Numéro d'ordre : 410

Centrale Lille

Thèse

pour obtenir le grade de :

Docteur

dans la spécialité

« Micro et nano technologies, acoustique et télécommunications »

par

Manon Benedito

Doctorat délivré par Centrale Lille

Mécanique statistique et thermodynamique des systèmes avec transitions conformationnelles : applications aux macromolécules biologiques

---

Statistical mechanics and thermodynamics of systems with conformational transitions: applications to biological macromolecules

Soutenance le 10 décembre 2020 devant le jury composé de :

M. Enrico Carlon	Professeur à KU Leuven, ITP	(Rapporteur)
M. John Palmeri	Directeur de recherche CNRS, Charles Coulomb	(Rapporteur)
Mme Hélène Montès	Professeur à l'ESPCI, SIMM	(Examinatrice, Présidente)
Mme Cendrine Moskalenko	Maître de conférence à l'ENS de Lyon, Lab. Phys.	(Examinatrice)
M. Jean-Marc Victor	Directeur de recherche CNRS, LPTMC	(Examinateur)
M. Stefano Giordano	Chargé de recherche CNRS, IEMN	(Directeur de thèse)
M. Pier Luca Palla	Maître de conférence à l'Université de Lille, IEMN	(Co-encadrant)
M. Philippe Pernod	Professeur à Centrale Lille, IEMN	(Invité)

Thèse préparée au Laboratoire International Associé (LIA) LEMAC/LICS

IEMN - Cité Scientifique - Avenue Henri Poincaré

CS 60069 - 59 652 Villeneuve d'Ascq Cedex

Ecole Doctorale SPI 072

This thesis is dedicated to our beloved Neelah and Barbabelle.

Ad impossibilia nemo tenetur

Acknowledgments

First of all, I would like to thank Stefano Giordano, my supervisor, for his infinite patience and knowledge. I deeply thank him for his unfailing support throughout the PhD, and for his presence, a constant source of motivation. Each moment spent with Stefano was a source of learning, reflection and open-mindedness. I appreciate the answers he provided each time to my many questions, scientific or not. I am deeply indebted to him for his humanity and his kindness. This PhD is the most interesting experiment of my life.

I sincerely acknowledge Pier Luca Palla for his numerical contribution and his kindness.

I would like as well to thank my thesis brother, Pierre, for listening, his precious advice and his perfect compliments. Many thanks also to Aurélien, for his serenity and his skills in LaTeX, to Romain, for his calm and his optimism, and to Nicolas Tiercelin, my office partner, for his kindness. I am really grateful to Olivier Bou-Matar, Philippe Pernod, Yannick Dusch, Karim Talbi, Marc Goueygou, Cécile Ghouila-Houri and Yuxin Liu for their help and their kindness during the courses at Centrale. I would like to sincerely acknowledge Michaël Baudoin and Farzam Zoueshtiagh for guiding me through my university career and allowing me to discover the IEMN. Finally, I would like to thank the whole AIMAN-FILMS team for making me feel welcome.

I would also like to extend my thanks for funding my PhD to the Ecole Centrale de Lille and the Hauts-de-France region.

I am also very grateful to Hélène Montès, Cendrine Moskalenko, Jean-Marc Victor and Philippe Pernod for their participation as members of my thesis committee, and to Enrico Carlon and John Palmeri for kindly accepting to review this manuscript.

The PhD is a very special period in someone's life. I was very fortunate to be surrounded by nice people, starting with my family. Special thanks to my wonderful mum, my dad, my two beloved sisters, my granny, my dear cousin and my cats. I would also

like to thank my friends from the bottom of my heart, Audrey, Carole, Eve-Anne, and Blandine, on who I can rely on for many years, and Lucie and Frettt, for their constant presence (unlike the French trains). I am infinitely grateful to Theo and Carmelo for their friendship and advice. I am deeply grateful to my favorite Auchan team, and particularly to Louise, Elaine, Rachel, Juliette et Morgane. I express all my gratitude to my super coach Malika, for her invigorating fitness classes, and to Isabelle and Eric, for their open-mindedness. I also thank my IT friends Benjamin and Cyril for the healthy and balanced meals we shared. Many thanks to Lore for kindly helping me improve my English.

Finally, I would like to express all my gratitude and love to my life partner Jeremy, whose unconditional support is a real strength. His humour and world view are a great source of inspiration, which I hope it will last for a very long time.

Résumé en français

Le travail présenté dans cette thèse est consacré à l'étude du comportement thermoélastique des macromolécules d'origine biologique, telles que l'ADN et les protéines. En effet, de nombreuses macromolécules présentent un comportement bistable avec des transitions conformationnelles lorsqu'elles sont soumises à des expériences de traction. La réponse de ces macromolécules à la déformation suscite un vaste intérêt scientifique, que ce soit au niveau de la modélisation mathématique, de la simulation numérique ou des expériences de spectroscopie de force. Les principaux dispositifs de spectroscopie de force auxquels nous nous réfèrerons pour la comparaison des données sont présentés dans l'introduction, ainsi que les principales macromolécules étudiées et les motivations et buts de notre étude. De plus, les transitions conformationnelles sont observées dans d'autres domaines tels que la mécanique des phénomènes plastiques de la rupture.

Dans cette thèse, nous nous intéressons à la modélisation mathématique par la mécanique statistique de chaînes composées d'unités bistables, c'est-à-dire d'unités présentant deux états (ou conformations) stables. Les chaînes d'unités bistables que nous étudions sont composées d'un faible nombre d'unités, de telle sorte que nos études se situent loin de la limite thermodynamique. Ceci implique la considération de deux ensembles statistiques, les ensembles de Gibbs et de Helmholtz, pouvant mener à des comportements différents. La modélisation mathématique de la réponse des macromolécules à la déformation et aux fluctuations thermiques permet de tester définitivement la validité de la mécanique statistique des petits systèmes, grâce à la comparaison avec les résultats expérimentaux obtenus par la spectroscopie de force à l'échelle de la molécule unique, qui fournit de précieuses informations sur les réponses statiques et dynamiques induites par les forces appliquées. Notamment, nous comparerons les réponses force-extension de la filamine et de la titine obtenues par spectroscopie de force avec celles obtenues à l'aide de notre modèle.

Le but des théories présentées dans ce manuscrit est d'obtenir les fonctions de partition pour les ensembles de Gibbs et de Helmholtz, à l'aide de la méthode des variables de spin. En effet, les fonctions de partition permettent d'accéder aux valeurs moyennes des observables physiques d'intérêt. La méthode des variables de spin consiste en la décomposition d'une énergie potentielle bistable en deux paraboles, chacune identifiée grâce à une variable de spin. Ainsi, chaque parabole correspond à un état, plié ou déplié, de la macromolécule.

La description de l'approche par les variables de spin fait l'objet du second chapitre, où sont également présentées la thermodynamique des petits systèmes avec transitions conformationnelles et la statistique complète des chaînes d'unités bistables.

Puis vient le troisième chapitre concernant l'extensibilité des liens entre les unités bistables d'une chaîne. La prise en compte de l'élasticité des unités est primordiale, car elle joue un rôle majeur dans la définition de la relation force-extension de la macromolécule. L'analyse détaillée d'une chaîne d'unités extensibles est fournie et le modèle de chaîne idéale a été utilisé pour l'étude. Nous obtenons la fonction de partition de Gibbs exacte en introduisant de relativement hautes valeurs de la constante élastique, en cohérence avec les macromolécules réelles. Comme la fonction de partition de Helmholtz ne peut pas être obtenue par intégration directe en raison des interactions implicites engendrées par la condition isométrique, elle est calculée à partir de la fonction de partition de Gibbs à l'aide de la transformée de Laplace. Sa forme définitive est obtenue en termes de polynômes d'Hermite à index négatif.

La quatrième partie permet, quant à elle, de considérer les interactions entre les unités bistables d'une chaîne, grâce au modèle d'Ising et à son coefficient λ. Ce modèle permet de traiter différents cas, comme celui d'une interaction positive ($\lambda > 0$, assimilable à une interaction ferromagnétique), c'est-à-dire que le dépliage d'une unité favorise celui d'unités adjacentes, ou encore celui d'une interaction négative ($\lambda < 0$, assimilable à une interaction antiferromagnétique), c'est-à-dire que le dépliage d'une unité empêche celui d'unités adjacentes. Les interactions sont prises en compte dans l'ensemble de Gibbs grâce à la technique des matrices de transfert. Quant à la fonction de partition de Helmholtz, elle est obtenue à partir de la fonction de partition de Gibbs à l'aide de la transformée de Laplace. Nous étudions dans un premier temps le système loin de la limite thermodynamique. Puis nous proposons l'exploration de cas asymptotiques (faible et forte interactions), ainsi que l'évolution de la relation force-extension à la limite thermodynamique pour l'ensemble de Gibbs. Enfin, nous généralisons la théorie pour prendre en compte l'élasticité des liens et les interactions d'Ising.

Dans la cinquième partie, il est question d'hétérogénéité, paramètre important pour déterminer la séquence de dépliage dans le repliement des protéines, mais également pour représenter au mieux les macromolécules réelles, tout comme l'ajout de l'élasticité et des interactions. Le caractère hétérogène des unités est introduit au niveau énergétique grâce au spin. En effet, introduire l'hétérogénéité au niveau énergétique des unités permet de casser la symétrie et crée une inégalité parmi les probabilités de dépliage. Ainsi, à chaque instance de dépliage, la probabilité de dépliage d'une unité donnée tend vers 1 et celle des autres tend vers 0. La fonction de partition de Gibbs est obtenue par intégration directe. Quant à celle d'Helmholtz, elle est obtenue sous une forme explicite grâce à la transformée de Laplace et à la forme basée sur le déterminant des formules de Newton-Girard. Finalement, le concept d'identifiabilité est proposé, afin de mesurer la capacité du système à identifier la séquence de dépliage la plus probable.

Enfin, il a été démontré que la vitesse de traction des molécules a une influence sur la hauteur des pics de force et donc sur la relation force-extension. Ainsi, dans le dernier chapitre, la dynamique de la déformation est décrite à l'aide de la méthode de Langevin, qui permet de prédire la réponse force-extension des macromolécules biologiques dépliées par les techniques de spectroscopie de force à une vitesse de traction donnée. L'approche par la méthode de Langevin peut être acceptée comme compromis entre les méthodes basées sur les simulations de dynamique moléculaire et d'autres résultats obtenus par des approximations analytiques, dans le but de considérer un rang plus large de vitesses de traction.

Publications

Some ideas and figures have appeared previously in the following publications:

1. M. Benedito and S. Giordano, J. Chem. Phys. **149**, 054901 (2018), DOI: 10.1063/1.5026386 [1],

2. M. Benedito and S. Giordano, Phys. Rev. E **98**, 052146 (2018), DOI: 10.1103/PhysRevE.98.052146 (Editor's suggestion) [2],

3. M. Benedito, F. Manca, and S. Giordano, Inventions **4**, 19 (2019), DOI: 10.3390/inventions4010019 [3],

4. M. Benedito and S. Giordano, Phys. Lett. A **384**, 1-9 (2020), DOI: 10.1016/j.physleta.2019.126124 [4],

5. M. Benedito, F. Manca, P. L. Palla, and S. Giordano, Phys. Biol. **17**, 056002 (2020), DOI: 10.1088/1478-3975/ab97a8 [5].

Contents

Table of acronyms

ABP	Acting-binding protein
AFM	Atomic force microscope
B-DNA	Nucleic acid double helix
ddFLN4	Fourth domain of Distyostelium discoideum filamin
DNA	Deoxyribonucleic acid
dsDNA	Double-stranded deoxyribonucleic acid
FJC	Freely jointed chain
GFP	Green fluorescent protein
HS-AFM	High-speed atomic force microscope
HS-FS	High-speed force spectroscopy
LOT	LASER optical tweezers
MD	Molecular dynamics
MEMS	Micro-electro-mechanical systems
mRNA	Messenger ribonucleic acid
NMR	Nuclear magnetic resonance
RNA	Ribonucleic acid
rRNA	Ribosomal ribonucleic acid
S-DNA	Stretched deoxyribonucleic acid
SMFS	Single-molecule force spectroscopy
SNT	Silicon nanotweezers
ssDNA	Single-stranded deoxyribonucleic acid
TPR	Tetratricopeptide repeat
tRNA	Transfer ribonucleic acid
TWLC	Twistable worm-like chain
WLC	Worm-like chain

Chapter 1

State of the art and motivations

1.1 Why nanomechanics of macromolecules?

Nanomechanics of macromolecules is an important area of study, involved in many research fields. The application spectrum of this activity is wide, ranging from theoretical developments in statistical mechanics to applications in biology and related areas. The importance of taking nanomechanics of macromolecules into account is demonstrated below by the different points discussed, which concern the structural stability of proteins, the dynamics of macromolecules, the thermodynamics of small systems and the mechanical consequences on health.

1.1.1 Structural stability of proteins

Proteins are polymers made up of units, also called monomers or amino acids (see Fig.1.1). Amino acids are organic molecules, of vital importance for our bodies. They serve, for instance, as hormones, enzyme precursors and neurotransmitters. They are needed for many of the metabolic processes which take place in our bodies every day. In biochemistry, proteins have several levels of structure, like the primary structure which represents the sequence of amino acids [6, 7]. A change in the amino acids sequence can affect the structure of a molecule and cause problems in its function. This can lead to diseases, e.g. sickle cell disease. Sickle cell disease is an inherited disorder affecting the hemoglobin of red blood cells. This protein is essential for respiratory function, since it allows the transport of oxygen in our body. It is also involved in the elimination of carbon dioxide. For people with sickle cell disease, hemoglobin is abnormal, as seen in Fig.1.2. This disease comes from a point mutation on the 6^{th} amino acid of the chromosome 11, where glutamine is replaced by valine. When the concentration of oxygen in the blood decreases, it deforms the red blood cells, which then take the shape of sickles, instead

of being biconcave. This results in several characteristic symptoms of the disease, like chronic anemia, painful vaso-occlusive attacks and increased risk of infection [8]. The second level of protein organisation is called secondary structure and is useful to identify folded regions of the protein. The most common folded structures are α-helix and β-sheet, which are controlled by hydrogen bonds.

Then, the tertiary or three-dimensional structure of a protein refers to its organisation and folding in space. This folding gives to the protein its functionality. A well-know structure-function relationship concerns antigen-antibody bond. An antigen is a natural or synthetic macromolecule that, when recognised by antibodies or cells of an organism's immune system, is able to trigger an immune response in the organism. If the antibody does not have the correct form, it can not fix to the antigen. Therefore, the specific bond between them can not be established and the immune response does not occur. So the tertiary structure of many macromolecules controls the relation between structure and functions of proteins and macromolecules. Under the action of some factors, the spatial configuration of proteins can be destroyed, leading to changes in their physical and chemical properties and to removal of their biological activity. The capacity of macromolecules to keep effective their spatial configuration against mechanical factors must be tested to evaluate their ability to conserve their functions, in particular thanks to the force spectroscopy (see Section 1.2). Generally, this phenomenon is called protein denaturation and these proteins are known as denatured proteins or inactive proteins. In other words, denatured proteins lose their biological activity and they can no longer perform their specific biological functions. For instance, if a macromolecule is pulled with a large force, it loses its biological function as the force reduces the structural stability of the protein. For an enzyme, it represents the loss of its catalytic capacity. For an antibody, it represents the loss of its ability to bind to an antigen. So, the mechanical actions on biological macromolecules can lead to crucial modifications in their functions, with crucial physiological consequences.

1.1.2 Dynamics of macromolecules

Many biological processes take place in the cell, like cell cycles, protein biosynthesis or replication. These processes have characteristic times, depending on the shape and the function of involved molecules and therefore on the underlying mechanics. The dynamic mechanical response is very important, especially for the characteristic times of specific internal processes and chemical processes. The classical pictures in books and the static measurements through NMR or X-ray diffraction could give the inaccurate impression

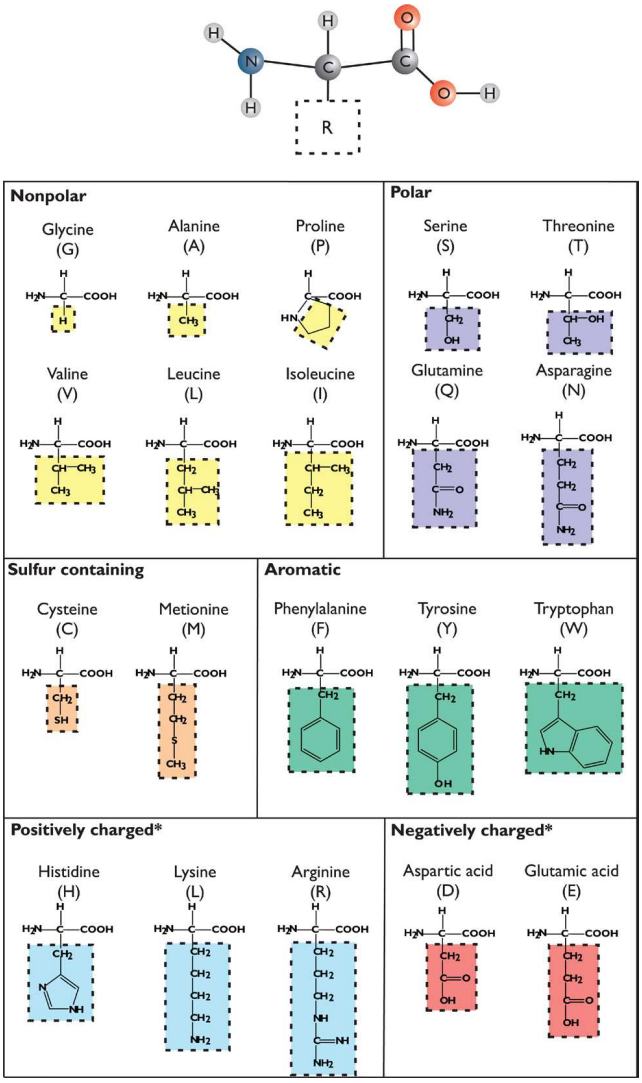

Figure 1.1 – Scheme of amino acids. Proteins are made up of subunits called amino acids. An amino acid is made up of a central carbon atom, known as the α-carbon, covalently bound to a hydrogen atom, an amino group (NH_2), a carboxyl group (COOH), and a side chain group (R group). The side chains of the 20 standard amino acids are shown in this figure, where the amino acids are grouped according to the properties of their side chains. Nonpolar, aliphatic amino acids (yellow) have hydrocarbon side chains and are typically hydrophobic. The aliphatic polar uncharged amino acids (purple) contain an amino or hydroxyl group and can form hydrogen bonds with atoms in other polar amino acids or water molecules. Aromatic amino acids (green) contain an aromatic ring and can be nonpolar or polar. The sulfur containing amino acids (orange) are named cysteine and methionine. The side chain of methionine is hydrophobic. The side chain of cysteine can form covalent disulphide bonds with other cysteine residues due to the sulfur-hydryl (SH) group found on its side-chain. At neutral pH the charged amino acids can be either charged negative, (acidic, red) or positive (basic, blue) [6].

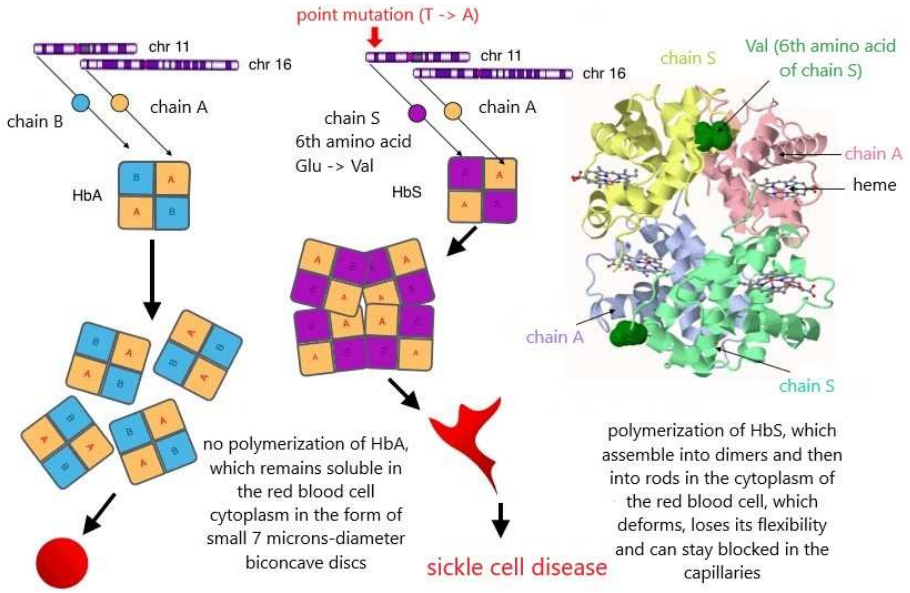

Figure 1.2 – Schematic diagram of normal and abnormal hemoglobin. Point mutation on chromosome 11 replacing glutamine by valine leads to abnormal hemoglobin. This mutation leads to sickle cell disease [9].

that proteins exist in a single form [10, 11]. However, proteins are dynamic and not static. Weber qualified proteins as "kicking and screaming stochastic molecules" [12] and it has been proved that a given protein can assume many different configurations, called conformational substates or metastable states [13–15]. These states represent different geometrical configurations of the chain in the space. The protein can be described by the energy landscape, a very important concept allowing to bring information about the structure, dynamics, and functions of biomolecules. To perform these functions, most proteins need to move. Protein motions allow to ligands like O_2 to enter and leave myoglobin or hemoglobin or again to substrates to reach the enzyme center in enzymatic reactions. These motions can either be thermal equilibrium fluctuations, since the temperature creates fluctuations in the system modifying the dynamic evolution of the system, or again non-equilibrium relaxations, for example induced by chemical reactions. In terms of the energy landscape, motions can be described as jumps of the system from a substate to another substate. At nano- and micro-scale, changes between states occur very fast and rate coefficients of fluctuations are in the range from fs^{-1} to s^{-1}, or even less. Hence, many different tools are necessary to study dynamic behaviours, like out-of-equilibrium statistical mechanics or methods based on Langevin and Fokker-Planck equations. When a molecule switches from one substate to another one, we can use Kramers theory [16] or its generalisations [17–19], which manages characteristic times to

switch between two states of bistability. This large spectrum of theories is very useful to describe dynamic biological systems, since it is often able to summarise the behaviours of complex systems with simple stochastic rules.

1.1.3 Thermodynamics of small systems

Why using statistical mechanics instead of classical mechanics to study macromolecules like DNA, proteins or other biological structures? In fact, for small systems, the energy related to the thermal fluctuations is comparable to the mechanical energy (for instance, energy accumulated in elastic bonds, i.e. the enthalpic contributions). Imagine a polymer whose one end is fixed, which only moves with thermal fluctuations. In this case, the polymer is randomly distributed on a sphere centred on the fixed end as it freely and isotropically explores the whole configurational phase space. To align polymer i.e. to extend it, the other end is now pulled. The polymer reacts against this deformation by creating an entropic force, as it prefers to be in the configuration of a sphere centred on fixed end. If the temperature of the system is high, the polymer wants to explore all configurations randomly. If a mechanical force is pulled on it, the polymer is stretched and an entropic force is created to go against the aligned configuration, which prevents the exploration of the phase space. Therefore, to study the behaviour of macromolecules and create some pertinent models, several forces have to be taken into account at nano-scale where thermal fluctuations play a crucial role. Entropic forces represent the paradigmatic example. Statistical mechanics allow to do so.
Hence, we can use statistical mechanics to study macromolecules made of monomers, like proteins or DNA. Here, we focus on the nanomechanics of macromolecules. To do so, we can consider two different systems, yielding two different results. We can consider a molecular chain with the first end-terminal tethered on a substrate. Firstly, we can imagine to apply a force at the second end of the macromolecule (Gibbs ensemble) or, secondly, to prescribe the spatial position of this second end (Helmholtz ensemble). Then, we suppose to measure the force-extension relation in both cases. If the system under consideration consists of infinitely many units, the statistical mechanics results under these different conditions are identical, as observed in Fig.1.3, when the number of particles increases [20]. In this case, the thermodynamic limit is attained and it represents the limit for a large number of monomers. However, if a system consists of few units as it is sometimes the case for proteins, equivalence between the ensembles is lost. The latter can be observed in real life biological experiments. In this case, single-molecule force spectroscopy techniques are useful to show non-equivalence between ensembles. Gibbs and

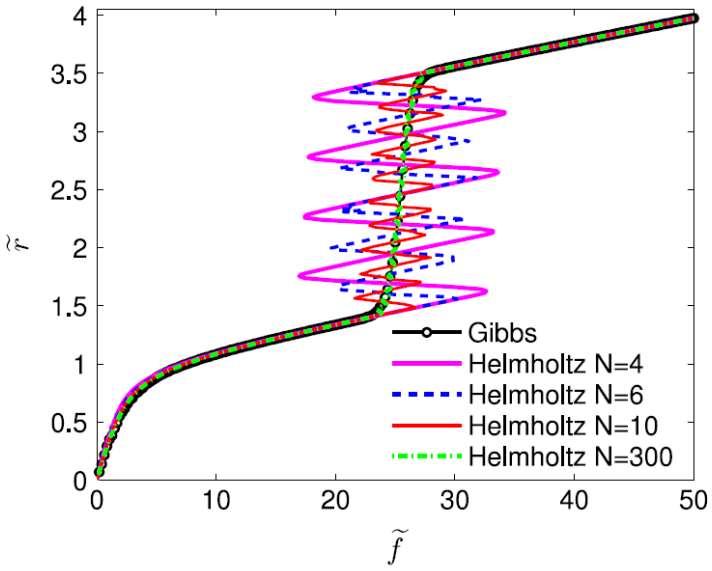

Figure 1.3 – Force–extension curves for Gibbs and ensembles. The response at constant applied force, i.e. in the Gibbs ensemble, shows a plateau force. This behaviour has been observed in the over-stretching of DNA, and in polysaccharides such as the dextran. Regarding the Helmholtz ensemble, with $N = 4, 6, 10, 300$ units, when the chain length is increased, the width of the peaks is decreased until, at a large enough N, the force–extension curve approaches again the plateau curve of the Gibbs ensemble. One can see the convergence between both ensembles for a sufficiently high N [20].

Helmholtz ensembles made up of few units are an example of ensembles which verify the non-equivalence. The Gibbs ensemble is typically referred to as the isotensional boundary condition. Typically, the force-extension response observed shows a plateau force, which the threshold force for which all units unfold at the same time. The Helmholtz ensemble, for its part, is referred to as the isometric boundary condition. Typically, the force-extension response shows a saw-tooth pattern, with force peaks, each peak corresponding to the unfolding of a unit of the chain. In the case of small systems, if the force is imposed or the extension is prescribed, different mechanical responses are observed. Theory predicts these differences and they are verified and tested in experiments with macromolecules, as observed in Fig.1.4 and Fig.1.5. It is important to underline that the differences in the ensembles behaviours can be appreciated both with equilibrium and non-equilibrium statistical mechanics. These differences will be studied in detail in this thesis for several particular cases. We note that the non-equivalence of the ensembles can be observed also in special cases at the thermodynamic limit (e.g., confined polymers or adhesion of polymers) [21–23].

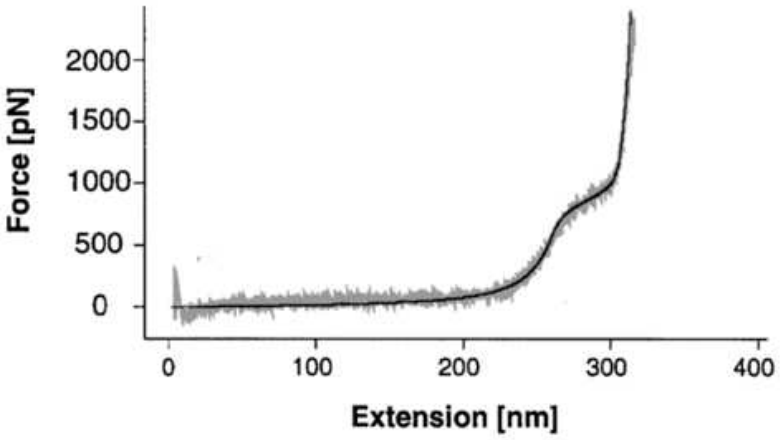

Figure 1.4 – Force-extension response of several dextran molecules. Extension curves of 20 different carboxymethylated dextran filaments obtained with an AFM device with various contour lengths from 50 nm to 2 μm measured on different samples with different cantilevers were normalised according to their length [24].

1.1.4 Mechanical consequences on health

As stated above, the structure of proteins determines their function. Misfolding of a protein or change in its primary structure can affect the tertiary structure. An incorrectly folded protein can have dramatic consequences on human body and health. For instance, misfolding of a protein can lead to type 2 diabetes, Alzheimer disease, Huntington disease (see Fig.1.6) and Parkinson disease [26]. In all these cases, a soluble protein that is normally secreted from the cell is misfolded and secreted as an insoluble protein. The latter form is called an amyloid fiber and all the diseases due to this misfolding are known as amyloidoses.

Another example concerns mad cow disease, transmissible to humans, that made us aware of the importance of the tertiary structure of the protein on health. Indeed, this disease is neither caused by a virus nor a bacterium. The pathological agent responsible for the disease is a protein called *prion*, found by the Nobel Prize for medicine 1997, Stanley Prusiner [28]. For a sick person, the primary sequence of this protein does not change. However, the tertiary structure is deeply modified, going from a structure mainly made of alpha helices to a very rich structure of beta sheets, as observed in Fig.1.7. This new structure is extremely stable and resistant to almost all disinfection techniques. These examples show the dramatic consequences of the structure-function relation on the health. Therefore, it is vitally important to have experimental methodologies to investigate the static and dynamic responses of macromolecules of biological origin.

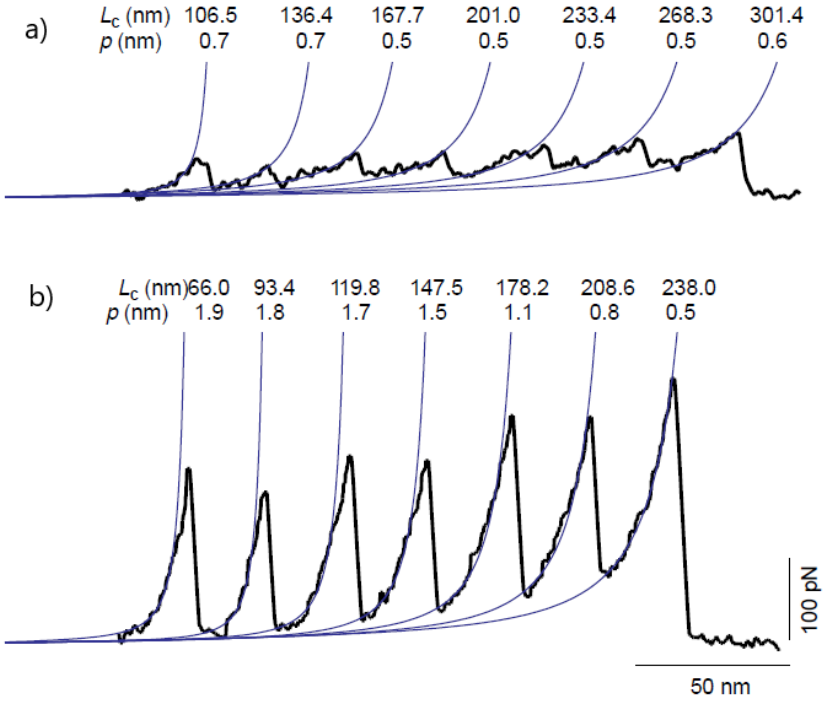

Figure 1.5 – Comparison of force–extension curves for spectrin and titin. On panel a), a force–extension curve for the cytoskeletal protein spectrin showing the contour length L_C and persistence length p for each of the domains is shown. On panel b), a force–extension curve for a recombinant fragment of titin consisting of titin Ig domains 27–34 showing the contour length and persistence length for each of the domains [25].

1.2 Single-molecule force spectroscopy

As described above, the functions of several macromolecules (for instance, proteins) are strongly related to the three dimensional conformation of their polymeric chain. Studying the relation between the three dimensional conformation of a macromolecule and its function can be directly done by using force spectroscopy methods (see Fig.1.8), which can be used to unfold the native folded structure of a macromolecule. The controlled unfolding leads to the estimate of the involved forces, the energy landscape and many other dynamic properties of the system under investigation.

Recent developments of mechanical experiments on single-molecule allowed to better understand intra- and intermolecular forces, by introducing important information about the thermodynamics and kinetics of several molecular processes. Single-molecule experiments are typically based on optical tweezers, magnetic tweezers, microelectromechanical systems (MEMS) and atomic force microscope (AFM) [31]. With such devices,

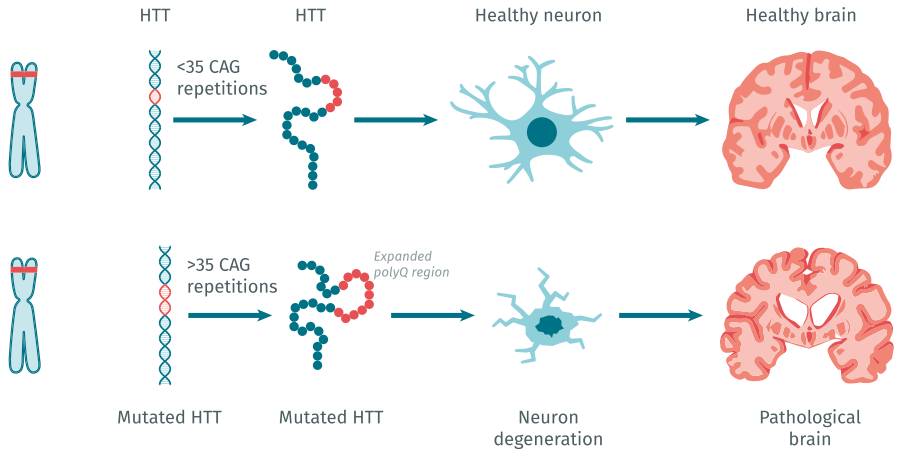

Figure 1.6 – Representation of the molecular genesis of Huntington's disease. Huntington's disease is a progressive and invariably fatal neurodegenerative genetic disorder. The gene of this disease, Huntingtin, contains a repeat of CAG codon coding for glutamine (a codon is a sequence of three nucleotides on a messenger ribonucleic acid). If the repeat contains 35 or more repeats, Huntington's disease develops and it results in the death of brain cells, leading to, inter alia, problems with mental abilities and a general lack of coordination [27].

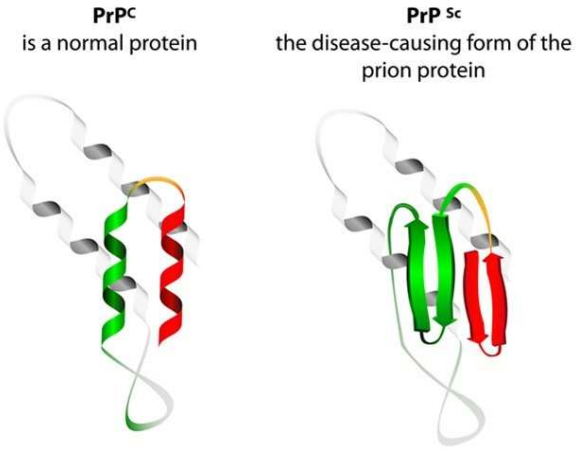

Figure 1.7 – Scheme of a normal prion and a disease-causing prion. In its native conformation, the prion protein consists of two domains, one is unstructured, the other one is globular made up of helices. However, during the course of the disease, if the primary sequence of the protein is preserved, the globular domain loses its alpha helices and forms beta sheets [29].

it is possible to measure the force-extension relation for a molecular chain under different conditions. These experiments can be applied to DNA, RNA, polysaccharides, or again

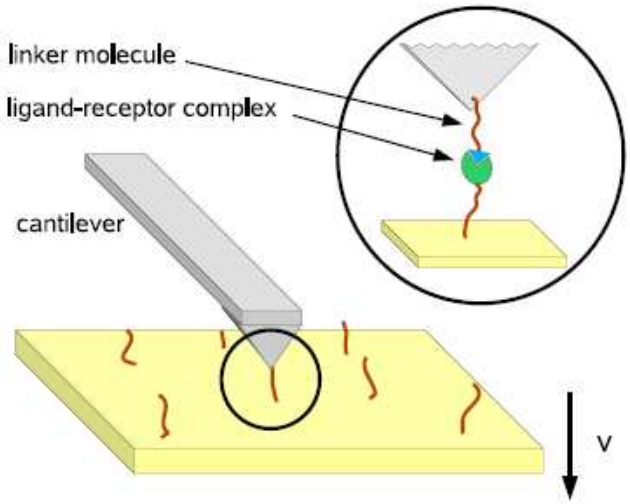

Figure 1.8 – Schematic illustration of a dynamic force spectroscopy experiment. A receptor is immobilized on the surface, and the ligand is connected via a linker to the tip of an AFM cantilever which serves as a force transducer. The distance between surface and tip can be controlled with a piezoelectric element [30].

proteins. The above techniques permit a clearer comprehension of the equilibrium and out-of-equilibrium thermodynamics of small systems and the experimental verification of the small systems thermodynamics. These devices explore a large range of stiffness. For example, LASER optical tweezers (LOT) and magnetic tweezers are considered as soft devices, with a stiffness from 10^{-4} to 10^{0} pN/nm, whereas AFM and MEMS are considered as a hard device, with a stiffness from 10^{0} to 10^{2} pN/nm.

1.2.1 Conventional and high-speed atomic force microscope

The atomic force microscope is a well-known technique invented in 1985 by Gerd Binnig, Calvin Quate and Christoph Gerber and commercialised for the first time in 1989 [32]. The AFM is a high-resolution scanning probe microscopy instrument allowing to reach the atomic resolution. In its primary operation mode, also known as "contact mode", AFM allows to visualise the topography of a sample surface by scanning it horizontally with a sharp tip placed at the extremity of a cantilever. As high-resolution imaging tool, it permits to measure the roughness of a sample surface.
The two main components of the AFM are the cantilever, which acts as a flexible sensor and a piezoelectric positioner, in order to control the sample position in the nanometric range. An AFM consists of a cantilever, mounted on a cantilever holder, whose position is controlled by a piezoelectric device. A focused laser beam is reflected off the surface

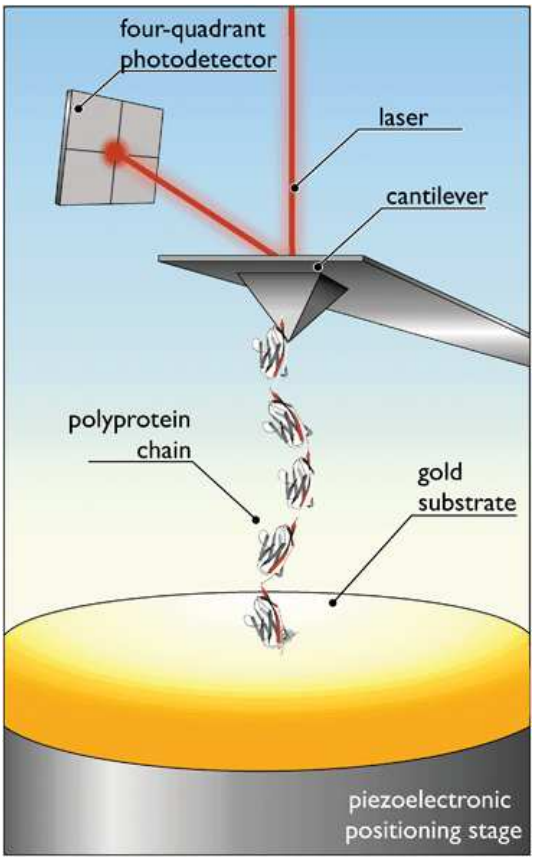

Figure 1.9 – Schematic diagram of principle of an AFM for SMFS. The protein is bound at one end to a substrate (here, gold). The gold substrate is placed on top of a piezoelectric device. A laser beam, reflected off of the cantilever tip into a photo-detector, is used to monitor the deflections of the cantilever [33].

of the cantilever on a photodetector. The moves of the cantilever can be monitored by following the movement of the laser spot on the photodiode. The incident light is converted into voltage by the photodiode output, which then outputs the voltage difference when the laser spot moves. The AFM measures the angular deviation of the laser spot allowing to obtain the forces exerted on the flexible cantilever. This technique has also been generalised to stretch a macromolecule and therefore to perform single-molecule force spectroscopy (SMFS), as observed in Fig.1.9. In this case, the pulling speeds used in SMFS experiments vary in the range of a few nm/s to about 10 μm s^{-1}, and the cantilever stiffness is typically between 6 and 100 pN nm^{-1}.

Recently, a high-speed AFM (HS-AFM) has been developed to unfold proteins at higher velocities than conventional AFM, allowing to reach pulling speeds of the order of 4000 μm/s [34–38]. This was made possible by using short cantilevers and miniature

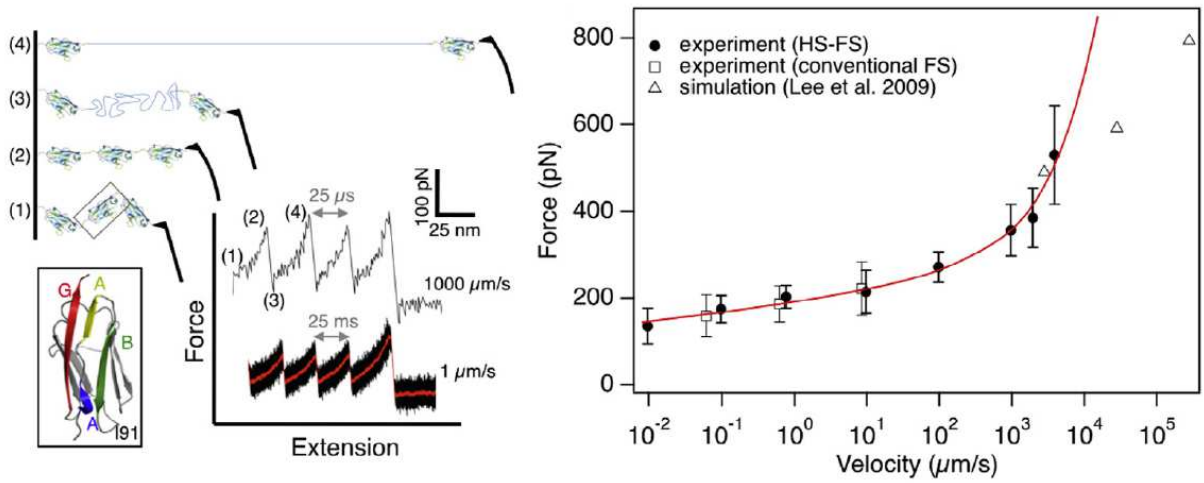

Figure 1.10 – High-speed force spectroscopy (HS-FS) of titin unfolding. On the left panel, one can see schematic process of titin forced unfolding showing the relevant steps: 1) relaxed polyprotein, 2) polyprotein stretching, 3) unfolding of one domain and 4) unfolded domain stretching. Two examples of force-extension curves revealing three unfolding peaks at 1 μm s^{-1} (bottom) and 1000 μm s^{-1} (top) are also shown. Gray arrows represent the time required to unfold and stretch a single domain. The inset shows the crystal structure of I91 domain with the relevant ß-strands in colour (PDB 1TIT). On the right panel, one can see the dynamic force spectrum of titin I91 unfolding using HS-FS (full circle), conventional AFM (square) and molecular dynamics simulations (triangle) [35].

piezoelectric actuator (see Fig.1.10). This is a significant development since it permits to SMFS experiments to be comparable to those probed in molecular dynamics simulations, which offer atomic-level descriptions of the forced unfolding [39, 40].

1.2.2 Magnetic tweezers

Magnetic tweezers are another well-known device used for force spectroscopy experiments. Most commonly, they are used to study mechanical properties of biological macromolecules like DNA or proteins in single-molecule experiments. They are also used for other applications like rheology of soft matter (study of flow) or study of force-regulated processes in living cells. In 1996, Strick, Bensimon and Croquette assembled the first magnetic tweezers to study elasticity of supercoiled DNA (over- or under-winding of a DNA strand) [41].

Magnetic tweezers device consists in magnetic micro-particles, which can be manipulated with the help of an external magnetic field (see Fig.1.11). This field is generally generated with a pair of permanent magnets (made up of rare earth), but generation based on elec-

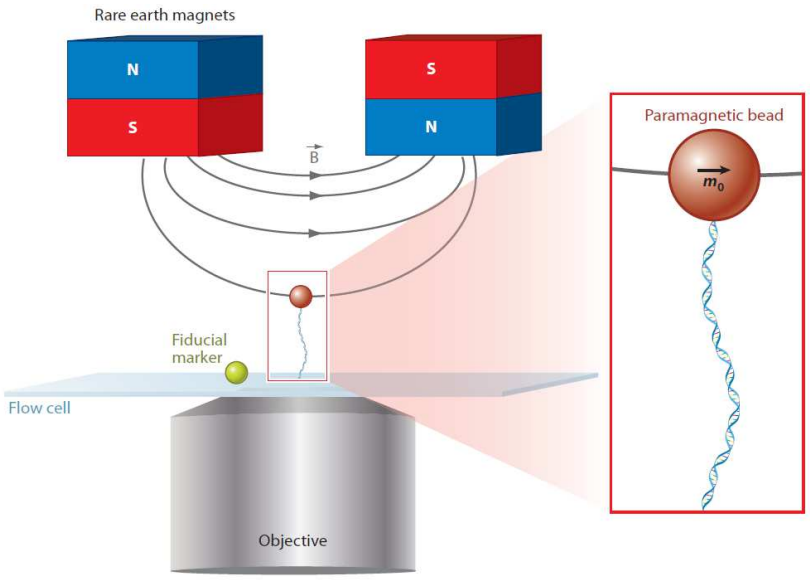

Figure 1.11 – Schematic of basic implementation of magnetic tweezers. A molecule is tethered between the surface of a flow cell and a paramagnetic bead. The magnetic field generated by a pair of magnets induces a magnetic moment $\vec{m}_0$ in the paramagnetic bead. The bead experiences a force proportional to the gradient of the field. The molecule can be coiled by rotating the external magnet. The fiducial marker serves as reference to measure the position of the bead [46].

tromagnets [42–44] or generation with near-field of a single permanent magnet [45] have also been implemented. The position of the magnetic particles is then determined by a microscopic objective with a camera. There are two main types of magnetic particles, the superparamagnetic beads and the ferromagnetic nanowires. For the superparamagnetic beads, one uses spherical particles, whose diameter is micrometric. The bead is made of a porous latex matrix, in which particles were inserted. The latex is autofluorescent, what may be an advantage. About the ferromagnetic nanowires, they allow to explore much larger applied forces. The length of the wires is in range from tens of nanometers up to tens of micrometers. Different ways of assembly of magnetic tweezers exist to fill diverse experimental needs, like the organisation of the magnetic field. Exerted forces are typically in the range from piconewton to nanonewton.

1.2.3 Optical tweezers

It is well known that light exerts forces on matter. The forces being weak, they remained mere curiosities until the invention of the laser in the 1960s. From then on, the evolution was rapid and continuous thanks to the work of Arthur Askin in particular [47]. The 1970s marked the first realisations of optical traps using two counter-propagating

beams, the first experiments of optical levitation of microspheres and the first realisation of a single laser beam strongly focused by a high digital aperture objective, creating an optical gradient force. The year 1986 is considered as the birth year of optical tweezers [48]. On the one hand, optical tweezers allow the precise handling of objects without any contact, with the consequence of remaining in a perfectly sterile environment during handling. On the other hand, the forces generated by optical tweezers are typically equivalent to the forces involved in a large number of cellular processes (adhesion, cytoskeletal mechanics, motricity, operation of molecular motors, etc.). Other fields than biology use optical tweezers, such as photochemistry or physics, such as the study and control of colloidal particles, the setting in motion and control by a light beam of micromotors or micropumps [49]. The essential elements for making optical tweezers are a laser beam, a high numerical aperture microscope objective, a sample containing the objects to be manipulated and a viewing device (see Fig.1.12). To account for the forces by the optical trap, it is necessary to rely on Lorentz-Mie's generalised theory that describes diffusion of the light by an object of any shape. The reflected rays contribute to the diffusion force that pushes the object in the direction of the laser beam, but the refracted beams incident at a high angle will keep the bead at the focus point, where the light intensity is highest, thanks to gradient forces. The trap is stable as soon as the gradient forces exceed the diffusion forces.

Dynamic studies of single molecules such as DNA or RNA molecules have progressed through manipulations using optical tweezers. It is now possible to measure the force applied to a DNA molecule that is attached at one end to the surface of a holder and at the other end, to a latex bead that is held in place with the optical tweezers, as seen in Fig.1.13. In this way, the elasticity of DNA can be measured directly. Separation of the DNA double helix is achievable through the action of a mechanical force exerted with optical tweezers. It was thus possible to determine the force with which the nitrogenous base pairs are bound and it was established that these forces vary according to the base pair sequence. Optical tweezers allowed the identification of defective DNA structures like base mismatches, missing bases or crosslinks. The latter occur in DNA with high frequency and must be efficiently identified and repaired to avoid direct consequences such as genetic mutations [51–54].

1.2.4 MEMS

Last but not least come microelectromechanical systems (MEMS) and the related micromechatronics and microsystems, which constitute the technology of microscopic de-

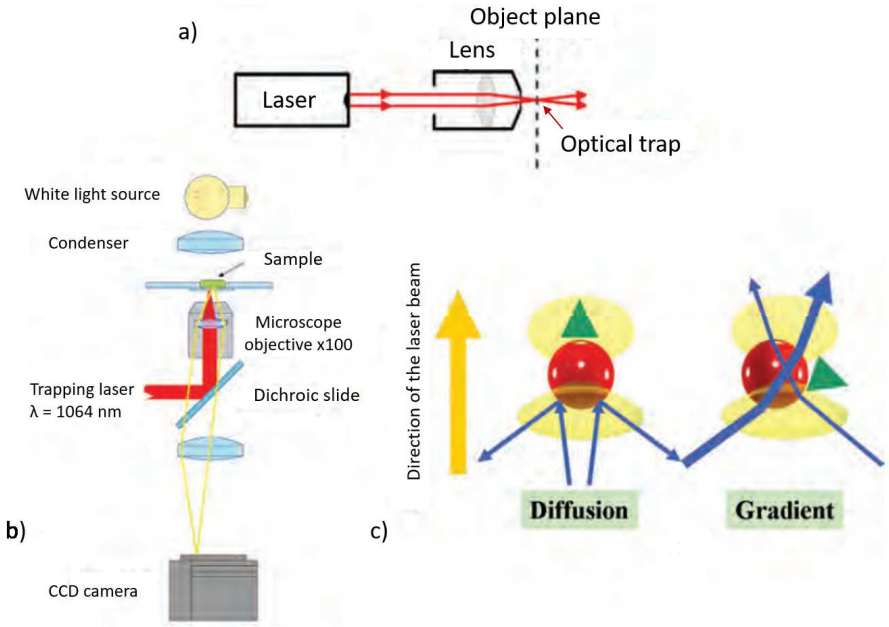

Figure 1.12 – Scheme of the general principle of optical tweezers. a) Essential elements for optical tweezers. b) Scheme of the general device. c) Physical principle of optical tweezers [50].

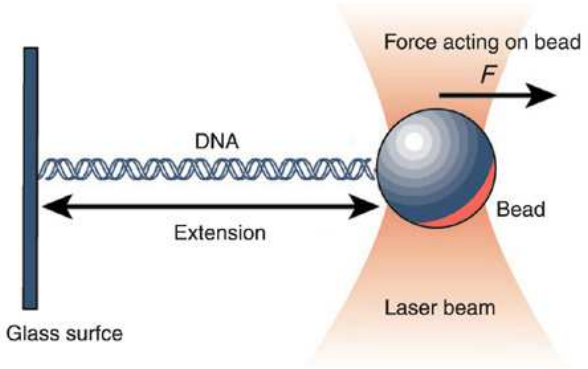

Figure 1.13 – Scheme of optical tweezers used to directly measure DNA elasticity [50].

vices, especially those with moving parts. MEMS were developed in the early 1970s as derivatives of microelectronics and were first commercialised in the 1980s with silicon pressure sensors, that quickly replaced older techniques and still form a significant part of the MEMS market. Since then, the MEMS field has been booming. These devices are used in many fields like automotive, aeronautics, medicine, biology, telecommunications and in several "everyday" applications such as high-definition television sets or car airbags.

Generally, MEMS are made up of components from 1 to 100 μm. Devices holding MEMS generally range in size from 20 micrometres to a millimetre. They usually consist of a central component that processes data (an integrated circuit chip such as microprocessor) and several components that interact with the environment (such as microsensors). Due to the large surface area to volume ratio of MEMS, forces created by ambient electromagnetism (e.g., electrostatic charges and magnetic moments) and fluid dynamics (e.g., surface tension and viscosity) have to be more taken into account for design than with larger scale mechanical devices. Considering surface chemistry makes the difference between MEMS technology and molecular nanotechnology or molecular electronics. Moreover, devices like AFM, optical tweezers and magnetic tweezers are bulky and rather expensive. Some experiments need to be realised in tiny or confined areas and MEMS allow to fill these gaps. The fabrication of MEMS evolved thanks to the process technology in semiconductor device fabrication, i.e. the basic techniques are deposition of material layers, patterning by photolithography and etching to obtain the required shapes. The materials used for MEMS manufacturing are silicon, polymers, metal and ceramics.

MEMS devices can be, *inter alia*, used to study macromolecules like DNA or proteins. For instance, one reports a micromachined DNA manipulation platform to stretch and rotate a single DNA molecule, as a novel micromachined magnetic tweezers for DNA manipulation, as seen in Fig.1.14 [55]. One could also cite a new hybrid field microfluidics (HFM) approach, employing both hydrodynamic forces and an electric field to regulate DNA initial conformations [56] or real time monitoring of the dynamics of the reactions without any surface or molecular modifications [57]. To work with nanoscale systems, it is possible to use nanoelectromechanics systems (NEMS), for instance, to realise mass spectrometry in real time [58].

MEMS technology is attracting a lot of interest, including that of the SMMIL-E, an international research program, whose goal is to transfer works on microsystems made by the University of Tokyo from Japan to France, in order to improve the research against cancer. To do so, French national research center CNRS and the University of Tokyo have established in Lille in 2016 a mirror site of the LIMMS, their common international laboratory installed in Tokyo since 1995. As part of this program, the Hauts-de-France region, the CHR of Lille, the CNRS and the LIMMS collaborate on the realisation of a specific MEMS device, the silicon nanotweezers (SNT), allowing to mechanically characterise a DNA bundle exposed to an ionising radiation beam, here delivered by a therapeutic linear

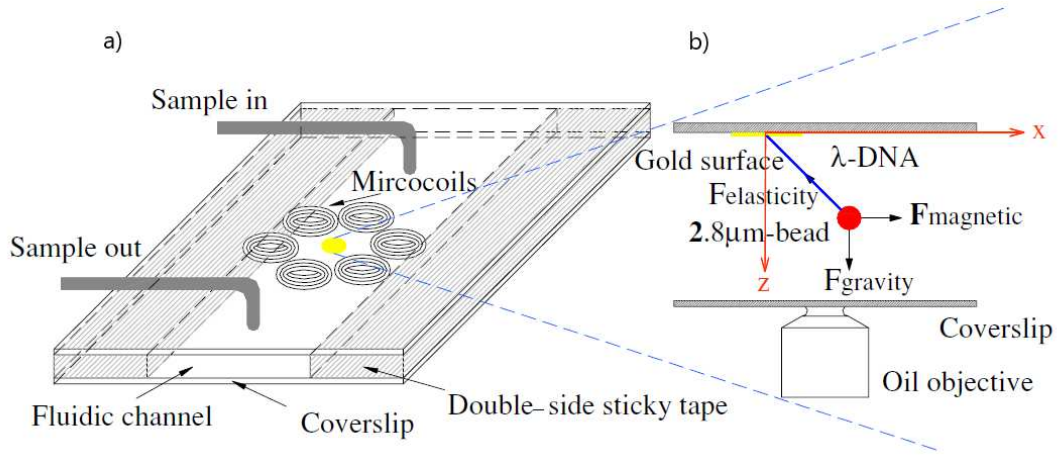

Figure 1.14 – Schematic diagram of the micromachined platform for DNA. a) Schematic of the DNA manipulation platform integrated with six microcoils, a fluidic channel and a gold-patterned surface. b) A tethered-DNA magnetic bead is in equilibrium while applying the magnetic force, DNA elastic force and the gravity force (which is the apparent weight of the bead in the buffer solution). The tethered-bead DNA could be stretched and rotated under the magnetic field [55].

particle accelerator (LINAC) (see Fig.1.15). The radiation induces a mechanical degradation i.e. a population of breaks in a DNA bundle, which can be quantified by measuring the elastic properties of the bundle itself. Hence, one could provide a direct relationship between radiation dose and its damaging effects. Silicon nanotweezers device resist against the environment of radiation beams, by still allowing molecular-level accuracy. The first real-time observations carried out using SNTs allow a better understanding of fundamental and clinical studies of the mechanisms of DNA degradation under ionising radiation, for better treatment of tumors [59, 60].

1.3 DNA, RNA and models

Now that the main techniques of single-molecule force spectroscopy have been presented, macromolecules whose they allow the study and models used for their representation are briefly introduced.

1.3.1 DNA and RNA

Deoxyribonucleic acid, also known as DNA, is a biological macromolecule present in all cells and in many viruses, bacteriae, and so on. DNA contains all the genetic information called *genome*, which allows living beings to develop, live and reproduce. In

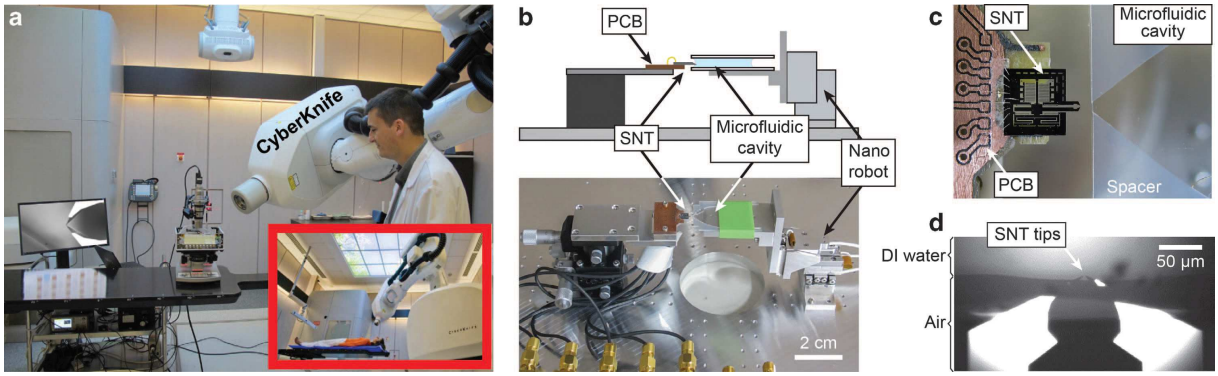

Figure 1.15 – SNT and microfluidic set-up in the hospital. a) Set-up on a patient bed support. The medical physicist focuses the beam direction of the CyberKnife (LINAC) on the tips of the SNT. b) The SNT is aligned in front of the microfluidic cavity. c) The top view of the SNT aligned to insert the tips into the cavity. d) Only the tips of the SNT enter the liquid so that the actuators and sensors can provide their in-air performance [60].

living cells, DNA molecules are made up of two anti parallel strands wrapped around each other to form a double helix, as observed in Fig.1.16. Hence, DNA is said *double-stranded*. Each of the two strands is called *polynucleotide*. The latter consists of monomers called *nucleotides*, linked to one another by covalent bonds between the deoxyribose of a nucleotide and the phosphate group of the following nucleotide, thus forming a chain in which polysaccharides and phosphates alternate, with nucleic bases each linked to a polysaccharide. A nucleotide is formed of a nucleic base or nitrogen base - there are four of them: adenine (A), cytosine (C) (both forming pyrimidines), guanine (G), and thymine (T) (both forming purines) - linked to a polysaccharide - here, deoxyribose - itself linked to a phosphate group. The *genetic code* is a set of correspondence rules allowing the genetic message made up of nucleotides to be translated by a cell into a polypeptide chain made up of amino acids. Each sequence of three consecutive bases carried by the mRNA corresponds to a single amino acid.

The order in which the nucleotides follow one another along a DNA strand constitutes the *sequence* of this strand. It is this sequence which carries the genetic information. The latter is structured into genes, which are expressed through transcription into ribonucleic acid (RNA). RNA plays an active role in cells, as it allows to control gene expression or to communicate responses to cellular signals for example. Cellular organisms use messenger RNA (mRNA, which is a coding RNA) to convey genetic information (using the RNA nitrogen bases, there are four of them: guanine (G), uracil (U), adenine (A), and cytosine (C)) which allows synthesis of specific proteins. Many viruses encode their genetic information using an RNA genome. One of the main and most important roles

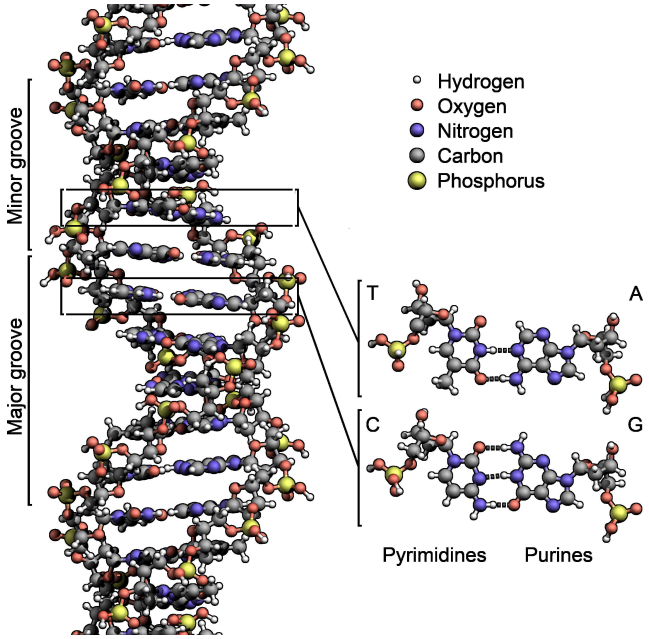

Figure 1.16 – Structure of DNA double helix showing the structure of the four nucleic bases: adenine, cytosine, guanine and thymine. The atoms in the structure are colour-coded by element and the detailed structures of two base pairs are shown in the bottom right [61].

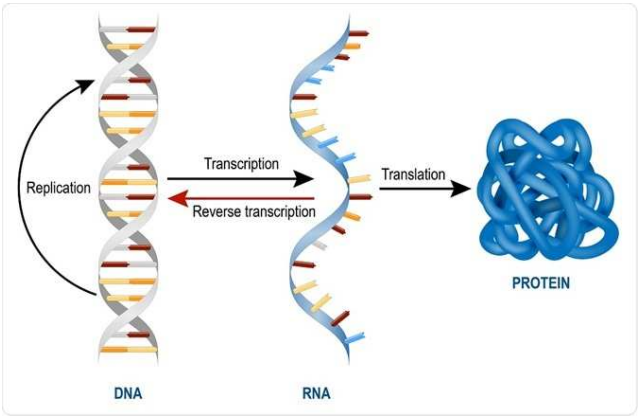

Figure 1.17 – Scheme of RNA role between DNA and protein [62].

is protein synthesis in ribosomes (macromolecules composed of RNA and proteins). This process uses transfer RNA (tRNA) molecules to deliver amino acids to the ribosome, where ribosomal RNA (rRNA, both non-coding) then links amino acids together to form coded proteins. Some RNA roles are shown in Fig.1.17.

The association of DNA and proteins in which DNA is compacted in the nucleus in eukaryotic cells is called *chromatin*. The latter consists of a combination of DNA and proteins of two types: histones and non-histones, as seen in Fig.1.18. On the one hand,

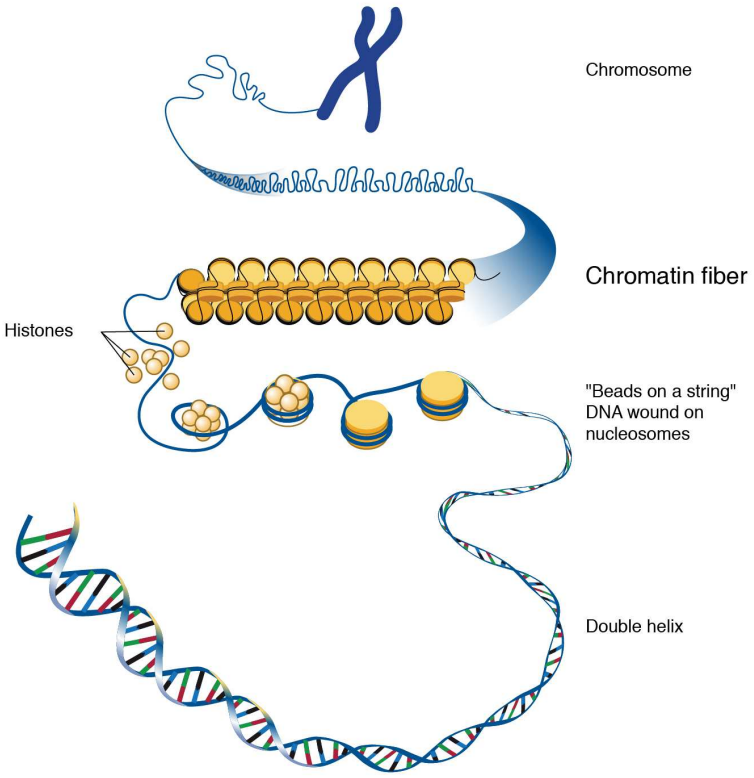

Figure 1.18 – Schematic diagram of DNA compaction [63].

non-histones proteins enable the compaction of DNA, but contrarily to histones proteins, they are not bound to DNA. On the other hand, histones are proteins located in the nuclei of eukaryotic cells and in archaea. They are the main protein components of chromosomes, closely associated with DNA since they allow to compact it. This compaction forms structures called *nucleosomes*, where DNA is wound around histones like thread around a coil. As DNA is structured in the chromatin with coils to build up information, knowing DNA mechanics allows to understand how to access to chromatin coils, in order to read genetic information.

Single molecule techniques allowed to study the force-extension response of chromatin, revealing the numerous conformational transitions undergone by the chromatin, as seen in Fig.1.19. The multiple conformations adopted by chromatin demonstrate a complex energy landscape due to the different levels of DNA compaction [64–68].

Force spectroscopy allows to study the response and the structure of DNA, which was first elucidated in 1953 by Watson and Crick, the Nobel Prize in Physiology or Medicine 1962 [69]. In 1992, Smith, Finzi, and Bustamante made direct experimental mechani-

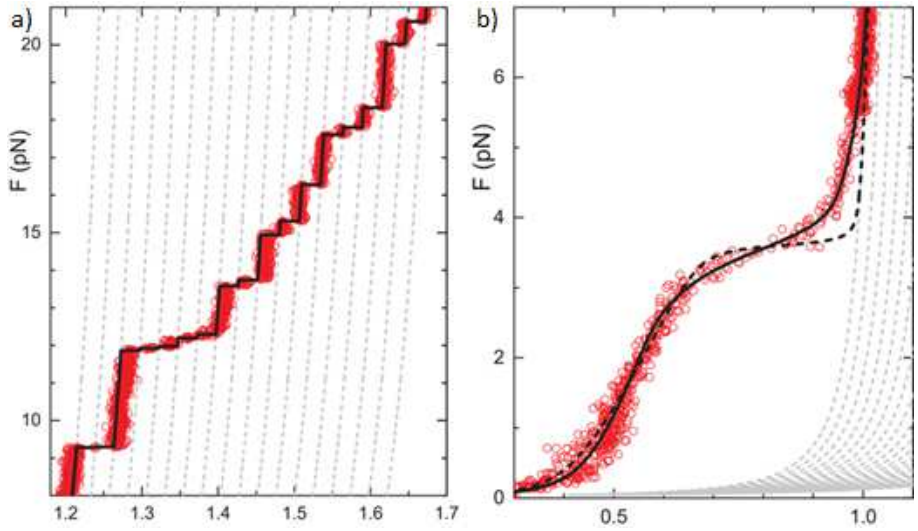

Figure 1.19 – Detailed analysis of the unfolding of a single chromatin fiber. a) A zoom in on the high-force region shows discrete steps in extension. Dashed gray lines represent the extensions of all states that are composed of extended and fully unwrapped nucleosomes. The black line shows the best match between individual data points and the various states of unwrapping. b) Unfolding of a 15*197 nucleosome repeat lengths chromatin fiber at low force. Below 7 pN the extension starts to deviate from a string of extended nucleosomes (gray dashed lines). A single transition (black dashed line) does not capture the force-extension data [64].

cal measurements on DNA by using magnetic beads [70]. They obtained extension versus force curves for individual DNA molecules at three different salt concentrations with forces in the range from 10^{-14} to 10^{-11} N. These results have been completely understood from the theoretical point of view thanks to the works of Marko and Siggia [71, 72]. In Fig.1.20, force-extension response of DNA is shown and compared to two elastic behaviour models, studied in the next Section. We anticipate that the DNA mechanism is well reproduced by the WLC model, rather than the FJC model [71, 72]. In the following, the DNA molecule has been pulled with larger forces and an *overstretching* phenomenon has been observed [73]. In particular, a force plateau at around 65 pN has been measured in the force-extension curve, as seen in Fig.1.21. This specific behaviour has been interpreted in terms of a transition similar to the one observed within the Gibbs ensemble in other macromolecules, such as several proteins and polysaccharides. The real molecular origin of this transition has been largely investigated and a debate exists on the DNA conformation after the overstretching: many researchers think that there is a simple mechanical denaturation leading to a transformation of the double-stranded DNA into two single-stranded DNA [74, 75]; however, other researchers have proposed the existence of a new DNA con-

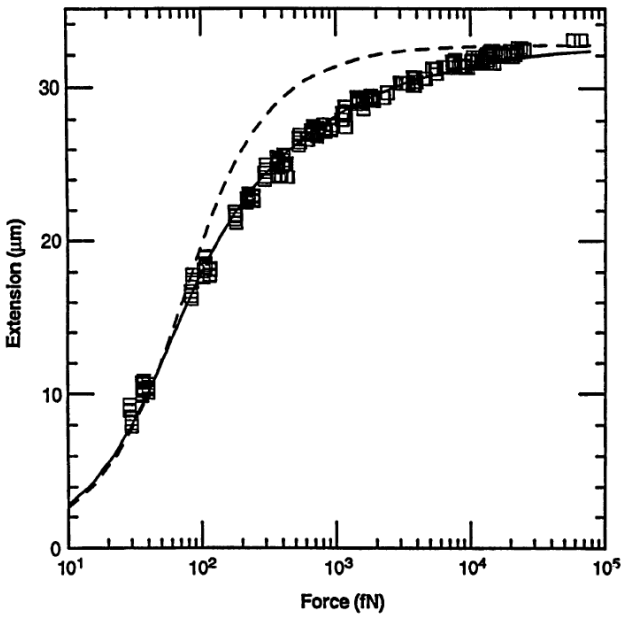

Figure 1.20 – Force-extension response of DNA. Squares are experimental force versus extension data for 97 kb λ-DNA dimers from figure 3 of [70]; solid line is a fit of the entropic force required to extend a worm-like polymer. The fit parameters are the DNA length ($L_C = 32.80 + 0.10\mu$m) and the persistence length ($p = 53.4 \pm 2.3$ nm). Shown for comparison (dashed curve) is the freely jointed chain model [70] with $L_C = 32.7\mu$m and a segment length $b = 100$ nm [72].

formation (called *S-DNA* for stretched DNA) which is an intermediate case between the double-stranded DNA and its denatured structure [76]; finally, other researchers thought that melting or S-DNA was depending on the composition in amino acids of the chain [77]. A discrete worm-like chain model has been implemented to describe the DNA stretching under force and to investigate dsDNA to ssDNA and dsDNA to S-DNA transitions [78].

1.3.2 Freely jointed chain model and worm-like chain model

The most important single-molecule experiment concerns the stretching of double-stranded DNA, in order to measure its elastic features, as previously described. Therefore, to do so, a model is needed to represent the behaviour of the macromolecule, like DNA, protein, etc. Depending on the structure of the macromolecule, several models can be considered. The simplest one is called *freely jointed chain* or *ideal chain*. This model allows a good representation of single-stranded DNA, RNA and other simple polymers. The freely jointed chain model (FJC) schematizes the macromolecule as a sequence of N segments of length b. The parameter b is the so-called *Kuhn length*, i.e. the length of a segment. The *contour length* L_C is the total length of the macromolecule at maximum

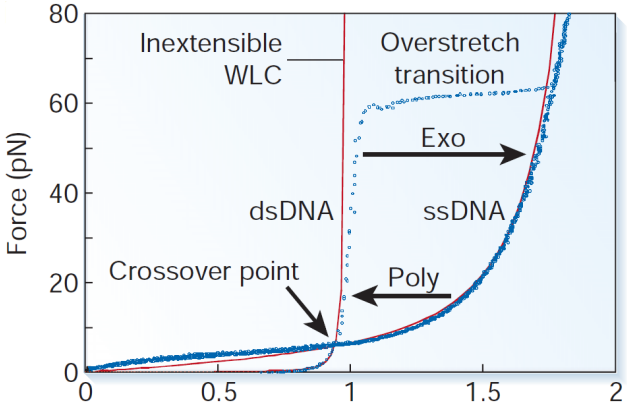

Figure 1.21 – B-DNA overstretching. Force versus extension for dsDNA and ssDNA molecules. Arrows show changes in extension observed at constant tension during polymerisation (Poly) or force-induced exonuclease activity (Exo) [79].

Figure 1.22 – Diagram of comparison between FJC (on the left panel) and WLC model (on the right panel).

physically possible extension. Hence, we obtain the following relation $L_C = Nb$. The FJC model does not take interactions among segments into account (see Fig.1.22, left panel). Moreover, in the FJC model, there is no preferred angle between two segments. This model is basic and simple, and due to its simplicity, it is easier to make exact analytic calculations. The force-extension response for the FJC model is given below:

$$\frac{x}{Nb} = \coth\left(\frac{fb}{k_BT}\right) - \frac{k_BT}{fb} = \mathcal{L}\left(\frac{fb}{k_BT}\right), \tag{1.1}$$

with x the extension of the chain, N the number of units, b the Kuhn length, f the exerted force, k_B the Boltzmann constant, T the temperature, and $\mathcal{L}$ the Langevin function, defined as $\mathcal{L}(x) = \coth(x) - \frac{1}{x}$.

However, the most used model is the *worm-like chain*, which perfectly mimics double-stranded DNA, as shown in Fig.1.20. Indeed, DNA flexibility, introduced in this scheme, has an important role in many biological processes. The worm-like chain model is also used to describe the physics of other polymeric chains, characterised by a given bending rigidity.

Compared to the previous model, an important parameter is added and it concerns the angles between the segments composing the polymer. In the ideal chain model, no forces are necessary to fold the chain. However, in the WLC model, we introduce an energy depending on the angles between adjacent segments (see Fig.1.22, right panel). More specifically, the energy is set to zero if all segments are aligned and therefore a force can be applied to fold or bend the chain. This property is taken into consideration by means of a new parameter called *persistence length*, p, which is defined by the ratio between the mechanical flexibility or bending stiffness and the energy of thermal fluctuations. Hence, it takes into account the balance between enthalpic and entropic contributions [71]. We remark that the bending stiffness can be calculated through the product of the Young modulus, E, multiplied by the moment of inertia, I. Moreover, in the WLC model, the length ℓ is set to zero and the number of segments N tends to infinity, so that the contour length remains constant: $L_C = N\ell$. A discrete version of the WLC has been developed which allows higher forces to be considered and the behaviour of DNA to be better approximated [80–82]. The partition function cannot be calculated exactly either. An example of use of the WLC model for the tenascin protein is shown in Fig.1.23. The behaviour of a chain with the WLC model is the following:

$$f = \frac{k_B T}{p}\left[\frac{1}{4}\left(1 - \frac{x}{L_C}\right)^2 - \frac{1}{4} + \frac{x}{L_C}\right], \tag{1.2}$$

with f the exerted force, k_B the Boltzmann constant, T the temperature, p the persistence length, x the extension of the chain, and L_C the contour length. From the physical point of view, the persistence length p can be defined as the length over which correlations in the direction of the tangent are lost.
Currently, the accepted model for the double-stranded DNA is the twistable worm-like chain model (TWLC), which allows to describe helicoidal double-stranded DNA response under both applied forces and torques [83]. However, this model is not in agreement with experiments realised on DNA with magnetic torque tweezers [84, 85]. Recently, a model was proposed to correctly take account of both bending and torsional stiffness by adding a coupling term between twist and bend deformations [86, 87].

A comparison between the FJC and the WLC to model the behaviour of a polyprotein can be observed in Fig.1.24. The approximation with the FJC model is reasonable, however, the approximation with the WLC is quantitatively better. Nevertheless, in our study, we develop models for chains composed of many units and therefore, we chose to consider simpler models, like FJC, to be able to take account of several extensions, such

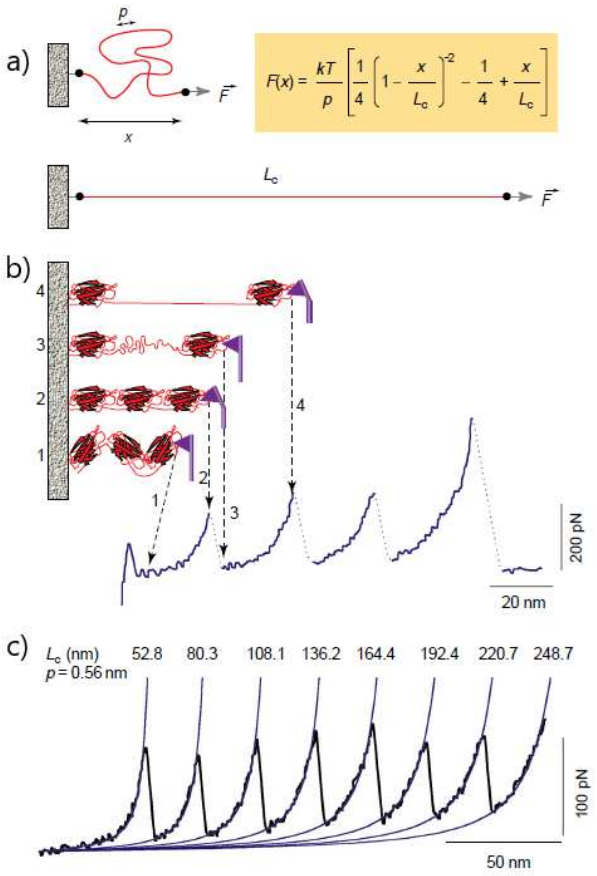

Figure 1.23 – The entropic elasticity of tenascin protein and domain unfolding. a) The entropic elasticity of proteins can be described by the WLC (worm-like chain) equation (inset), which expresses the relationship between force F and extension x of a protein using its persistence length p and its contour length L_C. k is Boltzmann's constant and T is the absolute temperature. b) The saw-tooth pattern of peaks that is observed when force is applied to extend the protein corresponds to sequential unravelling of individual domains of a modular protein. As the distance between substrate and cantilever increases (from state 1 to state 2) the protein elongates, generating a restoring force that bends the cantilever. When a domain unfolds (state 3) the free length of the protein increases, returning the force on the cantilever to near zero. Further extension again results in force on the cantilever (state 4). The last peak represents the final extension of the unfolded protein prior to detachment from the AFM tip. c) Consecutive unfolding peaks of recombinant human tenascin-C were fitted using the WLC model. The contour length (L_C) for each of the fits is shown; the persistence length p was fixed at 0.56 nm [25].

as the interactions among the units and to obtain analytical results. The WLC model allows for considering a single unit, whereas we use the FJC model to represent the whole curve of the force-extension response of the macromolecular chain. Thus, we favour the FJC, a simpler model.

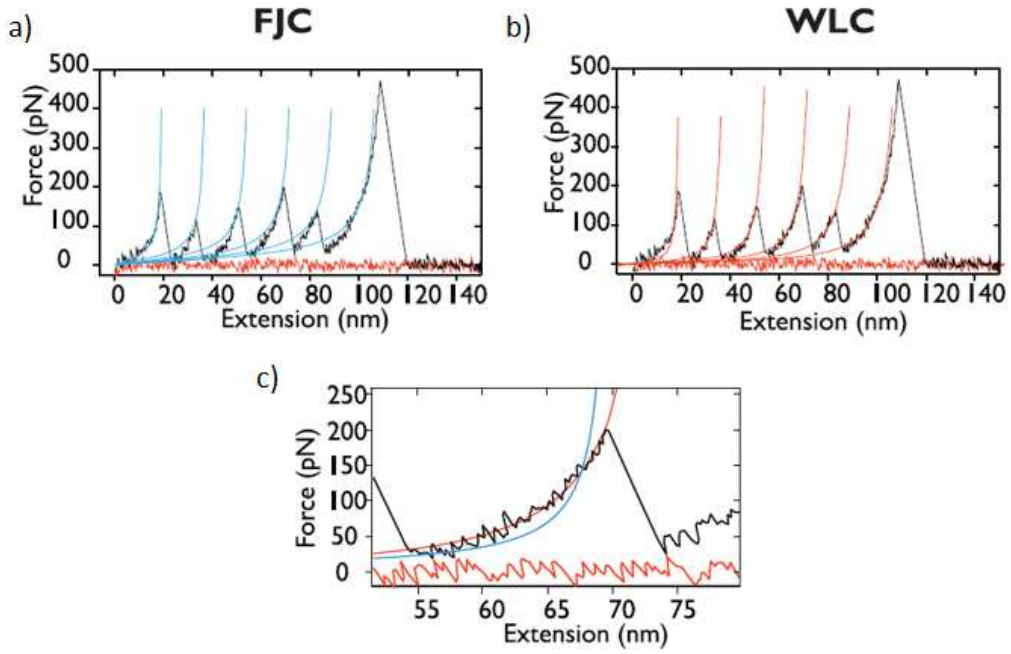

Figure 1.24 – Examples of the use of polymer elasticity models to fit to protein unfolding data from SMFS force versus extension experiments. a) The freely jointed chain (FJC) model is shown, with a Kuhn length of 0.22 nm for the force extension curve of the polyprotein $(pL)_5$. b) The worm like chain (WLC) model is shown, with a persistence length, p, of 0.39 nm for the same data, $(pL)_5$. c) A comparison of the model fit to the data for one protein unfolding event in the polyprotein chain for the FJC (blue) and WLC (red) models with an p of 0.4 nm and b $\approx$ 2p = 0.8 nm [6].

1.4 Proteins

Most of the functions in living beings depends on proteins, also called *polypeptides*, which are biological macromolecules present in all living cells. There are many different types of proteins, representing 50 % of the dry mass of cells and playing a variety of roles for the organism. For example, proteins allow to accelerate chemical reactions, to store amino acids to biosynthesize other proteins, to defence the organism, they also serve for structural support, cell communication and movement. A human body has tens of thousands of proteins. Each of these proteins has its own structure and function. These conformations are among the most complex biological structures and can be studied using SMFS techniques. As an example of complexity, if one observes the Fig.1.25, a pulling speed-dependence of the force peaks can be observed, meaning that the intensity of the force peaks increases when the pulling speed is increased. Such effects will be taken into account further in this manuscript.

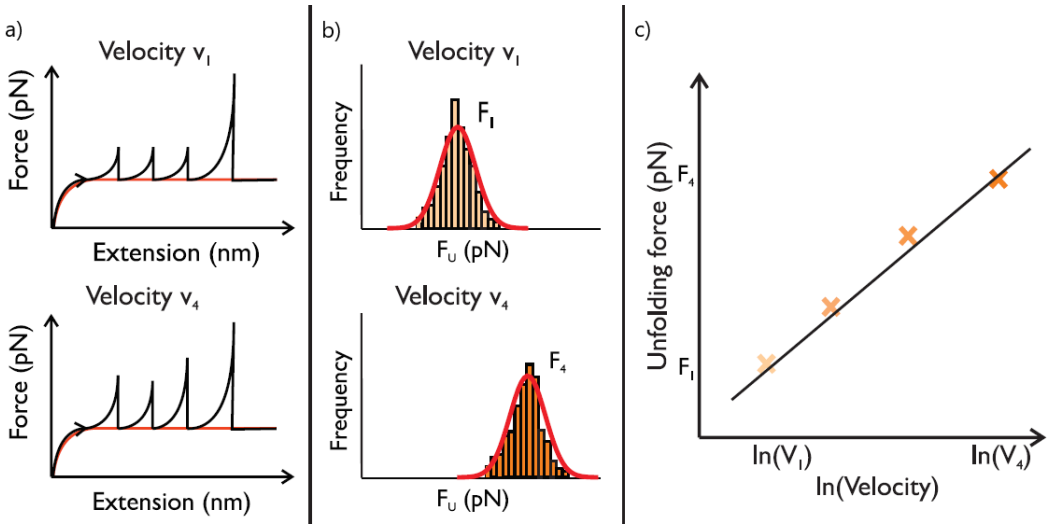

Figure 1.25 – Schematic diagram illustrating the dependence of pulling speed on force peaks. a) SMFS force-extension experiments can be completed at different pulling speeds. b) By measuring the unfolding force of each force peak and creating an unfolding force histogram, the unfolding force F_U can be obtained for a particular pulling velocity. The higher the pulling velocity, the higher the value of F_U. c) Completing force versus extension experiments at a range of pulling velocities allows the unfolding force F_U to be plotted as a function of the natural logarithm of the pulling velocity [6].

