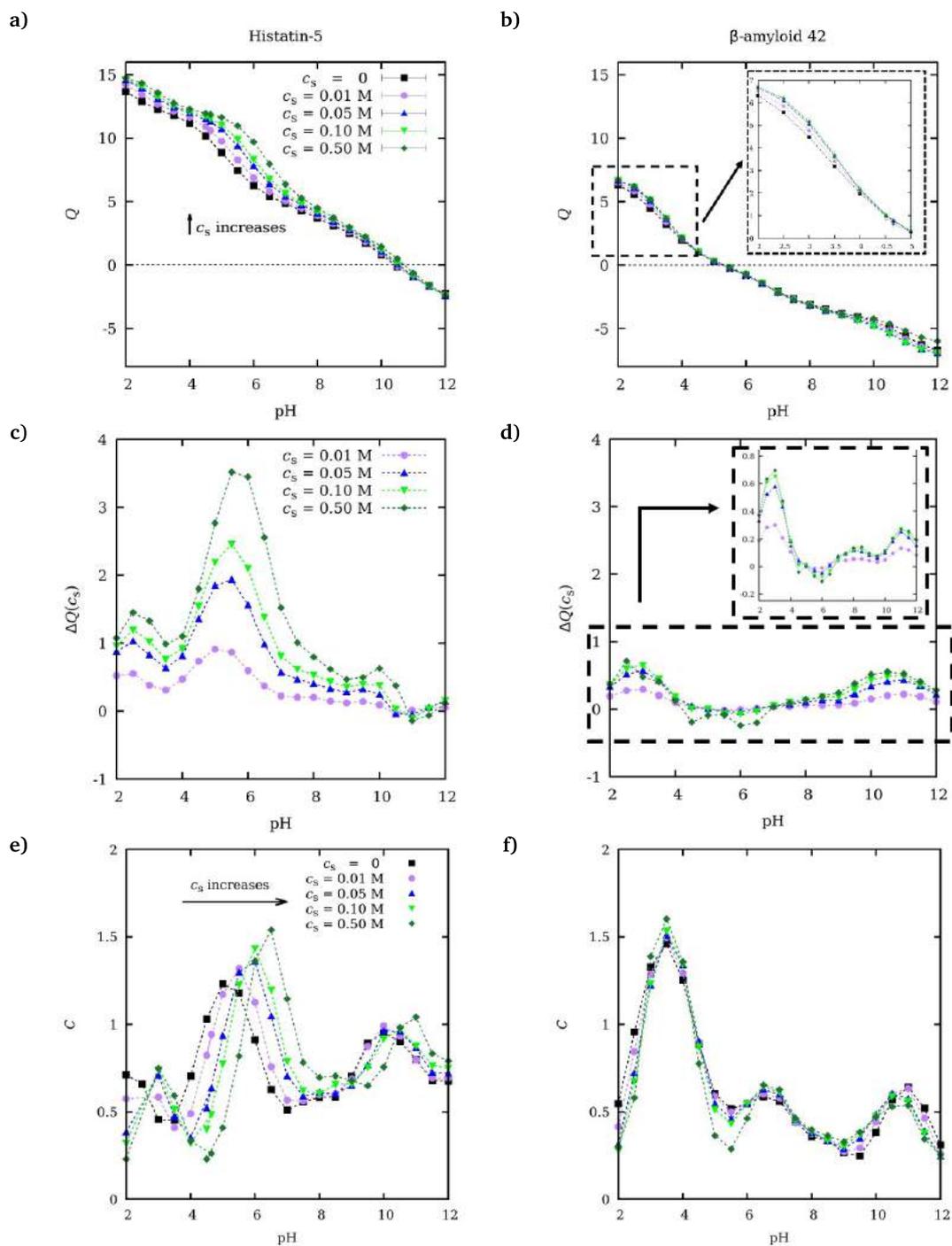


Fig. 3 Panels a) and b) show the average charge Q of the IDPs; panels c) and d) the charge increment due to the added salt $\Delta Q(c_s) = |Q(c_s)| - |Q(0)|$; and panels e) and f) the binding capacities C . The panels on the left-hand side refer to histatin-5 while the ones on the right to β -amyloid 42. The concentration of added salt are: no added salt (squares), $c_s = 0.01$ M (circles), 0.05 M (triangles), 0.1 M (inverted triangles) and 0.5 M (diamonds).

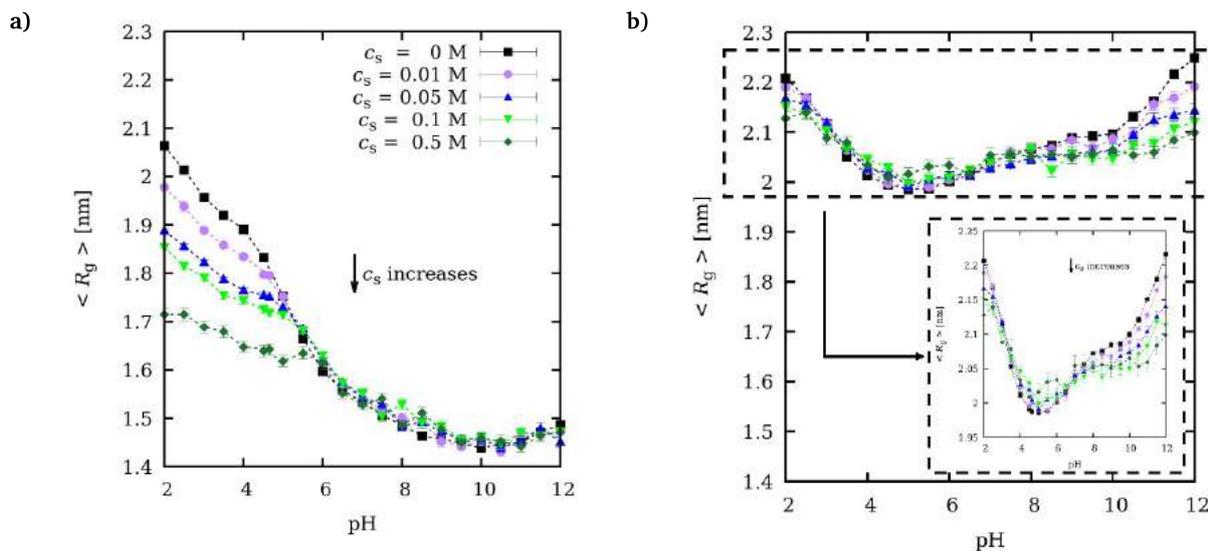


Fig. 4 Radius of gyration R_g of (a) histatin-5 and (b) β -amyloid 42 at salt concentrations ranging from $c_s = 0$ M (no added salt) to $c_s = 0.5$ M. The markers are the same as in Fig. 3.

$\phi = 0.1$ to $\phi = 0.5$. It can be observed that the impact of crowding is very different for the two studied IDPs. The average charge of histatin-5 increases with the crowder concentration at all the pH range, with a maximum increase of 0.6 units at pH = 6. Only for systems at pH = 11, a small decrease in the average charge observed. Conversely, β -amyloid 42 (Fig. 5b), presents a titration behavior that is nearly insensitive to the crowder concentration. It is only observed a small increase with ϕ at pH = 3 and pH = 10 and a small decrease at pH values from 4 to 7. The $\Delta Q'(\phi)$ vs. ϕ curves have been found to be very well fitted to an empiric power law in all the pH-range for both IDPs (See Section S4 in the Supporting Information[†]).

We also analyze the dependence of the average charge per aminoacid type q_A with ϕ by calculating the difference in charge with and without crowders $\Delta q'_A(\phi) = |q_A(\phi)| - |q_A(0)|$, where $q_A(\phi)$ is the average charge of an aminoacid type A at a given ϕ -value. In Figs. 5c (histatin-5) and 5d (β -amyloid 42), $\Delta q'_A(\phi)$ for aspartic acid (ASP), histidine (HIS) and lysine (LYS) are shown. These aminoacids are chosen as representatives of residues with acidic, neutral and basic isoelectric points, respectively. It can be observed that crowding affects the protonation/deprotonation of the aminoacids differently and $\Delta q'_A(\phi)$ of ASP, HIS, and LYS strongly depend on the identity of the IDP protein.

Interestingly, the variation of Q can be achieved by neutralizing aminoacids whose charges have opposite sign to that of the macromolecular charge. For instance, if an acidic residue, negatively charged, gets protonated, the global charge becomes more positive. An example can be observed at pH = 3, where both IDPs experience global charge increase (Figs. 5a and 5b), but the (absolute) charge of their acidic groups decrease (ASP in the Figs. 5c and 5d). Note that at this pH only the charge of the acidic residues is affected by crowding, because $\text{pH} \approx \text{p}K_{a,i} \approx 3$. This is, again, the pH region where their binding capacitance, and consequently the charge regulation ability of ASP, exhibits a maximum

(see Fig. S5 in the Supporting Information[†]). As discussed in the previous subsection, the maxima and minima in $\Delta Q'(\phi)$ coincide with the peaks in the charge capacitance C , as observed in Fig. 5e and Fig. 5f.

Hypothetically, the dependence of the macromolecular charge on ϕ could be regulated by two possible mechanisms. On the one side, larger excluded volumes lead to increase in the effective ionic strength I' of the solution (same number of salt ions in a reduced accessible volume). For a system with N charged chemical species, I' can be expressed in terms of the crowder excluded volume fraction ϕ as

$$I'(\phi) = \frac{1}{2} \sum_{i=1}^N c'_i(\phi) z_i^2 = \frac{c_s}{(1-\phi)}, \quad (13)$$

where $c'_i(\phi)$ is the “effective” concentration of the chemical specie i with charge z_i . For instance, for $\phi = 0.5$ and $c_s = 0.1$ M the effective ionic strength is $I' = 0.2$ M. As a result, the intra-molecular electrostatic interactions are weaker since they are more screened, and the average protein charge should be affected.^{34–36} On the other hand, in most of cases, excluded volume is found to induce more compact protein conformations.^{5,7,8} This compaction of the IDPs reduces the distance between charged groups and thus increases the intra-molecular electrostatic interactions.

Let us assess the relative importance of these two possible mechanisms. With this aim let us take as a reference situation $c_s = 0.1$ M. For $\phi = 0.5$, the effective ionic strength is $I' = 0.2$. If the variation in the IDPs charge is only due to the increase of I' induced by crowding, the corresponding charge variation in presence of crowding

$$\Delta Q'(\phi = 0.5) = |Q(\phi = 0.5, c_s = 0.1\text{M})| - |Q(\phi = 0, c_s = 0.1\text{M})| \quad (14)$$

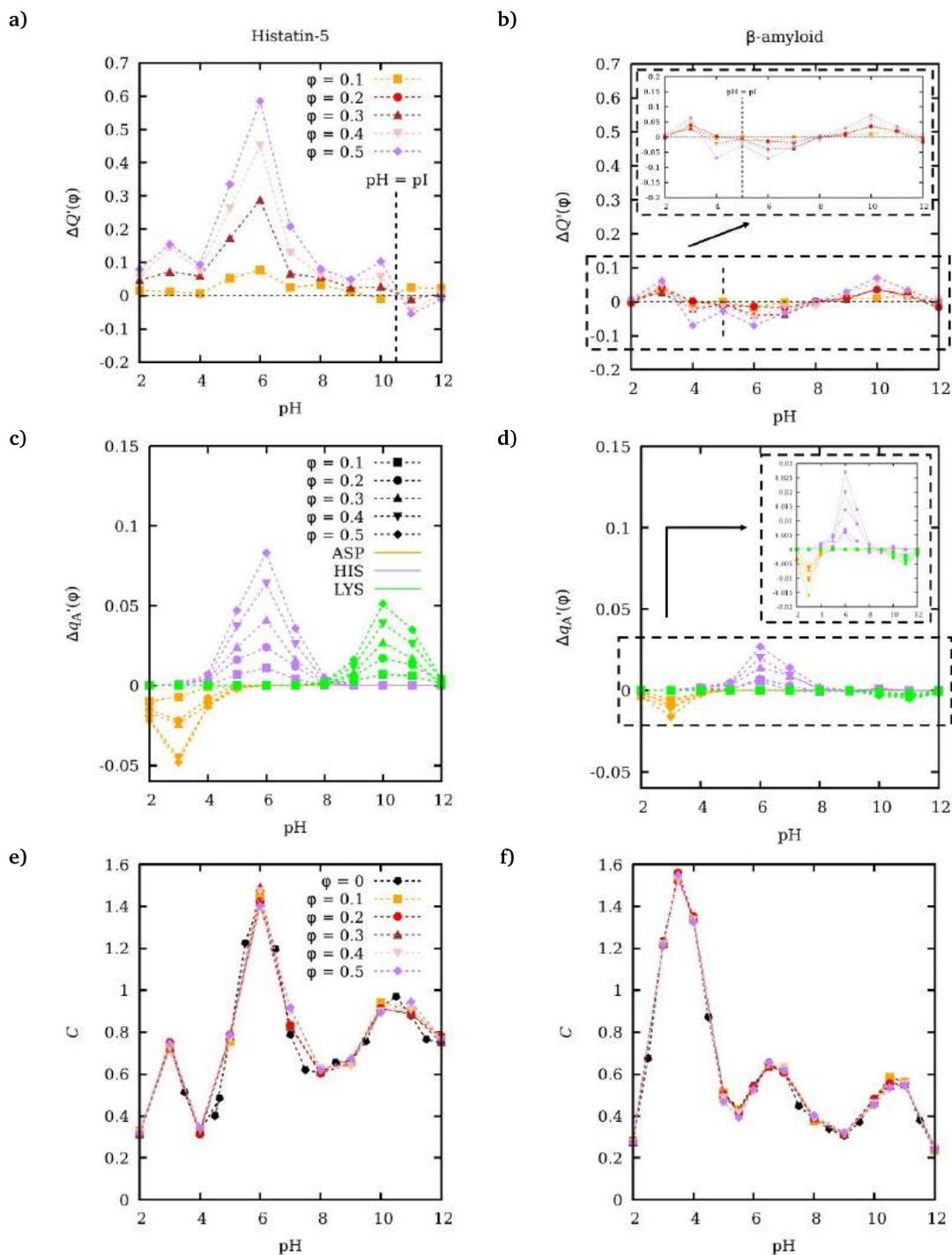


Fig. 5 Panels a and b: variation in the absolute macromolecular charge due to the presence of neutral crowders, $\Delta Q'(\phi) = |Q(\phi)| - |Q(0)|$. Panels c and d: variation in the average charge per aminoacid, $\Delta q_A'(\phi) = |q_A(\phi)| - |q_A(0)|$. Panels e and f: binding capacitance, C . Panels on the left-hand side refer to histatin-5 and the ones on the right-hand side to β -amyloid 42. The added salt concentration is $c_s = 0.1$ M of NaCl in all the simulations, while ϕ ranges from 0.1 to 0.5. The meaning of the markers and color of the lines used can be read inside the panels.

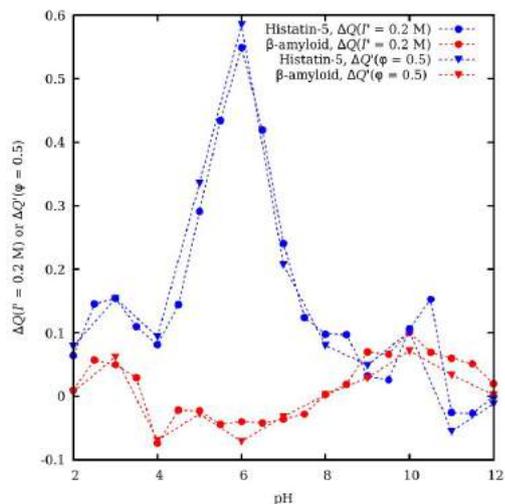


Fig. 6 Comparison between: i) charge variation produced when the added salt concentration is incremented from $c_s = 0.1$ M to 0.2 M without crowders $\Delta Q(I' = 0.2\text{M})$ (Eq. 15, circles); ii) charge variation produced when adding neutral crowders with resulting $\phi = 0.5$, $\Delta Q'(\phi = 0.5)$ (Eq. 14, triangles), to a protein solution with added salt concentration $c_s = 0.1$ M, i.e., with effective ionic strength $I' = 0.2$ M. Blue color refers to histatin-5 while red color refers to β -amyloid. As can be observed, $\Delta Q'(\phi = 0.5) \approx \Delta Q(I' = 0.2\text{M})$ evidencing that the main effect of the crowders is to increment the effective ionic strength.

should be the same as the charge variation without crowding but taking $c_s = I'$

$$\Delta Q(I' = 0.2\text{M}) = |Q(\phi = 0, c_s = 0.2\text{M})| - |Q(\phi = 0, c_s = 0.1\text{M})|, \quad (15)$$

that is

$$\Delta Q'(\phi = 0.5) \approx \Delta Q(I' = 0.2\text{M}) \quad (16)$$

This comparison is shown in Fig. 6. Clearly $\Delta Q(I' = 0.2\text{M})$ and $\Delta Q'(\phi = 0.5)$ are very similar for all the range of pH-values. This result strongly suggests that the main mechanism by which neutral crowders affect the charge regulation is increasing the effective ionic strength. Conversely, it seems that protein compaction induced by macromolecular crowding does not affect the IDPs charge.

Figs. 7a (histatin-5) and 7b (β -amyloid 42) show the normalized radii of gyration, $R_g(\phi)/R_g(0)$, where $R_g(0)$ is the radius of gyration in absence of crowders, at different pH and ϕ values. It can be seen that it slightly decreases with ϕ . This is in good agreement with the results obtained by Kang and co-authors⁹⁵, which found a reduction of the macromolecule dimension due to the presence of crowders of comparable size. They proposed to fit the normalized R_g to the scaling law^{95,96}

$$R_g(\phi)/R_g(0) = (1 - c\lambda\phi)^{1/5}, \quad (17)$$

where $\lambda = R_g(0)/R_c$ is the ratio between the IDP radii of gyration without crowders $R_g(0)$ and the crowder compact radius R_c , and c is an empiric parameter to be fitted. In this work, λ was found

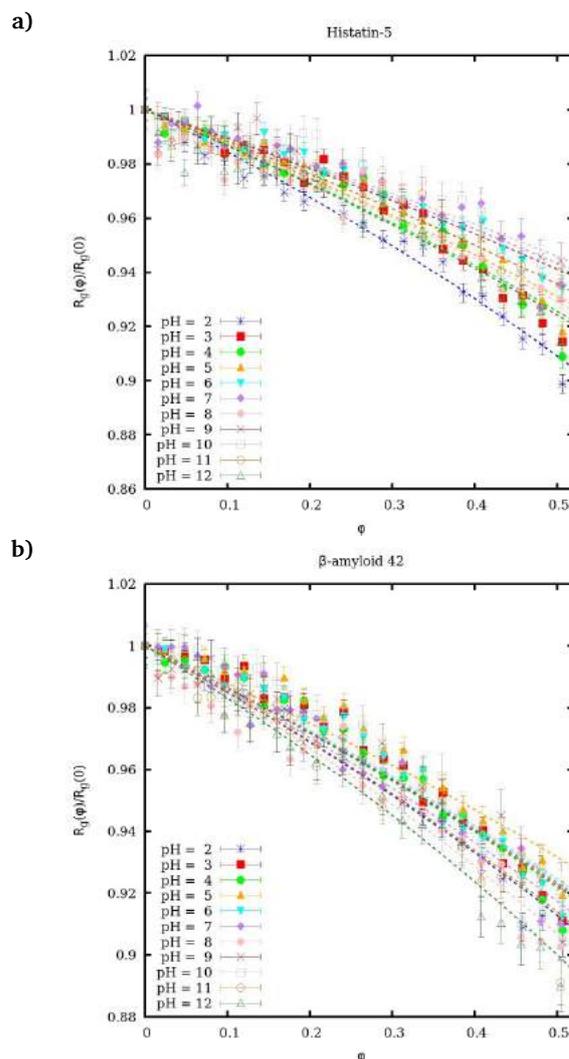


Fig. 7 Normalized radius of gyration, R_g , of (a) histatin 5 and (b) β -amyloid 42 vs. ϕ at pH-values ranging from 2 to 12 in the presence of neutral crowders. Dashed lines represent the best-fit to the scaling law (Eq. 17).

to range from 5.4 to 8.0, depending on the IDP and the pH conditions. The best fit curves to Eq. (17), are depicted as dashed lines in the Figs. 7a and 7b. The corresponding c -values (listed in Section S6 as Supporting Information[†]), are found to be nearly independent of the IDP and the pH value, with $c \approx 0.09$ in all the cases. It is also worth mentioning that if λ is set to 1, following the approach used in Ref.⁵³, where the crowders were IDPs, we obtain c -values for histatin-5 around ~ 0.6 , close to the reported value $c = 0.71$ for similar salt conditions.

3.3 IDPs titration in presence of charged crowders

The charged crowders are here modeled as a coarse grained version of Bovine Serum Albumin (BSA) proteins, which we assume to have the same characteristic radii (R_C and R_H) than the neutral crowders of the previous sub-section. The only difference is that we incorporate ionized groups at the surface, modeled as charged beads which mimic the ionized groups of the protein. The total charge of the surface corresponds to the average charge of BSA at the pH-value and salt concentration of the simulation. The latter information, shown in Tab. 2, is taken from Ref.⁷⁰, where the average charge of BSA for $c_s = 0.01$ M in a wide range of pH-values was experimentally determined. The charge of the crowders is compensated by the corresponding counter-ions in order to maintain the electroneutrality of the solution. We restrict ourselves to ϕ values up to 0.2 due to two reasons: (i) for very acidic/basic pH values the number of counter-ions increases dramatically with ϕ making the computations prohibitively expensive and (ii) the presence of charged hard beads in the crowder surface increases the excluded volume, which complicates the equilibration of systems at large ϕ -values.

Following the same order as in the previous sub-section, let us firstly analyze the variation in the IDP charge induced by the presence of charged crowders, $\Delta Q'(\phi) = |Q(\phi)| - |Q(0)|$, depicted in Figs. 8a (histatin 5) and 8b (β -amyloid 42). The average charges of the specific residues $\Delta q'_A(\phi) = |q_A(\phi)| - |q_A(0)|$ are reported in Figs. 8c and 8d. The simulations are performed at four ϕ -values: 0.05 (triangles), 0.1 (squares), 0.15 (inverted triangles) and 0.2 (circles). It can be observed that, for the same pH and ϕ values, charged crowders induce significantly larger variation in the charge of the individual aminoacids and in the IDP global charge. For example, for $\phi = 0.2$, histatin-5 exhibits a maximum $\Delta Q'(\phi)$ value of 0.2 units (pH = 6) with neutral crowders while in the presence of charged crowders the maximum $\Delta Q'(\phi)$ is 0.75. Likewise, for β -amyloid 42, $\Delta Q'(\phi)$ is very small for neutral crowders but increases significantly for almost all pH range with charged crowders. Again, the $\Delta Q'(\phi)$ vs. ϕ curves have been found to be very well fitted to an empiric power law in all the pH-range for both IDPs (See Section S4 in the Supporting Information[†]). The binding capacitance C in the presence of charged crowders is presented in Figs. 8e and 8f. It can be seen that crowding produces the shift of the first maximum towards larger pH values for histatin-5. A small increase in the height of the first maximum (at pH = 4) for β -amyloid 42 is also found. The same trend can be observed in the binding capacitance per aminoacid (Supporting Information[†], Fig. S7).

Note that, as happened with neutral crowders, in some cases (mostly at acidic pH-values), the protein charge increases by reducing the charge of the aminoacids with charge sign being opposite to the global charge. This fact can be observed, for example, at pH = 3 in Figs. 8a and 8c. Curiously, this phenomenon takes place for almost all the pH range in the case of β -amyloid 42 (See Figs. 8b and 8d). Note also that, for β -amyloid 42, the dependence of $\Delta q'_A$ of HIS with pH depends of the nature of the crowder. In the presence of neutral crowders, its charge slightly increases for larger ϕ values (See Fig. 5d), while the opposite situation is

observed for charged crowders.

In the previous sub-section, we highlighted that a fundamental mechanism of charge regulation upon the addition of neutral crowders was the increment of the effective ionic strength I' due to the reduction of the available volume. This effect should be even larger if the crowders are charged due to the presence of the counter-ions associated to them. I' can now be calculated by modifying Eq. 13 in order to include the counter-ions coming from the crowders. The resulting effective ionic strength I' reads

$$I'(\phi, \text{pH}) = \frac{c_s}{(1-\phi)} + \frac{3\phi|z(\text{pH})|}{4\pi R_c^3 N_A (1-\phi)}, \quad (18)$$

where $z(\text{pH})$ is the crowder charge at a given pH value (see Tab. 2), R_c is the crowder hard-core radius (we take $R_c = 2.73$ nm as in the previous sub-section) and N_A is the Avogadro number. The second term on the right-hand side represents the contribution of the counter-ions coming from the charged crowders. I' vs. ϕ curves (Eq. 18) for $c_s = 0.01$ M at different pH-values can be found as Supporting Information[†] (see Section S8). Depending on ϕ and pH, I' may amount to as much as 0.18 M, almost twenty times the c_s -value.

As in the previous section, let us investigate if the main mechanism for the IDP's charge regulation is the increment in I' . We want to compare the resulting charge variation $\Delta Q(I' = 0.02\text{M})$, calculated in the same way as in Eq. 15, with that of a system with charged crowders whose effective ionic strength is always $I' = 0.02\text{M}$, independently of the pH-value. That means that, for each pH-value, we have to calculate the ϕ -value for which Eq. 18 predicts $I' = 0.02$ M. The comparison between the charge variations in both situations, $\Delta Q(I' = 0.02\text{M})$ and $\Delta Q'(\phi)$, is shown in Fig. 9. As in the case of neutral crowders, both charge variations are very similar, evidencing that the main effect of the charged crowders is to increment the ionic strength. However, in this case it can be observed a shift between $\Delta Q(c_s)$ and $\Delta Q'(\phi)$ values which was not detected for neutral crowders (Fig. 6). This fact suggests that there is a significant contribution to charge regulation from the electrostatic interaction between the IDP and the charged crowders. This observation is in good agreement with previous research on the interactions between weak polyelectrolytes and charged nanoparticles, where the titration curves of weak polyelectrolytes was observed to be shifted as the surface charge of the nanoparticle increased⁹⁷. This effect seems to be, however, more modest for IDP's than for polyelectrolytes, probably because the charge density is smaller in the case of IDP's. The same analysis has been performed for $I' = 0.05$ M, *i.e.* larger excluded volumes, which is reported as Supporting Information[†] (see Section S9). The resulting shift is less pronounced in this case, suggesting that effect of increasing I' is now more pronounced and the IDP-charged crowder interaction plays a minor role. However, further research would probably be necessary to fully clarify this point.