To describe proteins, four levels of structure can be used: primary, secondary, tertiary, and quaternary. The basic elements of proteins are amino acids. The order of amino acids in the polypeptide chain represents the *sequence* of the protein and is known as the *primary structure*. These amino acids come from the DNA sequence following the *genetic code*, i.e the instructions to build a specific protein reading a corresponding gene (see Fig.1.26). Each gene's code combines the four bases in various ways to spell out three-letter "words" that specify which amino acid is needed at every step in making a protein.

Although proteins are different from each other, they are all made up of the same twenty amino acids, represented in Fig.1.1. Large polypeptide chains are created by forming peptide bonds between amino and carboxyl groups on two adjacent amino acids. These bonds are covalent. Then comes the secondary structure, whose most common are made of β-sheets, formed by parallel or anti-parallel β-strands or α-helices. The existence of these structures was found in 1951 by the Nobel prize in chemistry 1954, Linus Pauling [89]. Tertiary structure, for its part, refers to the three-dimensional shape of the protein, stabilised by interactions between the side chains, and can be classified into three

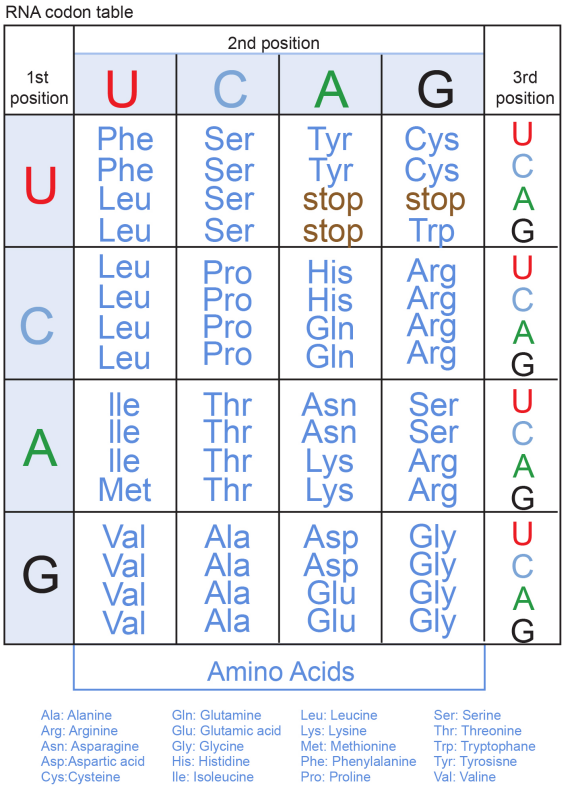

Figure 1.26 – Illustration of the genetic code, with the correspondence between nucleic acids and amino acids [88].

categories: all β-structures, all α-structures and α/β structures. The interactions allowing to held in place the three-dimensional structures are hydrogen bonds, i.e. dipole-dipole interaction between a hydrogen atom and an electronegative atom. Forα-helices, hydrogen bonding is made between the amine and carbonyl groups on the same polypeptide chain. Forβ-sheets, they are made between the amine groups of one polypeptide chain and carbonyl groups on a second adjacent chain. Finally, quaternary structure represents the assembly and interactions between two or more polypeptides. One notices that only the proteins consisting of more than one polypeptide have a quaternary structure. In Fig.1.27, one can see a schematic representation of the four different levels of the protein structure.

Other interactions can take place proteins or bigger macromolecules like hydrophobic and hydrophilic interactions or Van Der Walls interactions (distance-dependent interaction between atoms or molecules). Since most of these interactions are quite weak, their

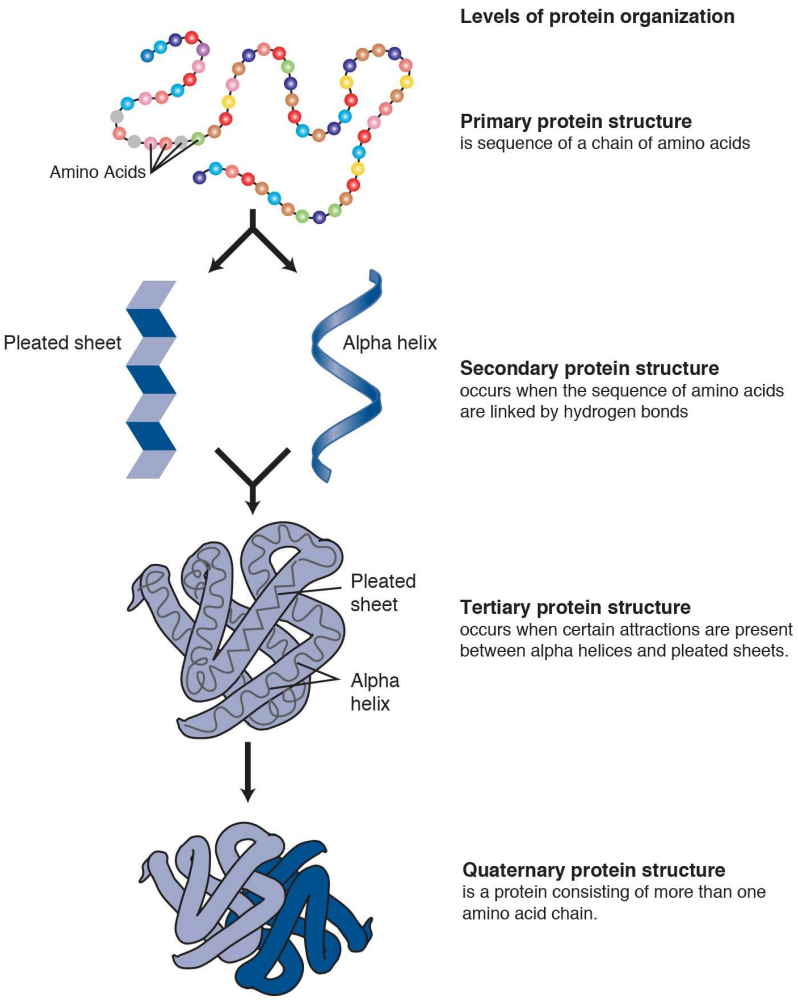

Figure 1.27 – Diagram of the four levels of structure of a protein [90].

energy may be comparable to the thermal energy fluctuations. This point explains the complexity of the proteins folding problem and justify the use of equilibrium and non-equilibrium statistical mechanics.

Once the structure of proteins has been studied, the latter can be classified in three main classes, represented in Fig.1.28, by the properties of their environment and the concerned interactions.

The first class concerns *globular proteins*. These proteins are water soluble, hence they are often studied. As the name suggests, they form a globular shape. This results from an assembly of different secondary structures. These secondary elements form to pro-

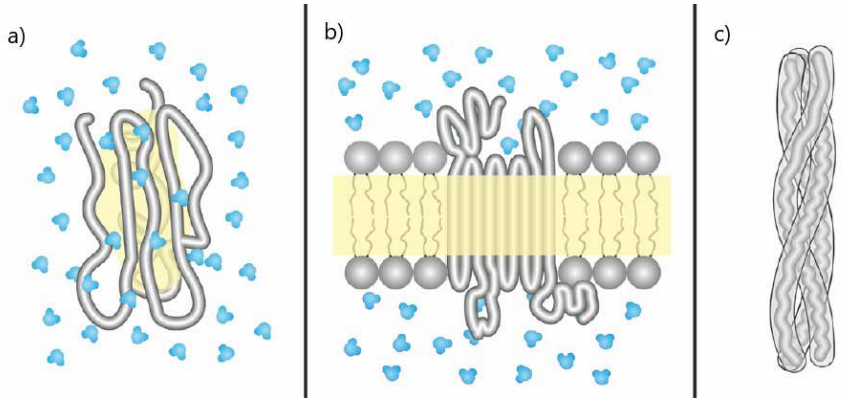

Figure 1.28 – Schematic diagram showing the three classes of proteins. Blue molecules represent water. On the panel a), a globular protein is shown. One can see hydrophobic amino acids in the yellow core of the protein, which is water soluble. On the panel b), a membrane protein traverses through the membrane (composed of lipids and proteins. Hydrophilic regions of the protein are exposed to the water molecules either side of the membrane. On the panel c), one can see an example of fibrous protein, collagen, which is a non-soluble protein. Collagen is a fibril made up of α-helices, formed by three intertwined polyproteins [6].

tect hydrophobic amino acids from contact with water. A well-know globular protein is hemoglobin, whose role is to transport oxygen from the respiratory system (lungs, gills) to the rest of the body. Globular proteins can have a structural role (e.g. actin G in the actin microfilaments). However, in contrast to fibrous proteins whose role is limited to a structural function, globular proteins can behave like enzymes, messengers (to transmit messages to regulate biological processes), stock of amino acids and so on.
The second class are *membrane proteins*. These proteins are in cell membranes and they separate internal components of the cell from external components and environment. They can contain multiple α-helices, which traverse several times through the membrane, or multiple β-sheets, which provide a channel through the membrane, or even just single α-helices allowing to traverse through the membrane once. Membrane proteins have many functions: transport of various elements across membranes, signalling, sensing, and so on.
Finally, the last class of proteins deals with *fibrous proteins*. The latter allow to provide structural support to cells and tissues. Fibrous proteins often associate to form filaments and fibrils. They can be separated into three categories: β-structural (made up of β structures, like silk fibre), α-structural (like superhelix formed by coiled helices) or collagen (non soluble fibrils found in the bone matrix). Fibrous proteins can contain smaller globular proteins, like titin, consisting of combination of regions alternating with both globular and fibrous proteins.

1.5 Structures with bistability

Most proteins can be viewed as composed of multistable units. Linus Pauling and Alfred Mirsky qualified in 1936 the denatured proteins they studied as "by the absence of a uniquely defined configuration" [91]. The modern statistical mechanical picture of protein folding is represented by a funneled energy landscape. In Fig.1.29, a simplified two-dimensional projection of the energy landscape is shown and nevertheless, it looks very complex [92]. To simplify the study of such macromolecules, a two-state energy landscape is often considered. Indeed, many macromolecules of biological origin show a two-state behaviour, like DNA, RNA or proteins. This two-state behaviour is associated to two configurational states, namely folded and unfolded states. Creating mathematical models is very useful to predict the behaviour of such structures. In practical terms, the first energy well is associated to the folded (or native) position, energetically the most stable position. The second potential well represents the unfolded state and is a metastable state.

At the thermodynamic equilibrium, if the system is far from the thermodynamic limit i.e. the number of units is small, different cases (or statistical ensembles) have to be studied. As seen before, differences between isometric and isotensional ensembles can be studied in small systems thanks to statistical mechanics. In the Gibbs ensemble, units are independent from each other. Their folding or unfolding only depends on applied forces or temperature effects. However, in the Helmholtz ensemble, an implicit interaction is created between the units because of the fixed length of the chain (isometric condition). Therefore, the calculation of the Helmholtz partition function is complicated due to implicit interactions leading to bonds between units. The mechanics of bistable chains embedded in a thermal bath is one of the most studied subject in this manuscript. The modelling of bistable systems is not only useful for the understanding of macromolecules, but also relevant for the understanding of the muscle operating principle [94,95]. In this case, bistable mechanical models are able to describe the physical mechanisms of the two passive and active regimes. Again, the study of bistable chains is also important in material science to investigate several plastic phenomena. For instance, a sawtooth-like pattern appears in the force-extension response of several non-biological systems undergoing discrete phase transformations, like ferromagnetic alloys, nano-intended substrates and plastic materials. A typical case concerns nanowires undergoing phase transformations [96–98]. When stretched, these structures present a first force peak higher than the others. This corresponds to an overtstress or a nucleation stress. The nucleation is the

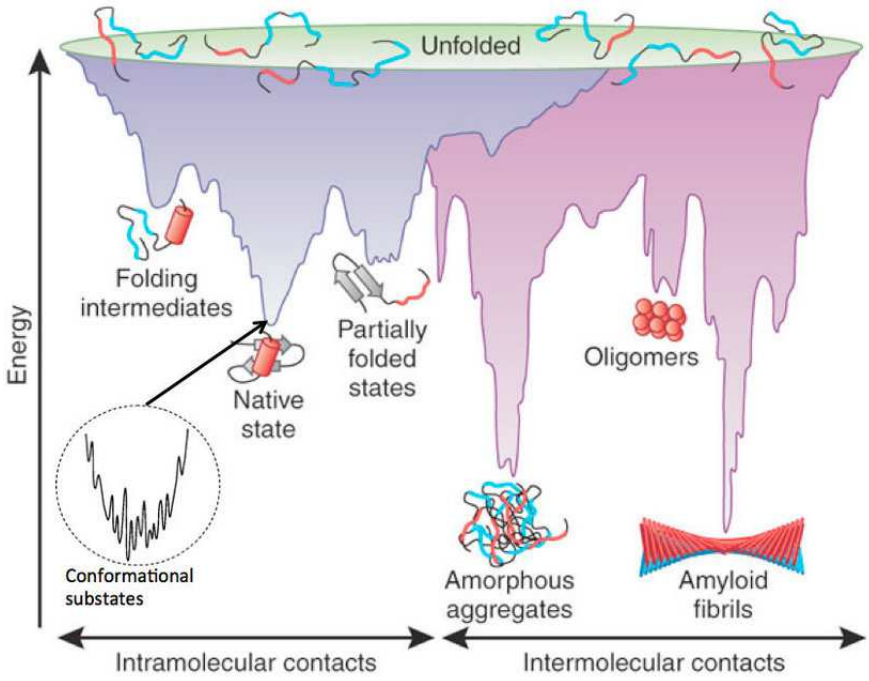

Figure 1.29 – Energy landscape of globular proteins. The protein energy landscape is represented by the free energy of the protein as a function of some reaction coordinates. Transiently unfolded proteins are in a higher state of energy. When exploring their energy landscape proteins transit spontaneously through different folding states, before eventually ending up in their native state, which is the lowest energy state. This phenomenon is known as the folding funnel. Within the minimum of the native state, a multitude of substates, known as conformational substates, are constantly explored [93].

creation of a new crystal structure and therefore of a phase between two interfaces. Using Ising-like models makes a lot of sense as it confers weight to interfaces. As a matter of fact, a parallel can be drawn between the first force peak observed in Fig.1.30 and the first force peak obtaining with bistable chain in the case of a ferromagnetic-like interaction under isometric condition (cf. Chapter 4). This is explained by the cost of creating interfaces. A higher force is required to counteract the stable state in which the system is in, a state in which it would remain if it were not constrained.

In several artificial systems, particular performances have obtained with the bistability, like bistable mechanical metamaterials with a negative Poisson ratio (auxetic media) [99] or systems to control the waves propagation [100]. Models based on chains of bistable units with transitions between two states have been also adopted to model hysteretic behaviours and martensitic transformations in continuum mechanics [101–107], or again energy dissipation with shock reduction [108], like protection in car crash thanks to energy

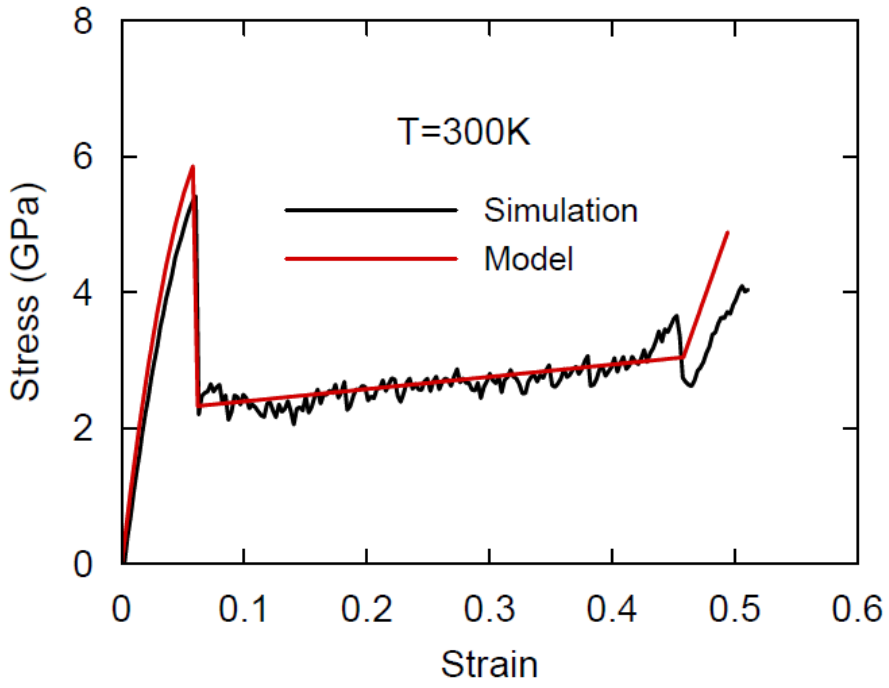

Figure 1.30 – Stress-strain behaviour of a 1.96×1.96 Cu nanowire at the temperature of 300 K. The comparison of the results of MD simulations with the model predictions is shown [97].

absorption by changing the energy state of a unit stepwise, which will reduce the impact of the shock.

1.6 Motivations and goals

To conclude this Introduction, we try here to summarise the motivations and goals of this thesis. The work presented in this manuscript is focused on the study of chains with bistable units by means of the spin variable approach. Some specific problems without a thorough understanding in existing literature are addressed here. These problems are interesting not only for the statistical mechanics theoretical aspects but also for a better comprehension of the physics underlying the force spectroscopy of biological macromolecules and other problems in material science (plasticity and so on, as briefly described above). From the historical point of view, the first biophysical model based on a discrete quantity, similar to a spin variable, has been performed to predict the response of skeletal muscles [109, 110]. This method has also recently been applied to different allosteric systems and macromolecular chains [94, 111, 112]. The spin variable technique has been developed as an equilibrium statistical mechanics methodology, based on the

introduction of a set of discrete variables, to describe the thermo-mechanical properties of a bistable chain under arbitrary external conditions [113]. The aim of this methodology is to obtain for both Gibbs and Helmholtz ensembles the partition functions, central functions of statistical mechanics allowing to obtain values like mean extension or mean force. A two-state freely jointed chain made up of bistable units is typically considered, with a stable folded state and a metastable unfolded state i.e. each unit is in only one of the two possible states (folded or unfolded). Introducing the spin variables permits to consider a simpler energy profile for the units. Indeed, the spin variable indicates in which of the two potential wells the unit is. Consequently, the potential energy, instead of being approximated by a bistable continuous function (with two minima), can be divided into two simpler parabolic functions. This point allows to do rather straightforward calculations. In this thesis, models for force-extension response of bistable chains are developed, permitting to take into account different extensions and generalisations of the existing spin models. These models are developed for both Gibbs and Helmholtz ensembles, as the force-extension response depends on the applied boundary condition for bistable units. We successively introduced intrinsic elasticity, interactions through the Ising model, heterogeneity and dynamics.

First of all, we add intrinsic elasticity to the existing model. In the classical bistable freely jointed model, the wells are considered with a very large stiffness (ideally infinite) in order to simplify the calculations. This assumption is coherent with the classical freely jointed chain where each bond is inextensible. We use in this part of the thesis the spin variables approach to identify the folded or unfolded state of the chain units by taking into account units with a finite elasticity, a property often neglected, while it plays a major role in determining the force-extension response. Indeed, the most of macromolecules of biological origin show an intrinsic elasticity, which can be hardly approximated with inextensible bonds. We obtain the complete solution for a specific case, namely the two-state freely jointed chain with extensible units stretched in the three-dimensional case. The partition function of the isotensional (Gibbs) ensemble is obtained in an exact analytic form by direct integration. Thanks to the generalisation with negative indices of the Hermite polynomials, also the partition function of the isometric (Helmholtz) can be written in an exact form. These original results allow to interpret stretching experiments, operated from the entropic regime up to the unfolding processes, by considering an extensible bistable chain.

Then, it seems obvious to add interactions among the units to the existing models. One of the main motivations to deal with interactions comes from tandem repeat proteins [114].

These proteins are defined as several (at least, two) adjacent copies having the same or very similar sequence motifs, generated by internal duplications. The repetition of the sequence ranges from a single amino-acid up to domains of 100 or more residues. Some tandem repeat proteins are far from each other in the sequence, however they are near in the configurational space and can interact despite the distance. Introducing interactions permits to take account of mutual interactions between units. Hence, the systems considered in this part are bistable and cooperative. It means that each unit can undergo a transition between the folded and unfolded states and the transition of one unit affects the transition of the others. To better understand the behaviour of both bistable and cooperative chains, we chose the simple Ising model. The Ising coefficient, λ, allows to consider several cases, like the case of a "positive" interaction i.e. the unfolding of one unit favours the unfolding of other units, or the case of a "negative" interaction, where the unfolding of unit encourages other units not to unfold. We provide evidence that the cooperativity, measured by the Ising interaction coefficient, strongly modifies the force-extension response of the chain and its configurational properties. Adding interactions makes our model more realistic. This scheme, in addition to represent the behaviour of some proteins, is able to explain the nucleation stress in nanowires, as briefly discussed above.

To bring the model even closer to reality, the heterogeneity of the chain has be to be taken into account. Until now, we considered the case of a bistable chain composed of the same units. The units sequence was homogeneous, at energetic levels for instance. For a homogeneous system (with all the units identical) in a given condition, the probability of unfolding is the same for all units, depending on the extension of the chain. If experiments (real or numerical) are made to find unfolding pathway of a bistable chain made of homogeneous bistable units, different unfolding pathways are observed having the same probability to occur. On the contrary, for a strongly heterogeneous system (units with different length, stiffness, energy level, etc), the probability of unfolding one unit tends to one and to zero for the others and the unfolding pathway can be obtained and identified with a given probability. Introducing heterogeneity of a macromolecular chain is really important to get closer from the behaviour of the real chains. Indeed, when a given unit unfolds, the value of the corresponding spin changes and we are theoretically able to obtain the sequence of transition. The latter will permit to reconstruct the unfolding pathway. As a matter of fact, the ultimate goal is to find thermomechanical properties but also more refined properties like the unfolding pathway, the relative probability of unfolding for a couple of units, etc. From the statistical mechanics point of view, the

heterogeneity can be considered as a quenched disorder introduced within the system. Quenched disorder is a disorder that remains in the system in spite of its evolution and is consequently not associated to the thermal fluctuations. In the Chapter dedicated to the heterogeneity, we investigate the behaviour of a chain made of bistable units with an heterogeneous distribution of energy jumps between the folded and unfolded states, with deterministic and random distributions of the energy jumps in both isotensional and isometric conditions. We obtained the closed form expression of the partition function thanks to the Laplace-Fourier relation between Gibbs and Helmholtz partition functions and the determinant form of the so-called Newton-Girard formulae (see details in Chapter 5). We observed that in the case of homogeneous chains, no unfolding pathway can be identified. On the contrary, we prove that the unfolding pathway is described by the ascending order of the energy jumps between folded and unfolded states in the case of heterogeneous units. The statistical properties of this pathway are studied in detail.
Finally, the transition from equilibrium to non-equilibrium statistical mechanics is investigated in the last Chapter of this manuscript. In previous Chapters, systems of bistable chains were studied at thermodynamic equilibrium. The thermodynamic equilibrium corresponds to a very low velocity (ideally, zero) of increasing extension of the chain i.e. each time the length of the chain increases, the chain is considered to reach a state of equilibrium (quasi-static transformations). In real life, with available devices of force spectroscopy, it is not possible to pull on the molecular chain with velocities low enough to exactly work at thermodynamic equilibrium. Currently, analytic solutions and molecular dynamics simulations are applied to solve the problem. However, analytic solutions permit to work only with very low extension speeds and molecular dynamics, with very high extension speeds, because of computational cost, limiting the total simulation time. Here, the aim is to develop theoretical and numerical models to be able to work with arbitrary extension speeds (of course in a reasonable physically given range). The main motivation comes from the experimental observation of measures of the peak forces depending on the pulling speed (see Fig.1.25). Instantaneous forces play an important role since they deform the energy landscape, modifying the energy jumps and consequently the frequency of folding and unfolding events. Indeed, we observe that the intensity of the peak forces measured within the Helmholtz ensemble are strongly dependent on the extension speed applied to the chain. This point will give the possibility to validate the model against real data of titin, a heavy chain and filamin, a light chain. To do so, we will work with the overdamped Langevin approximation, where the inertial term is neglected, leading up to a first order differential equation instead of a second order equation, classically describing

the motion evolution of the system. Its implementation allows for investigating the effect of the pulling rate and of the device intrinsic elasticity on the chain unfolding response. Our model allows for considering pulling speed ranging from the standard AFM to the HS-AFM. A stochastic Monte Carlo scheme is also used to introduce the bistability in an alternative way.

Chapter 2

Introduction to the thermodynamics of small systems and the spin variable method

2.1 Thermodynamics of small systems

In this Section, we show how the spin variable method can be used to obtain correct general results concerning the thermodynamics of chains with conformational transitions.

2.1.1 Introduction

As previously said, the whole stretching behaviour of macromolecules (composed of entropic, enthalpic, unfolding and over-stretching regimes) can be experimentally studied through SFMS methods. In particular, the unfolding of the units of a chain, governed by the conformational transitions between two (or more) states, is an important process, e.g., observed in polypeptides and nucleic acids. These transitions can modify parameters like the length, the energetic level and the elastic properties of the units of the chains, and they can be taken into account by means of a potential energy exhibiting two (or more) minima corresponding to the stable states or configurations.

Today, many efforts are devoted to the application of theoretical physics methods to the complexity of the biological context [115]. Accordingly, models introducing a specific Hamiltonian for the system and allowing to calculate the pertinent partition function and, eventually, the thermal and mechanical macroscopic properties, are developed for macromolecules of biological origin [116–121]. Classically, as previously introduced, macromolecules without configurational transitions can be studied through the FJC model [120, 122], the WLC model [71, 117, 123] and many other generalised ap-

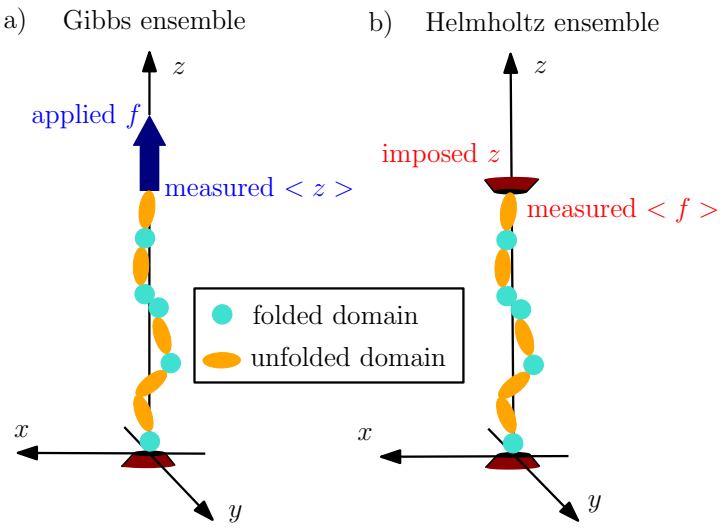

Figure 2.1 – Single-molecule stretching experiments conducted within the Gibbs a) and the Helmholtz b) ensembles. Each domain of the chain may experience a conformational transition between folded (length ℓ) and unfolded (length $\chi\ell$) states [1].

proaches [82, 124–127]. On the contrary, for macromolecules exhibiting folding/unfolding transitions, the boundary condition imposed to the system (see Fig.2.1 for details) has to be considered in the models. Indeed, the isotensional experiments (at constant applied force), typically performed with soft devices, correspond to the Gibbs statistical ensemble (Fig.2.1.a, and the isometric experiments (at prescribed displacement), performed by hard devices, represent a realisation of the Helmholtz statistical ensemble (Fig.2.1.b [128–131]. If the isotensional condition is applied to the system under consideration (Gibbs ensemble), the force-extension response typically presents a plateau-like shape, with a threshold force characterising the synchronised unfolding of all the units of the chain. This behaviour can be interpreted as a cooperative process, which induces the simultaneous transition of the macromolecular domains [73, 77, 79, 132, 133]. Contrarily, the force-extension curve under isometric conditions (Helmholtz ensemble) typically shows a sawtooth-like shape, corresponding to the sequential unfolding of the units in reaction to the increasing extension. This behaviour is therefore interpreted through a non-cooperative process, with a progressive unfolding of the units [24, 25, 134–136].

As explained before, the non-equivalence in the force-extension curves above mentioned (Gibbs versus Helmhotz ensembles) can be measured in the case of a system composed of a small number of units, or monomers, i.e. when the thermodynamic limit is not satisfied. On the other hand, whenever the number of units is very large (ideally, when it tends to infinity), the statistical ensembles become equivalent and they are described by the same constitutive force-extension response [20, 124, 137–139].

In order to get a wider comprehension of this complex scenario, several theoretical approaches have been so far proposed in literature. Concerning the Gibbs ensemble, the plateau-like response observed for double-stranded DNA has been largely investigated with thermodynamic approaches, statistical mechanics and molecular dynamics simulations [68, 74–77, 140]. A similar response has been also observed for long polysaccharides (e.g.,dextran) and modelled through a continuous two-state model [24, 128, 141].

In a like manner, theoretical developments have been adopted to model the sawtooth-like response observed within the Helmholtz ensemble. Theories for titin, RNA hairpins and other macromolecules have been elaborated through Landau-like free energies, first-order phase transition, Langevin equations, and Ising models [142–147]. A more general point of view about two-state systems driven by hard or soft devices can be found in the mechanical literature concerning discrete systems with multi-basin energy landscapes, Fermi-Pasta-Ulam chains of bistable elements, and structures undergoing discrete phase transformations [104, 105, 107, 148].

We present here a general methodology to study the thermodynamics of a system composed of two-state units and subjected to different boundary conditions. The main idea consists in associating to each unit a discrete variable (or spin variable), able to define the state of the unit itself. Such a variable represents a sort of "bit", assuming the values 0 and 1 for the folded and the unfolded states, respectively. The spin variables are considered as dynamic variables, thus belonging to the extended phase space of the system. From the historical point of view, the first biophysical model based on a discrete quantity, similar to a spin variable, has been performed to predict the response of skeletal muscles [109, 110]. This method has been recently applied to different allosteric systems and macromolecular chains as well [94, 111–113]. The introduction of the spin variables allows to strongly facilitate the calculation of the partition functions, preserving at the same time a good accuracy of results [113]. Briefly, an arbitrary potential function composed of two minima can be substituted by two quadratic potentials and the switching between them is governed by the corresponding spin variable. This method works for both Gibbs and Helmholtz ensembles, allowing to draw direct comparisons between isotensional and isometric conditions. We remark that the spin variables approach can be only adopted for systems at thermodynamic equilibrium. As a matter of fact, the quadratic potentials and the associated spin variables are not sufficient to describe the dynamics, since the characteristic relaxation times strongly depend on the energy barriers between the potential wells. This is a well-known result, encoded within the Kramers rate formula, originally formulated to study chemical reactions [16], and recently generalised

for arbitrary systems with nonconvex energy landscapes [149, 150]. Hence, we use in the following the spin variable approach both in Helmholtz and Gibbs ensembles and we show how it allows to obtain well-known general results of the thermodynamics.

2.1.2 Thermodynamics of chains with conformational transitions

We introduce here a general methodology to describe the problem of the conformational transitions in chains undergoing typical folding/unfolding processes. It means that each domain or unit of the chain may be in two states, corresponding to different chemical conformations. The potential energy of the system is therefore constituted of two wells, which can be explored during the dynamics of the system. To simplify the description of the system, considered at thermodynamic equilibrium, it is often convenient to introduce discrete variables or spins, which are able to indicate what basin is explored for each unit of the chain. Consequently, the more general form of the Hamiltonian for such a system can be written as

$$H = \sum_{i=1}^{N} \frac{\vec{p}_i \cdot \vec{p}_i}{2M} + U\left(\vec{r}_1, ..., \vec{r}_N, S_1, ..., S_N\right) - \sum_{i=1}^{N} \mu_i S_i, \tag{2.1}$$

where $S_i \in \{0, 1\}$ are the spin variables $\forall i = 1, ..., N$ introduced in the potential energy U. A domain is said *folded* (i.e. it is in the first energy well) when $S_i = 0$ and *unfolded* (i.e. it is in the second energy well) when $S_i = 1$. The first term represents the kinetic energy ($\vec{p}_i$ are the linear momenta of the units and M is their mass) and the second term represents the potential energy depending on the units positions $\vec{r}_i$ and on the spins S_i. Finally, the third term represents the effect of the external chemical environment on the state of the domains: if $\mu_i > 0$, the unfolded state is favoured, and conversely, if $\mu_i < 0$, the folded state is favoured. From the experimental point of view, the chemical potentials μ_i may mimic the presence of some chemical substances, able to induce or to impede the units unfolding. An interesting example concerns the effect of ethidium molecules on the force-induced melting of DNA [151]. As discussed below in our model, the chemical potential is able to shift the effective energy jump between folded and unfolded states of a given bistable system. This exactly corresponds to the experimental observations in response to an increasing concentration of ethidium, which reported an increase of the DNA transition force and of the effective energy jump between states before and after the transition [151].

It is important to remark that the introduced Hamiltonian function corresponds to a coarse-graining approach, applied to the actual molecular architecture of the system.

Indeed, the concept of units above introduced corresponds to a specific ensemble of atoms with a well-defined behaviour, summarised within the coarse-grained potential energy. Of course, we could have considered an *exact* Hamiltonian with a very large number of degrees of freedom (all the atomic coordinates and momenta as, *e.g.*, introduced in classical molecular dynamics), but we propose a simplification to provide analytic developments and corresponding physical interpretations. Therefore, when we mention a physical property of one unit (elastic constant or others), we have to keep in mind that it represents the collective response of the atoms of the unit, and not a real or intrinsic property of an *exact* Hamiltonian. Then, these properties of units may be considered temperature-dependent being defined at the mesoscopic scale. This is similar to the expansion of the Landau free energy used to describe critical phenomena, where the coefficients depend on temperature for the same reason. The coarse-graining approach also leads to the following issue in writing Eq.(2.1): since the extremities of a given unit $\vec{r}_{i-1}$ and $\vec{r}_i$ are associated to the momenta $\vec{p}_{i-1}$ and $\vec{p}_i$, it means that we assume the mass concentrated at points $\vec{r}_i$, while in the real case the mass is distributed between $\vec{r}_{i-1}$ and $\vec{r}_i$ on all the atoms composing the unit. This is a typical approximation used to describe a one-dimensional chain [152], but it should be kept in mind since each unit is finally identified by the two positions of extremities and by one spin variable.

As an example, a concrete case can be described by the following potential energy

$$\begin{aligned} U &= -\sum_{i=1}^{N-1} \lambda_i (2S_i - 1)(2S_{i+1} - 1) + \sum_{i=1}^{N} \frac{1}{2} k(S_i) \left[\|\vec{r}_i - \vec{r}_{i-1}\| - \ell(S_i) \right]^2 \\ &\quad + \sum_{i=1}^{N} v(S_i) + \sum_{i=1}^{N-1} \frac{1}{2} \kappa(S_i) \left(\vec{t}_{i+1} - \vec{t}_i \right)^2, \end{aligned} \tag{2.2}$$

where the first term represents an Ising interaction among the states of the units (λ_i being the interaction coefficients and $2S_i - 1 = \pm 1$ if $S_i = 0, 1$), the second and third terms represent the spring-like behaviour of the two-state domains (k, ℓ and v are the elastic constant, the equilibrium length and the basal energy, respectively, of the states), and the last term represents a discrete worm-like-chain interaction among the units (κ is the bending constant related to the persistence length $p = \frac{\ell\kappa}{k_B T}$ and $\vec{t}_i = (\vec{r}_i - \vec{r}_{i-1})/\|\vec{r}_i - \vec{r}_{i-1}\|$). The Ising scheme represents an interaction among the spin variables, with two adjacent units favoured to be in the same state if $\lambda_i > 0$ or favoured to be in opposite states if $\lambda_i < 0$. The typical case arising in protein structures corresponds to a positive interaction and allows the interpretation of cooperativity of the folding/unfolding process [153]. An interesting example of cooperativity in forced unfolding concerns the case of

tandem repeats proteins in red cell spectrin [114]. AFM-measured force-extension curves of thousands of constructs with few repeats of spectrin have shown a significant statistic of tandem repeat unfolding processes. It means that two adjacent units unfold at the same time (cooperatively), thus generating a single force peak in the observed response [114]. The mechanism at the base of this transition lies in the strongly interconnected repeats, which unfold in synchrony. This is the typical phenomenon modelled by an Ising model, as introduced in Eq.(2.2). Also, a similar observation has been made on Filamin A, where domain-domain interactions lead to a hierarchy of unfolding forces that may be studied by an Ising scheme [154]. An interesting system for studying interactions and cooperativity is given by the ankyrin repeat proteins. They present, contrarily to globular proteins, very few long-range contacts and therefore only local interactions are present [155]. However, these structures show a high degree of folding cooperativity, and are therefore very stimulating to analyse the folding/unfolding process. In this system, it is relatively easy to add or remove units, thus generating ideal models to compare experiments and theories [155]. We also note that the interactions among domains may be affected by chemical substances, which finally modify the folding pathway. As an example, we can cite the effects of chaperones on maltose binding proteins (MBP), recently observed by optical tweezers measurements [156, 157]. From the theoretical point of view, it means that the Ising interaction coefficients may depend on the external chemical conditions. We will analyse the Ising interactions in Chapter 4.

Concerning the WLC scheme, we recall that it represents a geometric interaction, which provides an elastic flexibility to a chain of objects, trying to align them along the same direction. As an example, in Eq.(2.2), this form of interaction has been introduced in a discrete form among the units of the chain. It means that the situation with aligned units is energetically favoured. Nevertheless, it is interesting to note that in most of protein structures, the WLC scheme can be used to model the force-extension behaviour of each unit. It means that it should be implemented at the level of the *exact* Hamiltonian (at atomic and/or molecular scale), above introduced. It is the basic model chosen to interpret the force-extension curves of force spectroscopy experiments, including the ones with folding/unfolding transitions [24, 128, 135]. In this thesis, the WLC energy term will be often neglected in order to obtain closed form solutions describing different types of Hamiltonian functions.

The folding/unfolding of the units, here described by the spin variables, is controlled by the mechanical boundary conditions applied to the chain and by the chemical potentials of the external environment. We consider here a chain with the first extremity tethered at

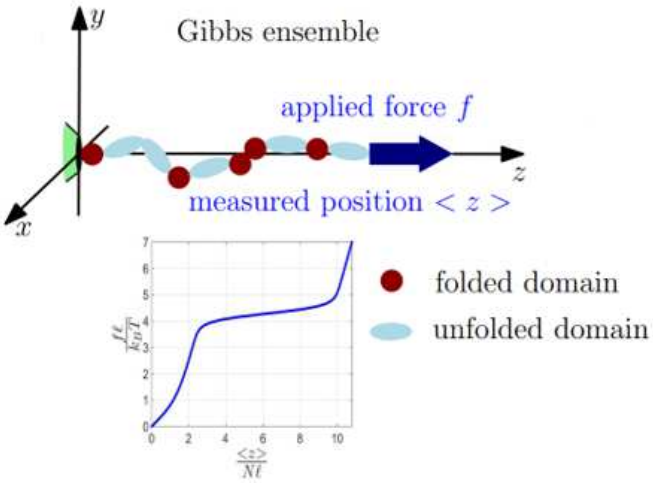

Figure 2.2 – Schematic representation of a stretching experiment within the Gibbs ensemble, with the typical force-extension response showing a force plateau.

the origin of the reference frame. In general, an arbitrary chain described by Eq.(2.1) can be studied within the Helmholtz ensemble or within the Gibbs ensemble of the statistical mechanics. In the first situation, we fix the second extremity at a given point of the space while, in the second situation, we apply an external force to the last unit. In the following, we briefly describe the statistical mechanics of both ensembles, eventually verifying that the proposed approach yields results, which are coherent with standard thermodynamics.

Within the Helmholtz ensemble we have a specific Hamiltonian given by

$$H_H = H(\vec{q}, \vec{p}, \vec{S}; \vec{r}, \vec{\mu}), \tag{2.3}$$

where we introduced $\vec{q} = (\vec{r}_1, ..., \vec{r}_{N-1})$ (we fixed $\vec{r}_N = \vec{r}$), $\vec{p} = (\vec{p}_1, ..., \vec{p}_{N-1})$ (we fixed $\vec{p}_N = 0$), $\vec{S} = (S_1, ..., S_N)$, and $\vec{\mu} = (\mu_1, ..., \mu_N)$ (see Fig.2.2). In this scheme, $\vec{r}$ and $\vec{\mu}$ are externally controlled variables and $\vec{q}, \vec{p}$ and $\vec{S}$ are the dynamical variables of the phase space. Therefore, the density probability of the canonical ensemble is given by [120, 158]

$$\rho_H(\vec{q}, \vec{p}, \vec{S}) = \frac{1}{Z_H(\vec{r}, \vec{\mu}, T)} \exp\left[-\frac{H_H(\vec{q}, \vec{p}, \vec{S}; \vec{r}, \vec{\mu})}{k_B T}\right], \tag{2.4}$$

where the Helmholtz partition function reads

$$Z_H(\vec{r}, \vec{\mu}, T) = \sum_{\vec{S} \in \{0,1\}^N} \int_{\mathcal{Q}} \int_{\mathcal{Q}} \exp\left[-\frac{H_H(\vec{q}, \vec{p}, \vec{S}; \vec{r}, \vec{\mu})}{k_B T}\right] \mathrm{d}\vec{q}\mathrm{d}\vec{p}, \tag{2.5}$$

and where $\mathcal{Q} = \mathbb{R}^{3(N-1)}$. By means of this probability density, one can determine the average value of any macroscopic variable. This is the reason why calculating the partition

function is of vital importance in the statistical mechanics. In particular, the average value of the force $\vec{f} = \frac{\partial H_H}{\partial \vec{r}}$ needed to fix $\vec{r}_N = \vec{r}$ and the average value of each spin variable S_i can be easily obtained through the introduction of the Helmholtz free energy $\mathcal{F} = -k_B T \log Z_H$. Indeed, straightforward calculations yield $\left\langle \vec{f} \right\rangle = \frac{\partial \mathcal{F}}{\partial \vec{r}}$ and $\left\langle \vec{S} \right\rangle = -\frac{\partial \mathcal{F}}{\partial \vec{\mu}}$, which represent two macroscopic or thermodynamic relations. Moreover, the first and second principles for quasi-static transformations can be derived as follows from previous statements. First of all, we define a quasi-static transformation through the time evolution of $\vec{r}(t)$, $\vec{\mu}(t)$ and $T(t)$, which are considered as externally controlled "slow-varying" variables. Under these hypotheses, we can assume that the canonical distribution in Eq.(2.4) remains valid also in this "weak" out-of-equilibrium regime. Therefore, we can evaluate the rate of change of the internal energy $\mathcal{U}$, defined as the average value of the Helmholtz Hamiltonian

$$\begin{aligned}
\frac{\mathrm{d}\mathcal{U}}{\mathrm{d}t} &= \sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} \frac{\mathrm{d}}{\mathrm{d}t} \left(H_H \rho_H \right) \mathrm{d}\vec{q}\mathrm{d}\vec{p} \\
&= \sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} \left(\frac{\partial H_H}{\partial \vec{\mu}} \cdot \frac{\mathrm{d}\vec{\mu}}{\mathrm{d}t} + \frac{\partial H_H}{\partial \vec{r}} \cdot \frac{\mathrm{d}\vec{r}}{\mathrm{d}t} \right) \rho_H \mathrm{d}\vec{q}\mathrm{d}\vec{p} + \sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} H_H \frac{\mathrm{d}\rho_H}{\mathrm{d}t} \mathrm{d}\vec{q}\mathrm{d}\vec{p} \\
&= -\left\langle \vec{S} \right\rangle \cdot \frac{\mathrm{d}\vec{\mu}}{\mathrm{d}t} + \left\langle \vec{f} \right\rangle \cdot \frac{\mathrm{d}\vec{r}}{\mathrm{d}t} + \sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} H_H \frac{\mathrm{d}\rho_H}{\mathrm{d}t} \mathrm{d}\vec{q}\mathrm{d}\vec{p}, \qquad (2.6)
\end{aligned}$$

where all the sums on $\vec{S}$ are performed over $\{0,1\}^N$. The first two terms represent the chemical and the mechanical work per unit of time, respectively, done on the system, while the third one describes the rate of heat entering the system. The latter can be further elaborated by defining $\eta = \frac{\mathcal{F} - H_H}{k_B T}$ and $\rho_H = e^{\eta}$. Since $\sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} e^{\eta} \mathrm{d}\vec{q}\mathrm{d}\vec{p} = 1$, the derivative with respect to the time gives $\sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} e^{\eta} \dot{\eta} \mathrm{d}\vec{q}\mathrm{d}\vec{p} = 0$, and since $\mathcal{F}$ is independent of $\vec{S}, \vec{q}$ and $\vec{p}$, we can also write that $\sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} \mathcal{F} e^{\eta} \dot{\eta} \mathrm{d}\vec{q}\mathrm{d}\vec{p} = 0$. The last term in Eq.(2.6) can be therefore rewritten as

$$\begin{aligned}
\sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} H_H \frac{\mathrm{d}\rho_H}{\mathrm{d}t} \mathrm{d}\vec{q}\mathrm{d}\vec{p} &= \sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} H_H e^{\eta} \dot{\eta} \mathrm{d}\vec{q}\mathrm{d}\vec{p} \\
&= \sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} \left(H_H - \mathcal{F} \right) e^{\eta} \dot{\eta} \mathrm{d}\vec{q}\mathrm{d}\vec{p}. \qquad (2.7)
\end{aligned}$$

Now, since $H_H - \mathcal{F} = -k_B T \eta$, we also have

$$\sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} H_H \frac{\mathrm{d}\rho_H}{\mathrm{d}t} \mathrm{d}\vec{q}\mathrm{d}\vec{p} = -k_B T \sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} e^{\eta} \eta \dot{\eta} \mathrm{d}\vec{q}\mathrm{d}\vec{p}. \qquad (2.8)$$

This expression can be easily interpreted by considering the average value of the variable η given by $\langle\eta\rangle = \sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} e^{\eta}\eta \mathrm{d}\vec{q}\mathrm{d}\vec{p}$ and its time derivatives

$$\frac{\mathrm{d}\langle\eta\rangle}{\mathrm{d}t} = \sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} e^{\eta}\dot{\eta} \mathrm{d}\vec{q}\mathrm{d}\vec{p} + \sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} e^{\eta}\eta\dot{\eta} \mathrm{d}\vec{q}\mathrm{d}\vec{p}. \tag{2.9}$$

Since the first integral is zero for previous calculations, we eventually obtain

$$\sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} H_H \frac{\mathrm{d}\rho_H}{\mathrm{d}t} \mathrm{d}\vec{q}\mathrm{d}\vec{p} \;=\; -k_B T \frac{\mathrm{d}\langle\eta\rangle}{\mathrm{d}t} = T\frac{\mathrm{d}\mathcal{S}}{\mathrm{d}t}, \tag{2.10}$$

where we introduced the entropy of the system as

$$\mathcal{S} = -k_B \langle \log \rho_H \rangle = -k_B \sum_{\vec{S}} \int_{\mathcal{Q}} \int_{\mathcal{Q}} \rho_H \log \rho_H \mathrm{d}\vec{q}\mathrm{d}\vec{p}. \tag{2.11}$$

The first and the second principles of the thermodynamics for quasi-static transformations can be finally summed up through the balance equation

$$\frac{\mathrm{d}\mathcal{U}}{\mathrm{d}t} = -\left\langle \vec{S} \right\rangle \cdot \frac{\mathrm{d}\vec{\mu}}{\mathrm{d}t} + \left\langle \vec{f} \right\rangle \cdot \frac{\mathrm{d}\vec{r}}{\mathrm{d}t} + T\frac{\mathrm{d}\mathcal{S}}{\mathrm{d}t}. \tag{2.12}$$

Hence, we proved that the macroscopic thermodynamics is perfectly coherent with the approach based on the spin variables, introduced to analyse systems with conformational transitions. In addition, the average value of the relation $H_H - \mathcal{F} = -k_B T\eta$ yields $\mathcal{F} = \mathcal{U} - T\mathcal{S}$, which corresponds to the macroscopic definition of Helmholtz free energy. By differentiating with respect to the time, we also have

$$\frac{\mathrm{d}\mathcal{F}}{\mathrm{d}t} = -\left\langle \vec{S} \right\rangle \cdot \frac{\mathrm{d}\vec{\mu}}{\mathrm{d}t} + \left\langle \vec{f} \right\rangle \cdot \frac{\mathrm{d}\vec{r}}{\mathrm{d}t} - \mathcal{S}\frac{\mathrm{d}T}{\mathrm{d}t}, \tag{2.13}$$

from which we can deduce the two already introduced equations $\left\langle \vec{f} \right\rangle = \frac{\partial \mathcal{F}}{\partial \vec{r}}$ and $\left\langle \vec{S} \right\rangle = -\frac{\partial \mathcal{F}}{\partial \vec{\mu}}$ and the important relation $\mathcal{S} = -\frac{\partial \mathcal{F}}{\partial T}$, giving the entropy in terms of the Helmholtz free energy.

Concerning the Gibbs ensemble, we can introduce the following extended Hamiltonian

$$H_G = H(\vec{q}, \vec{p}, \vec{S}; \vec{\mu}) - \vec{f} \cdot \vec{r}_N, \tag{2.14}$$

where the second term represents the energy associated to the external force applied to the last unit of the chain (see Fig.2.3). Here, we introduced $\vec{q} = (\vec{r}_1, ..., \vec{r}_N)$, $\vec{p} = (\vec{p}_1, ..., \vec{p}_N)$, $\vec{S} = (S_1, ..., S_N)$, and $\vec{\mu} = (\mu_1, ..., \mu_N)$. In this scheme, $\vec{f}$ and $\vec{\mu}$ are externally controlled

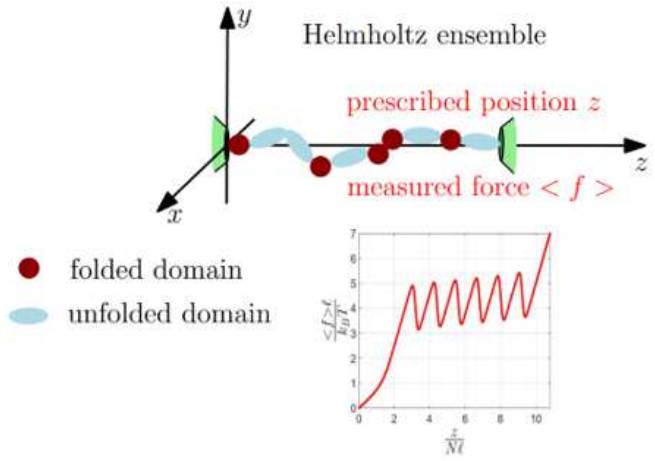

Figure 2.3 – Schematic representation of a stretching experiment within the Helmholtz ensemble, with the typical force-extension response showing a sawtooth-like behaviour.

variables and $\vec{q}, \vec{p}$ and $\vec{S}$ are the dynamical variables of the phase space. Therefore, the density probability of the canonical ensemble is given by [120, 158]

$$\rho_G(\vec{q}, \vec{p}, \vec{S}) = \frac{1}{Z_G(\vec{f}, \vec{\mu}, T)} \exp\left[-\frac{H_G(\vec{q}, \vec{p}, \vec{S}; \vec{f}, \vec{\mu})}{k_B T}\right], \tag{2.15}$$

where the Gibbs partition function reads

$$Z_G(\vec{f}, \vec{\mu}, T) = \sum_{\vec{S} \in \{0,1\}^N} \int_{\mathcal{P}} \int_{\mathcal{P}} \exp\left[-\frac{H_G(\vec{q}, \vec{p}, \vec{S}; \vec{f}, \vec{\mu})}{k_B T}\right] \mathrm{d}\vec{q}\mathrm{d}\vec{p}, \tag{2.16}$$

and where $\mathcal{P} = \mathbb{R}^{3N}$. As before, the most important expected values can be evaluated through the introduction of the Gibbs free energy $\mathcal{G} = -k_B T \log Z_G$. Simple calculations yield indeed $\langle \vec{r} \rangle = -\frac{\partial \mathcal{G}}{\partial \vec{f}}$ (where $\vec{r} = \vec{r}_N$) and $\left\langle \vec{S} \right\rangle = -\frac{\partial \mathcal{G}}{\partial \vec{\mu}}$, which represent two classical thermodynamic relations. Also in this isotensional case, we can introduce a quasi-static transformation through the time evolution of $\vec{f}(t)$, $\vec{\mu}(t)$ and $T(t)$, which are, as before, externally controlled "slow-varying" variables. We can assume that the canonical distribution in Eq.(2.15) remains valid and we can evaluate the rate of change of the average value of the Gibbs Hamiltonian, which is the so-called enthalpy $\mathcal{H}$

$$\begin{aligned}
\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} &= \sum_{\vec{S}} \int_{\mathcal{P}} \int_{\mathcal{P}} \frac{\mathrm{d}}{\mathrm{d}t} \left(H_G \rho_G\right) \mathrm{d}\vec{q}\mathrm{d}\vec{p} \\
&= \sum_{\vec{S}} \int_{\mathcal{P}} \int_{\mathcal{P}} \left(\frac{\partial H_G}{\partial \vec{\mu}} \cdot \frac{\mathrm{d}\vec{\mu}}{\mathrm{d}t} + \frac{\partial H_G}{\partial \vec{f}} \cdot \frac{\mathrm{d}\vec{f}}{\mathrm{d}t}\right) \rho_G \mathrm{d}\vec{q}\mathrm{d}\vec{p} + \sum_{\vec{S}} \int_{\mathcal{P}} \int_{\mathcal{P}} H_G \frac{\mathrm{d}\rho_G}{\mathrm{d}t} \mathrm{d}\vec{q}\mathrm{d}\vec{p} \\
&= -\left\langle \vec{S} \right\rangle \cdot \frac{\mathrm{d}\vec{\mu}}{\mathrm{d}t} - \langle \vec{r} \rangle \cdot \frac{\mathrm{d}\vec{f}}{\mathrm{d}t} + \sum_{\vec{S}} \int_{\mathcal{P}} \int_{\mathcal{P}} H_G \frac{\mathrm{d}\rho_G}{\mathrm{d}t} \mathrm{d}\vec{q}\mathrm{d}\vec{p}.
\end{aligned} \tag{2.17}$$

While the first two terms represent the power (chemical and mechanical) supplied to the system, the third term is the rate of heat transferred to the system. A calculation similar to the one developed for the Helmholtz ensemble allows us to rewrite this last term as follows

$$\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} = -\left\langle \vec{S} \right\rangle \cdot \frac{\mathrm{d}\vec{\mu}}{\mathrm{d}t} - \langle \vec{r} \rangle \cdot \frac{\mathrm{d}\vec{f}}{\mathrm{d}t} + T\frac{\mathrm{d}\mathcal{S}}{\mathrm{d}t}, \tag{2.18}$$

where the entropy for the Gibbs ensemble is given by

$$\mathcal{S} = -k_B \left\langle \log \rho_G \right\rangle = -k_B \sum_{\vec{S}} \int_{\mathcal{P}} \int_{\mathcal{P}} \rho_G \log \rho_G \mathrm{d}\vec{q}\mathrm{d}\vec{p}. \tag{2.19}$$

We can straightforwardly prove the thermodynamic relation $\mathcal{G} = \mathcal{H} - T\mathcal{S}$ and we obtain the rate of change of the Gibbs free energy as

$$\frac{\mathrm{d}\mathcal{G}}{\mathrm{d}t} = -\left\langle \vec{S} \right\rangle \cdot \frac{\mathrm{d}\vec{\mu}}{\mathrm{d}t} - \langle \vec{r} \rangle \cdot \frac{\mathrm{d}\vec{f}}{\mathrm{d}t} - \mathcal{S}\frac{\mathrm{d}T}{\mathrm{d}t}. \tag{2.20}$$

The last energy balance immediately delivers the relations $\langle \vec{r} \rangle = -\frac{\partial \mathcal{G}}{\partial \vec{f}}$ and $\left\langle \vec{S} \right\rangle = -\frac{\partial \mathcal{G}}{\partial \vec{\mu}}$, previously demonstrated, and the result $\mathcal{S} = -\frac{\partial \mathcal{G}}{\partial T}$, corresponding to the macroscopic definition of entropy.

The two schemes here outlined represent a complete description able to deliver thermo-mechanical response of a macromolecular chain undergoing conformational transitions under isometric or isotensional boundary conditions. We proved that the introduction of additional "spin" variables to describe the folding or unfolding of the units leaves unaltered the general results of the thermodynamics, thus being well-grounded and promising for applications. As an example, the application of this methodology to the bistable freely jointed chain can be found in the recent literature [113]. It is important to remark that the thermo-elastic response may be different for the two introduced ensembles if we consider chains composed of a small number N of units [120]. On the other hand, when the thermodynamic limit is attained (ideally for $N \to \infty$), the two ensembles become equivalent and they exhibit the same physical response [20, 124, 137–139]. This equivalence property is valid for non-branched single chains without confinements, i.e. freely fluctuating in the whole space [20]. Therefore, in our analysis, the only constraints consist in the punctual boundary conditions defining the Helmholtz and the Gibbs ensembles. However, it is noteworthy to mention that some particular cases on non-equivalence between dual canonical ensembles have been recently discussed for confined polymer chains [21–23].

2.2 Applications of the spin variable method

In the Section 2.1, we presented some general results of thermodynamics obtained with the help of the spin variable method. Here, the spin variable approach is described in detail, for specific one-dimensional and three-dimensional systems. The study of these two cases has already been published in Ref. [113]. They are here presented to introduce the spin variable method in the simplest cases, without any of the extensions which will be proposed in the following, like elasticity or again Ising interactions. Moreover, we show an example to prove that the spin variables method is not also useful to obtain average values of positions or forces but also to determine the complete probability density describing systems with bistability.

2.2.1 One dimensional system

In this Section, we present the spin variable approach applied to a system that can be analytically solved for both the Gibbs (isotensional) ensemble and the Helmholtz (isometric) ensemble. The aim is to obtain closed form expressions, useful to better understand the physics of the phenomena underlying the bistability in complex systems.

A one-dimensional system composed of N units with mechanical bistability is considered here. This means that each unit of the chain can be described by a symmetric potential energy function $U(x)$ exhibiting four minima (equilibrium points) at $x = \pm\ell$ and $x = \pm\chi\ell$, where χ represents the elongation ratio between the unfolded and folded configurations (see Fig.2.4). Moreover, the potential energy assumes the minimum value $U = 0$ for $x = \pm\ell$ and the minimum value $U = \Delta E$ for $x = \pm\chi\ell$. In order to simplify analysis and calculations, instead of considering the complex continuous potential function represented in Fig.2.4 (blue dashed line), we introduce an additional discrete variable y, which behaves as a spin. This variable belongs to the phase space of the system and, therefore, is a standard variable of the statistical mechanics. The variable y assumes the values in the set $\mathbf{S} = \{\pm 1, \pm\chi\}$ and allows to identify what basin the system is exploring. So, the multimodal energy function is substituted with this simpler expression

$$U(x, y) = v(y) + \frac{1}{2}k_0(y)(x - y\ell)^2, \tag{2.21}$$

where we consider $v(\pm 1) = 0$, $v(\pm\chi) = \Delta E$, $k_0(\pm 1) = k$ and $k_0(\pm\chi) = h$. By varying the value of the spin variable in $\mathbf{S}$, the potential energy in Eq.(2.21) provides the four parabolic wells represented in Fig.2.4 (red solid lines). The latter permit to fully describe the behaviour of the system represented by the multimodal energy profile when we assume

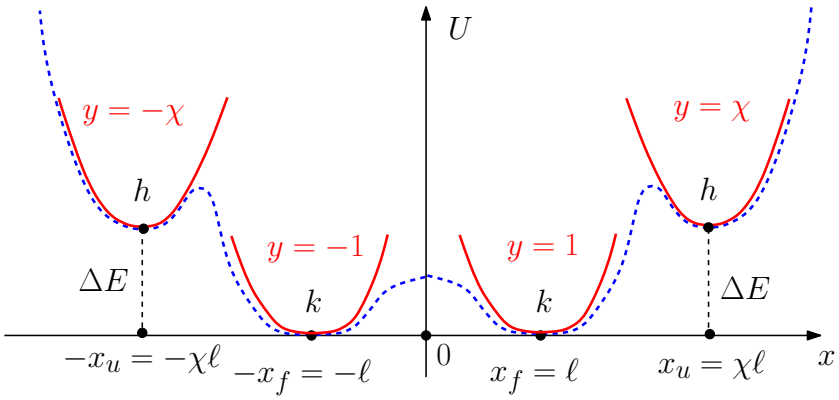

Figure 2.4 – Multistable symmetric potential energy of a single domain (blue dashed line) and its approximation by means of four parabolic profiles (red solid lines) [113].

to work at thermodynamic equilibrium. Similar assumptions have been introduced and analysed in recent literature [105, 150].

Here come some comments on the energy profile proposed in Fig.2.4. We introduced four potential wells (two folded and two unfolded) for the following reasons. Since we study a one-dimensional system, the consideration of positive ($x > 0$) and negative ($x < 0$) orientations of the elements allows to model the four constitutives regimes (entropic, enthalpic, unfolding and over-stretching). Indeed, if a system whose the first end-terminal is tethered at $x = 0$ and the second end-terminal free to fluctuate (without externally applied forces) is considered, we obtain that the mean value of the position of the second end-terminal is zero because of the random orientation of the domains (almost domains folded, 50% towards $x > 0$ and 50% towards $x < 0$). Incidentally, it means that the force-extension curve passes through the origin of the axes. If a weak applied force (positive or negative) is applied, the chain is in the entropic regime and the number of domains oriented in the direction of the force slightly increases. In these conditions, the mechanical reaction of the system is governed by the stochastic distribution of the differently oriented domains (entropy) and not by the actual spring-like behaviour of the elements. Then, if a larger force is applied, we act on the real elasticity of the domains and therefore the enthalpic regime gives way to the unfolding process up to the over-stretching regime, described below for both Gibbs and Helmholtz ensembles.

In conclusion, the entropic behaviour of a system under weak applied force can be represented by the four wells energy profile handled by the spin variable. This aspect is naturally introduced in two- or three-dimensional models by the higher dimensionality, which automatically induces a larger number of degrees of freedom (see Section 2.2.2 for details).

2.2.1.1 The Gibbs ensemble

The total potential energy of the system in the Gibbs ensemble (isotensional condition) can be written as

$$U_{tot}^{G}(\vec{x},\vec{y};f) = \sum_{i=1}^{N} U(x_i - x_{i-1}, y_i) - f x_N, \tag{2.22}$$

where f is the force applied to the last element, $\vec{x} = (x_1, x_2, ..., x_N)$ are continuous variables and $\vec{y} = (y_1, y_2, ..., y_N)$ are discrete spin variables. For this system, we can define the partition function Z_G, as follows

$$Z_G(f) = \sum_{y_1 \in \mathbf{S}} ... \sum_{y_N \in \mathbf{S}} \int_{\Re} ... \int_{\Re} e^{-\frac{U_{tot}^{G}}{k_B T}} \mathrm{d}x_1 ... \mathrm{d}x_N, \tag{2.23}$$

where the variable $\vec{x}$ is integrated whereas $\vec{y}$ is summed. Since we use orthogonal coordinates, the kinetic energy only depends on the linear momentum variables and the potential energy only on $\vec{x}$ and $\vec{y}$. Therefore, the kinetic energy has not been considered because its integral results in a non influential multiplicative constant in the partition function, $Z_G(f)$. Eq.(2.22) can now be substituted in Eq.(2.23). To calculate the integral, we apply the change of variables $x_1 - x_0 = \xi_1$, $x_2 - x_1 = \xi_2$,..., $x_N - x_{N-1} = \xi_N$, from which we get $x_n = \sum_{j=1}^{N} \xi_j$ (with $x_0 = 0$). The change of variables within the multiple integral must be applied by considering that $\mathrm{d}\vec{x} = J\mathrm{d}\vec{\xi}$, where the quantity J is the Jacobian of the transformation, i.e. $J = |\det[\partial\vec{x}/\partial\vec{\xi}]|$, where $[\partial\vec{x}/\partial\vec{\xi}]_{ij}$ is the matrix of the partial derivatives $\partial x_i/\partial\xi_j$. We can easily prove that $J = 1$ for our change of variables and, therefore, we get $\mathrm{d}\vec{\xi} = \mathrm{d}\vec{x}$. Hence, we obtain

$$\begin{aligned} Z_G(f) &= \sum_{\vec{y} \in \mathbf{S}^N} \int_{\Re^N} \exp\left\{ \sum_{j=1}^{N} \left[-\frac{U(\xi_j; y_j)}{k_B T} + \frac{f\xi_j}{k_B T} \right] \right\} \mathrm{d}\vec{\xi} \\ &= \left\{ \sum_{y \in \mathbf{S}} \int_{-\infty}^{+\infty} \exp\left[-\frac{U(\xi; y)}{k_B T} + \frac{f\xi}{k_B T} \right] \mathrm{d}\xi \right\}^N \\ &= \left\{ \sum_{y \in \mathbf{S}} \exp\left[-\frac{v(y)}{k_B T} \right] \mathcal{I}(f) \right\}^N, \end{aligned} \tag{2.24}$$

where the integral $\mathcal{I}(f)$ is defined as

$$\mathcal{I}(f) = \int_{-\infty}^{+\infty} \exp\left[-\frac{k_0(y)}{2k_B T}(\xi - y\ell)^2 + \frac{f\xi}{k_B T} \right] \mathrm{d}\xi. \tag{2.25}$$

This integral can be calculated in an explicit form by means of the well-known expression

$$\int_{-\infty}^{+\infty} e^{-\alpha x^2} e^{\beta x} \mathrm{d}x = \sqrt{\frac{\pi}{\alpha}} e^{\frac{\beta^2}{4\alpha}} \quad (\alpha > 0). \tag{2.26}$$

The final result can be written as follows

$$\mathcal{I}(f) = \sqrt{\frac{2\pi k_B T}{k_0(y)}} \exp\left[\frac{y\ell f}{k_B T} + \frac{f^2}{2k_B T k_0(y)}\right]. \tag{2.27}$$

Coming back to the partition function, we obtain

$$\begin{aligned} Z_G(f) &= (2\pi k_B T)^{\frac{N}{2}} \left\{ \sum_{y \in \mathbf{S}} \sqrt{\frac{1}{k_0(y)}} \exp\left[\frac{1}{k_B T}\left(\frac{f^2}{2k_0(y)} + y\ell f - v(y)\right)\right] \right\}^N \\ &= (8\pi k_B T)^{\frac{N}{2}} \left\{ \sum_{s \in \{1,\chi\}} \sqrt{\frac{1}{k_0(s)}} \exp\left[\frac{1}{k_B T}\left(\frac{f^2}{2k_0(s)} - v(s)\right)\right] \right. \\ &\quad \left. \times \cosh\frac{s\ell f}{k_B T} \right\}^N . \end{aligned} \tag{2.28}$$

The last sum can be evaluated as

$$Z_G(f) = (8\pi k_B T)^{\frac{N}{2}} \mathcal{D}^N, \tag{2.29}$$

by writing

$$\mathcal{D} = \sqrt{\frac{1}{k}} \exp\left(\frac{f^2}{2k_B T k}\right) \cosh\left(\frac{\ell f}{k_B T}\right) + \phi \sqrt{\frac{1}{h}} \exp\left(\frac{f^2}{2k_B T h}\right) \cosh\left(\frac{\chi \ell f}{k_B T}\right), \tag{2.30}$$

and $\phi = \exp\left(-\frac{\Delta E}{k_B T}\right)$. It is particularly noticeable that there is no interaction between the elements of the chain in the Gibbs ensemble. Consequently, we can write the partition function as a power with exponent N. A similar result can be found in Eq.(34) of Ref. [105], describing the statistical behaviour of a Fermi-Pasta-Ulam chain of bistable elements. Also in this case a soft device leads to the independence of the bistable units.

With the help of the expression $x_N = -\partial U_{tot}/\partial f$, the extension of the chain can be directly calculated and its average value is therefore $\langle x_N \rangle = \langle -\partial U_{tot}/\partial f \rangle$. It can be simply evaluated by means of the partition function, as $\langle x_N \rangle = k_B T\, \partial/\partial f (\log Z_G)$. The calculation eventually gives

$$\begin{aligned} \frac{\langle x_N \rangle}{N\ell} &= \left\{ \sqrt{\frac{1}{k}} \exp\left(\frac{f^2}{2k_B T k}\right) \left[\frac{f}{k\ell} \cosh\left(\frac{\ell f}{k_B T}\right) + \sinh\left(\frac{\ell f}{k_B T}\right)\right] \right. \\ &\quad \left. + \chi\phi \sqrt{\frac{1}{h}} \exp\left(\frac{f^2}{2k_B T h}\right) \left[\frac{f}{\chi \ell h} \cosh\left(\frac{\chi \ell f}{k_B T}\right) + \sinh\left(\frac{\chi \ell f}{k_B T}\right)\right] \right\} / \mathcal{D}. \end{aligned} \tag{2.31}$$

In the simpler case with $k = h$, we get

$$\frac{\langle x_N \rangle}{N\ell} = \frac{f}{k\ell} + \frac{\sinh\left(\frac{\ell f}{k_B T}\right) + \chi\phi \sinh\left(\frac{\chi \ell f}{k_B T}\right)}{\cosh\left(\frac{\ell f}{k_B T}\right) + \phi \cosh\left(\frac{\chi \ell f}{k_B T}\right)}. \tag{2.32}$$

We can also calculate the average value of the spin variable $\langle y \rangle = \langle y_i \rangle \, \forall i$, which is independent of the characteristics of the elements considered in the chain and is given by

$$\langle y \rangle = \langle y_i \rangle = \frac{\mathcal{N}}{\mathcal{D}} \quad \forall i, \tag{2.33}$$

where

$$\mathcal{N} = \sqrt{\frac{1}{k}} \exp\left(\frac{f^2}{2k_B T k}\right) \sinh\left(\frac{\ell f}{k_B T}\right) + \chi\phi \sqrt{\frac{1}{h}} \exp\left(\frac{f^2}{2k_B T h}\right) \sin\left(\frac{\chi \ell f}{k_B T}\right), \tag{2.34}$$

and $\mathcal{D}$ is given in Eq.(2.30). In the simpler case with $k = h$, we get

$$\langle y \rangle = \frac{\sinh\left(\frac{\ell f}{k_B T}\right) + \chi\phi \sinh\left(\frac{\chi \ell f}{k_B T}\right)}{\cosh \frac{\ell f}{k_B T} + \phi \cosh\left(\frac{\chi \ell f}{k_B T}\right)}. \tag{2.35}$$

By combining Eq.(2.32) with Eq.(2.35), we immediately obtain $\frac{\langle x_N \rangle}{N\ell} = \frac{f}{k\ell} + \langle y \rangle$ or, equivalently,

$$f = k\left(\frac{\langle x_N \rangle}{N} - \ell \langle y \rangle\right). \tag{2.36}$$

This constitutive equation represents a spring-like behaviour with an equilibrium length directly modulated by the average value of the spin variables.

An example of force-extension response and spin variable behaviour is represented in Fig.2.5 for $h = k$. In the curve of $\frac{f\ell}{k_B T}$ versus $\frac{\langle x_N \rangle}{N\ell}$, a force plateau corresponding to $f^* = \pm \frac{\Delta E}{(\chi - 1)\ell}$ or, equivalently, $\frac{f^* \ell}{k_B T} = \pm \frac{\Delta E}{(\chi - 1) k_B T}$ is observed. Equivalently, in the curve of $\frac{f\ell}{k_B T}$ versus $\langle y \rangle$, we can identify the regions $\langle y \rangle = \pm 1$ and $\langle y \rangle = \pm \chi$ with a transition corresponding to the same force f^*. This behaviour can be interpreted as a cooperative process, with a threshold force f^* leading to the transition of all the units of the chain at the same time. As far as DNA is concerned, it is possible to retrieve the value of its force plateau by using the following parameters: $\Delta E = 3,8 K_B T$, $\ell = 3,4$ Å, $\chi = 1,7$ and $(\chi - 1)\ell = 2,4$Å. The value of the threshold force thus obtained is about 65 pN, corresponding to the well-known experimental value [79].

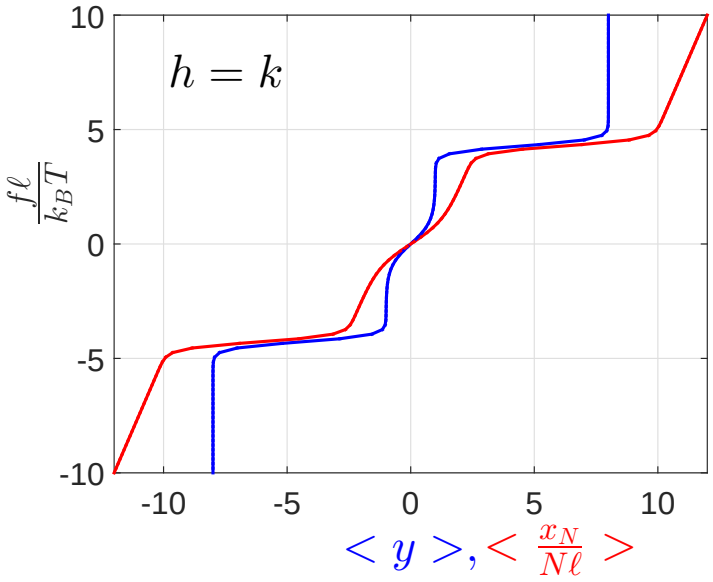

Figure 2.5 – Responses of the one-dimensional system with multistable elements under Gibbs conditions ($h = k$): average spin variable (blue line) and normalised extension (red line) versus dimensionless force. We adopted the parameters $\Delta E = 30 k_B T = 12.4 \times 10^{-20}$J, $\chi = 8$, $\ell = 0.5$nm and $h = k = 2.5 k_B T/\ell^2 = 0.0414$N/m (at $T = 300$K). The results are independent of N within the Gibbs ensemble [113].

The cooperative transition is manifest in both the force-extension curve and the spin variable behaviour. Of course, the slope of the plateau observed depends on the temperature. For higher values of the temperature, the response is smoother and the transition of the elements occurs in a less synchronised manner. Anyway, the value of the threshold force f^* can be explained as follows. As the response is symmetric, we limit the following discussion to the region with $f > 0$ and $\langle x_N \rangle > 0$. We consider the two following potential energies: $U_1(x) = \frac{1}{2}k(x - \ell)^2 - fx$ and $U_2(x) = \Delta E + \frac{1}{2}k(x - \chi\ell)^2 - fx$, corresponding to the potential wells of the system identified by $y = 1$ and $y = \chi$. The related equilibrium positions are defined by $\partial U_i/\partial x = 0$ and can be found as $x_1 = \ell + f/k$ and $x_2 = \chi\ell + f/k$. Thus, the unfolded configuration becomes more stable than the folded one if and only if $U_2(x_2) < U_1(x_1)$, that is to say $f < \frac{\Delta E}{(\chi-1)\ell}$, as observed in Fig.2.5. Notably, the value of the plateau force inducing the conformation transition does not depend on the spring constant, nor on the temperature. Such a result is readily interpreted in the framework of the Bell expression, originally derived in the context of the adhesion of cells [159].

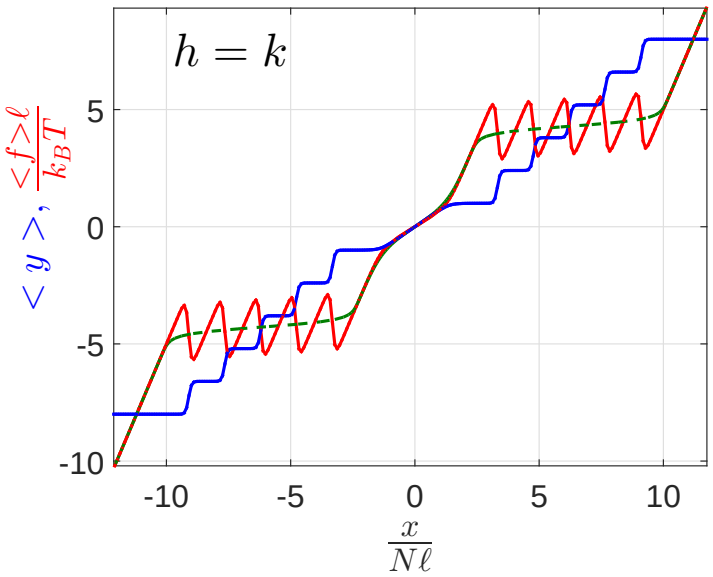

Figure 2.6 – Responses of the one-dimensional system with multistable elements under Helmholtz conditions for $N = 5$: average spin variable (blue line) and dimensionless force (red line) versus normalised extension. We adopted the parameters $\Delta E = 30 k_B T = 12.4 \times 10^{-20}$J, $\chi = 8$, $\ell = 0.5$nm and $h = k = 2.5 k_B T/\ell^2 = 0.0414$N/m (at $T = 300$K), as in Fig.2.5. For comparison, we also plotted the Gibbs force-extension response (green dashed line) [113].

2.2.1.2 The Helmholtz ensemble

We consider now the chain of multistable elements with both end-terminals tethered at the points $x_0 = 0$ and $x_N = x$, respectively. This leads to the following form for the potential energy under the isometric condition

$$U^H_{tot}(\vec{x}, \vec{y}; x_N) \quad = \quad \sum_{i=1}^{N} U(x_i - x_{i-1}, y_i), \tag{2.37}$$

where $x_N = x$ is the fixed extremity of the chain, $\vec{x} = (x_1, x_2, ..., x_{N-1})$ are continuous variable and $\vec{y} = (y_1, y_2, ..., y_N)$ are discrete spin variables. The potential energy $U(x, y)$ of a single unit in Eq.(2.37) is given in Eq.(2.21). Therefore, the partition function of the system can be written as

$$Z_H(x_N) = \sum_{y_1 \in \mathbf{S}} ... \sum_{y_N \in \mathbf{S}} \int_{\Re} ... \int_{\Re} e^{-\frac{U^H_{tot}}{k_B T}} d\mathrm{d}x_1 ... \mathrm{d}x_{N-1}. \tag{2.38}$$

The direct calculation of the integral in Eq.(2.38) is rather complicated, as the equality $x_N = x$ leads to interactions between the units of the chain. Nevertheless, by comparing

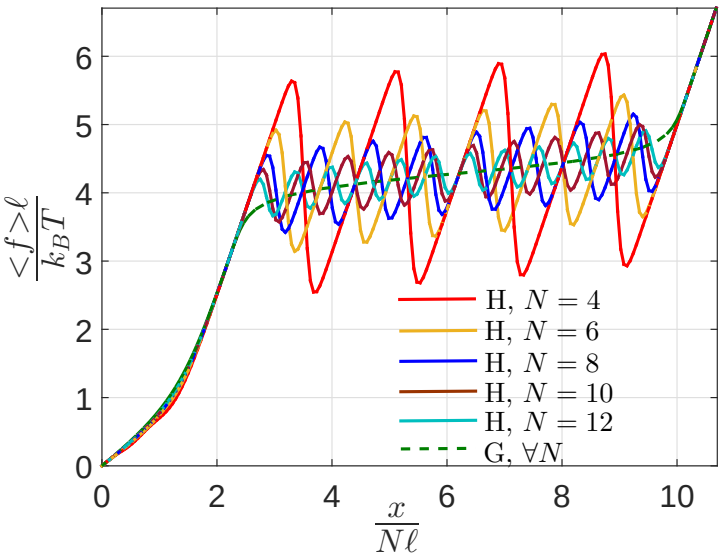

Figure 2.7 – Force-extension response of the one-dimensional system with multistable elements under Helmholtz (H) conditions for $N = 4, 6, 8, 10, 12$. We also reported the Gibbs (G) response to show the equivalence of the ensembles in the thermodynamic limit. We adopted the parameters $\Delta E = 30 k_B T = 12.4 \times 10^{-20}$J, $\chi = 8$, $\ell = 0.5$nm and $h = k = 2.5 k_B T/\ell^2 = 0.0414$N/m (at $T = 300$K), as in Figs.2.5 and 2.6 [113].

Eqs.(2.23) and (2.38), we deduce that the two partition functions Z_G and Z_H are related through a bilateral Laplace transform, as follows

$$Z_G(f) = \int_{-\infty}^{+\infty} Z_H(x) \exp\left(\frac{fx}{k_B T}\right) \mathrm{d}x. \tag{2.39}$$

Moreover, if we let $f = -i\omega k_B T$, we simply obtain

$$Z_G(-i\omega k_B T) = \int_{-\infty}^{+\infty} Z_H(x) \exp\left(-i\omega x\right) \mathrm{d}x, \tag{2.40}$$

which means that the Fourier transform of Z_H gives the analytical continuation of Z_G on the imaginary axis. Thus, we can invert the Fourier transform, by getting

$$Z_H(x) = \int_{-\infty}^{+\infty} Z_G(-i\omega k_B T) \exp\left(i\omega x\right) \mathrm{d}\omega. \tag{2.41}$$

Consequently, by considering the Gibbs partition function and with the help of Eq.(2.41), it is possible to study the behaviour of the system in the Helmholtz ensemble. Anyway, from Eqs.(2.29) and (2.30), we have

$$\begin{aligned} Z_G(-i\omega k_B T) \;=\; & c \sum_{p=0}^{N} \binom{N}{p} \sqrt{\frac{1}{h^p}} \sqrt{\frac{1}{k^{N-p}}} \phi^p \exp\left[-\frac{1}{2} k_B T \left(\frac{p}{h} + \frac{N-p}{k}\right) \omega^2\right] \\ & \times \cos^p(\chi \ell \omega) \cos^{N-p}(\ell \omega), \end{aligned} \tag{2.42}$$

where c is an unimportant multiplicative constant. By using the Newton development

$$\cos^n x = \frac{1}{2^n} e^{-inx} \sum_{t=0}^{n} \binom{n}{t} e^{2itx}, \tag{2.43}$$

we obtain from Eqs.(2.41) and (2.42)

$$\begin{aligned} Z_H(x) &= c \sum_{p=0}^{N} \sum_{q=0}^{p} \sum_{s=0}^{N-p} \binom{N}{p} \binom{p}{q} \binom{N-p}{s} \sqrt{\frac{1}{h^p}} \sqrt{\frac{1}{k^{N-p}}} \phi^p \\ &\quad \times \int_{-\infty}^{+\infty} \exp\left[-\frac{1}{2} k_B T \left(\frac{p}{h} + \frac{N-p}{k} \right) \omega^2 \right] \\ &\quad \times \exp\left[i\omega (2\chi \ell q - \chi \ell p + 2\ell s - \ell N + \ell p + x) \right] \mathrm{d}\omega. \end{aligned} \tag{2.44}$$

The integral in Eq.(2.44) can be evaluated in a closed form by means of the standard expression

$$\int_{-\infty}^{+\infty} e^{-\alpha x^2} e^{i\beta x} \mathrm{d}x = \sqrt{\frac{\pi}{\alpha}} e^{-\frac{\beta^2}{4\alpha}} \quad (\alpha > 0), \tag{2.45}$$

eventually giving

$$Z_H(x) = c \sum_{p=0}^{N} \sum_{q=0}^{p} \sum_{s=0}^{N-p} \binom{N}{p} \binom{p}{q} \binom{N-p}{s} \frac{1}{\sqrt{\frac{p}{h} + \frac{N-p}{k}}} \sqrt{\frac{1}{h^p}} \sqrt{\frac{1}{k^{N-p}}} \phi^p e^{-\varphi^2}, \tag{2.46}$$

where

$$\varphi = \frac{2\chi \ell q - \chi \ell p + 2\ell s - \ell N + \ell p + x}{\sqrt{2 k_B T \left(\frac{p}{h} + \frac{N-p}{k} \right)}}. \tag{2.47}$$

In this case, it is important to remark that the resulting partition function can not be written in terms of a power with exponent N. Therefore, the Helmholtz (isotensional) condition induces an effective interaction between the elements. The origin of this interaction is not explicitly defined in the potential energy of the system (as, e.g., in the Ising model), but comes from the specific boundary conditions characterising the Helmholtz ensemble. Our result can be compared with Eq.(20) of Ref. [105], obtained for a Fermi-Pasta-Ulam chain with bistable elements. This work permits to confirm that using hard devices leads to interaction among the elements (see also Fig.13 of Ref. [105] for the comparison of the mechanical response with soft or hard devices). Now, the average value of the overall force $\langle f \rangle = -k_B T \partial / \partial x \log Z_H$ applied to the system and the average value of the spin variables $\langle y \rangle = \left\langle \frac{1}{N} \sum_{i=1}^{N} y_i \right\rangle$ describing the transitions can be calculated as follows

$$\langle f\rangle = \frac{\displaystyle\sum_{p=0}^{N}\sum_{q=0}^{p}\sum_{s=0}^{N-p}\binom{N}{p}\binom{p}{q}\binom{N-p}{s}\frac{\sqrt{2k_BT}}{\frac{p}{h}+\frac{N-p}{k}}\sqrt{\frac{1}{h^p}}\sqrt{\frac{1}{k^{N-p}}}\phi^p\varphi e^{-\varphi^2}}{\displaystyle\sum_{p=0}^{N}\sum_{q=0}^{p}\sum_{s=0}^{N-p}\binom{N}{p}\binom{p}{q}\binom{N-p}{s}\frac{1}{\sqrt{\frac{p}{h}+\frac{N-p}{k}}}\sqrt{\frac{1}{h^p}}\sqrt{\frac{1}{k^{N-p}}}\phi^p e^{-\varphi^2}}, \quad (2.48)$$

$$\langle y\rangle = \frac{\frac{1}{N}\displaystyle\sum_{p=0}^{N}\sum_{q=0}^{p}\sum_{s=0}^{N-p}\binom{N}{p}\binom{p}{q}\binom{N-p}{s}\frac{1}{\sqrt{\frac{p}{h}+\frac{N-p}{k}}}\sqrt{\frac{1}{h^p}}\sqrt{\frac{1}{k^{N-p}}}\phi^p Q e^{-\varphi^2}}{\displaystyle\sum_{p=0}^{N}\sum_{q=0}^{p}\sum_{s=0}^{N-p}\binom{N}{p}\binom{p}{q}\binom{N-p}{s}\frac{1}{\sqrt{\frac{p}{h}+\frac{N-p}{k}}}\sqrt{\frac{1}{h^p}}\sqrt{\frac{1}{k^{N-p}}}\phi^p e^{-\varphi^2}}, \quad (2.49)$$

where $Q = (\chi p - 2\chi q + N - p - 2s)$.

An application of Eqs.(2.48) and (2.49) can be found in Fig.2.6 where the average spin variable and the dimensionless force are represented versus the normalised extension for $h = k$ and $N = 5$ (with the same parameters than those used in Fig.2.5). The force-extension response for the Gibbs condition is also shown for comparison. A non-cooperative behaviour is observed, showing that the units unfold progressively with the increasing total length of the system. This is a non-synchronised process, characterised by the picks visible in the Helmholtz force-extension relation and by the sequence of steps characterising the average value of the spin variable. In conclusion, whereas in the Gibbs ensemble, the unfolding of the domains is a collective (cooperative) process, leading to a force plateau in the force-extension curve, in the Helmholtz ensemble, the unfolding of the domains is a non-cooperative process, specific to each unit, leading to a sawtooth-like force-extension curve.

In Fig.2.7, the force-extension curve in the Helmholtz ensemble for different values of $N = 4, 6, 8, 10$ and 12 is shown. It is particularly noticeable that for an increasing number N of elements, the Helmholtz response converges to the Gibbs one, by progressively reducing the pick-to-pick distance in the sawtooth pattern. This is in perfect agreement with recent results concerning the ensembles equivalence in the thermodynamic limit [20, 139].

2.2.2 Bistable freely jointed chain

In this Section, we elaborate a generalisation of the classical freely jointed chain model in order to introduce the two-state behaviour in the elements of the system. The model

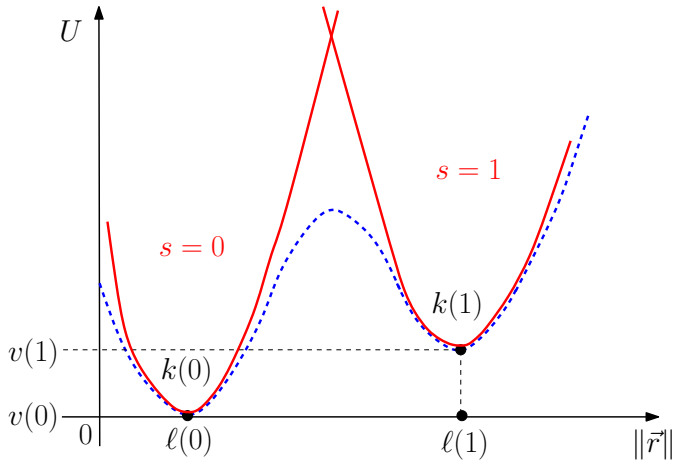

Figure 2.8 – Potential energy of a single element of the bistable freely jointed chain (dashed blue curve). The potential wells are approximated through two parabolic profiles identified by $s = 0$ and 1 [113].

is three-dimensional and the potential energy of the units of the chain is described by a bistable behaviour (see Fig.2.8, dashed blue line). The two potential wells can be written as

$$U(\vec{r}, s) = v(s) + \frac{1}{2}k(s)\left[\|\vec{r}\| - \ell(s)\right]^2, \tag{2.50}$$

where $\vec{r}$ is the end-to-end vector joining both ends of the unit (see Fig.2.8, red lines). The parameters $s \in \{0, 1\}$ represents a spin variable, permitting to identify what potential well is explored by the vector $\vec{r}$. The quantities $v(s)$, $k(s)$ and $\ell(s)$ stand for the basal energy, the elastic stiffness and the equilibrium length of the potential wells, respectively. It has to be noted that in the classical freely jointed chain, the elastic stiffness tends to infinity or, equivalently, the length of each element is keep constant. Accordingly, the mechanical behaviour is fully governed by entropic forces since the elastic contribution is not taken into account. Here, for mathematical convenience, it is better to proceed from Eq.(2.50), with finite elastic constants, and to analyse the limiting cases in a following phase. We will study this system within both the Gibbs and the Helmholtz ensembles.

2.2.2.1 The Gibbs ensemble

In the following, we consider a chain of N elements described by Eq.(2.50), with a force applied to the last one. Hence, the total potential energy of the system can be written as

$$U_{tot}(\vec{q}, \vec{s}; \vec{f}) = \sum_{i=1}^{N} U(\vec{r}_i - \vec{r}_{i-1}, s_i) - \vec{f} \cdot \vec{r}_N, \tag{2.51}$$

where $\vec{q} = (\vec{r}_1, ..., \vec{r}_N)$ is the generalised coordinates vector containing all positions $\vec{r}_1, ..., \vec{r}_N$, $\vec{s} = (s_1, ..., s_N)$ is the vector of all spin variables, and $\vec{f}$ is the force applied to the last element of the chain. The partition function can be therefore calculated by summing the discrete (spin-like) variables and integrating the continuous (coordinates) ones, as follows

$$Z_G(\vec{f}) = \sum_{s_1 \in \{0,1\}} ... \sum_{s_N \in \{0,1\}} \int_{\Re^{3N}} e^{-\frac{U_{tot}(\vec{q},\vec{s};\vec{f})}{k_B T}} \mathrm{d}\vec{q}. \tag{2.52}$$

We apply the change of variables $\vec{\xi}_1 = \vec{r}_1 - \vec{r}_0$, $\vec{\xi}_2 = \vec{r}_2 - \vec{r}_1$,..., $\vec{\xi}_N = \vec{r}_N - \vec{r}_{N-1}$ to evaluate the integral over the vector $\vec{q}$, leading to $\sum_{k=1}^{N} \vec{\xi}_k = \vec{r}_N - \vec{r}_0$ and $\mathrm{d}\vec{q} = \mathrm{d}\vec{\xi}_1 ... \mathrm{d}\vec{\xi}_N$. By fixing $\vec{r}_0$ at the origin of axes, we obtain

$$Z_G = \left\{ \sum_{s \in \{0,1\}} \int_{\Re^3} \exp\left[-\frac{U(\vec{\xi}, s)}{k_B T} + \frac{\vec{f} \cdot \vec{\xi}}{k_B T} \right] \mathrm{d}\vec{\xi} \right\}^N, \tag{2.53}$$

which means, giving the power of N, that the partition function is multiplicative with respect to the elements of the chain (as observed in Eq.(2.28)). Since the system spherically symmetric, we can choose an arbitrary direction for the applied force. Thus, to further simplify Eq.(2.53), we set $\vec{f} = (0, 0, f)$ and we change the variables according to $\vec{\xi} = (\xi \cos\varphi \sin\vartheta, \xi \sin\varphi \sin\vartheta, \xi \cos\vartheta)$. As $\mathrm{d}\vec{\xi} = \xi^2 \sin\vartheta \mathrm{d}\xi \mathrm{d}\varphi \mathrm{d}\vartheta$, $\|\vec{\xi}\| = \xi$ and $\vec{f} \cdot \vec{\xi} = f\xi \cos\vartheta$, we get the following simpler form of the partition function

$$Z_G = c \left\{ \sum_{s \in \{0,1\}} e^{-\frac{v(s)}{k_B T}} \int_0^\infty \exp\left[-\frac{k(s)}{2k_B T} \left[\xi - \ell(s)\right]^2 \right] \frac{\sinh\left(\frac{f\xi}{k_B T}\right)}{\frac{f\xi}{k_B T}} \xi^2 \mathrm{d}\xi \right\}^N, \tag{2.54}$$

where c stands represents a non influential multiplicative constant. As already noted in the case of the one-dimensional model within the Gibbs ensemble, the elements of the chain do not interact and this point leads to a partition function which is in the form of a power with exponent N.

The specific properties of the two potential wells can now be described, namely $v(0) = 0, \ell(0) = \ell, k(0) = K$ and $v(1) = \Delta E, \ell(1) = \chi\ell, k(1) = K$, where χ is the elongation ratio between the unfolded and folded equilibrium lengths. Then, we get

$$\begin{aligned} Z_G = c \Bigg\{ & \int_0^\infty e^{-\frac{K}{2k_B T}[\xi - \ell]^2} \frac{\sinh\left(\frac{f\xi}{k_B T}\right)}{\frac{f\xi}{k_B T}} \xi^2 \mathrm{d}\xi \\ & + \phi \int_0^\infty e^{-\frac{K}{2k_B T}[\xi - \chi\ell]^2} \frac{\sinh\left(\frac{f\xi}{k_B T}\right)}{\frac{f\xi}{k_B T}} \xi^2 \mathrm{d}\xi \Bigg\}^N, \end{aligned} \tag{2.55}$$

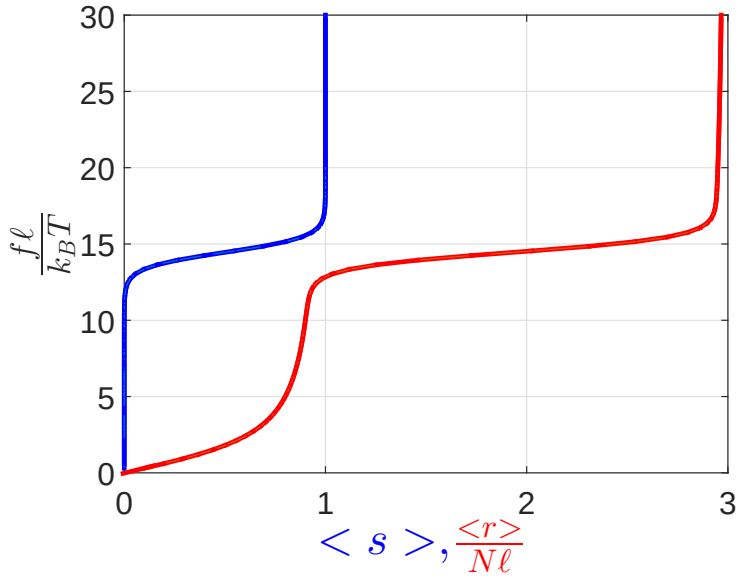

Figure 2.9 – Average normalised extension (red line) and normalised spin variable (blue line) versus the dimensionless force for the bistable freely jointed chain under Gibbs conditions. We adopted the parameters $\Delta E = 30 k_B T$ and $\chi = 3$ (at $T = 300$K). The curves are independent of N within the Gibbs ensemble [113].

where $\phi = \exp\left(-\frac{\Delta E}{k_B T}\right)$. To evaluate the limit for K approaching infinity, the latter form of the partition function can be used. This allows to properly define the bistable freely jointed chain. For this purpose, we can use the Dirac delta function property $\sqrt{\frac{\alpha}{\pi}} e^{-\alpha(x-x_0)^2} \underset{\alpha\to\infty}{\to} \delta(x - x_0)$, eventually yielding

$$Z_G = c \left\{ \frac{\sinh y}{y} + \chi\phi \frac{\sinh(\chi y)}{y} \right\}^N , \tag{2.56}$$

where we introduced the dimensionless force $y = \frac{\ell f}{k_B T}$. If the system under consideration is not bistable, the second term in Eq.(2.56) vanishes, and we obtain

$$Z_G = c \left\{ \frac{\sinh y}{y} \right\}^N , \tag{2.57}$$

which is the classical partition function of the freely jointed chain model [120, 122, 124]. The cases with finite values of the elastic constant will be considered in the next Chapter. The following standard relation permits to find the force-extension response for the bistable freely jointed $\langle r \rangle = k_B T \frac{\partial \log Z_G}{\partial f}$, producing the important result

$$\langle r \rangle = N\ell \frac{\mathcal{L}(y) + \chi^2 \phi \mathcal{L}(\chi y) \frac{\sinh(\chi y)}{\sinh y}}{1 + \chi\phi \frac{\sinh(\chi y)}{\sinh y}}, \tag{2.58}$$

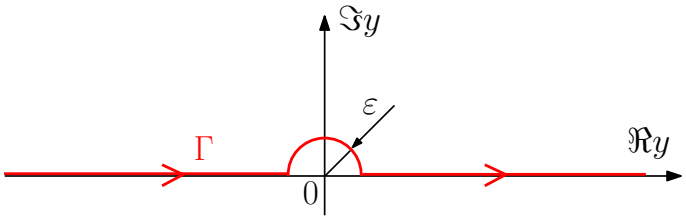

Figure 2.10 – Definition of the contour Γ on the complex plane [113].

where $\mathcal{L}(y) = \coth y - \frac{1}{y}$ is the Langevin function. If the average value of the spin variables is written as $\langle s \rangle = \frac{1}{N} \left\langle \sum_{i=1}^{N} s_i \right\rangle$, it is not difficult to prove that $\langle s \rangle = -k_B T \frac{\partial \log Z_G}{N \partial \Delta E}$. The second important achievement is therefore obtained

$$\langle s \rangle = \frac{\chi\phi \frac{\sinh(\chi y)}{\sinh y}}{1 + \chi\phi \frac{\sinh(\chi y)}{\sinh y}}. \tag{2.59}$$

By combining Eqs.(2.58) and (2.59), we can find the relationship

$$\langle r \rangle = N \left[(1 - \langle s \rangle) \, \ell \mathcal{L}(y) + \langle s \rangle \, \chi \ell \mathcal{L}(\chi y) \right], \tag{2.60}$$

affirming that the average extension of the bistable system is given by a linear combination of the responses of a FJC model with length ℓ and a FJC model with length $\chi\ell$, being the coefficients controlled by the average value of the spin variable. In other words, when $\langle s \rangle$ varies from 0 to 1, the effective equilibrium length of the element during the progressive unfolding increases from ℓ to $\chi\ell$.

An application of Eqs.(2.58) and (2.59) is shown in Fig.2.9, where the average normalised extension and the normalised spin variable are represented versus the dimensionless force. The force-extension curve shows a cooperative behaviour due to the collective unfolding of all the domains at the threshold value of the dimensionless force $\frac{f^* \ell}{k_B T} = \frac{\Delta E}{(\chi - 1) k_B T}$, as predicted in Section 2.2.1.1. The spin variable curve confirms this behaviour by exhibiting a clear transition from 0 to 1 at the same threshold value of the force.

2.2.2.2 The Helmholtz ensemble

Here, the system under consideration is the chain of bistable elements with both end-terminals tethered at the points $\vec{r}_0 = \vec{0}$ and $\vec{r}_N = \vec{r}$. This means that the total potential energy of the system can be written as

$$U^H_{tot}(\vec{q}, \vec{s}; \vec{r}_N) = \sum_{i=1}^{N} U(\vec{r}_i - \vec{r}_{i-1}, s_i), \tag{2.61}$$

where $\vec{r}_N = \vec{r}$ is the fixed extremity of the chain, $\vec{q} = (\vec{r}_1, ..., \vec{r}_{N-1})$ is the generalised coordinates vector containing all positions $\vec{r}_1, ..., \vec{r}_{N-1}$, and $\vec{s} = (s_1, ..., s_N)$ is the vector of all spin variables. The potential energy $U(\vec{r}, s)$ of a single unit in Eq.(2.61) is given in Eq.(2.50). The partition function of this system can be written as

$$Z_H(\vec{r}_N) = \sum_{s_1 \in \{0,1\}} ... \sum_{s_N \in \{0,1\}} \int_{\Re^{3(N-1)}} e^{-\frac{U^H_{tot}(\vec{q},\vec{s};\vec{r}_N)}{k_B T}} \mathrm{d}\vec{q}. \tag{2.62}$$

The comparison of Eqs.(2.52) and (2.62) allows to deduce the relation between the two partition functions Z_G and Z_H through a three-dimensional bilateral Laplace transform, as follows

$$Z_G(\vec{f}) = \int_{\Re^3} Z_H(\vec{r}) \exp\left(\frac{\vec{r} \cdot \vec{f}}{k_B T}\right) \mathrm{d}\vec{r}. \tag{2.63}$$

Moreover, since the problem is spherically symmetric, we easily get the following relationship

$$Z_H(r) = c \int_{-\infty}^{+\infty} Z_G(i\eta) \frac{\eta}{r} \sin\left(\frac{\eta r}{k_B T}\right) \mathrm{d}\eta, \tag{2.64}$$

and by substituting Eq.(2.56), we get the important integral expression

$$Z_H(r) = c \int_{-\infty}^{+\infty} \left\{\frac{\sin y}{y} + \chi\phi \frac{\sin(\chi y)}{y}\right\}^N \frac{y}{r} \sin\left(\frac{ry}{\ell}\right) \mathrm{d}y. \tag{2.65}$$

If the studied system is not bistable, the partition function can be formulated as

$$Z_H(r) = c \int_{-\infty}^{+\infty} \left\{\frac{\sin y}{y}\right\}^N \frac{y}{r} \sin\left(\frac{ry}{\ell}\right) \mathrm{d}y, \tag{2.66}$$

a result which has been largely studied by Rayleigh [160], Polya [161], Treloar [162], and Wang and Guth [163] to analyse the behaviour of chains and chains networks.

Here, in order to study the behaviour of the chain in the Helmholtz ensemble, we elaborate Eq.(2.65) in closed form. The function to integrate in Eq.(2.65) is regular on the real axis and analytical on a strip $|\Im m\, y| < M$ for an arbitrary $M \in \Re$. Then, instead of integrating on the whole real axis, the path Γ shown in Fig.2.10 is used. Therefore, we can write

$$Z_H(r) = -ic \int_{\Gamma} \left\{\frac{\sin y}{y} + \chi\phi \frac{\sin(\chi y)}{y}\right\}^N \frac{y}{r} e^{i\frac{ry}{\ell}} \mathrm{d}y. \tag{2.67}$$

By developing the power in previous expression and by using the expansion

$$\sin^n x = \frac{1}{(2i)^n} e^{inx} \sum_{t=0}^{n} \binom{n}{t} (-1)^t e^{-2itx}, \tag{2.68}$$

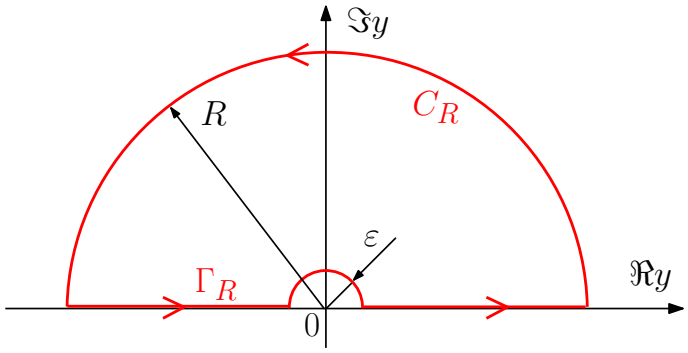

Figure 2.11 – Contour in the complex plane used to solve the integral for $a > 0$ [113].

we obtain

$$\begin{aligned} Z_H(r) &= \frac{c}{2^N i^{N+1} r} \sum_{k=0}^{N} \sum_{p=0}^{N-k} \sum_{q=0}^{k} \binom{N}{k} \binom{N-k}{p} \\ &\times \binom{k}{q} (-1)^{p+q} (\chi\phi)^k \int_\Gamma \frac{e^{-i\Lambda y}}{y^{N-1}} \mathrm{d}y, \end{aligned} \tag{2.69}$$

where we defined

$$\Lambda = k - N + 2p - \chi k + 2\chi q - \frac{r}{\ell}. \tag{2.70}$$

For the following calculations, we call a the parameter Λ and m the parameter $N-1$. We evaluate the integral $\int_\Gamma \frac{e^{iay}}{y^m} \mathrm{d}y$ over the contour Γ shown in Fig.2.10. To begin, we suppose $a \in \Re$ and $a > 0$. In this case, we observe that on the imaginary axis, we have $e^{iay} = e^{-a\Im m\, y}$, which is a decreasing to zero function for $\Im m\, y > 0$. So, we consider the contour in Fig.2.11 and we write

$$\oint_{\Gamma_R \cup C_R} \frac{e^{iay}}{y^m} \mathrm{d}y = 0, \tag{2.71}$$

since the function is holomorphic within $\Gamma_R \cup C_R$. We also have

$$\int_{\Gamma_R} \frac{e^{iay}}{y^m} \mathrm{d}y + \int_{C_R} \frac{e^{iay}}{y^m} \mathrm{d}y = 0. \tag{2.72}$$

Since $\lim_{R\to\infty} \int_{C_R} \frac{e^{iay}}{y^m} \mathrm{d}y = 0$ for the Jordan lemma and $\Gamma_R \to \Gamma$ when $R \to \infty$, we obtain the first result

$$\int_\Gamma \frac{e^{iay}}{y^m} \mathrm{d}y = 0 \ \text{ if } \ a > 0. \tag{2.73}$$

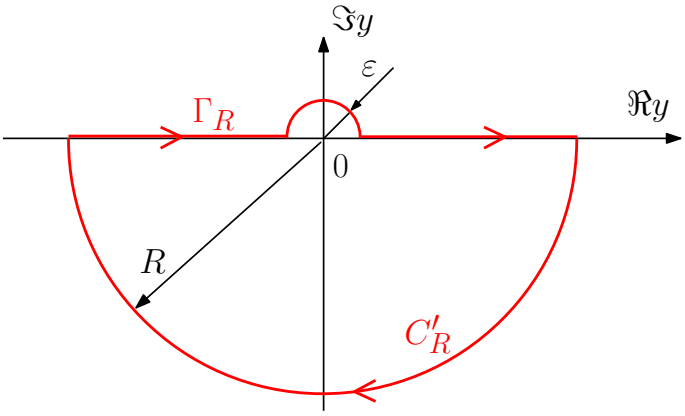

Figure 2.12 – Contour in the complex plane used to solve the integral for $a < 0$ [113].

We consider now the case with $a < 0$. On the imaginary axis, we have $e^{iay} = e^{-a\Im m\, y}$, which is a decreasing to zero function for $\Im m\, y < 0$. Therefore, we introduce the contour shown in Fig.2.12. The function is not holomorphic within $\Gamma_R \cup C'_R$ since it presents a pole of order m for $y = 0$. Hence, we have

$$\oint_{\Gamma_R \cup C'_R} \frac{e^{iay}}{y^m} \mathrm{d}y = -2\pi i \mathrm{Res}\left\{\frac{e^{iay}}{y^m}, 0\right\}, \tag{2.74}$$

or

$$\int_{\Gamma_R} \frac{e^{iay}}{y^m} \mathrm{d}y + \int_{C'_R} \frac{e^{iay}}{y^m} \mathrm{d}y = -2\pi i \mathrm{Res}\left\{\frac{e^{iay}}{y^m}, 0\right\}. \tag{2.75}$$

Since $a < 0$, for the Jordan lemma we have $\lim_{R\to\infty} \int_{C'_R} \frac{e^{iay}}{y^m} \mathrm{d}y = 0$. Moreover, the residue can be calculated as follows

$$\begin{aligned} \mathrm{Res}\left\{\frac{e^{iay}}{y^m}, 0\right\} &= \frac{1}{(m-1)!} \lim_{y\to 0} \frac{\mathrm{d}^{m-1}}{\mathrm{d}y^{m-1}} \left(y^m \frac{e^{iay}}{y^m}\right) \\ &= \frac{(ia)^{m-1}}{(m-1)!}. \end{aligned} \tag{2.76}$$

Summing up, we easily obtain the second result

$$\int_\Gamma \frac{e^{iay}}{y^m} \mathrm{d}y = -2\pi i^m \frac{a^{m-1}}{(m-1)!} \quad \text{if} \quad a \leq 0. \tag{2.77}$$

Hence, we proved in Eqs.(2.73) and (2.77) that

$$\int_\Gamma \frac{e^{iay}}{y^m} \mathrm{d}y = \begin{cases} 0 & \text{if} \quad a > 0 \\ -2\pi i^m \frac{a^{m-1}}{(m-1)!} & \text{if} \quad a \leq 0 \end{cases}, \tag{2.78}$$

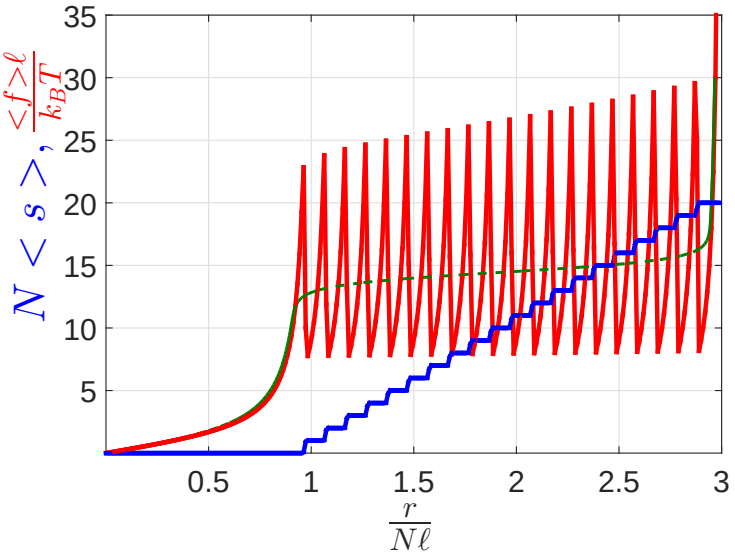

Figure 2.13 – Average dimensionless force (red line) and average spin variable (blue line) versus normalised extension for the bistable freely jointed chain under Helmholtz conditions. We adopted the parameters $\Delta E = 30 k_B T$, $\chi = 3$ and $N = 20$ (at $T = 300$ K). We also plotted for comparison the force-extension Gibbs response (dashed green line) [113].

and then we find from Eq.(2.69) the following result

$$Z_H(r) = \frac{\pi c}{2^{(N-1)}(N-2)!r} \sum_{k=0}^{N} \sum_{p=0}^{N-k} \sum_{q=0}^{k} \binom{N}{k} \binom{N-k}{p} \times \binom{k}{q} (-1)^{p+q} (\chi\phi)^k (-\Lambda)^{N-2} \mathbf{1}(\Lambda), \tag{2.79}$$

written in terms of the Heaviside step function $\mathbf{1}(x)$, defined as $\mathbf{1}(x) = 1$ if $x \geq 0$, and $\mathbf{1}(x) = 0$ if $x < 0$. As already noted in the one-dimensional case, the partition function can not be written in terms of a power with exponent N. The effective interaction among the elements in the Helmholtz ensemble induced by the boundary conditions is therefore confirmed.

The knowledge of the partition function allows us to obtain the force-extension response through the expression

$$\langle f \rangle = -k_B T \frac{\partial \log Z_H}{\partial r}, \tag{2.80}$$

and the average value of the spin variable, as follows

$$\langle s \rangle = \left\langle \frac{1}{N} \sum_{i=1}^{N} s_i \right\rangle = -\frac{1}{N} k_B T \frac{\partial \log Z_H}{\partial \Delta E}. \tag{2.81}$$

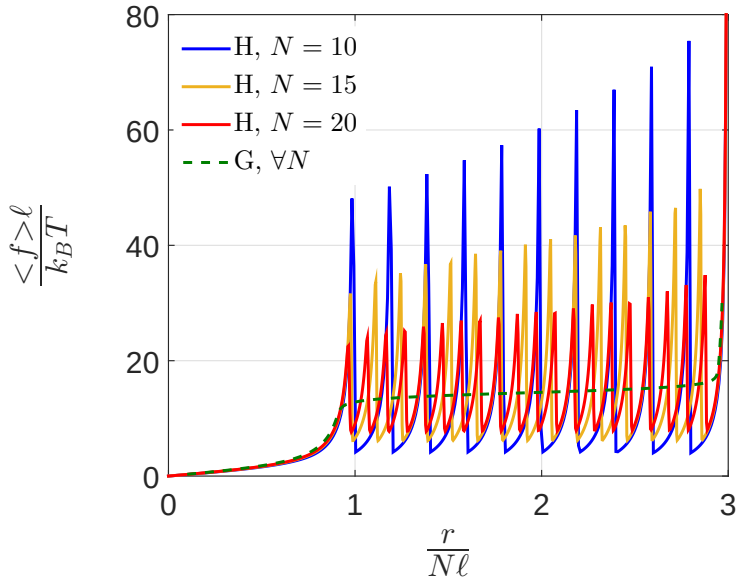

Figure 2.14 – Force-extension response of the bistable freely jointed chain with under Helmholtz (H) conditions for $N = 10, 15, 20$. We also reported the Gibbs (G) response to show the equivalence of the ensembles in the thermodynamic limit. We adopted the parameters $\Delta E = 30 k_B T$ and $\chi = 3$ [113].

For the sake of brevity, we omit the development of the derivatives and we show an example of application of Eqs.(2.80) and (2.81) in Fig.2.13. The typical sawtooth-like curve is observed in the force-extension response, corresponding to a non-cooperative process. This is explained by the individual successive unfolding the domains, one by one, as also confirmed by the average value of the spin variable, which exhibits a series of steps corresponding to each unfolding process.

In Fig.2.14, the force-extension curve in the Helmholtz ensemble for different values of $N = 10, 15$ and 20 is shown. It has to be noted that when the number N of elements increases, the Helmholtz response converges to the Gibbs one, by progressively reducing the pick-to-pick distance in the sawtooth pattern. Once again, the equivalence of the ensembles in the thermodynamic limit is confirmed, as recently demonstrated for a large class of non-confined polymer chains [20, 139].

Thus, we have described the spin variable method both in the Gibbs and the Helmholtz ensembles, allowing to consider a simpler form of the potential energy. The one-dimensional case has been studied, leading to closed forms of both Gibbs and Helmholtz partition functions. The partition function within the Gibbs ensemble has been calculated by direct integration, which allowed us to obtain the partition function within the Helmholtz ensemble by means of a bilateral Laplace transform. The three-dimensional case of a

bistable freely jointed chain has also been studied, yielding both Gibbs and Helmholtz partition functions in closed form. These results permitted us to give an interpretation of the cooperative (synchronised) and non-cooperative (non-synchronised) responses of the folding/unfolding processes. Finally, important values like average value of the extension (in the Gibbs ensemble), or the force (Helmholtz ensemble), or again the spin variable (in both ensembles) can now be easily evaluated. The three-dimensional case with a finite elasticity will be studied in Chapter 3.

2.3 Full statistics of conjugated thermodynamic ensembles in chains of two-state units

Typically, in the theoretical analyses conducted to study the behaviour of two-state systems under isotensional or isometric conditions, the considered quantities correspond to the average values of the fluctuating variables. Consequently, the average extension and force are respectively considered in the Gibbs and the Helmholtz ensembles. However, it is important to study the actual distributions of these fluctuating or stochastic variables, in order to better understand the random behaviour of these systems and to draw more refined comparisons with experiments. Indeed, it is important to underline that the experimental activities may probe not only the average values of the relevant quantities but also their actual distribution. Basically, this is achieved by a very large statistics of the trajectories of the system under investigation, which allows for a good exploration of the phase space and, consequently, for the determination of the pertinent probability densities. Therefore, we propose here a methodology to determine the exact distributions or probability densities of the pertinent quantities defined in both Gibbs and Helmholtz ensembles. In particular, for the Gibbs ensemble, we determine the distribution of the couple $(\dot{x}_N, x_N)$, where x_N is the extension of a chain of N two-state elements (under applied deterministic force), and, for the Helmholtz ensemble, the distribution of $(\dot{f}, f)$, where f is the measured force (under prescribed deterministic extension).

2.3.1 Configurational partition functions and force-extension relations in the Gibbs and the Helmholtz ensembles

We consider a very simple one-dimensional system (a *toy model*) composed of N elements with mechanical bistability, connected in series to compose a chain. Each element of the chain is represented by a symmetric potential energy function $U(x)$ showing two minima at $x = \pm\ell$ (see Fig.2.15). The variable y assumes its values in the set $\mathbf{S} = \{\pm 1\}$

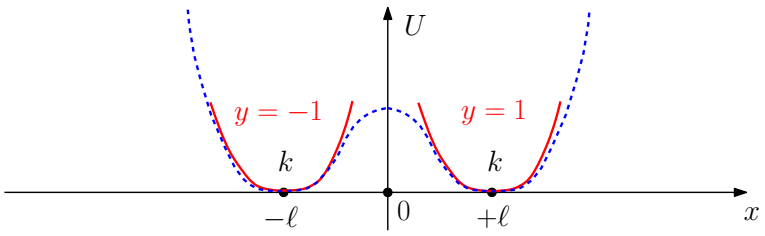

Figure 2.15 – Bistable symmetric potential energy of a single domain (blue dashed line) and its approximation by means of four parabolic profiles (red solid lines).

and is used to identify the basin or well explored by the system. In conclusion, the original bistable energy function is substituted with the simpler mathematical expression

$$U(x,y) = \frac{1}{2}k(x - y\ell)^2. \tag{2.82}$$

The potential energy in Eq.(2.82), by varying the value of the spin variable in $\mathbf{S}$, generates the two parabolic wells represented in Fig.2.15 (red solid lines). While without an applied stretching the units are in each basin with the same probability (the average value of the end-to-end distance is zero), an applied stretching induces a preferential direction in the extension of the chain. This stretching can be applied by imposing a force f (positive or negative) or prescribing the position x_N of the last element of the chain. Of course, in both cases, the first element is always tethered at the origin of the x-axis. These two possible mechanisms of stretching generate different stochastic mechanical behaviours of the system, which can be studied by calculating the corresponding configurational partition functions.

In the Gibbs ensemble, a force f is applied to the last chain unit, identified by its position x_N. The total potential energy of the system under the Gibbs condition (isotensional ensemble) is therefore given by

$$U^G_{tot}(\vec{x},\vec{y};f) = \sum_{i=1}^{N} U(x_i - x_{i-1}, y_i) - f x_N, \tag{2.83}$$

where f is the force applied to the last element, $\vec{x} = (x_1, x_2, ..., x_N)$ (continuous variables) and $\vec{y} = (y_1, y_2, ..., y_N)$ (discrete variables). For this system, we can define the configurational partition function Z_G, as follows

$$Z_G(f) = \sum_{y_1 \in \mathbf{S}} ... \sum_{y_N \in \mathbf{S}} \int_{\Re} ... \int_{\Re} e^{-\frac{U^G_{tot}}{k_B T}} \mathrm{d}x_1 ... \mathrm{d}x_N, \tag{2.84}$$

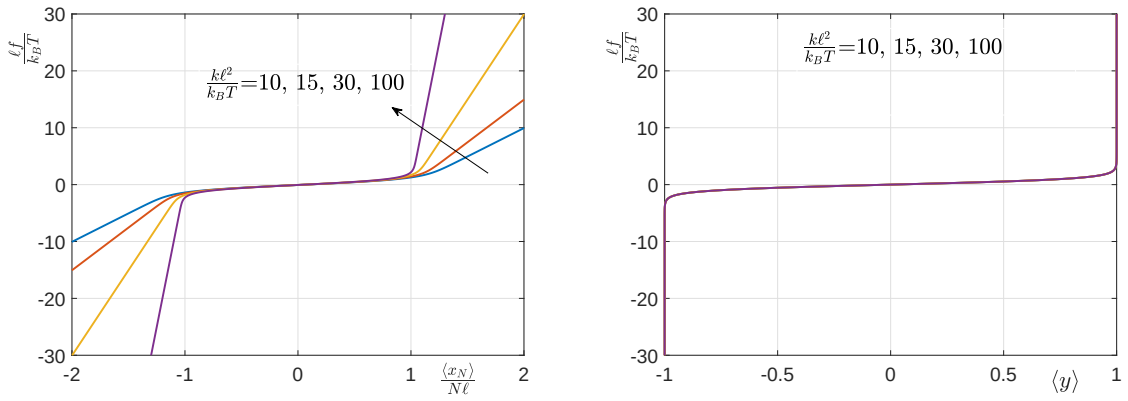

Figure 2.16 – Average force-extension curves and average spin variables (plotted by means of dimensionless quantities) for the Gibbs ensemble with $N = 5$ and $\frac{k\ell^2}{k_BT}$=10, 15, 30, 100 [3].

where the variable $\vec{x}$ is integrated whereas the variable $\vec{y}$ is summed. We can now substitute Eq.(2.83) in Eq.(2.84) and we eventually obtain the result

$$Z_G(f) = \left(\frac{8\pi k_B T}{k}\right)^{\frac{N}{2}} \left[\exp\left(\frac{f^2}{2k_B T k}\right) \cosh\left(\frac{\ell f}{k_B T}\right)\right]^N . \tag{2.85}$$

The extension of the chain can be directly calculated through the expression $x_N = -\partial U_{tot}/\partial f$ and its average value is therefore $\langle x_N\rangle = \langle -\partial U_{tot}/\partial f\rangle$. It can be simply evaluated by means of the configurational partition function, as $\langle x_N\rangle = k_B T\, \partial/\partial f(\log Z_G)$. The calculation eventually gives

$$\frac{\langle x_N\rangle}{N\ell} = \frac{f}{k\ell} + \tanh\left(\frac{\ell f}{k_B T}\right). \tag{2.86}$$

We can also calculate the average value of the spin variable $\langle y\rangle = \langle y_i\rangle\, \forall i$, which is independent of the element considered in the chain and is given by

$$\langle y\rangle = \langle y_i\rangle = \tanh\left(\frac{\ell f}{k_B T}\right) \quad \forall i. \tag{2.87}$$

An application of Eqs.(2.86) and (2.87) can be found in Fig.2.16. The force-extension curves have been plotted with dimensionless quantities and only one parameter defines the shape of the response, namely the elastic constant taken here into consideration through the dimensionless ratio $\frac{k\ell^2}{k_BT}$. It represents the ratio between the elastic (enthalpic) energy and the thermal energy. In these force-extension curves, we note a force plateau (for $f = 0$) corresponding to the synchronised switching (sometimes called cooperative) of the N units. This behaviour is confirmed by the average spin variable (which is independent

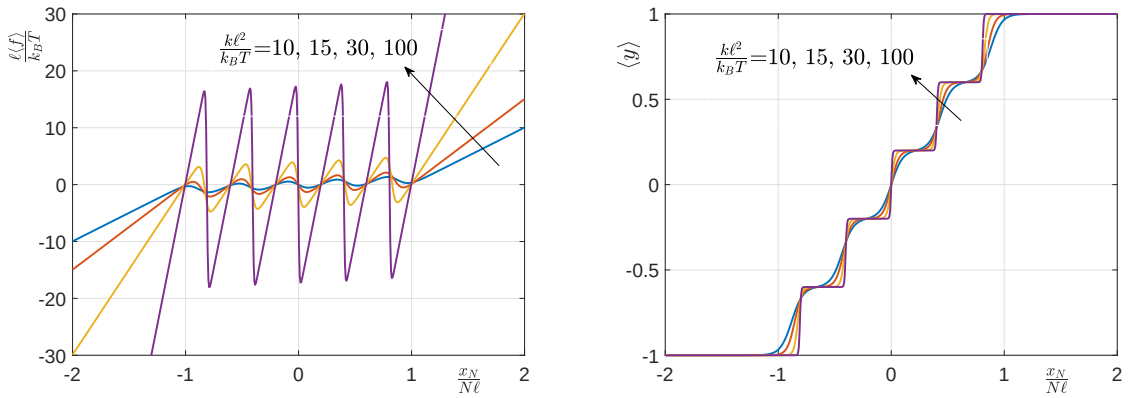

Figure 2.17 – Average force-extension curves and average spin variables (plotted by means of dimensionless quantities) for the Helmholtz ensemble with $N = 5$ and $\frac{k\ell^2}{k_BT}$=10, 15, 30, 100 [3].

of $\frac{k\ell^2}{k_BT}$), showing a transition from -1 to +1, at the same threshold force $f = 0$ as the previously mentioned plateau.

We can now introduce the second boundary condition corresponding to the Helmholtz ensemble. For imposing the isometric conditions, we consider the chain of bistable units with the two extremities tethered at the points $x_0 = 0$ and $x_N = x$, respectively. The total potential energy of the system can be therefore written as

$$U^H_{tot}(\vec{x}, \vec{y}; x_N) = \sum_{i=1}^{N} U(x_i - x_{i-1}, y_i), \tag{2.88}$$

where $x_N = x$ is the fixed extremity of the chain, $\vec{x} = (x_1, x_2, ..., x_{N-1})$ (continuous variables) and $\vec{y} = (y_1, y_2, ..., y_N)$ (discrete variables). In Eq.(2.88), the potential energy $U(x, y)$ of a single element is given in Eq.(2.82). The configurational partition function of this system can be written as

$$Z_H(x_N) = \sum_{y_1 \in \mathbf{S}} ... \sum_{y_N \in \mathbf{S}} \int_{\Re} ... \int_{\Re} e^{-\frac{U^H_{tot}}{k_BT}} \mathrm{d}x_1 ... \mathrm{d}x_{N-1}. \tag{2.89}$$

The two configurational partition functions Z_G and Z_H are related through a bilateral Laplace transform, and we have

$$Z_H(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} Z_G(-i\omega k_BT) \exp(i\omega x)\, \mathrm{d}\omega. \tag{2.90}$$

From Eq.(2.85), we have

$$Z_H(x) = \frac{1}{2\pi} \left(\frac{8\pi k_BT}{k}\right)^{\frac{N}{2}} \int_{-\infty}^{+\infty} \exp\left(-\frac{Nk_BT\omega^2}{2k}\right) \cos^N(\ell\omega) \exp(i\omega x)\, \mathrm{d}\omega, \tag{2.91}$$

and the calculation of the integral delivers

$$Z_H(x) = \frac{1}{2\pi}\left(\frac{2\pi k_B T}{k}\right)^{\frac{N}{2}} \sum_{t=0}^{N} \binom{N}{t} \sqrt{\frac{2k\pi}{Nk_BT}} \exp\left[-\frac{k}{2Nk_BT}(x+2t\ell-N\ell)^2\right]. \quad (2.92)$$

Now, we can evaluate the average value of the overall force $\langle f\rangle = -k_BT\partial/\partial x(\log Z_H)$ applied to the system and the average value of the spin variables $\langle y\rangle = \left\langle \frac{1}{N}\sum_{i=1}^{N} y_i \right\rangle$ describing the transitions, as follows

$$\langle f\rangle = \frac{\sum_{t=0}^{N}\binom{N}{t}\frac{k}{N}(x+2t\ell-N\ell)\exp\left[-\frac{k}{2Nk_BT}(x+2t\ell-N\ell)^2\right]}{\sum_{t=0}^{N}\binom{N}{t}\exp\left[-\frac{k}{2Nk_BT}(x+2t\ell-N\ell)^2\right]}, \quad (2.93)$$

and

$$\langle y\rangle = \frac{\frac{1}{N}\sum_{t=0}^{N}\binom{N}{t}(N-2t)\exp\left[-\frac{k}{2Nk_BT}(x+2t\ell-N\ell)^2\right]}{\sum_{t=0}^{N}\binom{N}{t}\exp\left[-\frac{k}{2Nk_BT}(x+2t\ell-N\ell)^2\right]}. \quad (2.94)$$

An example of application of Eqs.(2.93) and (2.94) can be found in Fig.2.17, where we show the force-extension response and the average spin variable for the Helmholtz ensemble. As before, the force-extension curves have been plotted with dimensionless quantities and only one parameter defines the shape of the response, namely the elastic constant taken here into consideration through the dimensionless ratio $\frac{k\ell^2}{k_BT}$. We observe that the force-extension curve is composed of a number of peaks corresponding to the non-synchronised (sequential) switching of the units. Sometimes, this behaviour is called non-cooperative in order to underline the independent transitions of the units. This is confirmed by the step-wise curve representing the average spin variable versus the chain extension. Each step corresponds to the switching of a unit induced by the increasing extension of the chain.

2.3.2 Complete probability densities in the Gibbs and the Helmholtz ensembles

We are interested here not only in the average value of the fluctuating quantities, but also in the complete probability densities. The knowledge of these more refined quantities allows for the determination of expected values of higher order such as variances, covariances and so on, very important to fully characterise the statistical properties of these systems. We define here the probability density of the system in the whole phase space within both the Gibbs and the Helmholtz statistical ensembles. These results will

be used in the following Sections to find the probability density of the specific quantities characterising the Gibbs and Helmholtz statistical ensembles.

Concerning the Gibbs ensemble, we can define the total energy of the system as

$$\begin{aligned} E_G(\vec{v},\vec{x},\vec{y};f) &= \sum_{i=1}^{N}\frac{1}{2}mv_i^2 + U_{tot}^G(\vec{x},\vec{y};f) \\ &= \sum_{i=1}^{N}\frac{1}{2}mv_i^2 + \sum_{i=1}^{N}U(x_i - x_{i-1}, y_i) - fx_N, \end{aligned} \tag{2.95}$$

where v_i is the velocity of the i-th particle of the chain and $\vec{v},\vec{x},\vec{y} \in \Re^N$ while $f \in \Re$. The complete probability density in the phase space is therefore given by the canonical distribution

$$\rho_G(\vec{v},\vec{x},\vec{y};f) = \frac{\exp\left[-\frac{1}{k_BT}E_G(\vec{v},\vec{x},\vec{y};f)\right]}{\left(\sqrt{\frac{2\pi k_BT}{m}}\right)^N Z_G(f)}, \tag{2.96}$$

where the term $\left(\sqrt{\frac{2\pi k_BT}{m}}\right)^N$ has been added to normalise the kinetic part of the Boltzmann factor and the configurational partition function $Z_G(f)$ is given in Eq.(2.85). Of course, we have that

$$\sum_{\vec{y}\in\{-1,+1\}^N}\int_{\vec{x}\in\Re^N}\int_{\vec{v}\in\Re^N}\rho_G(\vec{v},\vec{x},\vec{y};f)\mathrm{d}\vec{v}\mathrm{d}\vec{x} = 1 \forall f \in \Re. \tag{2.97}$$

Similarly, for the Helmholtz ensemble we can define the total energy as

$$E_H(\vec{v},\vec{x},\vec{y};x_N) = \sum_{i=1}^{N-1}\frac{1}{2}mv_i^2 + U_{tot}^H(\vec{x},\vec{y};x_N) = \sum_{i=1}^{N-1}\frac{1}{2}mv_i^2 + \sum_{i=1}^{N}U(x_i - x_{i-1}, y_i), \tag{2.98}$$

where, as before, v_i is the velocity of the i-th particle of the chain and $\vec{v},\vec{x} \in \Re^{N-1}$, $\vec{y} \in \Re^N$ while $x_N \in \Re$. In this case, the complete probability density in the phase space is given by the canonical distribution

$$\rho_H(\vec{v},\vec{x},\vec{y};x_N) = \frac{\exp\left[-\frac{1}{k_BT}E_H(\vec{v},\vec{x},\vec{y};x_N)\right]}{\left(\sqrt{\frac{2\pi k_BT}{m}}\right)^{N-1} Z_H(x_N)}, \tag{2.99}$$

where the term $\left(\sqrt{\frac{2\pi k_BT}{m}}\right)^{N-1}$ has been added to normalise the kinetic part of the Boltzmann factor and the configurational partition function $Z_H(x_N)$ is given in Eq.(2.92). Of

course, we have that

$$\sum_{\vec{y}\in\{-1,+1\}^N}\int_{\vec{x}\in\Re^{N-1}}\int_{\vec{v}\in\Re^{N-1}}\rho_H(\vec{v},\vec{x},\vec{y};x_N)\mathrm{d}\vec{v}\mathrm{d}\vec{x}=1\ \forall\ x_N\in\Re. \tag{2.100}$$

The two probability densities here described will be used to obtain a full statistics representing the behaviour of the two isotensional and isometric ensembles.

2.3.3 Probability density of the couple $(\dot{x}_N, x_N)$ versus f within the Gibbs ensemble

Since the force f is imposed within the Gibbs ensemble, we can measure the extension of the chain. The latter is a random variable that must be defined by its probability density, in order to have a complete description of its behaviour. Here, for the sake of completeness, we elaborate the probability density $\varrho_G(\dot{x}_N, x_N; f)$ for the couple $(\dot{x}_N, x_N)$, where we defined $\dot{x}_N = v_N$. In this case, to obtain the probability density, we have to sum or to integrate all the variables different from v_N and x_N in the complete density probability defined in Eq.(2.96). It means that we can write

$$\begin{aligned}\varrho_G(\dot{x}_N,x_N;f) &= \sum_{\vec{y}\in\{-1,+1\}^N}\int_{x_1}...\int_{x_{N-1}}\int_{v_1}...\int_{v_{N-1}}\rho_H(\vec{v},\vec{x},\vec{y};x_N)\\ &\quad\times\mathrm{d}v_{N-1}...\mathrm{d}v_1\mathrm{d}x_{N-1}...\mathrm{d}x_1.\end{aligned} \tag{2.101}$$

Now, it is not difficult to recognise that the integral over the positions x_1,...,x_{N-1} immediately leads to the configurational partition function of the Helmholtz ensemble and the integral over the velocities v_1,...,v_{N-1} can be directly calculated with the classical Gaussian integral. Eventually, we obtain

$$\varrho_G(\dot{x}_N,x_N;f)=\sqrt{\frac{m}{2\pi k_BT}}\exp\left(\frac{1}{2}m\dot{x}_N^2\right)\frac{\exp\left(\frac{fx_N}{k_BT}\right)Z_H(x_N)}{Z_G(f)}. \tag{2.102}$$

This is the most important result of this Section, which represents the probability density of the couple $(\dot{x}_N, x_N)$ for any value of the applied force f within the Gibbs ensemble.

We remark that this probability density can be factorised in two terms representing the density of $\dot{x}_N$ and the density of x_N. The first factor simply corresponds to the Maxwell distribution for the one-dimensional velocity

$$\varrho_G(\dot{x}_N)=\sqrt{\frac{m}{2\pi k_BT}}\exp\left(\frac{1}{2}m\dot{x}_N^2\right). \tag{2.103}$$

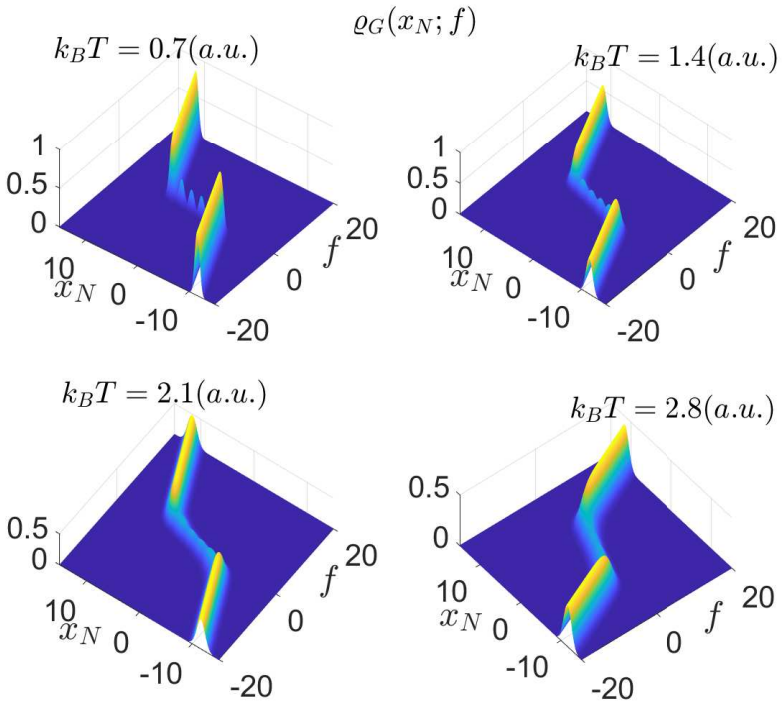

Figure 2.18 – Three-dimensional representation of the Gibbs density $\varrho_G(x_N; f)$ (see Eq.(2.104)) obtained with $N = 5$, $\ell = 1$ (a.u.), $k = 15$ (a.u.) and k_BT=0.7, 1.4, 2.1, 2.8 (a.u.) [3].

On the other hand, it is interesting to observe that the second configurational term depends on the ratio between the two partition functions

$$\varrho_G(x_N; f) = \exp\left(\frac{f x_N}{k_B T}\right) \frac{Z_H(x_N)}{Z_G(f)}. \tag{2.104}$$

This configurational density is correctly normalised because of the Laplace integral relationship between Gibbs and Helmholtz partition functions. The explicit form of $\varrho_G(\dot{x}_N, x_N; f)$ can be found by using the results given in Eqs.(2.85) and (2.92).

An example of application of the results obtained here is given in Figs.2.18, 2.19, 2.20 and 2.21. Since the kinetic component $\varrho_G(\dot{x}_N)$ is simply given by the Maxwell distribution, we focus our attention to the configurational part given by $\varrho_G(x_N; f)$. Accordingly, in Figs.2.18 and 2.19, we show a three-dimensional and a two-dimensional representation of the Gibbs density as function of x_N force f. These results are represented for four different levels of thermal agitation in order to understand the effects of the disorder on the switching behaviour between the states. The parameters used in this study are $N = 5$, $\ell = 1$ (a.u.), $k = 15$ (a.u.) and k_BT=0.7, 1.4, 2.1, 2.8 (a.u.). It is interesting to observe

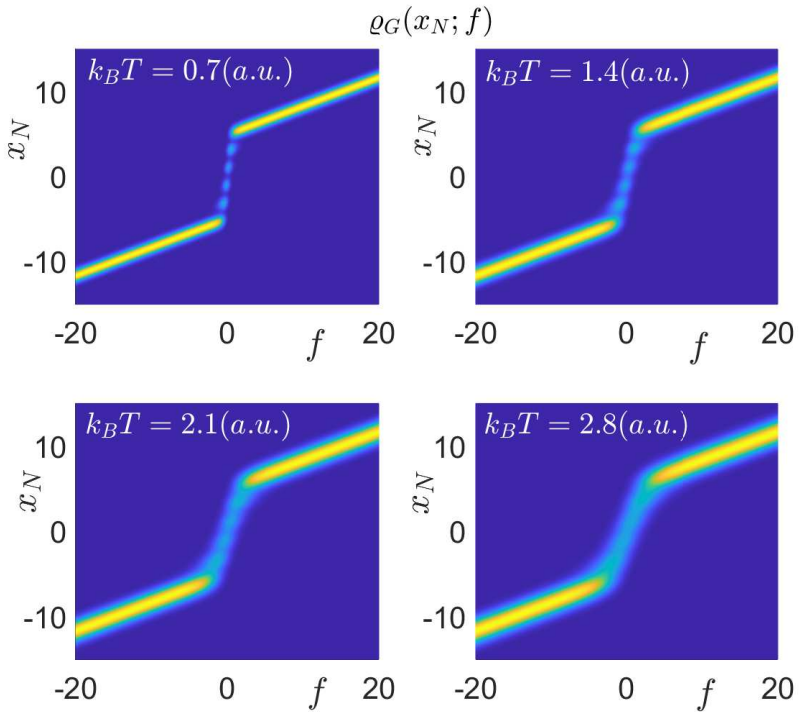

Figure 2.19 – Two-dimensional representation of the Gibbs density $\varrho_G(x_N; f)$ (see Eq.(2.104)) obtained with $N = 5$, $\ell = 1$ (a.u.), $k = 15$ (a.u.) and k_BT=0.7, 1.4, 2.1, 2.8 (a.u.) [3].

that, in spite of the simple shape of the force-extension response characterised by a force plateau at $f = 0$ with a synchronised switching of the units, the probability density of the quantity x_N is multimodal for the force range characterising the transition region. Indeed, in order to obtain the probability density of x_N for a given applied f, we have to section the plots in Figs.2.18 and 2.19 with a plane parallel to the x_N-axis and, at the same time, perpendicular to the f-axis. So doing, in the central transition region, we can observe the emergence of a series of peaks in the probability density confirming its multimodal character. This can be observed in Fig.2.20, where we plotted several curves $\varrho_G(x_N; f)$ (see Eq.(2.104)), for different values of the applied force f. We can observe the symmetric and multimodal profile of the probability density for $f = 0$ (at the center of the transition region) and the asymmetric and monomodal shape of the density for a large applied force (out of the transition region). We remark the multimodal character of the probability density of x_N in spite of the simple force plateau observed in the force-extension response. To conclude this analysis, we underline that the knowledge of the full statistics for the system allows us to determine all possible expected values. As an example, we show in

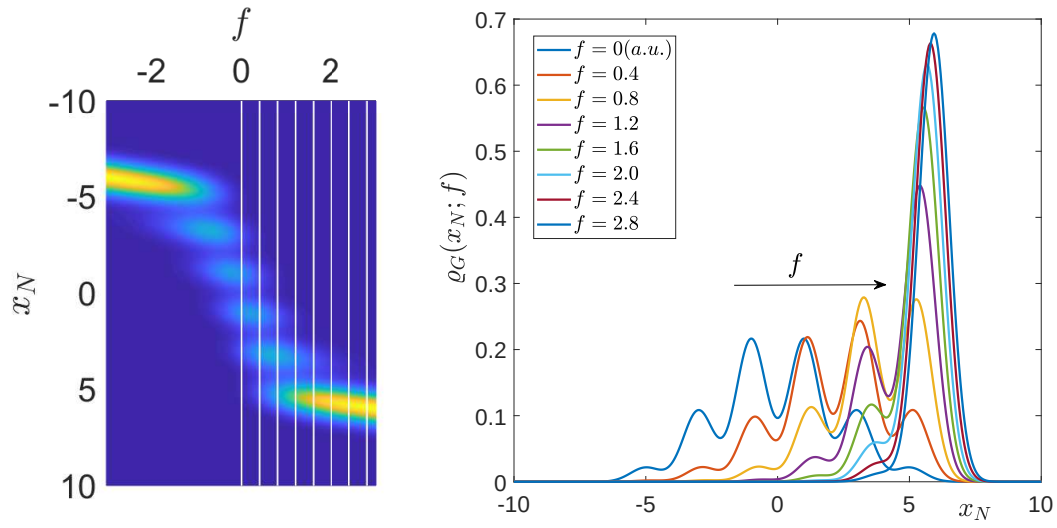

Figure 2.20 – Examples of multimodal curves obtained through the Gibbs density $\varrho_G(x_N; f)$ (see Eq.(2.104)). On the left panel, the two-dimensional representation of the Gibbs density is shown with the cuts corresponding to the curves plotted on the right panel. We used $N = 5$, $\ell = 1$ (a.u.), $k = 15$ (a.u.), k_BT=1 (a.u.) and different values of the applied force f, as indicated in the legend [3].

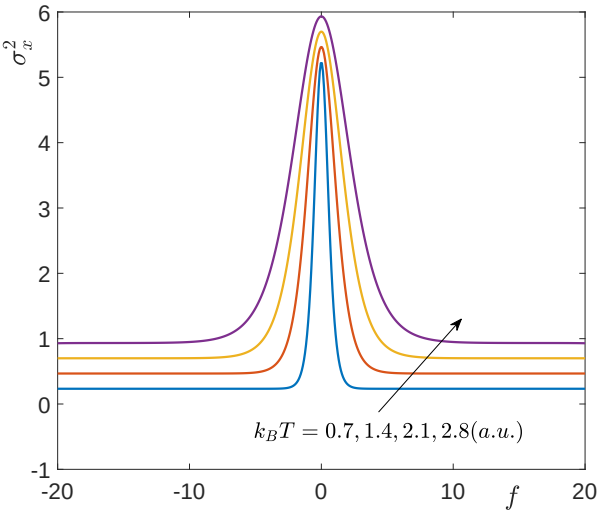

Figure 2.21 – Variance of x_N obtained by the Gibbs density $\varrho_G(x_N; f)$. As before, we used $N = 5$, $\ell = 1$ (a.u.), $k = 15$ (a.u.) and k_BT=0.7, 1.4, 2.1, 2.8 (a.u.) [3].

Fig.2.21 the behaviour of the variance of the position in terms of the applied force f and the thermal energy k_BT. We note that the variance is higher in the transition region,

where the two states of each unit can coexist. Moreover, we observe a larger variance for higher temperatures, as expected. Finally, we also note that the multimodal character of the probability density is smeared out by the integration process applied to calculate the variances. This behaviour will be shown to be dual with respect to the Helmholtz ensemble response, which is the subject of the next Section.

2.3.4 Probability density of the couple $(\dot{f}, f)$ versus x_N within the Helmholtz ensemble

The problem of finding the probability density for f and $\dot{f}$ when x_N is imposed is more complicated since, in this case, the variables f and $\dot{f}$ do not belong to the phase space and, therefore, we can not integrate the superfluous variables in order to get the searched density. To cope with this problem, we first introduce the standard technique to deal with a function of random variable. We suppose to have two random variables x and y, linked by a function $\mathsf{y} = g(\mathsf{x})$. If $F_{\mathsf{x}}(x)$ and $f_{\mathsf{x}}(x)$ are distribution function and probability density of the random variable x, we search for the same quantities $F_{\mathsf{y}}(y)$ and $f_{\mathsf{y}}(y)$ for $\mathsf{y} = g(\mathsf{x})$. We use the symbol ξ for the elements of the probability space and we can write

$$F_{\mathsf{x}}(x) = \Pr\left\{\xi : \mathsf{x}(\xi) \leq x\right\}, \tag{2.105}$$

and

$$f_{\mathsf{x}}(x) = \frac{\mathrm{d}}{\mathrm{d}x} F_{\mathsf{x}}(x). \tag{2.106}$$

Moreover, we can state that

$$\begin{aligned} F_{\mathsf{y}}(y) &= \Pr\left\{\xi : \mathsf{y}(\xi) \leq y\right\} = \Pr\left\{\xi : g(\mathsf{x}(\xi)) \leq y\right\} \\ &= \int_{g(x)\leq y} f_{\mathsf{x}}(x)\mathrm{d}x = \int_{-\infty}^{+\infty} \mathbf{1}(y - g(x)) f_{\mathsf{x}}(x)\mathrm{d}x, \end{aligned} \tag{2.107}$$

where $\mathbf{1}(z)$ is the Heaviside step function. Therefore, we can obtain the probability density of $\mathsf{y} = g(\mathsf{x})$ by differentiation

$$\begin{aligned} f_{\mathsf{y}}(y) &= \frac{\mathrm{d}}{\mathrm{d}y} F_{\mathsf{y}}(y) \\ &= \frac{\mathrm{d}}{\mathrm{d}y} \int_{-\infty}^{+\infty} \mathbf{1}(y - g(x)) f_{\mathsf{x}}(x)\mathrm{d}x \\ &= \int_{-\infty}^{+\infty} \delta(y - g(x)) f_{\mathsf{x}}(x)\mathrm{d}x, \end{aligned} \tag{2.108}$$

where we have introduced the Dirac delta function $\delta(z)$. This method based on the delta functions can be used to approach the problem of finding the Helmholtz probability density. To apply this technique, we need to write the variables f and $\dot{f}$ in terms of the variables of the phase space of the system. Given the total potential energy $U^H_{tot}(\vec{x},\vec{y};x_N)=\sum_{i=1}^{N}U(x_i-x_{i-1},y_i)$, we can simply write

$$f=\frac{\partial U^H_{tot}}{\partial x_N}=k(x_N-x_{N-1}-y_N\ell) \tag{2.109}$$

and

$$\begin{aligned}\dot{f} &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial U^H_{tot}}{\partial x_N}=\sum_{i=1}^{N-1}\frac{\partial^2 U^H_{tot}}{\partial x_i\partial x_N}\frac{\mathrm{d}x_i}{\mathrm{d}t}\\ &= \sum_{i=1}^{N-1}\frac{\partial^2 U^H_{tot}}{\partial x_i\partial x_N}v_i=\frac{\partial^2 U^H_{tot}}{\partial x_{N-1}\partial x_N}v_{N-1}=-kv_{N-1}.\end{aligned} \tag{2.110}$$

Now, given the complete probability density $\rho_H(\vec{v},\vec{x},\vec{y};x_N)$, we can obtain the density for the desired variables f and $\dot{f}$ as follows

$$\begin{aligned}\varrho_H(\dot{f},f;x_N) &= \sum_{\vec{y}}\int_{\vec{x}\in\Re^{N-1}}\int_{\vec{v}\in\Re^{N-1}}\delta\left(f-\frac{\partial U^H_{tot}}{\partial x_N}\right)\delta\left(\dot{f}-\frac{\partial^2 U^H_{tot}}{\partial x_{N-1}\partial x_N}v_{N-1}\right)\\ &\quad\times\rho_H(\vec{v},\vec{x},\vec{y};x_N)\mathrm{d}\vec{x}\mathrm{d}\vec{v}.\end{aligned} \tag{2.111}$$

This expression can be simplified delivering

$$\begin{aligned}\varrho_H(\dot{f},f;x_N) &= \sum_{\vec{y}}\int_{\vec{x}\in\Re^{N-1}}\int_{\vec{v}\in\Re^{N-1}}\delta\left(f-kx_N+kx_{N-1}+ky_N\ell\right)\delta\left(\dot{f}+kv_{N-1}\right)\\ &\quad\times\frac{\exp\left[-\frac{1}{k_BT}E_H(\vec{v},\vec{x},\vec{y};x_N)\right]}{\left(\sqrt{\frac{2\pi k_BT}{m}}\right)^{N-1}Z^N_H(x_N)}\mathrm{d}\vec{x}\mathrm{d}\vec{v},\end{aligned} \tag{2.112}$$

where we used the notation $Z_H(x_N)=Z^N_H(x_N)$ in order to specify that the Helmholtz partition function corresponds to a system with N units. Indeed, in the following calculations, we will also need the same function calculated for a system with $N-1$ units.

The elaboration of $\varrho_H(\dot{f}, f; x_N)$ can be continued as follows

$$\begin{aligned}
\varrho_H(\dot{f}, f; x_N) &= \frac{1}{\left(\sqrt{\frac{2\pi k_B T}{m}}\right)^{N-1} Z_H^N(x_N)} \\
&\times \sum_{\vec{y}} \int_{\vec{x}\in\Re^{N-1}} \int_{\vec{v}\in\Re^{N-1}} \delta\left(f - kx_N + kx_{N-1} + ky_N\ell\right) \delta\left(\dot{f} + kv_{N-1}\right) \\
&\times \exp\left(-\frac{m}{2k_BT}\sum_{i=1}^{N-1} v_i^2\right) \exp\left[-\frac{k}{2k_BT}\sum_{i=1}^{N}(x_i - x_{i-1} - y_i\ell)^2\right] \mathrm{d}\vec{x}\mathrm{d}\vec{v} \\
&= \frac{1}{\left(\sqrt{\frac{2\pi k_B T}{m}}\right)^{N-1} Z_H^N(x_N)} \int_{\Re^{N-2}} \exp\left(-\frac{m}{2k_BT}\sum_{i=1}^{N-2} v_i^2\right) \mathrm{d}v_1 ... \mathrm{d}v_{N-2} \\
&\times \int_{\Re} \frac{1}{k}\delta\left(\frac{1}{k}\dot{f} + v_{N-1}\right) \exp\left(-\frac{m}{2k_BT} v_{N-1}^2\right) \mathrm{d}v_{N-1} \\
&\times \sum_{\vec{y}} \int_{\vec{x}\in\Re^{N-1}} \frac{1}{k}\delta\left(\frac{f}{k} - x_N + x_{N-1} + y_N\ell\right) \exp\left[-\frac{k}{2k_BT}\right. \\
&\left.\times \sum_{i=1}^{N-2}(x_i - x_{i-1} - y_i\ell)^2\right] \exp\left[-\frac{k}{2k_BT}(x_{N-1} - x_{N-2} - y_{N-1}\ell)^2\right] \\
&\times \exp\left[-\frac{k}{2k_BT}(x_N - x_{N-1} - y_N\ell)^2\right] \mathrm{d}\vec{x}, \qquad (2.113)
\end{aligned}$$

where we used the property $\delta(ax) = \frac{1}{|a|}\delta(x)$. We remember now that $\int_{-\infty}^{+\infty} \exp(-\alpha x^2) = \sqrt{\frac{\pi}{\alpha}}$ for $\alpha > 0$, we perform the integrals of the delta functions over v_{N-1} and x_{N-1}, and we get

$$\begin{aligned}
\varrho_H(\dot{f}, f; x_N) &= \frac{1}{\sqrt{\frac{2\pi k_B T}{m}} Z_H^N(x_N)} \frac{1}{k} \exp\left(-\frac{m}{2k^2 k_B T}\dot{f}^2\right) \exp\left(-\frac{1}{2kk_BT} f^2\right) \\
&\times \sum_{\vec{y}} \int_{\vec{x}\in\Re^{N-2}} \frac{1}{k} \exp\left[-\frac{k}{2k_BT}\sum_{i=1}^{N-2}(x_i - x_{i-1} - y_i\ell)^2\right] \\
&\times \exp\left[-\frac{k}{2k_BT}(x_N - \frac{1}{k}f - y_N\ell - x_{N-2} - y_{N-1}\ell)^2\right] \\
&\times \mathrm{d}x_1 ... \mathrm{d}x_{N-2}. \qquad (2.114)
\end{aligned}$$

We can now recall the explicit definition of $Z_H^N(x_N)$ (see Eq.(2.89)), and we also introduce the exact expression for $Z_H^{N-1}(x_{N-1})$

$$\begin{aligned} Z_H^N(x_N) &= \sum_{y_1\in\mathbf{S}} ... \sum_{y_N\in\mathbf{S}} \int_{\Re} ... \int_{\Re} \exp\left[-\frac{k}{2k_BT}\sum_{i=1}^{N}(x_i - x_{i-1} - y_i\ell)^2\right] \\ &\quad \times \mathrm{d}x_1 ... \mathrm{d}x_{N-1}, \end{aligned} \tag{2.115}$$

$$\begin{aligned} Z_H^{N-1}(x_{N-1}) &= \sum_{y_1\in\mathbf{S}} ... \sum_{y_{N-1}\in\mathbf{S}} \int_{\Re} ... \int_{\Re} \exp\left[-\frac{k}{2k_BT}\sum_{i=1}^{N-1}(x_i - x_{i-1} - y_i\ell)^2\right] \\ &\quad \times \mathrm{d}x_1 ... \mathrm{d}x_{N-2}. \end{aligned} \tag{2.116}$$

So, in Eq.(2.114), we can identify the partition function $Z_H^{N-1}(x_{N-1})$ calculated for $x_{N-1} = x_N - \frac{1}{k}f - y_N\ell$, by obtaining

$$\begin{aligned} \varrho_H(\dot{f}, f; x_N) &= \frac{1}{k^2\sqrt{\frac{2\pi k_BT}{m}}} \exp\left(-\frac{m}{2k^2k_BT}\dot{f}^2\right) \exp\left(-\frac{1}{2kk_BT}f^2\right) \\ &\quad \times \frac{\sum_{y_N} Z_H^{N-1}\left(x_N - \frac{1}{k}f - y_N\ell\right)}{Z_H^N(x_N)}, \end{aligned} \tag{2.117}$$

or, equivalently,

$$\begin{aligned} \varrho_H(\dot{f}, f; x_N) &= \frac{1}{k^2\sqrt{\frac{2\pi k_BT}{m}}} \exp\left(-\frac{m}{2k^2k_BT}\dot{f}^2\right) \exp\left(-\frac{1}{2kk_BT}f^2\right) \\ &\quad \times \frac{Z_H^{N-1}\left(x_N - \frac{1}{k}f - \ell\right) + Z_H^{N-1}\left(x_N - \frac{1}{k}f + \ell\right)}{Z_H^N(x_N)}. \end{aligned} \tag{2.118}$$

This is the final result for the probability density within the Helmholtz ensemble. It is interesting to observe that it can be written in terms of the two partition functions $Z_H^N(x_N)$ and $Z_H^{N-1}(x_{N-1})$, corresponding to systems of size N and $N-1$, respectively.

We can split this probability density in two independent components describing separately $\dot{f}$ and f, as follows

$$\varrho_H(\dot{f}) = \frac{1}{k\sqrt{\frac{2\pi k_BT}{m}}} \exp\left(-\frac{m}{2k^2k_BT}\dot{f}^2\right), \tag{2.119}$$

$$\varrho_H(f; x_N) = \frac{1}{k} \exp\left(-\frac{1}{2kk_BT}f^2\right) \frac{Z_H^{N-1}\left(x_N - \frac{1}{k}f - \ell\right) + Z_H^{N-1}\left(x_N - \frac{1}{k}f + \ell\right)}{Z_H^N(x_N)}, \tag{2.120}$$

and it is not difficult to prove the normalisation of the two results.

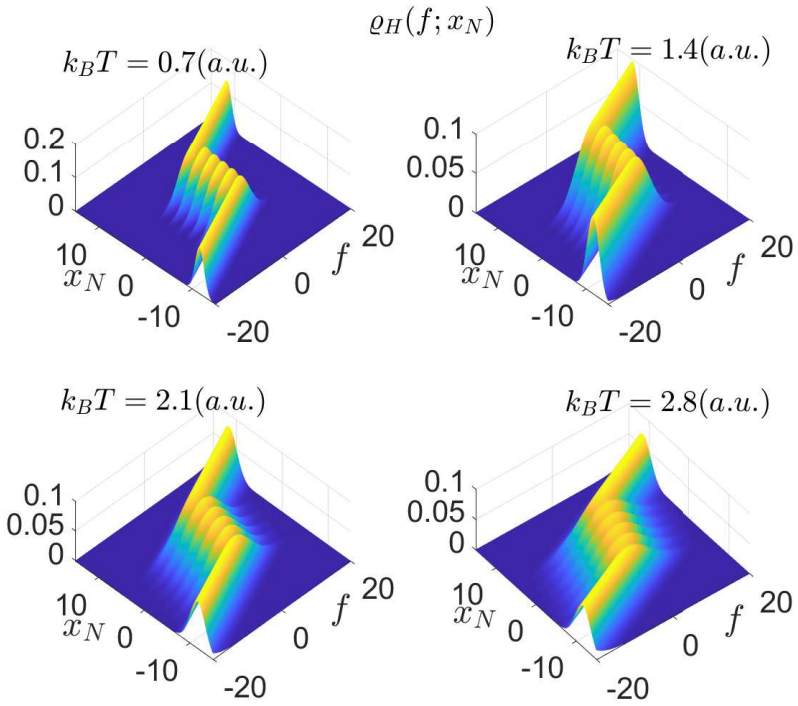

Figure 2.22 – Three-dimensional representation of the Helmholtz density $\varrho_H(f;x_N)$ (see Eq.(2.120)) obtained with $N = 5$, $\ell = 1$ (a.u.), $k = 15$ (a.u.) and k_BT=0.7, 1.4, 2.1, 2.8 (a.u.) [3].

A numerical application of these results can be found in Figs.2.22, 2.23, 2.24 and 2.25. Similarly to the Gibbs analysis, also in this case, we observe that the kinetic part of the probability density $\varrho_H(\dot{f})$ is a simple Gaussian function and therefore we study in more detail the configurational density $\varrho_H(f;x_N)$. Coherently with this planning, in Figs.2.22 and 2.23, we show the three-dimensional and the two-dimensional representation of the Helmholtz density as function of f and for a prescribed extension x_N. As before, the results have been obtained for four different temperatures to observe the effects of the thermal agitation on the transition processes. The parameters used in this study are the same already adopted for the Gibbs analysis, namely $N = 5$, $\ell = 1$ (a.u.), $k = 15$ (a.u.) and k_BT=0.7, 1.4, 2.1, 2.8 (a.u.). We give here a description of the behaviour of the system within the Helmholtz ensemble which is exactly dual with respect to the response of the Gibbs ensemble. Indeed, we observe that in spite of the saw-tooth shape of the force-extension response, the probability density of f is quite always monomodal. More precisely, it can be bimodal only with some sets of parameters and only for forces being in the transition region between two peaks of the force-extension curve. Anyway,

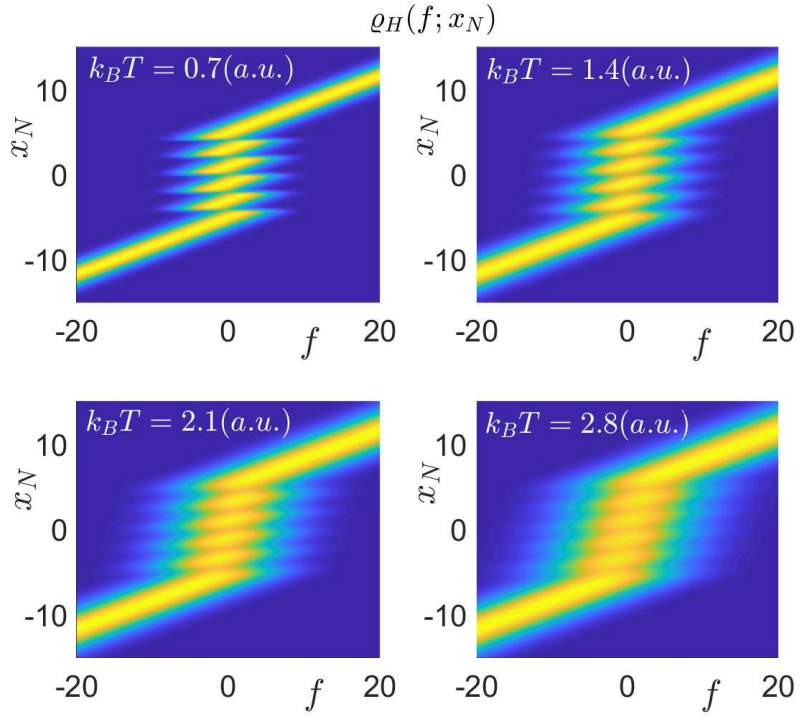

Figure 2.23 – Two-dimensional representation of the Helmholtz density $\varrho_H(f;x_N)$ (see Eq.(2.120)) obtained with $N = 5$, $\ell = 1$ (a.u.), $k = 15$ (a.u.) and k_BT=0.7, 1.4, 2.1, 2.8 (a.u.) [3].

we can affirm that this density is monomodal in the most cases of practical interest. To better explain this point, we observe that in order to obtain the probability density of f for a prescribed x_N, we have to section the plots in Figs.2.22 and 2.23 with a plane parallel to the f-axis and, at the same time, perpendicular to the x_N-axis. By performing this operation, in spite of the complex shape of $\varrho_H(f;x_N)$, we get monomodal functions (with the exceptions discussed above). This can be observed in Fig.2.24, where we plotted several curves $\varrho_H(f;x_N)$ for different values of the prescribed extension x_N. As before, we remark that the knowledge of the full probability density for the Helmholtz case can be used to determine the expected values of higher order. As an example, in Fig.2.25 we plotted the variance of the force f, necessary to impose the extension x_N. Interestingly enough, the variance is an increasing function of the temperature, as expected, and shows some peaks in correspondence to the switching of state of each unit. This is coherent with the general idea that the variance of the physical quantities is larger in proximity to a phase transition. Again, we underline the dual behaviour of the Gibbs and Helmholtz ensembles. Indeed, while the variance for the Gibbs case is given by a single peak corresponding to

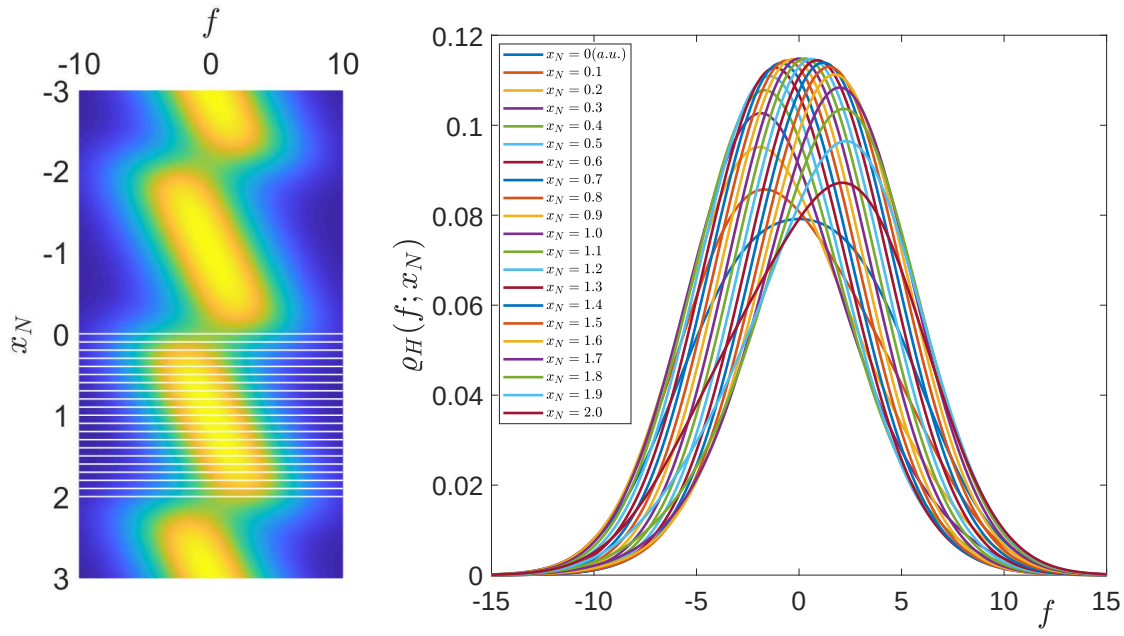

Figure 2.24 – Examples of monomodal curves obtained through the Helmholtz density $\varrho_H(f;x_N)$ (see Eq.(2.120)). On the left panel, the two-dimensional representation of the Helmholtz density is shown with the cuts corresponding to the curves plotted on the right panel. We used $N = 5$, $\ell = 1$ (a.u.), $k = 15$ (a.u.), k_BT=1 (a.u.) and different values of the prescribed position x_N, as indicated in the legend [3].

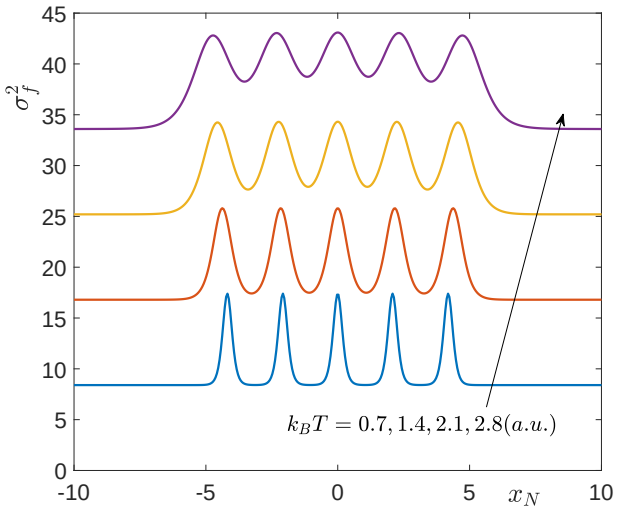

Figure 2.25 – Variance of f obtained by the Helmholtz density $\varrho_H(f;x_N)$. As before, we used $N = 5$, $\ell = 1$ (a.u.), $k = 15$ (a.u.) and k_BT=0.7, 1.4, 2.1, 2.8 (a.u.) [3].

the synchronised transition of the units, for the Helmholtz ensemble, we have a peak for each transition, underlying the sequential behaviour of this process.