The normalized radius of gyration R_g , as a function of ϕ for pH ranging from 3 to 11 is depicted in Figs. 10a and 10b. It is found that increasing ϕ leads to a slight decrease in R_g of both IDPs. This fact can be in part due to the steric hinderance caused by the presence charged crowders, which also took place in the case

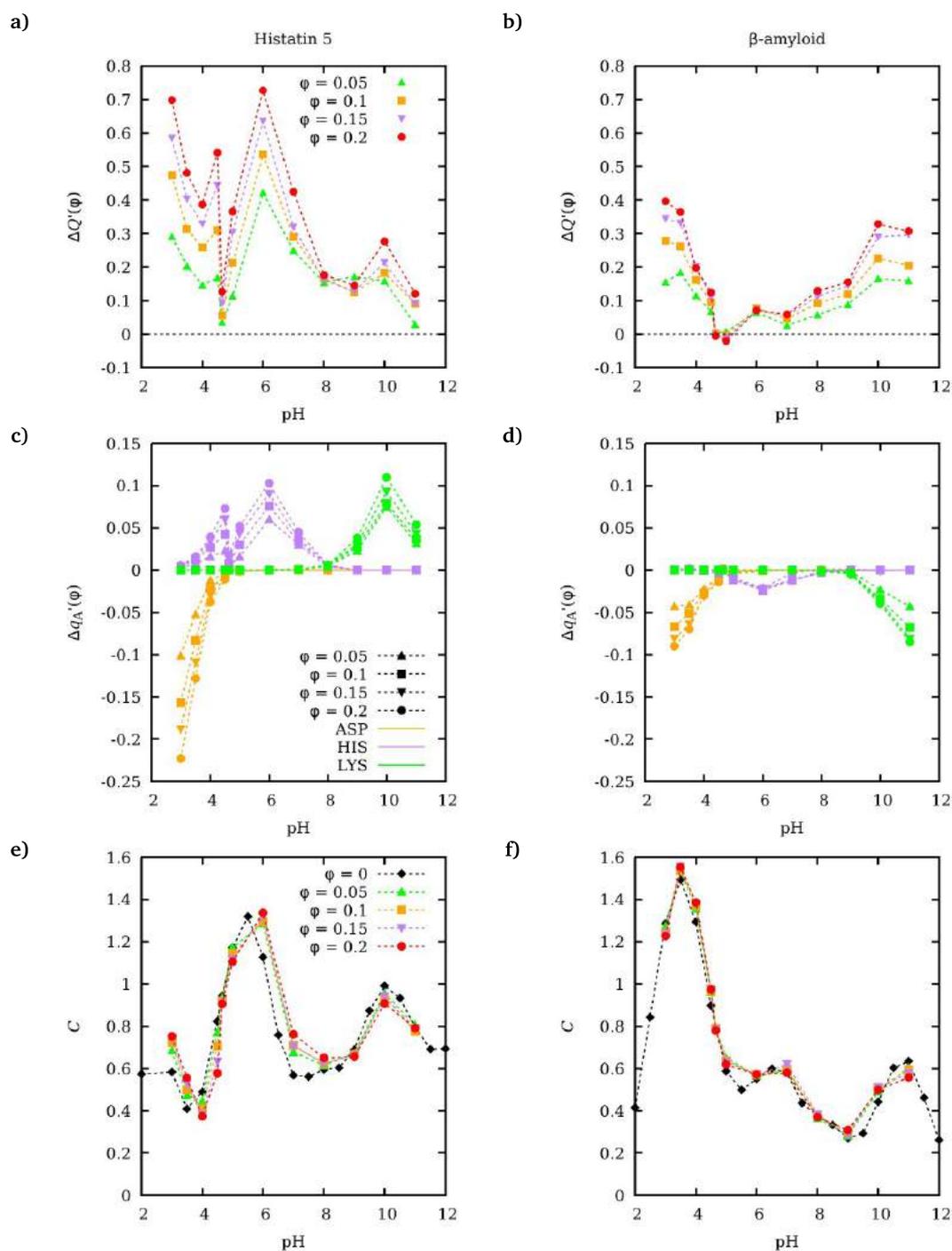


Fig. 8 Panels a and b: variation in the absolute macromolecular charge due to the presence of charged crowders, $\Delta Q'(\phi) = |Q(\phi)| - |Q(0)|$; panels c and d: variation in the average charge per aminoacid, $\Delta q_A'(\phi) = |q_A(\phi)| - |q_A(0)|$; and panels e and f: binding capacitance C . Panels on the left-hand side refer to histatin-5 and the ones on the right-hand side to β -amyloid 42. The added salt concentration is $c_s = 0.01$ M of NaCl in all the simulations, while ϕ ranges from zero to 0.2. The meaning of the markers and color of the curves can be read inside the panels.

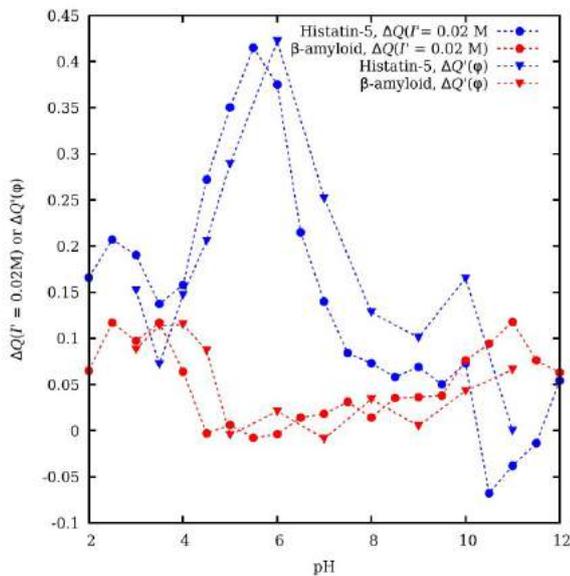


Fig. 9 Comparison between: i) charge variation produced when the added salt concentration is incremented from $c_s = 0.01$ M to 0.02 M without crowders ($\Delta Q(I' = 0.02\text{M}) = |Q(\phi = 0, c_s = 0.02\text{M})| - |Q(\phi = 0, c_s = 0.01\text{M})|$, circles); and ii) charge variation produced when adding charged crowders $\Delta Q'(\phi)$ (triangles) to a protein solution with added salt concentration $c_s = 0.01$ M. The excluded volume $\phi(\text{pH}, 0.01\text{M})$ is chosen such as the effective ionic strength I' (calculated using Eq.18) is always $I' = 0.02$ M. Blue color refers to Histatin-5 while red color refers to β -amyloid.

of neutral crowders. However, for the same ϕ , this effect is much stronger in the presence of charged crowders, as one concludes by comparing the normalized R_g vs. ϕ curves obtained with neutral (Figs. 7a and 7b) and with charged crowders. The difference is specially remarkable for histatin-5 at pH-values between pH = 3 and pH = 4, in which histatin-5 is highly charged. This difference can be explained again by the fact that increasing charged crowder concentration produces an extra increase in the effective ionic strength of the media, which weakens electrostatic repulsion and facilitates protein compaction. The normalized R_g obtained by the simulations (markers) have been fitted to the scaling law in Eq. 17 (dashed lines) and the resulting c -values are listed in Section S6 in the Supporting Information[†]). In general, very good fitting is obtained, with c -values ranging from 0.08 to 0.33. For pH values close to the IDPs isoelectric points (5 and 11 for histatin-5 and β -amyloid 42, respectively), the obtained c -values are similar to those corresponding to neutral crowders. Away from the isoelectric point, however, the resulting c -values for charged crowders are larger than those calculated in the presence of neutral crowders, specially for histatin-5 in the range pH = 3 to pH = 4.

4 Conclusions

The effect of macromolecular crowding in the charge regulation of two Intrinsically Disordered Proteins (IDPs), histatin-5 and β -amyloid 42, is studied by means of Semi-Grand Canonical Monte Carlo simulations. It is found that the titration curve of histatin-5 depends on the salt concentration. This effect is specially remarkable at pH-values for which the binding capacitance presents a maximum, *i.e.* the protein charge is more susceptible

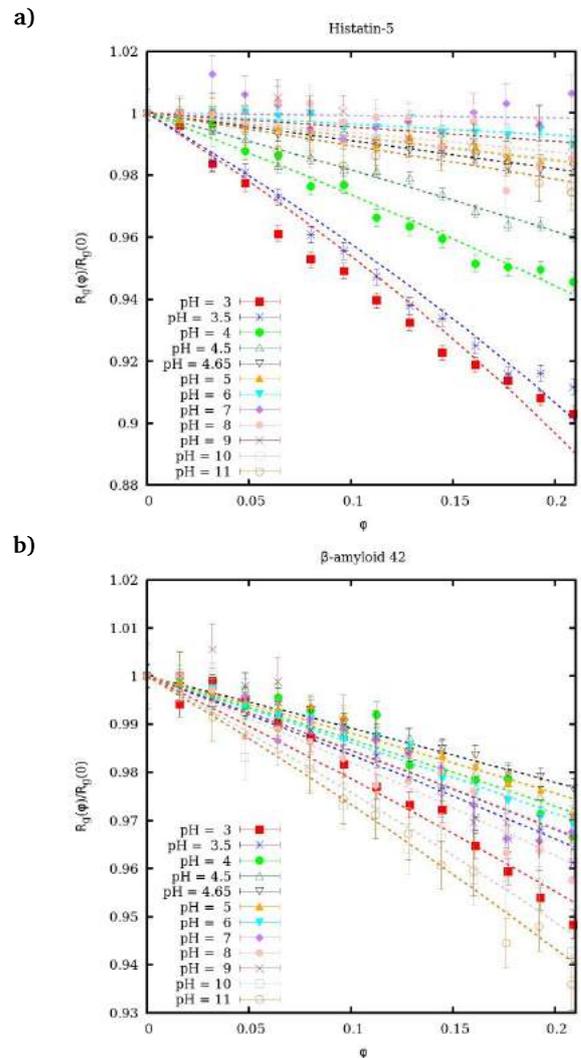


Fig. 10 Normalized radius of gyration, R_g , of (a) histatin 5 and (b) β -amyloid 42 vs. ϕ at pH-values ranging from 2 to 12 in the presence of charged crowders. Dashed lines represent the best fit to the scaling law (Eq. 17).

to be regulated. Conversely, the titration curve of β -amyloid 42 was nearly insensitive to salt concentration, probably due to its polyampholyte nature: repulsive and attractive intra-molecular forces are affected in opposite way. The radius of gyration R_g of the IDPs is significantly salt dependent for pH values far from their isoelectric point pI. In this case, R_g is reduced as the salt concentration increases, due to screening of the intra-molecular electrostatic repulsion.

Our results suggest that the main mechanism by which neutral crowders change the IDP charge is the effective increase in salt concentration due to excluded volume (which screens the IDP charge). The presence of neutral crowders induces charge regulation on histatin-5 whereas the charge of β -amyloid 42 is nearly unaffected. This can be again probably explained by the fact that in histatin-5 the positively charged basic residues are predominant. Conversely, β -amyloid 42 is a polyampholyte, and repulsive and attractive forces are affected in opposite manner.

The situation changes when charged crowders are introduced. For both IDPs, the global charge increases with ϕ for almost all the pH range. In this case, the effect of crowding in the IDPs charge is found to be not only dependent on the increase in the effective ionic strength, resulting from the volume reduction and the addition of counter-ions associated to the charged crowders, but also due to their direct interaction with the charged crowders, causing a shift when compared to the systems in absence of charged crowding agents at the same ionic strength.

For both neutral and charged crowders, the analysis of the average charge per aminoacid type reveals that, in some cases (specially at acidic pH-values), the mechanism through which the IDP regulates its global charge consists of neutralizing the charge of the aminoacids with charge sign opposite to that of the global charge. Moreover, the pH values for which the charge is more affected by crowding coincide with those of the maximum binding capacitance. Finally, the radius of gyration R_g decreases with ϕ , and its dependence on ϕ fits very well the power law proposed in the Ref.⁹⁵, containing only one fitting parameter.

We highlight that, even though the charge regulation of the IDPs induced by macromolecular crowding can be in principle considered moderate, it could have relevant biological implications. For example, charge regulation due to macromolecular crowding could be relevant in membrane-less organelles, which are formed by charged macromolecules.⁹⁸ Furthermore, the physicochemical effects here discussed could be probably more intense in systems with larger charge densities (polycarboxylates, polyamines, polysaccharides) or in the presence of multi-valent ions. It is worth to mention that in our simulations the salt concentration is kept constant. This situation resembles an *in vitro* experiment, where the titration of the IDP is performed in presence of a buffer at constant salt concentration. Alternatively, one could perform simulations at constant salt chemical potential, which could be more convenient, for instance, in describing dialysis experiments. Further work is necessary to clarify this point. To our knowledge, the present work represents the first attempt to study, at least by computational simulation, the interplay between charge regulation and macromolecular crowding.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

P. M. Blanco, S. Madurga, J. L. Garcés and F. Mas acknowledge the financial support from Generalitat de Catalunya (Grants 2017SGR1033 and 2017SGR1329). P. M. Blanco, S. Madurga and F. Mas acknowledge Spanish Structures of Excellence María de Maeztu program through (Grant MDM-2017-0767). J. L. Garcés also acknowledges the Spanish Ministry of Science and Innovation (project CTM2016-78798-C2-1-P). P. M. Blanco also acknowledges the financial support from grant (FI-2017) of Generalitat de Catalunya. The computations were performed on resources provided by the NTNU IDUN/EPIC computing cluster.

Notes and references

- 1 J.-P. Behr, *The lock-and-key principle: the state of the art-100 years on*, J. Wiley, 1994.
- 2 J. J. Ward, J. S. Sodhi, L. J. McGuffin, B. F. Buxton and D. T. Jones, *Journal of Molecular Biology*, 2004, **337**, 635–645.
- 3 V. N. Uversky, *Journal of Biomedicine and Biotechnology*, 2010, **2010**, 1–14.
- 4 V. N. Uversky, *Biochimica et Biophysica Acta - Proteins and Proteomics*, 2011, **1814**, 693–712.
- 5 A. V. Fonin, A. L. Darling, I. M. Kuznetsova, K. K. Turoverov and V. N. Uversky, *Cellular and Molecular Life Sciences*, 2018, **75**, 3907–3929.
- 6 V. N. Uversky and A. K. Dunker, *Biochimica et Biophysica Acta - Proteins and Proteomics*, 2010, **1804**, 1231–1264.
- 7 A. P. Minton, *Biopolymers*, 1981, **20**, 2093–2120.
- 8 A. P. Minton, *Biophysical Journal*, 1992, **63**, 1090–1100.
- 9 Y. Shen, J. M. Jacobs, D. G. Camp, R. Fang, R. J. Moore, R. D. Smith, W. Xiao, R. W. Davis and R. G. Tompkins, *Analytical Chemistry*, 2004, **76**, 1134–1144.
- 10 T. C. Laurent, *European Journal of Biochemistry*, 1971, **21**, 498–506.
- 11 R. J. Ellis, *Trends in Biochemical Sciences*, 2001, **26**, 597–604.
- 12 D. S. Banks and C. Fradin, *Biophysical Journal*, 2005, **89**, 2960–2971.
- 13 P. Mereghetti and R. C. Wade, *Journal of Physical Chemistry B*, 2012, **116**, 8523–8533.
- 14 J. Balbo, P. Mereghetti, D. P. Herten and R. C. Wade, *Biophysical Journal*, 2013, **104**, 1576–1584.
- 15 C. Balcells, I. Pastor, E. Vilaseca, S. Madurga, M. Cascante and F. Mas, *Journal of Physical Chemistry B*, 2014, **118**, 4062–4068.
- 16 I. Pastor, L. Pitulice, C. Balcells, E. Vilaseca, S. Madurga, A. Isvoran, M. Cascante and F. Mas, *Biophysical Chemistry*, 2014, **185**, 8–13.
- 17 C. Balcells, I. Pastor, L. Pitulice, M. Via, S. Madurga, E. Vilaseca, A. Isvoran, M. Cascante and F. Mas, *New Frontiers in Chemistry*, 2015, **24**, 3–16.
- 18 S. Kondrat, O. Zimmermann, W. Wiechert and E. V. Lieres, *Physical Biology*, 2015, **12**, 46003.
- 19 P. M. Blanco, M. Via, J. L. Garcés, S. Madurga and F. Mas, *Entropy*, 2017, **19**, 105.
- 20 M. Feig, I. Yu, P. H. Wang, G. Nawrocki and Y. Sugita, *Journal of Physical Chemistry B*, 2017, **121**, 8009–8025.
- 21 P. M. Blanco, J. L. Garcés, S. Madurga and F. Mas, *Soft Matter*, 2018, **14**, 3105–3114.
- 22 A. Kuzmak, S. Carmali, E. von Lieres, A. J. Russell and S. Kondrat, *Scientific Reports*, 2019, **9**, 1–7.
- 23 E. A. Cino, M. Karttunen and W. Y. Choy, *PLoS ONE*, 2012, **7**, e49876.
- 24 C. P. Brangwynne, P. Tompa and R. V. Pappu, *Nature Physics*, 2015, **11**, 899–904.
- 25 A. Banks, S. Qin, K. L. Weiss, C. B. Stanley and H. X. Zhou, *Biophysical Journal*, 2018, **114**, 1067–1079.
- 26 G. L. Dignon, W. Zheng, Y. C. Kim, R. B. Best and J. Mittal, *PLoS Computational Biology*, 2018, **14**, 1–23.

- 27 F. C. Zegarra, D. Homouz, A. G. Gasic, L. Babel, M. Kovermann, P. Wittung-Stafshede and M. S. Cheung, *Journal of Physical Chemistry B*, 2019, **123**, 3607–3617.
- 28 R. K. Das and R. V. Pappu, *Pnas*, 2013, **110**, 13392–13397.
- 29 R. K. Das, K. M. Ruff and R. V. Pappu, *Current opinion in structural biology*, 2015, **32**, 102–112.
- 30 M. Ullner and B. Jonsson, *Macromolecules*, 1996, **29**, 6645–6655.
- 31 S. Ulrich, M. Seijo and S. Stoll, *The Journal of Physical Chemistry B*, 2007, **111**, 8459–8467.
- 32 J. L. Garcés, S. Madurga and M. Borkovec, *Physical Chemistry Chemical Physics*, 2014, **16**, 4626–4638.
- 33 J. L. Garcés, S. Madurga, C. Rey-Castro and F. Mas, *Journal of Polymer Science, Part B: Polymer Physics*, 2017, **55**, 275–284.
- 34 P. M. Blanco, S. Madurga, F. Mas and J. L. Garcés, *Polymers*, 2018, **10**, 811.
- 35 P. M. Blanco, S. Madurga, F. Mas and J. L. Garcés, *Macromolecules*, 2019, **52**, 8017–8031.
- 36 P. M. Blanco, S. Madurga, C. F. Narambuena, F. Mas and J. L. Garcés, *Polymers*, 2019, **11**, 1962.
- 37 S. Uyaver and C. Seidel, *Europhysics Letters*, 2003, **64**, 536–542.
- 38 S. T. Whitten, B. Garcia-Moreno E. and V. J. Hilser, *Proceedings of the National Academy of Sciences*, 2005, **102**, 4282–4287.
- 39 D. S. Olander and A. Holtzer, *Journal of the American Chemical Society*, 1968, **90**, 4549–4560.
- 40 H. Boroudjerdi, Y. Kim, A. Naji, R. R. Netz, X. Schlagberger and A. Serr, *Physics Reports*, 2005, **416**, 129–199.
- 41 M. Muthukumar, *Macromolecules*, 2017, **50**, 9528–9560.
- 42 J. Landsgesell, L. Nová, O. Rud, F. Uhlík, D. Sean, P. Hebbeker, C. Holm and P. Košovan, *Soft Matter*, 2019, **15**, 1155–1185.
- 43 A. F. Jorge, R. S. Dias and A. A. Pais, *Biomacromolecules*, 2012, **13**, 3151–3161.
- 44 C. F. Narambuena, E. P. Leiva and E. Pérez, *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, 2015, **487**, 49–57.
- 45 S. K. Ramisetty and R. S. Dias, *Journal of Molecular Liquids*, 2015, **210**, 64–73.
- 46 M. Stornes, P. Linse and R. S. Dias, *Macromolecules*, 2017, **50**, 5978–5988.
- 47 P. Torres, L. Bojanich, F. Sanchez-Varretti, A. J. Ramirez-Pastor, E. Quiroga, V. Boeris and C. F. Narambuena, *Colloids and Surfaces B: Biointerfaces*, 2017, **160**, 161–168.
- 48 M. Stornes, B. Shrestha and R. S. Dias, *Journal of Physical Chemistry B*, 2018, **122**, 10237–10246.
- 49 P. B. Torres, E. Quiroga, A. J. Ramirez-pastor, V. Boeris and C. F. Narambuena, *The Journal of Physical Chemistry B*, 2019, **123**, 8617–8627.
- 50 K. Hyltegren, T. Nylander, M. Lund and M. Skepö, *Journal of Colloid and Interface Science*, 2016, **467**, 280–290.
- 51 K. Hyltegren and M. Skepö, *Journal of Colloid and Interface Science*, 2017, **494**, 266–273.
- 52 C. Cragnell, E. Rieloff and M. Skepö, *Journal of Molecular Biology*, 2018, **430**, 2478–2492.
- 53 E. Fagerberg, S. Lenton and M. Skepö, *Journal of Chemical Theory and Computation*, 2019, **15**, 6968–6983.
- 54 A. M. Baptista, V. H. Teixeira and C. M. Soares, *The Journal of Chemical Physics*, 2002, **117**, 4184–4200.
- 55 J. Mongan, D. A. Case and J. A. McCammon, *Journal of Computational Chemistry*, 2004, **25**, 2038–2048.
- 56 B. J. MacKay, L. Denepitiya, V. J. Iacono, S. B. Krost and J. J. Pollock, *Infection and Immunity*, 1984, **44**, 695–701.
- 57 J. J. Pollock, L. Denepitiya, B. J. MacKay and V. J. Iacono, *Infection and Immunity*, 1984, **44**, 702–707.
- 58 S. Puri and M. Edgerton, *Eukaryotic Cell*, 2014, **13**, 958–964.
- 59 K. Wróblewski, R. Muhandiram, A. Chakrabarty and A. Ben-nick, *European Journal of Biochemistry*, 2001, **268**, 4384–4397.
- 60 A. Bennick, *Critical Reviews in Oral Biology and Medicine*, 2002, **13**, 184–196.
- 61 J. Kang, H. G. Lemaire, A. Unterbeck, J. M. Salbaum, C. L. Masters, K. H. Grzeschik, G. Multhaupt, K. Beyreuther and B. Müller-Hill, *Nature*, 1987, **325**, 733–736.
- 62 P. M. Gorman, C. M. Yip, P. E. Fraser and A. Chakrabarty, *Journal of Molecular Biology*, 2003, **325**, 743–757.
- 63 G. F. Chen, T. H. Xu, Y. Yan, Y. R. Zhou, Y. Jiang, K. Melcher and H. E. Xu, *Acta Pharmacologica Sinica*, 2017, **38**, 1205–1235.
- 64 R. Jurij and L. Per, *Journal of Computational Chemistry*, 2015, **36**, 1259–1274.
- 65 F. G. Oppenheim, T. Xu, F. M. McMillian, S. M. Levitz, R. D. Diamond, G. D. Offner and R. F. Troxler, *Journal of Biological Chemistry*, 1988, **263**, 7472–7477.
- 66 Y. Nozaki and C. Tanford, *Methods in Enzymology*, 1967, **11**, 715–734.
- 67 J. L. Garcés, G. J. M. Koper and M. Borkovec, *J. Phys. Chem. B*, 2006, **110**, 10937–10950.
- 68 P. Vilaseca and G. Franzese, *Journal of Chemical Physics*, 2010, **133**, 1–11.
- 69 F. Leoni and G. Franzese, *Journal of Chemical Physics*, 2014, **141**, 174501.
- 70 M. K. Menon and A. L. Zydney, *Analytical Chemistry*, 1998, **70**, 1581–1584.
- 71 J. Bernhardt and H. Pauly, *Journal of Physical Chemistry*, 1975, **79**, 584–590.
- 72 J. R. Whitaker, *Analytical Chemistry*, 1963, **35**, 1950–1953.
- 73 M. L. Huber, R. A. Perkins, A. Laesecke, D. G. Friend, J. V. Sengers, M. J. Assael, I. N. Metaxa, E. Vogel, R. Mareš and K. Miyagawa, *Journal of Physical and Chemical Reference Data*, 2009, **38**, 101–125.
- 74 A. K. Gaigalas, J. B. Hubbard, M. McCurley and S. Woo, *Journal of Physical Chemistry*, 1992, **96**, 2355–2359.
- 75 T. Raj and W. H. Flygare, *Biochemistry*, 1974, **13**, 3336–3340.
- 76 F. Carlsson, M. Malmsten and P. Linse, *Journal of the American Chemical Society*, 2003, **125**, 3140–3149.
- 77 Y. Xu, M. Mazzawi, K. Chen, L. Sun and P. L. Dubin, *Biomacromolecules*, 2011, **12**, 1512–1522.

- 78 R. De Vries, *J. Chem. Phys.*, 2014, **120**, 3475–3481.
- 79 Y. Cemil, J. Heyda, M. Ballauff and J. Dzubiella, *J. Chem. Phys.*, 2015, **143**, 064905.
- 80 C. Yigit, M. Kanduc, M. Ballau and J. Dzubiella, *Langmuir*, 2017, **33**, 417–427.
- 81 M. J. Blandamer and M. C. Symons, *Journal of Physical Chemistry*, 1963, **67**, 1304–1306.
- 82 S. Madurga, A. Martín-Molina, E. Vilaseca, F. Mas and M. Quesada-Pérez, *Journal of Chemical Physics*, 2007, **126**, 234703–1–234703–11.
- 83 S. Madurga, C. Rey-Castro, I. Pastor, E. Vilaseca, C. David, J. L. Garcés, J. Puy and F. Mas, *Journal of Chemical Physics*, 2011, **135**, 184103.
- 84 B. Linse and P. Linse, *Journal of Chemical Physics*, 2014, **141**, 184114.
- 85 M. Sjalander, M. Jahre, G. Tufte and N. Reissmann, *EPIC: An Energy-Efficient, High-Performance GPGPU Computing Research Infrastructure*, 2019.
- 86 E. D. Cera, *J. Phys. Chem.*, 1991, **95**, 5082–5086.
- 87 M. Lund and B. Jo, *Quarterly Reviews of Biophysics*, 2013, **46**, 265–281.
- 88 E. J. Helmerhorshmt, W. V. Hof, P. Breeuwer, E. C. I. Veerman, T. Abee, R. F. Troxler, A. V. N. Amerongen and F. G. Oppenheim, *The Journal of Biological Chemistry*, 2001, **276**, 5643–5649.
- 89 H. Nikawa, H. Fukushima, S. Makihira, T. Hamada and L. P. Samaranayake, *Oral diseases*, 2004, **10**, 221–228.
- 90 A. Kurut, J. Henriques, J. Forsman, M. Skepo and M. Lund, *Proteins*, 2014, **82**, 657–667.
- 91 C. Cragnell, D. Durand, B. Cabane and M. Skepo, *Proteins*, 2016, 1–15.
- 92 G. M. Klug, D. Losic, S. S. Subasinghe, M. I. Aguilar, L. L. Martin and D. H. Small, *European Journal of Biochemistry*, 2003, **270**, 4282–4293.
- 93 L. Guldbrand, B. Jönsson, H. Wennerström and P. Linse, *The Journal of Chemical Physics*, 1984, **80**, 2221–2228.
- 94 M. Borkovec, B. Jönsson and G. J. M. Koper, *Surface and Colloid Science*, 2001, 99–339.
- 95 H. Kang, P. A. Pincus, C. Hyeon and D. Thirumalai, *Physical Review Letters*, 2015, **114**, 068303:1–5.
- 96 D. Thirumalai, *Physical Review A*, 1988, **37**, 269–276.
- 97 A. Laguerre and S. Stoll, *Polymer*, 2005, **46**, 1359 – 1372.
- 98 W. Stroberg and S. Schnell, *Journal of Theoretical Biology*, 2017, **434**, 42–49.

Conclusions

The present thesis consists of six studies that focus on different aspects of the conformational, binding, elastic and diffusive properties of weak polyelectrolytes and biopolymers. Although the particular conclusions and advances on the field resulting of each study have been already discussed on its corresponding chapter, here the general conclusions are outlined.

It is well known that the conformational and binding properties of weak polyelectrolytes are highly coupled.^{14,16} In this thesis, I have used the Site Binding Rotational Isomeric State (SBRIS) model to study both properties at the same foot using a Semi-Grand Canonical Monte Carlo (SGCMC) code programmed by myself. Using this program, I have proved that the conformational and binding properties of a model flexible polyelectrolyte can be analytically predicted using local parameters, that include in an effective way the long range intramolecular electrostatic interaction using variational techniques. The resulting approach, named the Local Effective Interaction Parameters, has potential applications to rationalize experimental titration curves of weak polyelectrolytes and in the development of pH-dependent torsional potential. Moreover, it could be extended to study the effect of a competitive binding of the proton with other chemical species such as metals or other ligands.

I have also used the SGCMC code to study the elastic response of a flexible weak polyelectrolyte, which is described using the SBRIS model. I have observed that the force-extension curves are dependent on the pH and ionic strength conditions. As a general trend, the polyelectrolyte is easier stretched in the conditions that promote the polyelectrolyte ionization due to intramolecular repulsive electrostatic interaction. This effect is only relevant in the intermediate force regime, where the force activates the conformational degrees of freedom. This is evidenced by a change in the persistence length and the *gauche* state probabilities in this force regime. Moreover, it is observed that the persistence length of the polyelectrolyte chain becomes dependent of the stretching force at the intermediate and high force regimes. In turn, the alteration of the conformational properties of the polyelectrolyte chain causes an according change in its average degree of ionization in the same force regime. At constant pH and ionic strength conditions, the average degree of ionization increases when the polyelectrolyte is

stretched since the intramolecular repulsive electrostatic interaction is reduced. This theoretical observation could be applicable in the rational design of pH-sensitive devices. For instance, one could conceive a system where several weak polyelectrolytes have each of its chain ends attached to two different surfaces. By separating those surfaces, the chains would be consequently stretched releasing/capturing protons or other binding ligands which in turn could be coupled to other chemical reactions.

All the aforementioned results are consequence of the inherent ability to regulate its charge in response of external stimuli, such as pH or salt variations or the action of a stretching force. I have assessed the relevance of the resulting charge fluctuation in the conformational and elastic properties of a model flexible polyelectrolyte by performing Monte Carlo simulations keeping the charge constant but switching on/off its fluctuations. I have found that charge fluctuation has a significant impact on the conformational and elastic properties of the polyelectrolyte. As a general trend, the presence of fluctuations in the charge diminish the intramolecular repulsive electrostatic interactions, causing the polyelectrolyte to be more folded than in its absence. This is because the charge is no longer correlated, which impede compact conformations of the polyelectrolyte chain. Moreover, I have observed that the intensity of the fluctuations, measured with the binding capacitance, is an excellent indicator to discern the conditions where the polyelectrolyte charge is easier regulated. For instance, the pH and ionic strength conditions where the polyelectrolyte exhibits larger values of the binding capacitance coincide with those where larger changes in the charge of the polyelectrolyte is observed when it is stretched.

Another external stimuli affecting the physico-chemical properties of polymers is the presence of high concentrations of other macromolecules in solution⁸⁶. This fact is specially relevant in biopolymers, since in biological media a large volume fraction (up to the 40%) is occupied by macromolecules, conditions known as macromolecular crowding.^{83,85} In this thesis, I have studied how the diffusive, conformational and binding properties of biopolymers are altered in presence of macromolecular crowding.

On one hand, I have investigated the diffusion of two globular proteins, α -chymiotrypsin and streptavidin, in presence of varying concentration of different-sized dextran macromolecules. As a first step, I modelled the macromolecules as hard spheres with an effective radius, following a super coarse-grained approach. I measured the diffusive properties of the proteins using a Brownian Dynamics code, which I programmed by myself, with different treatments of the hydrodynamics interactions. I found a qualitative agreement between the predictions of the simulations and the corresponding experimental results when treating the hydrodynamic interactions following the method proposed by Tokuyama.^{108–110} Then, I proposed a new single-sphere coarse grained model, named Chain Entanglement Softened Potential (CESP), that goes one step beyond the classic hard-sphere model. The CESP model describes the macromolecules

as spheres with two distinct regions: (i) a dense core, modeled as a hard-sphere; and (ii) a less dense outer part, which can be penetrated by other particles with an energetic penalty. The main novelty of the CESP model is that it allows to consider macromolecular compaction and entanglement in crowded media in an effective way. CESP model improved the previous results and it predicted quantitatively the streptavidin diffusive properties in different dextran solution. This result indicates that macromolecular entanglement and compaction in crowded media plays a relevant role in the biopolymer diffusion. The long-time diffusion coefficient D^{long} , the short-time diffusion coefficient (D^{short}) and the anomalous α are obtained from the BD simulations by using a new empirical expression, able to describe the full temporal evolution of the diffusion coefficient in crowded media.

On the other hand, I have assessed the effect of macromolecular crowding in the binding and structural properties of two Intrinsically Disordered Proteins (IDPs), histatin-5 and β -amyloid 42, which I modelled using a coarse-grained bead-and-spring model. The crowders are modelled after Bovin Serum Albumin (BSA) using the CESP model. When neutral crowders are added to the medium, the global charge of the IDPs is found to change depend on the effective increase in salt concentration due to excluded volume. When charged BSA crowders are introduced, the global charge of the IDPs is found to increase in all the pH range. In this case, the IDP charge is altered not only by the increase in the effective ionic strength but also by its direct electrostatic interaction with the charged crowders. For both neutral and charged crowders, the pH values for which the charge is more affected by crowding coincide with those of the maximum binding capacitance. This behaviour is common to the aforementioned finding that the charge of weak polyelectrolyte is more affected by the stretching force at the pH corresponding to those of maximum binding capacitance. Finally, the radius of gyration of the IDPs is found to decrease in presence of macromolecular crowding in all the pH-range.

List of Publications

Articles

- C. F. Narambuena, P. M. Blanco, A. Rodriguez, D. E. Rodriguez, S. Madurga, J. L. Garcés and F. Mas. Non-monotonic behavior of weak-polyelectrolytes adsorption on a cationic Surface: a Monte Carlo simulation study, *Polymer*, *accepted*.
- P. M. Blanco, S. Madurga, J. L. Garcés, F. Mas and R. S. Dias. Influence of Macromolecular Crowding in the Charge Regulation of Intrinsically Disordered Proteins, *Soft Matter*, 2020, DOI: 10.1039/D0SM01475C.
- P. M. Blanco, S. Madurga, C. F. Narambuena, F. Mas and J. L. Garcés. Role of Charge Regulation and Fluctuations in the Conformational and Mechanical Properties of Weak Flexible Polyelectrolytes. *Polymers*, 11:1962, 2019.
- P. M. Blanco, S. Madurga, F. Mas and J. L. Garcés. Effect of Charge Regulation and Conformational Equilibria in the Stretching Properties of Weak Polyelectrolytes. *Macromolecules*, 52:8017–8031, 2019.
- P. M. Blanco, S. Madurga, F. Mas and J. L. Garcés. Coupling of charge regulation and conformational equilibria in linear weak polyelectrolytes: Treatment of long-range interactions via effective short-ranged and pH-dependent interaction parameters. *Polymers*, 10(8):811, 2018.
- P. M. Blanco, J. L. Garcés, S. Madurga and F. Mas. Macromolecular diffusion in crowded media beyond the hard-sphere model. *Soft Matter*, 14(16):3105–3114, 2018.
- P. M. Blanco, M. Via, J. L. Garcés, S. Madurga, and F. Mas. Brownian dynamics computational model of protein diffusion in crowded media with dextran macromolecules as obstacles. *Entropy*, 19(3):105, 2017.

Book Chapters

- P. M. Blanco, S. Madurga, A. Isvoran, L. Pitulice and F. Mas. Fractal Dimension. In *New Frontiers in Nanochemistry: Concepts, Theories, and Trends*, Apple Academic Press, 171-186, 2020.

Appendix A

Supporting information to "Coupling of charge regulation and conformational equilibria in linear weak polyelectrolytes: treatment of long range interactions *via* effective short-ranged and pH-dependent interaction parameters"

A.1 Matricial expression for the SBRIS partition function for a linear polyelectrolyte

Let us express the SBRIS partition function as

$$\Xi_{\text{SBRIS}} = \sum_s \Xi_{\text{rot}}(s) \quad (\text{A.1})$$

where $\Xi_{\text{rot}}(s)$ is the partition function for the molecule in a binding state $s = \{s_1, s_2, \dots, s_N\}$. For simplicity, let us suppose that all the $N-1$ bonds hold two protonating sites at their ends.

The partition function for every 'frozen' binding configuration can then be expressed as a the RIS partition function but decorating the transfer matrices with suitable binding parameters. $\Xi_{\text{rot}}(s)$ adopts the form^{2,15,16}

$$\Xi_{\text{rot}}(s) = \mathbf{p} \prod_{i=1}^{N-1} \mathbf{U}_i^{s_i s_{i+1}} \mathbf{q}^T \quad (\text{A.2})$$

In the simplest case $\mathbf{U}_i^{00} = \mathbf{U}_i^{10}$; $\mathbf{U}_i^{01} = z \cdot \mathbf{U}_i$ and $\mathbf{U}_i^{11} = z \cdot \mathbf{U}_i \mathbf{u}_i$ where \mathbf{U}_i are the transfer matrices typical of the RIS model for a symmetric polymer

$$\mathbf{U}_i = \begin{pmatrix} 1 & \sigma & \sigma \\ 1 & \sigma\psi & \sigma\omega \\ 1 & \sigma\omega & \sigma\psi \end{pmatrix}_i \quad (\text{A.3})$$

where z is the reduced activity of the site and \mathbf{u} is a diagonal matrix containing the Boltzmann factors corresponding to the short range interactions between charged sites

$$\mathbf{u}_i = \begin{pmatrix} u_t & 0 & 0 \\ 0 & u_g & 0 \\ 0 & 0 & u_g \end{pmatrix}_i \quad (\text{A.4})$$

$-k_B T \ln u_t$ and $-k_B T \ln u_g$ represent the short range interaction energy between two sites separated by a bond in *trans* and *gauche* conformation, respectively. The next step is to calculate the sum in Eqn. (A.1), which can be done using the identity

$$\sum_s \left(\prod_{i=1}^{N-1} \mathbf{U}^{s_i s_{i+1}} \right) = (\mathbf{E} \mathbf{E}) \prod_{i=1}^{N-1} \begin{pmatrix} \mathbf{U}_i^{00} & \mathbf{U}_i^{01} \\ \mathbf{U}_i^{10} & \mathbf{U}_i^{11} \end{pmatrix} \begin{pmatrix} \mathbf{E} \\ \mathbf{E} \end{pmatrix} \quad (\text{A.5})$$

where \mathbf{E} is the 3×3 identity matrix. Combining Eqns. (A.5), (A.1) and (A.2), we obtain the SBRIS partition function.

$$\Xi_{\text{SBRIS}} = \mathbf{s} \prod_{i=1}^{N-1} \begin{pmatrix} \mathbf{U}_i^{00} & \mathbf{U}_i^{01} \\ \mathbf{U}_i^{10} & \mathbf{U}_i^{11} \end{pmatrix} \mathbf{t}^T \quad (\text{A.6})$$

where $\mathbf{s} = (\mathbf{p} \mathbf{p})$ and $\mathbf{t} = (\mathbf{q} \mathbf{q})$. Note that the SBRIS partition function is obtained from the RIS partition function by replacing

$$\mathbf{U} \rightarrow \begin{pmatrix} \mathbf{U} & \mathbf{U}_z \\ \mathbf{U} & \mathbf{U}_{uz} \end{pmatrix}; \mathbf{p} \rightarrow (\mathbf{p} \mathbf{p}); \mathbf{q} \rightarrow (\mathbf{q} \mathbf{q}) \quad (\text{A.7})$$

It can be easily shown that if some of the matrices in Eqn. (A.2) does not depend on any index s_i , the proper substitution is

$$\mathbf{U} \rightarrow \begin{pmatrix} \mathbf{U} & \mathbf{0} \\ \mathbf{0} & \mathbf{U} \end{pmatrix} \quad (\text{A.8})$$

which can be necessary for the transfer matrices corresponding to bonds which do not hold any ionizable site.

A.2 Calculation of the mean square distance between two nodes of the chain

The matrix summation trick used above can also be applied to calculate other physical quantities which become pH-dependent in the SBRIS approach. Let $f(c, s)$ be any quantity and $\langle f(c, s) \rangle$ its SBRIS thermal average. Then

$$\langle f(s, c) \rangle = \sum_{s,c} f(s, c) \frac{e^{-\beta F(s,c)}}{\Xi_{\text{SBRIS}}} = \sum_s \pi(s) \left(\sum_c \frac{f(s, c) e^{-\beta F(s,c)}}{\Xi_{\text{rot}}(s)} \right) = \sum_s \pi(s) \langle f(c, s) \rangle_c \quad (\text{A.9})$$

where

$$\pi(s) = \frac{\Xi_{\text{rot}}(s)}{\Xi_{\text{SBRIS}}} \quad (\text{A.10})$$

represents the probability of a protonation state s and $\langle f(c, s) \rangle_c$ is the average for that fixed protonation state. Note now that the quantity between brackets in Eqn. (A.9) represents a typical RIS average which can be calculated using the (conveniently decorated) transfer matrices $\mathbf{U}_i^{s_i s_{i+1}}$. In particular, matricial expressions for the mean square distance between two nodes, k and l , of a linear chain are available^{2,4}. For a fixed ionization state, the result is

$$\langle d_{kl}^2 \rangle(s) = \frac{2}{\Xi_{\text{rot}}(s)} \mathbf{p} \prod_{i=1}^k \mathbf{U}_i^{s_i s_{i+1}} [\mathbf{E} \mathbf{0} \mathbf{0} \mathbf{0} \mathbf{0}] \prod_{r=k+1}^l \mathbf{G}_r^{s_r s_{r+1}} \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \mathbf{E} \end{bmatrix} \prod_{j=l+1}^{N-1} \mathbf{U}_j^{s_j s_{j+1}} \mathbf{q} \quad (\text{A.11})$$

where the matrices \mathbf{G} are proper super-matrices which can be expressed in terms of the transfer matrices $\mathbf{U}_i^{s_i s_{i+1}}$, the translation matrices and the bond vectors. The details and derivations are given in the Chapter 4 of the classical Flory's book² (see Eqn. 35). Introducing (A.11) in (A.9) and using the summation (A.5), $\langle d_{kl}^2 \rangle$ is calculated as a function of the pH.

A.3 Minimal symmetric polyelectrolyte model

The model used in this work to illustrate the LEIP method consists of a polyelectrolyte like the one in Fig. 2b. In this molecule the ionizable sites are separated by three bonds **a**, **b** and **c**. The structure is very similar to the one of poly(oxyethylene) and the corresponding transfer matrices are given can be found in^{2,16}

$$\mathbf{U}_a = \begin{pmatrix} 1 & \sigma_a & \sigma_a \\ 1 & \sigma_a \psi & \sigma_a \omega \\ 1 & \sigma_a \omega & \sigma_a \psi \end{pmatrix}; \mathbf{U}_b = \begin{pmatrix} 1 & \sigma_a & \sigma_a \\ 1 & \sigma_a \psi' & \sigma_a \omega' \\ 1 & \sigma_a \omega' & \sigma_a \psi' \end{pmatrix}; \mathbf{U}_c = \begin{pmatrix} 1 & \sigma & \sigma \\ 1 & \sigma \psi & \sigma \omega \\ 1 & \sigma \omega & \sigma \psi \end{pmatrix}$$

The conformational model here used can be obtained by taking $\sigma_a = 0$, which assures that the **a** and **b** bonds are always in the *trans* state, and $\omega = \psi = \omega' = \psi' = 1$, which means that the bonds of the deprotonated molecule are independent.

Appendix B

Supporting information to “Effect of charge regulation and conformational equilibria in the stretching properties of weak polyelectrolytes”

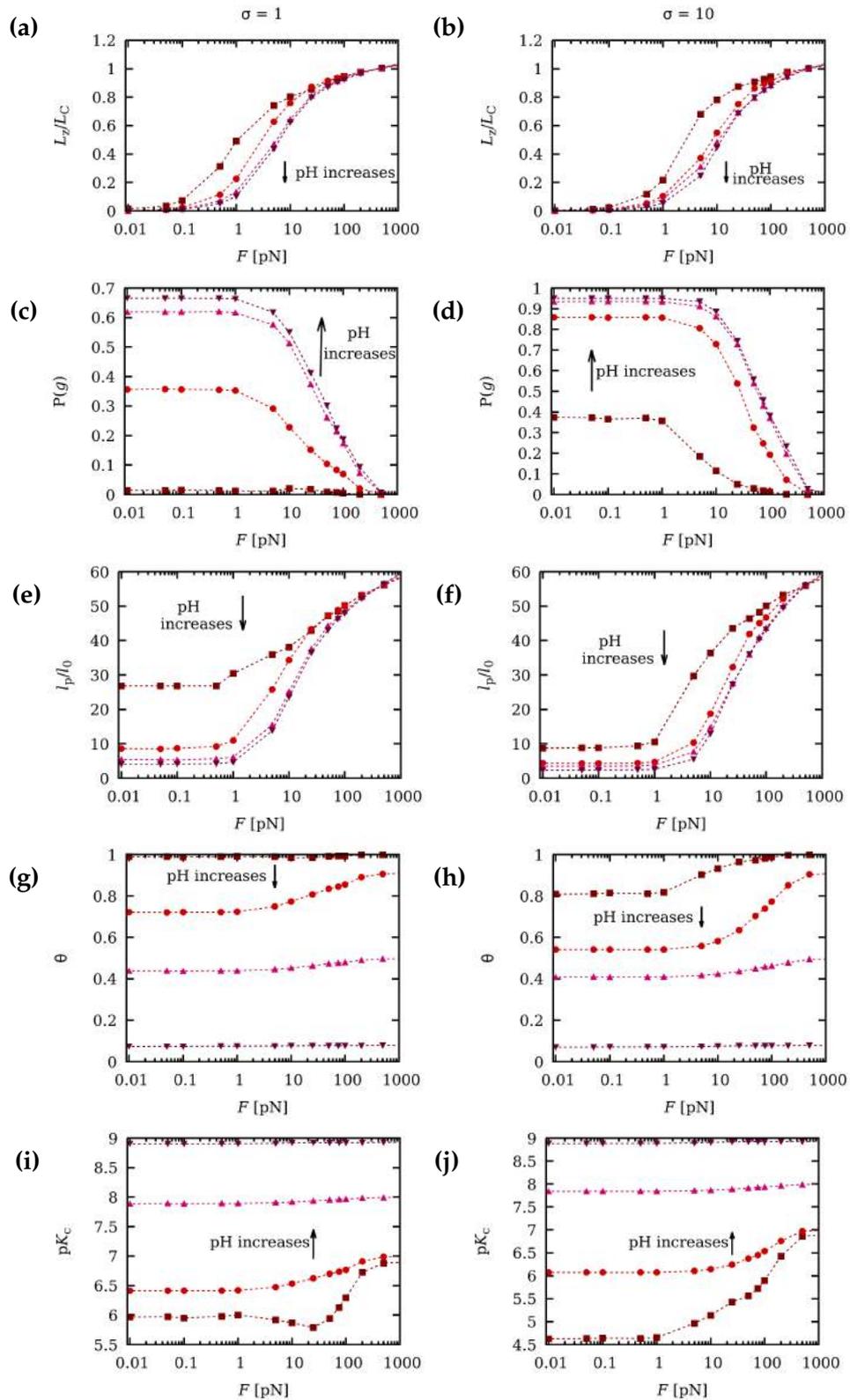


FIGURE B.1: (a and b) Chain extension (L_z), *gauche* state probability $P(g)$ (c and d), persistence length l_p (e and f), degree of protonation θ (g and h) and binding equilibrium constant K_c (i and j) as a function of the mechanical force F for $pK = 9$, $u_t = 10$, $u_g = 0$ and constant $I = 1M$. The simulations have been performed at four different pH-values: 4 (maroon squares), 6 (red circles), 8 (upwards pink triangles) and 10 (downwards purple triangles). The results on the left (a, c, e, g and i) correspond to the case with $\sigma = 1$ while the ones on the right (b, d, f, h and j) correspond to $\sigma = 10$. L_z is normalized by the polymer contour length $L_C = Nl_0 \cos(\frac{1}{2}(\pi - \alpha_0))$, where $l_0 = 1.5 \text{ \AA}$ and $\alpha_0 = 120^\circ$ are the equilibrium bond length and bond angle, respectively.

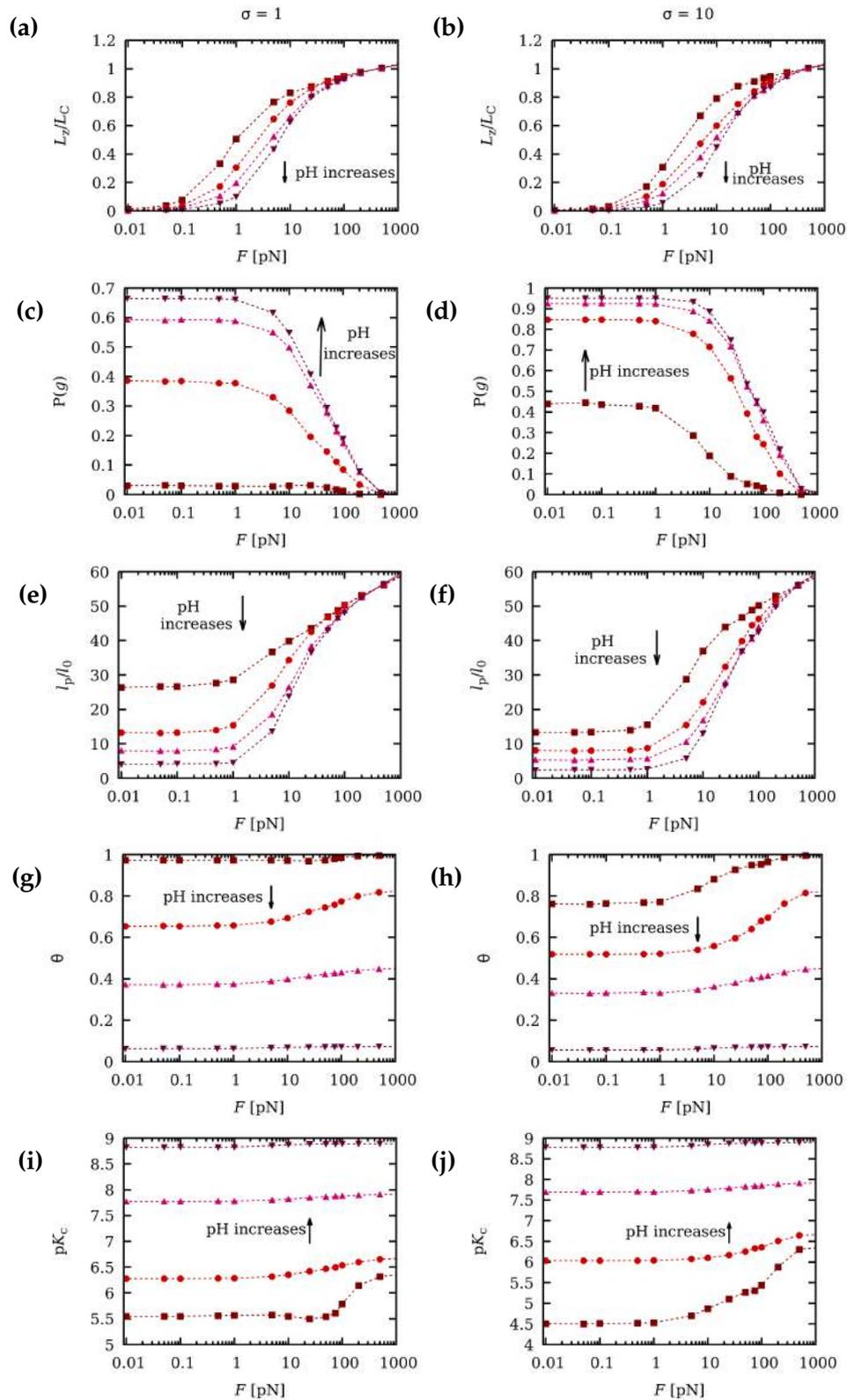


FIGURE B.2: (a and b) Chain extension (L_z), *gauche* state probability $P(g)$ (c and d), persistence length l_p (e and f), degree of protonation θ (g and h) and binding equilibrium constant K_c (i and j) as a function of the mechanical force F for $pK = 9$, $u_t = 10$, $u_g = 0$ and constant $I = 0.1M$. The simulations have been performed at four different pH-values: 4 (maroon squares), 6 (red circles), 8 (upwards pink triangles) and 10 (downwards purple triangles). The results on the left (a, c, e, g and i) correspond to the case with $\sigma = 1$ while the ones on the right (b, d, f, h and j) correspond to $\sigma = 10$. L_z is normalized by the polymer contour length $L_C = Nl_0 \cos(\frac{1}{2}(\pi - \alpha_0))$, where $l_0 = 1.5 \text{ \AA}$ and $\alpha_0 = 120^\circ$ are the equilibrium bond length and bond angle, respectively.

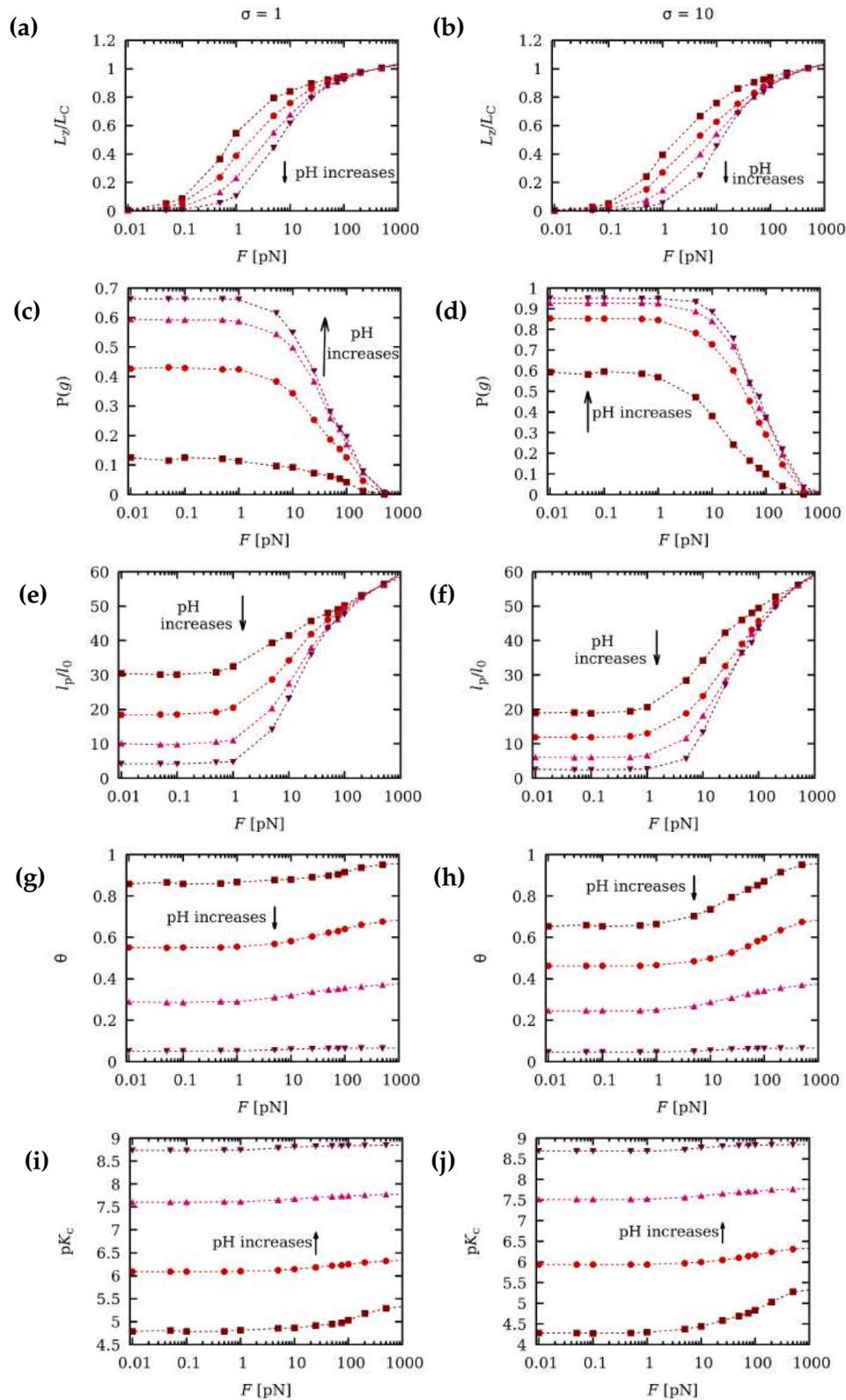


FIGURE B.3: (a and b) Chain extension (L_z), *gauche* state probability $P(g)$ (c and d), persistence length l_p (e and f), degree of protonation θ (g and h) and binding equilibrium constant K_c (i and j) as a function of the mechanical force F for $pK = 9$, $u_t = 10$, $u_g = 0$ and constant $I = 0.01M$. The simulations have been performed at four different pH-values: 4 (maroon squares), 6 (red circles), 8 (upwards pink triangles) and 10 (downwards purple triangles). The results on the left (a, c, e, g and i) correspond to the case with $\sigma = 1$ while the ones on the right (b, d, f, h and j) correspond to $\sigma = 10$. L_z is normalized by the polymer contour length $L_C = Nl_0 \cos(\frac{1}{2}(\pi - \alpha_0))$, where $l_0 = 1.5 \text{ \AA}$ and $\alpha_0 = 120^\circ$ are the equilibrium bond length and bond angle, respectively.