2.3.5 Final comparison

The results obtained for the specific case of a chain of two-state elements show the emergence of an intriguing duality between the two ensembles. For the isotensional condition, the force-extension curve is monotone with a characteristic force plateau and the density $\varrho(x_N; f)$ is multimodal in the transition region (near $x_N = 0$ and $f = 0$). Conversely, for the isometric condition, the force-extension curve is composed of a series of peaks while the density $\varrho(f; x_N)$ is simply monomodal. This duality is also reflected in the behaviour of the variances of these processes. In the Gibbs ensemble, we obtained a monomodal variance σ_x^2 with the symmetric peak at $f = 0$, whereas in the Helmholtz ensemble we obtained a multimodal variance σ_f^2 with a peak for each transition value of x_N. Of course, the peaks of variance must be explained through large fluctuations characterising the switching of the units states (classically, it is typical for the phase transitions).

Chapter 3

Extensible two-state freely jointed chain

3.1 Introduction

In Chapter 2, we introduced the spin variable approach. This method is used to study the thermo-elastic behaviour of chains or systems made up of bistable (or two-state) units, as previously described in Ref. [113]. In this thesis, several extensions are added to the spin variable model, in order to bring it closer to the reality of some biological macromolecules. The first generalisation added to the bistable freely jointed chain concerns the extensibility of the bonds between the chain units. Indeed, the classical freely jointed chain considers rigid segments of fixed length, and adding the elasticity is not a straightforward task. In this Chapter, we find the analytic solution of the partition functions for chains made of bistable units, taking account of extensibility, within both the Gibbs and the Helmholtz ensembles. Thus, we obtain the complete solution for a specific case, namely the two-state freely jointed chain with extensible units. It is necessary to take extensibility into account, since the finite elastic constant of the bonds between the units plays a major role in defining the force-extension response in both isotensional and isometric ensembles. Particularly, within the Helmholtz ensemble, the peak-to-peak force of the sawtooth-like curve strongly depends on the elastic constant. As this quantity is often considered as the main output of the force spectroscopy measurements, it is important to study the case with a finite value of the elastic stiffness. From the mathematical point of view, the most difficult issue concerns the calculation of the Helmholtz partition function. Eventually, its closed form expression has been found in terms of the Hermite polynomials, suitably generalised to negative indices.

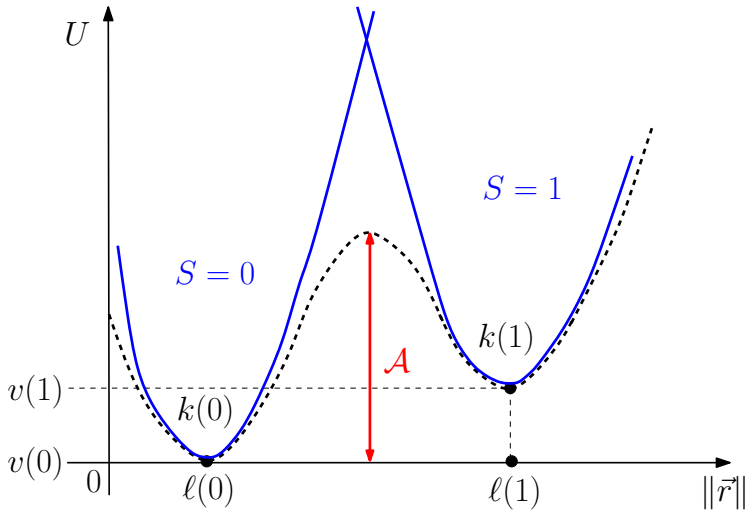

Figure 3.1 – Potential energy of a single unit of the bistable freely jointed chain with extensible bonds (dashed black curve). The potential wells are approximated through two parabolic profiles identified by $S = 0$ and 1 (solid blues curves) [1].

3.2 Two-state freely jointed chain with extensible units: the Gibbs ensemble

We take into consideration a two-state freely jointed chain composed of N domains. Here, instead of considering the units with an infinite elastic stiffness like in Chapter 2, we try to consider a finite elasticity for the bistable domains. The potential energy function of real domains is represented by a function with two minima, corresponding to the folded and unfolded configurations of the domain (dashed black curve in Fig.3.1). The spin variable approach consists in considering two quadratic potential curves, approximating the wells of this system and introducing a spin variable S_i for each unit, useful to identify the potential well explored during the system evolution (see again Fig.3.1). As previously discussed, the introduced spin variables belong to the phase space of the system and they must be considered as standard dynamic variables of the statistical mechanics. In other words, the switching of each spin variable between its two values is controlled by the boundary conditions applied to the system, here defined by the force $\vec{f}$ (Gibbs ensemble). We start the analysis of the Gibbs ensemble by writing the extended Hamiltonian of the

system as

$$\begin{aligned} H_G &= \sum_{i=1}^{N} \frac{\vec{p}_i \cdot \vec{p}_i}{2m} + \sum_{i=1}^{N} v(S_i) - \mu \sum_{i=1}^{N} S_i \\ &\quad + \sum_{i=1}^{N} \frac{1}{2} k(S_i) \left[\|\vec{r}_i - \vec{r}_{i-1}\| - \ell(S_i) \right]^2 - \vec{f} \cdot \vec{r}_N, \end{aligned} \tag{3.1}$$

where $S_i \in \{0, 1\}$ are the spin variables, m is the mass of each unit, $\vec{r}_i$ and $\vec{p}_i$ are positions and linear momenta of the units, and $\vec{f}$ is the force applied to the last element of the chain. The quantities $v(S_i)$, $k(S_i)$ and $\ell(S_i)$ stand for the basal energy, the elastic stiffness and the equilibrium length of the potential wells, respectively (see Fig.3.1 for details). Moreover, we introduce $\vec{q} = (\vec{r}_1, ..., \vec{r}_N)$ as the generalised vector containing all positions, $\vec{p} = (\vec{p}_1, ..., \vec{p}_N)$ as the generalised vector containing all momenta and $\vec{S} = (S_1, ..., S_N)$ as the generalised vector containing all spin variables. For the sake of simplicity, we assume that all chemical potentials μ_i take the same value μ. Since the dynamical variables are $\vec{q}, \vec{p}$ and $\vec{S}$, the partition function is given by

$$\begin{aligned} Z_G &= \sum_{S_1 \in \{0,1\}} ... \sum_{S_N \in \{0,1\}} \int_{\mathbb{R}^{3N}} \int_{\mathbb{R}^{3N}} \exp\left[-\frac{H_G(\vec{p}, \vec{q}, \vec{s})}{k_B T} \right] \mathrm{d}\vec{p} \mathrm{d}\vec{q} \\ &= \sum_{\vec{S}} \left[\int_{\mathbb{R}^3} \exp\left(-\frac{1}{2mk_BT} \vec{p} \cdot \vec{p} \right) \mathrm{d}\vec{p} \right]^N \\ &\quad \times \int_{\mathbb{R}^3} ... \int_{\mathbb{R}^3} \exp\left[-\sum_{i=1}^{N} \frac{v(S_i) - \mu S_i}{k_B T} + \frac{\vec{f} \cdot \vec{r}_N}{k_B T} \right. \\ &\quad \left. - \sum_{i=1}^{N} \frac{k(S_i)}{2k_B T} \left[\|\vec{r}_i - \vec{r}_{i-1}\| - \ell(S_i) \right]^2 \right] \mathrm{d}\vec{r}_1 ... \mathrm{d}\vec{r}_N. \end{aligned} \tag{3.2}$$

To begin the simplification, we remember that

$$\int_{\mathbb{R}^3} \exp\left(-\frac{1}{2mk_BT} \vec{p} \cdot \vec{p} \right) \mathrm{d}\vec{p} = (2\pi m k_B T)^{3/2}. \tag{3.3}$$

Then, we perform the change of variables $\vec{\xi}_1 = \vec{r}_1 - \vec{r}_0$, ..., $\vec{\xi}_N = \vec{r}_N - \vec{r}_{N-1}$. We have that $\mathrm{d}\vec{r}_1 ... \mathrm{d}\vec{r}_N = \mathrm{d}\vec{\xi}_1 ... \mathrm{d}\vec{\xi}_N$. By fixing $\vec{r}_0 = 0$ and knowing that $\vec{r}_N = \sum_{i=1}^{N} \vec{\xi}_i$, we obtain

$$\begin{aligned} Z_G &= (2\pi m k_B T)^{3N/2} \sum_{\vec{S}} \int_{\mathbb{R}^3} ... \int_{\mathbb{R}^3} \exp\left[-\sum_{i=1}^{N} \frac{v(S_i) - \mu S_i}{k_B T} - \sum_{i=1}^{N} \frac{k(S_i)}{2k_B T} \right. \\ &\quad \left. \times \left[\|\vec{\xi}_i\| - \ell(S_i) \right]^2 + \frac{1}{k_B T} \vec{f} \cdot \sum_{i=1}^{N} \vec{\xi}_i \right] \mathrm{d}\vec{\xi}_1 ... \mathrm{d}\vec{\xi}_N \\ &= (2\pi m k_B T)^{3N/2} \sum_{\vec{S}} \prod_{i=1}^{N} \int_{\mathbb{R}^3} \exp\left[-\frac{v(S_i) - \mu S_i}{k_B T} - \frac{k(S_i)}{2k_B T} \left[\|\vec{\xi}_i\| - \ell(S_i) \right]^2 \right. \\ &\quad \left. + \frac{\vec{f} \cdot \vec{\xi}_i}{k_B T} \right] \mathrm{d}\vec{\xi}_i. \end{aligned} \tag{3.4}$$

It is important to remark that within the Gibbs ensemble, the elements of the chain do not interact and this point leads to a factorised partition function, namely in the form of a power with exponent N. We can further simplify the calculation by means of the spherical coordinates $\vec{\xi} = (\xi \cos\varphi \sin\vartheta, \xi \sin\varphi \sin\vartheta, \xi \cos\vartheta)$. Because of the spherical symmetry of the problem, we set $\vec{f} = (0, 0, f)$. Since $\mathrm{d}\vec{\xi} = \xi^2 \sin\vartheta \mathrm{d}\xi \mathrm{d}\varphi \mathrm{d}\vartheta$, $\|\vec{\xi}\| = \xi$ and $\vec{f} \cdot \vec{\xi} = f\xi \cos\vartheta$, we get the following simpler form of the partition function

$$\begin{aligned} Z_G &= (2\pi m k_B T)^{3N/2} \left[\sum_{S} \int_{\mathbb{R}^3} \exp\left(-\frac{v(S) - \mu S}{k_B T} - \frac{k(S)}{2k_B T} \left[\|\vec{\xi}\| - \ell(S) \right]^2 \right. \right. \\ &\quad \left. \left. + \frac{\vec{f} \cdot \vec{\xi}}{k_B T} \right) \mathrm{d}\vec{\xi} \right]^N \\ &= (2\pi m k_B T)^{3N/2} \left[\sum_{S} \exp\left(-\frac{v(S) - \mu S}{k_B T} \right) \int_0^{+\infty} \int_0^{2\pi} \int_0^{\pi} \exp\left(-\frac{k(S)}{2k_B T} \right. \right. \\ &\quad \left. \left. \times \left[\xi - \ell(S) \right]^2 + \frac{f\xi \cos\theta}{k_B T} \right) \xi^2 \sin\theta \mathrm{d}\theta \mathrm{d}\varphi \mathrm{d}\xi \right]^N. \end{aligned} \tag{3.5}$$

We know that $\int_0^{2\pi} \mathrm{d}\varphi = 2\pi$, and by defining $\eta = \cos\theta$ and so $\mathrm{d}\eta = -\sin\theta \mathrm{d}\theta$, we obtain

$$\begin{aligned} & \int_0^{\pi} \sin\theta \exp\left(\frac{f\xi \cos\theta}{k_B T} \right) \mathrm{d}\theta \\ &= \int_{-1}^{1} \exp\left(\frac{f\xi\eta}{k_B T} \right) \mathrm{d}\eta \\ &= \frac{2k_B T}{f\xi} \sinh\left(\frac{f\xi}{k_B T} \right). \end{aligned} \tag{3.6}$$

The Gibbs partition function becomes

$$Z_G = (2\pi m k_B T)^{3N/2}(4\pi)^N \left[\sum_S \exp\left(-\frac{v(S)-\mu S}{k_B T}\right) \times \int_0^{+\infty} \exp\left(-\frac{k(S)}{2k_B T}\left[\xi-\ell(S)\right]^2\right) \frac{\sinh\left(\frac{f\xi}{k_B T}\right)}{\frac{f\xi}{k_B T}} \xi^2 \mathrm{d}\xi\right]^N . \quad (3.7)$$

Then, in order to specify the shape of the potential wells represented in Fig.3.1, we consider $\ell(0) = \ell, \ell(1) = \chi\ell, v(0) = 0, v(1) = \Delta E$ and $k(0) = k(1) = k$. In addition, from now on, we neglect the non influential multiplicative constant within the partition function. Hence, by replacing the corresponding values, we obtain

$$Z_G(f) = \left[\int_0^{+\infty} \exp\left[-\frac{k}{2k_B T}\left(\xi-\ell\right)^2\right] \frac{\sinh\left(\frac{f\xi}{k_B T}\right)}{\frac{f\xi}{k_B T}} \xi^2 \mathrm{d}\xi + \exp\left(-\frac{\Delta E-\mu}{k_B T}\right) \times \int_0^{+\infty} \exp\left[-\frac{k}{2k_B T}\left(\xi-\chi\ell\right)^2\right] \frac{\sinh\left(\frac{f\xi}{k_B T}\right)}{\frac{f\xi}{k_B T}} \xi^2 \mathrm{d}\xi\right]^N . \quad (3.8)$$

We note that the chemical potential μ acts on the system by simply shifting the energy jump ΔE. We can rewrite the expression of Z_G under the form

$$Z_G = \left\{I_1 + I_2\right\}^N , \quad (3.9)$$

with

$$\begin{cases} I_1 = \int_0^{+\infty} \exp\left[-\frac{k}{2k_B T}\left(\xi-\ell\right)^2\right] \frac{\sinh\left(\frac{f\xi}{k_B T}\right)}{\frac{f\xi}{k_B T}} \xi^2 \mathrm{d}\xi; \\ I_2 = \phi \int_0^{+\infty} \exp\left[-\frac{k}{2k_B T}\left(\xi-\chi\ell\right)^2\right] \frac{\sinh\left(\frac{f\xi}{k_B T}\right)}{\frac{f\xi}{k_B T}} \xi^2 \mathrm{d}\xi; \\ \phi = \exp\left(-\frac{\Delta E-\mu}{k_B T}\right). \end{cases} \quad (3.10)$$

To begin, we calculate the first integral

$$I_1 = \int_0^{+\infty} \exp\left[-\frac{k}{2k_B T}\left(\xi-\ell\right)^2\right] \frac{\sinh\left(\frac{f\xi}{k_B T}\right)}{\frac{f\xi}{k_B T}} \xi^2 \mathrm{d}\xi, \quad (3.11)$$

and then, we will determine the value of the second one by substituting ℓ with $\chi\ell$ and by multiplying it by ϕ.

$$\begin{aligned} I_1 &= \left\{\int_0^{+\infty} \exp\left(-\frac{k}{2k_BT}\xi^2\right)\exp\left(\frac{k}{k_BT}\xi\ell\right)\exp\left(-\frac{k}{2k_BT}\ell^2\right)\right. \\ &\quad \left.\times\frac{k_BT}{2f}\xi\left[\exp\left(\frac{f\xi}{k_BT}\right)-\exp\left(-\frac{f\xi}{k_BT}\right)\right]\right\}\mathrm{d}\xi \\ &= \frac{k_BT}{2f}\exp\left(-\frac{k}{2k_BT}\ell^2\right)\int_0^{+\infty}\exp\left(-\frac{k}{2k_BT}\xi^2\right) \\ &\quad \times\left[\exp\frac{(k\ell+f)\xi}{k_BT}-\exp\frac{(k\ell-f)\xi}{k_BT}\right]\xi\mathrm{d}\xi. \end{aligned} \tag{3.12}$$

We calculate the integral A, in order to deduce I_1

$$\begin{aligned} A &= \int_0^{+\infty}\exp\left(-\alpha x^2\right)\exp\left(\beta x\right)\mathrm{d}x \\ &= \exp\left(\frac{\beta^2}{4\alpha}\right)\int_0^{+\infty}\exp\left[-\alpha\left(x-\frac{\beta}{2\alpha}x\right)^2\right]\mathrm{d}x. \end{aligned} \tag{3.13}$$

Then, we use the change of variable $y = x - \frac{\beta}{2\alpha}$, from which we get

$$A = \exp\left(\frac{\beta^2}{4\alpha}\right)\int_{-\frac{\beta}{2\alpha}}^{+\infty}\exp\left(-\alpha y^2\right)\mathrm{d}y. \tag{3.14}$$

We let $t^2 = \alpha y^2$, which means that $t = \sqrt{\alpha}y$ and $\mathrm{d}y = \frac{1}{\sqrt{\alpha}}\mathrm{d}t$. We finally obtain

$$A = \frac{1}{\sqrt{\alpha}}\exp\left(\frac{\beta^2}{4\alpha}\right)\int_{-\frac{\beta}{2\sqrt{\alpha}}}^{+\infty}\exp\left(-t^2\right)\mathrm{d}t. \tag{3.15}$$

We know that the error function, $\mathrm{erf}(x)$, can be written as $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x \exp\left(-t^2\right)\mathrm{d}t$. Thus, we obtain

$$A = \frac{1}{\sqrt{\alpha}}\exp\left(\frac{\beta^2}{4\alpha}\right)\frac{\sqrt{\pi}}{2}\left[\mathrm{erf}\left(+\infty\right)-\mathrm{erf}\left(-\frac{\beta}{2\sqrt{\alpha}}\right)\right]. \tag{3.16}$$

With the help of the following results, namely $\mathrm{erf}\left(+\infty\right) = 1$ and $\mathrm{erf}(-x) = -\mathrm{erf}(x)$, we get

$$A = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}\exp\left(\frac{\beta^2}{4\alpha}\right)\left[1+\mathrm{erf}\left(\frac{\beta}{2\sqrt{\alpha}}\right)\right]. \tag{3.17}$$

Then, we try to obtain the value of the integral $B = \int_0^{+\infty} \exp\left(-\alpha x^2\right) \exp\left(\beta x\right) x \mathrm{d}x$ by deriving the integral A. Hence, we straightforwardly obtain

$$\begin{aligned} B &= \frac{\partial A}{\partial \beta} \\ &= \frac{\partial}{\partial \beta}\left(\frac{1}{2}\sqrt{\frac{\pi}{\alpha}}\exp\left(\frac{\beta^2}{4\alpha}\right)\left[1+\mathrm{erf}\left(\frac{\beta}{2\sqrt{\alpha}}\right)\right]\right) \\ &= \frac{1}{2\alpha}+\frac{\beta}{4\alpha}\sqrt{\frac{\pi}{\alpha}}\exp\left(\frac{\beta^2}{4\alpha}\right)\left[1+\mathrm{erf}\left(\frac{\beta}{2\sqrt{\alpha}}\right)\right]. \end{aligned} \tag{3.18}$$

We continue the calculation of the integral I_1, where we let $\alpha = \frac{k}{2k_BT}$, $\beta_1 = \frac{k\ell+f}{k_BT}$ and $\beta_2 = \frac{k\ell-f}{k_BT}$, to compare Eq.(3.12) with Eq.(3.18). We get

$$\begin{aligned} I_1 &= \frac{k_BT}{2f}\exp\left(-\frac{k}{2k_BT}\ell^2\right)\left[\frac{1}{4\alpha}\sqrt{\frac{\pi}{\alpha}}\left(\beta_1\exp\left(\frac{\beta_1^2}{4\alpha}\right)\right.\right. \\ &\quad \left.\left.\times\left[1+\mathrm{erf}\left(\frac{\beta_1}{2\sqrt{\alpha}}\right)\right]-\beta_2\exp\left(\frac{\beta_2^2}{4\alpha}\right)\left[1+\mathrm{erf}\left(\frac{\beta_2}{2\sqrt{\alpha}}\right)\right]\right)\right] \\ &= \frac{k_BT}{2f}\exp\left(-\frac{k}{2k_BT}\ell^2\right)\frac{k_BT}{2k}\sqrt{\frac{2\pi k_BT}{k}}\left\{\left(\frac{k\ell+f}{k_BT}\right)\right. \\ &\quad \times\exp\left(\frac{(k\ell+f)^2}{k_B^2T^2}\frac{k_BT}{2k}\right)\left[1+\mathrm{erf}\left(\frac{k\ell+f}{2k_BT}\sqrt{\frac{2k_BT}{k}}\right)\right] \\ &\quad \left.-\left(\frac{k\ell-f}{k_BT}\right)\exp\left(\frac{(k\ell-f)^2}{k_B^2T^2}\frac{k_BT}{2k}\right)\left[1+\mathrm{erf}\left(\frac{k\ell-f}{2k_BT}\sqrt{\frac{2k_BT}{k}}\right)\right]\right\}. \end{aligned} \tag{3.19}$$

We develop the calculation and we obtain

$$\begin{aligned} I_1 &= \frac{(k_BT)^2}{4kf}\sqrt{\frac{2\pi k_BT}{k}}\exp\left(\frac{f^2}{2kk_BT}\right)\left\{\left(\frac{k\ell+f}{k_BT}\right)\exp\left(\frac{2k\ell f}{2kk_BT}\right)\left[1+\mathrm{erf}\left(\frac{k\ell+f}{\sqrt{2kk_BT}}\right)\right]\right. \\ &\quad \left.-\left(\frac{k\ell-f}{k_BT}\right)\exp\left(-\frac{2k\ell f}{2kk_BT}\right)\left[1+\mathrm{erf}\left(\frac{k\ell-f}{\sqrt{2kk_BT}}\right)\right]\right\}. \end{aligned} \tag{3.20}$$

The final result can be written under the following simplified form

$$\begin{aligned} I_1 &= \frac{k_BT}{4f}\sqrt{\frac{2\pi k_BT}{k}}\exp\left(\frac{f^2}{2kk_BT}\right)\left\{\left(\ell+\frac{f}{k}\right)\exp\left(\frac{\ell f}{k_BT}\right)\left[1+\mathrm{erf}\left(\frac{k\ell+f}{\sqrt{2kk_BT}}\right)\right]\right. \\ &\quad \left.-\left(\ell-\frac{f}{k}\right)\exp\left(-\frac{\ell f}{k_BT}\right)\left[1+\mathrm{erf}\left(\frac{k\ell-f}{\sqrt{2kk_BT}}\right)\right]\right\}. \end{aligned} \tag{3.21}$$

The next step consists in considering extensible bonds with relatively high values of the elastic constant k. Typical values of the elastic constants obtained from experiments as well as *ab initio* calculations can be found in the literature and are perfectly compatible

with such approximations [164–167]. It is necessary to consider the case of relatively high values of the elastic constant in order to approximate the erf by an exponential function. Indeed, as the Fourier transform will be used to obtain the Helmholtz partition function from the Gibbs partition function, terms of the Gibbs partition function have to be integrable. In this case, we can use the asymptotic expression $\mathrm{erf}(x) \simeq 1 - \frac{1}{\sqrt{\pi}x} e^{-x^2}$, holding for large values of x. We consider $\frac{f}{k} \ll \ell$, so the argument of error function is

$$\frac{k\ell \pm f}{\sqrt{2kk_BT}} \approx \frac{k\ell}{\sqrt{2kk_BT}} = \frac{\sqrt{k}\ell}{\sqrt{2k_BT}}. \tag{3.22}$$

Therefore, we can write

$$\mathrm{erf}\left(\frac{\sqrt{k}\ell}{\sqrt{2k_BT}}\right) \simeq 1 - \frac{\sqrt{2k_BT}}{\sqrt{\pi}\sqrt{k}\ell} \exp\left(-\frac{k\ell^2}{2k_BT}\right). \tag{3.23}$$

For high values of k, we can finally consider that $\mathrm{erf}\left(\frac{k\ell \pm f}{\sqrt{2kk_BT}}\right) \simeq 1$. Hence,

$$\begin{aligned} I_1 &= \frac{k_BT}{4f}\sqrt{\frac{2\pi k_BT}{k}} \exp\left(\frac{f^2}{2kk_BT}\right) \\ &\quad \times \left[2\left(\ell + \frac{f}{k}\right)\exp\left(\frac{\ell f}{k_BT}\right) - 2\left(\ell - \frac{f}{k}\right)\exp\left(-\frac{\ell f}{k_BT}\right)\right] \\ &= \frac{k_BT}{4f}\sqrt{\frac{2\pi k_BT}{k}} \exp\left(\frac{f^2}{2kk_BT}\right)\left[\ell \sinh\left(\frac{\ell f}{k_BT}\right) + \frac{f}{k}\cosh\left(\frac{\ell f}{k_BT}\right)\right]. \end{aligned} \tag{3.24}$$

We define the adimensional force $y = \frac{\ell f}{k_BT}$, and the ratio $\alpha = \frac{k_BT}{k\ell^2}$ between thermal and elastic energies. So, we also have

$$\begin{aligned} I_1 &= \ell^2\sqrt{\frac{2\pi k_BT}{k}} \exp\left(\frac{f^2}{2kk_BT}\right)\left[\frac{k_BT}{f\ell^2}\ell \sinh\left(\frac{\ell f}{k_BT}\right)\right. \\ &\quad \left. + \left(\frac{k_BT}{f\ell^2}\right)\frac{f}{k}\cosh\left(\frac{\ell f}{k_BT}\right)\right] \\ &= \ell^2\sqrt{\frac{2\pi k_BT}{k}} \exp\left(\frac{f^2}{2kk_BT}\right)\left[\frac{\sinh y}{y} + \alpha\cosh y\right]. \end{aligned} \tag{3.25}$$

The total partition function of the extensible bistable freely jointed chain is composed of I_1 and I_2 with $I_2 = \phi I_1\,(\ell \to \chi\ell)$. It means that $y \to \chi y$ and $\alpha \to \frac{\alpha}{\chi^2}$. Then, I_2 becomes

$$\begin{aligned} I_2 &= \phi\ell^2\chi^2\sqrt{\frac{2\pi k_BT}{k}} \exp\left(\frac{f^2}{2kk_BT}\right)\left[\frac{\sinh(\chi y)}{\chi y} + \frac{\alpha}{\chi^2}\cosh\chi y\right] \\ &= \phi\ell^2\sqrt{\frac{2\pi k_BT}{k}} \exp\left(\frac{f^2}{2kk_BT}\right)\left[\frac{\chi\sinh(\chi y)}{y} + \alpha\cosh\chi y\right]. \end{aligned} \tag{3.26}$$

The exponential term can be written in terms of y and α (with $y = \frac{\ell f}{k_B T}$ and $\alpha = \frac{k_B T}{k\ell^2}$) as

$$\exp\left(\frac{f^2}{2kk_BT}\right) = \exp\left(\frac{k_B^2T^2y^2}{2\ell^2kk_BT}\right) = \exp\left(\frac{k_BTy^2}{2\ell^2k}\right) = \exp\left(\frac{\alpha y^2}{2}\right). \tag{3.27}$$

By replacing the corresponding terms, the partition function becomes

$$Z_G^{(1)}(f) = \left[\frac{\sinh y}{y} + \alpha\cosh y + \phi\chi\frac{\sinh(\chi y)}{y} + \phi\alpha\cosh(\chi y)\right]^N \exp\left(\frac{N\alpha y^2}{2}\right). \tag{3.28}$$

This is the first approximation that we obtained by simply neglecting the effect of the error function in the exact calculation. Another level of approximation can be obtained by developing the term into brackets to the first order in α. If we let

$$\begin{cases} a = \frac{\sinh y}{y} + \phi\chi\frac{\sinh(\chi y)}{y}; \\ b = \cosh y + \phi\cosh(\chi y); \end{cases} \tag{3.29}$$

and we suppose that $\alpha \ll 1$, then we can write

$$(a+\alpha b)^N = a^N\left(1+\alpha\frac{b}{a}\right)^N \approx a^N\left(1+N\alpha\frac{b}{a}\right) = a^N + N\alpha a^{N-1}b. \tag{3.30}$$

The second approximation for the partition function is therefore

$$\begin{aligned} Z_G^{(2)}(f) \;=\; & \left[\left(\frac{\sinh y}{y} + \phi\chi\frac{\sinh(\chi y)}{y}\right)^N + N\alpha\left(\frac{\sinh y}{y} + \phi\chi\frac{\sinh(\chi y)}{y}\right)^{N-1}\right. \\ & \left.\times\left(\cosh y + \phi\cosh(\chi y)\right)\right]\exp\left(\frac{N\alpha y^2}{2}\right). \end{aligned} \tag{3.31}$$

Finally, the third approximation can be obtained by completely neglecting the linear terms in α in previous expressions

$$Z_G^{(3)}(f) = \left(\frac{\sinh y}{y} + \phi\chi\frac{\sinh(\chi y)}{y}\right)^N \exp\left(\frac{N\alpha y^2}{2}\right). \tag{3.32}$$

The exponential term depending on α is responsible for the elasticity of the chain. The same result without bistability (i.e. with $\phi = 0$), has been discussed in recent literature [168, 169]. Moreover, the case with bistability but without extensibility of the units is considered in Chapter 2. In this situation, the partition function simply reduces to

$$Z_G(f) = \left(\frac{\sinh y}{y} + \phi\chi\frac{\sinh(\chi y)}{y}\right)^N. \tag{3.33}$$

As discussed previously, the macroscopic behaviour of this system is described by the Gibbs free energy $\mathcal{G} = -k_B T \log Z_G$. The force-extension response for the bistable freely jointed chain with extensible elements under isotensional condition can be found through

$$\begin{aligned} \langle r \rangle &= -\frac{\partial \mathcal{G}}{\partial f} \\ &= k_B T \frac{\partial \log Z_G}{\partial f} \\ &= \ell \frac{\partial \log Z_G}{\partial y}. \end{aligned} \tag{3.34}$$

On the other hand, we can define the average value of the spin variables as $\langle s \rangle = \frac{1}{N}\langle \sum_{i=1}^{N} S_i \rangle$. We can rewrite the definition of the Hamiltonian within the Gibbs ensemble

$$\begin{aligned} H_G\left(\vec{q}, \vec{p}, S\right) &= \sum_{i=1}^{N} \frac{\vec{p}_i \cdot \vec{p}_i}{2m} + \sum_{i=1}^{N} v(S_i) - \mu \sum_{i=1}^{N} S_i \\ &\quad + \sum_{i=1}^{N} \frac{1}{2} k(S_i) \left[\|\vec{r}_i - \vec{r}_{i-1}\| - \ell(S_i)\right]^2 - \vec{f} \cdot \vec{r}_N. \end{aligned} \tag{3.35}$$

The only variable depending on ΔE is the basal energy $v\left(S_i\right)$ and its value changes according to the spin. If the spin S takes the value 0, we have $v\left(0\right) = 0$ and if it takes the value 1, we have $v\left(1\right) = \Delta E$. This means that

$$\sum_{i=1}^{N} v\left(S_i\right) = \Delta E \sum_{i=1}^{N} S_i. \tag{3.36}$$

Therefore we get

$$\begin{aligned} H_G\left(\vec{q}, \vec{p}, S\right) &= \sum_{i=1}^{N} \frac{\vec{p}_i \cdot \vec{p}_i}{2m} + \Delta E \sum_{i=1}^{N} S_i - \mu \sum_{i=1}^{N} S_i \\ &\quad + \sum_{i=1}^{N} \frac{1}{2} k(S_i) \left[\|\vec{r}_i - \vec{r}_{i-1}\| - \ell(S_i)\right]^2 - \vec{f} \cdot \vec{r}_N. \end{aligned} \tag{3.37}$$

We can derive the Hamiltonian according to ΔE

$$\frac{\partial H_G}{\partial\left(\Delta E\right)} = \sum_{i=1}^{N} S_i. \tag{3.38}$$

In the same way, we can write the Gibbs partition function

$$Z_G = \sum_{S_1 \in \{0,1\}} ... \sum_{S_N \in \{0,1\}} \int_{\vec{p}} \int_{\vec{q}} \exp\left(-\frac{H_G}{k_B T}\right) \mathrm{d}\vec{p}\mathrm{d}\vec{q}. \tag{3.39}$$

We can derive Z_G according to ΔE

$$\frac{\partial Z_G}{\partial(\Delta E)} = \sum_{S_1\in\{0,1\}} ... \sum_{S_N\in\{0,1\}} \int_{\vec{p}}\int_{\vec{q}} \exp\left(-\frac{H_G}{k_BT}\right)\left(-\frac{1}{k_BT}\right)\sum_{i=1}^{N} S_i \mathrm{d}\vec{p}\mathrm{d}\vec{q}. \tag{3.40}$$

As $\langle s\rangle = \frac{1}{N}\left\langle\sum_{i=1}^{N} S_i\right\rangle$, we get

$$\begin{aligned}\langle s\rangle &= \frac{\left(-\frac{k_BT}{N}\right)\frac{\partial Z_G}{\partial(\Delta E)}}{Z_G} \\ &= \left(-\frac{k_BT}{N}\right)\frac{1}{Z_G}\frac{\partial Z_G}{\partial(\Delta E)}.\end{aligned} \tag{3.41}$$

And we finally obtain

$$\langle s\rangle = \left(-\frac{k_BT}{N}\right)\frac{\partial \log Z_G}{\partial(\Delta E)}. \tag{3.42}$$

We can also write it like

$$N\langle s\rangle = \frac{\partial \mathcal{G}}{\partial \Delta E} = -\frac{\partial \mathcal{G}}{\partial \mu}. \tag{3.43}$$

The explicit application of Eq.(3.34) to Eq.(3.32) yields

$$\langle r\rangle = N\ell\left[\frac{\mathcal{L}(y) + \chi^2\phi\mathcal{L}(\chi y)\frac{\sinh(\chi y)}{\sinh y}}{1+\chi\phi\frac{\sinh(\chi y)}{\sinh y}} + \alpha y\right], \tag{3.44}$$

where $\mathcal{L}(y) = \coth y - \frac{1}{y}$ is the Langevin function. On the other hand, by applying Eq.(3.43) to Eq.(3.32), we get

$$\langle s\rangle = \frac{\chi\phi\frac{\sinh\chi y}{\sinh y}}{1+\chi\phi\frac{\sinh\chi y}{\sinh y}}, \tag{3.45}$$

which does not depend on α. By combining Eqs.(3.44) and (3.45), we can find the relationship

$$\langle r\rangle = N\ell\left[(1-\langle s\rangle)\mathcal{L}(y) + \langle s\rangle\chi\mathcal{L}(\chi y) + \alpha y\right], \tag{3.46}$$

stating that the average extension of the extensible bistable system is given by a combination of the responses of two freely jointed chain models with lengths ℓ and $\chi\ell$, and by a linear term taking into account the elasticity of the system.

A first example of application of Eqs.(3.34) and (3.43) is given in Fig.3.2, where we compare a system with extensibility described by the exact Eq.(3.9) or by the approximated Eq.(3.32) and another system without extensibility described by Eq.(3.33). First of

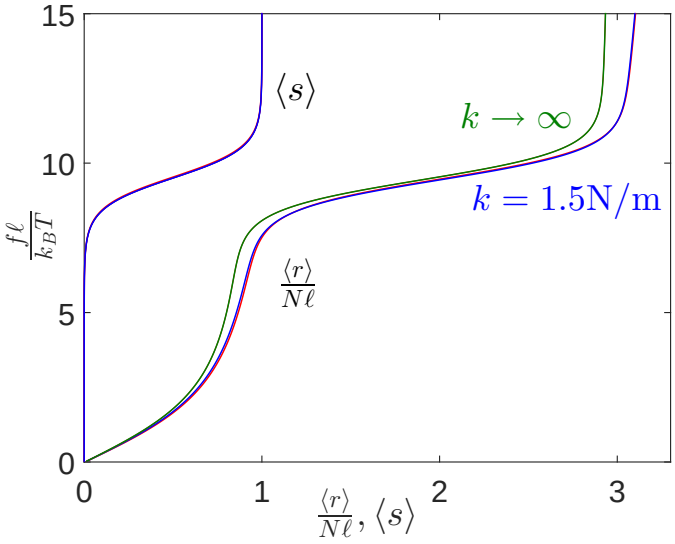

Figure 3.2 – Force-extension response and average spin for a system with extensibility ($k = 1.5\mathrm{N/m}$) and a system without extensibility ($k \to \infty$). The calculation for the system with extensibility is based on the exact result in Eq.(3.9) (red curves) and on the approximation in Eq.(3.32) (blue curves). On the other hand, the system without extensibility has been studied through Eq.(3.33). We adopted the parameters $N = 5$, $\ell = 0.5 \times 10^{-9}\mathrm{m}$, $T = 300\mathrm{K}$, $\chi = 3$, $\Delta E = 20 k_B T = 8.28 \times 10^{-20}\mathrm{J}$ and $\mu = 0$ [1].

all, we remark the very good agreement between Eqs.(3.9) and (3.32) for the system with extensibility (red and blues curves, respectively, in Fig.3.2). This proves the acceptability of the proposed approximation. Moreover, it is interesting to note the different asymptotic behaviour of systems with finite and infinite elastic constants. While the stiff system exhibits a vertical asymptote for large forces, the elastic one shows a tilted asymptote, which slope represents the effective stiffness of the chain. In the force-extension curve, we observe a force plateau corresponding to $f^* = \frac{\Delta E}{(\chi-1)\ell}$ (in the case of $\mu = 0$). Similarly, in the spin behaviour, we identify a transition from $S = 0$ to $S = 1$ corresponding to the same force f^*. We can explain these curves by means of a cooperative process, which generates the transition of all units at the same value of force f^*. Notably, the value of this plateau force does not depend on the spring constant, nor on the temperature. Such a result is readily interpreted in the framework of the Bell expression, originally derived in the context of the adhesion of cells [159]. A larger variation of the elastic constant is shown in Fig.3.3, where force-extension curves, spin variables and Gibbs free energies are plotted for several values of $k = 0.4, 0.8, 1.2, 1.6, 2.0$ and $2.4\mathrm{N/m}$. Also in this case, we underline the good agreement between exact and approximated results, with a small deviation appearing only for the softer chains. The origin of the force transition in the

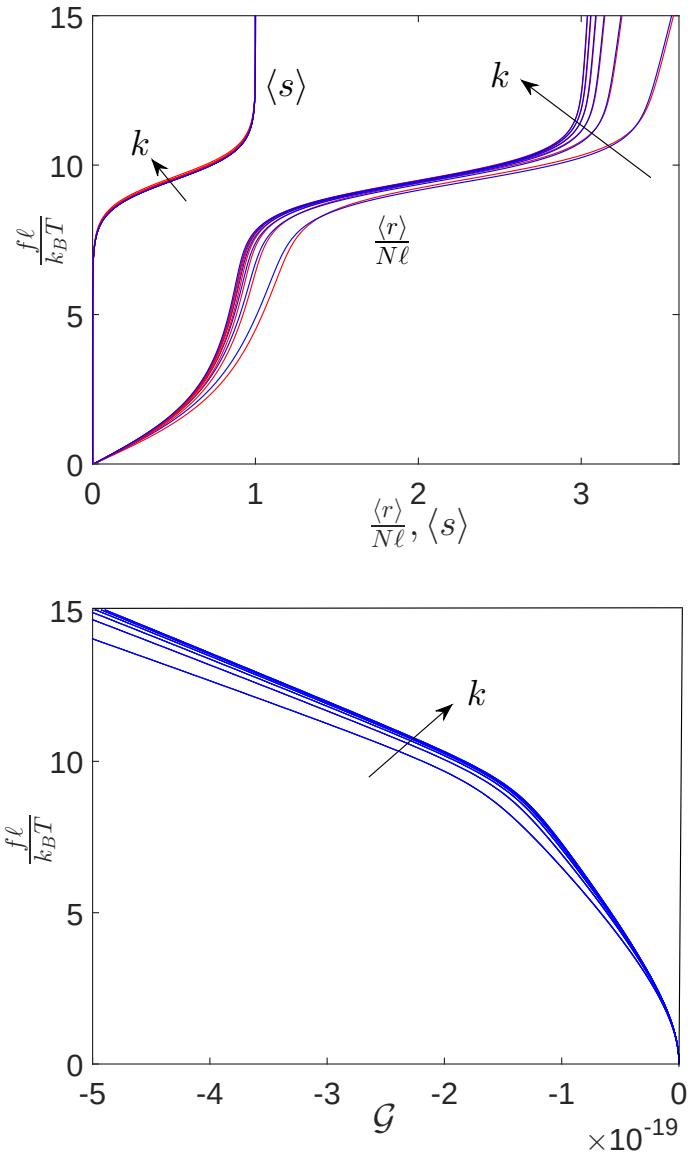

Figure 3.3 – Force-extension response and average spin (top panel) and Gibbs free energy (bottom panel) for a system with variable extensibility $k = 0.4, 0.8, 1.2, 1.6, 2.0$ and 2.4N/m. The calculation of the force-extension response and the average spin is based on the exact result in Eq.(3.9) (red curves) and on the approximation in Eq.(3.32) (blue curves). We adopted the parameters $N = 5$, $\ell = 0.5 \times 10^{-9}$m, $T = 300$K, $\chi = 3$, $\Delta E = 20k_BT = 8.28 \times 10^{-20}$J and $\mu = 0$ [1].

force-extension curve and in the spin behaviour can be recognised in the Gibbs free energy, which shows a slope change corresponding exactly to the same force. While the spin variable average remains unaltered by the elastic constant, the force-extension response is sensibly modified, showing a variable slope in the regions before and after the plateau

corresponding to the transition. The knowledge of chain behaviour under the isotensional or Gibbs ensemble is the starting point for the understanding of the isometric ensemble, as discussed in the next Section.

3.3 Two-state freely jointed chain with extensible units: the Helmholtz ensemble

We consider now the isometric condition (within the Helmholtz ensemble), where the last unit of the chain is considered as tethered at a given point $\vec{r}$ of the space. In this case, we can therefore write the Hamiltonian of the system under the form

$$\begin{aligned} H_H &= \sum_{i=1}^{N-1} \frac{\vec{p}_i \cdot \vec{p}_i}{2m} + \sum_{i=1}^{N} v(S_i) - \mu \sum_{i=1}^{N} S_i \\ &\quad + \sum_{i=1}^{N} \frac{1}{2} k(S_i) \left[\|\vec{r}_i - \vec{r}_{i-1}\| - \ell(S_i) \right]^2 , \end{aligned} \tag{3.47}$$

where $\vec{r}_N = \vec{r}$ is fixed. Consequently, we also consider $\vec{p}_N = 0$. In this situation, we introduce $\vec{q} = (\vec{r}_1, ..., \vec{r}_{N-1})$ as the generalised vector containing all positions, $\vec{p} = (\vec{p}_1, ..., \vec{p}_{N-1})$ as the generalised vector containing all momenta, and $\vec{S} = (S_1, ..., S_N)$ as the generalised vector of all spin variables. Similarly to Section 3.2, we assume that all chemical potentials μ_i take the same value μ. The partition function can be written in terms of these dynamic variables

$$\begin{aligned} Z_H &= \sum_{S_1 \in \{0,1\}} ... \sum_{S_N \in \{0,1\}} \int_{\mathbb{R}^{3N-3}} \int_{\mathbb{R}^{3N-3}} \exp\left[-\frac{H_H\left(\vec{p}, \vec{q}, \vec{S}\right)}{k_B T} \right] \mathrm{d}\vec{p}\mathrm{d}\vec{q} \\ &= \sum_{\vec{S}} \left[\int_{\mathbb{R}^3} \exp\left(-\frac{1}{2mk_BT} \vec{p} \cdot \vec{p} \right) \mathrm{d}\vec{p} \right]^{N-1} \int_{\mathbb{R}^3} ... \int_{\mathbb{R}^3} \exp\left[-\sum_{i=1}^{N} \frac{v(S_i) - \mu S_i}{k_B T} \right. \\ &\quad \left. - \sum_{i=1}^{N} \frac{k(S_i)}{2k_B T} \left[\|\vec{r}_i - \vec{r}_{i-1}\| - \ell(S_i) \right]^2 \right] \mathrm{d}\vec{r}_1 ... \mathrm{d}\vec{r}_{N-1}. \end{aligned} \tag{3.48}$$

It is not difficult to realise that the calculation of this partition function is much more complicated than the Gibbs partition function. Indeed, we can not apply a change of variables in order to factorise the multi-dimensional integral. From the physical point of view, this difficulty depends on the isometric condition, inducing an effective interaction among the units and so fixing the sum of all vectors $\vec{r}_i - \vec{r}_{i-1}$ (for i from 1 to N). This problem will produce a final result which can not be written under the form of a power

with exponent N. An useful technique to cope with this difficulty is the following. By comparing Eqs.(3.2) and (3.48), we deduce that the two partition functions Z_G and Z_H are related through a three-dimensional bilateral Laplace transform, as

$$Z_G(\vec{f}) = \int_{\mathbb{R}^3} Z_H(\vec{r}) \exp\left(\frac{\vec{r}\cdot\vec{f}}{k_B T}\right) \mathrm{d}\vec{r}, \tag{3.49}$$

where, as usual, we neglect the non-influential multiplicative constants in the partition function. Moreover, by considering the spherical symmetry of the problem, we easily obtain the inverse relationship

$$Z_H(r) = \int_{-\infty}^{+\infty} Z_G(i\eta) \frac{\eta}{r} \sin\left(\frac{\eta r}{k_B T}\right) \mathrm{d}\eta, \tag{3.50}$$

where we have introduced the analytic continuation of the function $Z_G(r)$ to the imaginary axis, leading to $Z_G(i\eta)$. We can also write that

$$Z_H(r) = \int_{-\infty}^{+\infty} Z_G\left(\frac{ik_B T}{y}\right) \frac{y}{r} \sin\left(\frac{ry}{\ell}\right) \mathrm{d}y. \tag{3.51}$$

By substituting Eq.(3.32) in Eq.(3.51), we get the important integral expression

$$Z_H(r) \;=\; \int_{-\infty}^{+\infty} \left\{\frac{\sin y}{y} + \chi\phi \frac{\sin(\chi y)}{y}\right\}^N \exp\left(-N\frac{\alpha y^2}{2}\right) \frac{y}{r} \sin\left(\frac{ry}{\ell}\right) \mathrm{d}y. \tag{3.52}$$

With the help of $e^x = \cos x + i \sin x$, we can also write it like

$$Z_H(r) \;=\; -i \int_{-\infty}^{+\infty} \left\{\frac{\sin y}{y} + \chi\phi \frac{\sin(\chi y)}{y}\right\}^N \exp\left(-N\frac{\alpha y^2}{2}\right) \frac{y}{r} \exp\left(i\frac{ry}{\ell}\right) \mathrm{d}y. \tag{3.53}$$

The analysis of such an integral for $\alpha = 0$ (i.e. without elasticity) has been performed in recent literature [113] and in Chapter 2. Previous investigations considered the case with $\alpha = 0$ and $\phi = 0$, thus dealing with the classical FJC model under isometric condition [160–163]. However, the presence of a finite elasticity of the units, quantified by the parameter $\alpha \neq 0$, completely modifies the approach to be used to obtain a closed form expression for $Z_H(r)$. Here, we elaborate Eq.(3.53) in order to get such a closed form expression, useful to better explain the chain behaviour within the Helmholtz ensemble. The function to integrate in Eqs.(3.52) or (3.53) is regular on the real axis and analytic (or holomorphic) on a strip $|\mathrm{Im}\, y| < M$ for an arbitrary $M \in \mathbb{R}$. Then, instead of integrating on the whole real axis, we can use the path Γ shown in Fig.3.4. This will be useful to elaborate the partition function integral and to write it in a form without singularities at the origin.

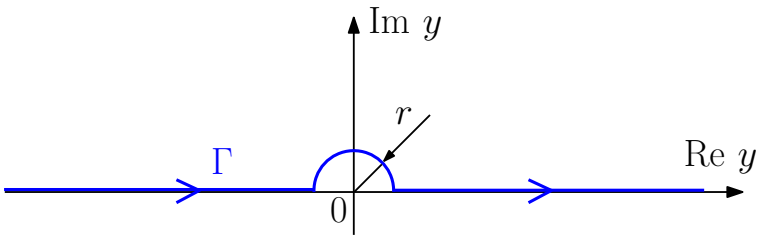

Figure 3.4 – Definition of the contour Γ on the complex plane with an arbitrary radius r [1].

We develop the power with the help of the Newton rule

$$\begin{aligned}
&\left(\frac{\sin y}{y}+\phi\chi\frac{\sin(\chi y)}{y}\right)^N \\
&= \sum_{k=0}^{N}\binom{N}{k}\left(\frac{\sin y}{y}\right)^{N-k}\left(\phi\chi\frac{\sin(\chi y)}{y}\right)^k \\
&= \frac{1}{y^N}\sum_{k=0}^{N}\binom{N}{k}\left(\sin^{N-k} y\right)\phi^k\chi^k\sin^k(\chi y). \qquad (3.54)
\end{aligned}$$

By developing the power in previous expression with the Newton rule and by using the following expansion

$$\begin{aligned}
\sin^n y &= \left(\frac{e^{iy}-e^{-iy}}{2i}\right)^n \\
&= \frac{1}{(2i)^n}\sum_{t=0}^{n}\binom{n}{t}\left(e^{iy}\right)^{n-t}\left(-e^{-iy}\right)^t \\
&= \frac{1}{(2i)^n}\sum_{t=0}^{n}\binom{n}{t}(-1)^t e^{iny}e^{-2iyt}, \qquad (3.55)
\end{aligned}$$

we obtain

$$\begin{aligned}
Z_H(r) &= \frac{1}{2^N i^{N+1} r}\sum_{k=0}^{N}\sum_{p=0}^{N-k}\sum_{q=0}^{k}\binom{N}{k}\binom{N-k}{p}\binom{k}{q} \\
&\quad\times(-1)^{p+q}(\chi\phi)^k\int_\Gamma e^{-i\Lambda y}e^{-N\frac{\alpha y^2}{2}}\frac{\mathrm{d}y}{y^{N-1}}, \qquad (3.56)
\end{aligned}$$

where we defined

$$\Lambda = k-N+2p-\chi k+2\chi q-\frac{r}{\ell}. \qquad (3.57)$$

The integral appearing in Eq.(3.56) is well defined since the path Γ excludes the singularity at the origin from the integration.

3.3.1 An integral calculation

To complete the calculation of the partition function, we have to determine the value of

$$I_m(a,b) = \int_\Gamma e^{iay} e^{-by^2} \frac{\mathrm{d}y}{y^m}, \tag{3.58}$$

where $a \in \mathbb{R}$, $b > 0$, $m \in \mathbb{N}$ and the path Γ is given in Fig.3.4. To this aim, we begin to introduce the closed paths ABCD shown in Fig.3.5. To understand the suitability of this approach, we firstly consider the segment CD on the paths shown in Fig.3.5. Here, we have $y = s + i\beta$ where $s \in \mathbb{R}$ is a variable and $\beta \in \mathbb{R}$ is a constant that we will determine with the purpose of simplifying the integral calculation. Therefore, we get

$$\begin{aligned} \int_{CD} e^{iay} e^{-by^2} \frac{\mathrm{d}y}{y^m} &= -\int_{-R}^{+R} e^{ia(s+i\beta)} e^{-b(s+i\beta)^2} \frac{ds}{(s+i\beta)^m} \\ &= -e^{-a\beta} e^{b\beta^2} \int_{-R}^{+R} e^{is(a-2b\beta)} e^{-bs^2} \frac{ds}{(s+i\beta)^m}, \end{aligned} \tag{3.59}$$

and we can let $\beta = a/(2b)$ to remove the imaginary exponential from the integral. Here, the value of R is defined in Fig.3.5. Hence, the result can be written as

$$\begin{aligned} \int_{CD} e^{iay} e^{-by^2} \frac{\mathrm{d}y}{y^m} &= e^{-\frac{a^2}{2b}} e^{\frac{ba^2}{4b^2}} \int_{-R}^{+R} \frac{e^{-bs^2} ds}{(s+i\beta)^m} \\ &= -e^{-\frac{a^2}{4b}} \int_{-R}^{+R} \frac{e^{-bs^2} ds}{(s+i\frac{a}{2b})^m}, \end{aligned} \tag{3.60}$$

and, by applying the limit for $R \to \infty$ and the change of variable $\eta = \sqrt{b}s$, it assumes the final form

$$\begin{aligned} \lim_{R\to\infty} \int_{CD} e^{iay} e^{-by^2} \frac{\mathrm{d}y}{y^m} &= -e^{-\frac{a^2}{4b}} \int_{-\infty}^{+\infty} \frac{e^{-\eta^2} \mathrm{d}\eta}{\sqrt{b}\left(\frac{\eta}{\sqrt{b}} + i\frac{a}{2b}\right)^m} \\ &= -e^{-\frac{a^2}{4b}} \int_{-\infty}^{+\infty} \frac{e^{-\eta^2} \mathrm{d}\eta}{\sqrt{b}\frac{1}{\sqrt{b}^m}\left(\eta + i\frac{a}{2\sqrt{b}}\right)^m} \\ &= -b^{\frac{m-1}{2}} e^{-\frac{a^2}{4b}} \int_{-\infty}^{+\infty} \frac{e^{-\eta^2} \mathrm{d}\eta}{\left(\eta + i\frac{a}{2\sqrt{b}}\right)^m}. \end{aligned} \tag{3.61}$$

This expression, of course, is valid for both negative and positive values of a.

For the calculation of the integral in Eq.(3.58), we have to consider two separate cases. If $a > 0$ or $\beta > 0$, the path shown in Fig.3.5.a does not contain the origin of the complex

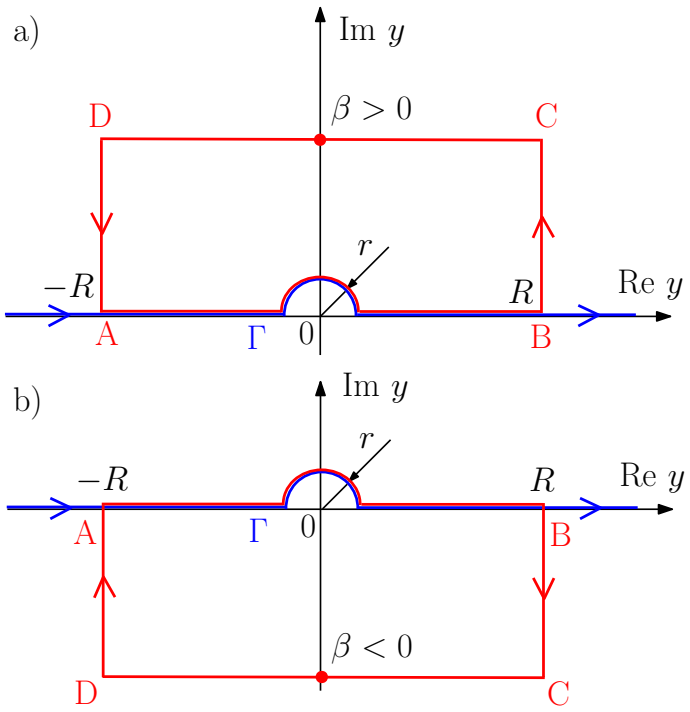

Figure 3.5 – Definition of the contours ABCD on the complex plane for $\beta > 0$ a) and $\beta < 0$ b) [1].

plane and therefore the function to integrate is holomorphic within the path. In this situation, we can directly apply the Cauchy theorem stating that

$$\left(\int_{AB} + \int_{BC} + \int_{CD} + \int_{DA}\right) e^{iay} e^{-by^2} \frac{\mathrm{d}y}{y^m} = 0. \tag{3.62}$$

Now, in the limit for $R \to \infty$, it is not difficult to prove that the integrals on the segments BC and DA vanish and, as the integral on CD is the opposite of the one on AB, we get the first result for $a > 0$

$$I_m(a,b) = b^{\frac{m-1}{2}} e^{-\frac{a^2}{4b}} \int_{-\infty}^{+\infty} \frac{e^{-\eta^2} \mathrm{d}\eta}{(\eta + i\frac{a}{2\sqrt{b}})^m}. \tag{3.63}$$

If $a < 0$ or $\beta < 0$, the path shown in Fig.3.5.b contains the origin of the complex plane, which corresponds to a singular point of the function to integrate. Hence, the residue theorem delivers

$$\left(\int_{AB} + \int_{BC} + \int_{CD} + \int_{DA}\right) e^{iay} e^{-by^2} \frac{\mathrm{d}y}{y^m} = -2\pi i \mathrm{Res}\left\{\frac{e^{iay} e^{-by^2}}{y^m}, 0\right\}, \tag{3.64}$$

where the residue can be easily calculated since the singular point is a pole or order m

$$\mathrm{Res}\left\{\frac{e^{iay} e^{-by^2}}{y^m}, 0\right\} = \frac{1}{(m-1)!} \frac{\mathrm{d}^{m-1}}{\mathrm{d}y^{m-1}} \left(e^{iay} e^{-by^2}\right)_{y=0}. \tag{3.65}$$

The derivative in Eq.(3.65) can be elaborated through the general rule

$$\frac{\mathrm{d}^{m-1}}{\mathrm{d}y^{m-1}}\left[f(y)g(y)\right]=\sum_{k=0}^{m-1}\binom{m-1}{k}\frac{\mathrm{d}^{k}f(y)}{\mathrm{d}y^{k}}\frac{\mathrm{d}^{m-1-k}g(y)}{\mathrm{d}y^{m-1-k}}. \tag{3.66}$$

If $g(y)=e^{iay}$, we simply have

$$\frac{\mathrm{d}^{m-1-k}g(y)}{\mathrm{d}y^{m-1-k}}=(ia)^{m-1-k}e^{iay}. \tag{3.67}$$

On the other hand, if $f(y)=e^{-by^2}$, the calculation of the derivative of order k is more involved. One method is based on the following result

$$\int_0^{+\infty}e^{-\alpha x^2}e^{\beta x}\mathrm{d}x=\frac{1}{2}\sqrt{\frac{\pi}{\alpha}}e^{\frac{\beta^2}{4\alpha}}\left[1+\mathrm{erf}\left(\frac{\beta}{2\sqrt{\alpha}}\right)\right], \tag{3.68}$$

proved in Section 3.2. Since we need

$$\int_{-\infty}^{+\infty}e^{-\alpha x^2}e^{\beta x}\mathrm{d}x, \tag{3.69}$$

we first elaborate

$$\begin{aligned}\int_{-\infty}^{0}e^{-\alpha x^2}e^{\beta x}\mathrm{d}x &= \int_{-\infty}^{0}e^{-\alpha y^2}e^{-\beta y}\left(-\mathrm{d}y\right)\\ &= \int_0^{+\infty}e^{-\alpha x^2}e^{-\beta x}\mathrm{d}x\\ &= \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}e^{\frac{\beta^2}{4\alpha}}\left[1+\mathrm{erf}\left(-\frac{\beta}{2\sqrt{\alpha}}\right)\right],\end{aligned} \tag{3.70}$$

having used Eq.(3.68), with $\beta\to-\beta$. Now, we sum these two equations and we get

$$\int_{-\infty}^{+\infty}e^{-\alpha x^2}e^{\beta x}\mathrm{d}x = \frac{1}{2}\sqrt{\frac{\pi}{\alpha}}e^{\frac{\beta^2}{4\alpha}}\left[2+\mathrm{erf}\left(\frac{\beta}{2\sqrt{\alpha}}\right)+\mathrm{erf}\left(-\frac{\beta}{2\sqrt{\alpha}}\right)\right]. \tag{3.71}$$

Here, we have that $\mathrm{erf}(x)=-\mathrm{erf}(-x)$, and therefore

$$\int_{-\infty}^{+\infty}e^{-\alpha x^2}e^{-\beta x}\mathrm{d}x=\sqrt{\frac{\pi}{\alpha}}e^{\frac{\beta^2}{4\alpha}}. \tag{3.72}$$

We can substitute $\beta=iy$ and $b=\frac{1}{4\alpha}$ to obtain

$$\int_{-\infty}^{+\infty}e^{-\frac{x^2}{4b}}e^{ixy}\mathrm{d}x=\sqrt{4\pi b}e^{-by^2}, \tag{3.73}$$

It means that the Fourier transform of $e^{\frac{-x^2}{4b}}$ is proportional to e^{-by^2}. Incidentally, if $b = \frac{1}{2}$ (i.e. $\frac{1}{4b} = b$), then we obtain the eigenvectors of the Fourier transform. We can calculate the derivative as

$$\begin{aligned}\frac{\mathrm{d}^k}{\mathrm{d}y^k}e^{-by^2} &= \frac{1}{\sqrt{4\pi b}}\frac{\mathrm{d}^k}{\mathrm{d}y^k}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{4b}}e^{ixy}\mathrm{d}x \\ &= \frac{1}{\sqrt{4\pi b}}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{4b}}\,(ix)^k\,e^{ixy}\mathrm{d}x. \end{aligned} \tag{3.74}$$

Therefore,

$$\frac{\mathrm{d}^k f(y)}{\mathrm{d}y^k} = \frac{1}{\sqrt{4\pi b}}\int_{-\infty}^{+\infty} e^{-\frac{x^2}{4b}}(ix)^k e^{ixy}\mathrm{d}x. \tag{3.75}$$

Hence, if we consider $y = 0$, we get

$$\left(\frac{\mathrm{d}^k f(y)}{\mathrm{d}y^k}\right)_{y=0} = \frac{i^k}{\sqrt{4\pi b}}\int_{-\infty}^{+\infty} x^k e^{-\frac{x^2}{4b}}\mathrm{d}x. \tag{3.76}$$

Then, we observe that the result is different from zero only for k even. Since $\int_{-\infty}^{+\infty} e^{-\alpha x^2}\mathrm{d}x = \sqrt{\frac{\pi}{\alpha}}$ (see Eq.(3.72) with $\beta = 0$), we get

$$\begin{aligned}\int_{-\infty}^{+\infty} -x^2 e^{-\alpha x^2}\mathrm{d}x &= \frac{\mathrm{d}}{\mathrm{d}\alpha}\sqrt{\frac{\pi}{\alpha}} \\ &= \sqrt{\pi}\left(-\frac{1}{2}\right)\alpha^{-3/2} \\ &= -\frac{1}{2}\sqrt{\frac{\pi}{\alpha^3}}, \end{aligned} \tag{3.77}$$

or we can also write

$$\int_{-\infty}^{+\infty} x^2 e^{-\alpha x^2}\mathrm{d}x = \frac{1}{2}\sqrt{\frac{\pi}{\alpha^3}} = \frac{1}{2\alpha}\sqrt{\frac{\pi}{\alpha}}. \tag{3.78}$$

Again

$$\begin{aligned}\int_{-\infty}^{+\infty} -x^4 e^{-\alpha x^2}\mathrm{d}x &= \frac{1}{2}\sqrt{\pi}\left(-\frac{3}{2}\right)\alpha^{-5/2} \\ &= -\frac{3}{4}\sqrt{\pi}\sqrt{\frac{1}{\alpha^5}}. \end{aligned} \tag{3.79}$$

We can sum it up like

$$\begin{aligned}\int_{-\infty}^{+\infty} x^{2n} e^{-\alpha x^2}\mathrm{d}x &= \frac{1}{\alpha^{n+1/2}}\frac{(2n)!}{(2^{2n})\,n!}\sqrt{\pi} \\ &= \frac{\Gamma\left(n+\frac{1}{2}\right)}{\alpha^{n+1/2}}, \end{aligned} \tag{3.80}$$

and

$$\int_{-\infty}^{+\infty} x^{2n+1} e^{-\alpha x^2} \mathrm{d}x = 0. \tag{3.81}$$

In particular, by fixing $\alpha = \frac{1}{4b}$, we eventually obtain

$$\left(\frac{\mathrm{d}^k e^{-by^2}}{\mathrm{d}y^k} \right)_{y=0} = \begin{cases} \frac{(-4b)^n}{\sqrt{\pi}} \Gamma\left(n + \frac{1}{2}\right) & \text{if} \quad k = 2n, \\ 0 & \text{if} \quad k = 2n+1, \end{cases} \tag{3.82}$$

where $\Gamma(z)$ is the Euler Gamma function. By combining Eqs.(3.66), (3.67) and (3.82), we can rewrite the residue defined in Eq.(3.65) as

$$\mathrm{Res}\left\{ \frac{e^{iay} e^{-by^2}}{y^m}, 0 \right\} = \sum_{k=0}^{m-1} \frac{(ia)^{m-1-k}}{k!(m-1-k)!} \frac{(-4b)^{k/2}}{2\sqrt{\pi}} \Gamma\left(\frac{k+1}{2}\right) \left[1 + (-1)^k\right]. \tag{3.83}$$

Therefore, the residue theorem written in the form of Eq.(3.64), and considered for $R \to \infty$, gives the final result for $a < 0$

$$\begin{aligned} I_m(a,b) &= b^{\frac{m-1}{2}} e^{-\frac{a^2}{4b}} \int_{-\infty}^{+\infty} \frac{e^{-\eta^2} \mathrm{d}\eta}{(\eta + i\frac{a}{2\sqrt{b}})^m} \\ &\quad - \sum_{k=0}^{m-1} \frac{2\pi i (ia)^{m-1-k}}{k!(m-1-k)!} \frac{(-4b)^{k/2}}{2\sqrt{\pi}} \Gamma\left(\frac{k+1}{2}\right) \left[1 + (-1)^k\right]. \end{aligned} \tag{3.84}$$

To conclude, Eq.(3.63) for $a > 0$ and Eq.(3.84) for $a < 0$ represent the integral $I_m(a,b)$ defined in Eq.(3.58) in terms of a second integral, which is the subject of the following discussion.

3.3.2 The Hermite elements with negative index

First of all, we observe that the integral

$$J_m(x) = \int_{-\infty}^{+\infty} \frac{e^{-\eta^2} \mathrm{d}\eta}{(\eta + ix)^m}, \tag{3.85}$$

appearing in Eqs.(3.63) and (3.84) with $x = a/(2\sqrt{b})$, is simpler than $I_m(a,b)$ since it is calculated over the real axis and there is no imaginary exponent in the function to integrate. To approach its calculation, we recall the following integral representation of the Hermite polynomials [170, 171]

$$\begin{aligned} H_m(z) &= \frac{2^m}{\sqrt{\pi}} \int_{-\infty}^{+\infty} (i\eta + z)^m e^{-\eta^2} \mathrm{d}\eta \\ &= \frac{(2i)^m}{\sqrt{\pi}} \int_{-\infty}^{+\infty} (\eta - iz)^m e^{-\eta^2} \mathrm{d}\eta. \end{aligned} \tag{3.86}$$

If we compare J_m with the last representation of H_m, we note that J_m is formally related to H_{-m}, which represents an Hermite element with negative index. We underline that the representation of the Hermite polynomials in Eq.(3.86) makes sense also for negative values of m. However, we will see that these functions are not polynomials but they exhibit several properties similar to those of the classical Hermite polynomials. To the author knowledge, these functions have been firstly introduced and studied in Ref. [172] but no other analysis or application can be found in the literature. Anyway, by replacing m with $-m$ in Eq.(3.86), we define

$$H_{-m}(z) = \frac{1}{(2i)^m\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{e^{-\eta^2}\mathrm{d}\eta}{(\eta - iz)^m}, \tag{3.87}$$

and we introduce their properties. Following Ref. [172], we analyse the first element $H_{-1}(z)$. We simply have

$$\begin{aligned} H_{-1}(z) &= \frac{1}{2i\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{e^{-\eta^2}\mathrm{d}\eta}{\eta - iz} \\ &= \frac{1}{2i\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{\eta + iz}{\eta^2 + z^2} e^{-\eta^2}\mathrm{d}\eta \\ &= \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{z}{\eta^2 + z^2} e^{-\eta^2}\mathrm{d}\eta. \end{aligned} \tag{3.88}$$

Indeed, the imaginary part is zero since it corresponds to an odd function integrated over a symmetric interval $(-\infty, +\infty)$. From Eq.(3.88), we deduce that $H_{-1}(z)$ is an odd function of z. Then we can study $H_{-1}(z)$ for $z > 0$. Under this hypothesis, we define $u = \eta/z$ and we obtain

$$H_{-1}(z) = \frac{1}{2\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{e^{-u^2z^2}du}{1 + u^2} \quad (z > 0), \tag{3.89}$$

from which we easily deduce the particular value $\lim_{z\to 0^+} H_{-1}(z) = \sqrt{\pi}/2$. In addition, Eq.(3.89) allows us to get a differential equation for $H_{-1}(z)$ $(z > 0)$. By differentiation, we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}z} H_{-1}(z) &= -\frac{z}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{u^2 e^{-u^2z^2}du}{1 + u^2} \\ &= -\frac{z}{\sqrt{\pi}} \int_{-\infty}^{+\infty} \left(\frac{1 + u^2}{1 + u^2} - \frac{1}{1 + u^2}\right) e^{-u^2z^2}du \\ &= -1 + 2zH_{-1}(z). \end{aligned} \tag{3.90}$$

Now, the differential equation $H'_{-1}(z) = -1 + 2zH_{-1}(z)$ is linear and can be directly solved for $z > 0$ with the initial condition $H_{-1}(0^+) = \sqrt{\pi}/2$ proved above. A straightforward

calculation yields

$$H_{-1}(z) = \frac{\sqrt{\pi}}{2} e^{z^2} \left[1 - \mathrm{erf}\,(z)\right], \quad z > 0. \tag{3.91}$$

Since $H_{-1}(z)$ is odd, i.e. $H_{-1}(-z) = -H_{-1}(z)$, we can also calculate its values for $z < 0$. By means of the first two functions $H_0(z) = 1\ \forall z$ and $H_{-1}(z)$ given in Eq.(3.91), we can determine all other Hermite elements with negative index through the following recursive formula

$$H_{-(n+1)}(z) = \frac{1}{2n} H_{-(n-1)}(z) - \frac{z}{n} H_{-n}(z). \tag{3.92}$$

Following again Ref. [172], Eq.(3.92) can be proved as

$$\begin{aligned} H_{-(n-1)}(z) &= \frac{1}{(2i)^{n-1}\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{e^{-\eta^2} \mathrm{d}\eta}{(\eta - iz)^{n-1}} \\ &= \frac{1}{(2i)^{n-1}\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{\eta - iz}{(\eta - iz)^{n}} e^{-\eta^2} \mathrm{d}\eta \\ &= \frac{1}{(2i)^{n-1}\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{\eta e^{-\eta^2} \mathrm{d}\eta}{(\eta - iz)^{n}} \\ &\quad - \frac{iz}{(2i)^{n-1}\sqrt{\pi}} \int_{-\infty}^{+\infty} \frac{e^{-\eta^2} \mathrm{d}\eta}{(\eta - iz)^{n}}. \end{aligned} \tag{3.93}$$

Now, the first integral in the result of Eq.(3.93) can be elaborated by parts, eventually obtaining

$$\int_{-\infty}^{+\infty} \frac{\eta e^{-\eta^2} \mathrm{d}\eta}{(\eta - iz)^{n}} = -\frac{n}{2} \int_{-\infty}^{+\infty} \frac{e^{-\eta^2} \mathrm{d}\eta}{(\eta - iz)^{n+1}}. \tag{3.94}$$

Thus, Eq.(3.93) becomes $H_{-(n-1)}(z) = 2nH_{-(n+1)}(z) + 2znH_{-n}(z)$, which is equivalent to Eq.(3.92). This recursive law allows us to affirm that $H_{-n}(z)$ is odd if n is odd and that $H_{-n}(z)$ is even if n is even. We underline that the Hermite elements with $n < 0$ are not polynomials, contrarily to the classical Hermite polynomials, defined with $n > 0$. Many other interesting properties of this sequence of functions can be found in Ref. [172]. For the sake of completeness, the first functions are reported below for $z > 0$

$$\begin{aligned} H_0(z) &= 1, \\ H_{-1}(z) &= \tfrac{\sqrt{\pi}}{2} e^{z^2} \mathrm{erfc}\,(z), \\ H_{-2}(z) &= \tfrac{1}{2} - \tfrac{\sqrt{\pi}}{2} z e^{z^2} \mathrm{erfc}\,(z), \\ H_{-3}(z) &= -\tfrac{1}{4} z + \tfrac{\sqrt{\pi}}{4} \left(\tfrac{1}{2} + z^2\right) e^{z^2} \mathrm{erfc}\,(z), \\ H_{-4}(z) &= \tfrac{1}{12} \left(1 + z^2\right) - \tfrac{\sqrt{\pi}}{4} z \left(\tfrac{1}{2} + \tfrac{1}{3} z^2\right) e^{z^2} \mathrm{erfc}\,(z), \end{aligned} \tag{3.95}$$

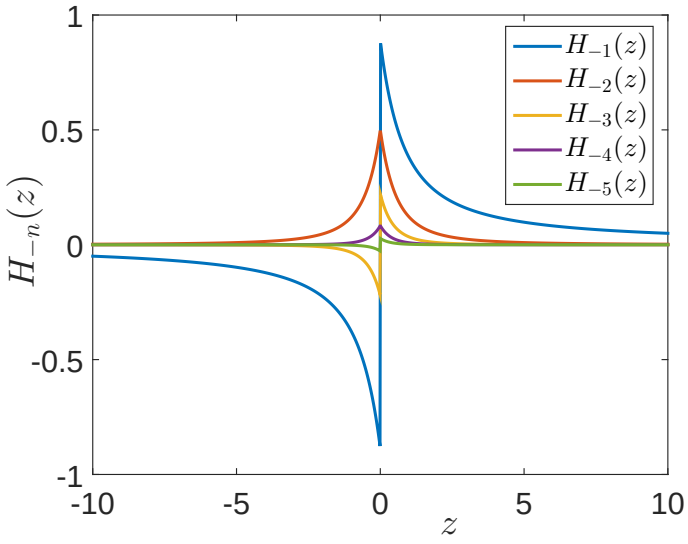

Figure 3.6 – Plot of the first Hermite elements with negative index. We observe that $H_{-n}(z)$ is odd if n is odd and that that $H_{-n}(z)$ is even if n is even [1].

where $\mathrm{erfc}(z) = 1 - \mathrm{erf}(z)$ is the complementary error function. Their plot can be found in Fig.3.6. From the computational point of view, the calculation of $e^{z^2}\mathrm{erfc}\,(z)$ must be done with the following asymptotic expansion (for $z > 4$) [170, 171]

$$\sqrt{\pi} z e^{z^2} \mathrm{erfc}\,(z) \sim 1 + \sum_{m=1}^{M} (-1)^m \frac{1 \cdot 3 \cdot ... (2m-1)}{(2z^2)^m}, \tag{3.96}$$

with a sufficiently large value of M.

3.3.3 The partition function and related results

We can now summarise the obtained results and address the determination of the Helmholtz partition function. Coming back to the calculation of the integral defined in Eq.(3.85), we have finally obtained the following result, written in terms of the functions $H_{-m}(z)$

$$J_m(x) = (2i)^m \sqrt{\pi} H_{-m}(-x). \tag{3.97}$$

The integral $I_m(a,b)$ defined in Eq.(3.58) can be therefore solved by Eq.(3.63) for $a>0$ and by Eq.(3.84) for $a<0$, rewritten and unified here in the compact form

$$\begin{aligned} I_m(a,b) &= \sqrt{\pi}\left\{ b^{\frac{m-1}{2}} e^{-\frac{a^2}{4b}} (2i)^m H_{-m}\left(-\frac{a}{2\sqrt{b}}\right) \right. \\ &\left. +\mathbf{1}(-a)\sum_{h=0}^{m-1} \frac{(ia)^{m-1-h}(-4b)^{h/2}}{ih!(m-1-h)!}\Gamma\left(\frac{h+1}{2}\right)\left[1+(-1)^h\right] \right\}, \end{aligned} \tag{3.98}$$

where $\mathbf{1}(x)$ represents the Heaviside step function, defined as $\mathbf{1}(x)=1$ if $x\geq 0$, and $\mathbf{1}(x)=0$ if $x<0$. We can write that

$$\begin{aligned} Z_H(r) &= \frac{1}{2^N}\frac{1}{i^{N+1}}\frac{1}{r}\sum_{k=0}^{N}\sum_{p=0}^{N-k}\sum_{q=0}^{k}\binom{N}{k}\binom{N-k}{p}\binom{k}{q}(-1)^{p+q} \\ &\times(\chi\phi)^k\left[\sqrt{\pi}\left(\frac{N\alpha}{2}\right)^{\frac{N-2}{2}}(2i)^{N_1}e^{-\frac{\Lambda^2}{2N\alpha}}H_{-N+1}\left(\frac{\Lambda}{\sqrt{2N\alpha}}\right)-\mathbf{1}(\Lambda)2\pi i\right. \\ &\left.\times\sum_{h=0}^{m-1}\frac{(ia)^{m-1-h}}{h!(m-1-h)!}\frac{(-2N\alpha)^{h/2}}{2\sqrt{\pi}}\Gamma\left(\frac{h+1}{2}\right)\left[1+(-1)^h\right]\right]. \end{aligned} \tag{3.99}$$

To conclude, we can simplify Eq.(3.98) to obtain the final form of the Helmholtz partition function

$$\begin{aligned} Z_H(r) &= -\frac{1}{2r}\sum_{k=0}^{N}\sum_{p=0}^{N-k}\sum_{q=0}^{k}\binom{N}{k}\binom{N-k}{p}\binom{k}{q}(-1)^{p+q} \\ &\times(\chi\phi)^k\left[\sqrt{\pi}\left(\frac{N\alpha}{2}\right)^{\frac{N-2}{2}}e^{-\frac{\Lambda^2}{2N\alpha}}H_{-N+1}\left(\frac{\Lambda}{\sqrt{2N\alpha}}\right)\right. \\ &\left.+\mathbf{1}(\Lambda)\sum_{h=0}^{\left[\frac{N}{2}-1\right]}\frac{(-1)^{N-1}\pi\left(\frac{\Lambda}{2}\right)^{N-2}}{h!(N-2-2h)!}\left(\frac{N\alpha}{2\Lambda^2}\right)^h\right], \end{aligned} \tag{3.100}$$

where $[x]$ represents the floor function giving the greatest integer that is less than or equal to x and Λ depends on k, p and q through Eq.(3.57). This result is the most important achievement of this Section and allows us to determine the complete response of the two-state freely jointed chain with extensible lengths in terms of force-extension curve and average value of the spin variables. Interestingly enough, Eq.(3.100) represents a closed form expression since all sums are performed over finite ranges and the Hermite elements are known recursively, as discussed above. As a check of the procedure, we can see that if $\alpha=0$ (i.e. infinite stiffness) the first term within the brackets vanishes and the second

one is given only by the addend with $k = 0$. Hence, the partition function simplifies to give

$$Z_H(r) = \frac{\pi}{2^{(N-1)}(N-2)!r} \sum_{k=0}^{N} \sum_{p=0}^{N-k} \sum_{q=0}^{k} \binom{N}{k} \binom{N-k}{p} \times \binom{k}{q} (-1)^{p+q} (\chi\phi)^k (-\Lambda)^{N-2} \mathbf{1}(\Lambda), \tag{3.101}$$

in perfect agreement with Chapter 2. In any case, we remark that Eqs.(3.100) and (3.101) can not be written as a power with exponent N. It means that within the Helmholtz ensemble there is an effective interaction among the elements induced by the isometric boundary conditions. However, this interaction is not explicitly written in the Hamiltonian of the system (as, e.g., in the Ising model) but comes from the prescribed positions of the two chain extremities.

All thermodynamic properties can be evaluated through the Helmholtz free energy $\mathcal{F} = -k_B T \log Z_H$. For instance, the force-extension response is given by the expression

$$\langle f \rangle = \frac{\partial \mathcal{F}}{\partial r}, \tag{3.102}$$

and the average value of the spin variable is as

$$N \langle s \rangle = \frac{\partial \mathcal{F}}{\partial \Delta E} = -\frac{\partial \mathcal{F}}{\partial \mu}. \tag{3.103}$$

Of course, both $\langle f \rangle$ and $\langle s \rangle$ could be written in closed form by performing the indicated derivatives. However, for the sake of brevity, we omit this development and we show an example of application of Eqs.(3.102) and (3.103) in Fig.3.7, where we compare an elastic system described by Eq.(3.100) with a stiff one described by Eq.(3.101). In the top panel, we show the force-extension curve for the stiff system with $k \to \infty$ and for the elastic one with $k = 1.5\text{N/m}$. Moreover, to better compare the responses, we added the Gibbs force-extension responses for both cases. In the bottom panel, the average value of the spin variables is also represented for the stiff and the soft systems. These results prove a non-cooperative behaviour characterised by a progressive unfolding of units in response to the increasing overall length. This behaviour corresponds to a series of peaks in the force-extension curves and to a staircase function for the spin variable. Any peak or step corresponds to an unfolding process induced by the increasing extension. The overall behaviour is therefore completely different from the Gibbs ensemble, where all units undergo the conformational transition at the same time. Anyway, for the soft system, we observe a sensibly reduced peak-to-peak force in the force-extension curve and,

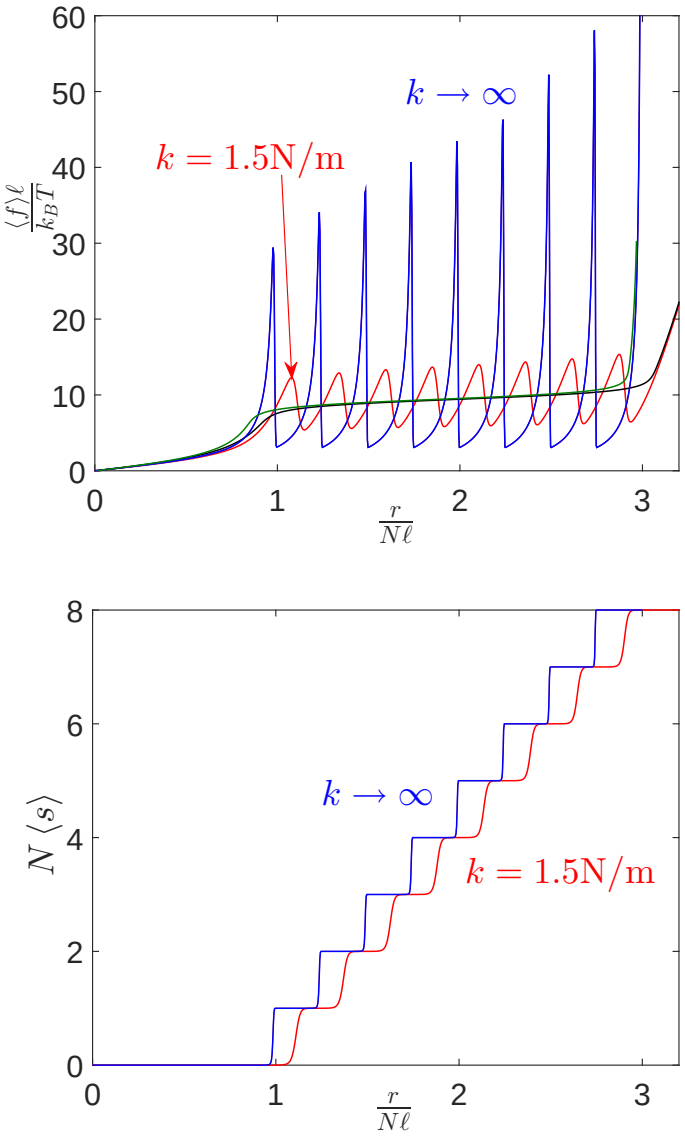

Figure 3.7 – Force-extension response (top panel) and average spin (bottom panel) for a system with extensibility ($k = 1.5\mathrm{N/m}$) and a system without extensibility ($k \to \infty$). The calculation for the system with extensibility is based on Eq.(3.100) (red curves) while the system without extensibility has been studied through Eq.(3.101) (blue curves). In the top panel, we also added the Gibbs force-extension responses for both cases. We adopted the parameters $N = 8$, $\ell = 0.5 \times 10^{-9}\mathrm{m}$, $T = 300\mathrm{K}$, $\chi = 3$, $\Delta E = 20 k_B T = 8.28 \times 10^{-20}\mathrm{J}$ and $\mu = 0$ [1].

correspondingly, a smoother transition in the steps of the average spin variable. Therefore, it is important to underline that relatively soft polymers (with small elastic constant) can have peak-to peak forces much smaller than ones predicted by the two-state freely jointed

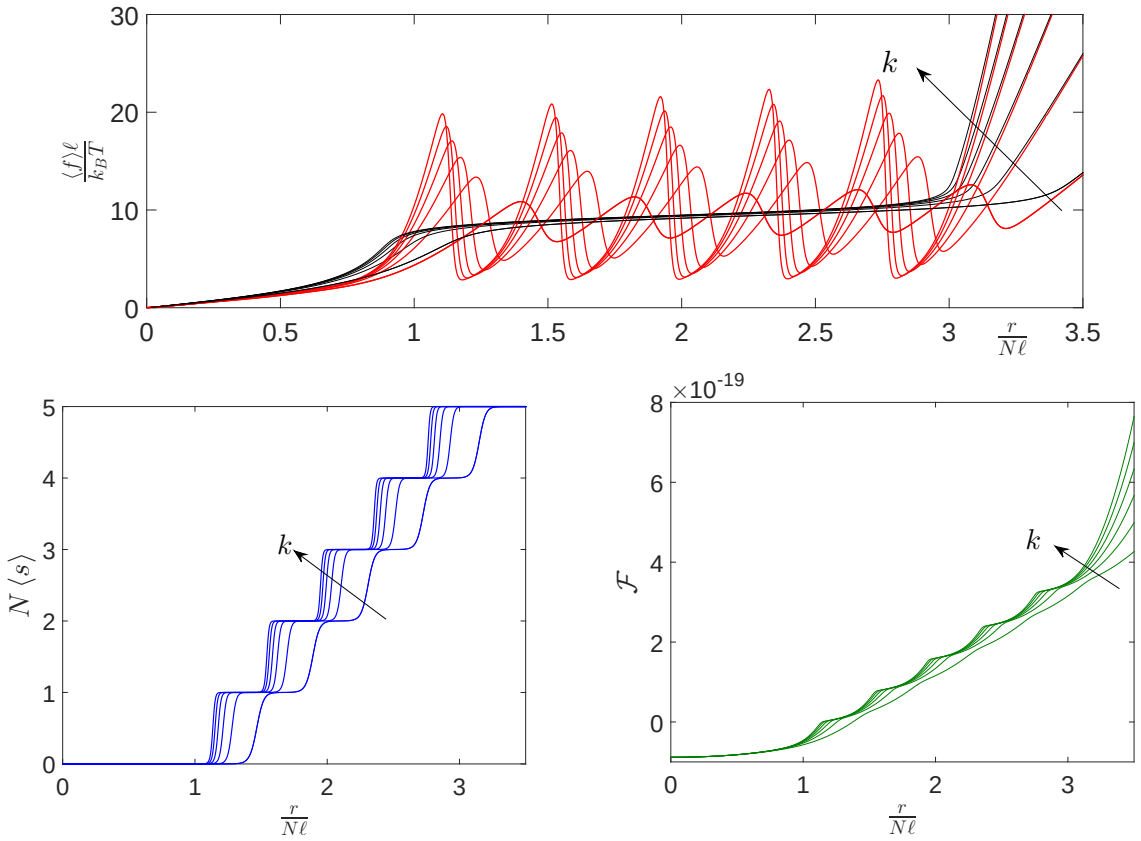

Figure 3.8 – Force-extension response (top panel), average spin variable (bottom-left panel) and Helmholtz free energy (bottom-right panel) for a system with variable extensibility $k = 0.4, 0.8, 1.2, 1.6, 2.0$ and 2.4N/m. The calculation is based on Eq.(3.100). We also added the Gibbs force-extension responses (black curves) to facilitate the comparison. We adopted the parameters $N = 5$, $\ell = 0.5 \times 10^{-9}$m, $T = 300$K, $\chi = 3$, $\Delta E = 20 k_B T = 8.28 \times 10^{-20}$J and $\mu = 0$ [1].

chain.