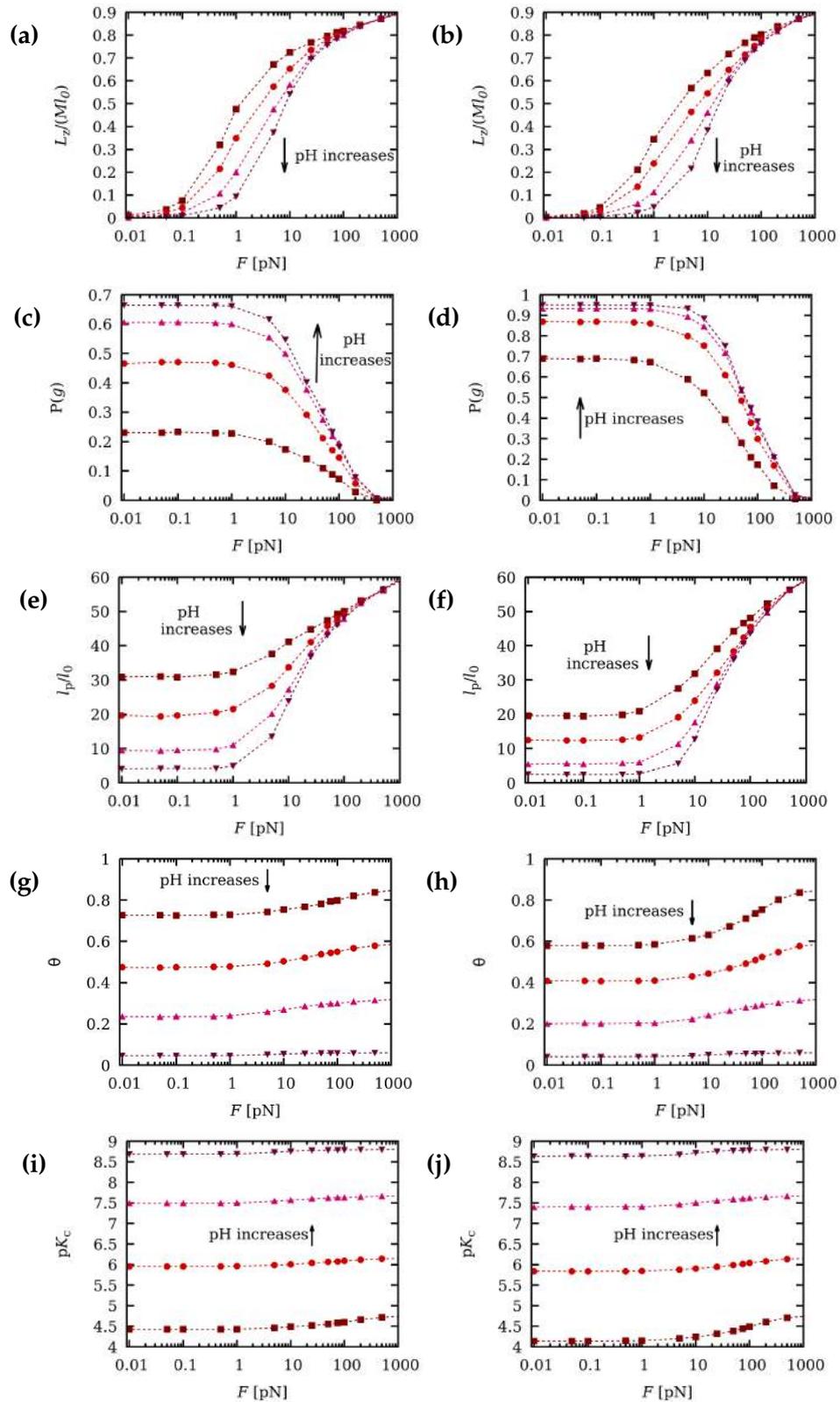


FIGURE B.4: (a and b) Chain extension (L_z), *gauche* state probability $P(g)$ (c and d), persistence length l_p (e and f), degree of protonation θ (g and h) and binding equilibrium constant K_c (i and j) as a function of the mechanical force F for $pK = 9$, $u_t = 10$, $u_g = 0$ and constant $I = 0.001M$. The simulations have been performed at four different pH-values: 4 (maroon squares), 6 (red circles), 8 (upwards pink triangles) and 10 (downwards purple triangles). The results on the left (a, c, e, g and i) correspond to the case with $\sigma = 1$ while the ones on the right (b, d, f, h and j) correspond to $\sigma = 10$. L_z is normalized by the polymer contour length $L_C = Nl_0 \cos(\frac{1}{2}(\pi - \alpha_0))$, where $l_0 = 1.5 \text{ \AA}$ and $\alpha_0 = 120^\circ$ are the equilibrium bond length and bond angle, respectively.

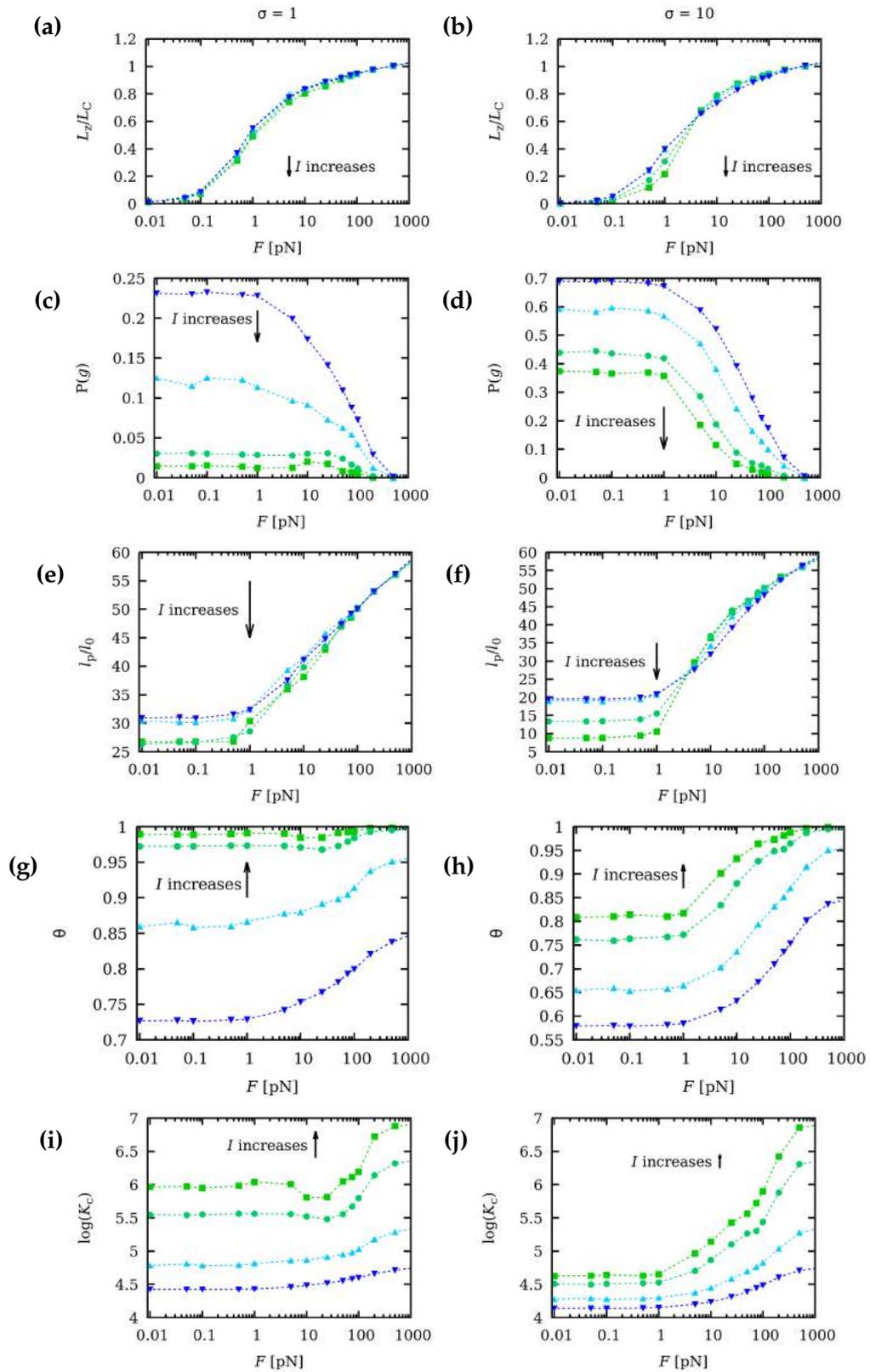


FIGURE B.5: (a and b) Chain extension (L_z), *gauche* state probability $P(g)$ (c and d), persistence length l_p (e and f), degree of protonation θ (g and h) and binding equilibrium constant K_C (i and j) as a function of the mechanical force F for $pK = 9$, $u_t = 10$, $u_g = 0$ and constant pH-value = 4. The simulations have been performed at four different ionic strengths: 1 M (green squares), 0.1 M (turquoise circles), 0.01 M (upwards cyan triangles) and 0.001 M (downwards blue triangles). The results on the left (a, c, e, g and i) correspond to the case with $\sigma = 1$ while the ones on the right (b, d, f, h and j) correspond to $\sigma = 10$. L_z is normalized by the polymer contour length $L_C = Nl_0 \cos(\frac{1}{2}(\pi - \alpha_0))$, where $l_0 = 1.5 \text{ \AA}$ and $\alpha_0 = 120^\circ$ are the equilibrium bond length and bond angle, respectively.

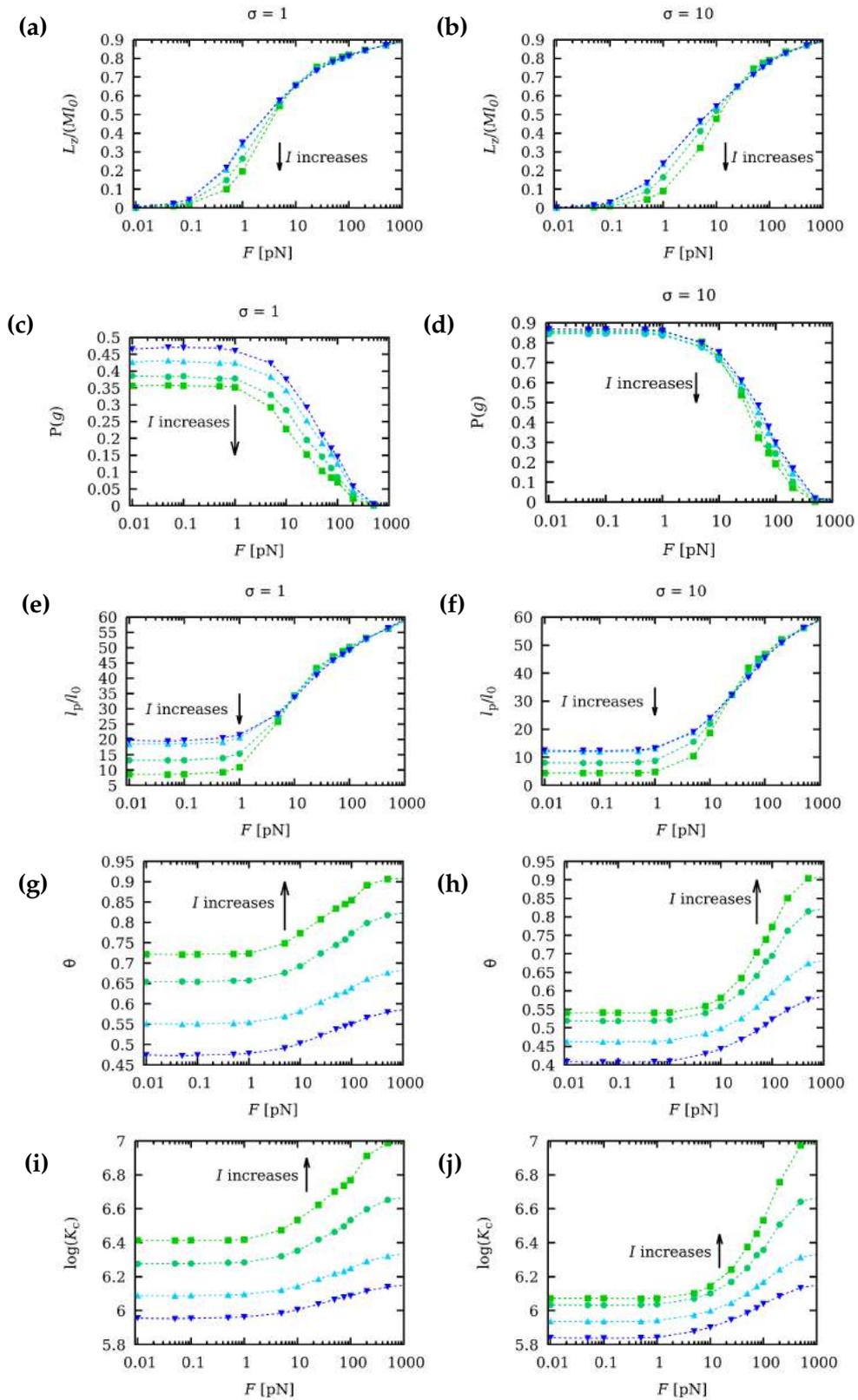


FIGURE B.6: (a and b) Chain extension (L_z), *gauche* state probability $P(g)$ (c and d), persistence length l_p (e and f), degree of protonation θ (g and h) and binding equilibrium constant K_c (i and j) as a function of the mechanical force F for $pK = 9$, $u_t = 10$, $u_g = 0$ and constant pH-value = 6. The simulations have been performed at four different ionic strengths: 1 M (green squares), 0.1 M (turquoise circles), 0.01 M (upwards cyan triangles) and 0.001 M (downwards blue triangles). The results on the left (a, c, e, g and i) correspond to the case with $\sigma = 1$ while the ones on the right (b, d, f, h and j) correspond to $\sigma = 10$. L_z is normalized by the polymer contour length $L_C = Nl_0 \cos(\frac{1}{2}(\pi - \alpha_0))$, where $l_0 = 1.5 \text{ \AA}$ and $\alpha_0 = 120^\circ$ are the equilibrium bond length and bond angle, respectively.

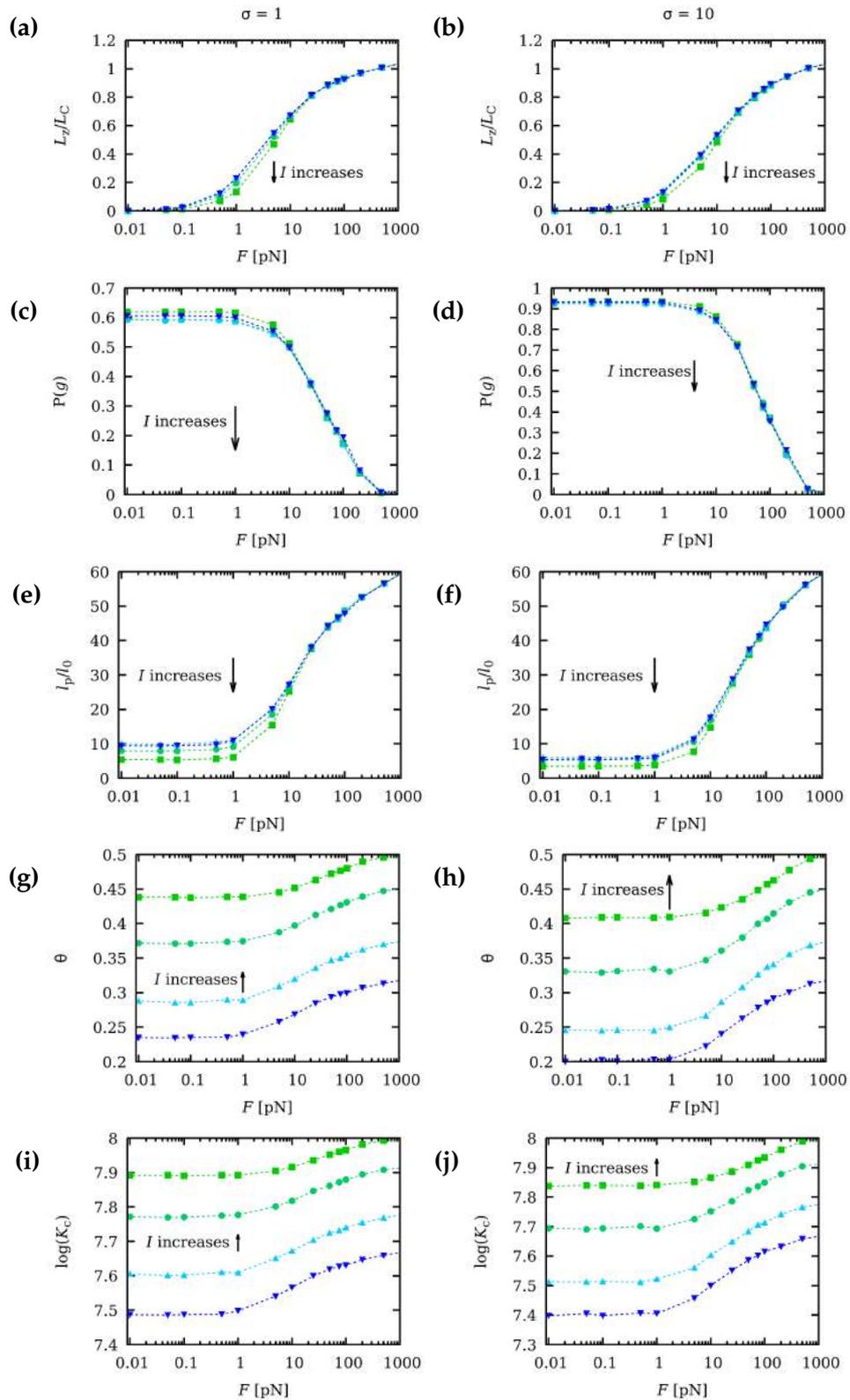


FIGURE B.7: (a and b) Chain extension (L_z), *gauche* state probability $P(g)$ (c and d), persistence length l_p (e and f), degree of protonation θ (g and h) and binding equilibrium constant K_C (i and j) as a function of the mechanical force F for $pK = 9$, $u_t = 10$, $u_g = 0$ and constant pH-value = 8. The simulations have been performed at four different ionic strengths: 1 M (green squares), 0.1 M (turquoise circles), 0.01 M (upwards cyan triangles) and 0.001 M (downwards blue triangles). The results on the left (a, c, e, g and i) correspond to the case with $\sigma = 1$ while the ones on the right (b, d, f, h and j) correspond to $\sigma = 10$. L_z is normalized by the polymer contour length $L_C = Nl_0 \cos(\frac{1}{2}(\pi - \alpha_0))$, where $l_0 = 1.5 \text{ \AA}$ and $\alpha_0 = 120^\circ$ are the equilibrium bond length and bond angle, respectively.

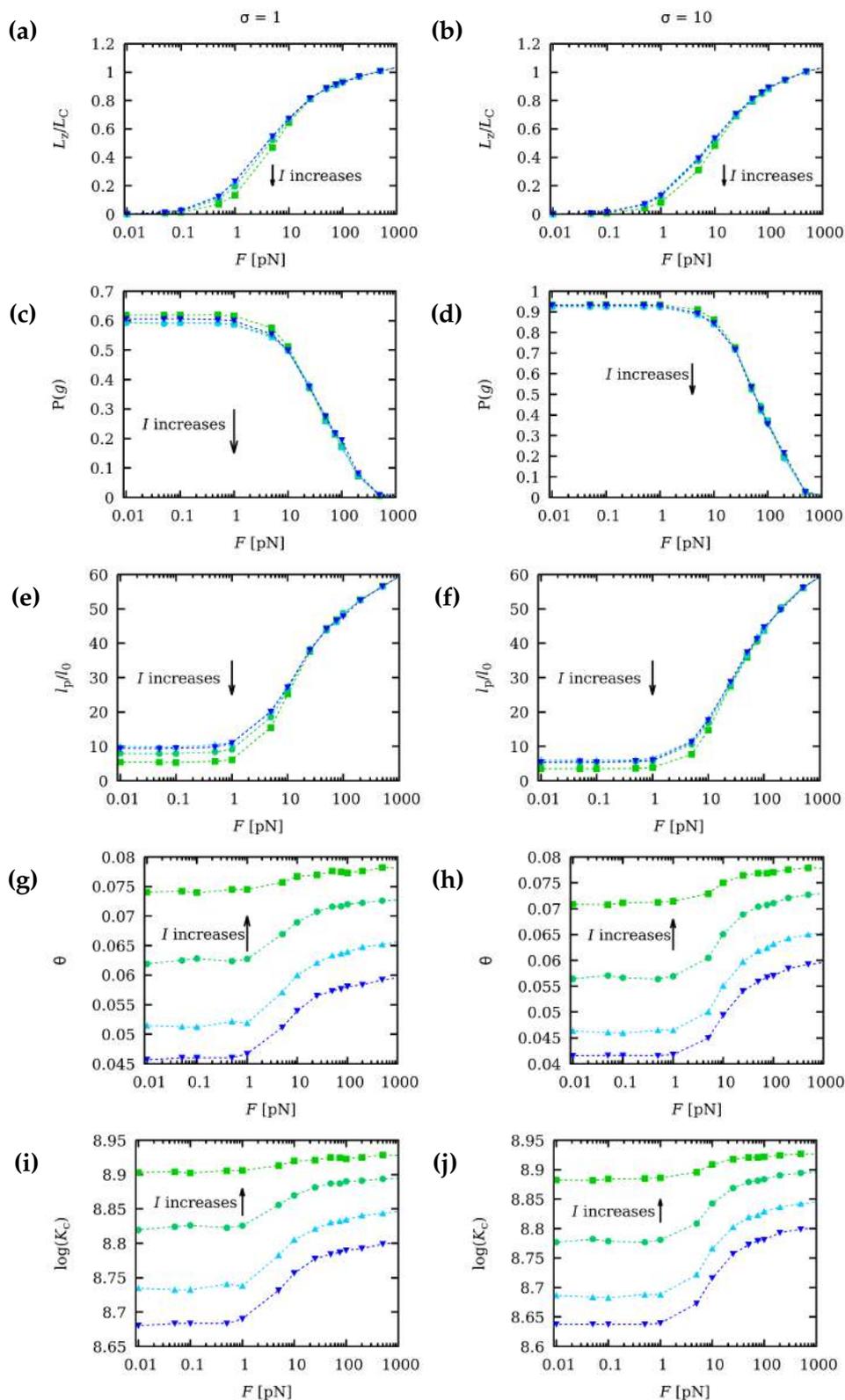


FIGURE B.8: (a and b) Chain extension (L_z), *gauche* state probability $P(g)$ (c and d), persistence length l_p (e and f), degree of protonation θ (g and h) and binding equilibrium constant K_C (i and j) as a function of the mechanical force F for $pK = 9$, $u_t = 10$, $u_g = 0$ and constant pH-value = 10. The simulations have been performed at four different ionic strengths: 1 M (green squares), 0.1 M (turquoise circles), 0.01 M (upwards cyan triangles) and 0.001 M (downwards blue triangles). The results on the left (a, c, e, g and i) correspond to the case with $\sigma = 1$ while the ones on the right (b, d, f, h and j) correspond to $\sigma = 10$. L_z is normalized by the polymer contour length $L_C = Nl_0 \cos(\frac{1}{2}(\pi - \alpha_0))$, where $l_0 = 1.5 \text{ \AA}$ and $\alpha_0 = 120^\circ$ are the equilibrium bond length and bond angle, respectively.

Appendix C

Supporting information to “Role of charge regulation and fluctuations in the conformational and mechanical properties of weak flexible polyelectrolytes”

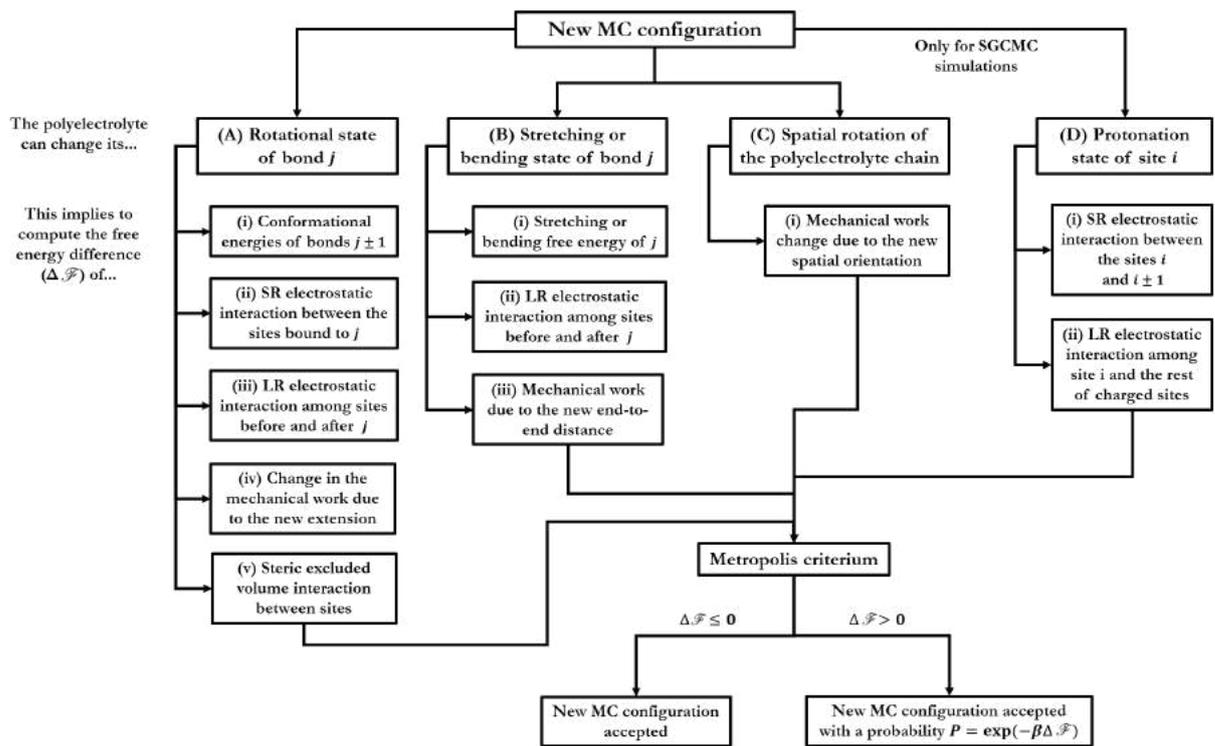


FIGURE C.1: Outline of the metropolis algorithm of the simulation code. In each new Monte Carlo (MC) configuration, the polyelectrolyte can change either (A) the rotational state of a bond, (B) the length or angle of a bond or (C) the spatial orientation of the polyelectrolyte chain in laboratory coordinate frame. In the case of Semi-Grand Canonical Monte Carlo (SGCMC) simulations, the polyelectrolyte can also alter the ionization state of a binding site (D).

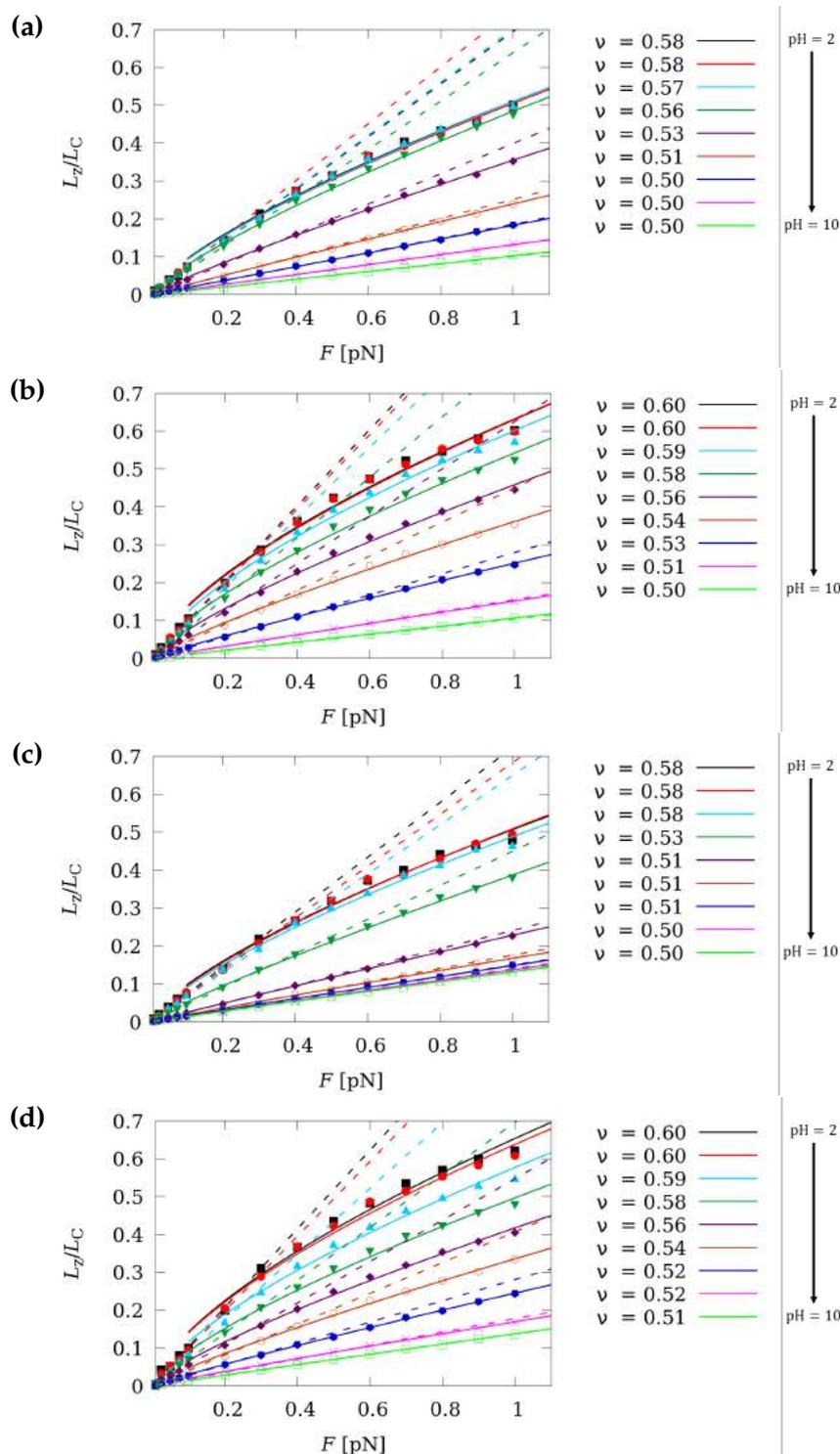


FIGURE C.2: Force F vs. extension L_z curves in the low force regime obtained at different pH-values of 2 (black squares), 3 (red circles), 4 (cyan upwards triangles), 5 (dark green downwards triangles), 6 (purple diamonds), 7 (ochre empty pentagons), 8 (blue filled pentagons), 9 (pink crosses) and 10 (light green empty squares). The markers show the ccMC(a and b) and the SMCMC with excluded volume (c and b) results at two different ionic strengths of 1 M (left) and 0.001 M (right). The dashed lines follow the force/extension linear prediction (Eq. 17) using the Kuhn length obtained from simulations at $F = 0$, whereas the continuous ones show the best fit of the computational data to the Pincus scaling law. L_z is normalized to the polyelectrolyte contour length $L_C = Nl_0 \cos((\pi - \alpha_0)/2)$.

Appendix D

Supporting information to

“Macromolecular diffusion in crowded media beyond the hard-sphere model”

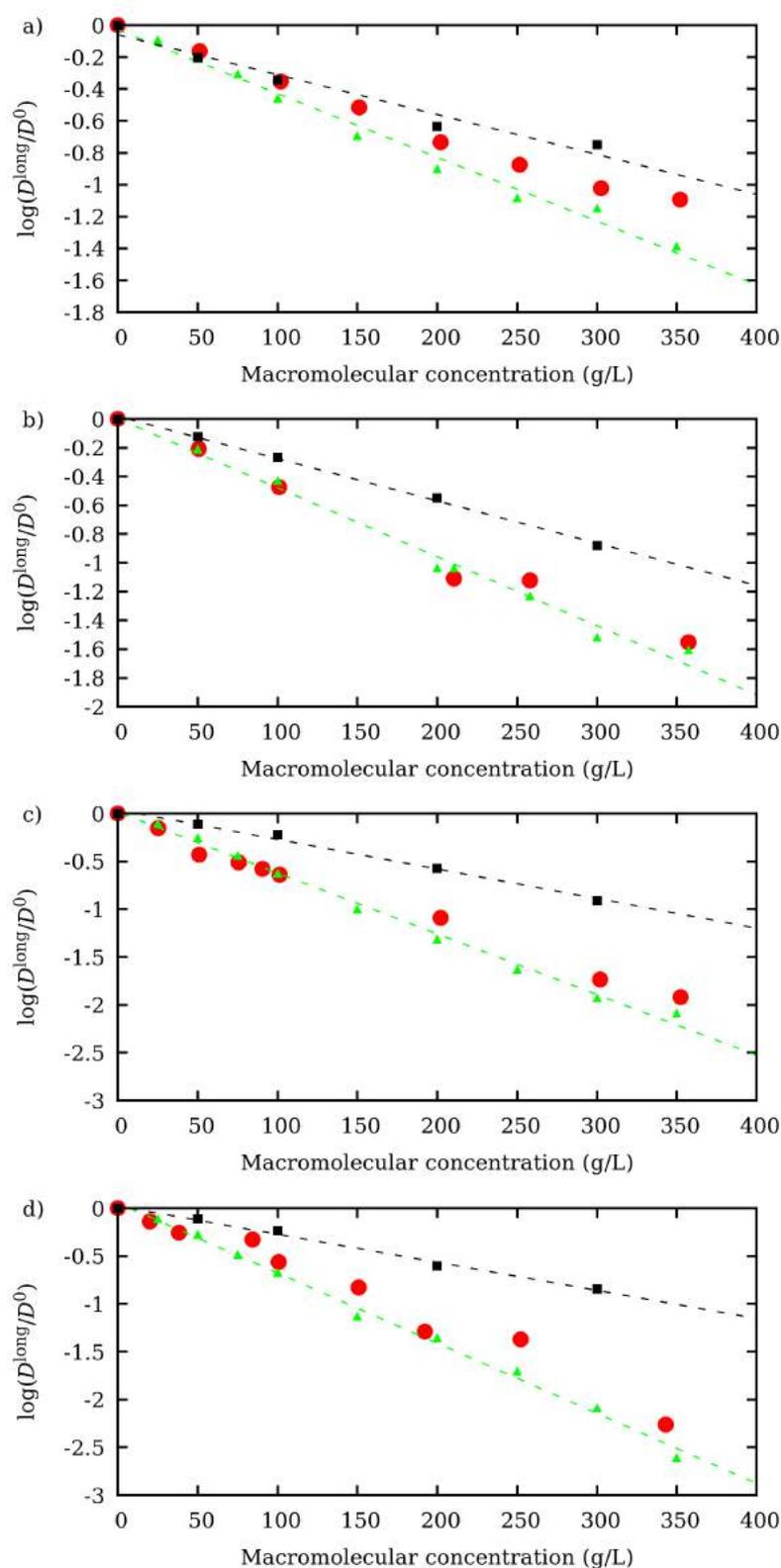


FIGURE D.1: $\log(D^{\text{long}}/D^0)$ for streptavidin *versus* macromolecular concentration for D10 (a), D50 (b), D400 (c) and D700 (d). Black squares represent BD simulation using hard-sphere model as explained in¹³⁷ and green triangles correspond to BD simulations using CESP model. Experimental values⁹⁷ are depicted as red circles. Dashed lines follow simulation results and are only to guide the eyes.

Appendix E

Supporting information to "Charge Regulation in Intrinsically Disordered Proteins due to Macromolecular Crowding"

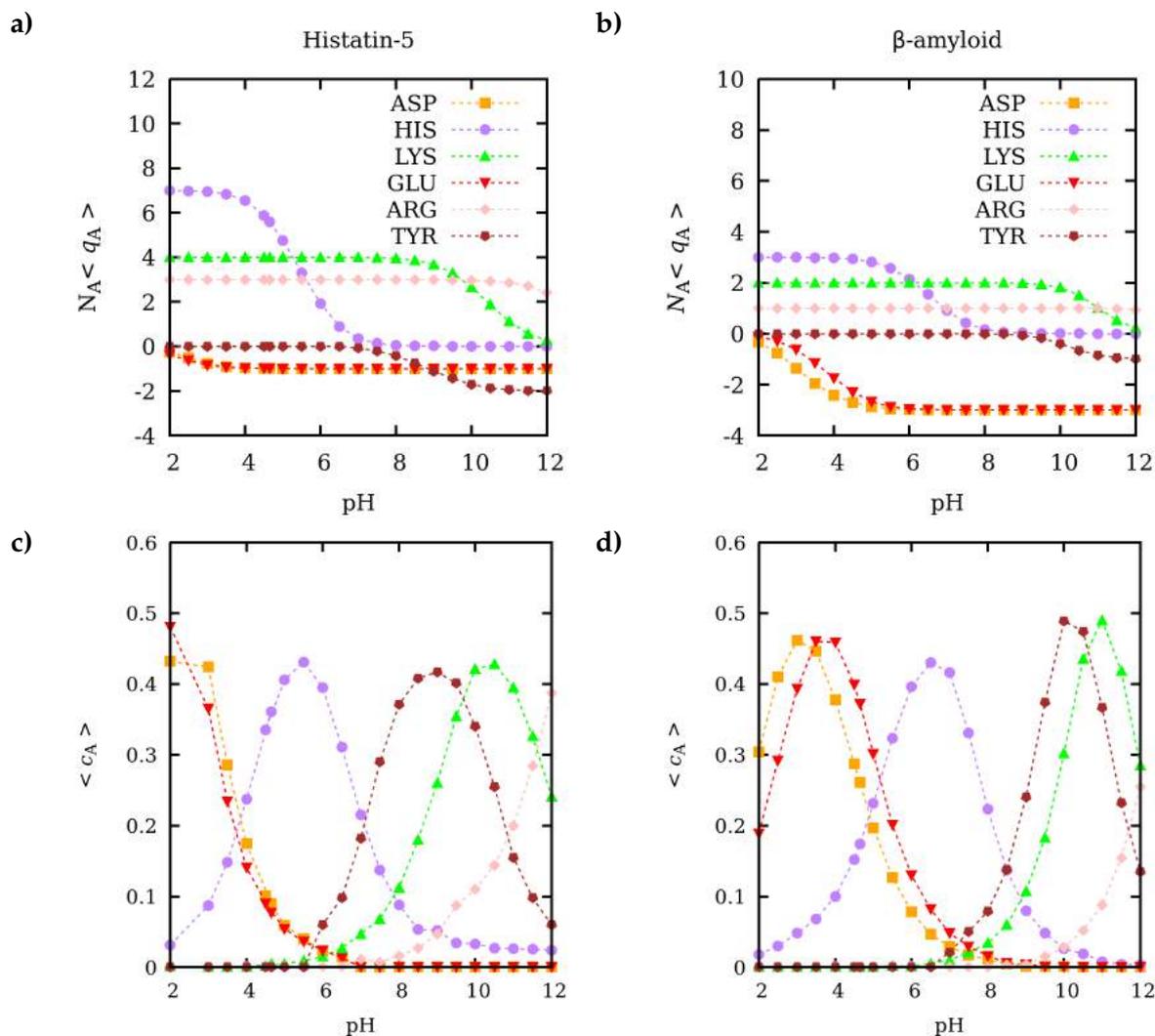
E.1 Aminoacid titration curves at $c_s = 0.01$ M without crowding agents

FIGURE E.1: Aminoacid total charge $N_A \langle q_A \rangle$ (panels a and b) and binding capacitance c_A (panels c and d) titration curves for the Intrinsically Disordered Proteins (IDPs) histatin-5 (panels a and c) and β -amyloid 42 (panels b and d). The obtained results for each aminoacid are depicted as follows: aspartic acid (ASP, orange squares), histidine (HIS, purple circles), lysine (LYS, green triangles), glutamic acid (GLU, red inverted triangles), arginine (ARG, pink diamonds) and tyrosine (TYR, pentagons). All data was computed by Semi-Grand Canonical Monte Carlo (SGCMC) simulations of each protein without crowding agents and an added salt concentration $c_s = 0.01$ M of NaCl.

E.2 Average charge variation of aspartic acid, histidine and lysine aminoacids for varying salt concentration

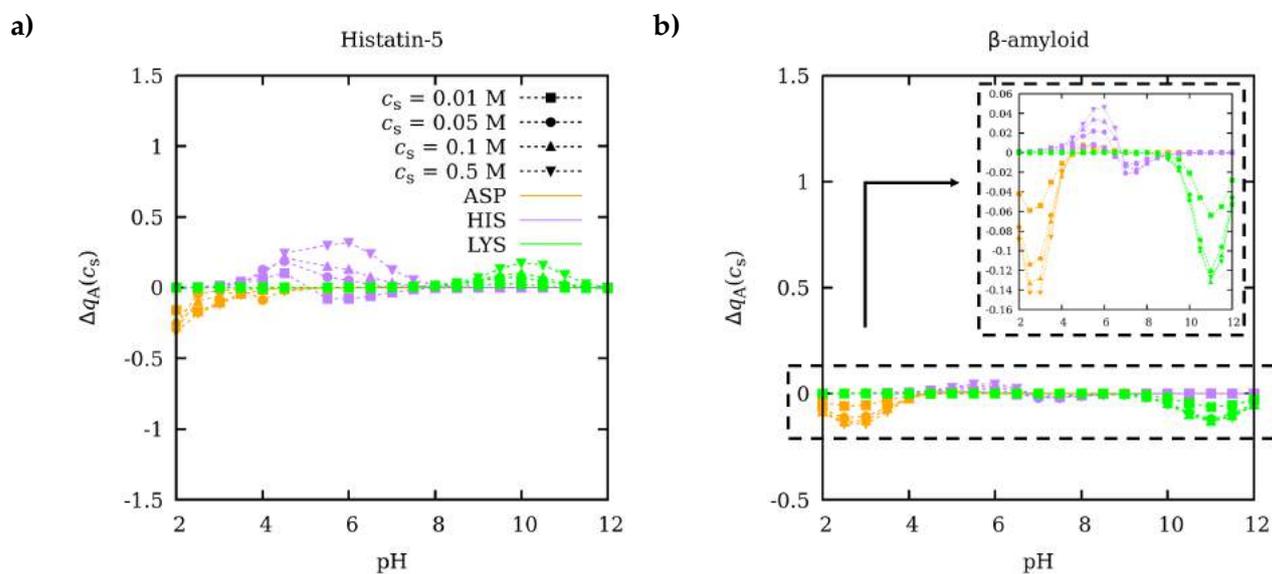


FIGURE E.2: Charge increment due to the added salt in the aminoacid type A $\Delta q_A(c_s) = |q_A(c_s)| - |q_A(0)|$ for aspartic acid (orange), histidine (purple) and lysine (green) as a function of the pH-value. The results were obtained with Semi-Grand Canonical Monte Carlo simulations without crowding agents, at constant added salt concentrations of 0.01 M (squares), 0.05 M (circles), 0.1 M (upwards triangles) and 0.5 M (downwards triangles).

E.3 Charge regulation in a medium with neutral crowders and salt concentration $c_s = 0.01$ M

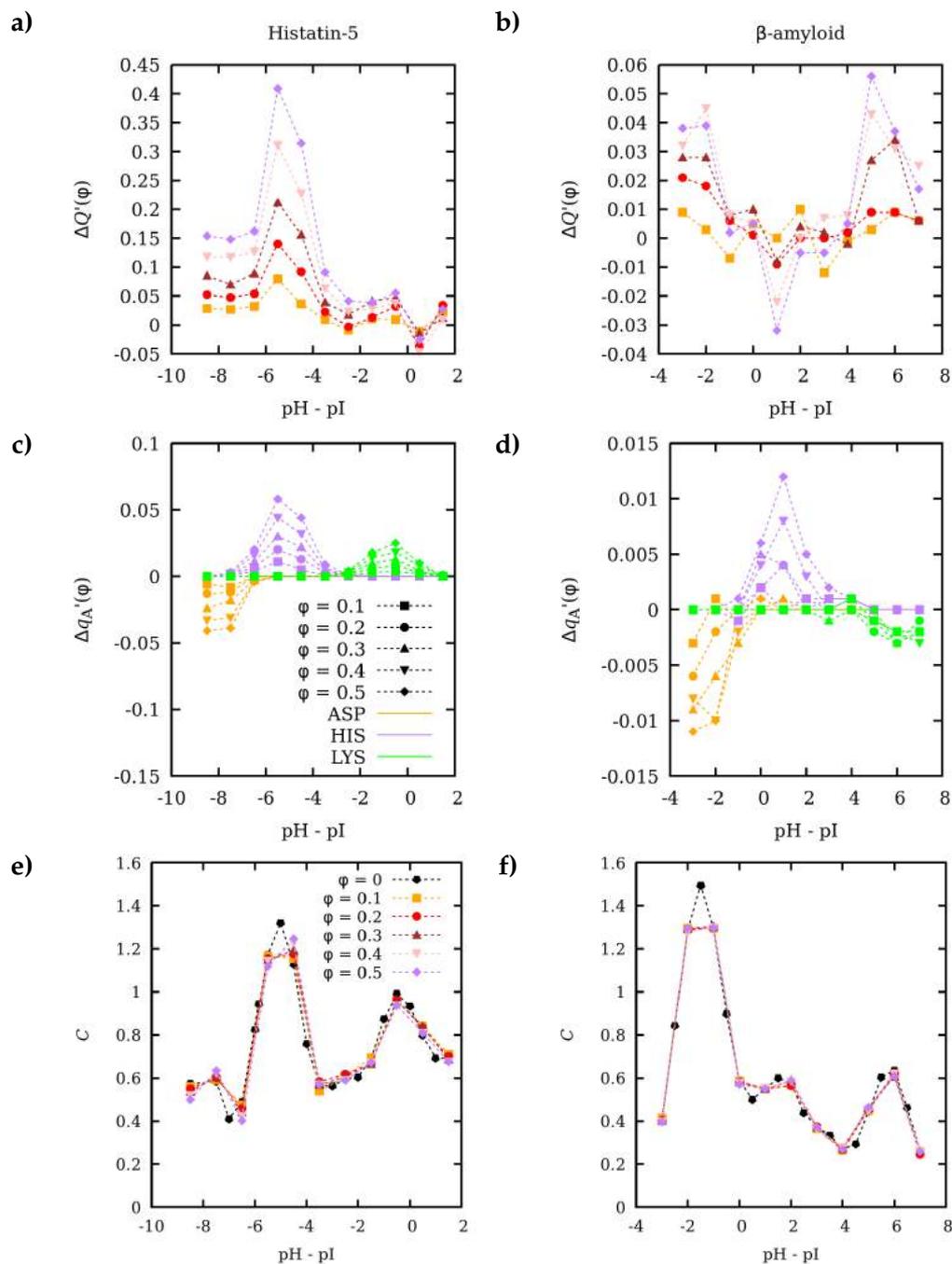


FIGURE E.3: (a and b) Increase in the absolute macromolecular charge due to the presence of neutral crowders, $\Delta Q'(\varphi) = |Q(\varphi)| - |Q(0)|$; (c and d) increase in the average charge per aminoacid, $\Delta q_A'(\varphi) = |q_A(\varphi)| - |q_A(0)|$; and (e and f) binding capacitance C . Panels on the left-hand side refer to histatin-5 and the ones on the right-hand side to β -amyloid 42. The added salt concentration is $c_s = 0.01$ M of NaCl in all the simulations, while φ ranges from 0.1 to 0.5.

The meaning of the markers used can be read inside the panels.

E.4 IDP global charge variation *vs.* excluded volume curves

For all the range of pH values, it is found that the curves can be very well fitted by the power law

$$\Delta Q'(\phi) = kC\phi^\alpha, \quad (\text{E.1})$$

where k and the α are empiric parameters (see the dashed lines in Figs. E.4 (neutral crowders) and (charged crowders)). For neutral crowders, the best fit k and α parameters are reported in Tab. E.1. It can be observed that $|k|$ ranges from 0.01 to 1.30 and α from 0.40 to 6.88. Interestingly, for the pH-values where crowding induces a significant charge regulation, α is close to one thus $\Delta Q'(\phi)$ scales almost linearly with ϕ . When charged crowders are considered, the best-fit parameters are listed in Tab. E.2. It is observed that k and α values range from 0.08 to 4.60 and from 0.21 to 2.0, respectively. It was not possible to properly fit the data corresponding to β -amyloid 42 at pH 4.65, since $\Delta Q'(\phi)$ was very small.

pH		2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00	11.00	12.0
Histatin 5	k	0.57	0.47	0.63	1.29	1.00	0.89	0.36	0.13	0.34	-0.59	-1.22
	α	1.02	1.15	1.06	1.27	1.31	1.41	0.918	0.934	1.41	2.65	7.90
β-amyloid	k	0.13	0.07	-1.01	-0.13	-0.31	-0.174	0.82	0.41	0.30	0.05	0.01
	α	1.50	0.67	4.48	1.45	1.25	1.14	6.88	1.48	1.06	0.40	1.56

TABLE E.1: k and α values obtained from the best fit of the $\Delta Q'(\phi)$ *vs.* ϕ curves for systems with neutral crowders (Fig. E.4), to the empiric power law (Eq. E.1). The standard error is always ~ 0.005 units

pH		3.00	3.50	4.00	4.50	4.65	5.00	6.00	7.00	8.00	9.00	10.0	11.0
Histatin 5	k	2.24	2.04	3.65	4.60	0.63	1.28	1.01	1.10	0.34	8.28	0.35	0.70
	α	0.55	0.52	0.80	1.00	1.00	0.84	0.38	0.31	0.10	12.69	0.22	0.91
β-amyloid	k	0.42	0.28	0.32	0.08	-	-0.21	0.04	0.12	0.43	0.63	1.68	1.88
	α	0.70	0.56	0.45	0.38	-	2.02	0.09	0.74	0.85	0.41	0.67	0.57

TABLE E.2: k and α values obtained by the best fit of $\Delta Q'(\phi)$ *vs.* ϕ curves (Fig. E.5) to the empiric power law (Eq. E.1) in the presence of charged crowders. The standard error is ~ 0.005 units.

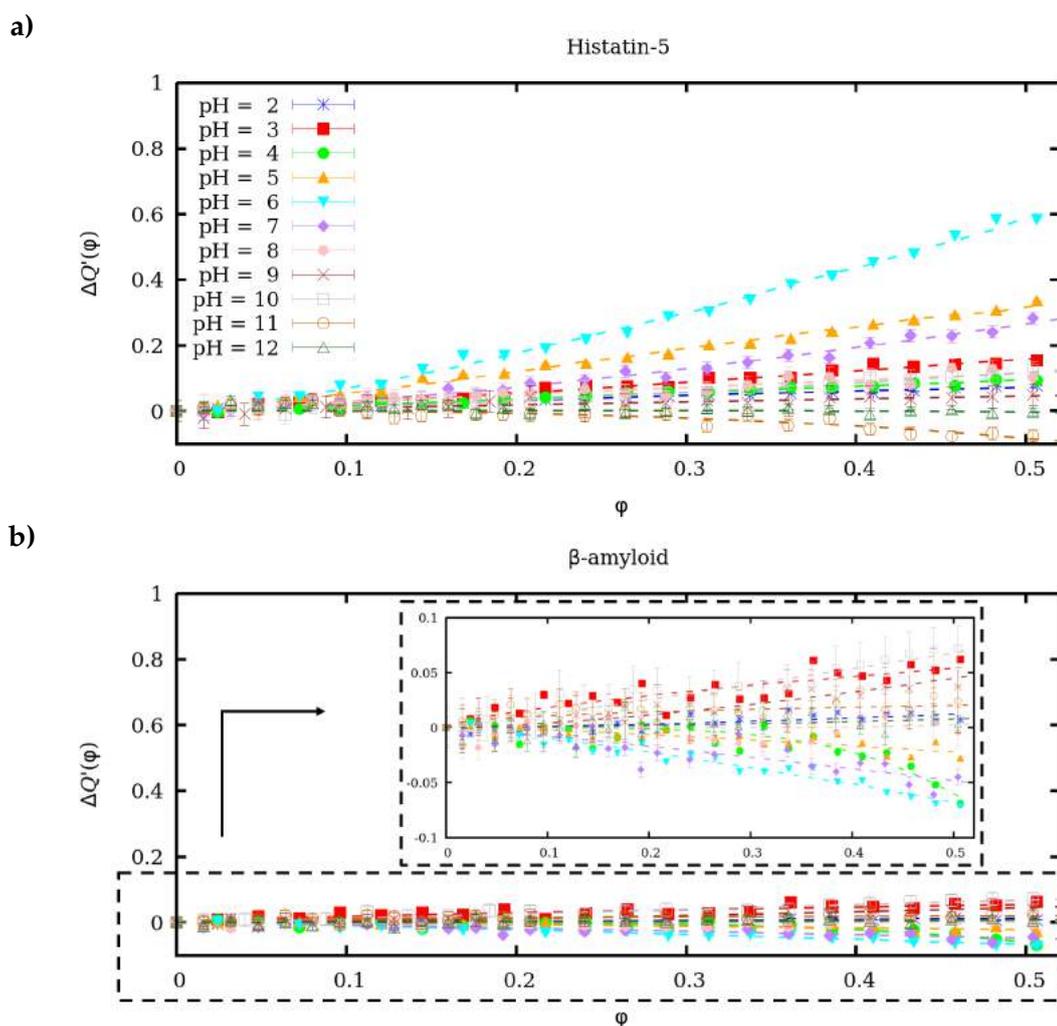


FIGURE E.4: Difference between absolute charges $\Delta Q'(\phi) = |Q(\phi)| - |Q(0)|$ for (a) histatin-5 and (b) β -amyloid 42 as a function of ϕ in the presence of neutral crowders at pH-values ranging from 3 to 11. The meaning of lines and markers is indicated in the top panel. Dashed lines correspond to the best-fit to the empiric power law (Eq. E.1). The resulting best-fit parameters are listed in Tab. E.1.