In Fig.3.8, one can find the force-extension response (top panel), the average spin variable (bottom-left panel) and the Helmholtz free energy (bottom-right panel) for a system with variable extensibility $k = 0.4, 0.8, 1.2, 1.6, 2.0$ and 2.4N/m. We can observe the progressive increase of the peak-to-peak force of each unfolding transition with increasing values of the elastic constant. At the same time, the transition are sharper for stiffer systems, as clearly visible in the average spin curves. The origin of the force peaks in the force-extension curve can be highlighted in the Helmholtz free energy curves, which are characterised by a series of "smooth cusps" able to generate the force peaks by derivation with respect to the position.

A further analysis concerns the variation of the peak-to-peak force with the elastic

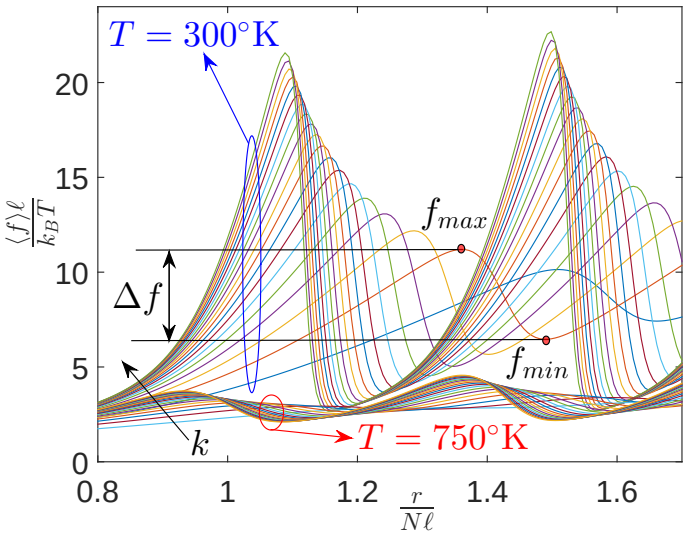

Figure 3.9 – Shape of the first peaks in the force-extension curves for two temperatures $T = 300\mathrm{K}$ and $T = 750\mathrm{K}$ and for several values of $k = (0.3 + 0.15n)\mathrm{N/m}$, $n = 0, ..., 18$. One can see the exact definition of the peak-to-peak force for the first peak holding for each force-extension curve. We adopted the parameters $N = 5$, $\ell = 0.5 \times 10^{-9}\mathrm{m}$, $\chi = 3$, $\Delta E = 8.28 \times 10^{-20}\mathrm{J}$ and $\mu = 0$ [1].

constant k and the temperature T. In Fig.3.9, we can observe the shape of the first peaks for two temperatures $T = 300\mathrm{K}$ and $T = 750\mathrm{K}$ and for several values of $k = (0.3 + 0.15n)\mathrm{N/m}$, $n = 0, ..., 18$. We can observe the strong reduction of the force peak with increasing elastic constant and/or temperature. Concerning the decrease of the peak with the temperature, we remark that the spinoidal behaviour of the force-extension curve (negative slope in some regions) is characterised by a critical temperature, defined as follows. For supercritical temperatures, the slope (or the effective stiffness) is always positive and for subcritical temperatures, it is negative in some regions. This concept of critical temperature for the chain within the Helmholtz ensemble is similar to the concept of Curie temperature governing the ferromagnetic-paramagnetic transition in magnetic materials. This point has been recently discussed in Refs. [111–113]. Interestingly enough, we notice that the chain with a negative effective stiffness (spinoidal behaviour), exhibited in specific working conditions, can be viewed as a metamaterial [111] and this property can be mimicked in bioinspired nanostructures [173]. In Fig.3.10, one can find the the variation of the peak-to-peak force (as defined in Fig.3.9) in terms of elastic constant and temperature. We can see that the force peak becomes negligible for low values of k and high values of T. This plot shows the importance of considering the correct value of k

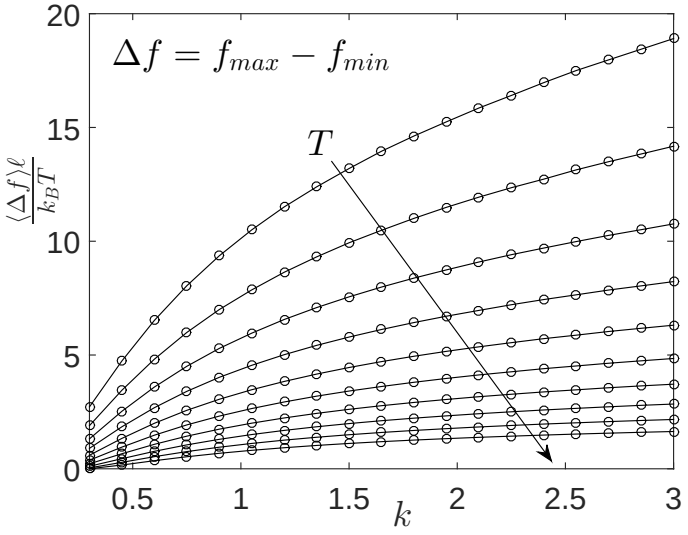

Figure 3.10 – Behaviour of the peak-to-peak force defined in Fig.3.9 versus elastic constant and temperature. We varied the elastic constant in the range $k = (0.3 + 0.15n)$N/m, $n = 0, ..., 18$ and the temperature as $T = (300 + 50m)$, $m = 0, ..., 9$. We observe the strong reduction of the force peak with the increasing of temperature and/or elastic constant. We adopted the parameters $N = 5$, $\ell = 0.5 \times 10^{-9}$m, $\chi = 3$, $\Delta E = 8.28 \times 10^{-20}$J and $\mu = 0$ [1].

for modelling the folding-unfolding processes of macromolecular chains. Indeed, the force peaks are typically considered as output of real experiments and their values are used to draw quantitative conclusions about mechanical and thermodynamic properties of the macromolecule.

Finally, we show in Fig.3.11 the behaviour of the system as a function of the number of units of the chain. This is an important analysis since it concerns the validity of the ensembles equivalence in the thermodynamic limit. In Fig.3.11, one can find the results for N from 2 to 8 for a system with an elastic constant $k = 1$N/m. It is interesting to remark that for an increasing value of N, the peak-to-peak force is progressively reduced, confirming the convergence of the Helmholtz ensemble to the Gibbs ensemble for $N \to \infty$. Indeed, if two ensembles are equivalent, the corresponding force-extension curves must coincide. This is perfectly coherent with known results concerning the ensembles equivalence in the thermodynamic limit, as discussed in recent literature [20, 139].

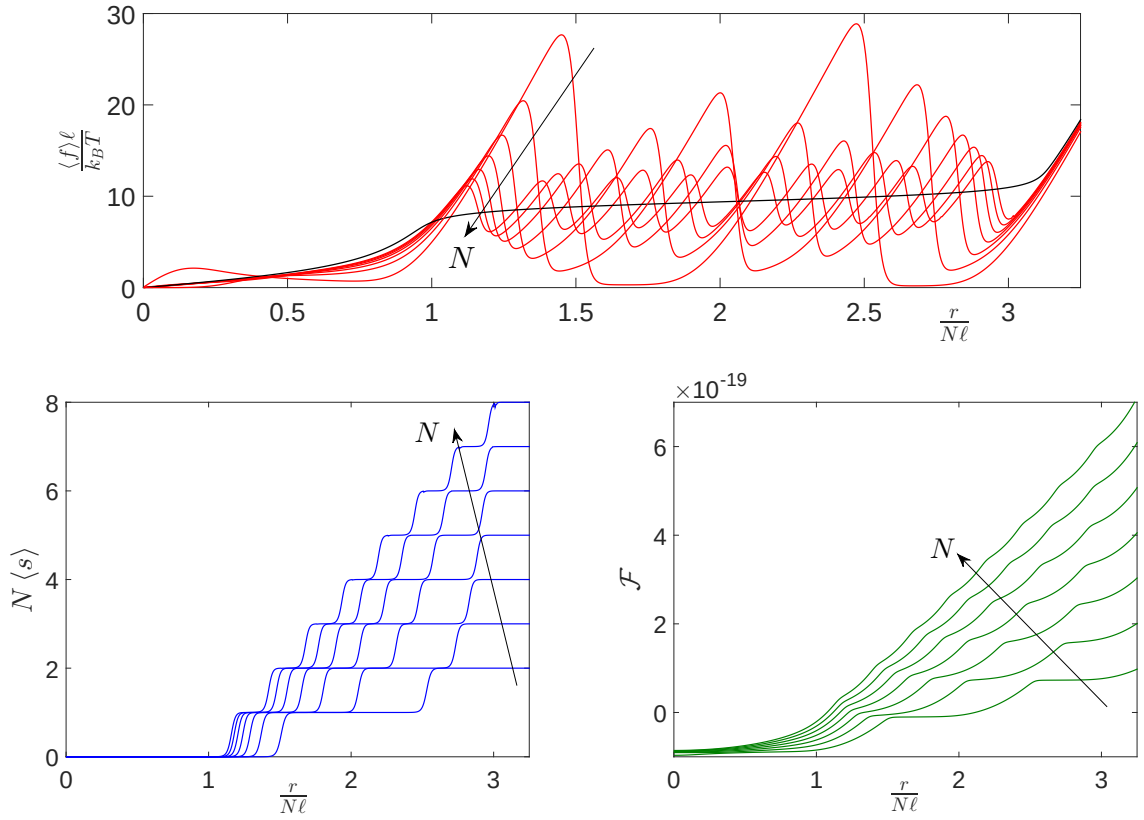

Figure 3.11 – Force-extension response (top panel), average spin variable (bottom-left panel) and Helmholtz free energy (bottom-right panel) for a system with a variable number of units $N = 2, 3, 4, 5, 6, 7$, and 8. The calculation is based on Eq.(3.100). We also added the Gibbs force-extension responses (black curves) to facilitate the comparison. We adopted the parameters $k = 1\text{N/m}$, $\ell = 0.5 \times 10^{-9}\text{m}$, $T = 300\text{K}$, $\chi = 3$, $\Delta E = 20 k_B T = 8.28 \times 10^{-20}\text{J}$ and $\mu = 0$ [1].

3.4 Conclusion

To conclude, we provided in detail the analysis of a two-state freely jointed chain with extensible bonds. Despite the complexity of the calculation of the partition function, especially within the Helmholtz ensemble, we added the extensibility to the units, an important point to describe real macromolecules. To exactly calculate the Gibbs partition function, we applied an approximation considering relatively high values of the elastic constant defining the extensibility of the units. This is coherent according to the parameters of most of macromolecules of biological interest. Concerning the Gibbs ensemble, we obtained the exact partition function by direct integration. The latter is then used to calculate the elastic response of the system and the average value of the spin variable. We observed that for a given threshold force, the spin exhibits a transi-

tion, showing the cooperative (or synchronised) behaviour of the chain within the Gibbs ensemble. Regarding the Helmholtz ensemble, the implicit interaction induced by the isometric condition complicates the direct calculation of the partition function. Indeed, the partition function cannot be factorised due to the interdependence of the units extensions. Both partition functions can be connected through a Laplace transform. This property is typically exploited in classical literature to prove the equivalence of the ensembles in the thermodynamic limit [120, 124]. Here, this property is used to obtain the closed form expression for the Helmholtz partition function in terms of the Hermite elements with a negative index. The Hermite polynomials have been introduced in literature in Ref. [172], however, to the author's knowledge, the study here presented is the first application to a physical problem. Contrarily to the Gibbs ensemble, the average spin variable, calculated with the help of the Helmholtz partition function, exhibits a stepwise behaviour, showing the non-cooperative (or non-synchronised) behaviour of the chain. Moreover, as seen in Fig.3.8, the force-extension response within the Helmholtz ensemble depends on the intrinsic stiffness of the units. A weak stiffness of the chain may strongly reduce the amplitude of the force peaks of the response. It means that for force-extension experiments, stiff macromolecules can be rather used than soft macromolecules to study cooperativity and other conformational properties. To conclude, in this Chapter, the spin variable approach is adopted to study bistable chains with extensibility. The next step, to bring our model closer to reality, involves the consideration of the possible interactions existing among units.

Chapter 4

Two-state freely jointed chain with Ising interactions

4.1 Introduction

The purpose of this Chapter is to introduce in the model interactions existing among the units of a macromolecular chain, with the help of the Ising scheme, combined with the spin variable method. We remember that, from the historical point of view, the statistical models with interactions have been introduced to explain phase transitions in condensed matter. For instance, Wilhelm Lenz invented the original Ising model and gave it as a problem to his student, Ernst Ising. The one-dimensional model was solved by Ising himself during his PhD in 1924. The obtained result, contrarily to the expectations, did not describe any phase transition. Two decades later, the two-dimensional model was solved by Lars Onsager (see Fig.4.1). This quite complex calculation showed for the first time a phase transition in an exactly solvable model of the statistical mechanics. In following studies, the Ising scheme (and many other interaction schemes) have been used to describe different features of the phase transitions but also to develop models with complex interactions in other physical and biophysical directions. In our context, we adopt the Ising scheme to introduce a cooperative or anti-cooperative behaviour in the folding and unfolding processes of the units of a chain. The study of interactions existing among units is mainly motivated by repeat tandem proteins [114]. These proteins are defined as several (at least, two) adjacent copies having the same or very similar sequence motifs, generated by internal duplications. The repetition of the sequence ranges from a single amino-acid up to domains of 100 or more residues. Interactions among units of the chain do not depend on the place of the units in the sequence. Indeed, even if units are far apart in the sequence, they can be next to each other in the configurational space, which may lead to interactions.

STATISTICAL PHYSICS

90 years of the Ising model

Ernst Ising's analysis of the one-dimensional variant of his eponymous model (*Z. Phys* **31,** 253–258; 1925) is an unusual paper in the history of early twentieth-century physics. Its central result — demonstrating that a linear chain of two-state spins cannot undergo a phase transition at finite temperature — is correct, if somewhat trivial compared with other physics breakthroughs published in the 1920s. But it is Ising's fateful extension of his conclusions to two and three dimensions that proved spectacularly wrong and, paradoxically, earned him an enduring association with the model that now bears his name.

A possible reason for Ising's unexpected celebrity is that his erroneous conclusions betray a superficial understanding of what turned out to be some of the deepest and far-reaching problems to be addressed in twentieth-century physics. The Hamiltonian of the model is simple to write down — it describes a network of spins interacting with each other through a coupling that only applies if the spins are next to each other — but the physics it displays is rich and non-trivial: not only does it provide an intuitive device for illustrating the essential features of phase transitions and critical phenomena, it neatly encapsulates the main traits of

the many-body problem that has come to dominate areas such as condensed-matter physics. The broader class of spin models it belongs to was used to uncover concepts such as universality, renormalization, symmetry-breaking and emergence. Ising can perhaps be forgiven for not predicting all of that.

Famously, the two-dimensional version for the model was solved analytically by Lars Onsager in the early 1940s (*Phys. Rev.* **65,** 117; 1944), a result that is rightly considered a towering achievement among many significant contributions made over the years by the likes of Peierls, Bethe, Yang, Kadanoff (see page 995) Fisher and Wilson, just to name a handful. But the three-dimensional lattice has never been solved exactly, in spite of a multitude of attempts and false dawns — including a claim by John Maddox (who would later become the editor of *Nature*) made at a conference in Paris in 1952.

Although the 3D model is thought by some to be analytically intractable (and has also been claimed to belong to the NP-complete category of computational decision problems), progress has continued and recent numerical techniques based on conformal field theory have shed further light on the structure of the problem (*J. Stat. Phys.* **157,** 869–914; 2014). Nevertheless, the real value of the Ising model and its many derivatives lies precisely in the complexity they encapsulate. These have found use in fields as disparate as condensed-matter physics, physical chemistry, neuroscience and, more broadly, the study of so-called complex systems.

Ising studied a deceptively simple model that, unknown to him at the time, captures the essential physics of an extremely wide category of problems. He may have been wrong in his 1925 paper, but he tripped over a veritable physics goldmine.

ANDREA TARONI

Figure 4.1 – Insert published on the occasion of the 90th anniversary of the Ising model [174].

A first reported case concerns the tandem repeats in red cell spectrin [114], where two units simultaneously unfold because of their strong interaction. Similar observations have been made on Filamin A [154], where domain-domain interactions lead to a hierarchy of unfolding forces and the Ising model is appropriate to study them. Indeed, we can introduce a positive Ising contribution which favours the same state (folded or unfolded) for adjacent units, or a negative Ising contribution which favours the different state for adjacent units. While these forms of interactions have been observed in specific couples of units in real protein chains, we introduce in our model the interaction term for all couples of adjacent units to simplify the analysis.

Hence, the systems considered in this Chapter are bistable and cooperative. Therefore, each unit can undergo transitions from folded to unfolded states or from unfolded to folded states, and the transition of one unit affects the transition of the others, favourably or not. The Ising model has been chosen to better understand both bistable and cooperative systems because it is one of the simplest interaction scheme and it directly works on the spin variables. The Ising coefficient, λ, is the parameter enabling to establish the desired type of interaction. This Ising coefficient allows to consider, for instance, the case of a "positive" interaction with $\lambda > 0$, leading to the unfolding of units favoured by

the previous unfolding of other units. A "negative" interaction with $\lambda < 0$ can also be considered, when the unfolding of units can prevent other units from unfolding. We provide evidence that the cooperativity, measured by the Ising interaction coefficient, strongly modifies the force-extension response of the chain, and its configurational properties. This scheme is not only useful to better understand the behaviour of real macromolecules, but it also enables to explain the nucleation stress in nanowires [96–98], an important topic in material science. While the Gibbs ensemble will be studied by means of the classical transfer matrix method [175], typically adopted for one-dimensional interacting models, the Helmholtz ensemble presents major difficulties and will be approached once again by exploiting the Laplace transform relationship between the Gibbs and Helmholtz partition functions [120].

We take into consideration a chain of N two-state elements (see Fig.4.2.a), each described by a bistable potential energy with a stable folded state and a metastable unfolded state (see Fig.4.2.b). The two potential wells in Fig.4.2.b can be characterised by the elastic constant $k(S_i)$, the equilibrium length $\ell(S_i)$ and the basal energy $v(S_i)$, where S_i is a discrete variable (or spin variable) assuming values in $\{-1, +1\}$, used to distinguish one well from the other. We state that $S_i = +1$ corresponds to unfolded elements, whereas $S_i = -1$ corresponds to folded ones. This description suggests that the bistable energy potential can be represented by two quadratic potentials approximating the real wells of the units (see again Fig.4.2.b) [113]. In this case, the discrete variables belong to the phase space of the system and allow to specify the explored well for each unit. The introduction of the discrete or spin variables also allows the direct implementation of an interaction between adjacent elements of the chain, *e.g.* described by a classical Ising Hamiltonian. The overall Hamiltonian of this system can be therefore written as

$$H = -\lambda \sum_{i=1}^{N-1} S_i S_{i+1} - \mu \sum_{i=1}^{N} S_i + \sum_{i=1}^{N} \left[v(S_i) + \frac{1}{2} k(S_i) \left[\|\vec{r}_i - \vec{r}_{i-1}\| - \ell(S_i) \right]^2 \right]. \qquad (4.1)$$

The first term in Eq.(4.1) represents the Ising interaction among spin variables, the second term represents the effect of a chemical potential, and the third one describes the spring-like behaviour of each unit placed between positions $\vec{r}_{i-1}$ and $\vec{r}_i$, $\forall i = 1, ..., N$. We remark that $\lambda > 0$ tries to force all elements to be folded or unfolded (ferromagnetic-like interaction), whereas $\lambda < 0$ tries to force all elements to be alternatively folded and unfolded (anti-ferromagnetic-like interaction). The parameter μ is a sort of external field or chemical potential ($\mu > 0$ tries to unfold the domains and $\mu < 0$ tries to fold them).

In the following, we suppose to embed the system in a thermal bath at the temperature T, we consider the system at thermodynamic equilibrium and we study the effects of

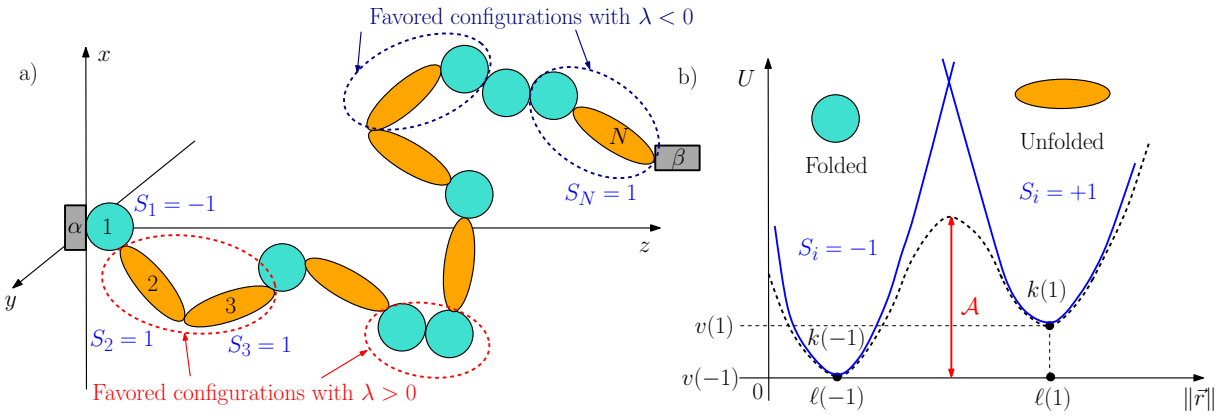

Figure 4.2 – a) Chain of N two-state units with Ising interactions. While the first end-terminal α is able to tether the first unit to a given substrate, the second one β is able either to apply a force (Gibbs condition) or a position (Helmholtz condition) to the last unit. b) Potential energy of a single unit of the chain (dashed black curve). The potential wells are approximated through two parabolic (i.e. quadratic) profiles (solid blues curves), identified by $S_i = -1$ (folded state) and $S_i = +1$ (unfolded state) [2].

the Ising interactions on the mechanical and configurational behaviour within the Gibbs (applied external force) or the Helmholtz (prescribed end-to-end distance) ensembles (see Fig.4.2.a).

4.2 Example of biological cooperativity

Tandem-repeat domains are frequently found protein architectures. In this Section, we briefly discuss the cooperativity in the solenoid class of the tandem-repeat proteins comprising repeats of approximately 12–40 amino acids [176]. The simplest solenoid proteins contain repeats of two secondary structure elements: α/α, α/β or β/β. More complex repeats have three or four secondary elements. The 'solenoid' term originally referred to a coil wound into a tightly packed helix. The repeats pack forms super-helices that differ greatly in their geometries, as shown in Fig.4.3.

Tandem-repeat protein structures are exclusively stabilised by local interactions either within a repeat or between adjacent repeats. By contrast, the stability of globular proteins originates from the high cooperativity between sequence-distant interactions. The simple topology of the repeat-protein architecture enables the use of a one-dimensional Ising model description to define the energetic values of each repeat under the assumption of all repeats being coupled. Indeed, their states are coupled to their nearest neighbours through an exchange interaction, a potential that favours parallel alignment between

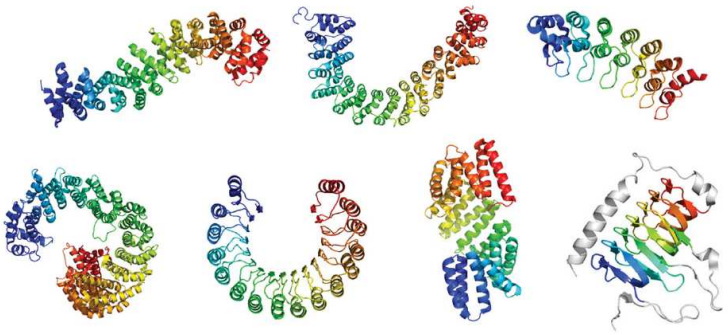

Figure 4.3 – Secondary structure representations of solenoid tandem-repeat proteins. From top left to bottom right (PDB identifiers in parenthesis): ARM-repeat protein β-catenin (2Z6H), HEAT-repeat protein PR65 (1B3U), Ankyrin-repeat protein gankyrin (1UOH), HEAT-repeat protein Importin-β (3ND2), leucine-rich repeat (LRR) protein Ribonuclease Inhibitor (1BNH), TPR protein RapI (4I1A), β-helical repeat protein carbonic anhydrase (1QRE) [176].

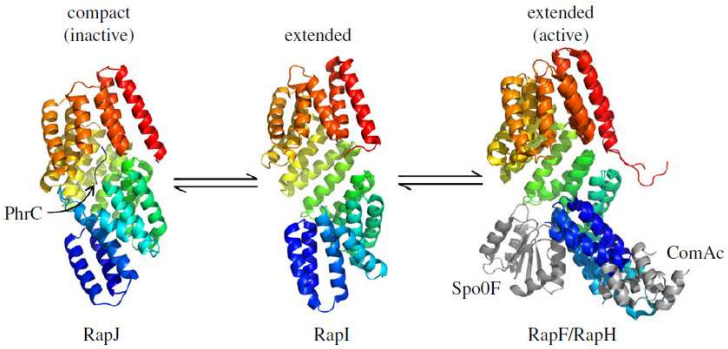

Figure 4.4 – Structures of different Rap proteins (C-terminus in red) depicting a possible mode of action. When the TPR domain binds to a signalling peptide, it causes the Rap protein to adopt a compact, or 'closed' conformation. Upon binding an interaction partner, however, conformational changes in the TPR domain are minimal, whereas the N-terminal three-helix bundle flips by approximately 180 ° [176].

states. Folding cooperativity of repeat proteins breaks down above approximately 100–150 amino acids, similar to the cooperativity limit of globular proteins.

In recent years, a family of bacterial regulators has been largely investigated: the RRNPP family. The name of these peptide-sensing regulators refers to the founding members of the family, Rap-Rgg-NprR-PlcR-Prgx. They all have the same domain organisation: an N-terminal three-helix bundle, a flexible helical linker and a C-terminal tratricopeptide repeats (TPR, as shown in Fig.4.3) capable of binding short peptides of five to eight residues [176]. The homologous structures of Rap proteins have been used to propose a mechanism of action for signal transduction (see Fig.4.4). The compact

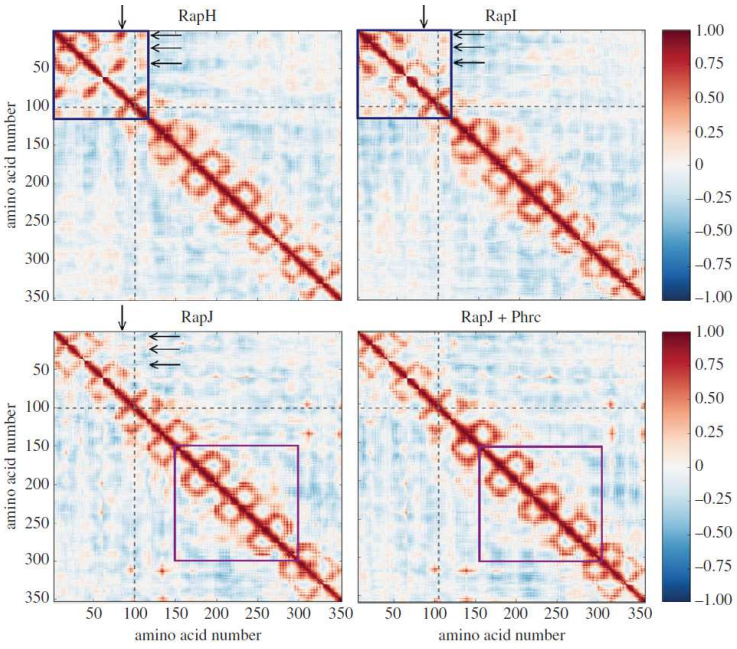

Figure 4.5 – Representative cross-correlation maps for the partner-bound, open and peptide-bound conformations. Cross-correlation between residues is a measure of how much these residues move in the same direction, where values of 1 and -1 represent perfectly correlated and anti-correlated motions, respectively. The TPR repeats exhibit correlated motions only with their nearest neighbours, giving rise to the distinctive pattern of squares along the diagonal. Movements of the N-terminal three-helix bundle, linker domain and first TPR motif (blue box) are non-TPR-like, exhibiting non-nearest-neighbour correlations, suggesting that they form a sub-domain relative to the rest of the TPR repeats. Some of these correlations are reduced in the open conformation, or even reversed, once a continuous TPR array is formed (arrows) and the distinction of this domain is lost. The global movement of peptide binding TPRs (purple box) and neighbouring repeats is only minimally affected in the presence of the peptide, which only causes a slight increase in the nearest-neighbour correlations. The N-terminal helix bundle and TPR repeats are divided by grey dashed lines and correlations are mirrored across the diagonal for clarity [176].

solenoid RapJ is the inactive configuration and the extended solenoid RapI is the active one. In its active configuration, the N-terminal helix bundle is capable of exposing the Spo0F- or ComA- binding regions of RapF and RapH, respectively. Peptide-bound Rap proteins undergo a conformational change locking the N-terminal domain in a compact configuration in which its binding sites are inaccessible.

The properties of RapJ are also studied when it is in complex with the PhrC peptide [176]. In Ref. [176], the authors analysed the correlation of motion between different residues of these structures and the results are reported in Fig.4.5. Such structures, as

clearly shown, exhibit correlated motion only with their nearest neighbours, giving rise to the distinctive pattern of squares along the diagonal and confirming the applicability of the Ising scheme. The binding of the peptide marginally increases nearest-neighbour correlation at the centre (purple box in Fig.4.5). Moreover, movements of the N-terminal three-helix bundle, linker domain and first TPRs repeat (blue box in Fig.4.5) are strongly correlated (non-nearest-neighbour), suggesting that they form a sub-domain relative to the rest of the TPR repeats. Finally, the Rap proteins are a good example of a system where nearest-neighbour interactions in a repeat array are sufficiently strong to justify a two-state modelling with a superposed Ising scheme.

4.3 Two-state freely jointed chain with Ising interactions: the Gibbs ensemble

We consider now the extended Hamiltonian

$$\begin{aligned} H_G &= -\lambda \sum_{i=1}^{N-1} S_i S_{i+1} - \vec{f} \cdot \vec{r}_N - \mu \sum_{i=1}^{N} S_i \\ &\quad + \sum_{i=1}^{N} \left[v(S_i) + \frac{1}{2} k(S_i) \left[\|\vec{r}_i - \vec{r}_{i-1}\| - \ell(S_i) \right]^2 \right]. \end{aligned} \tag{4.2}$$

Here, $\vec{f}$ is the applied force to the last element, identified by its position $\vec{r}_N$. We suppose that quantities $\vec{r}_i \in \mathbb{R}^3$ and $S_i \in \{-1, +1\} \, \forall i \in \{1...N\}$ belong to the phase space of the system. Moreover, to fix ideas, we always consider $\vec{r}_0 = 0$. The statistical mechanics of the system can be introduced by calculating the Gibbs partition function, as

$$\begin{aligned} Z_G(\vec{f}) &= \sum_{S_1 \in \{0,1\}} ... \sum_{S_N \in \{0,1\}} \int_{\mathbb{R}^3} ... \int_{\mathbb{R}^3} \exp\left[-H_G \frac{(\{S_i\}, \{\vec{r}_i\})}{k_B T} \right] \mathrm{d}\vec{r}_1 ... \mathrm{d}\vec{r}_N \\ &= \sum_{S_1 \in \{0,1\}} ... \sum_{S_N \in \{0,1\}} \exp\left(\frac{\lambda}{k_B T} \sum_{i=1}^{N-1} S_i S_{i+1} \right) \exp\left(\frac{\mu}{k_B T} \sum_{i=1}^{N} S_i \right) \\ &\quad \times \exp\left[-\frac{1}{k_B T} \sum_{i=1}^{N} v(S_i) \right] \int_{\mathbb{R}^{3N}} \exp\left(\frac{\vec{f} \cdot \vec{r}_N}{k_B T} \right) \\ &\quad \times \exp\left[-\frac{1}{2} \sum_{i=1}^{N} \frac{k(S_i)}{k_B T} \left[\|\vec{r}_i - \vec{r}_{i-1}\| - \ell(S_i) \right]^2 \right] \mathrm{d}\vec{r}_1 ... \mathrm{d}\vec{r}_N. \end{aligned} \tag{4.3}$$

The integral $I = \int_{\mathbb{R}^{3N}} ...\mathrm{d}\vec{r}_1...\mathrm{d}\vec{r}_N$ can be developed by means of the change of variables $\vec{\xi}_1 = \vec{r}_1 - \vec{r}_0; \vec{\xi}_2 = \vec{r}_2 - \vec{r}_1; ...; \vec{\xi}_N = \vec{r}_N - \vec{r}_{N-1}$, giving

$$I = \int_{\mathbb{R}^{3N}} \exp\left[-\frac{1}{2}\sum_{i=1}^{N}\frac{k(S_i)}{k_BT}\left[\|\xi_i\| - \ell(S_i)\right]^2\right]\exp\left(\frac{\vec{f}}{k_BT}\cdot\sum_{i=1}^{N}\vec{\xi}_i\right)\mathrm{d}\vec{\xi}_1...\mathrm{d}\vec{\xi}_N. \tag{4.4}$$

To further simplify the integral, we suppose that $\vec{f} = (0, 0, f)$ (the direction of the force can be fixed without limiting the generality of the calculation since the system is isotropic) and we introduce the spherical coordinates for the vectors $\vec{\xi}_i$: $\vec{\xi}_i = (\xi_i\cos\varphi_i\sin\theta_i; \xi_i\sin\varphi_i\sin\theta_i; \xi_i\cos\theta_i)$. Therefore, $\|\vec{\xi}_i\| = \xi$; $\vec{f}\cdot\vec{\xi}_i = f\xi_i\cos\theta_i$ and $\mathrm{d}\vec{\xi}_i = \xi_i^2\sin\theta_i\mathrm{d}\xi_i\mathrm{d}\varphi_i\mathrm{d}\theta_i$, and the integral becomes

$$\begin{aligned} I &= \int_{\mathcal{B}}\exp\left[-\frac{1}{2}\sum_{i=1}^{N}\frac{k(S_i)}{k_BT}\left[\xi_i - \ell(S_i)\right]^2\right]\exp\left(\sum_{i=1}^{N}\frac{f\xi_i\cos\theta_i}{k_BT}\right) \\ &\quad\times\prod_{i=1}^{N}\xi_i^2\sin\theta_i\mathrm{d}\xi_i\mathrm{d}\varphi_i\mathrm{d}\theta_i, \end{aligned} \tag{4.5}$$

with $\mathcal{B} = [(0;+\infty)\times(0;2\pi)\times(0;\pi)]^N$. The calculation can be developed as follows

$$\begin{aligned} I &= (2\pi)^N\int_{[0;+\infty]^N}\exp\left[-\frac{1}{2}\sum_{i=1}^{N}\frac{k(S_i)}{k_BT}\left[\xi_i - \ell(S_i)\right]^2\right]\left(\prod_{i=1}^{N}\xi_i^2\right) \\ &\quad\times\int_{[0;\pi]^N}\exp\left(\sum_{i=1}^{N}\frac{f\xi_i\cos\theta_i}{k_BT}\right)\prod_{i=1}^{N}\sin\theta_i\mathrm{d}\theta_i\mathrm{d}\xi_i \\ &= (2\pi)^N\prod_{i=1}^{N}\left\{\int_0^{+\infty}\exp\left[-\frac{1}{2}\frac{k(S_i)}{k_BT}\left[\xi_i - \ell(S_i)\right]^2\right]\xi_i^2\frac{2k_BT}{f\xi_i}\right. \\ &\quad\left.\times\sinh\left(\frac{f\xi_i}{k_BT}\right)\mathrm{d}\xi_i\right\} \\ &= (4\pi)^N\prod_{i=1}^{N}\left\{\int_0^{+\infty}\exp\left[-\frac{1}{2}\frac{k(S_i)}{k_BT}\left[\xi_i - \ell(S_i)\right]^2\right]\frac{\sinh\left(\frac{f\xi_i}{k_BT}\right)}{\frac{f\xi_i}{k_BT}}\xi_i^2\mathrm{d}\xi_i\right\}. \end{aligned} \tag{4.6}$$

Now, the integral I can be strongly simplified if we make the assumption to deal with a freely jointed chain model with elements of fixed lengths. It is equivalent to say that $k(+1) = k(-1) \to +\infty$. The case with finite stiffness of the wells will be considered in a following Section. If we use the property $\sqrt{\frac{\alpha}{\pi}}e^{-\alpha x^2} \to \delta(x)$, when $\alpha \to \infty$, we simplify the result for I, with c a constant, as

$$I = c\times\prod_{i=1}^{N}\int_0^{+\infty}\delta\left[\xi_i - \ell(S_i)\right]\frac{\sinh\left(\frac{f\xi_i}{k_BT}\right)}{\frac{f\xi_i}{k_BT}}\frac{\xi_i^2}{\ell^2}\mathrm{d}\xi_i, \tag{4.7}$$

where $\ell^2 = \ell^2(-1)$ i.e. $\ell = \ell(-1)$. We finally obtain

$$I = c \times \prod_{i=1}^{N} \frac{\sinh\left[\frac{f\ell(S_i)}{k_BT}\right]}{\frac{f\ell(S_i)}{k_BT}} \frac{\ell^2(S_i)}{\ell^2} \mathrm{d}\xi_i. \tag{4.8}$$

Eventually, the partition function assumes the simpler form

$$\begin{aligned} Z_G(\vec{f}) \;=\; & \sum_{S_1\in\{0,1\}} \ldots \sum_{S_N\in\{0,1\}} \left\{ \exp\left(\frac{\lambda}{k_BT}\sum_{i=1}^{N-1} S_iS_{i+1}\right) \exp\left(\frac{\mu}{k_BT}\sum_{i=1}^{N} S_i\right) \right. \\ & \left. \times \exp\left[-\frac{1}{k_BT}\sum_{i=1}^{N} v(S_i)\right] \prod_{i=1}^{N} \frac{\sinh\left[\frac{f\ell(S_i)}{k_BT}\right]}{\frac{f\ell(S_i)}{k_BT}} \frac{\ell^2(S_i)}{\ell^2} \mathrm{d}\xi_i \right\}. \end{aligned} \tag{4.9}$$

We have now to approach the problem of calculating the sums over the spin variables. To this aim, a more symmetric form of Eq.(4.9) can be obtained by observing that

$$\prod_{i=1}^{N} c_i = \sqrt{c_1}\left[\prod_{i=1}^{N-1}\sqrt{c_ic_{i+1}}\right]\sqrt{c_N}, \tag{4.10}$$

with $c_i > 0$, whatever i. So we have

$$\begin{aligned} Z_G(f) \;=\; & \sum_{S_1\in\{0,1\}} \ldots \sum_{S_N\in\{0,1\}} \exp\left(\frac{\mu}{2k_BT}S_1\right) \exp\left[\frac{-1}{2k_BT}v(S_1)\right]\sqrt{c_1} \prod_{i=1}^{N-1} \left\{ \exp\left(\frac{\lambda}{k_BT}S_iS_{i+1}\right) \right. \\ & \left. \times \exp\left[\frac{\mu}{2k_BT}(S_i+S_{i+1})\right] \exp\left[\frac{-1}{2k_BT}\left[v(S_i)+v(S_{i+1})\right]\right] \sqrt{c_ic_{i+1}} \right\} \\ & \times \exp\left(\frac{\mu}{2k_BT}S_N\right) \exp\left[\frac{-1}{2k_BT}v(S_N)\right]\sqrt{c_N}, \end{aligned} \tag{4.11}$$

where we defined

$$c_i = \frac{\sinh\left[\frac{f\ell(S_i)}{k_BT}\right]}{\frac{f\ell(S_i)}{k_BT}} \frac{\ell^2(S_i)}{\ell^2}. \tag{4.12}$$

We define $\ell(-1) = \ell; v(-1) = 0$ and $\ell(+1) = \chi\ell; v(+1) = \Delta E$, where χ is the ratio between unfolded and folded lengths, and ΔE is the energy jump between the wells. The next step to elaborate the Gibbs partition function consists in adopting the technique of the transfer matrix and then we can write

$$Z_G(f) = \vec{V}^T \mathsf{T}^{N-1} \vec{V}, \tag{4.13}$$

where we have

$$\vec{V} = \begin{bmatrix} e^{-\frac{\mu}{2k_BT}} \sqrt{\frac{\sinh\left(\frac{f\ell}{k_BT}\right)}{\frac{f\ell}{k_BT}}} \\ e^{\frac{\mu}{2k_BT}} e^{-\frac{\Delta E}{2k_BT}} \sqrt{\chi \frac{\sinh\left(\frac{f\ell\chi}{k_BT}\right)}{\frac{f\ell}{k_BT}}} \end{bmatrix} \tag{4.14}$$

or

$$\vec{V} = \begin{bmatrix} \sqrt{2p} \\ \sqrt{2q} \end{bmatrix} \tag{4.15}$$

if we define

$$p = \frac{1}{2} e^{-\frac{\mu}{k_BT}} \frac{\sinh \xi}{\xi}, \tag{4.16}$$

$$q = \frac{1}{2} e^{\frac{\mu}{k_BT}} \chi\phi \frac{\sinh(\chi\xi)}{\xi}, \tag{4.17}$$

where

$$\phi = e^{-\frac{\Delta E}{k_BT}}, \tag{4.18}$$

$$\xi = \frac{f\ell}{k_BT}. \tag{4.19}$$

Moreover,

$$\mathsf{T} = \begin{bmatrix} e^{\frac{\lambda}{k_BT}} e^{-\frac{\mu}{k_BT}} \frac{\sinh \xi}{\xi} & e^{-\frac{\lambda}{k_BT}} e^{-\frac{\Delta E}{2k_BT}} \sqrt{\frac{\sinh \xi}{\xi} \chi \frac{\sinh(\chi\xi)}{\xi}} \\ e^{-\frac{\lambda}{k_BT}} e^{-\frac{\Delta E}{2k_BT}} \sqrt{\frac{\sinh \xi}{\xi} \chi \frac{\sinh(\chi\xi)}{\xi}} & e^{\frac{\lambda}{k_BT}} e^{\frac{\mu}{k_BT}} \chi \frac{\sinh(\chi\xi)}{\xi} e^{-\frac{\Delta E}{k_BT}} \end{bmatrix} \tag{4.20}$$

or again

$$\mathsf{T} = 2 \begin{bmatrix} pe^{\frac{\lambda}{k_BT}} & \sqrt{pq} e^{\frac{-\lambda}{k_BT}} \\ \sqrt{pq} e^{\frac{-\lambda}{k_BT}} & qe^{\frac{\lambda}{k_BT}} \end{bmatrix}. \tag{4.21}$$

Since we are studying the thermodynamics of small systems (small N), we need to calculate the exact value of the partition function given in Eq.(4.13) and not its approximation evaluated for a large value of N, corresponding to the thermodynamic limit. Now, we can simply calculate the eigenvalues of the transfer matrix T, obtaining

$$\begin{aligned} \lambda_{1,2} &= e^{\frac{\lambda}{k_BT}} (p+q) \pm \sqrt{e^{\frac{2\lambda}{k_BT}} (p-q)^2 + 4e^{\frac{-2\lambda}{k_BT}} pq} \\ &= e^{\frac{\lambda}{k_BT}} \left[(p+q) \pm \sqrt{(p+q)^2 - 4pq\left(1 - e^{\frac{-4\lambda}{k_BT}}\right)} \right]. \end{aligned} \tag{4.22}$$

We note that if λ_1 corresponds to the sign "+" and λ_2 to the sign "−", then $\lambda_1 > \lambda_2 > 0$. Now, we need to explicitly determine the matrix power T^{N-1}. We introduce the general theory for an arbitrary matrix function: let $f(x) : \mathbb{R} \to \mathbb{R}$ a function with a Taylor development $f(x) = \sum_{k=0}^{+\infty} C_k x^k$, converging in a region $D \subset \mathbb{C}$. Then, we consider a matrix $\mathbf{A}$ $(M \times M)$, with all eigenvalues in D: $\lambda_j \in D$, whatever $j = 1, ..., M$. So, we define the value $f(\mathbf{A})$ by the series expansion $f(\mathbf{A}) = \sum_{k=0}^{+\infty} C_k \mathbf{A}^k$. We propose an efficient technique to determine $f(\mathbf{A})$ with a finite number of operations. We consider the characteristic polynomial $P_A(x)$ associated to the matrix $\mathbf{A}$

$$P_A(x) = det(\mathbf{A} - x\mathbf{1}) \tag{4.23}$$

with solutions being the eigenvalues of $\mathbf{A}$. We also remember the Cayley-Hamiltonian theorem stating that $P_A(\mathbf{A}) = 0$. In general, we know the procedure for dividing polynomials as

$$\frac{n(x)}{d(x)} = q(x) + \frac{r(x)}{d(x)} \tag{4.24}$$

or $n(x) = q(x)d(x) + r(x)$, where $\partial n = P, \partial d = D, \partial q = M - D$ and $\partial r = D - 1$. We say that $n(x)$ is the dividend, $d(x)$ is the divisor, $q(x)$ is the quotient and $r(x)$ is the remainder. The important point is that the remainder has a lower degree than the divisor. The quantities $q(x)$ and $r(x)$ are unique and the procedure is the so-called Euclidean division. Now, we apply this operation to $\frac{f(x)}{P_A(x)}$, getting

$$\frac{f(x)}{P_A(x)} = q(x) + \frac{r(x)}{P_A(x)} \tag{4.25}$$

or $f(x) = q(x)P_A(x) + r(x)$, with $\partial r = M - 1$ if $\mathbf{A}$ is a matrix $M \times M$. So, by using the Cayley-Hamilton theorem, we obtain that $f(\mathbf{A}) = r(\mathbf{A})$ with $\partial r = M - 1$. We finally have to find the M coefficient of the polynomial $r(x)$. If λ_i is a simple root of $P_A(x)$, we can write $f(\lambda_i) = q(\lambda_i)P_A(\lambda_i) + r(\lambda_i)$ or $f(\lambda_i) = r(\lambda_i)$, which is a condition to obtain the coefficients of $r(x)$. If λ_i has multiplicity 2 in $P_A(x)$, then we can write

$$\begin{aligned} f(x) &= q(x)P_A(x) + r(x), & (4.26)\\ f'(x) &= q'(x)P_A(x) + q(x)P_A'(x) + r'(x), & (4.27) \end{aligned}$$

and we have

$$\begin{aligned} f(\lambda_i) &= r(\lambda_i), & (4.28)\\ f'(\lambda_i) &= r'(\lambda_i). & (4.29) \end{aligned}$$

In conclusion, we have the following property: if $P_A(x) = (-1)^M(x-\lambda_1)^{m_1}...(x-\lambda_r)^{m_r}$, with $\sum_{i=1}^{r} m_i = M$, we can consider $r(x) = \sum_{k=0}^{M-1} r_k x^k$ where the coefficients r_k can be obtained through the system

$$f^{(\alpha)}(\lambda_i) = r^{(\alpha)}(\lambda_i), \tag{4.30}$$

with $0 < \alpha < m_i - 1$, whatever i from 1 to r.
In our case, we search for T^{N-1} with a 2 x 2 matrix $\mathsf{T}(M=2)$ having eigenvalues λ_1 and λ_2, as

$$\mathsf{T}^{N-1} = \alpha\mathbf{1} + \beta\mathsf{T}, \tag{4.31}$$

with

$$\lambda_1^{N-1} = \alpha + \beta\lambda_1, \tag{4.32}$$
$$\lambda_2^{N-1} = \alpha + \beta\lambda_2, \tag{4.33}$$

from which we get

$$\alpha = \frac{\lambda_1\lambda_2^{N-1} - \lambda_2\lambda_1^{N-1}}{\lambda_1 - \lambda_2}, \tag{4.34}$$
$$\beta = \frac{\lambda_1^{N-1} - \lambda_2^{N-1}}{\lambda_1 - \lambda_2}, \tag{4.35}$$

or equivalently

$$\mathsf{T}^{N-1} = \frac{\lambda_1^{N-1} - \lambda_2^{N-1}}{\lambda_1 - \lambda_2}\mathsf{T} + \frac{\lambda_1\lambda_2^{N-1} - \lambda_2\lambda_1^{N-1}}{\lambda_1 - \lambda_2}\mathbf{1}. \tag{4.36}$$

Therefore, the partition function becomes

$$Z_G(\vec{f}) = \vec{V}^T\mathsf{T}^{N-1}\vec{V} = \frac{\lambda_1^{N-1} - \lambda_2^{N-1}}{\lambda_1 - \lambda_2}\vec{V}^T\mathsf{T}\vec{V} + \frac{\lambda_1\lambda_2^{N-1} - \lambda_2\lambda_1^{N-1}}{\lambda_1 - \lambda_2}\vec{V}^T\vec{V}. \tag{4.37}$$

Thus, we have

$$\vec{V}^T\vec{V} = \begin{bmatrix}\sqrt{2p}\sqrt{2q}\end{bmatrix}\begin{bmatrix}\sqrt{2p}\\ \sqrt{2q}\end{bmatrix} = 2p + 2q = 2(p+q), \tag{4.38}$$

$$\begin{aligned}\vec{V}^T\mathsf{T}\vec{V} &= \begin{bmatrix}\sqrt{2p}\sqrt{2q}\end{bmatrix}\begin{bmatrix}2pe^{\frac{\lambda}{k_BT}} & 2e^{-\frac{\lambda}{k_BT}}\sqrt{pq}\\ 2e^{-\frac{\lambda}{k_BT}}\sqrt{pq} & 2qe^{\frac{\lambda}{k_BT}}\end{bmatrix}\begin{bmatrix}\sqrt{2p}\\ \sqrt{2q}\end{bmatrix}\\ &= 4p^2e^{\frac{\lambda}{k_BT}} + 4pqe^{-\frac{\lambda}{k_BT}} + 4pqe^{-\frac{\lambda}{k_BT}} + 4q^2e^{\frac{\lambda}{k_BT}}\\ &= 4e^{\frac{\lambda}{k_BT}}\left[(p+q)^2 - 2pq(1 - e^{-\frac{2\lambda}{k_BT}})\right].\end{aligned} \tag{4.39}$$

With some other modifications, we get

$$\lambda_1 + \lambda_2 = 2e^{\frac{\lambda}{k_B T}}(p+q) = e^{\frac{\lambda}{k_B T}}\vec{V}^T\vec{V} \tag{4.40}$$

$$\lambda_1 - \lambda_2 = 2e^{\frac{\lambda}{k_B T}}\sqrt{(p+q)^2 - 4pq\left(1 - e^{-\frac{4\lambda}{k_B T}}\right)}. \tag{4.41}$$

We try to write $\vec{V}^T\vec{V}$ and $\vec{V}^T\mathsf{T}\vec{V}$ in terms of λ_1, λ_2 and $e^{\frac{\lambda}{k_B T}}$. Therefore,

$$\vec{V}^T\vec{V} = x(\lambda_1 + \lambda_2), \tag{4.42}$$

with $x = \exp^{-\frac{\lambda}{k_B T}}$. From (4.41), we calculate $(\lambda_1 - \lambda_2)^2$

$$(\lambda_1 - \lambda_2)^2 = 4e^{\frac{2\lambda}{k_B T}}\left[\left[(\lambda_1 + \lambda_2)\frac{1}{2}e^{-\frac{\lambda}{k_B T}}\right]^2 - 4pq\left(1 - e^{-\frac{4\lambda}{k_B T}}\right)\right]. \tag{4.43}$$

So we get

$$\frac{x^2}{4}\left(\lambda_1^2 + 2\lambda_1\lambda_2 + \lambda_2^2 - \lambda_1^2 + 2\lambda_1\lambda_2 + \lambda_2^2\right) = 4pq\left(1 - e^{-\frac{4\lambda}{k_B T}}\right). \tag{4.44}$$

And finally

$$pq = \frac{1}{4}\frac{\lambda_1\lambda_2 x^2}{(1-x^2)(1+x^2)}. \tag{4.45}$$

Therefore, from (4.39), we obtain

$$\begin{aligned}\vec{V}^T\mathsf{T}\vec{V} &= \frac{4}{x}\left[\left(\frac{x}{2}\right)^2(\lambda_1+\lambda_2)^2 - \frac{2}{4}\frac{\lambda_1\lambda_2 x^2}{(1-x^4)}(1-x^2)\right] \\ &= \frac{x(1+x^2)(\lambda_1+\lambda_2)^2 - 2x\lambda_1\lambda_2}{1+x^2}.\end{aligned} \tag{4.46}$$

The final calculus of $Z_G(f)$ gives

$$\begin{aligned}Z_G(f) &= \frac{\lambda_1^{N-1} - \lambda_2^{N-1}}{\lambda_1 - \lambda_2}\frac{(x^3+x)(\lambda_1+\lambda_2)^2 - 2\lambda_1\lambda_2}{1+x^2} \\ &\quad + \frac{\lambda_1\lambda_2^{N-1} - \lambda_2\lambda_1^{N-1}}{\lambda_1 - \lambda_2}x(\lambda_1+\lambda_2) \\ &= \frac{x}{(\lambda_1-\lambda_2)(1+x^2)}\left[(\lambda_1^{N-1} - \lambda_2^{N-1})\left((1+x^2)(\lambda_1^2+\lambda_2^2+2\lambda_1\lambda_2) - 2\lambda_1\lambda_2\right)\right. \\ &\quad \left. + (\lambda_1+\lambda_2)(\lambda_1\lambda_2^{N-1} - \lambda_2\lambda_1^{N-1})(1+x^2)\right].\end{aligned} \tag{4.47}$$

Then, we calculate the term into brackets

$$\begin{aligned}&\left[(\lambda_1^{N-1} - \lambda_2^{N-1})\left((1+x^2)(\lambda_1^2+\lambda_2^2+2\lambda_1\lambda_2) - 2\lambda_1\lambda_2\right) + (\lambda_1+\lambda_2)\right. \\ &\left.\times(\lambda_1\lambda_2^{N-1} - \lambda_2\lambda_1^{N-1})(1+x^2)\right] \\ = \;& \lambda_1^{N-1}\lambda_1(\lambda_1-\lambda_2) - \lambda_2^{N-1}\lambda_2(\lambda_2-\lambda_1) \\ &+ x^2(\lambda_1^{N-1} - \lambda_2^{N-1})(\lambda_1+\lambda_2)^2 + x^2(\lambda_1^2\lambda_2^{N-1} - \lambda_2\lambda_1^N + \lambda_1\lambda_2^N - \lambda_2^2\lambda_1^{N-1}) \\ = \;& \lambda_1^N\left[\lambda_1 - \lambda_2 + x^2(\lambda_1+\lambda_2)\right] - \lambda_2^N\left[\lambda_2 - \lambda_1 + x^2(\lambda_1+\lambda_2)\right].\end{aligned} \tag{4.48}$$

Finally, we get the following form of the Gibbs partition function

$$Z_G(f) = \frac{x}{1+x^2}\left[\lambda_1^N\left(1+x^2\frac{\lambda_1+\lambda_2}{\lambda_1-\lambda_2}\right)+\lambda_2^N\left(1-x^2\frac{\lambda_1+\lambda_2}{\lambda_1-\lambda_2}\right)\right], \tag{4.49}$$

where

$$\begin{aligned} x &= e^{-\frac{\lambda}{k_BT}}, && (4.50)\\ \lambda_{1,2} &= \frac{1}{x}\left[p+q\pm\sqrt{(p+q)^2-4pq(1-x^4)}\right], && (4.51)\\ p &= \frac{1}{2}e^{-\frac{\mu}{k_BT}}\frac{\sinh\xi}{\xi}, && (4.52)\\ q &= \frac{1}{2}e^{\frac{\mu}{k_BT}}\chi\phi\frac{\sinh(\chi\xi)}{\xi}, && (4.53)\\ \xi &= \frac{f\ell}{k_BT}, && (4.54)\\ \phi &= e^{-\frac{\Delta E}{k_BT}}. && (4.55) \end{aligned}$$

As usual, we obtain the force-extension response as

$$\langle r\rangle = k_BT\frac{\partial\log Z_G(f)}{\partial f} = k_BT\frac{1}{Z_G}\frac{\partial Z_G}{\partial f}. \tag{4.56}$$

We also note that $\frac{S_i+1}{2}$ gives 0 for folded elements and 1 for unfolded elements. Therefore,the quantity $\left\langle\sum_{i=1}^N\frac{S_i+1}{2}\right\rangle$ is the average number of unfolded elements. The term $\sum_{i=1}^N v(S_i)$ of the Hamiltonian function previously seen can also be written as $\sum_{i=1}^N\frac{S_i+1}{2}\Delta E$, therefore we have

$$\sum_{i=1}^N\frac{S_i+1}{2} = \frac{\partial H_G}{\partial\Delta E}. \tag{4.57}$$

It follows that $\left\langle\sum_{i=1}^N\frac{S_i+1}{2}\right\rangle$ can be evaluated through the expression

$$\left\langle\sum_{i=1}^N\frac{S_i+1}{2}\right\rangle = \frac{\sum_{\vec{S}}\int\frac{\partial H_G}{\partial\Delta E}e^{-\frac{H_G}{k_BT}}\mathrm{d}\vec{r}}{\sum_{\vec{S}}\int e^{-\frac{H_G}{k_BT}}\mathrm{d}\vec{r}}, \tag{4.58}$$

which can be simplified to give

$$\left\langle\sum_{i=1}^N\frac{S_i+1}{2}\right\rangle = -k_BT\frac{\partial}{\partial\Delta E}\log[Z_G(f)], \tag{4.59}$$

corresponding to the average value of unfolded domains. It is useful to introduce here the Gibbs free energy of the system $\mathcal{G} = -k_BT\log Z_G$. The above expected values can be

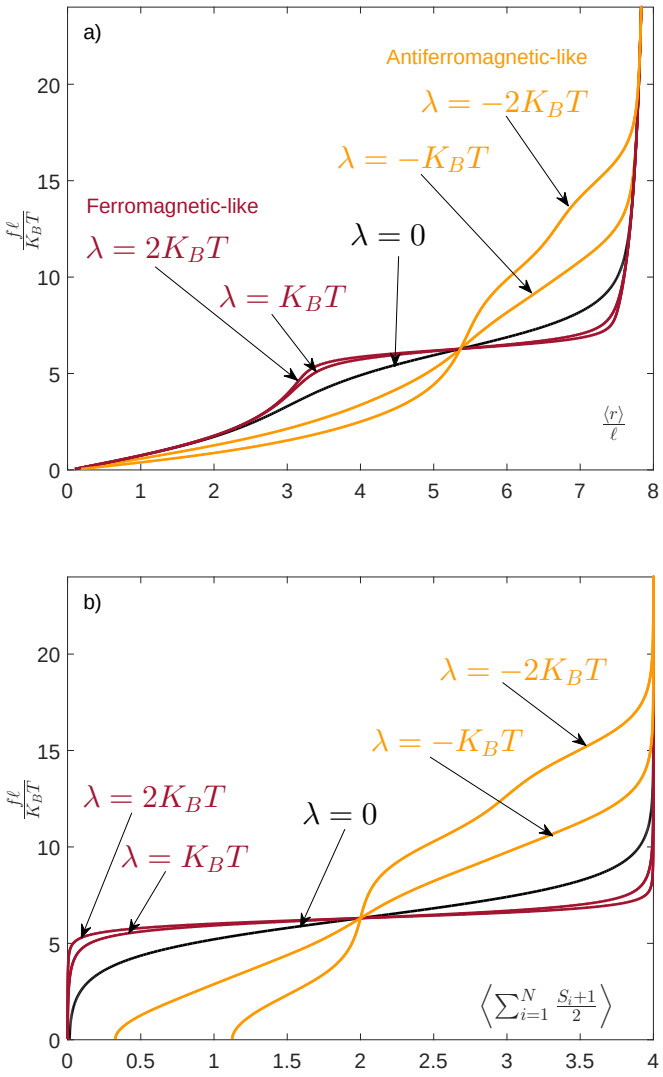

Figure 4.6 – Average normalised extension a) and average number of unfolded units b) versus the applied normalised force for a chain with Ising interactions under isotensional (Gibbs) condition. The curves have been obtained with $\Delta E = 7k_BT$, N=4, $\chi = 2$, $\mu = 0$, and $\lambda = 0, \pm 1, \pm 2k_BT$. The black dashed curves correspond to the chain without interactions ($\lambda = 0$), the dark red (or dark gray) curves to ferromagnetic-like Ising interactions ($\lambda > 0$) and the orange (or light gray) curves to anti-ferromagnetic-like interactions ($\lambda < 0$) [2].

reformulated in terms of this thermodynamic function as follows

$$\langle r \rangle = -\frac{\partial \mathcal{G}}{\partial f}, \tag{4.60}$$

$$\left\langle \sum_{i=1}^{N} \frac{S_i + 1}{2} \right\rangle = \frac{\partial \mathcal{G}}{\partial \Delta E}. \tag{4.61}$$

The knowledge of Z_G or $\mathcal{G}$ allows therefore the determination of both the average extension of the chain and the average number of unfolded units as function of the applied force and temperature.

An application of Eqs.(4.60) and (4.61) can be found in Fig.4.6. First of all, we discuss the curves for $\lambda = 0$, i.e. without Ising interactions: in the force-extension curve (Fig.4.6.a), we note a force plateau corresponding to the synchronised unfolding of the N units. This behaviour is confirmed by the number of unfolded units (Fig.4.6.b), which shows a transition from 0 to N, at the same threshold force as the previously mentioned plateau. This force plateau is the classical result of force spectroscopy experiments conducted with soft devices [24, 68, 76, 77, 133, 141] (see also previous Chapters for a wider discussion). The Ising interactions modify this scenario as follows: if $\lambda > 0$, the units are favoured to be in the same state and the transition is therefore sharper than the one observed for $\lambda = 0$; on the other hand, if $\lambda < 0$, it is easier to unfold the first units and the response is therefore smoother. It is interesting to note that, for strong anti-ferromagnetic-like interactions, even with $f = 0$, we can have a positive number of unfolded units because of the energetic compromise between λ and ΔE. Indeed, when we observe unfolded units with $f = 0$, the cooperativity effect is stronger than the energy jump between folded and unfolded states. We finally remark that the curves in Fig.4.6 are valid for any value of ℓ and T. If, as an example, we consider $\ell = 0.4$nm and $T = 300$K, we get a plateau force at $f = 70$ pN, which is coherent, *e.g.*, with the DNA overstretching transition [130].

4.4 Two-state freely jointed chain with Ising interactions: the Helmholtz ensemble

We consider now the Helmholtz ensemble where the last element of the chain is fixed at a given arbitrary position (isometric condition). The Hamiltonian reads

$$\begin{aligned} H_H &= \lambda \sum_{i=1}^{N-1} S_i S_{i+1} - \mu \sum_{i=1}^{N} S_i \\ &\quad + \sum_{i=1}^{N} \left[v(S_i) + \frac{1}{2} k(S_i) \left[\|\vec{r}_i - \vec{r}_{i-1}\| - \ell(S_i) \right]^2 \right], \end{aligned} \tag{4.62}$$

where $\vec{r}_N = \vec{r}$ is fixed. The phase space is therefore composed of $\vec{r}_i$ $\forall i = 1, ..., N-1$ and S_i $\forall i = 1, ..., N$. Hence, the partition function can be written as

$$\begin{aligned} Z_H(\vec{r}) &= \sum_{S_1 \in \{0,1\}} ... \sum_{S_N \in \{0,1\}} \int_{\mathcal{Q}} \exp\left[-\frac{H_H(\{S_i\}, \{\vec{r}_i\})}{k_B T} \right] \mathrm{d}\vec{r}_1 ... \mathrm{d}\vec{r}_{N-1} \\ &= \sum_{S_1 \in \{0,1\}} ... \sum_{S_N \{\in 0,1\}} \exp\left(\frac{\lambda}{k_B T} \sum_{i=1}^{N-1} S_i S_{i+1} \right) \\ &\quad \times \exp\left(\frac{\mu}{k_B T} \sum_{i=1}^{N} S_i \right) \exp\left[-\frac{1}{k_B T} \sum_{i=1}^{N} v(S_i) \right] \\ &\quad \times \int_{\mathcal{Q}} \exp\left[-\frac{1}{2} \sum_{i=1}^{N} \frac{k(S_i)}{k_B T} \left[\|\vec{r}_i - \vec{r}_{i-1}\| - \ell(S_i) \right]^2 \right] \\ &\quad \times \mathrm{d}\vec{r}_1 ... \mathrm{d}\vec{r}_{N-1}, \end{aligned} \tag{4.63}$$

where $\mathcal{Q} = \mathbb{R}^{3(N-1)}$. It is not difficult to realise that the calculation of $Z_H(\vec{r})$ is much more complicated than the one performed for $Z_G(\vec{f})$. Indeed, in this case, we can not apply a simple change of variables in order to factorise the multi-dimensional integral. From the physical point of view, this difficulty depends on the fact that the isometric condition induces an effective interaction among the units, fixing the sum of all vectors $\vec{r}_i - \vec{r}_{i-1}$ (for i from 1 to N). Then, we have now the combination of two forms of interaction among the units, being the first implicitly encoded in the isometric condition and the second explicitly implemented through the Ising scheme. An useful technique to cope with this difficulty is the following. By comparing Eqs.(4.3) and (4.63), we deduce again that the two partition functions $Z_G(\vec{f})$ and $Z_H(\vec{r})$ are related through a three-dimensional bilateral Laplace transform, as

$$Z_G(\vec{f}) = \int_{\mathbb{R}^3} Z_H(\vec{r}) \exp\left(\frac{\vec{r} \cdot \vec{f}}{k_B T} \right) \mathrm{d}\vec{r}, \tag{4.64}$$

where, as usual, we neglect the non-influential multiplicative constants in the partition functions. Moreover, by considering the spherical symmetry of the problem, we easily obtain the inverse relationship

$$Z_H(r) = \int_{-\infty}^{+\infty} Z_G(i\eta) \frac{\eta}{r} \sin\left(\frac{\eta r}{k_B T} \right) \mathrm{d}\eta, \tag{4.65}$$

where $Z_G(i\eta)$ is the analytic continuation of the partition function $Z_G(f)$ for the Gibbs ensemble, given in Eq.(4.49). The integral in Eq.(4.65) can be simplified by the change of variable $y = \frac{\eta \ell}{k_B T}$, leading to

$$Z_H(r) = \int_{-\infty}^{+\infty} Z_G\left(iy \frac{k_B T}{\ell} \right) \frac{y}{r} \sin\left(\frac{yr}{\ell} \right) \mathrm{d}y, \tag{4.66}$$

where, as before, we neglected the non-influential multiplicative constant. Coherently with our assumptions, the variables p and q assume the form

$$p = \frac{1}{2}e^{-\frac{\mu}{k_BT}}\frac{\sin y}{y} = \frac{\tilde{p}}{y}, \tag{4.67}$$

$$q = \frac{1}{2}e^{\frac{\mu}{k_BT}}\chi\phi\frac{\sin(\chi y)}{y} = \frac{\tilde{q}}{y}. \tag{4.68}$$

Accordingly, the eigenvalues of the transfer matrix become

$$\lambda_{1,2} = \frac{1}{xy}\left[\tilde{p}+\tilde{q} \pm \sqrt{(\tilde{p}+\tilde{q})^2 - 4\tilde{p}\tilde{q}(1-x^4)}\right] = \frac{\tilde{\lambda}_{1,2}}{y}, \tag{4.69}$$

where $\tilde{p} = py$, $\tilde{q} = qy$, and $\tilde{\lambda}_{1,2} = \lambda_{1,2}y$. Hence, the analytic continuation of the Gibbs partition function becomes

$$\begin{aligned} Z_G\left(iy\frac{k_BT}{\ell}\right) = \frac{x}{1+x^2}\frac{1}{y^N}\Bigg[&\tilde{\lambda}_1^m\left(1+x^2\frac{\tilde{\lambda}_1+\tilde{\lambda}_2}{\tilde{\lambda}_1-\tilde{\lambda}_2}\right) \\ &+\tilde{\lambda}_2^N\left(1-x^2\frac{\tilde{\lambda}_1+\tilde{\lambda}_2}{\tilde{\lambda}_1-\tilde{\lambda}_2}\right)\Bigg], \end{aligned} \tag{4.70}$$

where, importantly, $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ depend on y only through $\sin(y)$ and $\sin(\chi y)$. In particular, when χ is an integer (or also a rational number), Z_G is composed of a periodic function of y divided by y^N. So we have

$$Z_G\left(iy\frac{k_BT}{\ell}\right) = \frac{1}{y^N}P(y), \tag{4.71}$$

where $P(y) = P(y+L_y)$ for a given L_y and we have

$$\begin{aligned} P(y) = \frac{x}{1+x^2}\Bigg[&\tilde{\lambda}_1^N\left(1+x^2\frac{\tilde{\lambda}_1+\tilde{\lambda}_2}{\tilde{\lambda}_1-\tilde{\lambda}_2}\right) \\ &+\tilde{\lambda}_2^N\left(1-x^2\frac{\tilde{\lambda}_1+\tilde{\lambda}_2}{\tilde{\lambda}_1-\tilde{\lambda}_2}\right)\Bigg]. \end{aligned} \tag{4.72}$$

If we consider integer values of χ, $P(y)$ is periodic with a period of $L_y = 2\pi$, and it can be developed in Fourier series, as

$$P(y) = \sum_{k=-\infty}^{+\infty} C_k e^{iky}, \tag{4.73}$$

where

$$C_k = \frac{1}{2\pi}\int_0^{2\pi} P(y)e^{-iky}\mathrm{d}y. \tag{4.74}$$

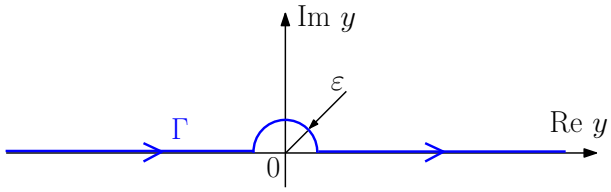

Figure 4.7 – Definition of the contour Γ on the complex plane with an arbitrary radius ε [2].

The values of C_k can be obtained numerically by calculating the integrals through classical numerical techniques (we verified that the simple Simpson's rule is sufficient to obtain accurate results). Once determined the C_k coefficients, the Helmholtz partition function can be obtained analytically as follows. To begin, we have from Eq.(4.65)

$$Z_H(r) = -i \int_\Gamma Z_G \left(iy \frac{k_B T}{\ell} \right) \frac{y}{r} e^{\frac{yr}{\ell}} \mathrm{d}y. \tag{4.75}$$

Here, we used the Euler formula $e^{\frac{iry}{\ell}} = \cos \frac{ry}{\ell} + i \sin \frac{ry}{\ell}$, and we observed that the integral with $\cos \frac{ry}{\ell}$ is zero since $Z_G \left(iy \frac{k_B T}{\ell} \right)$ is an even function of y. Moreover, since the function to integrate is regular on the real axis and holomorphic on a strip $|\mathrm{Im} y| < M$ for an arbitrary $M \in \mathbb{R}$, we can use the path Γ shown in Fig.4.7. This will be useful to elaborate the partition function integral and to write it in a form without singularities at the origin.

Indeed, we have

$$\begin{aligned} Z_H(r) &= -i \int_\Gamma \frac{P(y)}{y^N} \frac{y}{r} e^{\frac{iry}{\ell}} \mathrm{d}y \\ &= -i \int_\Gamma \frac{1}{y^N} \sum_{k=-\infty}^{+\infty} C_k e^{iky} \frac{y}{r} e^{\frac{iry}{\ell}} \mathrm{d}y \\ &= -i \sum_{k=-\infty}^{+\infty} C_k \frac{1}{r} \int_\Gamma \frac{1}{y^{N-1}} e^{i\left(k + \frac{r}{\ell}\right) y} \mathrm{d}y, \end{aligned} \tag{4.76}$$

where the last integral is well defined since the path Γ excludes the singularity at the origin from the integration. We know that an application of the residue theorem delivers [113]

$$\int_\Gamma \frac{e^{iay}}{y^m} \mathrm{d}y = \begin{cases} 0 & \text{if} \quad a > 0, \\ -2\pi i^m \frac{a^{m-1}}{(m-1)!} & \text{if} \quad a \leq 0, \end{cases} \tag{4.77}$$

as proved in Chapter 2, see Eq.(2.78). Therefore,

$$\begin{aligned} Z_H(r) &= i \sum_{k=-\infty}^{+\infty} \frac{C_k}{r} 2\pi i^{N-1} \frac{\left(k + \frac{r}{\ell}\right)^{N-2}}{(N-2)!} \mathbf{1}\left(-k - \frac{r}{\ell}\right) \\ &= \frac{2\pi i^N}{r} \sum_{h=-\infty}^{+\infty} C_{-h} \frac{\left(\frac{r}{\ell} - h\right)^{N-2}}{(N-2)!} \mathbf{1}\left(h - \frac{r}{\ell}\right), \end{aligned} \tag{4.78}$$

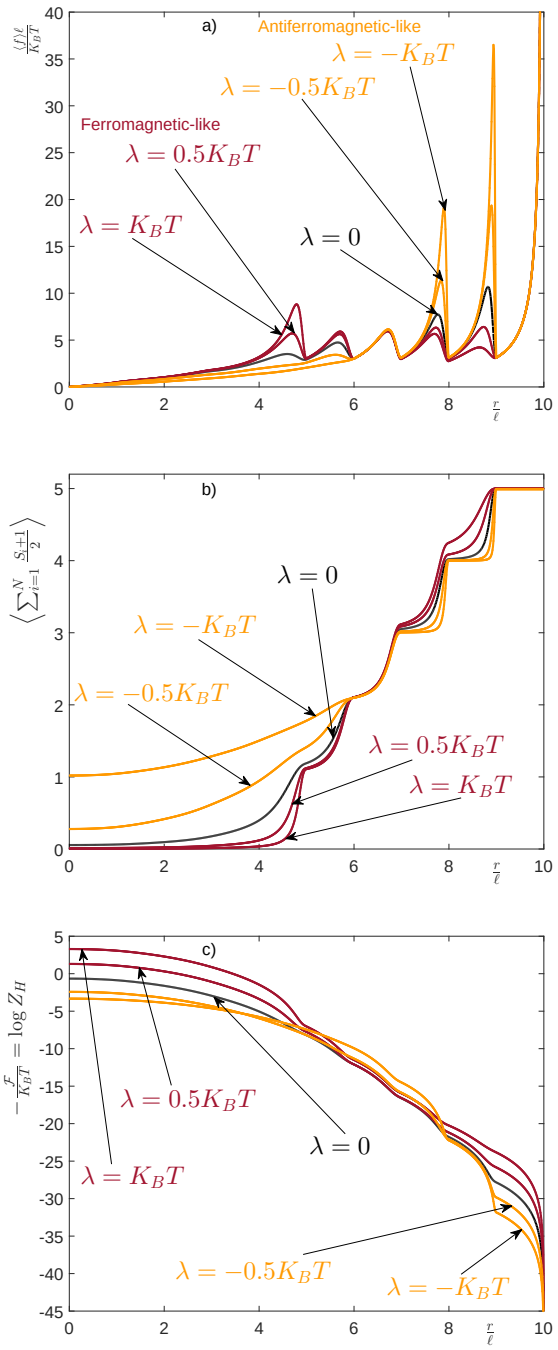

Figure 4.8 – Force-extension response a), average number of unfolded units b) and Helmholtz free energy c) for a chain with Ising interactions under isometric (Helmholtz) condition. The curves have been obtained with $\Delta E = 5k_BT$, N=5, $\chi = 2$, $\mu = 0$, and $\lambda = 0, \pm 0.5, \pm 1 k_BT$. The black curves correspond to the chain without interactions ($\lambda = 0$), the red curves to ferromagnetic-like Ising interactions ($\lambda > 0$) and the blue curves to anti-ferromagnetic-like interactions ($\lambda < 0$) [2].

where $\mathbf{1}(x)$ represents the Heaviside step function, defined as $\mathbf{1}(x) = 1$ if $x \geq 0$, and $\mathbf{1}(x) = 0$ if $x < 0$. Finally,

$$Z_H(r) = \frac{2\pi i^N}{r(N-2)!} \sum_{h=-\infty}^{+\infty} (C_h)^* \left(\frac{r}{\ell} - h\right)^{N-2} \mathbf{1}\left(h - \frac{r}{\ell}\right), \tag{4.79}$$

where we used the property stating that $C_{-h} = (C_h)^*$, which is valid for the Fourier coefficients of a real periodic function. The result obtained in Eq.(4.79) is exact for $\chi \in \mathbb{N}$, but it is based on the numerical computation of the coefficients C_k (semi-analytic procedure). The limitation introduced by considering integer values for χ does not restrict the physical interpretation of the results. Moreover, this procedure can be easily generalised in order to consider arbitrary rational values for χ (of course, the function $P(y)$ remains periodic with $\chi \in \mathbb{Z}$). Furthermore, in next Sections, we also discuss additional asymptotic results, which are not based on restrictions over the values of the parameter χ.

It is important to remark that our semi-analytic procedure, leading to Eq.(4.79) and based on the numerical implementation of Eq.(4.74), is very efficient for the determination of the Helmholtz partition function. Indeed, the direct numerical calculation of the original integral in Eq.(4.66), grounded on the knowledge of the Gibbs partition function given in Eq. (4.70), is a really hard – if not impossible – task since the integrand function is decreasing (as $1/y^{N-1}$) and oscillating for any r in the whole interval between 0 and $N\chi\ell$. Since we need the quantity $\log Z_H(r)$ to analyse the system behaviour, all the oscillations of the integrand function (also for large values of $|y|$) play an important role in defining the result. For this reason, our procedure leads to very accurate results, being based on the analytic determination of the integral over Γ and on the numerical evaluation of the integrals over $(0, 2\pi)$ defined in Eq.(4.74), which are much more stable than the one defined in Eq.(4.66).

By mean of $Z_H(r)$ given in Eq.(4.79), we can find the force-extension response through the expression

$$\langle f \rangle = -k_B T \frac{\partial}{\partial r} \log Z_H(r) = \frac{\partial \mathcal{F}}{\partial r}, \tag{4.80}$$

and the average value of unfolded domains with the relation

$$\left\langle \sum_{i=1}^{N} \frac{S_i + 1}{2} \right\rangle = -k_B T \frac{\partial}{\partial \Delta E} \log Z_H(r) = \frac{\partial \mathcal{F}}{\partial \Delta E}, \tag{4.81}$$

where we introduced the Helmholtz free energy of the system $\mathcal{F} = -k_B T \log Z_H$. An example of application can be found in Fig.4.8, where we show the force-extension response,

the average number of unfolded units and the Helmholtz free energy for a chain stretched under isometric condition and with a variable Ising interaction coefficient. First of all, concerning the case with $\lambda = 0$, we observe that the force-extension curve is composed of a number of peaks corresponding to the non-synchronised (sequential) unfolding of the units. This is confirmed by the step-wise curve representing the average number of unfolded units versus the chain extension. Each step corresponds to the unfolding of a unit induced by the increasing extension of the chain. This behaviour agrees with previous theoretical and experimental results [24,135,141–146] (see also previous Chapter for more details). This scenario is modified by the introduction of the Ising interactions. If $\lambda > 0$, the unfolding of the first units requires a larger force peak since the units are favoured to remain in the initial folded state. On the contrary, the unfolding of the last units requires a smaller force since most of the units are already unfolded and they prefer to be in the same state as the majority. This interpretation equally holds for the plot of the average value of unfolded units, where we can note that the anti-ferromagnetic-like behaviour may induce the unfolding of some units also without the applied force. The origin of the non-synchronised transitions can be highlighted in the Helmholtz free energy curves, characterised by a series of cusps able to induce the force peaks in the force-extension curve. To conclude, the Ising interactions induce a specific cooperativity, which can be detected in the modification of the hierarchy of forces in the sawtooth-like response, as recently observed in force spectroscopy experiments of proteins (*e.g.*, in Filamin A) [154].

A form of criticality can be noticed for the Helmholtz response of the bistable Ising chain. To do this, in the force-extension curves shown in Fig.4.8.a, we can identify the spinoidal regions, characterised by a negative slope or, equivalently, by a negative differential stiffness. It means that, for each force peak observed in Fig.4.8.a, we have a spinoidal interval with $\partial f/\partial r < 0$. It is interesting to study the evolution of these spinoidal regions in terms of the temperature. In general, we can say that the system is or not in a spinoidal phase depending on values of r and T. We can therefore determine a sort of phase diagram, as shown in Fig.4.9, where the end-points of each spinoidal interval (on the extension axis) are shown versus the temperature. While the left end-point corresponds to the maximum of the force peak, the right end-point corresponds to the following minimum. These curves have been represented for different values of the interaction coefficient λ to explore the effects of the Ising scheme on this critical behaviour. Importantly, we can observe that each spinoidal interval disappears for a given temperature, which is a critical temperature for the system. Hence, for a given chain composed of N units, there are N different critical temperatures, one for each unfolding process. We remark that, for a system

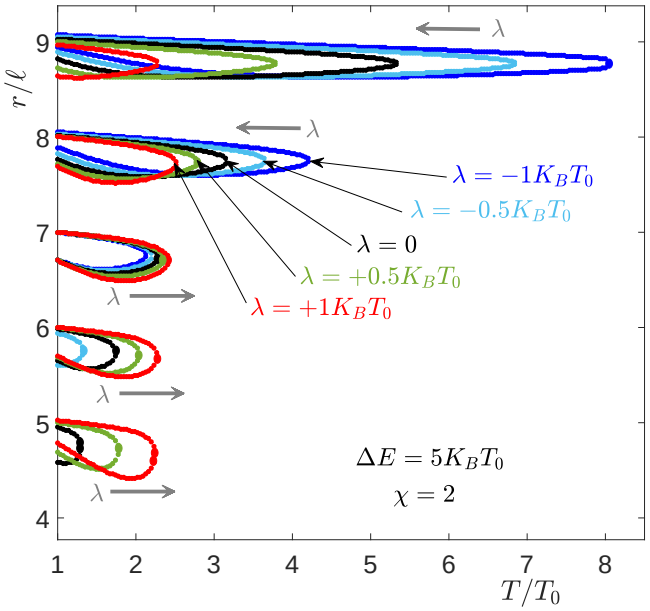

Figure 4.9 – Spinoidal phase diagram showing the end-points of the spinoidal intervals (with $\partial f/\partial r < 0$) versus the temperature of the system ($T > T_0$). Different Ising interaction coefficients have been considered. The curves have been obtained with $\Delta E = 5k_BT_0$, $N = 5$, $\chi = 2$ and $\mu = 0$. The parameters ℓ and T_0 are arbitrary (typical values are $\ell = 1$nm and $T_0 = 300$K) [2].

without Ising interactions, the critical temperature is larger for the last unfolded units. This contrast among critical temperatures is further amplified for anti-ferromagnetic-like systems. On the other hand, a given intensity of ferromagnetic-like interactions is able to equilibrate the critical temperatures among the unfolding processes (see, *e.g.*, the red curves in Fig.4.9). This point can be explained by observing that $\lambda > 0$ favours the simultaneous unfolding of the units, thus inducing a similar behaviour of these ones. We can say that the ferromagnetic-like interactions induce a resistance to fluctuations within the system. It means that all unfolding processes do not lose their snap-through response for a given range of temperature. The bistability resistant to fluctuations is an important concept for micro-and nano-mechanical systems with non-convex elastic energy, where one attempts to sustain the bistability at possibly large temperatures [177]. We remark that the observation of a negative differential stiffness for subcritical temperatures and of a positive differential stiffness for supercritical temperatures can be interpreted by stating that the system behaves as a metamaterial [111, 112]. In a following Section, we will also explore the effect of the intrinsic stiffness of the units on this critical behaviour.