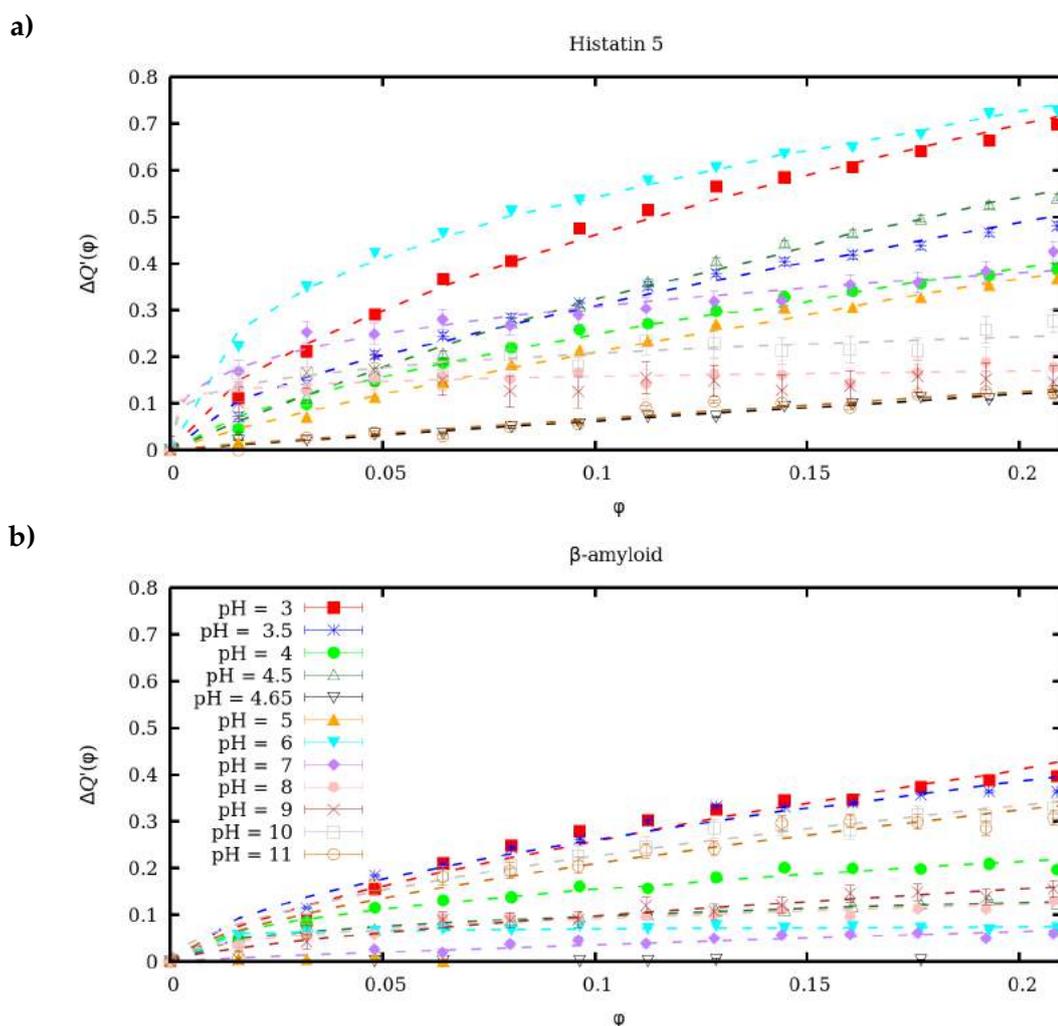
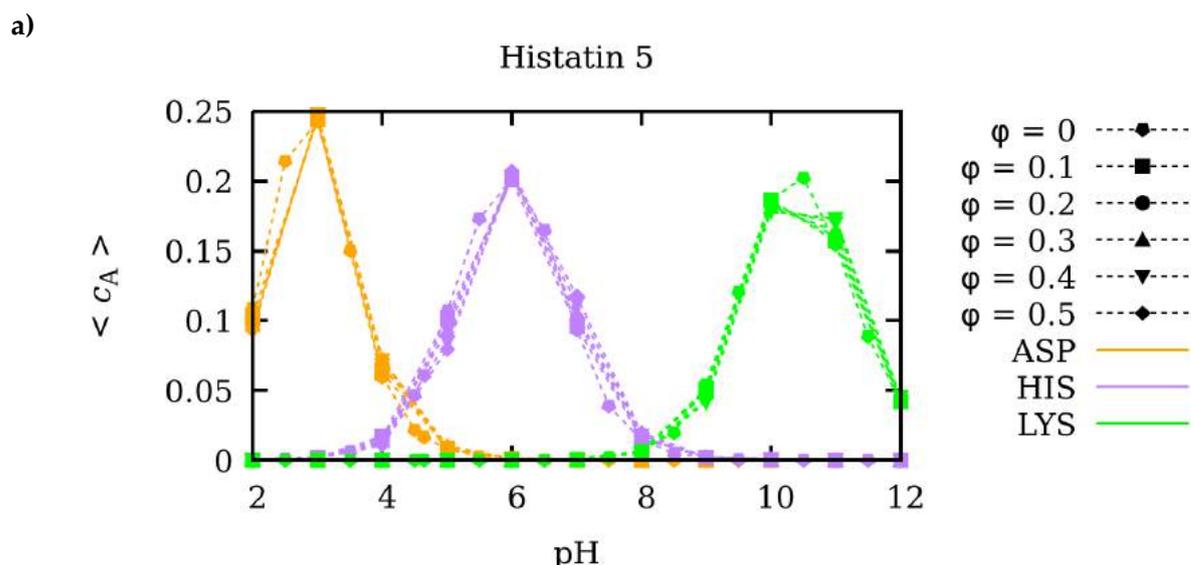


FIGURE E.5: Difference between absolute charges $\Delta Q'(\phi) = |Q(\phi)| - |Q(0)|$ for (a) histatin-5 and (b) β -amyloid 42 as a function of ϕ in the presence of charged crowders at pH-values ranging from 3 to 11. The meaning of the line colors and markers is indicated in the top panel. Dashed lines correspond to the best-fit to the empiric power law (Eq. E.1). The resulting best-fit parameters are listed in Tab. E.2.

E.5 Binding capacitance titrations curves per aminoacid type in a medium crowded with neutral crowders at $c_s = 0.1$ M



E.6 Best fit parameters to the scaling law (Eq. 13 in the main text) for the R_g vs. ϕ curves

pH	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00	10.00	11.00	12.0
Histatin 5	0.11	0.10	0.10	0.10	0.10	0.09	0.09	0.09	0.09	0.08	0.09
β-amyloid	0.09	0.09	0.09	0.08	0.09	0.09	0.09	0.09	0.09	0.09	0.09

TABLE E.3: Best fit c -values of the scaling law (Eq. 13 in the main text) to the computed R_g vs. ϕ curves for histatin 5 and β -amyloid 42 at different pH values and in presence of neutral crowders. The standard error is ~ 0.005 units

pH	3.00	3.50	4.00	4.50	4.65	5.00	6.00	7.00	8.00	9.00	10.0	11.0
Histatin 5	0.33	0.30	0.19	0.14	0.07	0.06	0.03	0.07	0.08	0.14	0.07	0.07
β-amyloid	0.15	0.12	0.10	0.09	0.07	0.08	0.10	0.12	0.13	0.12	0.15	0.17

TABLE E.4: List of the c -values obtained by the best-fit of the R_g vs. ϕ curves to the power law (Eq. 13 in the main text) in presence of charged crowders. The standard error is ~ 0.005 units

E.7 Binding capacitance per aminoacid titration curves in a media with charged crowders at $c_s = 0.01$ M

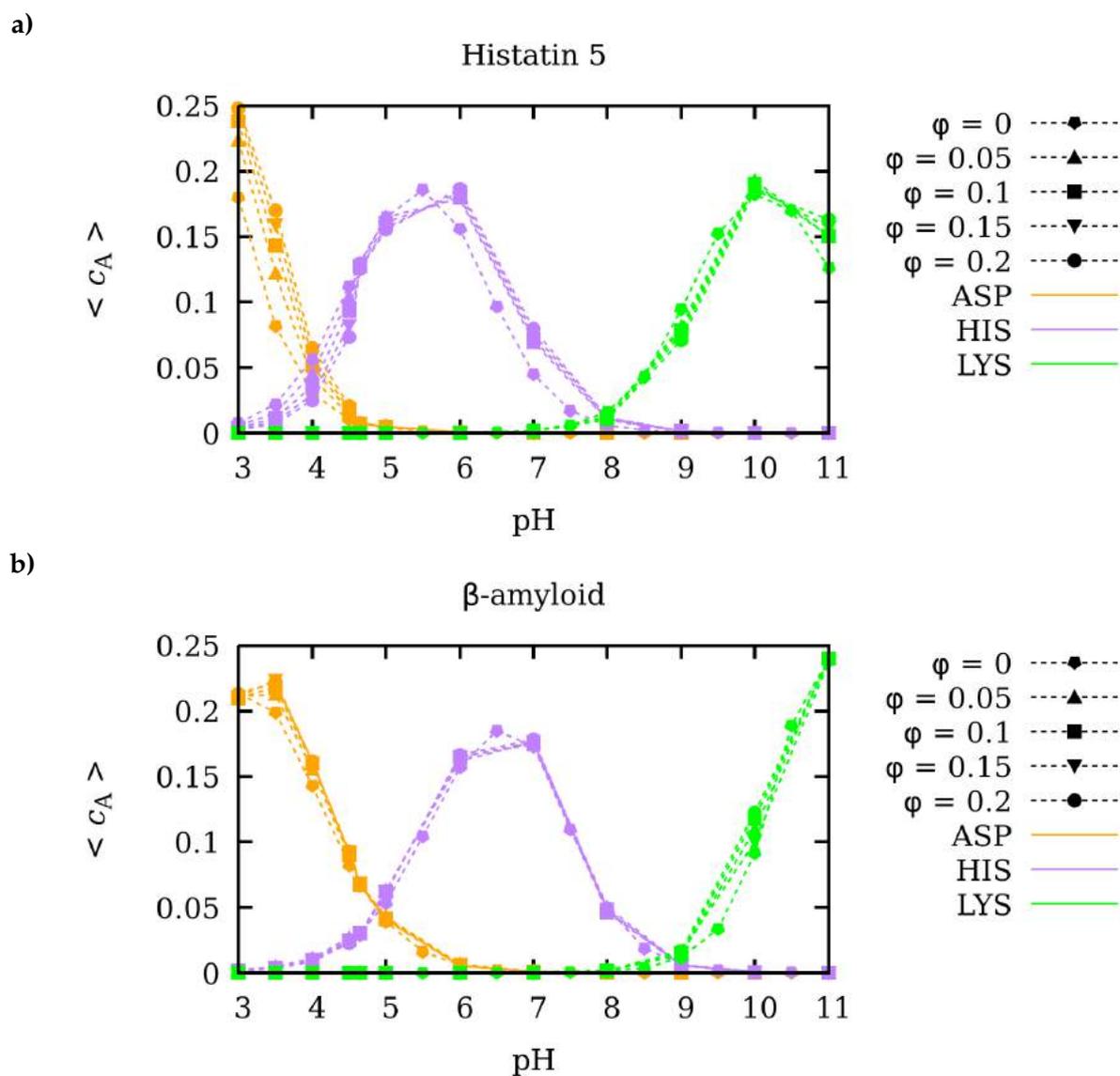


FIGURE E.7: Binding capacitance titration curves per aminoacid type $\langle c_A \rangle$ for histatin-5 (a) and β -amyloid 42 IDPs. For clarity, only the values of aspartic acid (orange), histidine (purple) and lysine (green) are shown as representatives of aminoacids with acidic, neutral and basic isoelectric points, respectively. The results were obtained by SGCMC simulations with a NaCl concentration $c_s = 0.01$ M and charged BSA crowders with excluded volume fraction ϕ of 0 (pentagons), 0.05 (upwards triangles), 0.1 (squares), 0.15 (downwards triangles) and 0.2 (circles).

E.8 Calculation of the effective ionic strength I' for the case with charged crowders

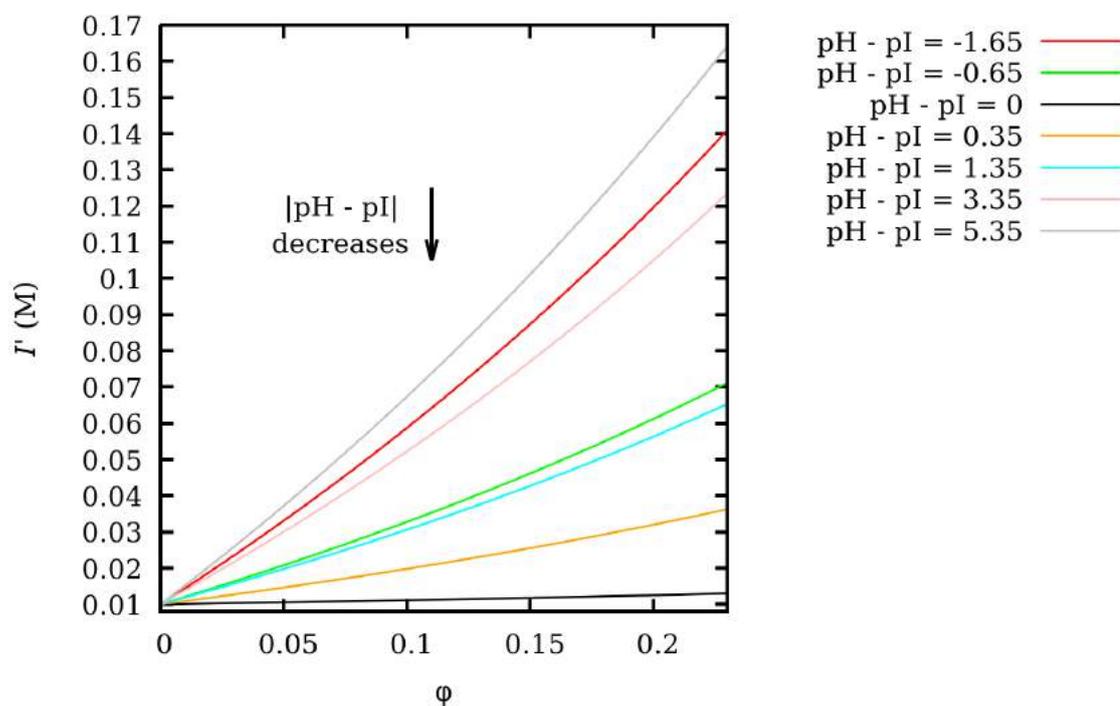


FIGURE E.8: Effective ionic strength I' as a function of the crowder excluded volume fraction ϕ at different pH-values. For clarity, the pH-values are referred to the Bovin Serum Albumin isoelectric point $\text{pI} = 4.65$.

E.9 Comparison of $\Delta Q(I' = 0.05\text{M})$ and $\Delta Q'(\phi)$ vs. pH for the case with charged crowders

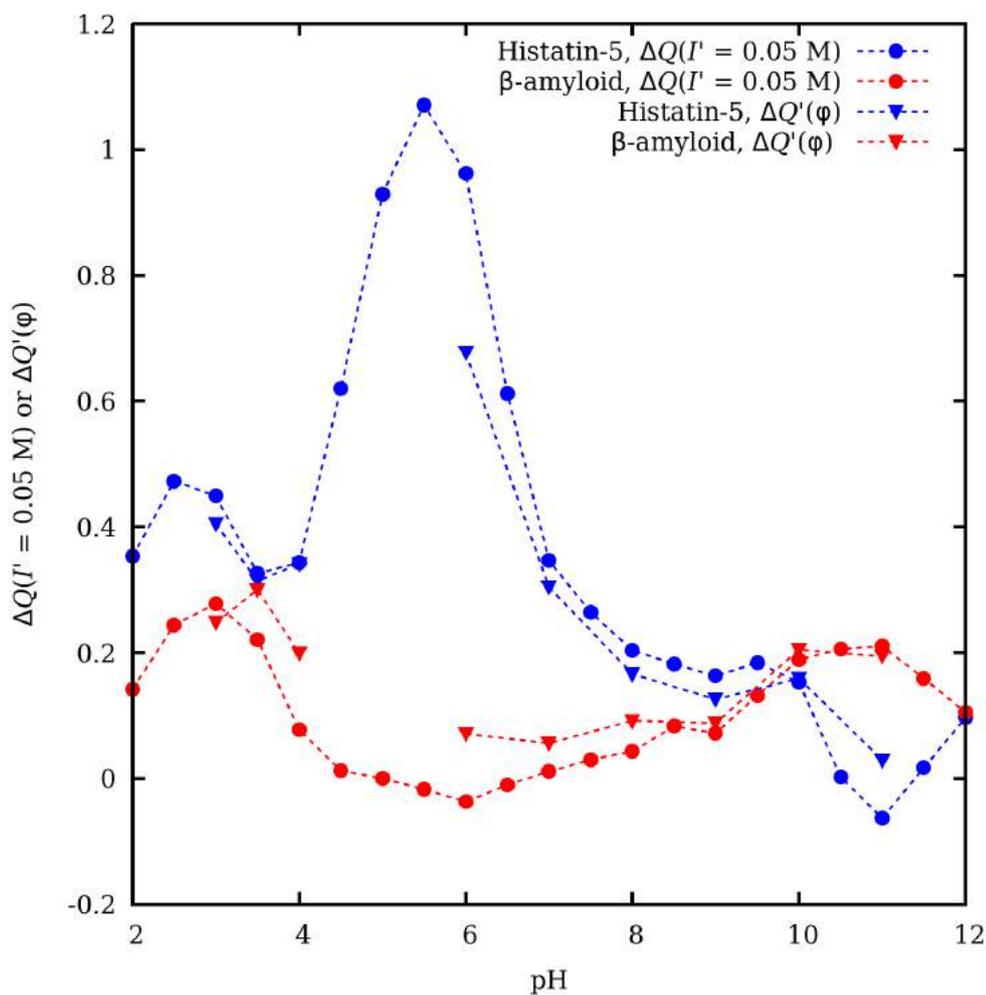


FIGURE E.9: Comparison between i) charge variation produced when the added salt concentration is incremented from $c_s = 0.01\text{ M}$ to 0.05 M without crowders ($\Delta Q(I' = 0.05\text{M}) = |Q(\phi = 0, c_s = 0.05\text{M})| - |Q(\phi = 0, c_s = 0.01\text{M})|$, circles); and ii) charge variation produced when adding charged crowders $\Delta Q'(\phi)$ (triangles) to a protein solution with added salt concentration $c_s = 0.01\text{ M}$. The excluded volume $\phi(\text{pH}, 0.05\text{M})$ is chosen such as the effective ionic strength I' (calculated using Eq.18) is always $I' = 0.05\text{ M}$ (circles). Blue color refers to Histatin-5 while red color refers to β -amyloid.

Bibliography

- [1] N. G. McCrum, C. P. Buckley, and C. B. Bucknall, *Principles of polymer engineering*. Oxford University Press, 1997.
- [2] P. J. Flory, *Statistical Mechanics of Chain Molecules*. Wiley, 1969.
- [3] B. A. Miller-chou and J. L. Koenig, "A review of polymer dissolution," *Progress in polymer science*, vol. 28, pp. 1223–1270, 2003.
- [4] P. J. Flory and R. L. Jernigan, "Second and Fourth Moments of Chain Molecules," *J. Chem. Phys.*, vol. 42, pp. 3509–3519, 1965.
- [5] L. Livadaru, R. R. Netz, and H. J. Kreuzer, "Stretching response of discrete semiflexible polymers," *Macromolecules*, vol. 36, no. 10, pp. 3732–3744, 2003.
- [6] G. Porod, "The mean separation of chain ends in threadlike molecules," *Monatsh. Chem.*, vol. 80, p. 251, 1949.
- [7] O. Kratky and G. Porod, "Röntgenuntersuchung gelöster fadenmoleküle," *Rec. Trav. Chim.*, vol. 68, p. 1106, 1949.
- [8] O. A. Saleh, "Single polymer mechanics across the force regimes," *Journal of Chemical Physics*, vol. 142, no. 19, pp. 1–9, 2015.
- [9] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*. Academic Press, 1982.
- [10] D. Chandler, *Introduction to Modern Statistical Mechanics*. Oxford University Press, 1987.
- [11] P. J. Flory, "Foundations of Rotational Isomeric State Theory and General Methods for Generating Configurational Averages," *Macromolecules*, vol. 7, no. 3, pp. 381–392, 1974.
- [12] M. Muthukumar, "50th Anniversary Perspective: A Perspective on Polyelectrolyte Solutions," *Macromolecules*, vol. 50, no. 24, pp. 9528–9560, 2017.
- [13] A. V. Dobrynin, "Polyelectrolytes: On the doorsteps of the second century," *Polymer*, vol. 202, p. 122714, 2020.

- [14] J. L. Garcés, G. J. Koper, and M. Borkovec, "Ionization Equilibria and Conformational Transitions in Polyprotic Molecules and Polyelectrolytes," *J. Phys. Chem. B*, vol. 110, pp. 10937–10950, 2006.
- [15] J. Lluís Garcés, C. Rey-Castro, C. David, S. Madurga, F. Mas, I. Pastor, and J. Puy, "Model-independent link between the macroscopic and microscopic descriptions of multidentate macromolecular binding: relationship between stepwise, intrinsic, and microscopic equilibrium constants.," *The journal of physical chemistry. B*, vol. 113, no. 46, pp. 15145–15155, 2009.
- [16] J. L. Garcés, S. Madurga, and M. Borkovec, "Coupling of conformational and ionization equilibria in linear poly(ethylenimine): A study based on the site binding/rotational isomeric state (SBRIS) model," *Physical Chemistry Chemical Physics*, vol. 16, no. 10, pp. 4626–4638, 2014.
- [17] J. L. Garcés, S. Madurga, C. Rey-Castro, and F. Mas, "Dealing with long-range interactions in the determination of polyelectrolyte ionization properties. Extension of the transfer matrix formalism to the full range of ionic strengths," *Journal of Polymer Science, Part B: Polymer Physics*, vol. 55, no. 3, pp. 275–284, 2017.
- [18] G. S. Manning, "Polyelectrolytes," *Annual Review of Physical Chemistry*, vol. 23, no. 1, pp. 117–140, 1972.
- [19] J. Landsgesell, C. Holm, and J. Smiatek, "Simulation of weak polyelectrolytes: a comparison between the constant pH and the reaction ensemble method," *European Physical Journal: Special Topics*, vol. 226, no. 4, pp. 725–736, 2017.
- [20] M. Borkovec and G. J. Koper, "Affinity Distributions of Polyampholytes with Interacting Acid-Base Groups," *Langmuir*, vol. 10, no. 9, pp. 2863–2865, 1994.
- [21] M. Borkovec and G. J. Koper, "A cluster expansion method for the complete resolution of microscopic ionization equilibria from NMR titrations," *Analytical Chemistry*, vol. 72, no. 14, pp. 3272–3279, 2000.
- [22] M. Borkovec, F. Jonsson, and G. J. M. J. Koper, *Surface and Colloid Science*. Plenum Press, 2001.
- [23] G. Koper and M. Borkovec, "Binding of metal ions to polyelectrolytes and their oligomeric counterparts: An application of a generalized potts model," *Journal of Physical Chemistry B*, vol. 105, no. 28, pp. 6666–6674, 2001.
- [24] M. Borkovec, G. J. M. Koper, and C. Piguet, "Ion binding to polyelectrolytes," *Current Opinion in Colloid and Interface Science*, vol. 11, no. 5, pp. 280–289, 2006.

- [25] M. Borkovec, D. Čakara, and G. J. M. Koper, "Resolution of microscopic protonation enthalpies of polyprotic molecules by means of cluster expansions," *Journal of Physical Chemistry B*, vol. 116, no. 14, pp. 4300–4309, 2012.
- [26] G. J. Koper and M. Borkovec, "Proton binding by linear, branched, and hyperbranched polyelectrolytes," *Polymer*, vol. 51, pp. 5649–5662, nov 2010.
- [27] H. Boroudjerdi, Y. Kim, A. Naji, R. R. Netz, X. Schlagberger, and A. Serr, "Statics and dynamics of strongly charged soft matter," *Physics Reports*, vol. 416, pp. 129–199, 2005.
- [28] M. Lund and B. Jönsson, "Charge regulation in biomolecular solution," *Quarterly Reviews of Biophysics*, vol. 46, no. 03, pp. 265–281, 2013.
- [29] C. F. Narambuena, G. S. Longo, and I. Szleifer, "Lysozyme adsorption in pH-responsive hydrogel thin-films: the non-trivial role of acid-base equilibrium," *Soft Matter*, vol. 11, no. 33, pp. 6669–6679, 2015.
- [30] F. Carnal, A. Clavier, and S. Stoll, "Polypeptide-nanoparticle interactions and corona formation investigated by monte carlo simulations," *Polymers*, vol. 8, no. 6, 2016.
- [31] G. Trefalt, S. H. Behrens, and M. Borkovec, "Charge Regulation in the Electrical Double Layer: Ion Adsorption and Surface Interactions," *Langmuir*, vol. 32, no. 2, pp. 380–400, 2016.
- [32] J. Lipfert, S. Doniach, R. Das, and D. Herschlag, "Understanding Nucleic Acid–Ion Interactions," *Annual Review of Biochemistry*, vol. 83, no. 1, pp. 813–841, 2014.
- [33] S. Uyaver and C. Seidel, "First-order conformational transition of annealed polyelectrolytes in a poor solvent," *Europhysics Letters*, vol. 64, no. 4, pp. 536–542, 2003.
- [34] D. S. Olander and A. Holtzer, "The Stability of the Polyglutamic Acid alpha Helix," *Journal of the American Chemical Society*, vol. 1194, no. 3, pp. 4549–4560, 1959.
- [35] S. T. Whitten, B. Garcia-Moreno E., and V. J. Hilser, "Local conformational fluctuations can modulate the coupling between proton binding and global structural transitions in proteins," *Proceedings of the National Academy of Sciences*, vol. 102, no. 12, pp. 4282–4287, 2005.
- [36] R. K. Das and R. V. Pappu, "Conformations of intrinsically disordered proteins are influenced by linear sequence distributions of oppositely charged residues," *Pnas*, vol. 110, no. 33, pp. 13392–13397, 2013.

- [37] R. K. Das, K. M. Ruff, and R. V. Pappu, "Relating sequence encoded information to form and function of intrinsically disordered proteins.," *Current opinion in structural biology*, vol. 32, pp. 102–112, apr 2015.
- [38] E. D. Cera, "Stochastic linkage : Effect of random fluctuations on a twostate process
Stochastic linkage : Effect of random fluctuations on a two-state process," *J. Phys. Chem.*, vol. 95, no. 7, pp. 5082–5086, 1991.
- [39] S. A. Cannas, "One-dimensional Ising model with long-range interactions: A renormalization-group treatment," *Physical Review B*, vol. 52, no. 5, pp. 1–4, 1995.
- [40] M. Ullner and B. Jonsson, "Monte Carlo study of titrating polyelectrolytes in the presence of salt," *Macromolecules*, vol. 29, no. 20, pp. 6645–6655, 1996.
- [41] G. J. Koper and M. Borkovec, "Exact affinity distributions for linear polyampholytes and polyelectrolytes," *Journal of Chemical Physics*, vol. 104, no. 11, pp. 4204–4213, 1996.
- [42] A. A. Lee, S. Kondrat, and A. A. Kornyshev, "Single-file charge storage in conducting nanopores," *Physical Review Letters*, vol. 113, no. 4, pp. 1–5, 2014.
- [43] V. B. Teif, "General transfer matrix formalism to calculate DNA-protein-drug binding in gene regulation: application to ORoperator of phage λ ," *Nucleic Acids Research*, vol. 35, no. 11, pp. 3519–3867, 2007.
- [44] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications*. Elsevier, 1969.
- [45] C. E. Reed and W. F. Reed, "Monte Carlo study of titration of linear polyelectrolytes," *The Journal of Chemical Physics*, vol. 96, no. 2, p. 1609, 1992.
- [46] M. Stornes, P. Linse, and R. S. Dias, "Monte Carlo Simulations of Complexation between Weak Polyelectrolytes and a Charged Nanoparticle. Influence of Polyelectrolyte Chain Length and Concentration," *Macromolecules*, vol. 50, no. 15, pp. 5978–5988, 2017.
- [47] M. Stornes, B. Shrestha, and R. S. Dias, "PH-Dependent Polyelectrolyte Bridging of Charged Nanoparticles," *Journal of Physical Chemistry B*, vol. 122, no. 44, pp. 10237–10246, 2018.
- [48] K. Hyltegren, T. Nylander, M. Lund, and M. Skepö, "Adsorption of the intrinsically disordered saliva protein histatin 5 to silica surfaces. A Monte Carlo simulation and ellipsometry study," *Journal of Colloid and Interface Science*, vol. 467, pp. 280–290, 2016.

- [49] K. Hyltegren and M. Skepö, "Adsorption of polyelectrolyte-like proteins to silica surfaces and the impact of pH on the response to ionic strength. A Monte Carlo simulation and ellipsometry study," *Journal of Colloid and Interface Science*, vol. 494, pp. 266–273, 2017.
- [50] P. Torres, L. Bojanich, F. Sanchez-Varretti, A. J. Ramirez-Pastor, E. Quiroga, V. Boeris, and C. F. Narambuena, "Protonation of β -lactoglobulin in the presence of strong polyelectrolyte chains: a study using Monte Carlo simulation," *Colloids and Surfaces B: Biointerfaces*, vol. 160, pp. 161–168, 2017.
- [51] P. B. Torres, E. Quiroga, A. J. Ramirez-Pastor, V. Boeris, and C. F. Narambuena, "Interaction between β -Lactoglobuline and Weak Polyelectrolyte Chains : A Study Using Monte Carlo Simulation," *The Journal of Physical Chemistry B*, vol. 123, pp. 8617–8627, 2019.
- [52] S. Madurga, A. Martín-Molina, E. Vilaseca, F. Mas, and M. Quesada-Pérez, "Effect of the surface charge discretization on electric double layers: A Monte Carlo simulation study," *Journal of Chemical Physics*, vol. 126, no. 23, 2007.
- [53] S. Madurga, J. L. Garcés, E. Companys, C. Rey-Castro, J. Salvador, J. Galceran, E. Vilaseca, J. Puy, and F. Mas, "Ion binding to polyelectrolytes: Monte Carlo simulations versus classical mean field theories," *Theoretical Chemistry Accounts*, vol. 123, no. 1-2, pp. 127–135, 2009.
- [54] M. I. Giannotti and G. J. Vancso, "Interrogation of single synthetic polymer chains and polysaccharides by AFM-based force spectroscopy," *ChemPhysChem*, vol. 8, no. 16, pp. 2290–2307, 2007.
- [55] C. Bustamante, J. Macosko, and G. Wuite, "Grabbing the cat by the tail: manipulating molecules one by one.," *Nat. Rev. Mol. Cell. Biol.*, vol. 1, pp. 130–136, 2000.
- [56] P. E. Marszalek, A. F. Oberhauser, Y. P. Pang, and J. M. Fernandez, "Polysaccharide elasticity governed by chair-boat transitions of the glucopyranose ring," *Nature*, vol. 396, no. 6712, pp. 661–664, 1998.
- [57] M. I. Giannotti, M. Rinaudo, and G. J. Vancso, "Force spectroscopy of hyaluronan by atomic force microscopy: From hydrogen-bonded networks toward single-chain behavior," *Biomacromolecules*, vol. 8, no. 9, pp. 2648–2652, 2007.
- [58] J. Camunas-Soler, M. Ribezzi-Crivellari, and F. Ritort, "Elastic Properties of Nucleic Acids by Single-Molecule Force Spectroscopy," *Annual Review of Biophysics*, vol. 45, no. 1, pp. 65–84, 2016.