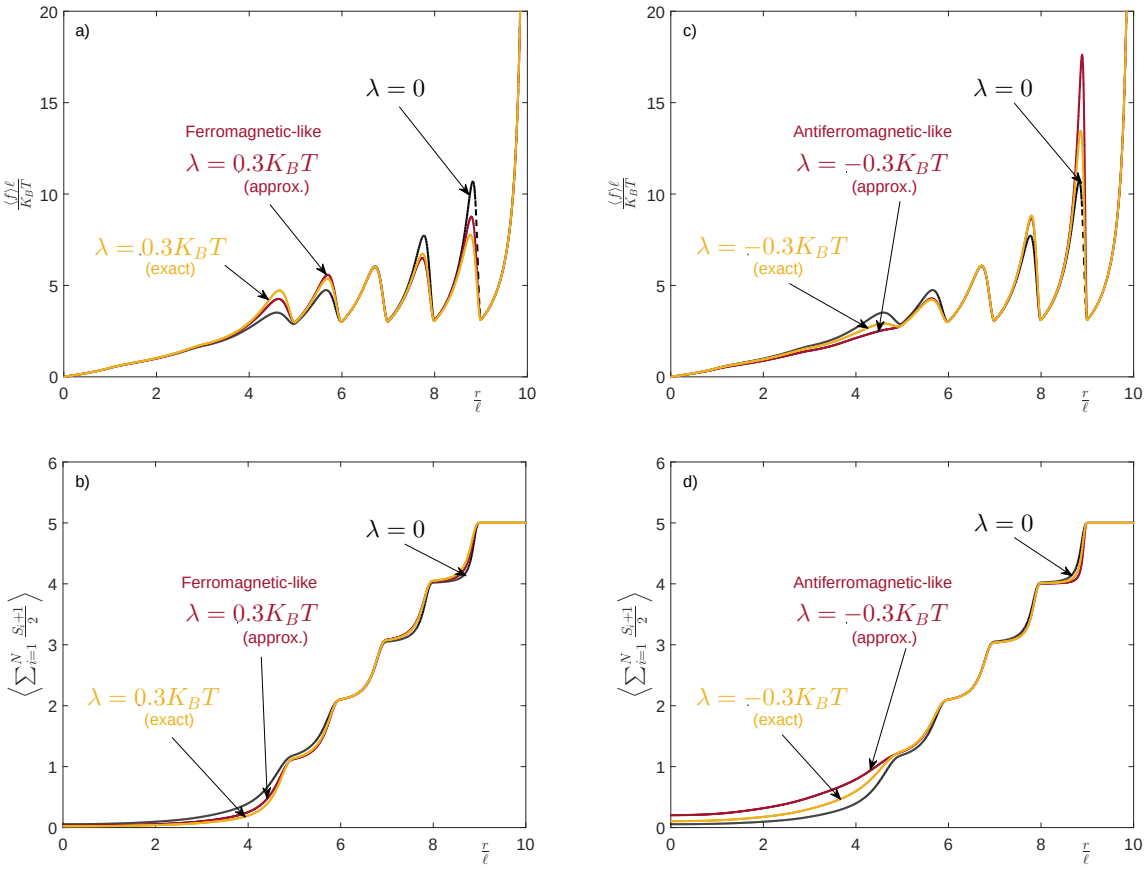

Figure 4.10 – Force-extension response a)-c) and average number of unfolded units b)-d) for weak Ising interactions with ferromagnetic-like behaviour a)-b) and anti-ferromagnetic-like behaviour c)-d). In each panel, the response without interactions ($\lambda = 0$, black dashed curves) is shown together with the exact result (orange or light gray curves) obtained through Eq.(4.79) and the first order approximation (dark red or dark gray curves) given in Eq.(4.98). The curves have been obtained with $\Delta E = 5k_BT$, $N = 5$, $\chi = 2$, $\mu = 0$ and $\lambda = \pm 0.3k_BT$ [2].

4.5 Explicit expression for the Helmholtz response under weak Ising interaction: $|\lambda| \ll k_BT$

We investigate in more detail the particular case with weak Ising interaction, i.e. $|\lambda| \ll k_BT$, by supposing both ferromagnetic and anti ferromagnetic interactions. We know that $Z_H(r)$ can be written like

$$Z_H(r) = \int_{-\infty}^{+\infty} Z_G\left(iy\frac{k_BT}{\ell}\right)\frac{y}{r}\sin\left(\frac{yr}{\ell}\right)\mathrm{d}y. \tag{4.82}$$

We can develop $Z_G\left(iy\frac{k_BT}{\ell}\right)$ in Taylor series with respect to the parameter λ controlling the Ising interaction. We easily find

$$\begin{aligned} Z_G\left(iy\frac{k_BT}{\ell}\right) &= \left[1+\frac{\lambda}{k_BT}(N-1)\right](a+b)^N - \frac{4\lambda}{k_BT}(N-1)\,ab \\ &\quad \times(a+b)^{N-2} + 0(\lambda^2). \end{aligned} \tag{4.83}$$

This expression, which is valid for $|\lambda| \ll k_BT$, can be easily integrated to obtain the Helmholtz partition function. Here, we defined

$$a = \frac{\sin y}{y}, \tag{4.84}$$

$$b = \chi\phi\frac{\sin(\chi y)}{y}. \tag{4.85}$$

We note that, for $\lambda = 0$, we obtain the partition function of the system without Ising interaction. In this Section, we consider $\mu = 0$ to simplify the calculation. We have to calculate

$$Z_H(r) = -i\int_\Gamma Z_G\left(iy\frac{k_BT}{\ell}\right)\frac{y}{r}e^{i\frac{ry}{\ell}}\mathrm{d}y, \tag{4.86}$$

where $Z_G(f)$ is the approximation given previously. We divide the calculation into two steps, the first one is called (1) and the second one is called (2). We firstly calculate

$$\begin{aligned} (1) &= -i\int_\Gamma (a+b)^N\left(\frac{y}{r}\right)e^{i\frac{ry}{\ell}}\mathrm{d}y \\ &= -i\int_\Gamma \sum_{k=0}^{N}\binom{N}{k}\left(\frac{\sin y}{y}\right)^{N-k}\left[\chi\phi\frac{\sin(\chi y)}{y}\right]^k\left(\frac{y}{r}\right)e^{i\frac{ry}{\ell}}\mathrm{d}y. \end{aligned} \tag{4.87}$$

Now, we have

$$\sin^{N-k} y = \frac{1}{(2i)^{N-k}}\sum_{p=0}^{N-k}\binom{N-k}{p}(-1)^p\, e^{iy(N-k-2p)}, \tag{4.88}$$

$$\sin^k(\chi y) = \frac{1}{(2i)^k}\sum_{q=0}^{k}\binom{k}{q}(-1)^q\, e^{i\chi y(k-2q)}. \tag{4.89}$$

Then, we find

$$\begin{aligned}(1) &= -i\int_\Gamma \frac{1}{y^N}\sum_{k=0}^{N}\sum_{p=0}^{N-k}\sum_{q=0}^{k}\binom{N}{k}\binom{N-k}{p}\binom{k}{q}\frac{1}{(2i)^N}(-1)^{p+q}\,(\chi\phi)^k \\ &\quad\times e^{iy(N-k-2p+\chi k-2q\chi)}\frac{y}{r}e^{i\frac{ry}{\ell}}\mathrm{d}y \\ &= -\frac{1}{2^N i^{N-1}}\frac{1}{r}\sum_{k=0}^{N}\sum_{p=0}^{N-k}\sum_{q=0}^{k}\binom{N}{k}\binom{N-k}{p}\binom{k}{q}(-1)^{p+q}\,(\chi\phi)^k \\ &\quad\times\int_\Gamma \frac{e^{iy(N-k-2p+\chi k-2q\chi+\frac{r}{\ell})}}{y^{N-1}}\mathrm{d}y. \end{aligned} \tag{4.90}$$

Now, we use the integral

$$\int_\Gamma \frac{e^{iay}}{y^N}\mathrm{d}y = \begin{cases} 0 & \text{if } a>0, \\ -2\pi i^N\frac{a^{N-1}}{(N-1)!} & \text{if } a\le 0. \end{cases} \tag{4.91}$$

Hence, we obtain

$$\begin{aligned}(1) &= \frac{\pi}{2^{N-1}}\frac{1}{r(N-2)!}\sum_{k=0}^{N}\sum_{p=0}^{N-k}\sum_{q=0}^{k}\binom{N}{k}\binom{N-k}{p}\binom{k}{q} \\ &\quad\times(-1)^{p+q}\,(\chi\phi)^k\,(-\lambda)^{N-2}\,\mathbf{1}(\lambda), \end{aligned} \tag{4.92}$$

where

$$-\lambda = N-k-2p+\chi k-2\chi q+\frac{r}{\ell}. \tag{4.93}$$

Now, we calculate the second integral given by

$$\begin{aligned}(2) &= -i\int_\Gamma ab\,(a+b)^{N-2}\;\frac{y}{r}e^{i\frac{ry}{\ell}}\mathrm{d}y \\ &= -i\int_\Gamma ab\sum_{k=0}^{N-2}\binom{N-2}{k}b^k a^{N-2k}\;\frac{y}{r}e^{i\frac{ry}{\ell}}\mathrm{d}y \\ &= -i\int_\Gamma \sum_{k=0}^{N-2}\binom{N-2}{k}b^{k+1}a^{N-1-k}\;\frac{y}{r}e^{i\frac{ry}{\ell}}\mathrm{d}y. \end{aligned} \tag{4.94}$$

Here again, we develop the powers

$$
\begin{aligned}
(2) &= -i\int_\Gamma \sum_{k=0}^{N-2}\sum_{q=0}^{k+1}\sum_{p=0}^{N-1-k}\binom{N-2}{k}\binom{k+1}{q}\binom{N-1-k}{p}\,\frac{y}{r}e^{i\frac{ry}{\ell}} \\
&\quad\times(-1)^{p+q}\frac{1}{(2i)^N}\frac{1}{y^N}\,(\chi\phi)^{k+1}\,e^{i\chi y(k+1-q)}e^{-i\chi yq}e^{iy(N-1-k-p)}e^{-iyp}\mathrm{d}y \\
&= -i\int_\Gamma \sum_{k=0}^{N-2}\sum_{q=0}^{k+1}\sum_{p=0}^{N-1-k}\binom{N-2}{k}\binom{k+1}{q}\binom{N-1-k}{p}(-1)^{p+q} \\
&\quad\times\frac{(\chi\phi)^{k+1}}{(2i)^N}\,\frac{1}{r\,y^{N-1}}\,e^{\left(iy\left[N-1-k-p-p-\chi q-\chi q+\chi k+\chi+\frac{r}{\ell}\right]\right)}\mathrm{d}y. \qquad (4.95)
\end{aligned}
$$

By using Eq.(4.91), we get

$$
\begin{aligned}
(2) &= \sum_{k=0}^{N-2}\sum_{q=0}^{k+1}\sum_{p=0}^{N-1-k}\binom{N-2}{k}\binom{k+1}{q}\binom{N-1-k}{p}(-1)^{p+q} \\
&\quad\times(\chi\phi)^{k+1}\,\frac{\pi}{2^{N-1}}\frac{1}{r(N-2)!}\,(-\lambda_0)^{N-2}\,\mathbf{1}\,(\lambda_0)\,, \qquad (4.96)
\end{aligned}
$$

where

$$
-\lambda_0 = N-1-k-2p-2\chi q+\chi k+\chi+\frac{r}{\ell}. \qquad (4.97)
$$

Finally, the partition function reads

$$
\begin{aligned}
Z_H(r) &= \frac{\pi}{2^{N-1}(N-2)!r}\sum_{k=0}^{N}\sum_{p=0}^{N-k}\sum_{q=0}^{k}\binom{N}{k}\binom{N-k}{p}\binom{k}{q}\left[1+\frac{\lambda}{k_BT}(N-1)\right](-1)^{p+q} \\
&\quad\times(\chi\phi)^k\,(-\lambda)^{N-2}\,\mathbf{1}(\lambda)-\frac{\pi}{2^{N-1}(N-2)!r}\sum_{k=0}^{N-2}\sum_{q=0}^{k+1}\sum_{p=0}^{N-1-k}\binom{N-2}{k}\binom{k+1}{q} \\
&\quad\times\binom{N-1-k}{p}(-1)^{p+q}\,(\chi\phi)^{k+1}\,\frac{4\lambda}{k_BT}(N-1)\,(-\lambda_0)^{N-2}\,\mathbf{1}(\lambda_0). \qquad (4.98)
\end{aligned}
$$

This is the final form of the Helmholtz partition function, calculated under the hypothesis of weak Ising interaction. In this expression, there are no limitations concerning the parameter χ. We note that, for $\lambda = 0$, we obtain the partition function of the system without Ising interactions, as discussed in recent literature [113]. In Fig.4.10, one can find some results for $\lambda = 0$, $\lambda = +0.3k_BT$ and $\lambda = -0.3k_BT$. We plotted the force-extension curves and the average number of unfolded units for both ferromagnetic-like and anti-ferromagnetic-like interactions. In each case, we compared the approximated result stated in Eq.(4.98) (dark red or dark gray curves), the exact result given in Eq.(4.79) (orange

or light gray curves), and the response without Ising interactions (black dashed curves). We note a good agreement between approximated and exact results both for $\lambda < 0$ and $\lambda > 0$. Concerning the interpretation of the curves, the discussion reported at the end of Section 4.4 remains valid for all results of Fig.4.10.

4.6 Explicit expression for the Helmholtz response under strong Ising ferromagnetic interaction: $\lambda \gg k_B T$

We introduce here an asymptotic development concerning the case of a strong Ising ferromagnetic-like interaction. If $\lambda \to +\infty$, then $e^{-\frac{\lambda}{k_B T}} \to 0$, and therefore it is not difficult to obtain the asymptotic expression for $Z_G(f)$ when $\lambda \to \infty$. The result is

$$Z_G\left(iy\frac{k_BT}{\ell}\right) = \exp\left[\frac{\lambda(N-1)}{k_BT}\right]\left[\left(\frac{\sin y}{y}\right)^N + \left(\chi\phi\frac{\sin(\chi y)}{y}\right)^N\right]. \tag{4.99}$$

Therefore, we easily determine the Helmholtz partition function, as

$$\begin{aligned}
Z_H(r) &= -i\int_\Gamma Z_G\left(iy\frac{k_BT}{\ell}\right)\frac{y}{r}e^{i\frac{ry}{\ell}}\mathrm{d}y \\
&= -i\int_\Gamma \exp\left[\frac{\lambda(N-1)}{k_BT}\right]\left[\sin^N y + (\chi\phi)^N\sin^N(\chi y)\right]\frac{1}{y^{N-1}r}e^{i\frac{ry}{\ell}}\mathrm{d}y \\
&= -i\frac{1}{(2i)^N}\exp\left[\frac{\lambda(N-1)}{k_BT}\right]\sum_{k=0}^{N}\binom{N}{k}(-1)^k\left(\frac{1}{r}\right)\left[\frac{\left(N-2k+\frac{r}{\ell}\right)^{N-2}}{(N-2)!}\right. \\
&\quad \times\mathbf{1}\left(2k-N-\frac{r}{\ell}\right)(-2\pi)i^{N-1} + (\chi\phi)^N(-2\pi)i^{N-1} \\
&\quad \left.\times\frac{\left(\chi N-2\chi k+\frac{r}{\ell}\right)^{N-2}}{(N-2)!}\mathbf{1}\left(2\chi k-\chi N-\frac{r}{\ell}\right)\right].
\end{aligned} \tag{4.100}$$

Finally,

$$\begin{aligned}
Z_H(r) &= \frac{\pi}{2^{N-1}r(N-2)!}\exp\left[\frac{\lambda(N-1)}{k_BT}\right]\sum_{k=0}^{N}\binom{N}{k}(-1)^k \\
&\quad \times\left[\left(N-2k+\frac{r}{\ell}\right)^{N-2}\mathbf{1}\left(2k-N-\frac{r}{\ell}\right)\right. \\
&\quad \left.+(\chi\phi)^N\left(\chi N-2\chi k+\frac{r}{\ell}\right)^{N-2}\mathbf{1}\left(2\chi k-\chi N-\frac{r}{\ell}\right)\right].
\end{aligned} \tag{4.101}$$

which is valid for strong ferromagnetic-like Ising interactions.
An application of this expression is shown in Fig.4.11. In particular, we compare the

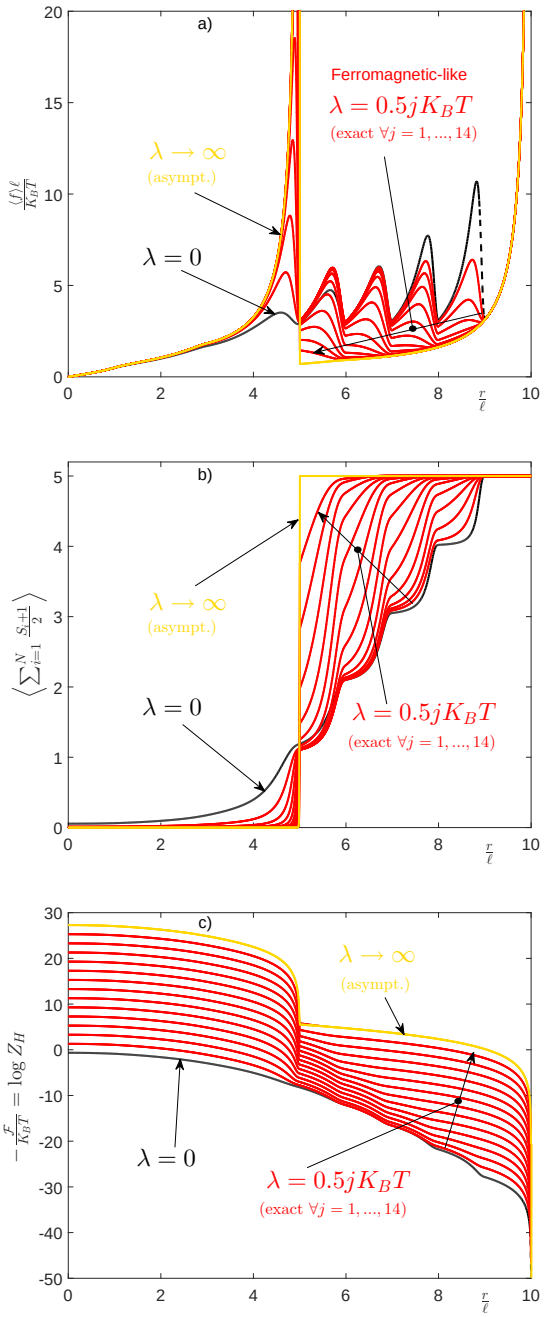

Figure 4.11 – Force-extension response a), average number of unfolded units b) and Helmholtz free energy c) for a chain with strong ferromagnetic-like interactions. In each panel, the response without interactions ($\lambda = 0$, black dashed curves) is shown together with the exact results for an increasing ferromagnetic interaction ($\lambda = 0.5jk_BT\forall j = 1, ..., 14$, red or gray curves) and the asymptotic expansion ($\lambda \to \infty$, yellow or light gray curves) given in Eq.(4.101). The curves have been obtained with $\Delta E = 5k_BT$, $N = 5$, $\chi = 2$ and $\mu = 0$ [2].

approximated result in Eq.(4.101) (yellow or light gray curve) with the exact response obtained from Eq.(4.79) for $\lambda = 0.5, 1, 1.5, ..., 7k_BT$ (red or gray curves) and with the response without Ising interactions (black dashed curve). In Fig.4.11, one can find the force-extention curves, the average number of unfolded units and the Helmholtz free energy. It is interesting to discuss the evolution of the overall behaviour of the system with an increasing interaction coefficient. Indeed, as λ is increased, the units are progressively favoured to be in the same state, and therefore there is an increasing average number of units which unfolds at $r = N\ell$. It means that the number of unfolding processes at $r = N\ell$ is a growing function of the Ising coefficient λ, going from 1 with $\lambda = 0$ to N with λ approaching infinity. This can be seen in Fig.4.11.b, where this process is represented by the series of red curves (or gray) with increasing λ, and it ends with the yellow (or light gray) curve obtained through Eq.(4.101). The latter means that all units unfold at the same time at $r = N\ell$ when $\lambda \to \infty$, and this behaviour is perfectly caught by the asymptotic development. Accordingly, the peaks in the force-extension curve are strongly modified by increasing λ: while the first peak becomes more and more pronounced, the others are progressively reduced, as shown in Fig.4.11.a. As a matter of fact, the first peak corresponds to the simultaneous unfolding of the units when λ is very large. Hence, in the limiting case of $\lambda \to \infty$, the force-extension curve is composed of only one peak (yellow or light gray curve), as one can see in Fig.4.11.a. Of course, the origin of peaks modification in the force-extension curve and of the steps structure in the average number of unfolded units can be observed in the plot of the Helmholtz free energy, shown in in Fig.4.11.c. Here, we can see the evolution of the typical cusps with the increasing Ising coefficient. As an example, the collapse of all the force peaks into a single unfolding event explains the tandem repeats behaviour in red cell spectrin, where two units unfold simultaneously because of a strong cooperativity [114].

4.7 Explicit expression for the Helmholtz response under strong Ising anti-ferromagnetic interaction: $\lambda \ll -k_BT$

We discuss here the development of the theory under strong Ising anti-ferromagnetic-like interactions. As before, we can develop the Gibbs partition function in Taylor series for $\lambda \to -\infty$, i.e. $x = e^{-\frac{\lambda}{k_BT}} \to +\infty$.

The result is

$$\begin{aligned} Z_G\left(iy\frac{k_BT}{\ell}\right) &= x^{N-1}P^{\frac{N}{2}}\left[\left(1+\frac{1}{2}\frac{S}{P^{\frac{1}{2}}}\right)+(-1)^N\left(1-\frac{1}{2}\frac{S}{P^{\frac{1}{2}}}\right)\right] \\ &= x^{N-1}\left[P^{\frac{N}{2}}\left[1+(-1)^N\right]+\frac{1}{2}P^{\frac{N-1}{2}}S\left[1-(-1)^N\right]\right]. \end{aligned} \tag{4.102}$$

And so, we can summarise it as

$$Z_G\left(iy\frac{k_BT}{\ell}\right)=\begin{cases} x^{N-1}P^{\frac{N}{2}} & \text{if } N \text{ is even,} \\ x^{N-1}SP^{\frac{N-1}{2}} & \text{if } N \text{ is odd.} \end{cases} \tag{4.103}$$

where

$$S = \frac{\sin y}{y}+\chi\phi\frac{\sin(\chi y)}{y}=a+b, \tag{4.104}$$

$$P = \frac{\sin y}{y}\chi\phi\frac{\sin(\chi y)}{y}=ab. \tag{4.105}$$

We first elaborate the Helmholtz partition function for N odd as

$$\begin{aligned} Z_H(r) &= -i\int_\Gamma Z_G\left(iy\frac{k_BT}{\ell}\right)\frac{y}{r}e^{i\frac{ry}{\ell}}\mathrm{d}y \\ &= -i\int_\Gamma \frac{x^{N-1}}{y^{N-1}}\left[\sin y+\chi\phi\sin(\chi y)\right]\left[\chi\phi\sin y\sin(\chi y)\right]^{\frac{N-1}{2}}\frac{1}{r}e^{i\frac{ry}{\ell}}\mathrm{d}y \\ &= -i\frac{x^{N-1}}{r}\left(\chi\phi\right)^{\frac{N-1}{2}}\int_\Gamma\frac{1}{y^{N-1}}\left[\sin^{\frac{N+1}{2}}(y)\sin^{\frac{N-1}{2}}(\chi y)\right. \\ &\quad \left.+\chi\phi\sin^{\frac{N-1}{2}}(y)\sin^{\frac{N+1}{2}}(\chi y)\right]e^{i\frac{ry}{\ell}}\mathrm{d}y. \end{aligned} \tag{4.106}$$

Now, we use the relation

$$\sin^k y=\frac{1}{(2i)^k}\sum_{p=0}^{k}\binom{k}{p}(-1)^p e^{iy(k-2p)}, \tag{4.107}$$

and we have

$$\begin{aligned} Z_H(r) &= -i\frac{x^{N-1}}{r}\left(\chi\phi\right)^{\frac{N-1}{2}}\int_\Gamma\frac{1}{y^{N-1}}\left[\frac{1}{(2i)^{\frac{N+1}{2}}}\sum_{p=0}^{\frac{N+1}{2}}\binom{\frac{N+1}{2}}{p}(-1)^p e^{iy\left(\frac{N+1}{2}-2p\right)}\right. \\ &\quad\times\frac{1}{(2i)^{\frac{N-1}{2}}}\sum_{q=0}^{\frac{N-1}{2}}\binom{\frac{N-1}{2}}{q}(-1)^q e^{iy\chi\left(\frac{N-1}{2}-2q\right)}+\chi\phi\frac{1}{(2i)^{\frac{N-1}{2}}}\sum_{q=0}^{\frac{N-1}{2}}\binom{\frac{N-1}{2}}{q}(-1)^q \\ &\quad\left.\times e^{iy\left(\frac{N-1}{2}-2q\right)}\frac{1}{(2i)^{\frac{N+1}{2}}}\sum_{p=0}^{\frac{N+1}{2}}\binom{\frac{N+1}{2}}{p}(-1)^p e^{iy\chi\left(\frac{N+1}{2}-2p\right)}\right]e^{i\frac{ry}{\ell}}\mathrm{d}y. \end{aligned} \tag{4.108}$$

This expression can be reorganised like

$$\begin{aligned} Z_H(r) &= -i\frac{x^{N-1}}{r}\left(\chi\phi\right)^{\frac{N-1}{2}}\frac{1}{(2i)^N}\sum_{p=0}^{\frac{N+1}{2}}\sum_{q=0}^{\frac{N-1}{2}}\binom{\frac{N+1}{2}}{p}\binom{\frac{N-1}{2}}{q}(-1)^{p+q} \\ &\times\left[\int_\Gamma\frac{1}{y^{N-1}}e^{iy\left(\frac{N+1}{2}-2p+\chi\frac{N-1}{2}-2\chi q+\frac{r}{\ell}\right)}\mathrm{d}y\right. \\ &\left.+\chi\phi\int_\Gamma\frac{1}{y^{N-1}}e^{iy\left(\frac{N-1}{2}-2q+\chi\frac{N+1}{2}-2\chi p+\frac{r}{\ell}\right)}\mathrm{d}y\right]. \end{aligned} \tag{4.109}$$

Now, since we know that

$$\int_\Gamma\frac{e^{iay}}{y^m}\mathrm{d}y = -2\pi i^m\frac{a^{m-1}}{(m-1)!}\mathbf{1}(-a), \tag{4.110}$$

we get

$$\begin{aligned} Z_H(r) &= -i\frac{x^{N-1}}{r}\left(\chi\phi\right)^{\frac{N-1}{2}}\frac{1}{(2i)^N}\sum_{p=0}^{\frac{N+1}{2}}\sum_{q=0}^{\frac{N-1}{2}}\binom{\frac{N+1}{2}}{p}\binom{\frac{N-1}{2}}{q}(-1)^{p+q} \\ &\times\left[-2\pi i^{N-1}\frac{(-\lambda_a)^{N-2}}{(N-2)!}\mathbf{1}\left(\lambda_a\right)+\chi\phi(-2\pi)i^{N-1}\frac{(-\lambda_b)^{N-2}}{(N-2)!}\mathbf{1}\left(\lambda_b\right)\right], \end{aligned} \tag{4.111}$$

where

$$\lambda_a = 2p-\frac{N+1}{2}-\chi\frac{N-1}{2}+2\chi q-\frac{r}{\ell}, \tag{4.112}$$

$$\lambda_b = 2q-\frac{N-1}{2}-\chi\frac{N+1}{2}+2\chi p-\frac{r}{\ell}. \tag{4.113}$$

We simplify the expression by reorganising it and we obtain the following expression for the Helmholtz partition function under strong Ising anti-ferromagnetic interaction, for N odd

$$\begin{aligned} Z_H(r) &= \frac{2\pi}{2^N}\frac{\exp\left[-\frac{\lambda}{k_BT}(N-1)\right]}{(N-2)!}\frac{1}{r}\left(\chi\phi\right)^{\frac{N-1}{2}}\sum_{p=0}^{\frac{N+1}{2}}\sum_{q=0}^{\frac{N-1}{2}}\binom{\frac{N+1}{2}}{p}\binom{\frac{N-1}{2}}{q} \\ &\times(-1)^{p+q}\left[\left(-\lambda_a\right)^{N-2}\mathbf{1}\left(\lambda_a\right)+\chi\phi\left(-\lambda_b\right)^{N-2}\mathbf{1}\left(\lambda_b\right)\right]. \end{aligned} \tag{4.114}$$

We calculate now the same quantity for N even

$$\begin{aligned}
Z_H(r) &= -i\int_\Gamma Z_G\left(iy\frac{k_BT}{\ell}\right)\frac{y}{r}e^{i\frac{ry}{\ell}}\mathrm{d}y \\
&= -i\int_\Gamma x^{N-1}\left(\frac{\sin y}{y}\chi\phi\frac{\sin(\chi y)}{y}\right)^{\frac{N}{2}}\frac{y}{r}e^{i\frac{ry}{\ell}}\mathrm{d}y \\
&= -ix^{N-1}\left(\chi\phi\right)^{\frac{N}{2}}\frac{1}{r}\int_\Gamma\frac{1}{y^{N-1}}\sin^{\frac{N}{2}}(y)\sin^{\frac{N}{2}}(\chi y)\,e^{i\frac{ry}{\ell}}\mathrm{d}y \\
&= -ix^{N-1}\left(\chi\phi\right)^{\frac{N}{2}}\frac{1}{r}\frac{1}{(2i)^N}\sum_{p=0}^{\frac{N}{2}}\sum_{q=0}^{\frac{N}{2}}\binom{\frac{N}{2}}{p}\binom{\frac{N}{2}}{q}(-1)^{p+q} \\
&\quad\times(-2\pi)i^{N-1}\left(-\lambda_c\right)^{N-2}\frac{1}{(N-2)!}\mathbf{1}\left(-\lambda_c\right),
\end{aligned} \tag{4.115}$$

where $\lambda_c = 2p - \frac{N}{2} + 2\chi q - \chi\frac{N}{2} - \frac{r}{\ell}$. Finally, we get the following expression for the Helmholtz partition function under strong Ising anti-ferromagnetic interaction

$$\begin{aligned}
Z_H(r) &= \frac{2\pi}{2^N}\frac{\exp\left[-\frac{\lambda}{k_BT}(N-1)\right]}{(N-2)!}\frac{1}{r}\left(\chi\phi\right)^{\frac{N}{2}}\sum_{p=0}^{\frac{N}{2}}\sum_{q=0}^{\frac{N}{2}}\binom{\frac{N}{2}}{p}\binom{\frac{N}{2}}{q} \\
&\quad\times(-1)^{p+q}\left(-\lambda_c\right)^{N-2}\mathbf{1}\left(\lambda_c\right),
\end{aligned} \tag{4.116}$$

which is valid for N even. The solutions given in Eqs.(4.114) and (4.116) represent the most important result of this Section and must be discussed as follows.

An example of application of Eq.(4.114) for N odd can be found in Fig.4.12, where force-extension curves, average number of unfolded units and Helmholtz free energy are represented for several values of λ. When we consider an increasing value of $|\lambda|$ ($\lambda < 0$), we observe that the first force peaks tend to disappear while the last ones become more and more pronounced (see Fig.4.12.a). This is coherent with the assumption that, in an anti-ferromagnetic-like system, the favoured states are alternatively folded and unfolded. Accordingly, with an increasing value of $|\lambda|$ ($\lambda < 0$), we have an increasing number of unfolded units in the initial configuration with $r = 0$. Clearly, the maximum value of this number of unfolded units with $r = 0$ is $(N-1)/2$ and not $(N+1)/2$ since the unfolded units are costly from the energetic point of view and the system chooses the configuration with the smallest number of unfolded units between $(N-1)/2$ and $(N+1)/2$. This overall interpretation of Fig.4.12 must be improved to better understand the yellow (or light gray) curves corresponding to $\lambda \to \infty$. So, if we look at the equation for N odd, we note that $Z_H(r)$ is different from 0 where $\lambda_a > 0$ or $\lambda_b > 0$. It means that

$$2p - \frac{N+1}{2} - \chi\frac{N-1}{2} + 2\chi q - \frac{r}{\ell} > 0 \tag{4.117}$$

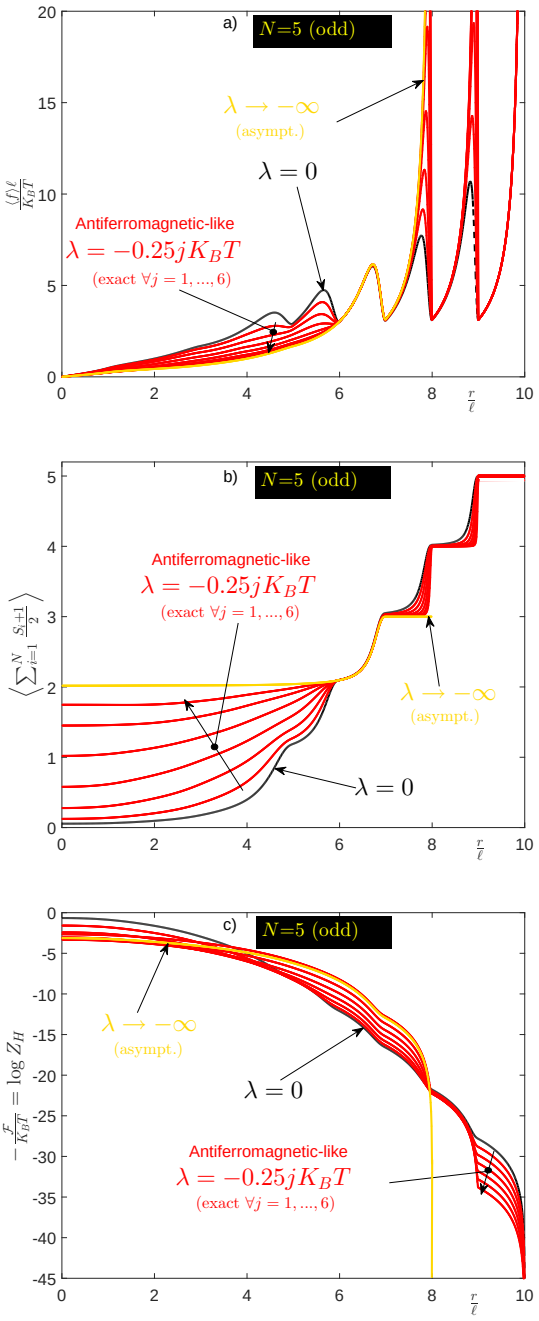

Figure 4.12 – Force-extension response a), average number of unfolded units b) and Helmholtz free energy c) for a chain with strong anti-ferromagnetic-like interactions and an odd number of units. In each panel, the response without interactions ($\lambda = 0$, black dashed curve) is shown together with the exact results for an increasing anti-ferromagnetic interaction ($\lambda = -0.25jk_BT\ \forall j = 1, ..., 6$, red or gray curves) and the asymptotic expansion ($\lambda \to -\infty$, yellow or light gray curve) given in Eq.(4.114). The curves have been obtained with $\Delta E = 5k_BT$, $N = 5$, $\chi = 2$ and $\mu = 0$ [2].

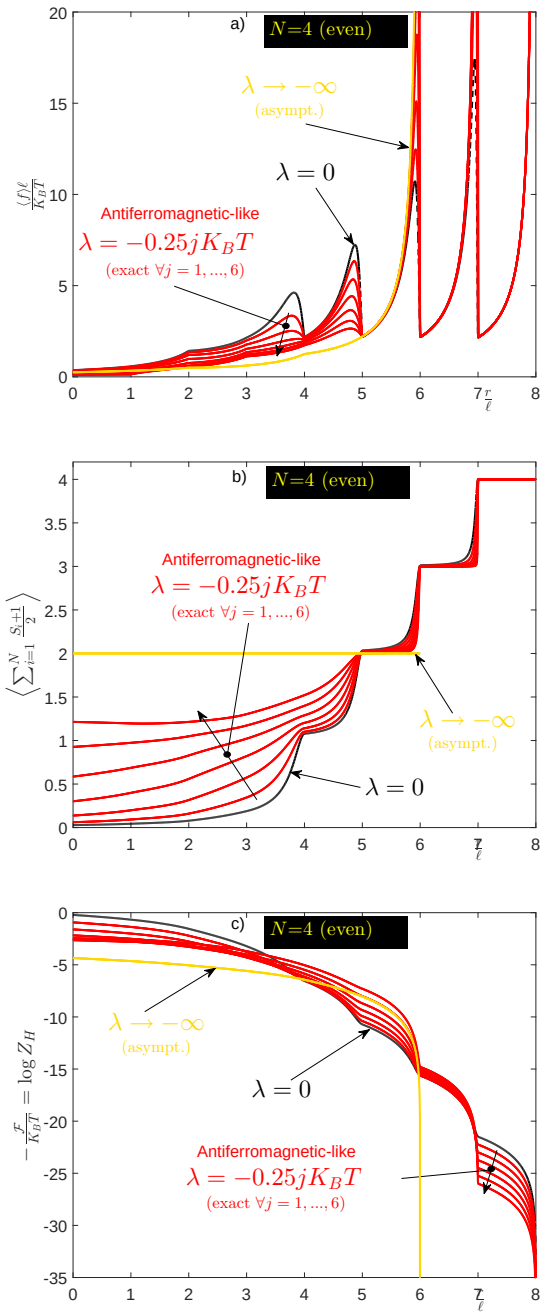

Figure 4.13 – Force-extension response a), average number of unfolded units b) and Helmholtz free energy c) for a chain with strong anti-ferromagnetic-like interactions and an even number of units. In each panel, the response without interactions ($\lambda = 0$, black dashed curve) is shown together with the exact results for an increasing anti-ferromagnetic interaction ($\lambda = -0.25jk_BT\ \forall j = 1, ..., 6$, red or gray curves) and the asymptotic expansion ($\lambda \to -\infty$, yellow or light gray curve) given in Eq.(4.116). The curves have been obtained with $\Delta E = 5k_BT$, $N = 4$, $\chi = 2$ and $\mu = 0$ [2].

or

$$2q - \frac{N-1}{2} - \chi\frac{N+1}{2} + 2\chi p - \frac{r}{\ell} > 0. \tag{4.118}$$

Now, the maximum values of p and q are $\frac{N+1}{2}$ and $\frac{N-1}{2}$, respectively. So we have

$$\frac{r}{\ell} < 2\frac{N+1}{2} - \frac{N+1}{2} - \chi\frac{N-1}{2} + 2\chi\frac{N-1}{2} \tag{4.119}$$

or

$$\frac{r}{\ell} < 2\frac{N-1}{2} - \frac{N-1}{2} - \chi\frac{N+1}{2} + 2\chi\frac{N+1}{2}. \tag{4.120}$$

Equivalently,

$$\frac{r}{\ell} < \frac{N+1}{2} + \chi\frac{N-1}{2} < N\chi \tag{4.121}$$

or

$$\frac{r}{\ell} < \frac{N-1}{2} + \chi\frac{N+1}{2} < N\chi, \tag{4.122}$$

where $N\chi$ is the upper limit of $\frac{r}{\ell}$, attained when all elements are unfolded. It means that the total length of the unit can not exceed $\frac{N-1}{2}+\chi\frac{N+1}{2}$, which is the highest value between Eqs.(4.121) and (4.122). The value $\frac{N+1}{2} + \chi\frac{N-1}{2}$ corresponds to $\frac{N+1}{2}$ folded domains and $\frac{N-1}{2}$ unfolded domains: this is the starting configuration for the strong anti-ferromagnetic system, which can be represented as $\vec{S} = \{\uparrow\downarrow\uparrow\downarrow\uparrow\}$ for $N = 5$, where $\{\uparrow\}$ represents a folded unit and $\{\downarrow\}$ an unfolded one. This configuration is stable with $r = 0$. When we apply a sufficient extension, the unfolded domains become more stable, and we have an inversion in the alternating disposition leading to $\vec{S} = \{\downarrow\uparrow\downarrow\uparrow\downarrow\}$. This transition can be observed through the single peak in the force-extension yellow (or light gray) curve (see Fig.4.12.a) and in the shift from 2 to 3 of the average number of unfolded units (yellow or light gray curve in Fig.4.12.b). As a matter of fact, in this case, we can not unfold all elements because of the hypothesis of large (ideally infinite) anti-ferromagnetism ($\lambda \to -\infty$). For this reason, $Z_H(r)$ is defined for $r < \frac{N-1}{2}\ell + \frac{N+1}{2}\chi\ell$ and not for $r < N\chi\ell$, as in previous cases.

A similar discussion holds for the case with N even. An example of application of Eq.(4.116) is shown in Fig.4.13. While the red (or gray) curves with an increasing value of $|\lambda|$ ($\lambda < 0$) are similar to those shown in Fig.4.12, the yellow (or light gray) curves representing the asymptotic behaviour for $\lambda \to -\infty$ are different and require some comments. In this case, $Z_H(r)$ is different from 0 if $\Lambda_c > 0$, or if $2p - \frac{N}{2} + 2\chi q - \chi\frac{N}{2} - \frac{r}{\ell} > 0$.

Now, p and q assume the same maximum value $\frac{N}{2}$ and therefore we have $\frac{r}{\ell} < 2\frac{N}{2} - \frac{N}{2} + 2\chi\frac{N}{2} - \chi\frac{N}{2}$ or equivalently, $\frac{r}{\ell} < \frac{N}{2} + \chi\frac{N}{2}$. It is evident that $\left(\frac{N}{2} + \chi\frac{N}{2}\right)\ell$ is the length of the chain with $\frac{N}{2}$ folded domains and $\frac{N}{2}$ unfolded domains. So, if $\lambda \to -\infty$, the number of unfolded domains is a constant equal to $\frac{N}{2}$ and $Z_H(r)$ is defined only for $r < \left(\frac{N}{2} + \chi\frac{N}{2}\right)\ell$. In this case, we have no inversion of the configuration, which corresponds to $\vec{S} = \{\downarrow\uparrow\downarrow\uparrow\}$ or to $\vec{S} = \{\uparrow\downarrow\uparrow\downarrow\}$ for $\lambda \to -\infty$ and $N = 4$. These two configuration are indeed completely equivalent from the energetic point of view and they are therefore indistinguishable. As a conclusion, if $\lambda \to -\infty$, the force-extension curve is a monotonically increasing function of r (without peaks), the average number of unfolded units is a constant equal to $N/2$ and the Helmholtz free energy is without cusps (yellow or light gray curves in Fig.4.13).

4.8 The thermodynamic limit

If we consider now the system at the thermodynamic limit, the number of elements of the chain, N, tends to be infinite. In this case, for most of the chain models, the Gibbs and the Helmholtz ensembles are equivalent [20, 124, 138, 139]. It means that the same constitutive equation describes both statistical ensembles. To be precise, it must be admitted that the known criteria to ensure the equivalence of the ensembles in the thermodynamic limit are not applicable in our case with the interactions introduced through the Ising scheme. Nevertheless, we can reasonably assume that this property is also satisfied in this case. However, we are unable to provide a rigorous demonstration of this statement at this time and we leave this research to the near future. In any case, we can study the thermodynamic limit within the Gibbs ensemble. Then, we try to obtain a closed form expression for $\langle r \rangle$ and $\langle \sum_{i=1}^{N} \frac{1}{N} \frac{S_i+1}{2} \rangle$, starting from the Gibbs partition function evaluated for $N \to +\infty$. If $N \to +\infty$ and $\lambda_1 > \lambda_2$, we have from Eq.(4.49)

$$Z_G \approx \frac{x}{1+x^2} \lambda_1^N \left(1 + x^2 \frac{\lambda_1 + \lambda_2}{\lambda_1 - \lambda_2}\right) \tag{4.123}$$

and therefore

$$\log Z_G \approx \log \frac{x}{1+x^2} \left(1 + x^2 \frac{\lambda_1 + \lambda_2}{\lambda_1 - \lambda_2}\right) + N \log \lambda_1 \approx N \log \lambda_1, \tag{4.124}$$

where

$$\lambda_1 = e^{\frac{\lambda}{k_B T}} \left[(p+q) + \sqrt{(p+q)^2 - 4pq\left(1 - e^{-\frac{4\lambda}{k_B T}}\right)}\right]. \tag{4.125}$$

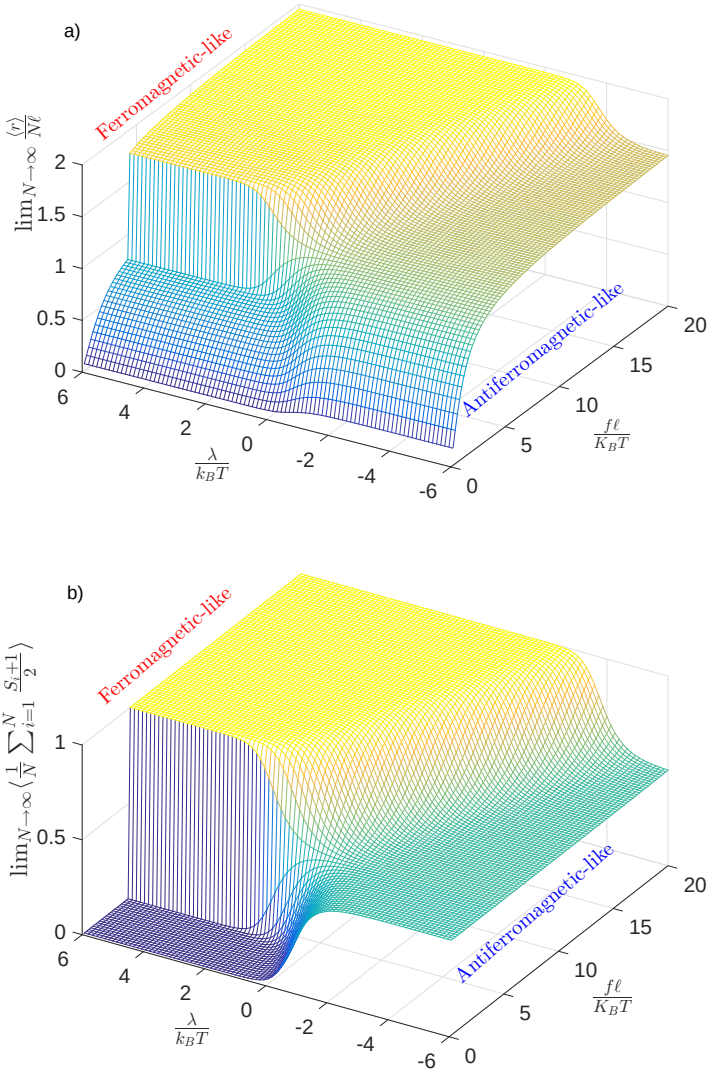

Figure 4.14 – Average normalised extension a) and average number of unfolded units b) versus the applied normalised force and the coefficient $\frac{\lambda}{k_BT}$ for a chain with Ising interactions at the thermodynamic equilibrium within the Gibbs ensemble.

We have

$$\langle r\rangle = k_BT\frac{\partial}{\partial f}\log Z_G \approx k_BT\frac{\partial}{\partial f}N\log\lambda_1. \tag{4.126}$$

Here,

$$p = \frac{1}{2}e^{-\frac{\mu}{k_BT}}\frac{\sinh\xi}{\xi}, \tag{4.127}$$

$$q = \frac{1}{2}e^{\frac{\mu}{k_BT}}\frac{\sinh(\chi\xi)}{\xi}\chi\phi, \tag{4.128}$$

with $\phi = e^{-\frac{\Delta E}{k_B T}}$ and $\xi = \frac{f\ell}{k_B T}$. We can write

$$\begin{aligned}\langle r \rangle &= \ell N \frac{\partial}{\partial \xi} \log \lambda_1 \\ &= \ell N \frac{\partial}{\partial \xi} \log \left[(p+q) + \sqrt{(p+q)^2 - 4pq\left(1 - e^{-\frac{4\lambda}{k_B T}}\right)} \right]. \end{aligned} \tag{4.129}$$

This derivative is calculated first

$$\begin{aligned} D\frac{\sinh(\chi\xi)}{\xi} &= \frac{\chi\xi\cosh(\chi\xi) - \sinh(\chi\xi)}{\xi^2} \\ &= \chi\frac{\sinh(\chi\xi)}{\xi}\left[\frac{\cosh(\chi\xi)}{\sinh(\chi\xi)} - \frac{1}{\chi\xi}\right] \\ &= \chi\frac{\sinh(\chi\xi)}{\xi}\mathcal{L}(\chi\xi), \end{aligned} \tag{4.130}$$

with $D = \frac{\partial}{\partial \xi}$, and $\mathcal{L}$ the Langevin function. Similarly, we can write that

$$D\frac{\sinh\xi}{\xi} = \frac{\sinh\xi}{\xi}\mathcal{L}(\xi). \tag{4.131}$$

In other words, we can say that

$$Dp = p\mathcal{L}(\xi), \tag{4.132}$$

$$Dq = q\chi\mathcal{L}(\chi\xi). \tag{4.133}$$

Hence, we get

$$\begin{aligned}\langle r \rangle &= \ell N \frac{Dp + Dq + \frac{2(p+q)(Dp+Dq) - 4(pDq+qDp)\left(1-e^{-\frac{4\lambda}{k_B T}}\right)}{2\sqrt{(p+q)^2 - 4pq\left(1-e^{-\frac{4\lambda}{k_B T}}\right)}}}{p+q+\sqrt{(p+q)^2 - 4pq\left(1-e^{-\frac{4\lambda}{k_B T}}\right)}} \\ &= \ell N \left\{ \frac{\left[p\mathcal{L}(\xi) + q\chi\mathcal{L}(\chi\xi)\right]\left[p+q+\sqrt{(p+q)^2 - 4pq\left(1-e^{-\frac{4\lambda}{k_B T}}\right)}\right]}{(p+q)\sqrt{(p+q)^2 - 4pq\left(1-e^{-\frac{4\lambda}{k_B T}}\right)} + (p+q)^2 - 4pq\left(1-e^{-\frac{4\lambda}{k_B T}}\right)} \right. \\ &\quad \left. - \frac{2pq\left[\mathcal{L}(\xi) + \chi\mathcal{L}(\chi\xi)\right]\left(1-e^{-\frac{4\lambda}{k_B T}}\right)}{(p+q)\sqrt{(p+q)^2 - 4pq\left(1-e^{-\frac{4\lambda}{k_B T}}\right)} + (p+q)^2 - 4pq\left(1-e^{-\frac{4\lambda}{k_B T}}\right)} \right\}. \end{aligned} \tag{4.134}$$

To simplify, we introduce

$$R = \sqrt{1 - 4\frac{pq}{(p+q)^2}\left(1 - e^{-\frac{4\lambda}{k_B T}}\right)}, \tag{4.135}$$

and

$$\nu = \frac{q}{p+q}. \tag{4.136}$$

Then, we have

$$1-\nu = 1 - \frac{q}{p+q} = \frac{p}{p+q}, \tag{4.137}$$

and we can write

$$R = \sqrt{1 - 4\nu(1-\nu)\left(1 - e^{-\frac{4\lambda}{k_B T}}\right)}. \tag{4.138}$$

Therefore, $p+q+\sqrt{(p+q)^2 - 4pq\left(1-e^{-\frac{4\lambda}{k_BT}}\right)} = (p+q)(1+R)$, and we obtain

$$\begin{aligned}
\langle r \rangle &= N\ell \frac{\frac{p\mathcal{L}(\xi)+q\chi\mathcal{L}(\chi\xi)}{p+q} - \frac{2}{1+R}\frac{pq}{(p+q)^2}\left[\mathcal{L}(\xi)+\chi\mathcal{L}(\chi\xi)\right]\left(1-e^{-\frac{4\lambda}{k_BT}}\right)}{1 - \frac{4}{1+R}\frac{pq}{(p+q)^2}\left(1-e^{-\frac{4\lambda}{k_BT}}\right)} \\
&= N\ell \frac{\frac{p\mathcal{L}(\xi)+q\chi\mathcal{L}(\chi\xi)}{p+q} - \frac{1-R^2}{1+R}\frac{[\mathcal{L}(\xi)+\chi\mathcal{L}(\chi\xi)]}{2}}{1-\frac{1-R^2}{1+R}},
\end{aligned} \tag{4.139}$$

where we used the simple property $1-R^2 = 4\frac{pq}{(p+q)^2}\left(1-e^{-\frac{4\lambda}{k_BT}}\right)$.
We finally obtain

$$\frac{\langle r \rangle}{N\ell} = \frac{1}{R}\left[(1-\nu)\mathcal{L}(\xi) + \nu\chi\mathcal{L}(\chi\xi)\right] + \left(1-\frac{1}{R}\right)\frac{\mathcal{L}(\xi)+\chi\mathcal{L}(\chi\xi)}{2}, \tag{4.140}$$

which is the first important result in this Section. We can perform a similar calculation to find

$$\begin{aligned}
\left\langle \sum_{i=1}^{N} \frac{S_i+1}{2} \right\rangle &= -k_B T \frac{\partial}{\partial \Delta E} \log Z_G \\
&\approx -k_B T N \frac{\partial}{\partial \Delta E} \log \lambda_1.
\end{aligned} \tag{4.141}$$

Since $\phi = e^{-\frac{\Delta E}{k_B T}}$, we have that

$$\frac{\partial \Delta}{\partial \Delta E} = \frac{\partial \Delta}{\partial \phi}\frac{\partial \phi}{\partial \Delta E} = \frac{\partial \Delta}{\partial \phi}\phi\left(\frac{-1}{k_B T}\right), \tag{4.142}$$

where Δ is an arbitrary function. So, we have

$$\left\langle \frac{1}{N}\sum_{i=1}^{N}\frac{S_i+1}{2} \right\rangle = \phi\frac{\partial}{\partial\phi}\log\lambda_1. \tag{4.143}$$

Therefore,

$$\left\langle \frac{1}{N}\sum_{i=1}^{N}\frac{S_i+1}{2}\right\rangle = \phi\frac{\partial}{\partial\phi}\log\left[p+q+\sqrt{(p+q)^2-4pq\left(1-e^{-\frac{4\lambda}{k_BT}}\right)}\right]$$
$$= \phi\frac{q'+\frac{2(p+q)q'-4pq'\left(1-e^{-\frac{4\lambda}{k_BT}}\right)}{2\sqrt{(p+q)^2-4pq\left(1-e^{-\frac{4\lambda}{k_BT}}\right)}}}{p+q+\sqrt{(p+q)^2-4pq\left(1-e^{-\frac{4\lambda}{k_BT}}\right)}}, \quad (4.144)$$

where $q'=\frac{\partial q}{\partial\phi}=\frac{q}{\phi}$. Hence,

$$\left\langle \frac{1}{N}\sum_{i=1}^{N}\frac{S_i+1}{2}\right\rangle = \frac{q\sqrt{(p+q)^2-4pq\left(1-e^{-\frac{4\lambda}{k_BT}}\right)}+(p+q)q-2pq\left(1-e^{-\frac{4\lambda}{k_BT}}\right)}{\left[p+q+\sqrt{(p+q)^2-4pq\left(1-e^{-\frac{4\lambda}{k_BT}}\right)}\right]\sqrt{(p+q)^2-4pq\left(1-e^{-\frac{4\lambda}{k_BT}}\right)}}$$
$$= \frac{q(q+p)R+(p+q)q-2pq\left(1-e^{-\frac{4\lambda}{k_BT}}\right)}{(p+q)(R+1)(p+q)R}. \quad (4.145)$$

Since $1-R^2=4\frac{pq}{(p+q)^2}\left(1-e^{-\frac{4\lambda}{k_BT}}\right)$, we obtain

$$\left\langle \frac{1}{N}\sum_{i=1}^{N}\frac{S_i+1}{2}\right\rangle = \frac{q(p+q)(R+1)-\frac{1}{2}(1-R^2)(p+q)^2}{(p+q)^2R(1+R)}$$
$$= \frac{q-\frac{1}{2}(1-R)(p+q)}{(p+q)R}$$
$$= \frac{1}{R}\left(\frac{q}{p+q}\right)+\frac{1}{2}\left(\frac{R-1}{R}\right). \quad (4.146)$$

Finally,

$$\left\langle \frac{1}{N}\sum_{i=1}^{N}\frac{S_i+1}{2}\right\rangle = \frac{1}{R}\nu+\frac{1}{2}\left(1-\frac{1}{R}\right), \quad (4.147)$$

which is the second important result of this Section. Eq.(4.140) and Eq.(4.147), combined with Eq.(4.136) and Eq.(4.138), give the complete solution for the system in thermodynamic limit (i.e. $N\to\infty$). A numerical application of these equations can be found in Fig.4.14. Here, we show the average normalised extension and the average number of unfolded units versus the applied normalised force and the Ising coefficient $\frac{\lambda}{k_BT}$. We can observe that for large and positive values of λ, all the units unfold cooperatively for a given threshold force. This can be seen in the heavily pronounced force plateau in Fig.4.14.a

and in the sharp transition from 0 to 1 in Fig.4.14.b. On the other hand, for large and negative values of λ, the units remain alternatively folded and unfolded for any value of the applied force. This can be seen in the smooth force-extension curve corresponding to a freely jointed chain response with half units folded and half units unfolded (see Fig.4.14.a). Moreover, this behaviour is confirmed by the average value of the number of unfolded unit, which is equal to $\frac{1}{2}$ (see Fig.4.14.b). For intermediate values of λ, we can observe a continuous transitional behaviour of the system from the anti-ferromagnetic to the ferromagnetic response.

4.9 Ising interactions coupled with extensible units

We consider now a chain of bistable units characterised by a finite intrinsic elastic constant. We start the analysis by considering the Gibbs ensemble defined through the extended Hamiltonian given in Eq.(4.2). Here, for the sake of simplicity, we suppose that the folded and unfolded basins of the potential energy shown in Fig.4.2 exhibit the same finite elastic constant $k(-1) = k(+1) = \kappa$. In this case, it is not difficult to prove that Eq.(4.8) of Section 4.3 can be substituted by the approximated result

$$I = \exp\left(\frac{f^2}{2\kappa k_B T}\right) \prod_{i=1}^{N} \frac{\sinh\left[\frac{f\ell(S_i)}{k_B T}\right]}{\frac{f\ell(S_i)}{k_B T}} \frac{\ell^2(S_i)}{\ell^2}, \tag{4.148}$$

which can be easily proved by calculating the exact solution of Eq.(4.6) and by considering its approximation for relatively high values of κ. The details of this calculation can be found in recent literature [1] and in Chapter 3. Typical values of κ for real macromolecules are compatible with such approximations [168]. Coherently with Eq.(4.148), the Gibbs partition function can be obtained as

$$\begin{aligned} Z_G(f) &= \frac{x}{1+x^2} \exp\left(\frac{N\alpha\xi^2}{2}\right) \\ &\times \left[\lambda_1^N \left(1 + x^2 \frac{\lambda_1 + \lambda_2}{\lambda_1 - \lambda_2}\right) + \lambda_2^N \left(1 - x^2 \frac{\lambda_1 + \lambda_2}{\lambda_1 - \lambda_2}\right)\right], \end{aligned} \tag{4.149}$$

where, with respect to Eq.(4.49) of Section 4.3, we only added an exponential term, which is quadratic in the normalised force $\xi = \frac{f\ell}{k_B T}$ and defined through a coefficient $\alpha = \frac{k_B T}{\kappa \ell^2}$. This latter represents the ratio between the thermal energy and the elastic one, thus measuring the compromise between the enthalpic contribution and fluctuations. The parameters x, λ_1 and λ_2 remain defined as in Section 4.3. Concerning the Helmholtz ensemble, the system is described by the Hamiltonian in Eq.(4.62), where, as before,

we assume that $k(-1) = k(+1) = \kappa$. The calculation of the corresponding partition function is based on the Laplace transform relation between the Gibbs and the Helmholtz ensembles, summed up in Eq.(4.66). If we consider integer values for the parameter χ (ratio between unfolded and folded lengths), the analytic continuation of the Gibbs partition function assumes the form

$$Z_G\left(iy\frac{k_BT}{\ell}\right) = \frac{1}{y^N}\exp\left(-\frac{N\alpha y^2}{2}\right)P(y), \tag{4.150}$$

where $P(y)$ is the periodic function defined in Eq.(4.72) and described by the Fourier coefficients given in Eq.(4.74). The Helmholtz partition function can be finally determined as

$$\begin{aligned} Z_H(r) &= -i\int_\Gamma Z_G\left(iy\frac{k_BT}{\ell}\right)\frac{y}{r}e^{\frac{yr}{\ell}}\mathrm{d}y \\ &= -i\int_\Gamma \frac{P(y)}{y^N}\frac{y}{r}e^{-\frac{N\alpha y^2}{2}}e^{\frac{iry}{\ell}}\mathrm{d}y \\ &= -i\sum_{k=-\infty}^{+\infty} C_k\frac{1}{r}\int_\Gamma \frac{1}{y^{N-1}}e^{-\frac{N\alpha y^2}{2}}e^{i\left(k+\frac{r}{\ell}\right)y}\mathrm{d}y \\ &= -i\sum_{k=-\infty}^{+\infty} C_k\frac{1}{r}J_{N-1}\left(k+\frac{r}{\ell},\frac{N\alpha}{2}\right) \end{aligned} \tag{4.151}$$

where we used the Fourier development of the function $P(y)$ in order to perform the calculation. To complete the task, we have to calculate a sequence of integral of the form

$$J_N(a,b) = \int_\Gamma e^{iay}e^{-by^2}\frac{\mathrm{d}y}{y^N}, \tag{4.152}$$

where $a \in \mathbb{R}$, $b > 0$, $N \in \mathbb{N}$ and the path Γ is given in Fig.4.7. As largely discussed in Chapter 3, an application of the complex variable method allows us to obtain the closed form expression for this integral, as [1]

$$\begin{aligned} J_N(a,b) &= \sqrt{\pi}\left\{b^{\frac{N-1}{2}}e^{-\frac{a^2}{4b}}(2i)^N H_{-N}\left(-\frac{a}{2\sqrt{b}}\right)\right. \\ &\left. +\mathbf{1}(-a)\sum_{h=0}^{N-1}\frac{(ia)^{N-1-h}(-4b)^{h/2}}{ih!(N-1-h)!}\Gamma\left(\frac{h+1}{2}\right)\left[1+(-1)^h\right]\right\}, \end{aligned} \tag{4.153}$$

where $\mathbf{1}(x)$ represents the Heaviside step function, and $H_{-N}(z)$ are the generalisation of the Hermite polynomials, obtained by considering negative indices (see Section 3.3.2 in Chapter 3). We remember that they can be defined as [1, 172]

$$H_{-N}(z) = \frac{1}{(2i)^N\sqrt{\pi}}\int_{-\infty}^{+\infty}\frac{e^{-\eta^2}\mathrm{d}\eta}{(\eta - iz)^N}, \tag{4.154}$$

and they can be obtained recursively through the following formula [1,172]

$$H_{-(N+1)}(z) = \frac{1}{2N} H_{-(N-1)}(z) - \frac{z}{N} H_{-N}(z), \tag{4.155}$$

initialised with $H_0(z) = 1 \; \forall z \in \mathbb{R}$ and $H_{-1}(z) = \frac{\sqrt{\pi}}{2} e^{z^2} \left[1 - \mathrm{erf}\,(z)\right]$ for $z > 0$ and $H_{-1}(z)$ odd, i.e. $H_{-1}(-z) = -H_{-1}(z)$ [1,172]. This recursive law allows us to affirm that $H_{-N}(z)$ is odd if N is odd and that $H_{-N}(z)$ is even if N is even. We also remember that the Hermite elements with $N < 0$ are not polynomials, contrarily to the classical Hermite polynomials, defined with $N > 0$. Many other interesting properties of this sequence of functions can be found in Ref. [172].

An application of the Gibbs and Helmholtz partition functions, stated in Eqs.(4.149) and (4.151), respectively, is presented in Fig.4.15, where we show the force-extension curves for two values of the constant κ, and for three values of the coefficient λ. Firstly, we note that the constant slope of the final part of the force-extension curves represents the finite effective stiffness of the chain, after the unfolding processes. Moreover, it is interesting to remark that the softer systems exhibit a sensibly reduced force peaks in the Helmholtz response. This point can be also noticed by drawing a comparison between Fig.4.8.a, obtained for $\kappa \to \infty$, and Fig.4.15, corresponding to finite values of κ. A similar phenomenon can be also observed in the phase diagram showing the critical behaviour of the spinoidal response of the system. Indeed, we plotted in Fig.4.16 four phase diagrams corresponding to four different values of the elastic constant. We observe that the critical temperatures of the unfolding processes are an increasing function of κ, similarly to the previously discussed force peaks of the Helmholtz response. Besides, as already seen in Fig.4.9 concerning the case with $\kappa \to \infty$, anti-ferromagnetic-like interactions amplify the dissimilarity among the critical temperatures, while ferromagnetic-like interactions reduce this contrast, eventually producing a more uniform response of the unfolding processes.

4.10 Conclusion

In this Chapter, we investigated the behaviour of a chain made of two-state units coupled through an Ising interaction scheme with the coefficient λ, providing a description of the effects of bistability and cooperativity in biological and artificial micro- and nano-systems. Accordingly, we firstly studied our model by means of the statistical mechanics of small systems, i.e. far from the thermodynamic limit. We analysed that the force-extension response exhibits sharp or smooth transitions within the Gibbs ensemble

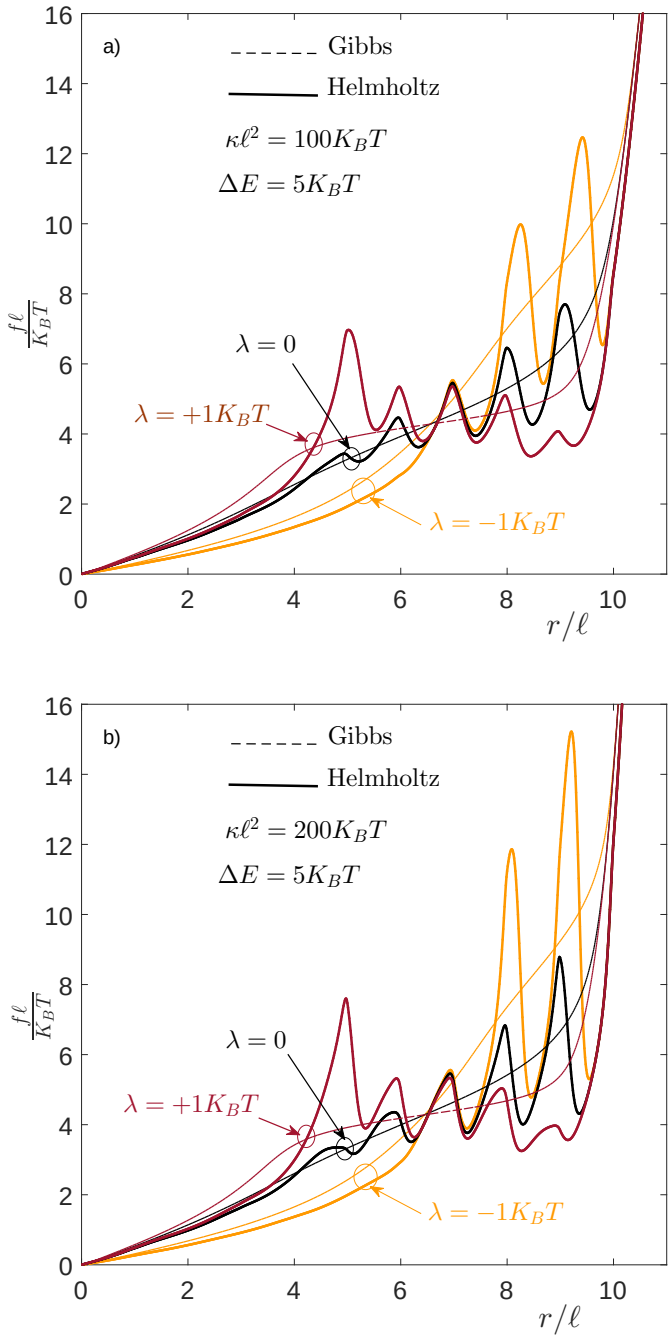

Figure 4.15 – Force-extension response for a chain of interacting units with finite intrinsic stiffness κ. We used the values $\kappa\ell^2 = 100k_BT$ a), and $\kappa\ell^2 = 200k_BT$ b). In each panel, the response without interactions ($\lambda = 0$) is shown together with the results with $\lambda = \pm 1k_BT$. The curves have been obtained with $\Delta E = 5k_BT$, $N = 5$, $\chi = 2$ and $\mu = 0$ [2].

depending on λ, and, within the Helmholtz ensemble, the force peaks present a hierarchy depending on the "positive" or "negative" cooperativity. The spin variable approach

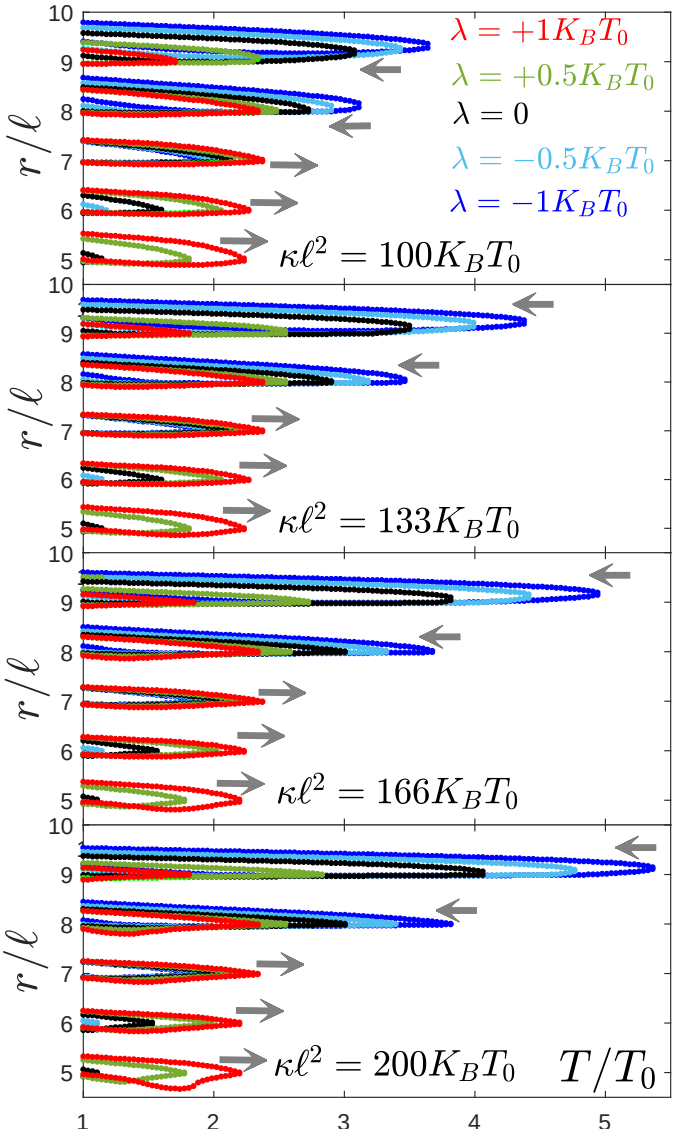

Figure 4.16 – Spinoidal phase diagrams obtained for four different values of the intrinsic stiffness κ of the units (defined by $\kappa\ell^2 = 100, 133, 166, 200 k_BT_0$). They show the end-points of the spinoidal intervals versus the temperature of the system ($T > T_0$). The curves have been obtained with $\Delta E = 5k_BT_0$, $N = 5$, $\chi = 2$, $\mu = 0$ and five different values of the Ising interaction coefficient $\lambda = 0, \pm 0.5, \pm 1 k_BT_0$. The parameters ℓ and T_0 are arbitrary (typical values are $\ell = 1$nm and $T_0 = 300$K) [2].

permitted to elaborate semi-analytic or closed-form expressions for the relevant observables. Within the Gibbs ensemble, interactions are studied in the model with the help of the transfer matrix technique. Within the Helmholtz ensemble, the partition function is elaborated from the Gibbs partition function with the help of the Laplace transform.

In addition, we propose results exploring asymptotic cases, and describing the behaviour of the system under weak and strong Ising interactions (for both ferromagnetic-like and anti-ferromagnetic-like schemes).
We also investigated a form of criticality exhibited by the system. Especially, our analysis focused on the critical behaviour of the spinoidal regions, characterising the part of the isometric response, showing a negative differential elastic stiffness. We proved that each unfolding process exhibits a critical temperature defined by stating that we measure a negative differential stiffness for subcritical temperatures and a positive differential stiffness for supercritical temperatures. This behaviour is influenced by the cooperativity, which has the capability to make the critical temperatures of the unfolding processes more uniform. We can therefore state that a positive cooperativity increases the resistance to fluctuations, making the spinoidal intervals equally stable to temperature variations.
We finally investigated the thermodynamic limit within the Gibbs ensemble and the generalisation of the theory with a finite intrinsic elasticity of the units. A next improvement to represent more realistic situations concerns the heterogeneity of all parameters defining the properties of the chain units. Indeed, for the macromolecules of biological origin, such as proteins, parameters should be freely chosen. Thus, it seems important to introduce this point, in order to determine the full unfolding pathway, depending on the system microstructure. As an example, this is directly related to the biological function of a protein.

Chapter 5

Two-state heterogeneous chains

5.1 Introduction

In this Chapter, we investigate the behaviour of a chain of two-state units with a heterogeneous distribution of the energy jumps between the folded and unfolded states. For the sake of generality, we also consider the heterogeneous elastic constants, but we will observe that they do not modify the qualitative behaviour of the chain. Previously, we considered that all units were identical, being described by the potential energy of Fig.2.8 with the results discussed in previous Chapters 2, 3 and 4. The bimodal energy potential of each unit is therefore composed of two potential wells with different basal energies, separated by a given energy barrier. For a chain made of homogeneous units, loaded by soft or hard devices, all units have the same probability to unfold at each switching occurrence. Therefore, it is impossible to identify an unfolding pathway, i.e. a sequence of well defined unfolding processes. Conversely, the heterogeneity represents a quenched disorder from the statistical mechanics point of view and is able to break the symmetry among the units. Introducing an heterogeneous distribution of energy jumps will allow to obtain an unfolding pathway, as the heterogeneity will generate different unfolding probabilities at each transition, eventually producing a preferential order in the units unfolding. It is important to underline that the process corresponding to the complete unfolding of a chain is a probabilistic event. Consequently, the observed pathway assumes a statistical character. It means that if we repeat the experiment many times, we can observe, as a result, different unfolding pathways. As a matter of fact, the heterogeneity defines different probabilities for each pathway and therefore, it is possible to identify the most probable unfolding pathway for a given chain. We will discuss this concept by introducing the unfolding pathway identifiability. It is worth noticing that the symmetry breaking can also be obtained with non-local interactions between the units of the chain [178, 179],

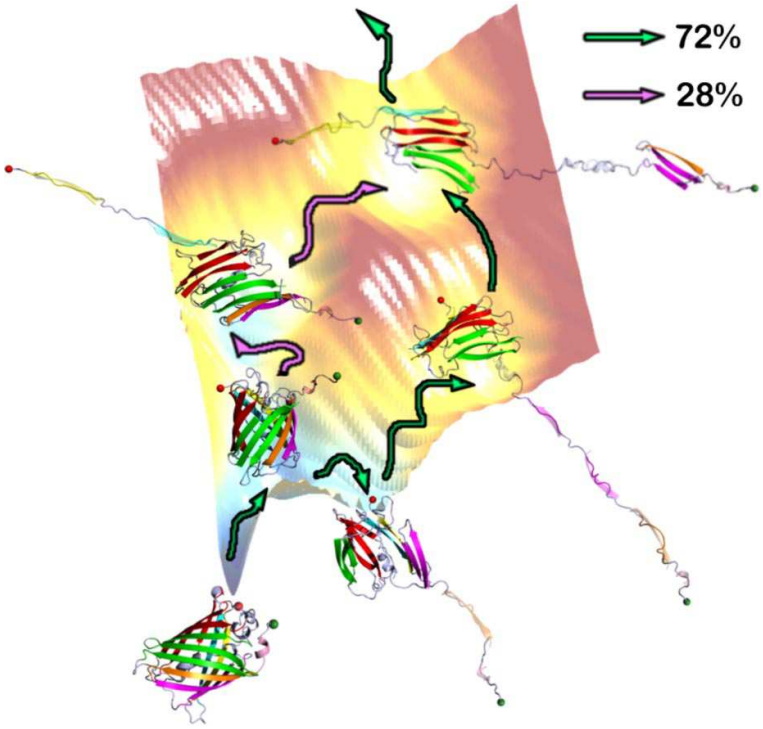

Figure 5.1 – The complex energy landscape for GFP unfolding is constructed based on the results of simulations and experiments. Starting from the folded structure, which corresponds to the native basin of attraction, unfolding occurs by bifurcation in the pathway after the rupture of the α-helix. The molecules that unfold by the dominant pathway are shown by green arrows and the purple arrows show the fate of GFP molecules that follow the minor pathway. The structures of the intermediates in the various basins are explicitly shown. The approximate fraction of the molecules along each pathway is indicated. These numbers can be altered by mutations (cross-link in this study) and by changing the force direction. Thus, the energy landscape is not only rugged, but also can be manipulated [180].

similarly to the case of the Ising scheme discussed in Chapter 4. We propose here a mathematical model explaining why the heterogeneity breaks the symmetry and generates the unfolding pathways.

5.2 Examples of unfolding pathway

The nanomanipulation of biomolecules by using single-molecule force spectroscopy and computer simulations allowed to study the energy landscape of macromolecules undergoing conformational transitions. These studies are very useful to understand the different structures appearing during the unfolding processes. As discussed in Section 1.1, the functions of a protein are strongly related to its structural conformations. Therefore, forced

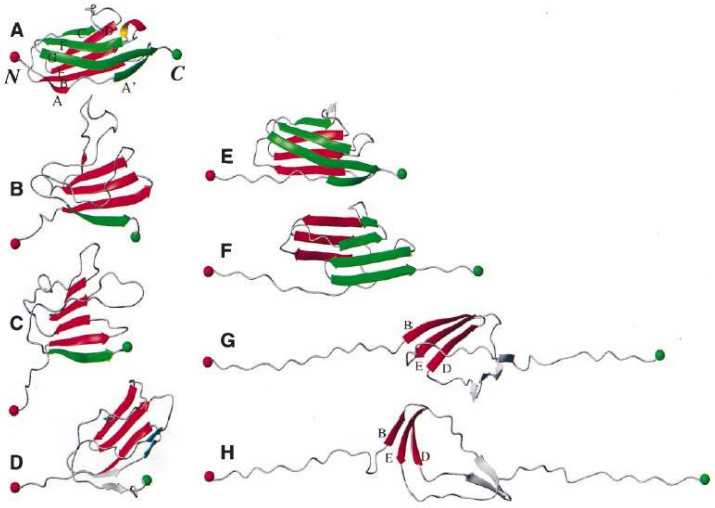

Figure 5.2 – Pathway for the unfolding of the 1TIT domain. (A) X-ray structure, $r_{NC} = 45$ Å; B–D temperature-induced unfolding pathway; (B) 0.6 ns, $r_{NC} = 34$ Å, (C) 1.5 ns, $r_{NC} = 37$ Å; (D) 4.8 ns, $r_{NC} = 38$ Å. E–H forced unfolding pathway; (E) $r_{NC} = 54$ Å; (F) $r_{NC} = 75$ Å; (G) $r_{NC} = 126$ Å; (H) $r_{NC} = 138$ Å. Native elements of secondary structure are shown by using the following colour code. The N- and C-terminal atoms are represented by red and green spheres, respectively; the same colours are used to represent β-strands belonging to the N- and C-terminal sheets of the β-sandwich proteins. Nonnative elements are in grey [181].

unfolding is of particular interest for proteins that are under mechanical stress in their biological function. We present here some examples of unfolding pathway taken from the recent literature. In Ref. [180], simulations and single-molecule force spectroscopy are used to reveal the complex energy landscape of *green fluorescent protein*, also known as *GFP*. This protein is expressed in many species, among which human, bacteria and fish. Its gene is used as a reporter gene, because it can be observed in the laboratory due to its fluorescence and its detectable enzymatic activity. It can also be used in different forms to make biosensors. In Fig.5.1, one can observe the complex energy landscape of the GFP and its numerous conformational transitions. It is interesting to observe that two different unfolding pathways can be followed during the unfolding, with two different probabilities given in this case by 0.72 and 0.28. This point proves the fact that the unfolding pathway is a statistical concept and justifies the introduction of the concept of identifiability in Section 5.4.

It is also possible to study the unfolding of a protein made of α-helices or β-sheets [181]. Concerning the example in Fig.5.2, we can see the comparison between a thermal induced unfolding and a force reduced unfolding for an immunoglobulin (Ig) domain (1TIT), which is a β-sandwich protein. Interestingly, an hairpin not present in the native structure

is formed in the stage H of the unfolding process induced by a force. Anyway, we can observe that the overall unfolding process is composed of a well ordered sequence of minor unfolding occurrences. Other similar examples are thoroughly described and discussed in Ref. [181]. As previously discussed, we consider in this Chapter the case of a chain of two-state heterogeneous units. Introducing the heterogeneity is of vital importance, since it will allow to identify the different conformational states. We prove, in particular, that a small heterogeneity is sufficient to generate the unfolding pathways.

5.3 Two-state heterogeneous one-dimensional system

The study of the effect of the heterogeneity on the mechanical and configurational behaviour of bistable chains will be performed on one-dimensional system in order to simplify the mathematical complexity of the problem and therefore to better appreciate the mechanisms at the origin of the unfolding pathway generation. We will analyse both the Gibbs and the Helmholtz ensembles.

5.3.1 The Gibbs ensemble

We consider a one-dimensional chain aligned with the x-axis of a reference frame and composed of N units, which are bistable (in each direction, $x < 0$ and $x > 0$). Hence, they can be either folded or unfolded (in each direction). These two-state correspond to a length ℓ for the folded state and to a length $\chi\ell$ for the unfolded one. The potential energy of the units is described by

$$U_i(x, y_i) = v_i(y_i) + \frac{1}{2}k_i \left(x - y_i\ell\right)^2 , \tag{5.1}$$

where $v_i(\pm 1) = 0$ and $v_i(\pm\chi) = \Delta E_i$ are the energy jumps between folded and unfolded states (arbitrarily varying with i). This is coherent with the analysis performed in Section 2.2.1. The parameters k_i represent the elastic constants of the units, which are independent of the folded or unfolded state but possibly heterogeneous along the chain. Moreover, y_i is the spin variable and assumes the values in $\mathbf{S} = \{\pm 1, \pm\chi\}$. The value of $y_i \in \mathbf{S}$ allows the identification of the energy well explored by the i-th unit. We remark that the potential energy in Eq.(5.1) is symmetric with respect to the coordinate x, thus describing the folding-unfolding process in both direction of $x > 0$ and $x < 0$ (see Fig.2.4). While the model should be three-dimensional to exactly represent, e.g., the behaviour of real macromolecules, we adopt a one-dimensional scheme for the sake of simplicity. In this regard, the passage at $x = 0$ is somewhat unphysical, but we will study extensions and forces only

in the positive direction. Since we consider a one-dimensional system, we introduced four potential wells (two folded and two unfolded). Indeed, the consideration of positive and negative orientations of the elements allows modelling entropic, enthalpic, unfolding and over-stretching regimes, as discussed in Ref. [113] and in Section 2.2.1. As in previous Chapters, it is also important to underline that the approximation introduced by the spin variables allows us to perform an analytic study of the system under the hypothesis of thermodynamic equilibrium [1, 2, 111–113]. However, if we consider the out-of-equilibrium regime, the dynamics of the folding-unfolding process is also influenced by the energy barrier between the states, as classically described by the reaction-rate Kramers theory [16]. Concerning the Gibbs ensemble, the total potential energy of the one-dimensional system is given by

$$U_G\left(\vec{x},\vec{y},f\right)=\sum_{i=1}^{N}U_i\left[(x_i-x_{i-1}),y_i\right]-f\cdot x_N-\sum_{i=1}^{N}g_iy_i, \tag{5.2}$$

where $x_0=0$ and f is the applied force to the last element of the chain. Here, we defined $\vec{x}=(x_1,...,x_N)\in\mathbb{R}^N$ and $\vec{y}=(y_1,...,y_N)\in\mathbf{S}^N$. Moreover, the quantities g_i represent an external field $\vec{g}$ directly acting on the configurational state of the elements (it acts as a chemical potential) [1]. While we will consider $\vec{g}=0$ in the examples considered in this Chapter, the vector $\vec{g}$ is very important from the mathematical point of view, to easily calculate the heterogeneous average values of the spin variables (see Eqs.(5.10), (5.11), (5.16) and (5.17) below). Therefore, we can write the partition function in the Gibbs ensemble by summing the discrete spins and integrating the continuous coordinates We have more explicitly

$$U_G\left(\vec{x},\vec{y},f\right)=\sum_{i=1}^{N}\left[v_i(y_i)+\frac{1}{2}k_i\left(x_i-x_{i-1}-y_i\ell\right)^2\right]-f\cdot x_N-\sum_{i=1}^{N}g_iy_i. \tag{5.3}$$

Therefore, we can write the partition function in the Gibbs ensemble

$$Z_G(f)=\sum_{y_1\in\mathbf{S}}...\sum_{y_N\in\mathbf{S}}\int_{\mathbb{R}}...\int_{\mathbb{R}}e^{-\frac{U_G(\vec{x},\vec{y},f)}{k_BT}}\mathrm{d}x_1...\mathrm{d}x_N. \tag{5.4}$$

We apply now the following change of variables: $x_1 - x_0 = \xi_1$, $x_2 - x_1 = \xi_2$, ..., $x_N - x_{N-1} = \xi_N$, from which we get $x_N = \sum_{j=1}^{N} \xi_j$, with $x_0 = 0$. Hence, we obtain

$$\begin{aligned} Z_G(f) &= \sum_{\vec{y}\in\mathbf{S}^N} \int_{\mathbb{R}^N} \exp\left\{ \sum_{i=1}^{N} \left[-\frac{v_i(y_i)}{k_B T} - \frac{1}{2}\frac{k_i}{k_B T}\left(x_i - x_{i-1} - y_i\ell\right)^2 \right] \right. \\ &\quad \left. + \sum_{i=1}^{N} \frac{g_i y_i}{k_B T} + \frac{f x_N}{k_B T} \right\} \mathrm{d}x_1 ... \mathrm{d}x_N \\ &= \sum_{\vec{y}\in\mathbf{S}^N} \int_{\mathbb{R}^N} \exp\left\{ \sum_{i=1}^{N} \left[-\frac{v_i(y_i)}{k_B T} - \frac{1}{2}\frac{k_i}{k_B T}\left(\xi_i - y_i\ell\right)^2 + \frac{g_i y_i}{k_B T} + \frac{f\xi_i}{k_B T} \right] \right\} \mathrm{d}\xi_1 ... \mathrm{d}\xi_N \\ &= \prod_{i=1}^{N} \left\{ \sum_{y_i\in\mathbf{S}} e^{-\frac{v(y_i)}{k_B T}} \int_{\mathbb{R}} \exp\left[-\frac{1}{2}\frac{k_i}{k_B T}(\xi_i - y_i\ell)^2 + \frac{f\xi_i}{k_B T} + \frac{g_i y_i}{k_B T} \right] \mathrm{d}\xi_i \right\}. \end{aligned} \tag{5.5}$$

We focus now on the integral, which can be evaluated as [113]

$$\begin{aligned} &\int_{\mathbb{R}} \exp\left[-\frac{1}{2}\frac{k_i}{k_B T}(\xi_i - y_i\ell)^2 + \frac{f\xi_i}{k_B T} + \frac{g_i y_i}{k_B T} \right] \mathrm{d}\xi_i \\ &= \sqrt{\frac{2\pi k_B T}{k_i}} \exp\left[\frac{y_i(\ell f + g_i)}{k_B T} + \frac{f^2}{2k_B T k_i} \right]. \end{aligned} \tag{5.6}$$

Then, we get

$$\begin{aligned} Z_G(f) &= \prod_{i=1}^{N} \left\{ \sum_{y_i\in\{\pm 1, \pm\chi\}} e^{-\frac{v(y_i)}{k_B T}} \sqrt{\frac{2\pi k_B T}{k_i}} \exp\left[\frac{y_i(\ell f + g_i)}{k_B T} + \frac{f^2}{2k_B T k_i} \right] \right\} \\ &= (8\pi k_B T)^{N/2} \prod_{i=1}^{N} \left\{ \sqrt{\frac{1}{k_i}} \exp\left(\frac{f^2}{2k_B T k_i} \right) \left[\cosh\left(\frac{\ell f + g_i}{k_B T} \right) \right. \right. \\ &\quad \left. \left. + \phi_i \cosh\left(\frac{\chi\ell f + \chi g_i}{k_B T} \right) \right] \right\}, \end{aligned} \tag{5.7}$$

where $\phi_i = e^{-\frac{\Delta E_i}{k_B T}}$ are the Boltzmann factors calculated with the energy jumps ΔE_i. Finally, we get

$$\begin{aligned} Z_G(f) &= (8\pi k_B T)^{N/2} \left(\prod_{i=1}^{N} \frac{1}{\sqrt{k_i}} \right) \exp\left(\frac{f^2}{2k_B T} \sum_{i=1}^{N} \frac{1}{k_i} \right) \\ &\quad \times \prod_{i=1}^{N} \left[\cosh\left(\frac{\ell f + g_i}{k_B T} \right) + \phi_i \cosh\left(\frac{\chi\ell f + \chi g_i}{k_B T} \right) \right]. \end{aligned} \tag{5.8}$$

We define the effective elastic stiffness as $k_{\text{eff}} = N\left(\sum_{i=1}^{N} \frac{1}{k_i} \right)^{-1}$ and we obtain

$$Z_G(f) = c \times \exp\left(\frac{N f^2}{2k_B T k_{\text{eff}}} \right) \prod_{i=1}^{N} \left[\cosh\left(\frac{\ell f + g_i}{k_B T} \right) + \phi_i \cosh\left(\frac{\chi\ell f + \chi g_i}{k_B T} \right) \right], \tag{5.9}$$

with c a constant. The macroscopic behaviour of the chain is described by the force-extension response and by the average value of the spin variables, which can be obtained through the Gibbs partition function as follows [1, 113]

$$\langle x_N \rangle = k_B T \frac{\partial}{\partial f} \log Z_G, \tag{5.10}$$

$$\langle\, y_i \,\rangle = k_B T \frac{\partial}{\partial g_i} \log Z_G. \tag{5.11}$$

When the external field $\vec{g}$ is 0, these results may be evaluated as

$$\langle x_N \rangle = \frac{f}{k_{\text{eff}}} + \sum_{i=1}^{N} \ell \frac{\sinh\left(\frac{\ell f}{k_B T}\right) + \chi\phi_i \sinh\left(\frac{\chi \ell f}{k_B T}\right)}{\cosh\left(\frac{\ell f}{k_B T}\right) + \phi_i \cosh\left(\frac{\chi \ell f}{k_B T}\right)}, \tag{5.12}$$

$$\langle\, y_i \,\rangle = \frac{\sinh\left(\frac{\ell f}{k_B T}\right) + \chi\phi_i \sinh\left(\frac{\chi \ell f}{k_B T}\right)}{\cosh\left(\frac{\ell f}{k_B T}\right) + \phi_i \cosh\left(\frac{\chi \ell f}{k_B T}\right)}. \tag{5.13}$$

While the first expression represents the macroscopic mechanical response of the system, the second one describes the configurational state (folded or unfolded) of the units in terms of the applied force. Interestingly enough, Eqs.(5.12) and (5.13) can be combined to give $\langle x \rangle = \frac{Nf}{k_{\text{eff}}} + \ell \sum_{i=1}^{N} \langle\, y_i \,\rangle$, which represents a spring-like behaviour with the equilibrium length controlled by the spin variables. An application of Eqs.(5.12) and (5.13) can be found in Figs.5.3.b, 5.3.c, and 5.3.d, where the force-extension relation, the average spin variables and the quantities $\partial\langle\, y_i \,\rangle / \partial\langle x \rangle$ versus $\langle x \rangle$ are represented for a homogeneous chain with $\Delta E_i = \Delta E \ \forall i$ (blue curves). These results describe a synchronous unfolding of the units for a given threshold force given by $f^* = \Delta E / [(\chi - 1)\ell]$ [113, 130, 159]. It is a well-known behaviour observed in DNA [71, 73, 182], and in other molecules of biological origin [132, 183]. A second example can be found in Figs.5.3.j, 5.3.k, and 5.3.l, where the same quantities have been shown for a heterogeneous chain, as represented in Fig.5.3.i. In this case, while the force-extension curve is only slightly modified, the spin variables assume different behaviours for the different units, proving the emergence of an unfolding pathway induced by the heterogeneity of the metastable states energy levels. Also, the blue curve in Fig.5.3.k shows the mean value of the numbered curves and is similar to the Gibbs response of the homogeneous case. This is true since we used the same parameter χ for all the units of the chain. It is important to remark that, for an heterogeneous parameter χ, we can have a different behaviour between $\langle y_i \rangle$ of the homogeneous chain and $\frac{1}{N} \sum_{i=1}^{N} \langle\, y_i \,\rangle$ of the heterogeneous chain. To conclude, we observe that the quantity $\partial\langle\, y_i \,\rangle / \partial\langle x \rangle$ can be considered as an approximated measure of the probability density

of the position x at which a transition occurs between the states of the i-th unit (see below for details). Therefore, Fig.5.3.l confirms the identification of a unfolding sequence induced by the heterogeneity. The analysis of the Gibbs ensemble is the starting point to study the behaviour of the Helmholtz ensemble, which is the core of our investigation as discussed below.

5.3.2 The Helmholtz ensemble

We are now interested in the case of a two-state heterogeneous one-dimensional system in the Helmholtz ensemble. To analyse this system, we use the Fourier relation linking the Gibbs and Helmholtz partition functions [20, 124]

$$Z_H(x) = \int_{-\infty}^{+\infty} Z_G(-i\omega k_B T)\exp(i\omega x)\mathrm{d}\omega. \tag{5.14}$$

First of all, we have from Eq.(5.9)

$$\begin{aligned} Z_G\left(-i\omega k_B T\right) &= c \times \exp\left(-\frac{N\omega^2 k_B T}{2k_{\text{eff}}}\right)\prod_{i=1}^{N}\left[\cosh\left(\frac{g_i - i\omega\ell k_B T}{k_B T}\right)\right. \\ &\left. +\phi_i \cosh\left(\frac{\chi g_i - i\chi\omega\ell k_B T)}{k_B T}\right)\right]. \end{aligned} \tag{5.15}$$

with c a constant. However, it is better to continue the calculation by observing that we have to determine the following quantities, both for $\vec{g} = 0$

$$\langle f\rangle = -k_B T\frac{1}{Z_H}\frac{\partial Z_H}{\partial x}, \tag{5.16}$$

$$\langle y_j\rangle = k_B T\frac{1}{Z_H}\frac{\partial Z_H}{\partial g_j}. \tag{5.17}$$

Therefore, we develop both Z_H and $\frac{\partial Z_H(x)}{\partial g_j}$ with $\vec{g} = 0$, as follows

$$Z_H(x) = \int_{-\infty}^{+\infty} e^{-\frac{N\omega^2 k_B T}{2k_{\text{eff}}}} e^{i\omega x}\prod_{i=1}^{N}(a + \phi_i b)\,\mathrm{d}\omega, \tag{5.18}$$

$$\begin{aligned} \frac{\partial Z_H(x)}{\partial g_j} &= \left(-\frac{i}{k_B T}\right)\int_{-\infty}^{+\infty}\left[\sin(\omega\ell) + \phi_j\chi\sin(\omega\chi\ell)\right] \\ &\times \prod_{k=1,k\neq j}^{N}(a + \phi_k b)\, e^{-\frac{N\omega^2 k_B T}{2k_{\text{eff}}}} e^{i\omega x}\mathrm{d}\omega, \end{aligned} \tag{5.19}$$

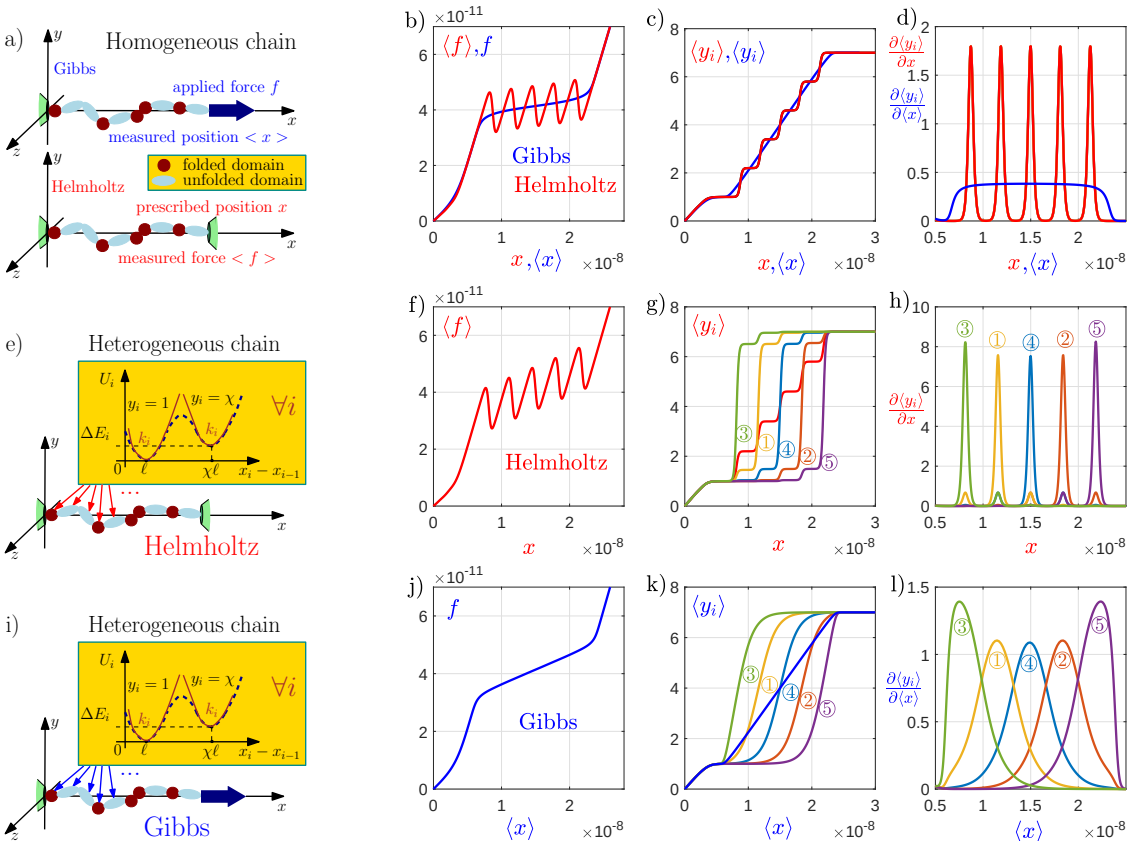

Figure 5.3 – Folding/unfolding processes in homogeneous and heterogeneous chains. Panel a): Gibbs and Helmholtz conditions in homogeneous chains. Panel b): force-extension response for isotensional (blue) and isometric (red) conditions. Panel c): average spin variables (homogeneous chain). All spins have the same behaviour (no unfolding pathway). Panel d): plot of the quantities $\partial\langle\, y_i \,\rangle/\partial x$ (or $\partial\langle\, y_i \,\rangle/\partial\langle x\rangle$) versus x (or $\langle x\rangle$) for homogeneous chains (proportional to the probability density of the position of the unfolding transition of the i-th unit). Panel e): Scheme of an heterogeneous chain under isometric condition (different ΔE_i and k_i). Panel f): Helmholtz force-extension response, slightly modified by the heterogeneity. Panel g): the average spin variables under isometric condition show the pathway given by the ascending order of the values ΔE_i: ③,①,④,②,⑤ (in this example $\Delta E_3 < \Delta E_1 < \Delta E_4 < \Delta E_2 < \Delta E_5$). Red curve: mean value $\frac{1}{N}\sum_{i=1}^{N}\langle\, y_i \,\rangle$ of the averaged spins. Panel h): plot of the quantities $\partial\langle\, y_i \,\rangle/\partial x$, confirming the symmetry breaking induced by the heterogeneity. Panel i): scheme of an heterogeneous chain under isotensional condition (different ΔE_i and k_i). Panel j): Gibbs force-extension response, slightly modified by the introduction of the heterogeneity. Panel k): the average spin variables under isotensional condition show the unfolding pathway given by the ascending order of the values ΔE_i: ③,①,④,②,⑤ (in this example $\Delta E_3 < \Delta E_1 < \Delta E_4 < \Delta E_2 < \Delta E_5$, as before). Blue curve: mean value $\frac{1}{N}\sum_{i=1}^{N}\langle\, y_i \,\rangle$ of the averaged spins. Panel l): plot of the quantities $\partial\langle\, y_i \,\rangle/\partial\langle x\rangle$, confirming the symmetry breaking induced by the heterogeneity. We adopted the following parameters: $N = 5$, $\ell = 0.5$nm, $T = 300$K, $\chi = 7$, $k_i = 0.04$N/m $\forall i$, $\Delta E_i = 30 k_B T$ $\forall i$ in the homogeneous chain and $\Delta E_1 = 27.5 k_B T$, $\Delta E_2 = 32.5 k_B T$, $\Delta E_3 = 25 k_B T$, $\Delta E_4 = 30 k_B T$, $\Delta E_5 = 35 k_B T$ in the heterogeneous case [4].

where we omitted the unimportant multiplicative constant and we defined $a = \cos(\omega\ell)$ and $b = \cos(\omega\chi\ell)$. Hence, we calculate

$$\begin{aligned}\prod_{i=1}^{N}(a+\phi_i b) &= a^N + a^{N-1}b\sum_{j=1}^{N}\phi_j + \frac{1}{2}a^{N-2}b^2\sum_{i\neq j}\phi_i\phi_j + \frac{1}{3!}a^{N-3}b^3 \\ &\quad \times \sum_{i\neq j,k\neq j,i\neq k}\phi_i\phi_j\phi_k + ... + b^N\phi_1\times...\times\phi_N \\ &= \sum_{k=0}^{N}a^{N-k}b^k S_k, \end{aligned} \tag{5.20}$$

with $S_0 = 1$ and

$$S_k = \frac{1}{k!}\sum_{J_a\neq J_b\forall a\neq b}\phi_{j1}\times...\times\phi_{jk}. \tag{5.21}$$

The quantities S_k are *called elementary symmetric polynomials* in the variables $\phi_1...\phi_N$. We describe here the theory for the calculation of such quantities and we postpone the calculation of $\frac{\partial Z_H(x)}{\partial g_j}$. We consider a polynomial $f(x)$ defined as

$$f(x) = \sum_{k=0}^{N}a_k x^k = a_0 + a_1 x + a_2 x^2 + ... + a_N x^N, \tag{5.22}$$

and we suppose that it has the N zeros $\phi_1, ..., \phi_N$. It means that

$$\begin{aligned} f(x) &= a_N(x-\phi_1)\times...\times(x-\phi_N) \\ &= a_N\prod_{k=1}^{N}(x-\phi_k) \\ &= a_N\left[x^N - x^{N-1}\sum_{i=1}^{N}\phi_i + \frac{1}{2}x^{N-2}\sum_{i\neq j}^{N}\phi_i\phi_j - \frac{1}{3!}x^{N-3}\sum_{i\neq j\neq k}^{N}\phi_i\phi_j\phi_k + ...\right. \\ &\quad \left. +(-1)^N\phi_1\times...\times\phi_N\right] \\ &= a_N\left[x^N - x^{N-1}S_1 + x^{N-2}S_2 - x^{N-3}S_3 + ... + (-1)^N S_N\right]. \end{aligned} \tag{5.23}$$

We can also write

$$\left\{\begin{array}{lll} \sum_{i=1}^{N}\phi_i & = -\frac{a_{N-1}}{a_N} & = S_1, \\ \frac{1}{2}\sum_{i\neq j}\phi_i\phi_j & = \frac{a_{N-2}}{a_N} & = S_2, \\ \frac{1}{3!}\sum_{i\neq j\neq k}\phi_i\phi_j\phi_k & = -\frac{a_{N-3}}{a_N} & = S_3, \\ ... & & \\ \phi_1\times...\times\phi_N & = (-1)^N\frac{a_0}{a_N} & = S_N. \end{array}\right. \tag{5.24}$$

The problem is to easily calculate all these quantities, which can be computationally expensive because of the very important number of permutations. One possibility to simplify the calculation is based on the introduction of the so-called *power sums*, defined as

$$P_h = \sum_{t=1}^{N} \phi_t^h. \tag{5.25}$$

We consider, by definition, $P_0 = N$. We search for the relation between the power sums and the elementary symmetric polynomials. We have that

$$\sum_{k=0}^{N} a_k \phi_t^k = 0 \ \ \forall t = 1, ..., N, \tag{5.26}$$

since ϕ_t is a zero of $f(x)$. By multiplying this expression by ϕ_t^m, we get

$$\sum_{k=0}^{N} a_k \phi_t^{k+m} = 0 \ \ \forall t = 1, ..., N, \forall m \geq 0. \tag{5.27}$$

We can sum over t from 1 to N

$$\sum_{k=0}^{N} a_k P_{k+m} = 0 \ \ \forall m \geq 0. \tag{5.28}$$

These equations allow the determination of all P_h if we know the first $P_1, ...P_{N-1}$, and we remember that $P_0 = N$. To get equations for $P_1, ...P_{N-1}$, we consider

$$f(x) = a_N(x - \phi_1) \times ... \times (x - \phi_N), \tag{5.29}$$

and

$$\begin{aligned} f'(x) &= \frac{f(x)}{x - \phi_1} + ... + \frac{f(x)}{x - \phi_N} \\ &= \sum_{q=1}^{N} \frac{f(x)}{x - \phi_q}. \end{aligned} \tag{5.30}$$

We can also write

$$f'(x) = \sum_{q=1}^{N} \frac{f(x)}{x - \phi_q} = \sum_{q=1}^{N} \frac{\sum_{k=0}^{N} a_k x^k}{x - \phi_q} \tag{5.31}$$

and, since ϕ_q is a zero of $\sum_{k=0}^{N} a_k x^k$, we can also get the expression

$$f'(x) = \sum_{q=1}^{N} \frac{\sum_{k=0}^{N} a_k (x^k - \phi_q^k)}{x - \phi_q}. \tag{5.32}$$

Moreover, the term with $k = 0$ is zero and we obtain

$$f'(x) = \sum_{q=1}^{N} \frac{\sum_{k=1}^{N} a_k (x^k - \phi_q^k)}{x - \phi_q}. \tag{5.33}$$

Now, we have to develop the expression $\frac{x^k - y^k}{x-y}$. To do this, we need to calculate the geometrical sum

$$\tilde{S} = 1 + x + x^2 + ... + x^{N-1}. \tag{5.34}$$

We multiply it by x and we get

$$x\tilde{S} = x + x^2 + ... + x^N, \tag{5.35}$$

or again

$$\tilde{S} - x\tilde{S} = 1 - x^N, \tag{5.36}$$

eventually giving

$$\tilde{S} = \frac{1 - x^N}{1 - x}. \tag{5.37}$$

We substitute x with $\frac{y}{x}$, getting

$$1 + \frac{y}{x} + \left(\frac{y}{x}\right)^2 + ... + \left(\frac{y}{x}\right)^{N-1} = \frac{1 - \left(\frac{y}{x}\right)^N}{1 - \frac{y}{x}}, \tag{5.38}$$

or, in another way of writing, we have

$$x^{N-1} + yx^{N-2} + y^2x^{N-3} + ... + y^{N-1} = \frac{x^N - y^N}{x - y}. \tag{5.39}$$

Therefore, we can finally write

$$\frac{x^k - y^k}{x - y} = \sum_{j=0}^{k-1} y^j x^{k-j-1}. \tag{5.40}$$

We use this result in Eq.(5.33) and we have

$$f'(x) = \sum_{q=1}^{N} \sum_{k=1}^{N} a_k \sum_{j=0}^{k-1} \phi_q^j x^{k-j-1}. \tag{5.41}$$

Now, since $P_j = \sum_{q=1}^{N} \phi_q^j$, we finally get

$$f'(x) = \sum_{k=1}^{N} \sum_{j=0}^{k-1} a_k P_j x^{k-j-1}. \tag{5.42}$$

And, if we let $i = k - j - 1$ becoming $j = k - i - 1$, we can find the dual expression

$$f'(x) = \sum_{k=1}^{N} \sum_{i=0}^{k-1} a_k P_{k-i-1} x^i, \tag{5.43}$$

that we can compare with

$$f'(x) = \sum_{i=0}^{N} a_i i x^{i-1} = \sum_{i=1}^{N} i a_i x^{i-1}. \tag{5.44}$$

To do this comparison, we observe that

$$\sum_{k=1}^{N} \sum_{i=0}^{k-1} = \sum_{i=0}^{N-1} \sum_{k=i+1}^{N} . \tag{5.45}$$

The previous observation allows us to draw a comparison between

$$f'(x) = \sum_{k=1}^{N} \sum_{i=0}^{k-1} a_k P_{k-i-1} x^i = \sum_{i=0}^{N-1} \sum_{k=i+1}^{N} a_k P_{k-i-1} x^i \tag{5.46}$$

and

$$f'(x) = \sum_{i=1}^{N} i a_i x^{i-1} = \sum_{i=0}^{N-1} (i+1) a_{i+1} x^i. \tag{5.47}$$

We find

$$(i+1) a_{i+1} = \sum_{k=i+1}^{N} a_k P_{k-i-1} \ \ \forall i = 0, 1, ..., N-1. \tag{5.48}$$

For $i = N - 1$, we have a trivial identity since we get $N a_N = a_N P_0$, hence $P_0 = N$. So, we can write

$$(i+1) a_{i+1} = N a_{i+1} + \sum_{k=i+2}^{N} a_k P_{k-i-1} \ \ \forall i = 0, 1, ..., N-2, \tag{5.49}$$

and we finally have $(N-1)$ equations, given by

$$-(N-i-1) a_{i+1} = \sum_{k=i+2}^{N} a_k P_{k-i-1} \ \ \forall i = 0, 1, ..., N-2. \tag{5.50}$$

We can express it in a more explicit way

$$\left\{ \begin{array}{lllll} -(N-1)a_1 & = a_2 P_1 & +a_3 P_2 & + \ ... & +a_N P_{N-1}, \\ -(N-2)a_2 & = a_3 P_1 & +a_4 P_2 & + \ ... & +a_N P_{N-2}, \\ ... & & & & \\ -3a_{N-3} & = a_{N-2} P_1 & +a_{N-1} P_2 & +a_N P_3, & \\ -2a_{N-2} & = a_{N-1} P_1 & +a_N P_2, & & \\ -a_{N-1} & = a_N P_1. & & & \end{array} \right. \tag{5.51}$$

The previous system of equations represents the link between $(P_1, ..., P_{N-1})$ and $\left(\frac{a_1}{a_N}, ..., \frac{a_{N-1}}{a_N}\right) = \left((-1)^{N-1}S_{N-1}, ..., S_2, -S_1\right)$. We can also add the equation Eq. (5.28) for $m = 0$, namely

$$-Na_0 = a_1P_1 + a_2P_2 + ... + a_NP_N. \tag{5.52}$$

With this equation, the system links $(P_1, ..., P_N)$ with $\left(\frac{a_0}{a_N}, ..., \frac{a_{N-1}}{a_N}\right) = \left((-1)^N S_N, ..., -S_1\right)$. Introducing the equation for $m = 0$ allows us to write the complete system as

$$\begin{cases} -Na_0 & -a_1P_1 & -a_2P_2 & - \ ... & -a_{N-1}P_{N-1} & = a_NP_N, \\ -(N-1)a_1 & -a_2P_1 & -a_3P_2 & - \ ... & -a_{N-1}P_{N-2} & = a_NP_{N-1}, \\ -(N-2)a_2 & -a_3P_1 & -a_4P_2 & - \ ... & -a_{N-1}P_{N-3} & = a_NP_{N-2}, \\ ... & & & & & \\ -3a_{N-3} & -a_{N-2}P_1 & -a_{N-1}P_2 & = a_NP_3, & & \\ -2a_{N-2} & -a_{N-1}P_1 & = a_NP_2, & & & \\ -a_{N-1} & = a_NP_1. & & & & \end{cases} \tag{5.53}$$

We can also write it in a matrix form

$$\begin{bmatrix} N & P_1 & P_2 & P_3 & ... & P_{N-3} & P_{N-2} & P_{N-1} \\ 0 & N-1 & P_1 & P_2 & ... & P_{N-4} & P_{N-3} & P_{N-2} \\ 0 & 0 & N-2 & P_1 & ... & P_{N-5} & P_{N-4} & P_{N-3} \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & 0 & 0 & ... & 3 & P_1 & P_2 \\ 0 & 0 & 0 & 0 & ... & 0 & 2 & P_1 \\ 0 & 0 & 0 & 0 & ... & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{a_0}{a_N} \\ \frac{a_1}{a_N} \\ \frac{a_2}{a_N} \\ ... \\ \frac{a_{N-3}}{a_N} \\ \frac{a_{N-2}}{a_N} \\ \frac{a_{N-1}}{a_N} \end{bmatrix} = - \begin{bmatrix} P_N \\ P_{N-1} \\ P_{N-2} \\ ... \\ P_3 \\ P_2 \\ P_1 \end{bmatrix}. \tag{5.54}$$

We apply the Kramer theory to the subsystem composed of the last $(N - i)$ equations and we get

$$\frac{a_i}{a_N} = -\frac{\det \begin{bmatrix} P_{N-i} & P_1 & ... & P_{N-i-3} & P_{N-i-2} & P_{N-i-1} \\ P_{N-i-1} & N-i-1 & ... & P_{N-i-4} & P_{N-i-3} & P_{N-i-2} \\ ... & ... & ... & ... & ... & ... \\ P_3 & 0 & ... & 3 & P_1 & P_2 \\ P_2 & 0 & ... & 0 & 2 & P_1 \\ P_1 & 0 & ... & 0 & 0 & 1 \end{bmatrix}}{\det \begin{bmatrix} N-i & P_1 & ... & P_{N-i-3} & P_{N-i-2} & P_{N-i-1} \\ 0 & N-i-1 & ... & P_{N-i-4} & P_{N-i-3} & P_{N-i-2} \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & 3 & P_1 & P_2 \\ 0 & 0 & ... & 0 & 2 & P_1 \\ 0 & 0 & ... & 0 & 0 & 1 \end{bmatrix}}. \tag{5.55}$$

In the first determinant, we move the first column to the last position, getting

$$\frac{a_i}{a_N} = \frac{(-1)^{N-i-1}(-1)}{(N-i)!}\det\begin{bmatrix} P_1 & P_2 & ... & P_{N-i-2} & P_{N-i-1} & P_{N-i} \\ N-i-1 & P_1 & ... & P_{N-i-3} & P_{N-i-2} & P_{N-i-1} \\ 0 & N-i-2 & ... & P_{N-i-4} & P_{N-i-3} & P_{N-i-2} \\ ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & P_1 & P_2 & P_3 \\ 0 & 0 & ... & 2 & P_1 & P_2 \\ 0 & 0 & ... & 0 & 1 & P_1 \end{bmatrix} \quad (5.56)$$

Finally, rotations of 180°do not alter the value of a determinant and therefore

$$\frac{a_i}{a_N} = \frac{(-1)^{N-i}}{(N-i)!}\det\begin{bmatrix} P_1 & 1 & 0 & 0 & ... & 0 & 0 \\ P_2 & P_1 & 2 & 0 & ... & 0 & 0 \\ P_3 & P_2 & P_1 & 3 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ P_{N-i-1} & P_{N-i-2} & P_{N-i-3} & P_{N-i-4} & ... & P_1 & N-i-1 \\ P_{N-i} & P_{N-i-1} & P_{N-i-2} & P_{N-i-3} & ... & P_2 & P_1 \end{bmatrix}. \quad (5.57)$$

We remember that $S_i = (-1)^i \frac{a_{N-i}}{a_N}$ or, equivalently, $S_{N-i} = (-1)^{N-i} \frac{a_i}{a_N}$, and we obtain the final form of the elementary symmetric polynomials

$$S_{N-i} = \frac{1}{(N-i)!}\det\begin{bmatrix} P_1 & 1 & 0 & 0 & ... & 0 & 0 \\ P_2 & P_1 & 2 & 0 & ... & 0 & 0 \\ P_3 & P_2 & P_1 & 3 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ P_{N-i-1} & P_{N-i-2} & P_{N-i-3} & P_{N-i-4} & ... & P_1 & N-i-1 \\ P_{N-i} & P_{N-i-1} & P_{N-i-2} & P_{N-i-3} & ... & P_2 & P_1 \end{bmatrix}, \quad (5.58)$$

or again

$$S_j = \frac{1}{j!}\det\begin{bmatrix} P_1 & 1 & 0 & 0 & ... & 0 & 0 \\ P_2 & P_1 & 2 & 0 & ... & 0 & 0 \\ P_3 & P_2 & P_1 & 3 & ... & 0 & 0 \\ ... & ... & ... & ... & ... & ... & ... \\ P_{j-1} & P_{j-2} & P_{j-3} & P_{j-4} & ... & P_1 & j-1 \\ P_j & P_{j-1} & P_{j-2} & P_{j-3} & ... & P_2 & P_1 \end{bmatrix}. \quad (5.59)$$

This is the closed form relation between the elementary symmetric polynomials and the power sums. These expressions can be named *determinant form of the Newton-Girard formulae* [184–187]. With this result, we can now use the determinant form of the terms S_j to develop the expressions of Z_H and $\frac{\partial Z_H}{\partial g_j}$, as indicated previously. We start with $Z_H(x)$ and we can write

$$\begin{aligned} Z_H(x) &= \int_{-\infty}^{+\infty} e^{-\frac{N\omega^2 k_B T}{2k_{\text{eff}}}} e^{i\omega x} \prod_{i=1}^{N}(a+\phi_i b)\mathrm{d}\omega \\ &= \int_{-\infty}^{+\infty} e^{-\frac{N\omega^2 k_B T}{2k_{\text{eff}}}} e^{i\omega x} \sum_{k=0}^{N} a^{N-k} b^k S_k \mathrm{d}\omega, \end{aligned} \quad (5.60)$$

with $S_0 = 1$. So, we firstly need the calculation of

$$\begin{aligned} A &= \int_{-\infty}^{+\infty} e^{-\frac{N\omega^2 k_B T}{2k_{\rm eff}}} e^{i\omega x} \left[\cos(\omega\ell)\right]^{N-k} \left[\cos(\chi\omega\ell)\right]^k \mathrm{d}\omega \\ &= \int_{-\infty}^{+\infty} e^{-\frac{N\omega^2 k_B T}{2k_{\rm eff}}} e^{i\omega x} \left(\frac{e^{i\omega\ell}+e^{-i\omega\ell}}{2}\right)^{N-k} \left(\frac{e^{i\chi\omega\ell}+e^{-i\chi\omega\ell}}{2}\right)^k \mathrm{d}\omega \\ &= \sum_{s=0}^{N-k}\sum_{q=0}^{k} \frac{1}{2^N}\binom{N-k}{s}\binom{k}{q} \int_{-\infty}^{+\infty} e^{i\omega[x+\ell s-\ell(N-k-s)+\chi\ell q-\chi\ell(k-q)]} \\ &\quad \times e^{-\frac{N\omega^2 k_B T}{2k_{\rm eff}}} \mathrm{d}\omega. \end{aligned} \tag{5.61}$$

To solve this integral, we use the well-known relation

$$\int_{-\infty}^{+\infty} e^{-\alpha x^2} e^{i\beta x} \mathrm{d}x = \sqrt{\frac{\pi}{\alpha}} e^{-\frac{\beta^2}{4\alpha}}, \tag{5.62}$$

so the equation becomes

$$A = \sum_{s=0}^{N-k}\sum_{q=0}^{k} \frac{1}{2^N}\binom{N-k}{s}\binom{k}{q} \sqrt{\frac{2\pi k_{\rm eff}}{N k_B T}} e^{-\frac{k_{\rm eff}[x+\ell(2s-N+k+2\chi q-\chi k)]^2}{2N k_B T}}. \tag{5.63}$$

Finally, the first result is

$$Z_H(x) = \frac{1}{2^N}\sqrt{\frac{2\pi k_{\rm eff}}{N k_B T}} \sum_{k=0}^{N}\sum_{s=0}^{N-k}\sum_{q=0}^{k} \binom{N-k}{s}\binom{k}{q} e^{-\frac{k_{\rm eff}[x+\ell(2s-N+k+2\chi q-\chi k)]^2}{2N k_B T}} S_k, \tag{5.64}$$

where the S_k coefficients are given by the determinants in Eq.(5.59). Concerning the calculation of $\frac{\partial Z_H(x)}{\partial g_j}$, we need to evaluate as in Eq.(5.19)

$$\begin{aligned} \frac{\partial Z_H(x)}{\partial g_j} &= \left(-\frac{i}{k_B T}\right) \int_{-\infty}^{+\infty} e^{-\frac{N\omega^2 k_B T}{2k_{\rm eff}}} e^{i\omega x} \prod_{k=1, k\neq j}^{N} \left[\cos(\omega\ell) + \phi_k \cos(\omega\chi\ell)\right] \\ &\quad \times \left[\sin(\omega\ell) + \phi_j \chi \sin(\omega\chi\ell)\right] \mathrm{d}\omega. \end{aligned} \tag{5.65}$$

We can define

$$S_k^{(j)} = \frac{1}{k!} \sum_{j_a \neq j_b \forall a\neq b, j_c \neq j \forall j} \phi_{j1} \times ... \times \phi_{jk}, \tag{5.66}$$

where the quantities $S_k^{(j)}$ $(k = 1, ..., N-1)$ are defined similarly to the quantities S_k $(k = 1, ..., N)$ but are based on the set containing all $\phi_1, ..., \phi_N$ except ϕ_j. They can

be simply calculated with the same technique based on the determinants, as shown in Eq.(5.59). Hence we can write, after developing the power on N

$$\begin{aligned}
\frac{\partial Z_H}{\partial g_j} &= \left(-\frac{i}{k_B T}\right) \int_{-\infty}^{+\infty} e^{-\frac{N\omega^2 k_B T}{2k_{\rm eff}}} e^{i\omega x} \sum_{k=0}^{N-1} \left[\cos(\omega\ell)\right]^{N-1-k} \left[\cos(\chi\omega\ell)\right]^k S_k^{(j)} \\
&\quad \times \left[\sin(\omega\ell) + \phi_j \chi \sin(\chi\omega\ell)\right] \mathrm{d}\omega \\
&= \frac{1}{k_B T} \int_{-\infty}^{+\infty} e^{-\frac{N\omega^2 k_B T}{2k_{\rm eff}}} e^{i\omega x} \sum_{k=0}^{N-1} \left[\cos(\omega\ell)\right]^{N-1-k} \left[\cos(\chi\omega\ell)\right]^k S_k^{(j)} \\
&\quad \times \left(\frac{e^{-i\omega\ell} - e^{i\omega\ell}}{2}\right) \mathrm{d}\omega + \frac{1}{k_B T} \int_{-\infty}^{+\infty} e^{-\frac{N\omega^2 k_B T}{2k_{\rm eff}}} e^{i\omega x} \sum_{k=0}^{N-1} \left[\cos(\omega\ell)\right]^{N-1-k} \\
&\quad \times \left[\cos(\chi\omega\ell)\right]^k S_k^{(j)} \chi\phi_j \left(\frac{e^{-i\omega\chi\ell} - e^{i\omega\chi\ell}}{2}\right) \mathrm{d}\omega. \qquad (5.67)
\end{aligned}$$

We separately calculate the two integrals, called β_k and γ_k. We begin the calculation of β_k by developing the powers

$$\begin{aligned}
\beta_k &= \int_{-\infty}^{+\infty} e^{-\frac{N\omega^2 k_B T}{2k_{\rm eff}}} e^{i\omega x} \left[\cos(\omega\ell)\right]^{N-1-k} \left[\cos(\chi\omega\ell)\right]^k \left(\frac{e^{-i\omega\ell} - e^{i\omega\ell}}{2}\right) \mathrm{d}\omega \\
&= \int_{-\infty}^{+\infty} e^{-\frac{N\omega^2 k_B T}{2k_{\rm eff}}} e^{i\omega x} \sum_{s=0}^{N-k-1} \binom{N-k-1}{s} e^{i\omega\ell s} e^{-i\omega\ell(N-1-k-s)} \frac{1}{2^{N-1-k}} \\
&\quad \times \sum_{q=0}^{k} \binom{k}{q} e^{i\chi\omega\ell q} e^{-i\chi\omega\ell(k-q)} \frac{1}{2^k} \left(\frac{e^{-i\omega\ell} - e^{i\omega\ell}}{2}\right) \mathrm{d}\omega \\
&= \frac{1}{2^N} \sum_{s=0}^{N-k-1} \sum_{q=0}^{k} \binom{N-k-1}{s} \binom{k}{q} \left[\int_{-\infty}^{+\infty} e^{-\frac{N\omega^2 k_B T}{2k_{\rm eff}}} e^{i\omega[x+\ell(2s-N+k+2\chi q-\chi k)]} \mathrm{d}\omega \right. \\
&\quad \left. - \int_{-\infty}^{+\infty} e^{-\frac{N\omega^2 k_B T}{2k_{\rm eff}}} e^{i\omega[x+\ell(2s+2-N+k-2\chi q-\chi k)]} \mathrm{d}\omega \right]. \qquad (5.68)
\end{aligned}$$

Finally, we get

$$\begin{aligned}
\beta_k &= \frac{1}{2^N} \sqrt{\frac{2\pi k_{\rm eff}}{N k_B T}} \sum_{s=0}^{N-k-1} \sum_{q=0}^{k} \binom{N-k-1}{s} \binom{k}{q} \left[e^{-\frac{k_{\rm eff}[x+\ell(2s-N+k+2\chi q-\chi k)]^2}{2N k_B T}} \right. \\
&\quad \left. - e^{-\frac{k_{\rm eff}[x+\ell(2s-N+2+k+2\chi q-\chi k)]^2}{2N k_B T}} \right]. \qquad (5.69)
\end{aligned}$$

Now, we calculate the second integral, called γ_k

$$\gamma_k = \int_{-\infty}^{+\infty} e^{-\frac{N\omega^2 k_B T}{2k_{\rm eff}}} e^{i\omega x} \left[\cos(\omega\ell)\right]^{N-1-k} \left[\cos(\chi\omega\ell)\right]^k \left(\frac{e^{-i\omega\chi\ell} - e^{i\omega\chi\ell}}{2}\right) \mathrm{d}\omega. \qquad (5.70)$$

We do the same calculation developed for β_k and we obtain

$$\gamma_k = \frac{1}{2^N}\sqrt{\frac{2\pi k_{\text{eff}}}{Nk_BT}}\sum_{s=0}^{N-k-1}\sum_{q=0}^{k}\binom{N-k-1}{s}\binom{k}{q}\left[e^{-\frac{k_{\text{eff}}[x+\ell(2s-N+1-\chi+k+2\chi q-\chi k)]^2}{2Nk_BT}}\right.$$
$$\left.-e^{-\frac{k_{\text{eff}}[x+\ell(2s-N+1+\chi+k+2\chi q-\chi k)]^2}{2Nk_BT}}\right]. \tag{5.71}$$

We get the final form of $\frac{\partial Z_H(x)}{\partial g_j}$

$$\frac{\partial Z_H(x)}{\partial g_j} = \frac{1}{k_BT}\sum_{k=0}^{N-1}\left[\left(\beta_k+\chi\gamma_k\phi_j\right)S_k^{(j)}\right]. \tag{5.72}$$

We can finally deduce the expression of the force-extension response from Eqs.(5.16) and (5.64) as

$$\langle f\rangle = \frac{\sum_{k=0}^{N}\sum_{s=0}^{N-k}\sum_{q=0}^{k}\binom{N-k}{s}\binom{k}{q}e^{-\frac{k_{\text{eff}}\varphi^2}{2Nk_BT}}k_{\text{eff}}\varphi S_k}{\sum_{k=0}^{N}\sum_{s=0}^{N-k}\sum_{q=0}^{k}\binom{N-k}{s}\binom{k}{q}e^{-\frac{k_{\text{eff}}\varphi^2}{2Nk_BT}}S_k}, \tag{5.73}$$

with $\varphi = x+\ell(2s-N+k+2\chi q-\chi k)$ and the coefficients S_k are given in Eq.(5.59). Similarly, the expression for the average spin variable $\langle y_j\rangle$ can be obtained from Eqs.(5.17) and (5.72) as

$$\langle y_j\rangle = \frac{\sum_{k=0}^{N-1}\left(\beta_k S_k^{(j)}+\gamma_k\phi_j\chi S_k^{(j)}\right)}{Z_H}, \tag{5.74}$$

where the coefficients $S_k^{(j)}$ are given in Eq.(5.66). Although the quenched disorder strongly complicates the analysis of the system within the Helmholtz ensemble, the application of the determinant expression in Eq.(5.59) or Eq.(5.66) (determining the elementary symmetric polynomials S_k or $S_k^{(j)}$ related to the Boltzmann factors with arbitrary energy jumps) allows for an analytic treatment of the problem. In particular, we are able to obtain the partition function in Eq.(5.64) (along with its derivative in Eq.(5.72)) in closed form and to derive expressions for the macroscopic observables.

A first simple application of these results can be found in Figs.5.3.b and 5.3.c, where we plot $\langle f\rangle$ and $\langle y_i\rangle$ versus the prescribed extension x for a homogeneous chain under isometric condition (red curves). The behaviour of $\langle f\rangle$ and $\langle y_i\rangle$ can be interpreted through a sequential unfolding of the units. This response is in good qualitative agreement with several force spectroscopy measurements performed on proteins and other macromolecules [135, 142–145]. We also plotted $\partial\langle y_i\rangle/\partial x$ versus x in Fig.5.3.d. This quantity satisfies two

crucial properties: firstly, $\int_0^{+\infty}(\partial\langle y_i\rangle/\partial x)\mathrm{d}x = \langle y_i\rangle(+\infty) - \langle y_i\rangle(0) = \chi$ (which means that it can be normalised); secondly, $\langle y_i\rangle$ is always non-decreasing, leading to a non-negative function $\partial\langle y_i\rangle/\partial x$. Hence, it follows that $\partial\langle y_i\rangle/\partial x$ can be used as a quantity approximately measuring the probability density of the position x at which a transition occurs between the states of the i-th unit. We remark that it is not a rigorous statement but a useful practical approach to quantify the statistics of the unfolding processes. Fig.5.3.d shows that at each transition occurrence, the switching probability is the same for all units of the homogeneous chain. Therefore, no unfolding pathway can be identified. We describe now the behaviour of a heterogeneous chain, as represented in Fig.5.3.e. While the force-extension curve in Fig.5.3.f is slightly modified with respect to Fig.5.3.b, we observe that the heterogeneity of the chain, which is a quenched disorder embedded in the system, is able to perform a symmetry breaking generating an unfolding pathway. Indeed, the numbered curves of the average spin variables in Fig.5.3.g are able to precisely identify what unit is unfolded at each transition, indicating the actual sequence of unfolding processes. Interestingly enough, the red curve in the same panel shows the average value of these numbered curves and is similar to the Helmhotz response of the homogeneous case. This is true since we used the same parameter χ for all the units of the chain. It is important to remark that, for an heterogeneous parameter χ, we can have a different behaviour between $\langle y_i\rangle$ of the homogeneous chain and $\frac{1}{N}\sum_{i=1}^{N}\langle\, y_i\,\rangle$ of the heterogeneous chain. Importantly, it follows that through the spin variables, we can now analyse the unfolding pathway generated by the heterogeneity. This is further confirmed by the plots of $\partial\langle y_i\rangle/\partial x$ in Fig.5.3.h, where each curve is characterised by one pronounced peak corresponding to the actual switching, and other smaller peaks measuring the uncertainty in the pathway identification (see next Section for details). Moreover, the numerical results show that the unfolding pathway corresponds to the ascending order of the values ΔE_i. This implies that the equilibrium statistical mechanics of this system implements a statistical sorting procedure. This result is independent of the heterogeneity of the elastic constants k_i since $\langle f\rangle$ and $\langle y_j\rangle$ depend only on the effective stiffness k_{eff}.

5.4 Unfolding pathway identifiability

While the described identification of the unfolding pathway may seem a simple and expected result, it is important to observe that: (i) the model elaborated mathematically explains how the bistable chain can implement the above introduced statistical sorting procedure, and the same methodology can be also applied to more realistic situations with additional heterogeneous geometrical and/or physical parameters; (ii) the knowledge of

the average spin variables given in Eq.(5.74) is also useful to quantitatively evaluate the identifiability of the most probable unfolding pathway, i.e. of the most likely observed unfolding sequence in a given experiment. This concept measures to what extend we are able to identify the most probable unfolding pathway, which represents the sequence of unfolding process observed the largest number of times if we conduct several identical experiments. For the sake of brevity, we develop this concept only by considering the Helmholtz ensemble. If we look at the density-like curve $\partial\langle y_i\rangle/\partial x$ for a given unit (see Fig.5.3.h), the identifiability can be defined as the relative difference between the largest peak and the second largest peak. Indeed, this difference measures the capability to properly identify the transition of that unit with respect to the other ones. Consequently, if we consider the i-th unit, we can define

$$\mathcal{I}_i = \frac{\mathcal{F}_m\left\{\frac{\partial\langle y_i\rangle}{\partial x}\right\} - \mathcal{S}_m\left\{\frac{\partial\langle y_i\rangle}{\partial x}\right\}}{\mathcal{F}_m\left\{\frac{\partial\langle y_i\rangle}{\partial x}\right\}}, \tag{5.75}$$

where $\mathcal{F}_m$ and $\mathcal{S}_m$ are operators extracting the largest peak and the second largest peak, respectively, of a given function. The identifiability of the whole unfolding process can be therefore defined by the average value of these quantities over the N units

$$\mathcal{I} = \frac{1}{N}\sum_{i=1}^{N}\mathcal{I}_i. \tag{5.76}$$

We first apply this concept to a chain composed of $N = 5$ units with uniformly distributed energy jumps $\Delta E_i/(k_BT) \in \{\alpha, \alpha \pm \varepsilon, \alpha \pm 2\varepsilon\}$, where the parameter ε measures their dispersion. The resulting identifiability $\mathcal{I}$ can be found in Fig.5.4, where it is plotted versus ε and parametrized by the temperature T. We note that $\mathcal{I} = 0$ for the homogeneous case with $\varepsilon = 0$, and $\mathcal{I} \to 1$ for increasing value of ε, being the limiting value $\mathcal{I} = 1$ related to the pathway identification without uncertainty. We also observe that increasing values of the temperature reduce the identifiability, for a constant ε. This is coherent with the idea that the thermal fluctuations are able to reduce the knowledge on the configurational state of the system.

A plateau at around $\epsilon = 0.5$ can be observed and is due to a particular phenomenon observed for small values of ϵ. We explain this point as follows. We consider only the case with $T = 300K$ and $N = 5$. Then, we assume $\Delta E_i = \{30, 30+\epsilon, 30-\epsilon, 30+2\epsilon, 30-2\epsilon\}$ $\times k_BT$ as in Fig.5.4 (the other parameters used being exactly as quoted in the caption of Fig.5.4). We take a small value of ϵ, namely $\epsilon = 0.4$. In Fig.5.5, we can now observe the

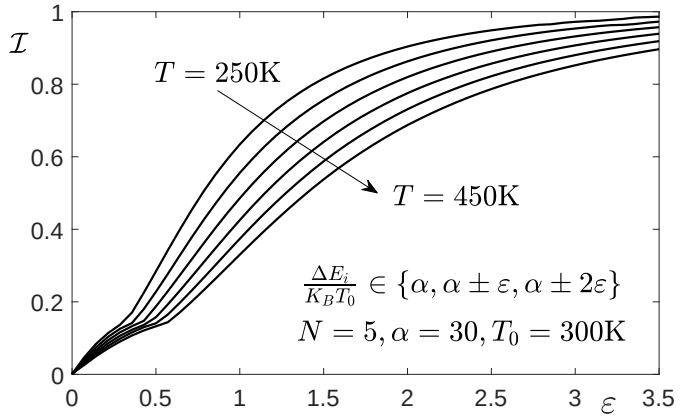

Figure 5.4 – Identifiability $\mathcal{I}$ defined as in Eq.(5.76) as function of the distribution of the energy jumps ΔE_i (described by the parameter ε) and the temperature T. We adopted the following parameters: $N = 5$, $\ell = 0.5$nm, $T_0 = 300$K, $\chi = 7$, $k_i = 0.04$N/m $\forall i$, $\frac{\Delta E_i}{k_B T_0} \in \{\alpha, \alpha \pm \varepsilon, \alpha \pm 2\varepsilon\}$ with $\alpha = 30$ and ε variable [4].

behaviour of the average spin variables. However, if we look at the derivatives of the spin variables, we obtain the plots shown in Fig.5.6.

Therefore, we see that the green curve show the greatest peak at the first transition, as expected, whereas the yellow curve, unexpectedly, does not show the greatest peak at the second transition but at the first one. This is at the origin of the specific behaviour shown by the curves in Fig.5.4 for small ϵ. We can observe that when, increasing ϵ, the peaks assume the "normal" configuration, and we are at the right extremity of the small plateau region. The conclusion is that, probably, the identification of the unfolding sequence on the base of the largest peaks in the derivatives of the spins is not the optimal approach for very small values of ϵ, or, equivalently, for not sufficiently spaced energy jumps. Consequently, also the definition of the identifiability on the base of the peaks intensity is not well adapted for very low values of ϵ. Nevertheless, this phenomenon is observed only for differences between the energy jumps lower than 1-2%.

A more complex example deals with a non-linear distribution of energy jumps of the units. More specifically, we can consider an exponential distribution described by $\Delta E_i = \mathcal{P} e^{i\varphi} + \mathcal{Q}$, where $\mathcal{P}$ and $\mathcal{Q}$ are fixed by imposing the values ΔE_1 and ΔE_N, φ is a free parameter defining the non-linearity of the energy jumps, and i is the index enumerating the units. It means that the distribution of energy jumps can be written as

$$\Delta E_i = \frac{\Delta E_1 \left(e^{N\varphi} - e^{i\varphi}\right) + \Delta E_N \left(e^{i\varphi} - e^{\varphi}\right)}{e^{N\varphi} - e^{\varphi}}. \tag{5.77}$$

As a particular case, we observe that if $\varphi \to 0$, we obtain the linear distribution of energy

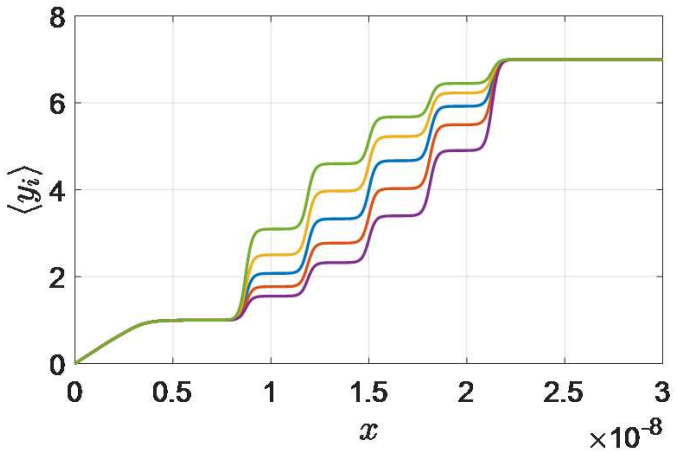

Figure 5.5 – Average spin variable versus x. We see that, although the parameter ϵ assume a small value, the spin responses are sufficiently separated to distinguish the unfolding process. In particular, the green curve corresponds to the lower energy jump, the yellow curve to the following energy jump, and so forth.

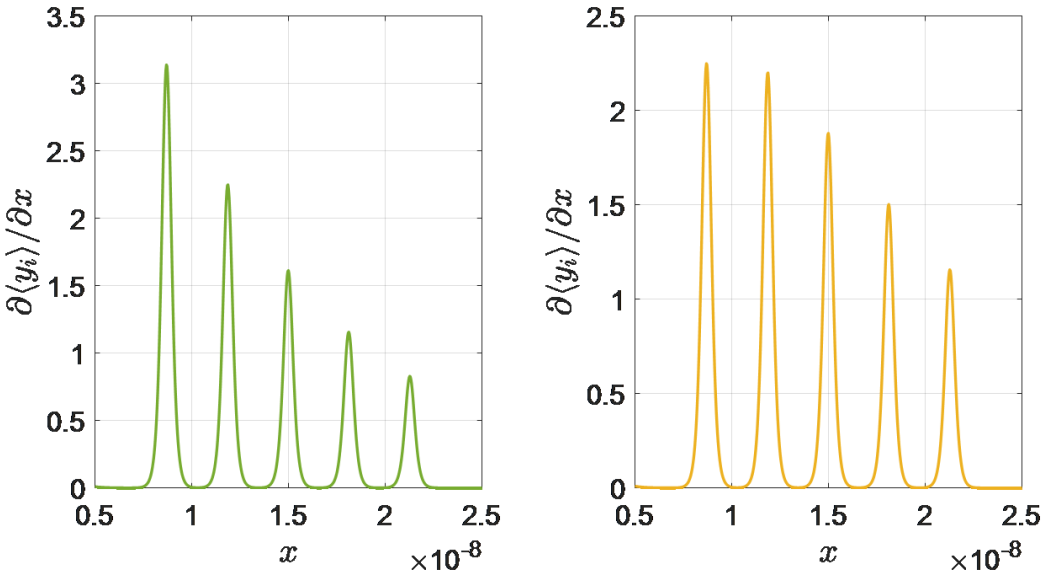

Figure 5.6 – Derivatives of Z_H versus x. The green curve corresponds to the spin variable of the unit with the lower energy jump and the yellow curve to the spin variable of the unit with the following energy jump (the colours correspond to the previous Figure with the averages of y_i).

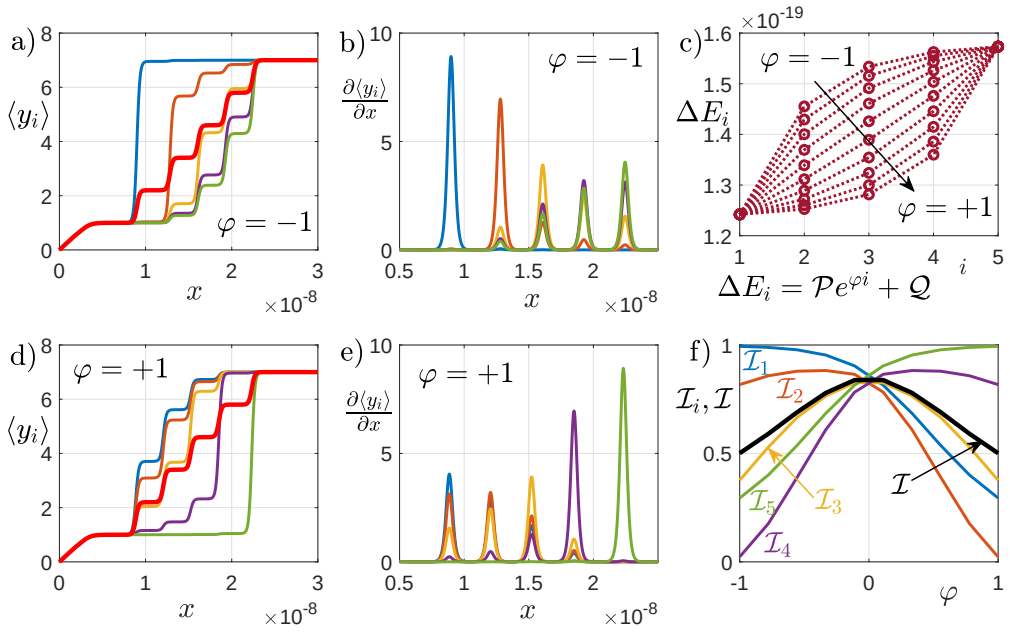

Figure 5.7 – Unfolding pathway in heterogeneous chains with non-uniform distribution of energy jumps. We adopted the distribution $\Delta E_i = \mathcal{P} e^{i\varphi} + \mathcal{Q}$ with $\mathcal{P} = (\Delta E_N - \Delta E_1)/(e^{N\varphi} - e^{\varphi})$ and $\mathcal{Q} = (\Delta E_1 e^{N\varphi} - \Delta E_N e^{\varphi})/(e^{N\varphi} - e^{\varphi})$ (see Eq.(5.77)). Panels a) and d): average spin variables $\langle y_i \rangle$ for $\varphi = \pm 1$. Panels b) and e): derivatives $\partial \langle y_i \rangle / \partial x$ for $\varphi = \pm 1$. Panel c): distributions of the energy jumps for $-1 \leq \varphi \leq +1$ and with $\Delta E_1 = 30 k_B T$ and $\Delta E_N = \Delta E_5 = 38 k_B T$. Panel f): identifiabilities $\mathcal{I}_i$ and $\mathcal{I}$ versus φ. We adopted the following parameters: $N = 5$, $\ell = 0.5$nm, $T = 300$K, $\chi = 7$, and $k_i = 0.04$N/m $\forall i$ [4].

jumps $\Delta E_i = \Delta E_1 + \Delta E_N (i-1)/(N-1)$, already considered in the previous analysis. The results based on these assumptions can be found in Fig.5.7. In panels a) and b), one can find the average spin variables and their derivatives, respectively, for the case with $\varphi = -1$. In this case, the energy jumps are given by a non-linear concave distribution, as plotted in panel c). Similarly, in panels d) and e), we show the results for $\varphi = 1$, corresponding to a non-linear convex distribution, which is shown in panel c), as well. Finally, in panel f), the identifiabilities are represented for each unit and for the whole chain. It is important to observe that the non-linearity of the energy jumps is reflected in the spread or dispersion of the values $\mathcal{I}_i$, especially for φ approaching ± 1. Indeed, it is more difficult to identify the unfolding pathway of units with similar energy jumps (see, e.g., $\mathcal{I}_4$ and $\mathcal{I}_5$ for $\varphi = -1$ or $\mathcal{I}_1$ and $\mathcal{I}_2$ for $\varphi = +1$) than the unfolding pathway of units with largely spaced energy jumps (see, e.g., $\mathcal{I}_1$ and $\mathcal{I}_2$ for $\varphi = -1$ or $\mathcal{I}_4$ and $\mathcal{I}_5$ for $\varphi = +1$). We also note that in the limiting case with $\varphi \to 0$, the distribution becomes linear, as previously anticipated, and all the quantities $\mathcal{I}_i$ assume approximately the same

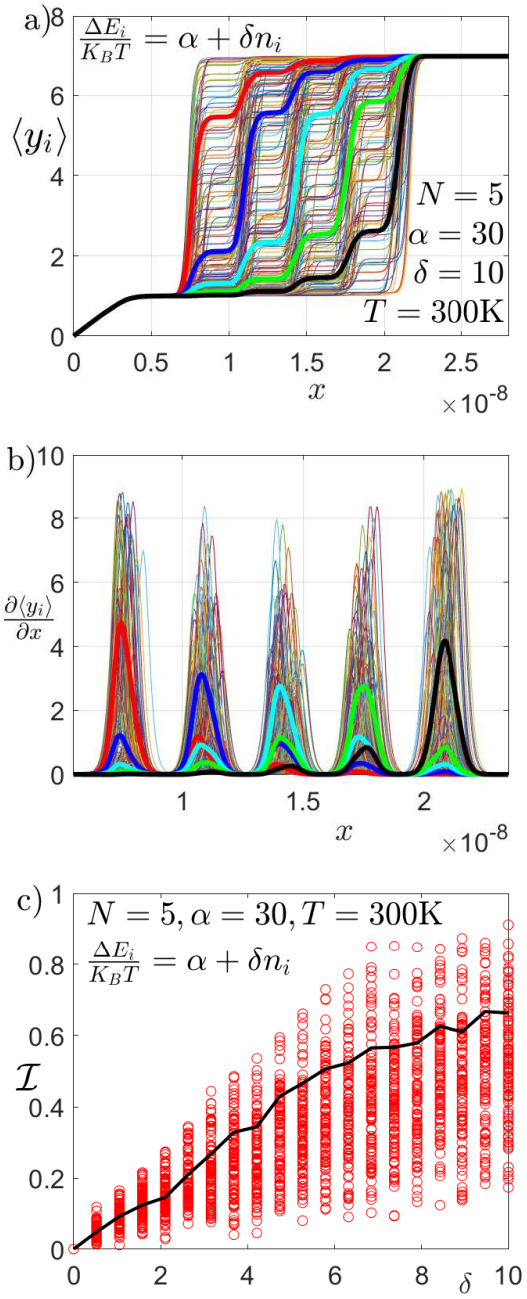

Figure 5.8 – Unfolding pathway in randomly heterogeneous chains. Panel a): average spin variables $\langle y_i \rangle$. Panel b): derivatives $\partial \langle y_i \rangle / \partial x$. In both panels, 50 Monte-Carlo realisations have been plotted (thin lines) together with their average values (thick lines) for a system with $\Delta E_i/(k_B T) = \alpha + \delta n_i$, where n_i are independent and uniform random variables in the interval $(-1, 1)$. Panel c): identifiability $\mathcal{I}$ versus δ. We plotted 100 Monte-Carlo realisations (red circles) and the average value (black solid line). We adopted the following parameters: $N = 5$, $\ell = 0.5\text{nm}$, $T = 300\text{K}$, $\chi = 7$, $k_i = 0.04\text{N/m}$ $\forall i$, $\alpha = 30$ and $\delta = 10$ (in panels a) and b) only) [4].

value. Coherently, the black curve in panel f), representing the average value $\mathcal{I}$, shows a maximum for $\varphi \to 0$, proving that the largest identifiability is achieved for linearly spaced or distributed energy jumps.

The applicability of the identifiability concept is twofold. From one side, it allows a better understanding of the unfolding pathways of proteins and other bio-macromolecules, typically measured through force spectroscopy techniques. In particular, it can explain the statistical modifications or variability of the unfolding pathway, which is sometimes depending on several experimental conditions [188–190]. On the other side, the identifiability concept may be useful to improve the design of heterogeneous micro- and nano-systems based on bi-and multi-stability, where folding and unfolding sequences represent the response of the system and should be therefore stable to temperature variations and to other structural or external parameters [177]. The important point for the applications is that the identifiability can be calculated for any set of parameters describing the chain, and allows therefore a parametric analysis of the stability of the most probable unfolding pathway, observed when we conduct several identical experiments. Even if we limited the analysis of the identifiability only to the Helmholtz case, we can compare the two ensembles as follows. From panels h) and l) of Fig.5.3, it is not difficult to realise that for a fixed chain, the identifiability in the two ensembles is not the same and the Helmholtz ensemble is capable to maximise this quantity. Indeed, the variance of the peaks in panel l) is much larger than the variance of the peaks in panel h). This result can be easily interpreted in terms of the differences between the Helmholtz and the Gibbs ensemble. As a matter of fact, also for a simple homogeneous chain, in the Gibbs ensemble, we have a synchronised unfolding of the units whereas in the Helmholtz ensemble, we have a sequential unfolding. Therefore, the Helmholtz unfolding, being sequential, is more adapted to separate the unfolding events and to eventually improve the identifiability.

To conclude this discussion, we determine $\mathcal{I}$ for a bistable chain with random energy jumps between the folded and unfolded states (under isometric conditions). We assume that the energy jumps are given by $\Delta E_i/(k_B T) = \alpha + \delta n_i$, where α and δ are fixed parameters while n_i are independent and uniform random variables in the interval $(-1, 1)$. Of course, the parameter δ measures the stochastic spread of the energy jumps distribution. We generate a given number of chains (Monte-Carlo realisations) using the previous rule to assign the energy jumps of the units. Then, we sort the units in each chain in such a way as to have the energy jumps in ascending order. This is simply useful to easily compare the spin variables of different chains, corresponding to units of the same ordered position. For any chain, we are able to calculate the average spin variables $\langle y_i \rangle$

and their derivatives $\partial\langle y_i\rangle/\partial x$ with respect to the increasing extension x of the chain. Consequently, for each chain, we can determine the corresponding identifiability through Eq.(5.76). These calculations can be repeated for all generated Monte-Carlo realisations and the means values can be eventually evaluated (sampling Monte-Carlo approach). The results can be found in Fig.5.8. In panel a) and b), we show the average spin variables $\langle y_i\rangle$ and their derivatives $\partial\langle y_i\rangle/\partial x$, respectively, for 50 Monte-Carlo realisations of the system. The results of the single realisations (thin solid lines) are plotted together with their sample mean values (thick solid lines). We can see that, with the adopted parameters, the average spin variables allow the unfolding pathway identification also with random energy jumps. While panels a) and b) of Fig.5.8 concern a fixed value of δ, we can perform a more complete analysis where δ is variable over a given range. Hence, being the standard deviation of the energy jumps $\sigma_{\Delta E_i} = k_B T\delta/\sqrt{3}$ proportional to δ, we plot in panel c) the identifiability $\mathcal{I}$ versus δ. We used 100 Monte-Carlo realisations for each value of δ (20 values of δ in the range $[0, 10]$), represented by the red circles in panel c), and we calculated the sample mean values of the identifiability, represented by the solid black line. We repeated the whole protocol several times and we proved that the solid black line, describing the behaviour of $\mathcal{I}$, remains stable within an maximal error bar of around ± 0.1. It means that the Monte-Carlo sample with 100 chains is large enough to give acceptable results. We observe that, for random chains, relatively large values of $\delta < \alpha$ are necessary to obtain a good average identifiability of the unfolding path.

5.5 Conclusion

In this Chapter, we considered the statistical and mechanical behaviour of chains of heterogeneous bistable units. As this system is useful to represent situations of practical interest, we thoroughly analysed its behaviour in both isotensional and isometric conditions. The most important achievement concerns the exact calculation of the partition function (and related quantities) within the Helmholtz ensemble (isometric condition). In this case, the heterogeneity represents a quenched disorder, whose analysis is a difficult task of the statistical mechanics. We obtained the closed form expression of the Helmholtz partition function thanks to the Laplace-Fourier relation between Z_H and Z_G [20,124] and using the determinant form of the so-called Newton-Girard formulae [184–187]. This original approach represents the core of our analysis. We observed that for a homogeneous chain, no unfolded pathway can be identified since all units have the same switching probability at each transition occurrence. On the other hand, the heterogeneity breaks this symmetry and we can identify an unfolded pathway, which is described by the ascending

order of the energy jumps between folded and unfolded states of each unit. It means that the system implements a statistical sorting procedure when we simply prescribe an increasing distance between first and last units. Since this process has a statistical character, we can define the concept of identifiability, which measures the capability to identify the most probable unfolding pathway. This concept has been applied to deterministic (linear and exponential) and random distributions of energy jumps. The results of this work can be applied to the better interpretation of the force spectroscopy measurements of biological macromolecules [188–190] and to the accurate design of micro- and nano-systems based on bistable chains with specific properties [177].

Chapter 6

Pulling speed dependence of the force-extension response of bistable chains

6.1 Introduction

The development of single-molecule force spectroscopy techniques has widely promoted the study of the mechanical and kinetic properties of several biological macromolecules, such as proteins [6, 25, 33, 134, 191], nucleic acids (DNA [70, 71, 73, 182] and RNA [192, 193]) and polysaccharides [24, 194, 195]. The static and dynamic responses of such macromolecules are crucial for assessing the equilibrium and out-of-equilibrium thermodynamics of small systems. The typical single-molecule experiment is conducted by tethering the first end of the macromolecule and by moving its second end by means of the force spectroscopy device, such as described in Chapter 1. For practical reasons, these experiments are typically conducted by imposing a given pulling speed to the second end of the chain. This point has crucial effects on the force-extension relation, which shows a strong speed-dependent behaviour. Therefore, it is necessary to consider the influence of the pulling speed on the force-extension response of a macromolecule, as the force peaks increase with increasing speed [34, 196] (see Fig.6.1). Moreover, the recent increasing of the pulling speeds used in force spectroscopy experiments with high-speed AFM [34,35,37,38] requires the consideration of pulling speeds higher than in the existing models. In Fig.6.1.a, it can be observed that the force peaks are nearly proportional to the logarithm of the used pulling speed. The exploration of the response for different pulling speeds allows the reconstruction of the potential energy, as seen in Fig.6.1.b, and as discussed below. The histograms in Fig.6.1.c clearly show the increasing of the force peaks when the pulling speed increases. Similarly, in Fig.6.2, a high-speed force spectroscopy

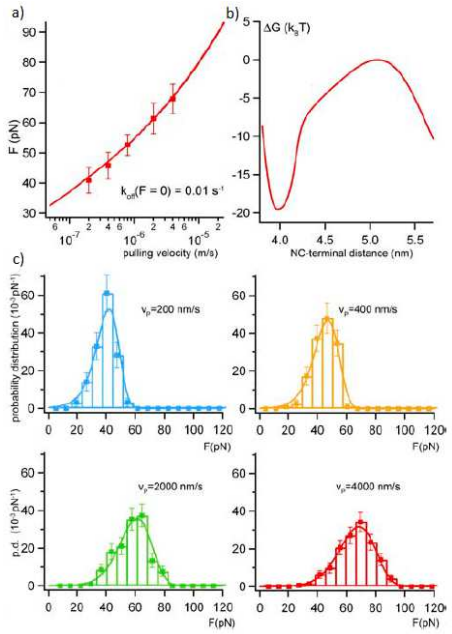

Figure 6.1 – a) Pulling velocity data obtained through the unfolding of the Immunoglobulin-like domain 4 (ddFLN4) compared with the theoretical curve given by the Kramers' theory. b) The reconstructed energy landscape shows detailed curvature along the unfolding/folding pathway. c) The characteristic behaviour of the experimental unfolding force distributions with increasing pulling velocity is reproduced well using the Kramers' model [196].

device is presented, which allows to attain high pulling speed of around $10^3 - 10^4 \mu m/s$. The increase in peak force values is also pronounced.

On the one hand, for very low pulling speeds, the system is not far from the thermodynamic equilibrium and its response can be studied by means of the classical canonical distribution of the statistical mechanics (see Chapters 2, 3, 4 and 5). On the other hand, it is important to consider the case of a system stretched with a finite pulling speed. Its typical values in standard AFM force spectroscopy experiments range between 10^{-2} and $10^2 \mu$m/s [196–201], and are even larger for high-speed AFM, as shown in Fig.6.2. Along with experiments, molecular dynamics simulations have been performed to study the speed-dependent unfolding of bistable macromolecules [39]. However, this approach is limited to pulling speeds larger than $10^3 \mu$m/s and, therefore, it is not possible to draw a comparison with standard AFM experiments [40]. Indeed, for lower values of the pulling speed, the total computational time to observe the unfolding is too extended and the simulation is not realisable with the available computational resources [35–40]. The advantage of the molecular dynamics is that the bistable character of the units directly comes from the molecular architecture of the system. Nevertheless, to reduce the com-

plexity of the molecular dynamics simulations, a coarse grained model can be adopted and implemented through a kinetic Monte Carlo method [128, 202]. In this case, the bistability of each unit is simply described by a two-state potential energy characterised by the rates of the folding and unfolding processes. Hence, the reduced complexity of the model allows to consider a larger range of pulling speeds. Nonetheless, in this approach, only the folding and unfolding rates are considered to describe the bistable characters of the units.

By considering the Kramers theory [16], the Bell approximation [159], and the Evans-Ritchie more elaborated formalism [204], we can assert that the folding and unfolding rates are directly related to the energy barrier ΔB and the energy jump ΔE between the two states, as well as the mechanical action applied to the system. This point allows to affirm that we can estimate energetic features of the macromolecule (e.g., ΔB and/or ΔE) from the force-extension response measured at different pulling speeds (energy landscape reconstruction). In fact, the observed unfolding forces are influenced by the dynamic interplay between the transition rates and the applied pulling speed. Of course, much theoretical work has been done to apply the Kramers, Bell and Evans-Ritchie theories for decoding the force spectroscopy data and yielding the energetic parameters that govern the biomolecular processes [205]. This approach has generated a powerful class of approximated results, which are of simple application in many different regimes and configurations [149,150,203,206–209]. In the present Chapter, we propose an approach based on the numerical implementation of the Langevin methodology. The classical Langevin equations are equipped with the full description of an arbitrary potential energy mimicking the bistable units. This method can be used to analyse force spectroscopy experimental data in order to estimate biological features of the macromolecules, such as ΔE, ΔB or other biochemical parameters. It allows for considering the whole range of pulling speeds used in real experiments, including those of standard and high-speed AFM experiments. Finally, the Langevin approach will be used to elucidate the role of the device in the single-molecule measurements and to interpret experimental data concerning filamin and titin proteins.

6.2 Out-of-equilibrium statistical mechanics through the Langevin approach

To introduce the out-of-equilibrium statistical mechanics of the pulling process, we use the Langevin approach. It means that a friction term and a noise term are added to

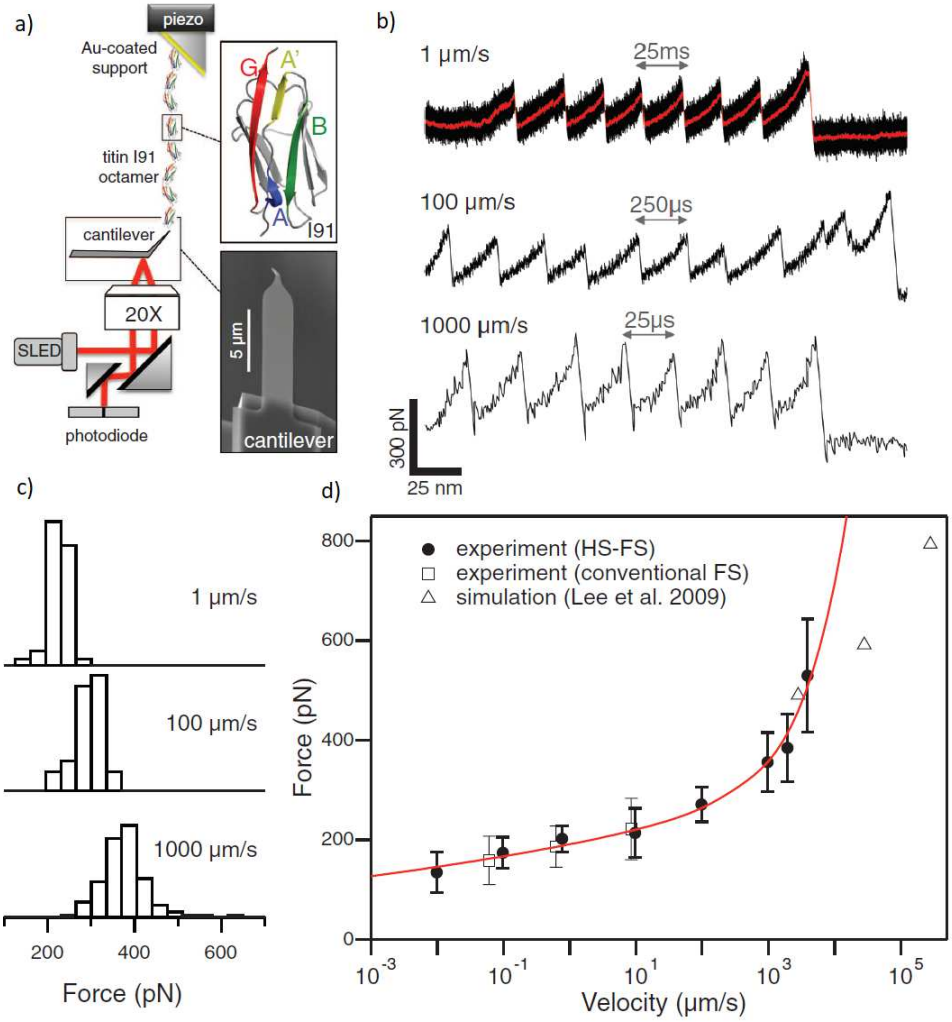

Figure 6.2 – a) High-speed force spectroscopy (HS-FS) setup. An objective focuses the beam of the superluminescent light-emitting diode (SLED) and collects the light reflected by the cantilever, finally detected by a segmented photodiode. Titin I91 concatemers of eight domains are immobilized on a tilted gold-coated surface via C-terminal cysteines. They are pulled by their N-terminal histidine tag with a nickel-coated tip at the end of a short cantilever. Tilting the sample surface further reduces hydrodynamic forces. Top inset shows a titin I91 domain (PDB ID: 1TIT) with relevant β-strands coloured in blue (A), yellow (A'), green (B), and red (G). Bottom inset shows a scanning electron micrograph of a short cantilever. b) Force-extension curves acquired at three different retraction velocities: 1, 100, and 1000 μm/s. The 1 μm/s curve is moving average–filtered (red trace, 65-μs time window). Times to unfold a single I91 domain are indicated by arrows. c) Unfolding force histograms of the 1, 100, and 1000 μm/s retraction velocity experiments. d) Average unfolding forces versus retraction velocity obtained using HS-FS, conventional FS (error bars denote standard deviation), and steered molecular dynamics simulations (data from Ref. [40]). Solid red line is the fit to the entire dynamic range of HS-FS with the full microscopic model (from Ref. [203]) [34].

the equation of motion of each degree of freedom of the system. This stochastic evolution is coherent with the asymptotic behaviour of the system, which is represented by the classical canonical distribution of the statistical mechanics [210, 211]. For a particle in motion within the three-dimensional space, the exact Newton dynamic equation is

$$m\frac{\mathrm{d}^2\vec{r}}{\mathrm{d}t^2} = -\frac{\partial V}{\partial \vec{r}} - m\beta\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} + \sqrt{Dm}\vec{n}, \tag{6.1}$$

where V is the potential energy describing the force field applied to the particle, β is the friction coefficient (per unit mass, the so-called collision frequency) and D is the diffusion coefficient (per unit mass) given by the Einstein relation $D = k_B T\beta$ [212, 213]. Moreover, in Eq.(6.1), $\vec{r}$ is the position vector of the particle, m is its mass and $\vec{n}$ is a random process. As usual, we assume the following hypotheses on $\vec{n}$: $\vec{n}(t) \in \mathbb{R}^3$ is a Gaussian stochastic process, $E\{\vec{n}(t)\} = 0$, and $E\{n_i(t_1)n_j(t_2)\} = 2\delta_{ij}\delta(t_1 - t_2)$ (here E means "expected value", δ_{ij} is the Kronecker delta, and $\delta(.)$ is the Dirac delta function). It is well known that these properties are sufficient to obtain a correct thermodynamic behaviour for the system [210–213]. In typical biological environments, the coefficient $m\beta$ describing the friction is quite large. It means that we can neglect the inertial term within the Newton equation of motion. This simplification is sometimes named Smoluchowski approximation and it is valid in the so-called *overdamped regime* [212, 213]. Hence, for large values of β, we can write

$$m\frac{\mathrm{d}\vec{r}}{\mathrm{d}t} = -\frac{1}{\beta}\frac{\partial V}{\partial \vec{r}} + \frac{\sqrt{Dm}}{\beta}\vec{n}, \tag{6.2}$$

being the inertial term negligible.