- [59] A. Bosco, J. Camunas-Soler, and F. Ritort, "Elastic properties and secondary structure formation of single-stranded DNA at monovalent and divalent salt conditions," *Nucleic Acids Research*, vol. 42, no. 3, pp. 2064–2074, 2014.
- [60] A. P. Wiita, R. Perez-Jimenez, K. A. Walther, F. Gräter, B. J. Berne, A. Holmgren, J. M. Sanchez-Ruiz, and J. M. Fernandez, "Probing the chemistry of thioredoxin catalysis with force," *Nature*, vol. 450, no. 7166, pp. 124–127, 2007.
- [61] M. Radiom, P. Kong, P. Maroni, M. Schäfer, A. F. M. Kilbinger, and M. Borkovec, "Mechanically induced cis-to-trans isomerization of carbon-carbon double bonds using atomic force microscopy," *Phys. Chem. Chem. Phys.*, vol. 18, no. 45, pp. 31202–31210, 2016.
- [62] C. Friedsam, H. E. Gaub, and R. R. Netz, "Probing surfaces with single-polymer atomic force microscope experiments," *Biointerphases*, vol. 1, no. 1, pp. MR1–MR21, 2006.
- [63] S. Krysiak, S. Liese, R. R. Netz, and T. Hugel, "Peptide desorption kinetics from single molecule force spectroscopy studies," *Journal of the American Chemical Society*, vol. 136, no. 2, pp. 688–697, 2014.
- [64] P. Pincus, "Excluded Volume Effects and Stretched Polymer Chains," *Macromolecules*, vol. 9, no. 3, pp. 386–388, 1975.
- [65] P. Pincus, "Excluded Volume Effects and Stretched Polymer Chains," *Macromolecules*, vol. 9, no. 3, pp. 386–388, 1976.
- [66] J. F. Marko and E. D. Siggia, "Stretching DNA," *Macromolecules*, vol. 28, no. 26, pp. 8759–8770, 1995.
- [67] P. De Gennes, "Dynamics of entangled polymer solutions. i. the rouse model," *Annual Review of Physical Chemistry*, vol. 9, pp. 587–593, 1976.
- [68] M. Radiom and M. Borkovec, "Influence of ligand-receptor interactions on force-extension behavior within the freely jointed chain model," *Physical Review E*, vol. 96, no. 6, pp. 1–7, 2017.
- [69] T. Odijk, "Stiff chains and filaments under tension," *Macromolecules*, vol. 28, pp. 7016–7018, 1995.
- [70] A. V. Dobrynin, J.-M. Y. Carrillo, and M. Rubinstein, "Chains are more flexible under tension," *Macromolecules*, vol. 43, no. 21, pp. 9181–9190, 2010. PMID: 21415940.
- [71] D. R. Jacobson, D. B. McIntosh, M. J. Stevens, M. Rubinstein, and O. A. Saleh, "Single-stranded nucleic acid elasticity arises from internal electrostatic tension," *PNAS*, vol. 114, no. 20, pp. 5095–5100, 2017.

- [72] R. R. Netz, "Strongly stretched semiflexible extensible polyelectrolytes and DNA," *Macromolecules*, vol. 34, no. 21, pp. 7522–7529, 2001.
- [73] M. N. Dessinges, B. Maier, Y. Zhang, M. Peliti, D. Bensimon, and V. Croquette, "Stretching Single Stranded DNA, a Model Polyelectrolyte," *Physical Review Letters*, vol. 89, no. 24, pp. 2–5, 2002.
- [74] Y. Seol, G. M. Skinner, and K. Visscher, "Elastic properties of a single-stranded charged homopolymeric ribonucleotide," *Physical Review Letters*, vol. 93, no. 11, pp. 8–11, 2004.
- [75] O. A. Saleh, D. B. McIntosh, P. Pincus, and N. Ribeck, "Nonlinear low-force elasticity of single-stranded DNA molecules," *Physical Review Letters*, vol. 102, no. 6, pp. 1–4, 2009.
- [76] D. B. Mcintosh and O. A. Saleh, "Salt Species-Dependent Electrostatic Effects on ssDNA Elasticity," *Macromolecules*, vol. 44, pp. 2328–2333, 2011.
- [77] M. J. Stevens, D. B. McIntosh, and O. A. Saleh, "Simulations of stretching a strong, flexible polyelectrolyte: Using Long chains to access the pincus scaling regime," *Macromolecules*, vol. 46, no. 15, pp. 6369–6373, 2013.
- [78] M. J. Stevens, J. P. Berezney, and O. A. Saleh, "The effect of chain stiffness and salt on the elastic response of a polyelectrolyte," *Journal of Chemical Physics*, vol. 149, no. 16, p. 163328, 2018.
- [79] D. R. Jacobson, D. B. McIntosh, and O. A. Saleh, "The snakelike chain character of unstructured RNA," *Biophysical Journal*, vol. 105, no. 11, pp. 2569–2576, 2013.
- [80] M. Stevens, D. McIntosh, and O. Saleh, "Simulations of stretching a strong, flexible polyelectrolyte," *Macromolecules*, vol. 45, no. 14, pp. 5757–5765, 2012.
- [81] M. Ullner, "Comments on the Scaling Behavior of Flexible Polyelectrolytes within the Debye-Huckel Approximation," *J. Phys. Chem. B*, vol. 107, pp. 8097–8110, 2003.
- [82] H.-X. Zhou, G. Rivas, and A. P. Minton, "Macromolecular Crowding and Confinement: Biochemical, Biophysical, and Potential Physiological Consequences," *Annual Review of Biophysics*, vol. 37, no. 1, pp. 375–397, 2008.
- [83] A. P. Minton, "Excluded volume as a determinant of protein structure and stability," *Biopolymers*, vol. 20, pp. 2093–2120, 1981.
- [84] S. B. Zimmerman and S. O. Trach, "Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of *Escherichia coli*," *Journal of Molecular Biology*, vol. 222, no. 3, pp. 599–620, 1991.

- [85] R. J. Ellis, "Macromolecular crowding: Obvious but underappreciated," *Trends in Biochemical Sciences*, vol. 26, no. 10, pp. 597–604, 2001.
- [86] T. C. Laurent, "Enzyme Reactions in Polymer Media," *European Journal of Biochemistry*, vol. 21, pp. 498–506, 1971.
- [87] C. Balcells, I. Pastor, E. Vilaseca, S. Madurga, M. Cascante, and F. Mas, "Macromolecular crowding effect upon in vitro enzyme kinetics: Mixed activation-diffusion control of the oxidation of NADH by pyruvate catalyzed by lactate dehydrogenase," *Journal of Physical Chemistry B*, vol. 118, no. 15, pp. 4062–4068, 2014.
- [88] I. Yu, T. Mori, T. Ando, R. Harada, J. Jung, Y. Sugita, and M. Feig, "Biomolecular interactions modulate macromolecular structure and dynamics in atomistic model of a bacterial cytoplasm," *eLife*, vol. 5, no. NOVEMBER2016, pp. 1–22, 2016.
- [89] C. Balcells, I. Pastor, L. Pitulice, M. Via, S. Madurga, E. Vilaseca, A. Isvoran, M. Cascante, and F. Mas, "Macromolecular crowding upon in-vivo-like enzyme-kinetics: effect of enzyme-obstacle size ratio," *New Frontiers in Chemistry*, vol. 24, no. 1, pp. 3–16, 2015.
- [90] A. Kuzmak, S. Carmali, E. von Lieres, A. J. Russell, and S. Kondrat, "Can enzyme proximity accelerate cascade reactions?," *Scientific Reports*, vol. 9, no. 1, pp. 1–7, 2019.
- [91] E. A. Cino, M. Karttunen, and W. Y. Choy, "Effects of Molecular Crowding on the Dynamics of Intrinsically Disordered Proteins," *PLoS ONE*, vol. 7, no. 11, 2012.
- [92] C. P. Brangwynne, P. Tompa, and R. V. Pappu, "Polymer physics of intracellular phase transitions," *Nature Physics*, vol. 11, no. 11, pp. 899–904, 2015.
- [93] A. Banks, S. Qin, K. L. Weiss, C. B. Stanley, and H. X. Zhou, "Intrinsically Disordered Protein Exhibits Both Compaction and Expansion under Macromolecular Crowding," *Biophysical Journal*, vol. 114, no. 5, pp. 1067–1079, 2018.
- [94] G. L. Dignon, W. Zheng, Y. C. Kim, R. B. Best, and J. Mittal, "Sequence determinants of protein phase behavior from a coarse-grained model," *PLoS Computational Biology*, vol. 14, no. 1, pp. 1–23, 2018.
- [95] A. V. Fonin, A. L. Darling, I. M. Kuznetsova, K. K. Turoverov, and V. N. Uversky, "Intrinsically disordered proteins in crowded milieu: when chaos prevails within the cellular gumbo," *Cellular and Molecular Life Sciences*, vol. 75, no. 21, pp. 3907–3929, 2018.
- [96] F. C. Zegarra, D. Homouz, A. G. Gasic, L. Babel, M. Kovermann, P. Wittung-Stafshede, and M. S. Cheung, "Crowding-Induced Elongated Conformation of Urea-Unfolded Apoazurin: Investigating the Role of Crowder Shape in Silico," *Journal of Physical Chemistry B*, vol. 123, no. 17, pp. 3607–3617, 2019.

- [97] D. S. Banks and C. Fradin, "Anomalous diffusion of proteins due to molecular crowding," *Biophysical Journal*, vol. 89, no. 5, pp. 2960–2971, 2005.
- [98] I. Pastor, E. Vilaseca, S. Madurga, M. Cascante, and F. Mas, "Diffusion of alpha-Chymiotrypsin in Solution-Crowded Media," *Journal of Physical Chemistry B*, vol. 114, pp. 4028–4034, 2010.
- [99] I. Pastor, E. Vilaseca, S. Madurga, M. Cascante, and F. Mas, "Erratum to Diffusion of alpha-Chymiotrypsin in Solution-Crowded Media," *Journal of Physical Chemistry B*, vol. 114B, p. 12182, 2010.
- [100] E. Vilaseca, A. Isvoran, S. Madurga, I. Pastor, J. L. Garcés, and F. Mas, "New insights into diffusion in 3D crowded media by Monte Carlo simulations: Effect of size, mobility and spatial distribution of obstacles," *Physical Chemistry Chemical Physics*, vol. 13, no. 16, pp. 7396–7407, 2011.
- [101] E. Vilaseca, I. Pastor, A. Isvoran, S. Madurga, J. L. Garcés, and F. Mas, "Diffusion in macromolecular crowded media: Monte Carlo simulation of obstructed diffusion vs. FRAP experiments," *Theoretical Chemistry Accounts*, vol. 128, no. 4, pp. 795–805, 2011.
- [102] P. Mereghetti and R. C. Wade, "Atomic detail brownian dynamics simulations of concentrated protein solutions with a mean field treatment of hydrodynamic interactions," *Journal of Physical Chemistry B*, vol. 116, no. 29, pp. 8523–8533, 2012.
- [103] J. Balbo, P. Mereghetti, D. P. Herten, and R. C. Wade, "The shape of protein crowders is a major determinant of protein diffusion," *Biophysical Journal*, vol. 104, no. 7, pp. 1576–1584, 2013.
- [104] I. Pastor, L. Pitulice, C. Balcells, E. Vilaseca, S. Madurga, A. Isvoran, M. Cascante, and F. Mas, "Effect of crowding by Dextrans in enzymatic reactions," *Biophysical Chemistry*, vol. 185, pp. 8–13, 2014.
- [105] S. Kondrat, O. Zimmermann, W. Wiechert, and E. V. Lieres, "The effect of composition on diffusion of macromolecules in a crowded environment," *Physical Biology*, vol. 12, p. 046003, 2015.
- [106] M. J. Saxton, "Anomalous diffusion due to obstacles: a Monte Carlo study," *Biophysical Journal*, vol. 66, no. 2, pp. 394–401, 1994.
- [107] J. K. G. Dhont, *An introduction to dynamics of colloids*. Elsevier, 1996.
- [108] M. Tokuyama and I. Oppenheim, "Dynamics of hard-sphere suspensions," *Physical Review E*, vol. 50, no. 1, pp. 10–14, 1994.

- [109] M. Tokuyama, "Mean-field theory of glass transitions," *Physica A: Statistical Mechanics and its Applications*, vol. 364, pp. 23–62, 2006.
- [110] M. Tokuyama, T. Moriki, and Y. Kimura, "Self-diffusion of biomolecules in solution," *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, vol. 83, no. 5, pp. 1–8, 2011.
- [111] T. Ando, E. Chow, and J. Skolnick, "Dynamic simulation of concentrated macromolecular solutions with screened long-range hydrodynamic interactions: Algorithm and limitations," *Journal of Chemical Physics*, vol. 139, no. 12, 2013.
- [112] J. P. Bouchaud and A. Georges, "Anomalous diffusion in disordered media: Statistical mechanisms, models and physical applications," *Physics Reports*, vol. 195, no. 4-5, pp. 127–293, 1990.
- [113] N. A. Baker, "Improving implicit solvent simulations : a Poisson-centric view," *Current opinion in structural biology*, vol. 15, pp. 137–143, 2005.
- [114] S. Bucciarelli, J. S. Myung, B. Farago, S. Das, G. A. Vliegthart, O. Holderer, R. G. Winkler, P. Schurtenberger, G. Gompper, and A. Stradner, "Dramatic influence of patchy attractions on short-time protein diffusion under crowded conditions.," *Science advances*, vol. 2, no. 12, p. e1601432, 2016.
- [115] P. hung Wang, I. Yu, M. Feig, and Y. Sugita, "Influence of protein crowder size on hydration structure and dynamics in macromolecular crowding," *Chemical Physics Letters*, vol. 671, pp. 63–70, 2017.
- [116] T. Sentjabrskaja, E. Zaccarelli, C. De Michele, F. Sciortino, P. Tartaglia, T. Voigtmann, S. U. Egelhaaf, and M. Laurati, "Anomalous dynamics of intruders in a crowded environment of mobile obstacles," *Nature Communications*, vol. 7, p. 11133, 2016.
- [117] S. Smith and R. Grima, "Fast simulation of Brownian dynamics in a crowded environment," *Journal of Chemical Physics*, vol. 146, no. 2, p. 024105, 2017.
- [118] D. T. Gillespie, "The multivariate Langevin and Fokker–Planck equations," *American Journal of Physics*, vol. 64, no. 10, pp. 1246–1257, 1996.
- [119] D. L. Ermak and J. A. McCammon, "Brownian dynamics with hydrodynamic interactions," *The Journal of Chemical Physics*, vol. 69, no. 4, pp. 1352–1360, 1978.
- [120] J. Sun and H. Weinstein, "Toward realistic modeling of dynamic processes in cell signaling: Quantification of macromolecular crowding effects," *Journal of Chemical Physics*, vol. 127, no. 15, pp. 1–10, 2007.

- [121] R. J. Phillips, J. F. Brady, and G. Bossis, "Hydrodynamic transport properties of hard-sphere dispersions. II. Porous media," *Physics of Fluids*, vol. 31, no. 12, p. 3473, 1988.
- [122] R. R. Schmidt, J. G. Cifre, and J. G. De La Torre, "Comparison of Brownian dynamics algorithms with hydrodynamic interaction," *Journal of Chemical Physics*, vol. 135, no. 8, 2011.
- [123] J. Rotne and S. Prager, "Variational Treatment of Hydrodynamic Interaction in Polymers," *Journal of Chemical Physics*, vol. 50, no. 11, pp. 4831–4837, 1969.
- [124] H. Yamakawa, "Transport Properties of Polymer Chains in Dilute Solution: Hydrodynamic Interaction," *The Journal of Chemical Physics*, vol. 53, no. 1, pp. 436–443, 1970.
- [125] J. F. Brady, "Dynamic simulation of bounded suspensions of hydrodynamically interacting particles," *Journal of Fluid Mechanics*, vol. 200, no. -1, pp. 39–67, 1989.
- [126] J. Brady, "Stokesian Dynamics," *Annual Review of Fluid Mechanics*, vol. 20, no. 1, pp. 111–157, 1988.
- [127] P. J. Zuk, E. Wajnryb, K. A. Mizerski, and P. Szymczak, "Rotne-Prager-Yamakawa approximation for different-sized particles in application to macromolecular bead models," *Journal of Fluid Mechanics*, vol. 741, pp. 1–13, 2014.
- [128] T. Hugel, M. Rief, M. Seitz, H. E. Gaub, and R. R. Netz, "Highly stretched single polymers: Atomic-force-microscope experiments versus ab-initio theory," *Physical Review Letters*, vol. 94, no. 4, pp. 1–4, 2005.
- [129] C. F. Narambuena, D. M. Beltramo, E. P. M. Leiva, C. F. Narambuena, D. M. Beltramo, and E. P. M. Leiva, "Polyelectrolyte Adsorption on a Charged Surface. Free Energy Calculation from Monte Carlo Simulations Using Jarzynski Equality Polyelectrolyte Adsorption on a Charged Surface. Free Energy Calculation from Monte Carlo Simulations Using Jarzynski Equality," *Macromolecules*, pp. 8267–8274, 2008.
- [130] S. Madurga, C. Rey-Castro, I. Pastor, E. Vilaseca, C. David, J. L. Garcés, J. Puy, and F. Mas, "A semi-grand canonical Monte Carlo simulation model for ion binding to ionizable surfaces: Proton binding of carboxylated latex particles as a case study," *Journal of Chemical Physics*, vol. 135, no. 18, pp. 1–10, 2011.
- [131] M. Radiom, P. Maroni, and W. Tomasz A., "Size extensivity of elastic properties of alkane fragments," *Journal of Molecular Modeling*, vol. 24, no. 36, pp. 1–8, 2018.
- [132] M. Karplus and J. Kuriyan, "Molecular dynamics and protein function," *PNAS*, vol. 102, no. 19, pp. 6679–6685, 2005.

- [133] J. Landsgesell, C. Holm, and J. Smiatek, "Wang-Landau Reaction Ensemble Method : Simulation of Weak Polyelectrolytes and General Acid-Base Reactions," *Journal of Chemical Theory and Computation*, vol. 13, pp. 852–862, 2016.
- [134] P. M. Blanco, S. Madurga, F. Mas, and J. L. Garcés, "Coupling of charge regulation and conformational equilibria in linear weak polyelectrolytes: Treatment of long-range interactions via effective short-ranged and pH-dependent interaction parameters," *Polymers*, vol. 10, no. 8, p. 811, 2018.
- [135] P. M. Blanco, S. Madurga, F. Mas, and J. L. Garcés, "Effect of Charge Regulation and Conformational Equilibria in the Stretching Properties of Weak Polyelectrolytes," *Macromolecules*, vol. 52, pp. 8017–8031, 2019.
- [136] P. M. Blanco, S. Madurga, C. F. Narambuena, F. Mas, and J. L. Garcés, "Role of Charge Regulation and Fluctuations in the Conformational and Mechanical Properties of Weak Flexible Polyelectrolytes," *Polymers*, vol. 11, p. 1962, 2019.
- [137] P. M. Blanco, M. Via, J. L. Garcés, S. Madurga, and F. Mas, "Brownian dynamics computational model of protein diffusion in crowded media with dextran macromolecules as obstacles," *Entropy*, vol. 19, no. 3, p. 105, 2017.
- [138] P. M. Blanco, J. L. Garcés, S. Madurga, and F. Mas, "Macromolecular diffusion in crowded media beyond the hard-sphere model," *Soft Matter*, vol. 14, no. 16, pp. 3105–3114, 2018.
- [139] R. Jurij and L. Per, "MOLSIM: A modular molecular simulation software," *Journal of Computational Chemistry*, vol. 36, no. 16, pp. 1259–1274, 2015.
- [140] J. Wyman and S. J. Gill, *Binding and linkage: functional chemistry of biological macromolecules*. University Science Books, 2009.
- [141] A. Warshel, P. K. Sharma, M. Kato, and W. W. Parson, "Modeling electrostatic effects in proteins," *Biochimica et Biophysica Acta - Proteins and Proteomics*, vol. 1764, no. 11, pp. 1647–1676, 2006.
- [142] J. Hamacek, M. Borkovec, and C. Piguet, "Simple thermodynamics for unravelling sophisticated self-assembly processes," *Dalton Transactions*, vol. 12, p. 1473, 2006.
- [143] Y. Li, T. Zhao, C. Wang, Z. Lin, G. Huang, B. D. Sumer, and J. Gao, "Molecular basis of cooperativity in pH-triggered supramolecular self-assembly," *Nature Communications*, vol. 7, no. May, pp. 1–9, 2016.
- [144] W. Stumm and J. J. Morgan, *Aquatic chemistry: Chemical equilibria and rates in natural waters*. Wiley, 1996.

- [145] B. L. Rivas, E. D. Pereira, and I. Moreno-Villoslada, "Water-soluble polymer – metal ion interactions," *Progress in Polymer Science*, vol. 28, pp. 173–208, 2003.
- [146] A. V. Dobrynin, "Theory and simulations of charged polymers: From solution properties to polymeric nanomaterials," *Current Opinion in Colloid and Interface Science*, vol. 13, no. 6, pp. 376–388, 2008.
- [147] J. P. Chapel and J. F. Berret, "Versatile electrostatic assembly of nanoparticles and polyelectrolytes: Coating, clustering and layer-by-layer processes," *Current Opinion in Colloid and Interface Science*, vol. 17, no. 2, pp. 97–105, 2012.
- [148] S. M. Hartig, R. R. Greene, M. M. Dikov, A. Prokop, and J. M. Davidson, "Multifunctional nanoparticulate polyelectrolyte complexes," *Pharmaceutical Research*, vol. 24, no. 12, pp. 2353–2369, 2007.
- [149] T. Nishio, "Monte Carlo studies on potentiometric titration of poly(glutamic acid)," *Biophysical Chemistry*, vol. 71, no. 2-3, pp. 173–184, 1998.
- [150] P. Zarzycki, "Monte Carlo study of the topographic effects on the proton binding at the energetically heterogeneous metal oxide/electrolyte interface," *Langmuir*, vol. 22, no. 26, pp. 11234–11240, 2006.
- [151] P. Zarzycki, "Effective adsorption energy distribution function as a new mean-field characteristic of surface heterogeneity in adsorption systems with lateral interactions," *Journal of Colloid and Interface Science*, vol. 311, no. 2, pp. 622–627, 2007.
- [152] S. Ulrich, M. Seijo, and S. Stoll, "A Monte Carlo study of weak polyampholytes: Stiffness and primary structure influences on titration curves and chain conformations," *Journal of Physical Chemistry B*, vol. 111, no. 29, pp. 8459–8467, 2007.
- [153] E. Companys, J. L. Garcés, J. Salvador, J. Galceran, J. Puy, and F. Mas, "Electrostatic and specific binding to macromolecular ligands. A general analytical expression for the Donnan volume," *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, vol. 306, no. 1-3 SPEC. ISS., pp. 2–13, 2007.
- [154] F. Carnal, S. Ulrich, and S. Stoll, "Influence of explicit ions on titration curves and conformations of flexible polyelectrolytes: A Monte Carlo study," *Macromolecules*, vol. 43, no. 5, pp. 2544–2553, 2010.
- [155] F. Carnal and S. Stoll, "Chain stiffness, salt valency, and concentration influences on titration curves of polyelectrolytes: Monte Carlo simulations," *Journal of Chemical Physics*, vol. 134, no. 4, pp. 044904:1–11, 2011.

- [156] A. Aleksandrov, S. Polydorides, G. Archontis, and T. Simonson, "Predicting the Acid / Base Behavior of Proteins : A Constant-pH Monte Carlo Approach with Generalized Born Solvent," *J. Phys. Chem. B*, vol. 114, pp. 10634–10648, 2010.
- [157] D. Poland and H. A. Scheraga, *Theory of helix-coil transitions in biopolymers: statistical mechanical theory of order-disorder transitions in biological macromolecules*. Academic Press, 1978.
- [158] J. Vila, "Statistical mechanics in homopolypeptides. I. Theoretical approach for titration process in α helix," *The Journal of Chemical Physics*, vol. 84, no. 11, pp. 6421–6425, 1986.
- [159] T. Alfrey and H. Morawetz, "Amphoteric Polyelectrolytes. I. 2-Vinylpyridine-Methacrylic Acid Copolymers," *Journal of the American Chemical Society*, vol. 74, no. 2, pp. 436–438, 1952.
- [160] J. Mazur, A. Silberberg, and A. Katchalsky, "Potentiometric behavior of polyampholytes," *J. Polym. Sci.*, vol. 35, no. 128, pp. 43–70, 1959.
- [161] C. David, E. Companys, J. Galceran, J. L. Garcés, F. Mas, C. Rey-Castro, J. Salvador, and J. Puy, "Competitive ion complexation to polyelectrolytes: Determination of the stepwise stability constants. the $\text{Ca}^{2+}/\text{H}^{+}$ /polyacrylate system," *Journal of Physical Chemistry B*, vol. 111, no. 35, pp. 10421–10430, 2007.
- [162] C. David, E. Companys, J. Galceran, J. L. Garcés, F. Mas, C. Rey-Castro, J. Salvador, and J. Puy, "Competitive $\text{Cd}^{2+}/\text{H}^{+}$ complexation to polyacrylic acid described by the stepwise and intrinsic stability constants," *Journal of Physical Chemistry B*, vol. 112, no. 33, pp. 10092–10100, 2008.
- [163] Y. Burak and R. R. Netz, "Charge Regulation of Interacting Weak Polyelectrolytes," *Journal of Physical Chemistry B*, vol. 108, pp. 4840–4849, 2004.
- [164] A. M. Baptista, V. H. Teixeira, C. M. Soares, M. Baptista, and V. H. Teixeira, "Constant-pH molecular dynamics using stochastic titration Constant-p H molecular dynamics using stochastic titration," *J. Chem. Phys.*, vol. 117, pp. 4184–4199, 2002.
- [165] J. Mongan, D. A. Case, and J. A. McCammon, "Constant pH molecular dynamics in generalized Born implicit solvent," *Journal of Computational Chemistry*, vol. 25, no. 16, pp. 2038–2048, 2004.
- [166] A. Valiaev, W. L. Dong, T. G. Oas, A. Chilkoti, and S. Zauscher, "Force-induced prolyl cis-trans isomerization in elastin-like polypeptides," *Journal of the American Chemical Society*, vol. 129, no. 20, pp. 6491–6497, 2007.

- [167] L. Rognoni, T. Most, G. Oldak, and M. Rief, "Force-dependent isomerization kinetics of a highly conserved proline switch modulates the mechanosensing region of filamin," *Proceedings of the National Academy of Sciences*, vol. 111, no. 15, pp. 5568–5573, 2014.
- [168] H. M. Klukovich, T. B. Kouznetsova, Z. S. Kean, J. M. Lenhardt, and S. L. Craig, "A backbone lever-arm effect enhances polymer mechanochemistry," *Nature Chemistry*, vol. 5, no. 2, pp. 110–114, 2013.
- [169] J. Wang, T. B. Kouznetsova, Z. Niu, M. T. Ong, H. M. Klukovich, A. L. Rheingold, T. J. Martinez, and S. L. Craig, "Inducing and quantifying forbidden reactivity with single-molecule polymer mechanochemistry," *Nature Chemistry*, vol. 7, no. 4, pp. 323–327, 2015.
- [170] Y. Liu, K. Liu, Z. Wang, and X. Zhang, "Host-enhanced π - π Interaction for water-soluble supramolecular polymerization," *Chemistry - A European Journal*, vol. 17, no. 36, pp. 9930–9935, 2011.
- [171] D. Schütze, K. Holz, J. Müller, M. K. Beyer, U. Lüning, and B. Hartke, "Pinpointing mechanochemical bond rupture by embedding the mechanophore into a macrocycle," *Angewandte Chemie - International Edition*, vol. 54, no. 8, pp. 2556–2559, 2015.
- [172] T. R. Strick, M. N. Dessinges, G. Charvin, N. H. Dekker, J. F. Allemand, D. Bensimon, and V. Croquette, "Stretching of macromolecules and proteins," *Reports on Progress in Physics*, vol. 66, no. 1, pp. 1–45, 2003.
- [173] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub, "Immunoglobulin Domains by AFM Reversible Unfolding of Individual Titin Immunoglobulin Domains by AFM," *Science*, vol. 276, pp. 1109–1112, 1997.
- [174] C. Bustamante, J. F. Marko, E. D. Siggia, and S. Smith, "Entropic Elasticity of X-Phage DNA Explicit and Implicit Learning and Maps of Cortical Motor Output," *Science*, vol. 265, pp. 1599–1600, 1994.
- [175] J. Kierfeld, O. Niamploy, V. Sa-Yakanit, and R. Lipowsky, "Stretching of semiflexible polymers with elastic bonds," *European Physical Journal E*, vol. 14, no. 1, pp. 17–34, 2004.
- [176] H. J. Kreuzer and M. Grunze, "Stretching of single polymer strands: A first-principles theory," *Europhysics Letters (EPL)*, vol. 55, no. 5, pp. 640–646, 2001.
- [177] F. Hanke, A. Serr, H. J. Kreuzer, and R. R. Netz, "Stretching single polypeptides: The effect of rotational constraints in the backbone," *EPL (Europhysics Letters)*, vol. 92, no. 5, p. 53001, 2010.
- [178] G. Neuert, T. Hugel, R. R. Netz, and H. E. Gaub, "Elasticity of poly(azobenzene-peptides)," *Macromolecules*, vol. 39, no. 2, pp. 789–797, 2006.

- [179] F. Oesterhelt, M. Rief, and H. E. Gaub, "Single molecule force spectroscopy by AFM indicates helical structure of poly (ethylene-glycol) in water," *New Journal of Physics*, vol. 1, pp. 1–11, 1999.
- [180] S. Liese, M. Gensler, S. Krysiak, R. Schwarzl, A. Achazi, B. Paulus, T. Hugel, J. P. Rabe, and R. R. Netz, "Hydration Effects Turn a Highly Stretched Polymer from an Entropic into an Energetic Spring," *ACS Nano*, vol. 11, no. 1, pp. 702–712, 2017.
- [181] P. M. Blanco, S. Madurga, F. Mas, and J. L. Garcés, "Coupling of charge regulation and conformational equilibria in linear weak polyelectrolytes: Treatment of long-range interactions via effective short-ranged and pH-dependent interaction parameters," *Polymers*, vol. 10, no. 8, p. 811, 2018.
- [182] D. S. Olander and A. Holtzer, "The Stability of the Polyglutamic Acid alpha Helix," *Journal of the American Chemical Society*, vol. 90, pp. 4549–4560, 1968.
- [183] J. P. Berezney and O. A. Saleh, "Electrostatic Effects on the Conformation and Elasticity of Hyaluronic Acid, a Moderately Flexible Polyelectrolyte," *Macromolecules*, vol. 50, no. 3, pp. 1085–1089, 2017.
- [184] C. F. Narambuena, "On the reasons for α -lactalbumin adsorption on a charged surface: a study by monte carlo simulation," *Colloids and Surfaces B: Biointerfaces*, vol. 174, pp. 511 – 520, 2019.
- [185] S. Boudon and G. Wipff, "Conformational analysis of protonated ethylenediamine in the gas phase and in water," *Journal of Molecular Structure: THEOCHEM*, vol. 228, no. C, pp. 61–70, 1991.
- [186] Y. R. Sliozberg, M. Kröger, and T. L. Chantawansri, "Fast equilibration protocol for million atom systems of highly entangled linear polyethylene chains," *Journal of Chemical Physics*, vol. 144, no. 15, p. 154901, 2016.
- [187] J. L. Garcés, F. Mas, J. Puy, J. Galceran, and J. Salvador, "Use of activity coefficients for bound and free sites to describe metal–macromolecule complexation," *Journal of the Chemical Society-Faraday Transactions*, vol. 94, pp. 2783–2794, 1998.
- [188] J. L. Garcés, F. Mas, J. Cecília, E. Companys, J. Galceran, J. Salvador, J. Puy, and J. Galceran, "Complexation isotherms in metal speciation studies at trace concentration levels. Voltammetric techniques in environmental samples," *Physical Chemistry Chemical Physics*, vol. 4, pp. 3764–3773, 2002.
- [189] F. Hanakam, G. Günther, S. Lotz, T. Alt, and A. Seelig, "Binding of Hisactophilin I and II to Lipid Membranes Is Controlled by a," *Biochemistry*, vol. 35, pp. 11036–11044, 1996.

- [190] M. Lund and B. Jönsson, "On the Charge Regulation of Proteins," *Biochemistry*, vol. 44, pp. 5722–5727, 2005.
- [191] A. C. Mason and J. H. Jensen, "Protein-protein binding is often associated with changes in protonation state," *Proteins: Structure, Function, and Bioinformatics*, vol. 71, pp. 81–91, 2008.
- [192] B. Svensson, B. Jonsson, and E. Thulin, "Binding of Ca²⁺ to Calmodulin and Its Tryptic Fragments : Theory and Experiment," *Biochemistry*, vol. 32, pp. 2828–2834, 1993.
- [193] B. Aguilar, R. Anandakrishnan, J. Z. Ruscio, and A. V. Onufriev, "Statistics and physical origins of pK and ionization state changes upon protein-ligand binding," *Biophysical Journal*, vol. 98, no. 5, pp. 872–880, 2010.
- [194] M. I. Giannotti and G. J. Vancso, "Interrogation of single synthetic polymer chains and polysaccharides by AFM-based force spectroscopy," *ChemPhysChem*, vol. 8, no. 16, pp. 2290–2307, 2007.
- [195] P. M. Blanco, S. Madurga, F. Mas, and J. L. Garcés, "Effect of Charge Regulation and Conformational Equilibria in the Stretching Properties of Weak Polyelectrolytes," *Macromolecules*, vol. 52, pp. 8017–8031, 2019.
- [196] J. G. Kirkwood and J. B. Shumaker, "Forces between Protein Molecules in Solution Arising from Fluctuations in Proton Charge and Configuration," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 38, no. 10, pp. 863–871, 1952.
- [197] T. Odijk, "Polyelectrolytes Near the Rod Limit," *J Polym Sci Polym Phys Ed*, vol. 15, no. 3, pp. 477–483, 1977.
- [198] J. L. Barrat and J. F. Joanny, "Persistence length of polyelectrolyte chains," *Epl*, vol. 24, no. 5, pp. 333–338, 1993.
- [199] A. H. Elcock, "Models of macromolecular crowding effects and the need for quantitative comparisons with experiment," *Current Opinion in Structural Biology*, vol. 20, no. 2, pp. 196–206, 2010.
- [200] S. Hasnain, C. L. McClendon, M. T. Hsu, M. P. Jacobson, and P. Bandyopadhyay, "A new coarse-grained model for E. coli cytoplasm: Accurate calculation of the diffusion coefficient of proteins and observation of anomalous diffusion," *PLoS ONE*, vol. 9, no. 9, 2014.
- [201] D. Gomez and S. Klumpp, "Biochemical reactions in crowded environments: revisiting the effects of volume exclusion with simulations," *Frontiers in Physics*, vol. 3, no. June, pp. 1–14, 2015.

- [202] A. H. Elcock, "Molecule-centered method for accelerating the calculation of hydrodynamic interactions in Brownian dynamics simulations containing many flexible biomolecules," *Journal of Chemical Theory and Computation*, vol. 9, no. 7, pp. 3224–3239, 2013.
- [203] M. Fixman, "Implicit Algorithm for Brownian Dynamics of Polymers," *Macromolecules*, vol. 19, no. 4, pp. 1195–1204, 1986.
- [204] T. Geyer and U. Winter, "An $O(N^2)$ approximation for hydrodynamic interactions in Brownian dynamics simulations," *Journal of Chemical Physics*, vol. 130, no. 11, pp. 1–9, 2009.
- [205] H. Mori, "Transport, Collective Motion, and Brownian Motion," 1965.
- [206] J. Schöneberg and F. Noé, "ReaDDy—a software for particle-based reaction-diffusion dynamics in crowded cellular environments.," *PLoS one*, vol. 8, no. 9, p. e74261, 2013.
- [207] F. Roosen-Runge, M. Hennig, F. Zhang, R. M. Jacobs, M. Sztucki, H. Schober, T. Seydel, and F. Schreiber, "Protein self-diffusion in crowded solutions," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 29, pp. 11815–11820, 2011.
- [208] S. Qin and H. X. Zhou, "Atomistic modeling of macromolecular crowding predicts modest increases in protein folding and binding stability," *Biophysical Journal*, vol. 97, no. 1, pp. 12–19, 2009.
- [209] S. Sastry, T. M. Truskett, P. G. Debenedetti, S. Torquato, and F. H. Stillinger, "Free volume in the hard sphere liquid," *Molecular Physics*, vol. 95, no. 2, pp. 289–297, 1998.
- [210] J. F. Brady, R. J. Phillips, and G. Bossis, "Dynamic simulation of hydrodynamically interacting suspensions," *Journal of Fluid Mechanics*, vol. 195, pp. 257–280, 1988.
- [211] C. W. J. Beenakker, "Ewald sum of the Rotne–Prager tensor," *The Journal of Chemical Physics*, vol. 85, no. 3, pp. 1581–1582, 1986.
- [212] A. J. Ladd, "Hydrodynamic transport coefficients of random dispersions of hard spheres," *The Journal of Chemical Physics*, vol. 93, no. 5, pp. 3484–3494, 1990.
- [213] J. K. Armstrong, R. B. Wenby, H. J. Meiselman, and T. C. Fisher, "The hydrodynamic radii of macromolecules and their effect on red blood cell aggregation," *Biophysical Journal*, vol. 87, no. 6, pp. 4259–4270, 2004.
- [214] G. Fundueanu, C. Nastruzzi, A. Carpov, J. Desbrieres, and M. Rinaudo, "Physicochemical characterization of Ca-alginate microparticles produced with different methods," *Biomaterials*, vol. 20, no. 15, pp. 1427–1435, 1999.

- [215] J. W. Sloan, B. H. Alexander, R. L. Lohmar, I. A. Wolff, and C. E. Rist, "Determination of Dextran Structure by Periodate Oxidation Techniques," *Journal of the American Chemical Society*, vol. 76, no. 17, pp. 4429–4434, 1954.
- [216] J. Sabatie, L. Choplin, M. Moan, J. L. Doublier, F. Paul, and P. Monsan, "The effect of synthesis temperature on the structure of dextran NRRL B 512F," *Carbohydrate Polymers*, vol. 9, no. 2, pp. 87–101, 1988.
- [217] D. W. Schaefer, "A unified model for the structure of polymers in semidilute solution," *Polymer (United Kingdom)*, vol. 25, no. 3, pp. 387–394, 1984.
- [218] P. Vilaseca and G. Franzese, "Softness dependence of the anomalies for the continuous shouldered well potential," *Journal of Chemical Physics*, vol. 133, no. 8, pp. 1–11, 2010.
- [219] A. Minton, "Confinement as a determinant of macromolecular structure and reactivity," *Biophysical Journal*, vol. 63, no. 4, pp. 1090–1100, 1992.
- [220] G. Rivas and A. P. Minton, "Macromolecular Crowding In Vitro, In Vivo, and In Between," *Trends in Biochemical Sciences*, vol. 41, no. 11, pp. 970–981, 2016.
- [221] M. C. Konopka, I. A. Shkel, S. Cayley, M. T. Record, and J. C. Weisshaar, "Crowding and confinement effects on protein diffusion in vivo," *Journal of Bacteriology*, vol. 188, no. 17, pp. 6115–6123, 2006.
- [222] J. A. Dix and A. Verkman, "Crowding Effects on Diffusion in Solutions and Cells," *Annual Review of Biophysics*, vol. 37, no. 1, pp. 247–263, 2008.
- [223] M. Feig, I. Yu, P.-h. Wang, G. Nawrocki, and Y. Sugita, "Crowding in Cellular Environments at an Atomistic Level from Computer Simulations," *The Journal of Physical Chemistry B*, pp. 1–17, 2017.
- [224] A. Irbäck and S. Mohanty, "Protein folding/unfolding in the presence of interacting macromolecular crowders," *European Physical Journal: Special Topics*, vol. 226, no. 4, pp. 627–638, 2017.
- [225] D. Bedrov, C. Ayyagari, and G. D. Smith, "Multiscale Modeling of Poly (ethylene oxide) - Poly (propylene oxide) -Poly (ethylene oxide) Triblock Copolymer Micelles in Aqueous Solution," *J.Chem.Theory Comput.*, vol. 2, pp. 598–606, 2006.
- [226] D. Bedrov, G. D. Smith, and J. Yoon, "Structure and Interactions in Micellar Solution: Molecular Simulations of Pluronic L64 Aqueous Solutions," *Langmuir*, vol. 23, no. 24, pp. 12032–12041, 2007.

- [227] C. N. Likos, H. Löwen, A. Poppe, L. Willner, J. Roovers, B. Cubitt, and D. Richter, "Ordering phenomena of star polymer solutions approaching the Θ state," *Physical Review E - Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics*, vol. 58, no. 5, pp. 6299–6307, 1998.
- [228] C. N. Likos, H. Löwen, M. Watzlawek, B. Abbas, O. Jucknischke, J. Allgaier, and D. Richter, "Star Polymers Viewed as Ultrasoft Colloidal Particles," *Physical Review Letters*, vol. 80, no. 20, pp. 4450–4453, 1998.
- [229] B. Loppinet, G. Fytas, D. Vlassopoulos, C. N. Likos, G. Meier, and J. L. Guo, "Dynamics of dense suspensions of star-like micelles with responsive fixed cores," *Macromolecular Chemistry and Physics*, vol. 206, no. 1, pp. 163–172, 2005.
- [230] L. Liu, W. K. den Otter, and W. J. Briels, "Coarse grain forces in star polymer melts," *Soft Matter*, vol. 10, no. 39, pp. 7874–7886, 2014.
- [231] E. Locatelli, B. Capone, and C. N. Likos, "Multiblob coarse-graining for mixtures of long polymers and soft colloids," *Journal of Chemical Physics*, vol. 145, no. 17, p. 174901, 2016.
- [232] S. T. Milner, T. A. Witten, and M. E. Cates, "Theory of the Grafted Polymer Brush," *Macromolecules*, vol. 21, no. 8, pp. 2610–2619, 1988.
- [233] J. Mewis, W. J. Frith, T. A. Strivens, and W. B. Russel, "The rheology of suspensions containing polymerically stabilized particles," *AIChE Journal*, vol. 35, no. 3, pp. 415–422, 1989.
- [234] U. Genz, B. D'Aguzzo, J. Mewis, and R. Klein, "Structure of Sterically Stabilized Colloids," *Langmuir*, vol. 10, no. 7, pp. 2206–2212, 1994.
- [235] F. Leoni and G. Franzese, "Structural behavior and dynamics of an anomalous fluid between attractive and repulsive walls: Templating, molding, and superdiffusion," *Journal of Chemical Physics*, vol. 141, no. 17, p. 174501, 2014.
- [236] A. Ortega, D. Amorós, and J. García De La Torre, "Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models," *Biophysical Journal*, vol. 101, no. 4, pp. 892–898, 2011.
- [237] M. Fairhead, D. Krndija, E. D. Lowe, and M. Howarth, "Plug-and-play pairing via defined divalent streptavidins," *Journal of Molecular Biology*, vol. 426, no. 1, pp. 199–214, 2014.
- [238] T. Terai, M. Kohno, G. Boncompain, S. Sugiyama, N. Saito, R. Fujikake, T. Ueno, T. Komatsu, K. Hanaoka, T. Okabe, Y. Urano, F. Perez, and T. Nagano, "Artificial Ligands of Streptavidin (ALiS): Discovery, Characterization, and Application for Reversible Control

- of Intracellular Protein Transport," *Journal of the American Chemical Society*, vol. 137, no. 33, pp. 10464–10467, 2015.
- [239] L. Baugh, I. Le Trong, P. S. Stayton, R. E. Stenkamp, and T. P. Lybrand, "A Streptavidin Binding Site Mutation Yields an Unexpected Result: An Ionized Asp128 Residue Is Not Essential for Strong Biotin Binding," *Biochemistry*, vol. 55, no. 37, pp. 5201–5203, 2016.
- [240] C. E. Pähler, A. and Hendrickson, W. A. and Gawinowicz Kolks, M. A. and Argaña and C. R. Cantor, "Characterization and Crystallization of Core Streptavidin," *The Journal of Biological Chemistry*, vol. 262, pp. 13933–13937, 1987.
- [241] J.-P. Behr, *The lock-and-key principle: the state of the art-100 years on*. J. Wiley, 1994.
- [242] J. J. Ward, J. S. Sodhi, L. J. McGuffin, B. F. Buxton, and D. T. Jones, "Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life," *Journal of Molecular Biology*, vol. 337, no. 3, pp. 635–645, 2004.
- [243] V. N. Uversky, "The mysterious unfoldome: Structureless, underappreciated, yet vital part of any given proteome," *Journal of Biomedicine and Biotechnology*, vol. 2010, pp. 1–14, 2010.
- [244] V. N. Uversky, "Intrinsically disordered proteins may escape unwanted interactions via functional misfolding," *Biochimica et Biophysica Acta - Proteins and Proteomics*, vol. 1814, no. 5, pp. 693–712, 2011.
- [245] V. N. Uversky and A. K. Dunker, "Understanding protein non-folding," *Biochimica et Biophysica Acta - Proteins and Proteomics*, vol. 1804, no. 6, pp. 1231–1264, 2010.
- [246] Y. Shen, J. M. Jacobs, D. G. Camp, R. Fang, R. J. Moore, R. D. Smith, W. Xiao, R. W. Davis, and R. G. Tompkins, "Ultra-High-Efficiency Strong Cation Exchange LC/RPLC/MS/MS for High Dynamic Range Characterization of the Human Plasma Proteome," *Analytical Chemistry*, vol. 76, no. 4, pp. 1134–1144, 2004.
- [247] J. Landsgesell, L. Nová, O. Rud, F. Uhlík, D. Sean, P. Hebbeker, C. Holm, and P. Kořovan, "Simulations of ionization equilibria in weak polyelectrolyte solutions and gels," *Soft Matter*, vol. 15, pp. 1155–1185, 2019.
- [248] A. F. Jorge, R. S. Dias, and A. A. Pais, "Enhanced condensation and facilitated release of DNA using mixed cationic agents: A combined experimental and Monte Carlo study," *Biomacromolecules*, vol. 13, no. 10, pp. 3151–3161, 2012.
- [249] S. K. Ramisetty and R. S. Dias, "Synergistic role of DNA-binding protein and macromolecular crowding on DNA condensation. An experimental and theoretical approach," *Journal of Molecular Liquids*, vol. 210, pp. 64–73, 2015.

- [250] C. Cragnell, E. Rieloff, and M. Skepö, "Utilizing Coarse-Grained Modeling and Monte Carlo Simulations to Evaluate the Conformational Ensemble of Intrinsically Disordered Proteins and Regions," *Journal of Molecular Biology*, vol. 430, no. 16, pp. 2478–2492, 2018.
- [251] E. Fagerberg, S. Lenton, and M. Skepö, "Evaluating Models of Varying Complexity of Crowded Intrinsically Disordered Protein Solutions Against SAXS," *Journal of Chemical Theory and Computation*, vol. 15, pp. 6968–6983, 2019.
- [252] B. J. MacKay, L. Denepitiya, V. J. Iacono, S. B. Krost, and J. J. Pollock, "Growth-inhibitory and bactericidal effects of human parotid salivary histidine-rich polypeptides on *Streptococcus mutans*," *Infection and Immunity*, vol. 44, no. 3, pp. 695–701, 1984.
- [253] J. J. Pollock, L. Denepitiya, B. J. MacKay, and V. J. Iacono, "Fungistatic and fungicidal activity of human parotid salivary histidine-rich polypeptides on *Candida albicans*," *Infection and Immunity*, vol. 44, no. 3, pp. 702–707, 1984.
- [254] S. Puri and M. Edgerton, "How does it kill?: Understanding the candidacidal mechanism of salivary histatin 5," *Eukaryotic Cell*, vol. 13, no. 8, pp. 958–964, 2014.
- [255] K. Wróblewski, R. Muhandiram, A. Chakrabartty, and A. Bennick, "The molecular interaction of human salivary histatins with polyphenolic compounds," *European Journal of Biochemistry*, vol. 268, no. 16, pp. 4384–4397, 2001.
- [256] A. Bennick, "Interaction of plant polyphenols with salivary proteins," *Critical Reviews in Oral Biology and Medicine*, vol. 13, no. 2, pp. 184–196, 2002.
- [257] J. Kang, H. G. Lemaire, A. Unterbeck, J. M. Salbaum, C. L. Masters, K. H. Grzeschik, G. Multhaup, K. Beyreuther, and B. Müller-Hill, "The precursor of Alzheimer's disease amyloid A4 protein resembles a cell-surface receptor," *Nature*, vol. 325, no. 6106, pp. 733–736, 1987.
- [258] P. M. Gorman, C. M. Yip, P. E. Fraser, and A. Chakrabartty, "Alternate aggregation pathways of the Alzheimer β -amyloid peptide: $A\beta$ association kinetics at endosomal pH," *Journal of Molecular Biology*, vol. 325, no. 4, pp. 743–757, 2003.
- [259] G. F. Chen, T. H. Xu, Y. Yan, Y. R. Zhou, Y. Jiang, K. Melcher, and H. E. Xu, "Amyloid beta: Structure, biology and structure-based therapeutic development," *Acta Pharmacologica Sinica*, vol. 38, no. 9, pp. 1205–1235, 2017.
- [260] F. G. Oppenheim, T. Xu, F. M. McMillian, S. M. Levitz, R. D. Diamond, G. D. Offner, and R. F. Troxler, "Histatins, a novel family of histidine-rich proteins in human parotid secretion. Isolation, characterization, primary structure, and fungistatic effects on *Candida albicans*," *Journal of Biological Chemistry*, vol. 263, no. 16, pp. 7472–7477, 1988.

- [261] Y. Nozaki and C. Tanford, "Examination of titration behavior," *Methods in Enzymology*, vol. 11, no. C, pp. 715–734, 1967.
- [262] M. K. Menon and A. L. Zydney, "Measurement of Protein Charge and Ion Binding Using Capillary Electrophoresis," *Analytical Chemistry*, vol. 70, no. 8, pp. 1581–1584, 1998.
- [263] J. Bernhardt and H. Pauly, "Partial specific volumes in highly concentrated protein solutions. I. Water-bovine serum albumin and water-bovine hemoglobin," *Journal of Physical Chemistry*, vol. 79, no. 6, pp. 584–590, 1975.
- [264] J. R. Whitaker, "Determination of Molecular Weights of Proteins by Gel Filtration of Sephadex.," *Analytical Chemistry*, vol. 35, no. 12, pp. 1950–1953, 1963.
- [265] M. L. Huber, R. A. Perkins, A. Laesecke, D. G. Friend, J. V. Sengers, M. J. Assael, I. N. Metaxa, E. Vogel, R. Mareš, and K. Miyagawa, "New international formulation for the viscosity of H₂O," *Journal of Physical and Chemical Reference Data*, vol. 38, no. 2, pp. 101–125, 2009.
- [266] A. K. Gaigalas, J. B. Hubbard, M. McCurley, and S. Woo, "Diffusion of bovine serum albumin in aqueous solutions," *Journal of Physical Chemistry*, vol. 96, no. 5, pp. 2355–2359, 1992.
- [267] T. Raj and W. H. Flygare, "Diffusion studies of bovine serum albumin by quasielastic light scattering," *Biochemistry*, vol. 13, no. 16, pp. 3336–3340, 1974.
- [268] F. Carlsson, M. Malmsten, and P. Linse, "Protein-polyelectrolyte cluster formation and redissolution: A monte carlo study," *Journal of the American Chemical Society*, vol. 125, no. 10, pp. 3140–3149, 2003. PMID: 12617682.
- [269] Y. Xu, M. Mazzawi, K. Chen, L. Sun, and P. L. Dubin, "Protein purification by polyelectrolyte coacervation: Influence of protein charge anisotropy on selectivity," *Biomacromolecules*, vol. 12, no. 5, pp. 1512–1522, 2011. PMID: 21413681.
- [270] R. De Vries, "Monte Carlo simulations of flexible polyanions complexing with whey proteins at their isoelectric point," *J. Chem. Phys.*, vol. 120, pp. 3475–3481, 2014.
- [271] Y. Cemil, J. Heyda, M. Ballauff, and J. Dzubiella, "Like-charged protein-polyelectrolyte complexation driven by charge patches," *J. Chem. Phys.*, vol. 143, p. 064905, 2015.
- [272] C. Yigit, M. Kanduc, M. Ballau, and J. Dzubiella, "Interaction of Charged Patchy Protein Models with Like-Charged Polyelectrolyte Brushes," *Langmuir*, vol. 33, pp. 417–427, 2017.

- [273] M. J. Blandamer and M. C. Symons, "Significance of new values for ionic radii to solvation phenomena in aqueous solution," *Journal of Physical Chemistry*, vol. 67, no. 6, pp. 1304–1306, 1963.
- [274] B. Linse and P. Linse, "Tuning the smooth particle mesh Ewald sum: Application on ionic solutions and dipolar fluids," *Journal of Chemical Physics*, vol. 141, no. 18, 2014.
- [275] M. Sjalander, M. Jahre, G. Tufte, and N. Reissmann, "EPIC: An energy-efficient, high-performance GPGPU computing research infrastructure," 2019.
- [276] E. J. Helmerhorshmt, W. V. Hof, P. Breeuwer, E. C. I. Veerman, T. Abee, R. F. Troxler, A. V. N. Amerongen, and F. G. Oppenheim, "Characterization of Histatin 5 with Respect to Amphipathicity, Hydrophobicity, and Effects on Cell and Mitochondrial Membrane Integrity Excludes a Candidacidal Mechanism of Pore Formation *," *The Journal of Biological Chemistry*, vol. 276, no. 8, pp. 5643–5649, 2001.
- [277] H. Nikawa, H. Fukushima, S. Makihira, T. Hamada, and L. P. Samaranayake, "Fungicidal effect of three new synthetic cationic peptides against *Candida albicans*," *Oral diseases*, vol. 10, no. February, pp. 221–228, 2004.
- [278] A. Kurut, J. Henriques, J. Forsman, M. Skepo, and M. Lund, "Role of histidine for charge regulation of unstructured peptides at interfaces and in bulk," *Proteins*, vol. 82, no. September, pp. 657–667, 2014.
- [279] C. Cragnell, D. Durand, B. Cabane, and M. Skepo, "Coarse-grained modeling of the intrinsically disordered protein Histatin 5 in solution: Monte Carlo simulations in combination with SAXS," *Proteins*, no. February, pp. 1–15, 2016.
- [280] G. M. Klug, D. Losic, S. S. Subasinghe, M. I. Aguilar, L. L. Martin, and D. H. Small, " β -Amyloid protein oligomers induced by metal ions and acid pH are distinct from those generated by slow spontaneous ageing at neutral pH," *European Journal of Biochemistry*, vol. 270, no. 21, pp. 4282–4293, 2003.
- [281] L. Guldbrand, B. Jönsson, H. Wennerström, and P. Linse, "Electrical double layer forces. A Monte Carlo study," *The Journal of Chemical Physics*, vol. 80, no. 5, pp. 2221–2228, 1984.
- [282] H. Kang, P. A. Pincus, C. Hyeon, and D. Thirumalai, "Effects of macromolecular crowding on the collapse of biopolymers," *Physical Review Letters*, vol. 114, no. 6, pp. 1–5, 2015.
- [283] D. Thirumalai, "Isolated polymer molecule in a random environment," *Physical Review A*, vol. 37, no. 1, pp. 269–276, 1988.

-
- [284] A. Laguerre and S. Stoll, "Adsorption of a weakly charged polymer on an oppositely charged colloidal particle: Monte carlo simulations investigation," *Polymer*, vol. 46, no. 4, pp. 1359 – 1372, 2005. Collected papers from PDM 2004. *Polymers in Dispersed Media - Colloids: from Preparation to Application*.
- [285] W. Stroberg and S. Schnell, "On the origin of non-membrane-bound organelles, and their physiological function," *Journal of Theoretical Biology*, vol. 434, pp. 42–49, 2017.