This approach can be adopted to develop the model for a one-dimensional chain of N bistable units, unfolded through a force spectroscopy device (see Fig.6.3.a). In this case, we can write the system of stochastic differential equations as

$$m\frac{\mathrm{d}x_i}{\mathrm{d}t} = -\frac{1}{\beta}U'(x_i - x_{i-1}) + \frac{1}{\beta}U'(x_{i+1} - x_i) + \sqrt{\frac{k_B T m}{\beta}}n_i(t), \tag{6.3}$$

where $i = 1, ..., N-1$ if we consider N identical units described by the potential energy U. This approach has been recently adopted to study the effect of the pulling velocity on the unfolding pathway of heterogeneous chains [152, 214], and it is used here to investigate the effect of the pulling velocity on the unfolding forces. These equations must be solved with the boundary conditions $x_0(t) = 0$ and $x_N(t) = v_0 t + N\ell$, where ℓ is the length of one folded unit (before the application of the traction). For $t = 0$, we have $x_N(0) = N\ell$, which means that all the units are folded at the beginning of the process. In addition,

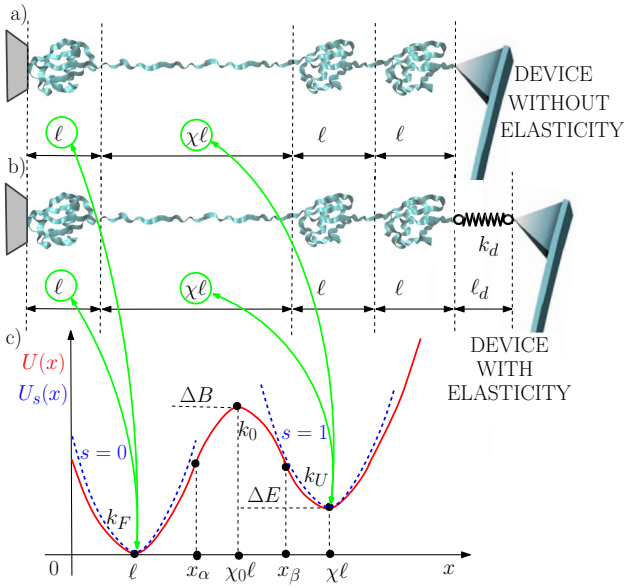

Figure 6.3 – Scheme of the single-molecule force spectroscopy pulling experiments. Panel a): chain of bistable units stretched by a device without intrinsic elasticity and with $x_N(t) = v_0 t + N\ell$. Panel b): chain of bistable units stretched by a device with intrinsic stiffness k_d, equilibrium length ℓ_d, and $x_{N+1}(t) = v_0 t + N\ell + \ell_d$. Panel c): bistable potential energy $U(x)$ of a chain unit and its approximation $U_s(x)$ based on the spin variable s ($s = 0$ for the folded state and $s = 1$ for the unfolded state) [4].

the latter condition corresponds to a device without intrinsic elasticity, able to perfectly prescribe the trajectory of the last element of the chain. On the one hand, when v_0 is very small, the system can be studied with the equilibrium statistical mechanics. This is discussed in previous Chapters. On the other hand, when v_0 assumes arbitrary values, the problem can be approached by the direct integration of Eq.(6.3). To do this, we have to specify the energy profile $U(x)$, which represents the bistable character of the system units. As a particularly simple example, we will analyse a real bistable system described by the following potential energy (see Fig.6.3.c)

$$U(x) = \begin{cases} \frac{1}{2}k_F(x-\ell)^2 & \text{if } x \le x_\alpha, \\ \Delta B - \frac{1}{2}k_0(x-\chi_0\ell)^2 & \text{if } x_\alpha < x \le x_\beta, \\ \Delta E + \frac{1}{2}k_U(x-\chi\ell)^2 & \text{if } x > x_\beta. \end{cases} \tag{6.4}$$

The parameter k_F represents the elastic constant of the folded state, k_0 the (reverse) elastic constant of the concave barrier between the wells, and k_U the elastic constant of the unfolded state. Here, $k_F > 0$, $k_0 > 0$, $k_U > 0$, $1 < \chi_0 < \chi$, ΔB represents the energy barrier between the wells, and ΔE measures the energy jump between the two states. While ℓ is the position of the first well, the parameters χ_0 and χ are geometrical factors fixing the position of the barrier and of the second well, respectively. The points

x_α and x_β represent the boundaries separating the three different regions (the first well, the spinodal region, and the second well). If the elastic constants of the two wells assume the same value (see Fig.6.3.c) with $k_U = k_F$), the conditions of continuity and derivability of U at points x_α and x_β give the relations

$$\chi_0 = \frac{\sqrt{\Delta B - \Delta E} + \chi\sqrt{\Delta B}}{\sqrt{\Delta B - \Delta E} + \sqrt{\Delta B}}, \tag{6.5}$$

$$\frac{1}{k_0} = \frac{\ell^2(\chi-1)^2}{2\left(\sqrt{\Delta B - \Delta E} + \sqrt{\Delta B}\right)^2} - \frac{1}{k_F}, \tag{6.6}$$

$$x_\alpha = \frac{k_F + k_0\chi_0}{k_F + k_0}\ell, \tag{6.7}$$

$$x_\beta = \frac{k_F\chi + k_0\chi_0}{k_F + k_0}\ell. \tag{6.8}$$

Therefore, once fixed ΔE, ΔB, k_F, χ and ℓ, we can easily find x_α, x_β, k_0 and χ_0. In order to further investigate the behaviour of the bistable chain, we will also analyse the case with a different elastic constant in the two wells (see Fig.6.3.c) with $k_U \neq k_F$). The conditions of continuity and derivability of U at points x_α and x_β give now

$$(\Delta B - \Delta E)(\chi_0 - 1)^2 - \Delta B(\chi - \chi_0)^2 = \frac{2\Delta B(\Delta B - \Delta E)(k_U - k_F)}{k_U k_F \ell^2}, \tag{6.9}$$

which is a second degree equation for χ_0, and

$$k_0 = \frac{2\Delta B k_F}{k_F\ell^2(\chi_0 - 1)^2 - 2\Delta B}, \tag{6.10}$$

$$x_\alpha = \frac{k_F + k_0\chi_0}{k_F + k_0}\ell, \tag{6.11}$$

$$x_\beta = \frac{k_U\chi + k_0\chi_0}{k_U + k_0}\ell. \tag{6.12}$$

Therefore, once fixed ΔE, ΔB, k_F, k_U, χ and ℓ, we can easily find x_α, x_β, k_0 and χ_0. The behaviour of the whole system (with either $k_U = k_F$ or $k_U \neq k_F$) can be studied through the numerical solution of the Langevin equation. The results of this analysis will be described in Sections 6.3 and 6.4.

Thus, the numerical solution of Eq.(6.3), with $U(x)$ defined in Eq.(6.4), can be implemented by means of the following procedure. First of all, we integrate Eq.(6.3) over an arbitrary interval $(t, t + \Delta t)$. This operation delivers for $i = 1, ..., N - 1$

$$m\left[x_i(t+\Delta t) - x_i(t)\right] \simeq -\frac{1}{\beta}U'\left[x_i(t) - x_{i-1}(t)\right]\Delta t + \frac{1}{\beta}U'\left[x_{i+1}(t) - x_i(t)\right]\Delta t + \sqrt{\frac{k_B T m}{\beta}}\int_t^{t+\Delta t} n_i(t)\mathrm{d}t, \tag{6.13}$$

where $\Delta w_i = \int_t^{t+\Delta t} n_i(t)\mathrm{d}t$ are independent increments of the Wiener process [212, 213]. It follows that the quantities Δw_i are random Gaussian variables and fulfil the properties $E\{\Delta w_i\} = 0$ and $E\{\Delta w_i \Delta w_j\} = 2\Delta t \delta_{ij}$. We can therefore introduce the random variables $P_i = \frac{1}{\sqrt{2\Delta t}}\Delta w_i$, which are independent and normal Gaussian, thus described by the probability density $f(P_i) = \frac{1}{\sqrt{2\pi}} e^{-P_i^2/2}$. The motion equations can be therefore written as

$$\begin{aligned} x_i(t+\Delta t) - x_i(t) &\simeq -U'\left[x_i(t) - x_{i-1}(t)\right]\frac{\Delta t}{m\beta} + U'\left[x_{i+1}(t) - x_i(t)\right]\frac{\Delta t}{m\beta} \\ &\quad + \sqrt{2k_BT}\sqrt{\frac{\Delta t}{m\beta}}P_i. \end{aligned} \tag{6.14}$$

We can now observe that the time step Δt controls the convergence of the numerical scheme and the solution must be stable for decreasing values of Δt. If we define $\eta = \frac{\Delta t}{m\beta}$, we obtain the final version of the finite difference scheme of the Langevin equations for $i = 1, ..., N-1$

$$\begin{aligned} x_i(t+\Delta t) - x_i(t) &\simeq -U'\left[x_i(t) - x_{i-1}(t)\right]\eta + U'\left[x_{i+1}(t) - x_i(t)\right]\eta \\ &\quad + \sqrt{2k_BT}\sqrt{\eta}P_i. \end{aligned} \tag{6.15}$$

We draw the random numbers P_i [215], and we generate M realisations of the trajectory to finally determine their mean value [216].

The model so far introduced is represented by a chain of N bistable units with the last unit directly connected to the device (without the intrinsic elasticity) able to apply a traction defined by the uniform motion $x_N(t) = v_0 t + N\ell$ (see Fig.6.3.a). Actually, a force spectroscopy device is typically characterised by its intrinsic elasticity, which is rather important for the understanding of the system behaviour [217, 218]. Therefore, we also introduce a second model where the last unit of the chain is connected to a device with a given finite elasticity (see Fig.6.3.b). It means that Eq.(6.3) must be substituted by the following system of equations

$$m\frac{\mathrm{d}x_i}{\mathrm{d}t} = -\frac{1}{\beta}U'(x_i - x_{i-1}) + \frac{1}{\beta}U'(x_{i+1} - x_i) + \sqrt{\frac{k_BTm}{\beta}}n_i(t), \tag{6.16}$$

for $i = 1, ..., N-1$ and

$$m\frac{\mathrm{d}x_N}{\mathrm{d}t} = -\frac{1}{\beta}U'(x_N - x_{N-1}) + \frac{1}{\beta}k_d(x_{N+1} - x_N - \ell_d) + \sqrt{\frac{k_BTm}{\beta}}n_N(t), \tag{6.17}$$

for the device performing the force spectroscopy measurement. Here, k_d represents the intrinsic elasticity of the device, ℓ_d its intrinsic length, and the traction is described by $x_{N+1}(t) = v_0 t + N\ell + \ell_d$.

6.3 Analytical and numerical results

In this Section, we provide the analyses of four different cases dealing with a bistable chain (typically representing a protein) having the following properties: i) the same elastic constants k_F and k_U in the folded and unfolded states, and a device without intrinsic elasticity; ii) different elastic constants k_F and k_U, and a device without intrinsic elasticity; iii) the same elastic constants k_F and k_U, and a realistic device with k_d and ℓ_d; iv) different elastic constants k_F and k_U, and a realistic device with k_d and ℓ_d. The comparison of these four different cases is important to deduce some general features of the force-extension relation, useful to better interpret and decipher the single-molecule experimental results. In particular, this analysis will be able to shed light on the role of the force spectroscopy device in the measured force-extension curves. We underline that in this Section, the numerical calculations have been performed with reasonable parameters with respect to real macromolecules but we decided not to refer to specific biological systems, in order to show some general trends. However, we will show in Section 6.4 a direct comparison with experiments conducted on filamin and titin proteins.

6.3.1 Device without intrinsic elasticity

In Fig.6.4, we can find the results for the first case dealing with a chain with equal elastic constants for the two wells and unfolded through a device without intrinsic elasticity. We considered a macromolecule with $N = 4$ domains. In Fig.6.4.a, we show the energy profile of the bistable units, where we can observe the energy barrier $\Delta B = 18 k_B T$ and the energy jump $\Delta E = 10 k_B T$. Then, in Fig.6.4.b, we show the force-extension curves for different pulling velocities v_0. The dashed black curve corresponds to the theoretical result obtained with $v_0 = 0$. The other coloured curves represent the response of the chain with an increasing pulling velocity. Each force-extension curve is obtained by averaging the solution of the Langevin equation (see Eqs.(6.3) or (6.15)) over $M = 2000$ Monte Carlo realisations of the process. The deviation between each coloured curve and the dashed black curve measures how the process is far from the thermodynamic equilibrium. We remark that, while the case with $v_0 = 0$ is accessible with closed form mathematical expressions based on the canonical distribution of the statistical mechanics, the out-of-equilibrium behaviour can be only numerically explored by means of the Langevin formalism. The important point is that the force peaks, representing the sequential unfolding of the chain units (or protein domains), are more pronounced for larger values of the pulling velocity. Therefore, following the experimental protocol, it is interesting to analyse the behaviour

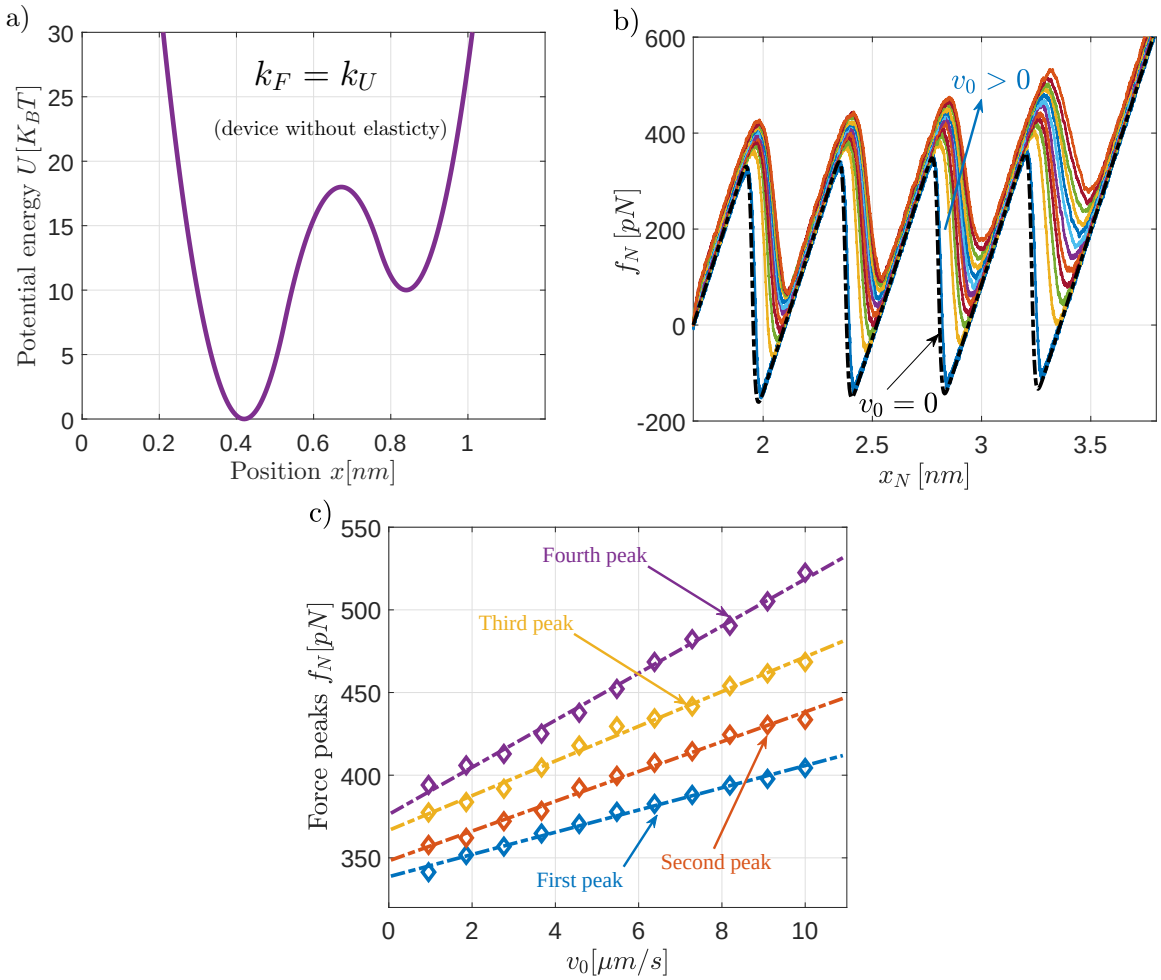

Figure 6.4 – Results for a chain with equal elastic constants for the two wells and unfolded through a device without intrinsic elasticity. Bistable potential energy of the chain units a), average force exerted on the N-th unit for different pulling velocities $v_0 \geq 0$ (average curves determined over M=2000 trajectories) b), and force peaks versus the applied pulling velocity (symbols: data; dashed lines: linear least squares approximations) c). In panels b) and c), we adopted different pulling velocities from 0.95 to 10 μm/s. The dotted black curve in panel b) represents the force-extension response at thermodynamic equilibrium. The results have been obtained with $\Delta B = 18k_BT$, $\Delta E = 10k_BT$, $T = 300$K, N=4, $\ell = 0.42$nm, $\chi = 2$, $k_F = k_U = 5.63$N/m, and η variable ranging from $\eta = 2.5 \times 10^{-3}$m/N with $\Delta t = 1.47 \times 10^{-8}$s for $v = 0.95$ μm/s, to $\eta = 2.36 \times 10^{-4}$m/N with $\Delta t = 1.4 \times 10^{-9}$s for $v = 10$ μm/s [4].

of these force peaks in terms of the applied pulling velocity. This result can be found in Fig.6.4.c, where the intensity of the four peaks (symbols) are plotted versus v_0.

The error bars of these results are of the same order of magnitude of the fluctuations observed in the curves of Fig.6.4.b. Then, they are not represented in Fig.6.4.c since

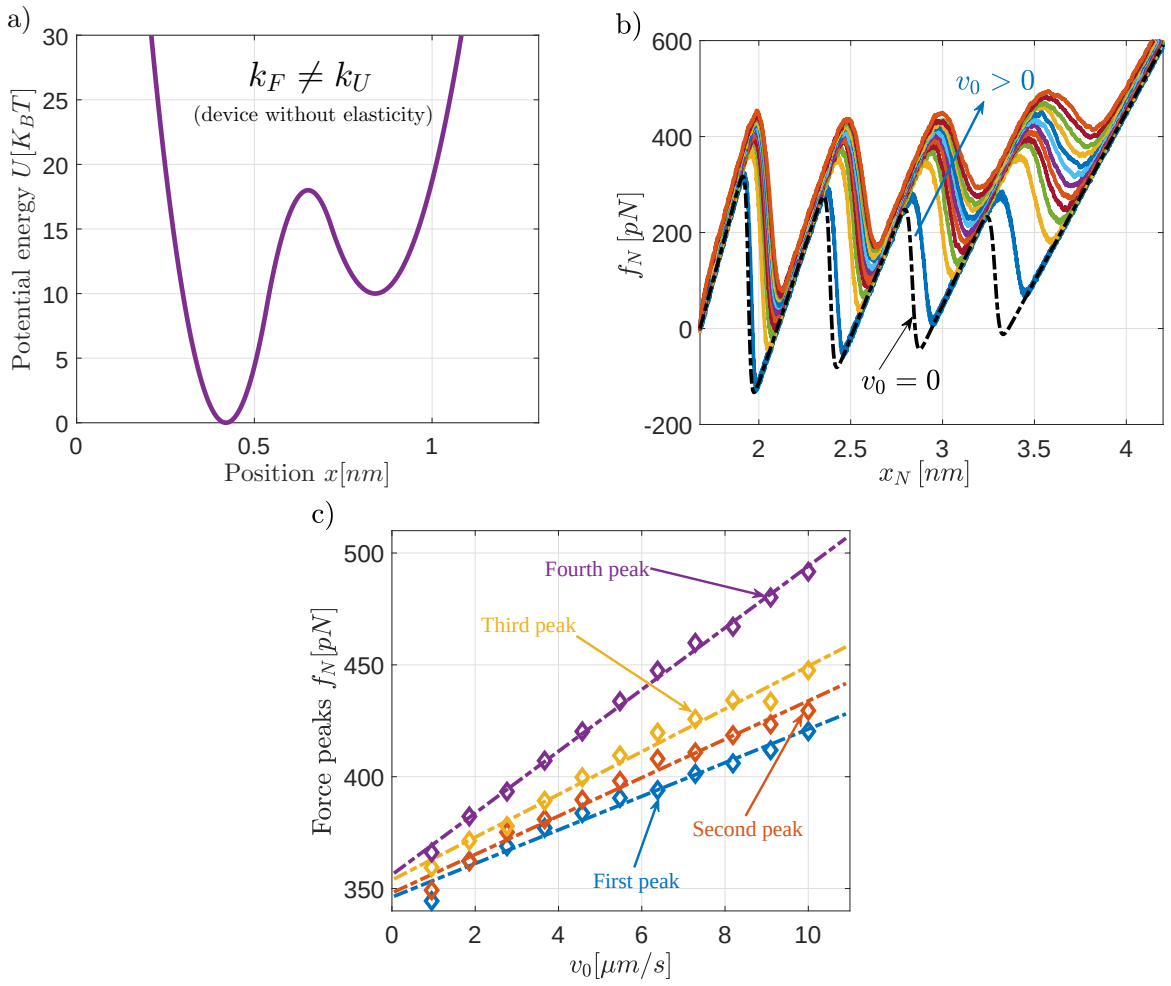

Figure 6.5 – Results for a chain with different elastic constants for the two wells and unfolded through a device without intrinsic elasticity. Bistable potential energy of the chain units a), average force exerted on the N-th unit for different pulling velocities $v_0 \geq 0$ (average curves determined over M=2000 trajectories) b), and force peaks versus the applied pulling velocity (symbols: data; lines: dashed linear least squares approximations) c). In panels b) and c), we adopted different pulling velocities from 0.95 to 10 μm/s. The dotted black curve in panel b) represents the force-extension response at thermodynamic equilibrium. The results have been obtained with $\Delta B = 18k_BT$, $\Delta E = 10k_BT$, $T = 300$K, N=4, $\ell = 0.42$nm, $\chi = 2$, $k_F = 5.63$N/m, $k_U = 2.81$N/m, and η variable ranging from $\eta = 2.5 \times 10^{-3}$m/N with $\Delta t = 1.47 \times 10^{-8}$s for $v = 0.95$ μm/s, to $\eta = 2.36 \times 10^{-4}$m/N with $\Delta t = 1.4 \times 10^{-9}$s for $v = 10$ μm/s [4].

they are smaller than the symbols used to represent the force peaks. This remains true for all results of Sections 6.3 and 6.4. The first important result is that the four force peaks are linearly increasing with the pulling velocity for a chain with $k_F = k_U$, unfolded through a device without intrinsic elasticity. This is confirmed by the linear least squares

approximations (dashed straight lines), shown in Fig.6.4.c. We remark that this behaviour does not correspond to the typical experimental observation, where the force peaks show a logarithmic trend with the pulling velocity. This trend, sometimes called Evans and Ritchie law, is based on the classical Bell theory [159], and it is valid with the extension speed varying over several orders of magnitude [135, 204, 219]. However, we remark that such a logarithmic trend can sometimes be inexact, especially for quite large pulling velocities [34–38]. In order to understand the specific features characterising the response observed in single-molecule force spectroscopy experiments, we therefore explore the cases with $k_F \neq k_U$ and/or with a real device.

In Fig.6.5, we can find the results for a chain with $k_F \neq k_U$, unfolded by means of a device without intrinsic elasticity. We considered $k_U = k_F/2$ and we assumed all the other parameters exactly as in the case shown in Fig.6.4. The results shown in Fig.6.5 are quite similar to the ones presented in Fig.6.4 and, in particular, the linear trend between force peaks and pulling velocity is confirmed also in this case with $k_F \neq k_U$. It means that the different elastic constants of the two energy wells describing the bistable units are not at the origin of the dynamic behaviour experimentally observed (with the force peaks proportional to the logarithm of the pulling velocity). Consequently, we consider in the following Section the cases with a real device characterised by an intrinsic elasticity and an intrinsic equilibrium length.

A final comment concerns the shape of the force-extension curves represented in Figs.6.4 and 6.5. In both cases, we observe that some lower peaks of the curves correspond to a negative force. This is a specific behaviour induced by the fact that we are studying one-dimensional systems. In fact, for one-dimensional systems, the idea of a protein chain as a random coil with a complex distribution of domains is degenerated to a simple alignment of units on one axis with a preferred total length given by $j\ell+(N-j)\chi\ell$ (with $j = 0, ..., N$). Since the total length is exactly imposed by the device within the Helmholtz ensemble, it is not difficult to imagine configurations with either a pushing force (negative) or a pulling one (positive), applied to the device. Of course, this apparently paradoxical behaviour disappears for two- or three-dimensional geometries, where the exploration of the phase space is more pertinent to polymer models [131]. Indeed, in these real multi-dimensional cases, the pulling (negative) forces induce the geometrical rearrangement of the chain and are not directly observed in the total force measured on the device. It is important to underline that also the theoretical results obtained with the spin variables for $v_0 \to 0$ exhibit the same behaviour confirming that it is related to the one-dimensional geometry of our system. In spite of this limitation of our approach, we underline that the model is

able to correctly evaluate the force versus pulling velocity relation since at the transition points the macromolecule is quite aligned along the traction direction.

6.3.2 Device with intrinsic elasticity

In Fig.6.6, we show the results for a chain with equal elastic constants for the two wells and unfolded through a realistic device with intrinsic elasticity. Typically, the elastic constant of the devices is much smaller than the elastic constant characterising the protein domains. For the sake of definiteness, in our case, we assume that $k_d = k_F/8 = k_U/8$. The comparison between $k_F = k_U$ and k_d can be observed in Fig.6.6.a, where the potential energy of the units and of the device are represented. In Fig.6.6.b, the force-extension response is shown with different values of the pulling velocity. We remark that in this figure the curves are represented by plotting the force f_N directly applied to the macromolecule versus the traction x_{N+1} prescribed by the device.

The dashed black curve in Fig.6.6.b represents the force-extension relation at thermodynamic equilibrium ($v_0 \to 0$) by means of the spin variables approach. The peaks of force shown in Fig.6.6.b represent the sequential unfolding of the units and are quantified in Figs.6.6.c and 6.6.d. Here, these peaks are represented in terms of the pulling speed in both linear and logarithmic scales, respectively.

From Fig.6.6.c, characterised by the linear scale of pulling velocity, we deduce that the linearity between peaks and velocity is not confirmed in the presence of a device with intrinsic elasticity. Indeed, by observing Fig.6.6.d, we conclude that we have in this case a linear dependence between force peaks and the logarithm of the pulling velocity, as observed in most of experiments. We also observe that this linear relation is not verified for large values of the pulling velocity. Coherently, deviations from the linearity between force peaks and the logarithm of the pulling velocity have been experimentally observed in high-speed single-molecule measurements [34–38]. The dashed straight lines in Fig.6.6.d represent the linear least squares approximations of the peaks data (for $v_0 < 2.5\mu$m/s) and confirm the linear relation between the force peaks and $\log_{10} v_0$, at least for small values of v_0. This result underlines the crucial role of the realistic device, with its specific stiffness, in determining the force-extension response and the force peaks observed during the molecule unfolding.

This scenario is also confirmed for the case with $k_F \neq k_U$, as reported in Fig.6.7. Finally, the model developed on the base of the Langevin equation is able to thoroughly take into consideration the bistable character of the units of a given macromolecule as well as the elasticity of the device used to generate the unfolding process. This point is

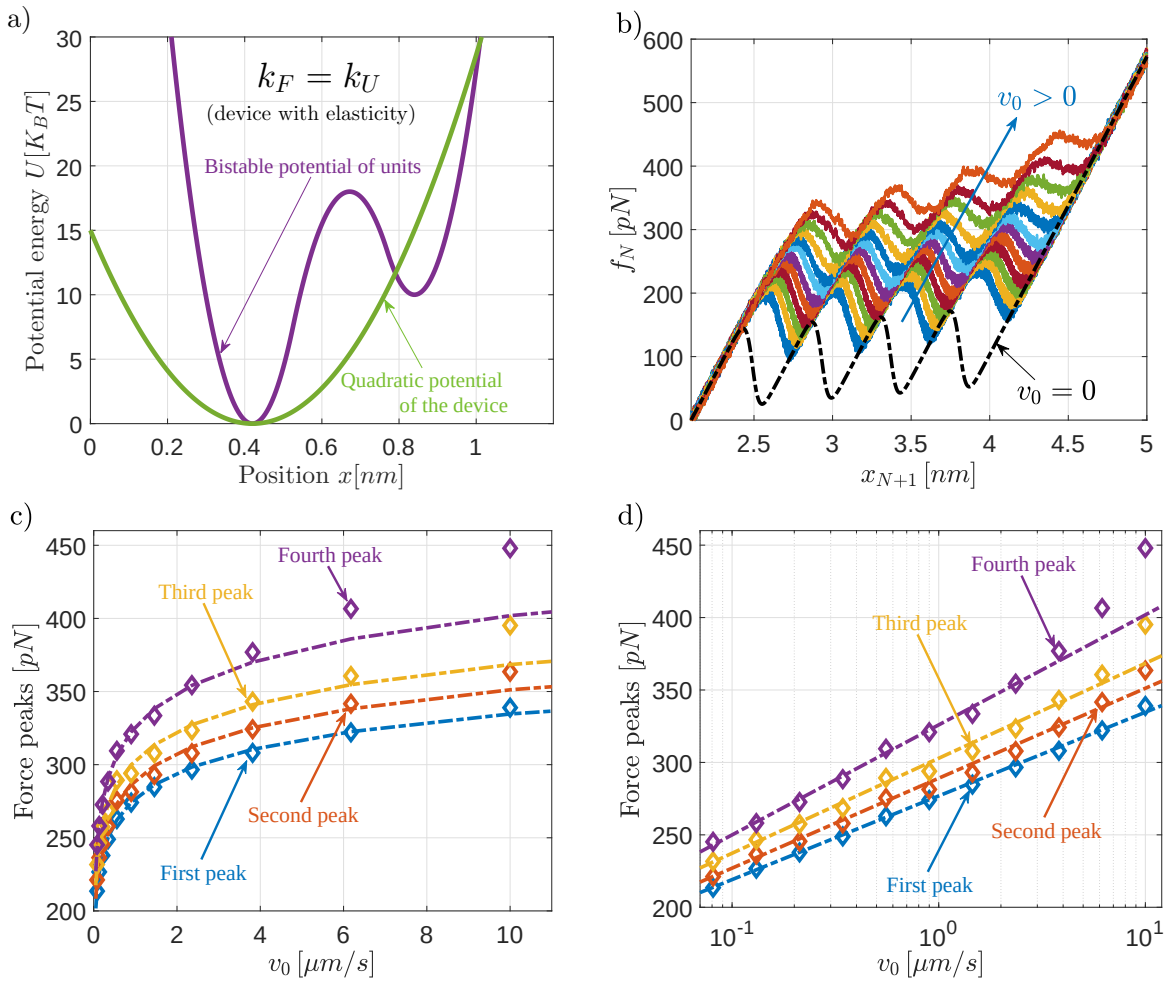

Figure 6.6 – Results for a chain with equal elastic constants for the two wells and unfolded through a device with intrinsic elasticity. Potential energy of units and device a), average force exerted on the N-th unit of the chain versus the position x_{N+1} of the device (average curves determined over M=2000 trajectories) b), force peaks versus the applied pulling velocity in linear scale c), and force peaks versus the applied pulling velocity in semi-log scale (symbols: data; dashed lines: linear least squares approximations) d). In panels b), c) and d), we adopted different pulling velocities from 0.081 to 10 μm/s. The dotted black curve in panel b) represents the force-extension response at thermodynamic equilibrium. The results have been obtained with $\Delta B = 18k_BT$, $\Delta E = 10k_BT$, $T = 300$K, N=4, $\ell = \ell_d = 0.42$nm, $\chi = 2$, $k_F = k_U = 5.63$N/m, $k_d = 0.704$N/m, and η variable ranging from $\eta = 3.2\times10^{-2}$m/N with $\Delta t = 1.9\times10^{-7}$s for $v = 0.081$ μm/s, to $\eta = 2.6\times10^{-4}$m/N with $\Delta t = 1.54\times10^{-9}$s for $v = 10$ μm/s [4].

further substantiated by a direct comparison with two different experiments, as discussed in the following Section.

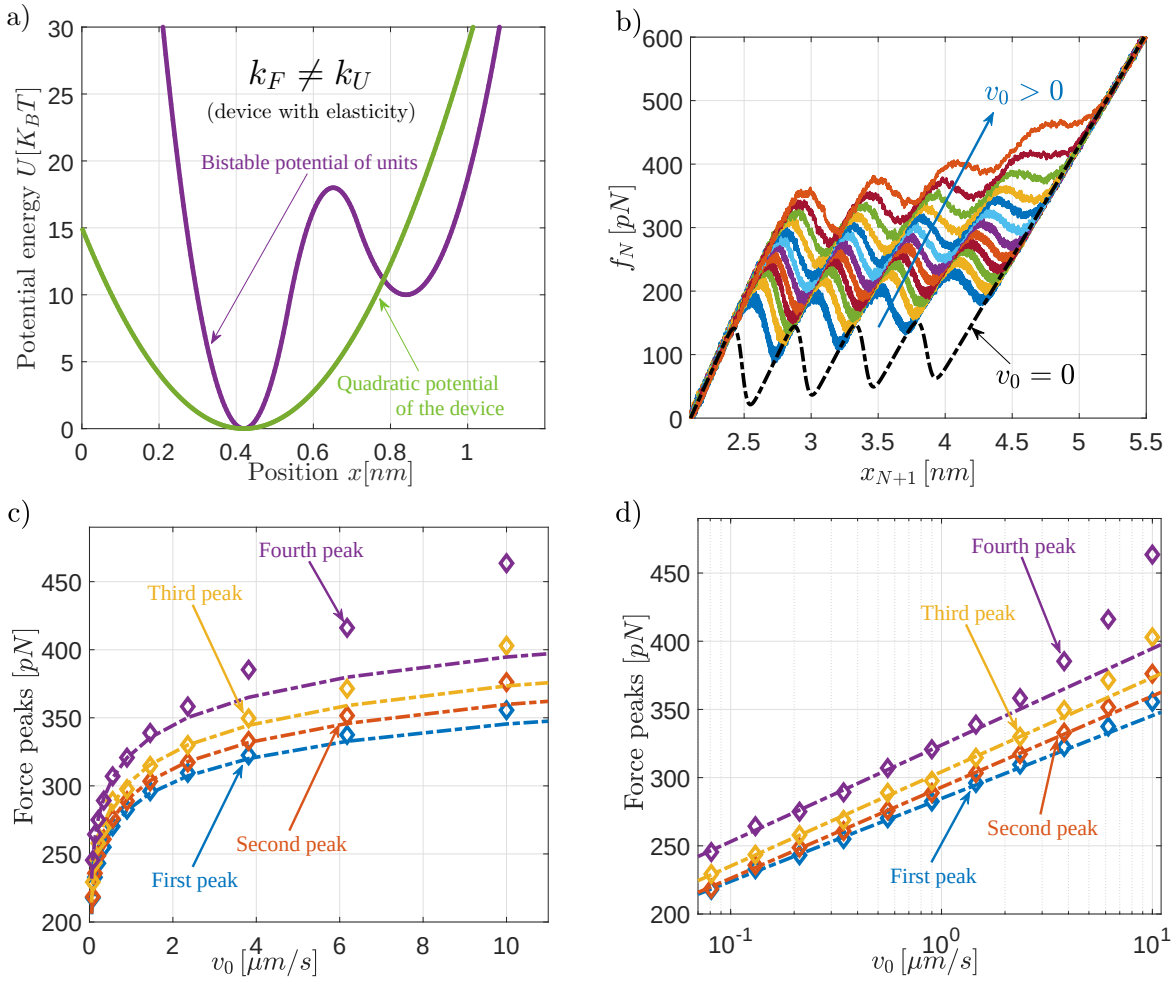

Figure 6.7 – Results for a chain with different elastic constants for the two wells and unfolded through a device with intrinsic elasticity. Potential energy of units and device a), average force exerted on the N-th unit of the chain versus the position x_{N+1} of the device (average curves determined over M=2000 trajectories) b), force peaks versus the applied pulling velocity in linear scale c), and force peaks versus the applied pulling velocity in semi-log scale (symbols: data; dashed lines: linear least squares approximations) d). In panels b), c) and d), we adopted different pulling velocities from 0.081 to 10 μm/s. The dotted black curve in panel b) represents the force-extension response at thermodynamic equilibrium. The results have been obtained with $\Delta B = 18k_BT$, $\Delta E = 10k_BT$, $T = 300$K, N=4, $\ell = \ell_d = 0.42$nm, $\chi = 2$, $k_F = 5.63$N/m, $k_U = 2.81$N/m, $k_d = 0.704$N/m, and η variable ranging from $\eta = 4.1 \times 10^{-2}$m/N with $\Delta t = 2.42 \times 10^{-7}$s for $v = 0.081$ μm/s, to $\eta = 3.3 \times 10^{-4}$m/N with $\Delta t = 1.96 \times 10^{-9}$s for $v = 10$ μm/s [4].

6.4 Theory meets experiments

In this Section, we draw a comparison between the results obtained through the numerical solution of the Langevin equations and two experiments concerning the force

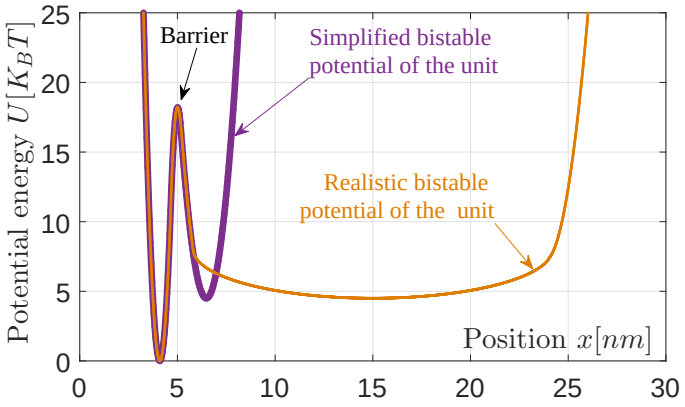

Figure 6.8 – Realistic and simplified potential energy for the filamin domain. The simplified potential has been obtained with $\Delta B = 18.2k_BT$, $\Delta E = 4.5k_BT$, $\ell = 4.1$nm, $\chi = 1.58$, $k_F = 0.295$N/m and $k_U = 0.059$N/m [4].

spectroscopy analysis of filamin and titin. The investigated filamin protein is composed of a single unit, which represents the fourth domain of *Dictyostelium discoideum* filamin (ddFLN4) [196, 197]. On the other hand, the studied titin molecule is composed by 8 repetitions of the 27th domain from the I-band of the human cardiac titin (I27) [198, 203].

6.4.1 Modelling the dynamic stretching of filamin

Filamin belongs to the family of the actin-crosslinking proteins, involved in the constant reorganisation of the cytoskeleton, consisting in the moving and the multiplication of cells [221]. These proteins attach to the actin and stabilise the structure of the latter, and they are therefore called actin-binding-protein (ABP). As a reminder, actin plays with myosin a very important role in the contracting of the muscle. We focus our attention on the filamin found in *Dictyostelium discoideum* (ddFLN), a species of amoeba living in soil, commonly referred to as slime mold whose many of genes are homologous to those of humans and share similar organisations. The ddFLN structure is also known as *gelation factor* or *ABP-120*. Experiments were made on several domains of ddFLN, especially on ddFLN4, the fourth domain of ddFLN, entirely composed of β-sheets [197]. When investigated through force spectroscopy, domain 4 unfolds at a lower force than all the other domains in the *Dictyostelium discoideum* filamin [222]. Therefore, it has been selected to study the unfolding forces in terms of the pulling speed [196].

Our approach is a strong simplification of the reality. To begin, our model is one-dimensional whereas the real structure exhibits a complex three-dimensional geometry. Moreover, since we are interested in understanding the relation between the force peak and

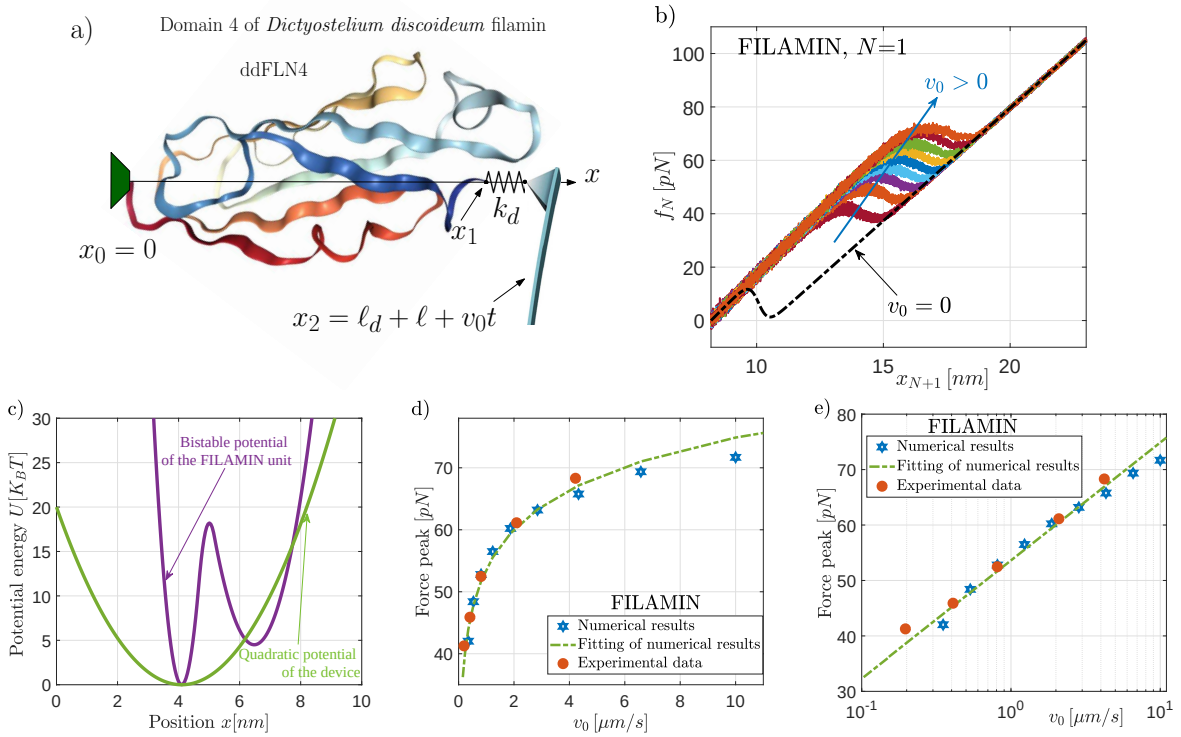

Figure 6.9 – Comparison between numerical results and experimental data for the filamin protein. Panel a): scheme of the force spectroscopy experiment conducted on the filamin unit ($N = 1$). Image from the RCSB PDB (rcsb.org) of PDB ID 1KSR generated by the NGL viewer with NMR data from Ref. [220]. Panel b): average force exerted on the filamin unit versus the prescribed device position (average curves determined over M=2000 trajectories). Panel c): assumed bistable potential energy of the filamin and potential energy of the AFM device. Panel d) and e): force peak versus the applied pulling velocity in linear scale, and in semi-log scale, respectively. In panels b), d) and e), we adopted different pulling velocities from 0.35 to 10 μm/s, coherently with experimental data [196]. The dashed black curve in panel b) represents the force-extension response at thermodynamic equilibrium. The curves have been obtained with $\Delta B = 18.2 k_B T$, $\Delta E = 4.5 k_B T$, $T = 300$K, N=1, $\ell = 4.1$nm, $\chi = 1.58$, $k_F = 0.295$N/m, $k_U = 0.059$N/m, $k_d = 0.00985$N/m, and η variable ranging from $\eta = 2.6$m/N with $\Delta t = 2.6 \times 10^{-7}$s for $v = 0.35$ μm/s, to $\eta = 0.09$m/N with $\Delta t = 9.05 \times 10^{-9}$s for $v = 10$ μm/s [4].

the pulling velocity, the important features are concentrated in the transition region of the dynamics. It means that, for our application, the relevant parts of the potential energy are the first energy well and the following energy barrier. In fact, the peak force is determined by the barrier crossing. Since, for the sake of simplicity, we represent the potential energy with the three parabola expression given in Eq.(6.4), we are able to correctly represent the first well and the barrier and we are forced to neglect the exact geometry of the second well. This simplification should not affect the force-velocity relation and can be found in Fig.6.8, where we compare the realistic and the simplified potential energy. We remark that several theoretical approximations discussed in the Introduction consider only k_F and ΔB (and k_d) for estimating the force-velocity relation. It is therefore important to remember that our model is quite accurate from the point of view of this force-velocity relation, but it does not take into account the correct geometry as previously discussed. The parameters used to model the first energy well and the barrier correspond to the values of the recent literature, obtained by means AFM experiments (under isometric condition) [196, 197]. Importantly, the realistic geometry of the filamin domain has been completely detected by using magnetic tweezers force spectroscopy (under isotensional condition) [223, 224]. It has been clearly proved that the step size (difference between the energy minima corresponding to the folded and unfolded states) is around 12-18nm and the total contour length of the unfolded domain is around 25-30nm, as schematically shown in Fig.6.8.

In Fig.6.9.a, one can find the force spectroscopy AFM configuration applied to the ddFLN4 domain. The structure represented in Fig.6.9.a has been generated with the nuclear magnetic resonance (NMR) spectroscopy data reported in Ref. [220]. The comparison between the numerical Langevin approach and the experimental data obtained by force spectroscopy can be found in the other panels of Fig.6.9. In Fig.6.9.c, we can observe the potential energies of domain (simplified as discussed above) and device with intrinsic elasticity, and we can observe that the device is much softer that the protein under investigation. From Fig.6.9.b, we deduce the shape of the force-extension curves for case with a single unit. We remark that in these curves, the force intensity is correct but the geometry of the domain (its step size) has not been respected as discussed above. Moreover, we can observe that the force spectroscopy experiment is conducted under conditions quite far from thermodynamic equilibrium. Indeed, the coloured curves corresponding to the different pulling speeds are rather far from the dashed black curve representing the thermodynamic equilibrium. Nevertheless, we can see in Figs.6.9.d and e that the force peaks are linearly depending on the logarithm of the applied pulling speed.

We underline the good agreement between our numerical results based on the Langevin equation and the experimental ones, confirmed by the linear least squares approximation shown in Fig.6.9.e. The important point emerging from the agreement between numerical and experimental dynamical results is the following. When we perform an experiment at very low pulling speed, ideally at thermodynamic equilibrium, the force-extension curve measured depends only on the energy jump ΔE and not on the energy barrier ΔB between stable and metastable states. However, when a finite pulling speed is applied, the response depends also on ΔB, which governs the transition rates between the folded and unfolded configurations. Then, performing experiments in the out-of-equilibrium regime allows the indirect measurement of the energy barrier ΔB, which can be obtained by comparing the experimental results with numerical ones. In our specific case, the agreement has been obtained with $\Delta B = 18.2 k_B T$ (at room temperature), confirming the value reported in the literature [196, 197]. It is interesting to point out that the dynamic force spectroscopy method may reveal important features of the energy landscape of a protein, without the necessity to induce the chemical or thermal unfolding of the molecule.

6.4.2 Modelling the dynamic stretching of titin

Titin, also known as *connectin*, is the largest protein of the human body, whose role is to assembly and stabilise the sarcomere, the unit of contraction of the muscle, composed of three systems of filaments: actin, myosin and titin. Data of force spectroscopy show that titin is able to store and to provide energy, mainly by folding and unfolding its multiple immunoglobulin-like domains [226]. We especially focus our attention on domain immunoglobulin-like 27 (I27) of the I-band of the human cardiac titin, also known as domain I91. Each domain has 89 amino-acids and is composed of β-sheets.

As before, our analysis is strongly simplified with respect to the reality. Firstly, the model is one-dimensional as previously discussed. In addition, since we are interested in the force-velocity relationship, we use a simplified geometry, as depicted in Fig.6.8 for the filamin. As a matter of fact, also the energy profile of the titin domain is composed of a very narrow first well, followed by a quite high barrier and a long smooth second well (see Fig.7B of Ref. [198] for details). Given that we search for the force peaks as function of the pulling velocity, we use Eq.(6.4), which allows us to correctly represent the first well and the barrier, but we cannot represent the correct geometry of the second well. This point, however, as previously discussed, should not affect the transition forces. While the physical parameters of the first well and the barrier have been deduced from AFM experiments (under isometric condition) [198, 203], the complete geometrical description of the

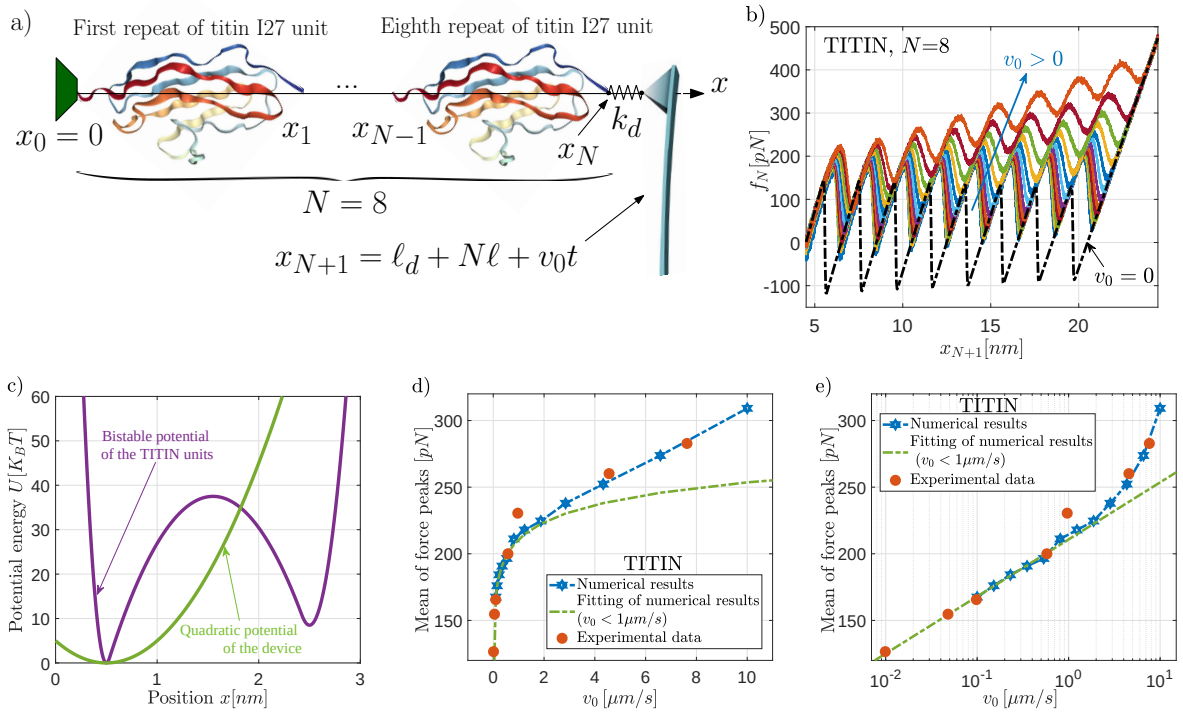

Figure 6.10 – Comparison between numerical results and experimental data for the titin protein. Panel a): scheme of the force spectroscopy experiment conducted on the titin protein ($N = 8$). Image from the RCSB PDB (rcsb.org) of PDB ID 1TIT generated by the NGL viewer with NMR data from Ref. [225]. Panel b): average force exerted on the titin molecule versus the prescribed device position (average curves determined over M=2000 trajectories). Panel c): assumed bistable potential energy of the titin protein and potential energy of the AFM device. Panel d) and e): mean of the $N = 8$ force peaks versus the applied pulling velocity in linear scale, and in semi-log scale, respectively. In panels b), d) and e), we adopted different pulling velocities from 0.1 to 10 μm/s, coherently with experimental data [198]. The dashed black curve in panel b) represents the force-extension response at thermodynamic equilibrium. The curves have been obtained with $\Delta B = 37.5 k_B T$, $\Delta E = 8.5 k_B T$, $T = 300$K, N=8, $\ell = 0.5$nm, $\chi = 5$, $k_F = 9.94$N/m, $k_U = 3.32$N/m and $k_d = 0.166$N/m, and η variable ranging from $\eta = 8.45 \times 10^{-2}$m/N with $\Delta t = 4.2 \times 10^{-7}$s for $v = 0.1$ μm/s, to $\eta = 8.45 \times 10^{-4}$m/N with $\Delta t = 4.2 \times 10^{-9}$s for $v = 10$ μm/s [4].

titin domain has been obtained with magnetic tweezers spectroscopy (under isotensional condition) [227, 228]. In particular, a step size of about 10-20nm and a total contour length of about 30nm have been measured.

Force spectroscopy experiments on mechanical unfolding of eight tethered I27 domains were realised with a pulling speed varying over three orders of magnitude (from 10^{-2} to $10^{1}\mu$m/s) [198, 203]. The NMR spectroscopy allowed to determine the stable I27 structure, which is represented in Fig.6.10.a [225]. In Fig.6.10.b, one can find the force-extension curves numerically obtained for different values of the pulling velocity and with the potential energy of protein units and device represented in Fig.6.10.c. We remember that here, the geometry of the second well has not been respected, being the real step size of about 10-20nm. Each curve in Fig.6.10.b has been obtained as the average value of $M = 2000$ independent Langevin trajectories. As before, the black dashed curve in Fig.6.10.b corresponds to the thermodynamic equilibrium. We remark that some parts of the force-extensions curves are negative: this behaviour comes from the one-dimensional geometry of the system, as discussed at the end of Section 6.3.1. Furthermore, we can underline that the shape of the force peaks in the saw-tooth pattern is not as sharp as observed in the experimental AFM results. This is mainly due to the fact that, in our plot, we show the mean value of several trajectories whereas typical experimental curves correspond to a single realisation of the process. Moreover, in our simplified approach, we have not implemented the classical worm-like chain model, which is able to correctly represent the stretching of most macromolecules and shows a sharper force-extension curve. Indeed, in our one-dimensional chain, we only described the bistable character of the units without introducing a specific persistence length. The observed force peaks are related to the sequential unfolding of the eight titin domains. Then, in Figs.6.10.d and e, the mean value of these eight force peaks is represented versus the pulling velocity in linear and logarithmic scales, respectively. First of all, we observe that the good agreement between numerical and experimental data has been obtained with $\Delta B = 37.5 k_B T$, which is the value accepted in the literature [198, 203]. Moreover, it is interesting to observe that the relation between the mean unfolding force and the logarithm of v_0 is not linear and we can measure a deviation between the linear fitting (calculated for $v_0 < 1\mu$m/s) and the numerical and experimental data for the largest pulling velocities adopted. This behaviour is coherent with previous investigations [34, 203].

6.5 Conclusion

We presented an approach based on the overdamped Langevin formalism able to predict the force-extension response of biological macromolecules unfolded through a force spectroscopy device at a given pulling speed. In particular, we proposed a model able to work in the out-of-equilibrium regime of the statistical mechanics. This model is useful to analyse the experimental data in order to estimate the main important biological and biochemical features of the macromolecule under investigation.

The model presented considers an arbitrary continuous bistable potential energy for describing the units of the macromolecule. This coarse-grained description has the advantage to require less computational resources than those required by the molecular dynamics approach. As a matter of fact, each unit is defined by an effective bistable energy, without taking into account the complete atomistic description of the molecular architecture. On the other hand, the continuous description of the bistable potential energy is more complete than the simplified representation based on a limited set of parameters, which is often adopted in several analytical models. In this regard, the Langevin approach can be seen as a good compromise between methods based on molecular dynamics simulations and others based on analytical approximations. Its implementation allows for considering pulling speeds ranging from the standard AFM to the HS-AFM. We show here a good agreement between the results obtained with this approach and the experimental data concerning the unfolding of filamin and titin.

Conclusions and perspectives

The work presented in this manuscript provides a detailed analysis of bistable chains through the spin variable approach within the Gibbs and the Helmholtz ensembles.
First of all, we present the state of the art and the motivations of our studies. An overview of the single-molecule force spectroscopy techniques is exposed. We also introduce descriptions of the most important macromolecules, such as DNA, RNA or again proteins. We demonstrate why it is important to study the structures with bistability.

The thermodynamics of small systems is discussed in Chapter 2, allowing to introduce both the Gibbs and the Helmholtz ensembles of the statistical mechanics. On the one hand, the Gibbs ensemble is referred to as the isotensional condition, i.e. a force is applied to fix the second end of the chain. On the other hand, the Helmoltz ensemble is referred to as the isometric condition, i.e. the position of the second end is prescribed to fix the chain. This discussion allows to definitely test and verify the statistical mechanics of small systems. Then, a detailed presentation of the spin variable approach, used in Chapters 3, 4 and 5, is given. Finally, we provide in detail the full statistics of conjugated thermodynamic ensembles in bistable chains, enabling to consider not only the average values, but also the complete probability densities of the fluctuating variables.

The Chapter 3 provides the first development of the existing spin variable method. It enables to consider the extensibility of the bonds between the bistable units. Taking account of extensibility is necessary since the finite elastic constant of these bonds plays a major role in the definition of the force-extension response, within both the isotensional and isometric ensembles. A detailed analysis of a two-state extensible freely jointed chain model is proposed. We obtained the exact Gibbs partition function by introducing an approximation considering relatively high values of the elastic constant, coherently with real macromolecules. As the Helmholtz partition function cannot be directly factorised due to the implicit interaction induced by the isometric condition, it is calculated through the Laplace transform of the Gibbs partition function, eventually obtaining its closed form in terms of the Hermite elements with a negative index.

The Chapter 4 deals with the existing interactions among the units, with the help of the Ising scheme. We investigate the behaviour of a chain of two-state units coupled through an Ising interaction scheme. The coefficient λ, representing the Ising interaction coefficient in the final partition functions provides a description of the effects of bistability and cooperativity in biological and artificial micro- and nano-systems. We consider the case of a positive interaction when $\lambda > 0$, which can be associated to a ferromagnetic-like interaction, i.e. the unfolding a unit favours the unfolding of adjacent units. Conversely, we also consider the case of a negative interaction when $\lambda < 0$, which can be associated to an anti-ferromagnetic-like interaction, i.e. the folding of a unit prevents adjacent units from unfolding. The interactions in the Gibbs partition function are studied in the model with the help of the transfer matrix technique. Then, the Helmholtz partition function is elaborated from the Gibbs partition function with the help of the Laplace transform. We firstly study the model far from the thermodynamic limit. In addition, we propose results exploring asymptotic cases, such as weak and strong Ising interactions (for both ferromagnetic-like and anti-ferromagneticlike schemes). We also provide the study of bistable chains with Ising interactions at the thermodynamic limit within the Gibbs ensemble. Finally, we generalise the theory for both extensible bonds and Ising interactions.

The Chapter 5 provides the last development, performed at the thermodynamic equilibrium, for a one-dimensional bistable chain through the spin variables method. The heterogeneity is introduced within the energetic levels. The Helmholtz partition function is obtained in closed form thanks to the Laplace transform and the determinant form of the so-called Newton-Girard formulae. The heterogeneity of energetic levels of two-state units allows to obtain the important unfolding sequence of the bistable chain, i.e. the so-called unfolding pathway. Indeed, introducing various energetic levels of the units breaks the symmetry and creates an inequality among the unfolding probabilities. Therefore, at each switching occurence, once the heterogeneity introduced, the probability of unfolding one given unit tends to 1, whereas the probability of the others tends to 0. It means that the system implements a statistical sorting procedure. Eventually, the concept is identifiability is proposed, to measure the capability to identify the most probable unfolding pathway.

Finally, the Chapter 6 considers the case of an out-of-equilibrium bistable chain. Indeed, it has been proved that the pulling speed exerts an influence on the amplitude of the force peaks measured during stretching experiments on macromolecules. The presented

approach is based on the Langevin formalism, allowing to predict the force-extension response of biological macromolecules unfolded through a force spectroscopy device at a given pulling speed. The units of the macromolecule are described by an arbitrary continuous bistable potential. The Langevin approach can be seen as a good compromise between methods based on molecular dynamics simulations and others based on analytical approximations, in order to consider a wider range of pulling speeds (from standard to high-speed AFM). Finally, we show a good agreement between the results obtained with this approach and the experimental data concerning the unfolding of filamin and titin.

The models introduced in this thesis are rather simple and paradigmatic, i.e. introduced to study the effect of a single chain property to the final thermodynamic and mechanical response. This choice permits to better discuss the theoretical origin of the specific features of the force-extension curves observed in the different situations and their statistical character. Nevertheless, the models can be further improved to take into account other relevant physical and geometrical features. For instance, concerning the application to macromolecules, the following generalisations can be considered: the dynamical approach (out-of-equilibrium) should be applied to the case with Ising interactions; the study of the heterogeneity should be generalised to deal with three-dimensional bistable freely jointed chains (also with extensibility); the Ising interaction scheme could be generalised to the case where the interactions are not only between neighbouring domains but also long-range to better describe globular proteins and more complex system. Moreover, recent works have provided evidence that the traction velocity applied to the chain with a hard device plays an important role in defining the unfolding pathway [152, 229–231]. Therefore, it is important to fully analyse the interplay between the distribution of energy jumps and the applied traction velocity on the unfolding pathway. To do this, the Langevin methodology could be combined with the spin variables technique in order to fully describe the dynamics of the system.

In addition to the applications discussed in the present thesis, force spectroscopy allows the study of many different problems including, e.g., mechanical resistance of chemical bonds. In this case, the force at which the bond breaks up can be directly measured [232]. Further, it is possible to map the energy landscape of the chemical bond under mechanical force [233]. This is very important to investigate the antibody-antigen bond, and the protein-protein or protein-living cell interactions [234]. Moreover, the adhesion of cells on tissues or substrates is an interesting similar field of research concerning the so-called Focal Adhesions, which are large macromolecular assemblies through which mechanical forces are transmitted between the extracellular matrix and cells [235–237] (see Fig.7.1).

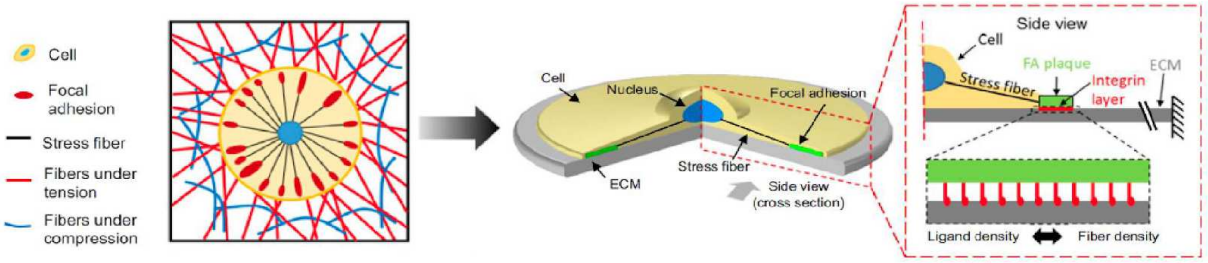

Figure 7.1 – Coarse-grained model for the mechanical response of the Focal Adhesions (FA). Left panel: schematic of a cell adhered to fibrous Extra Cellular Matrix (ECM). The FA are formed at the periphery of the cell. Right panel: schematic of a coarse-grained model: stress fibers connect the FA and the nucleus. The FA is connected to ECM through an integrin layer whose density is correlated with the fiber density under the cell [238].

Also these problems can be approached with the statistical mechanics methodologies introduced in this thesis.

On the other hand, in mechanics of materials, and more generally in material science, similar behaviours have been observed in different situations. First of all, in the plastic behaviour of materials, each single plastic event can be explained through an instability, which is able to switch the system from a ground state to a permanently deformed or metastable state. These forms of micro-instabilities can be seen in different shapes of the force-extension (or stress-strain) response of plastic materials. For example, Lüders bands are localised bands of plastic deformation observed in metals experiencing tensile stresses, which usually starts at one end of the specimen and propagates toward the other one [239]. During the bands propagation, the stress-strain curve is quite flat with some fluctuations due to the internal microinstabilities (see Fig.7.2, left). Another example is given by the Portevin-Le Chatelier effect, which can be observed as an unstable plastic flow during tensile tests of some dilute alloys (under certain regimes of strain rate and temperature) [240]. In this case, a series of peaks can be seen in a region with a positive slope of the stress-strain curve. These peaks represent the sequence of instabilities of the plastic phenomenon (see Fig.7.2, center). Other manifestations of discontinuous deformation are staircase creep and the Savart–Masson effect. They imply the formation of steps on loading curves with a constant stress rate in a soft testing machine (see Fig.7.2, right). There are relatively few studies on the Savart–Masson effect [241], although it was discovered long before the Portevin-Le Chatelier effect.

Another important situation where the micro-instabilities between two states play a central role is the adhesion of films and their peeling from a substrate [243] (see Fig.7.3,

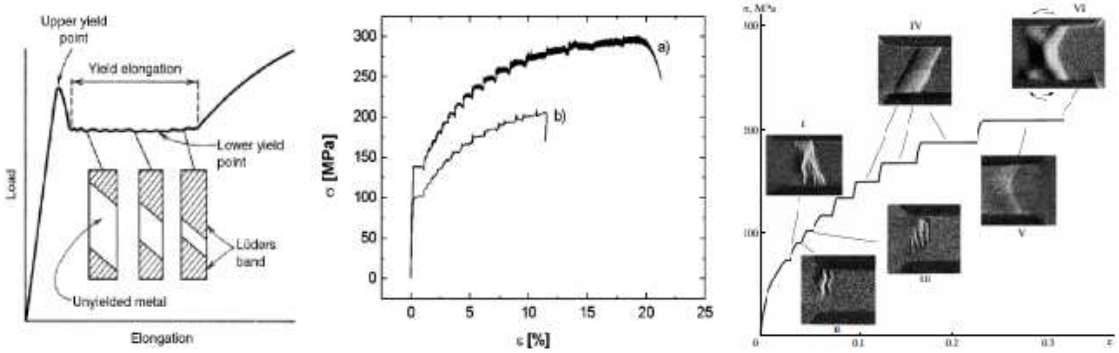

Figure 7.2 – Three examples of stress strain curves of materials showing the Lüders bands in the first case, the Portevin–Le Chatelier effect in the second case and the Savart–Masson effect in the third one. In all these phenomena, we have a series of micro-instabilities generating the plastic behaviour [239, 241, 242].

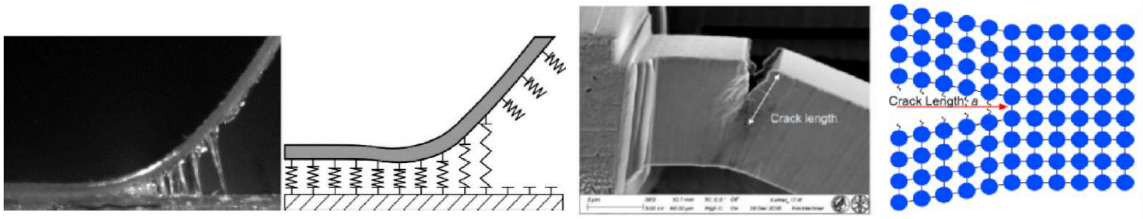

Figure 7.3 – Microinstabilities in the peeling process (left) and in the crack propagation (right) [243, 244].

left). In this context, one idea to improve and tailor the adhesion properties of artificial materials is to use bioinspired geometries and mechanisms [244]. The combination of thermal fluctuation and micro-instabilities is also at the origin of the fracture phenomena in materials. If we think to a crystalline material, the propagation of a crack front must be identified by the gradual breaking of the atomic interactions along a line of the crystal structure (see Fig.7.3, right). Therefore, each single crack propagation step is a form of instability, which can be defined by assuming a maximum length of the interatomic distance before rupture. Of course, the thermal fluctuations modify the crack propagation, fostering the breaking of the interatomic bonds. In some cases, the temperature effects can cause the crack generation and therefore we say that we have a thermally activated crack propagation [245, 246].

The observed similarities between the problems concerning micro-instabilities in biophysics and material science is of great importance, not only from the cultural point of view, but also for the possibility to adopt the same techniques to approach very different themes. Indeed, the common denominator in all these examples, from both biophysics and

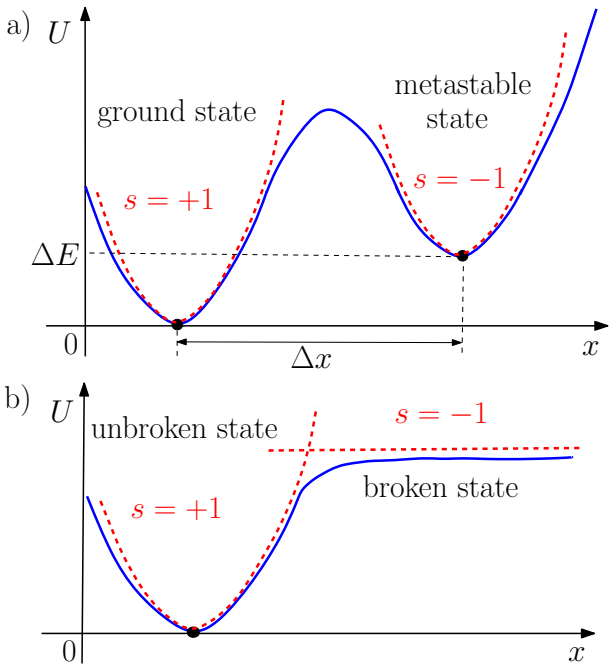

Figure 7.4 – Paradigmatic potential energies of the system components for the two different classes of micro-instabilities discussed in the main text: bistability between a folded (ground) state and an unfolded (metastable) state a), and damage process between the intact (unbroken) state and the damaged (broken) state b).

material science, is that some units of the system under investigation exhibit a two-state or multi-state character. In particular, we can identify two classes of micro-instabilities. On the one side, we can have a bistable (or multistable) behaviour between one ground state and one (or more) metastable state. These admitted states represent different conformations or configurations of the intact yet deformable system units. For instance, this scheme corresponds to conformational transitions in macromolecules, as studied in this thesis, or plastic transformations in solids. On the other side, we can also have transitions between the broken or unbroken states of some breakable units of the system. In this case, the system unit can switch from the intact condition to the damaged condition, and this process can be reversible or not depending on the specific situation. Examples of this scheme include deadhesion of cells, peeling of films or crack propagation.

To better appreciate the real difference between these two classes of micro-instabilities, we can find in Fig.7.4 the potential energy $U(x)$ describing the units behaviour in both cases (we consider a one-dimensional geometry for the sake of simplicity). In the first case (Fig.7.4.a), we have an elastic unit with two configurations (folded and unfolded) corresponding to the two wells of the potential energy, having different equilibrium lengths and elastic constants. In the second case (Fig.7.4.b), we have a breakable mechanical

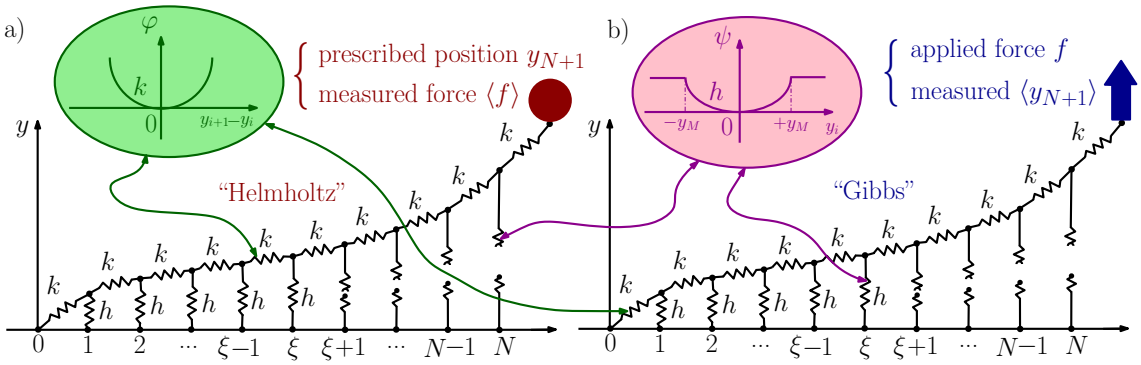

Figure 7.5 – Scheme of the cohesion-decohesion process within both the Helmholtz a) and the Gibbs b) ensembles. While in the first case, we prescribe the position of the last unit and we measure the average force $\langle r \rangle$, in the second case, we apply a force f and we measure the average position of the last unit. In both cases, we consider ξ unbroken vertical springs and $N - \xi$ broken vertical springs. The energy potentials φ and Ψ correspond to the horizontal and the vertical springs, respectively.

unit with the unbroken configuration corresponding to the potential well, and the broken configuration corresponding to horizontal portion of the curve. Indeed, in this region the exerted force is zero, which means that we are in the broken state of the unit. Peyrard, Bishop and Dauxois have extensively studied the damage process between the intact (unbroken) and damaged (broken) states. The adhesion-deadhesion model has been inspired by their PBD model of DNA denaturation [247–250]. The method of the "spin variables" discussed in this thesis has been largely adopted to model units with transitions between ground and metastable states (see Fig.7.4.a). However, the case of units undergoing damage processes between unbroken and broken states (Fig.7.4.b) must be further examined and can be investigated by the same approach. As a matter of fact, the thermal effects on the damage processes are still far from being fully understood. Therefore, an important perspective of the spin variable approach is the analysis of the cohesion-decohesion processes for complex systems, with the crucial effect of the temperature.

To do this, we can imagine a paradigmatic system representing the adhesion-deadhesion of a film from a substrate. This basic system that may be approached is shown in Fig.7.5, where one can find the film represented by the horizontal mass-spring chain, and grounded to the substrate through a series of breakable springs. This configuration mimics adhesion, cohesion, and the rupture phenomena in general. The system is embedded in a thermal bath to consider the thermal fluctuations and their effects on the adhesion mechanism. This point makes the model well adapted to represent adhesion of cells, unzipping of hairpins, denaturation of DNA and other macromolecules, peeling of films from a substrate,

fracture propagation in solids, rupture phenomena and so on.

In this scheme, the vertical breakable elements remain intact if the extension is lower than a given threshold and become broken if the extension is larger than the same threshold. With this definition, we can study the adhesion process with an external mechanical action. The spin variables method allows for describing the system response, at a fixed temperature, within the two different statistical ensembles. A further envisaged generalisation concern the model shown in Fig.7.5. Here, we can find the same geometry of Fig.7.4 with however a different behaviour in the response of the breakable elements. Indeed, their response is represented in Fig.7.5.d, where potential energy and force are plotted. At a given applied extension $\pm y_p$ (points A or B), we can observe a softening mechanism, which produces a change of the elastic constant from h_e to h_p with $h_p < h_e$. Then, when the extension increases to an even larger value $\pm y_b$, we have the complete breaking of the elements (points C or D).

This scheme perfectly reproduces a biological situation corresponding to the so-called sacrificial bond [251]. Sacrificial bonds are defined as bonds that break before the main structural link (often the molecular backbone) is broken, thus producing a change in the effective elastic constant. These bonds are frequently weaker than the covalent bond of molecular backbones. Sacrificial bonds in structural molecules and composites have been found to greatly increase the fracture toughness of biomaterials by providing a reversible, molecular-scale energy-dissipation mechanism. In many biological cases, the breaking of sacrificial bonds has been found to be reversible, thereby additionally providing a "self-healing" property to the material. Due to the nanoscopic nature of this mechanism, single-molecule force spectroscopy using an atomic force microscope has been a useful tool to investigate this mechanism.

Also in material science, the softening mechanism is relevant to describe the strength of materials as a function of the temperature. Indeed, in many crystalline materials, the failure is obtained through the propagation of the crack in correspondence of the application of the material strength. For higher values of the temperature, however, the crack propagation occurs after the generation of a population of dislocations (thermally activated), which are able to reduce the elastic moduli of the material [252]. Then, the proposed model is relevant to study the strength of materials at high temperature values.

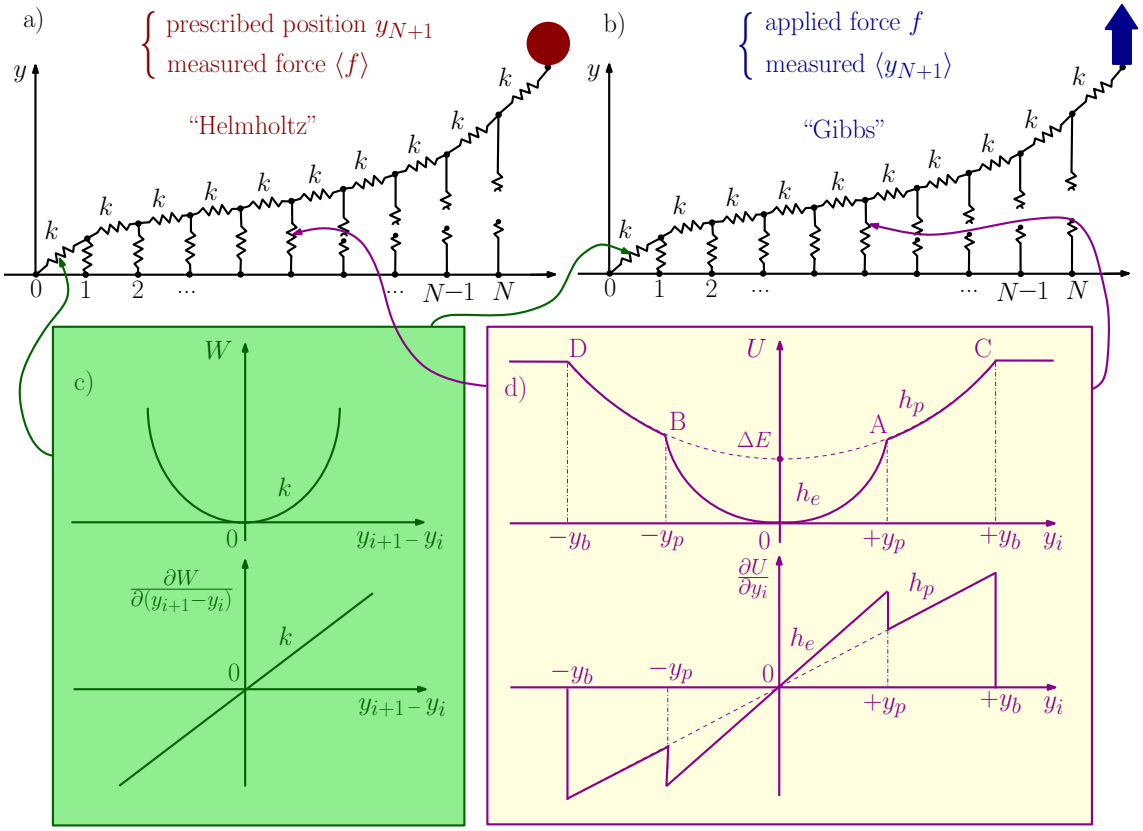

Figure 7.6 – Scheme of the cohesion-decohesion process within both the Helmholtz a) and the Gibbs b) ensembles. While in the first case, we prescribe the position of the last unit of the chain and we measure the average force $\langle f \rangle$, in the second case, we apply a force f and we measure the average position. In both cases, we consider a linear elastic behaviour for the horizontal springs c) and a breakable response with softening mechanism d) for the vertical elements. Coherently, the energy potentials W and U correspond to the horizontal and the vertical springs, respectively.

Bibliography

[1] M. Benedito and S. Giordano. Thermodynamics of small systems with conformational transitions: The case of two-state freely jointed chains with extensible units. *The Journal of Chemical Physics*, 149(5):054901, August 2018.

[2] M. Benedito and S. Giordano. Isotensional and isometric force-extension response of chains with bistable units and Ising interactions. *Physical Review E*, 98(5):052146, November 2018.

[3] M. Benedito, F. Manca, and S. Giordano. Full Statistics of Conjugated Thermodynamic Ensembles in Chains of Bistable Units. *Inventions*, 4(1):19, March 2019.

[4] M. Benedito and S. Giordano. Unfolding pathway and its identifiability in heterogeneous chains of bistable units. *Physics Letters A*, 384(5):126124, February 2020.

[5] M. Benedito, F. Manca, P. L. Palla, and S. Giordano. Rate-dependent force–extension models for single-molecule force spectroscopy experiments. *Physical Biology*, 17(5):056002, August 2020.

[6] M. L. Hughes and L. Dougan. The physics of pulling polyproteins: A review of single molecule force spectroscopy using the AFM to study protein unfolding. *Reports on Progress in Physics. Physical Society (Great Britain)*, 79(7):076601, July 2016.

[7] 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*, 66(1):1–45, December 2002.

[8] D. C. Rees, T. N. Williams, and M. T. Gladwin. Sickle-cell disease. *The Lancet*, 376(9757):2018–2031, December 2010.

[9] Sickle cell disease. https://www.inserm.fr/information-en-sante/dossiers-information/drepanocytose.

[10] K. Wüthrich. NMR with Proteins and Nucleic Acids. *Europhysics News*, 17(1):11–13, 1986.

[11] D. Ringe and G. A. Petsko. Study of protein dynamics by X-ray diffraction. In *Methods in Enzymology*, volume 131 of *Enzyme Structure Part L*, pages 389–433. Academic Press, January 1986.

[12] G. Weber. Energetics of Ligand Binding to Proteins. In C. B. Anfinsen, John T. Edsall, and Frederic M. Richards, editors, *Advances in Protein Chemistry*, volume 29, pages 1–83. Academic Press, January 1975.

[13] R. H. Austin, K. W. Beeson, L. Eisenstein, H. Frauenfelder, and I. C. Gunsalus. Dynamics of ligand binding to myoglobin. *Biochemistry*, 14(24):5355–5373, December 1975.

[14] H. Frauenfelder, G. A. Petsko, and D. Tsernoglou. Temperature-dependent X-ray diffraction as a probe of protein structural dynamics. *Nature*, 280(5723):558–563, August 1979.

[15] H. Frauenfelder, S. G. Sligar, and P. G. Wolynes. The energy landscapes and motions of proteins. *Science*, 254(5038):1598–1603, December 1991.

[16] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7(4):284–304, April 1940.

[17] J. S. Langer. Theory of the condensation point. *Annals of Physics*, 41(1):108–157, January 1967.

[18] J. S. Langer. Theory of Nucleation Rates. *Physical Review Letters*, 21(14):973–976, September 1968.

[19] J. S. Langer. Statistical theory of the decay of metastable states. *Annals of Physics*, 54(2):258–275, September 1969.

[20] F. Manca, S. Giordano, P. L. Palla, and F. Cleri. On the equivalence of thermodynamics ensembles for flexible polymer chains. *Physica A: Statistical Mechanics and its Applications*, 395:154–170, February 2014.

[21] A. M. Skvortsov, L. I. Klushin, and F. a. M. Leermakers. Negative compressibility and nonequivalence of two statistical ensembles in the escape transition of a polymer chain. *The Journal of Chemical Physics*, 126(2):024905, January 2007.

[22] D. I. Dimitrov, L. I. Klushin, A. Skvortsov, A. Milchev, and K. Binder. The escape transition of a polymer: A unique case of non-equivalence between statistical ensembles. *The European Physical Journal E*, 29(1):9–25, May 2009.

[23] A. M. Skvortsov, L. I. Klushin, A. A. Polotsky, and K. Binder. Mechanical desorption of a single chain: Unusual aspects of phase coexistence at a first-order transition. *Physical Review E*, 85(3):031803, March 2012.

[24] M. Rief, F. Oesterhelt, B. Heymann, and H. E. Gaub. Single Molecule Force Spectroscopy on Polysaccharides by Atomic Force Microscopy. *Science*, 275(5304):1295–1297, February 1997.

[25] T. E. Fisher, A. F. Oberhauser, M. Carrion-Vazquez, P. E. Marszalek, and J. M. Fernandez. The study of protein mechanics with the atomic force microscope. *Trends in Biochemical Sciences*, 24(10):379–384, October 1999.

[26] I. Rehman and S. Botelho. Biochemistry, Tertiary Structure, Protein. December 2017.

[27] Huntington's disease. https://fr.cisbio.eu/content/new-hope-for-huntingtons-disease-drug-discovery/.

[28] S. B. Prusiner. Molecular biology of prion diseases. *Science*, 252(5012):1515–1522, June 1991.

[29] Madcow disease. https://www.news-medical.net/health/How-Do-Prion-Diseases-Spread-(French).aspx.

[30] S. Getfert. Escape processes far from thermal equilibrium : Path integrals and force spectroscopy. 2009.

[31] S. S. Mandal. Force Spectroscopy on Single Molecules of Life. *ACS Omega*, 5(20):11271–11278, May 2020.

[32] G. Binnig, C. F. Quate, and C. Gerber. Atomic Force Microscope. *Physical Review Letters*, 56(9):930–933, March 1986.

[33] T. Hoffmann and L. Dougan. Single molecule force spectroscopy using polyproteins. *Chemical Society Reviews*, 41(14):4781–4796, June 2012.

[34] F. Rico, L. Gonzalez, I. Casuso, M. Puig-Vidal, and S. Scheuring. High-Speed Force Spectroscopy Unfolds Titin at the Velocity of Molecular Dynamics Simulations. *Science*, 342(6159):741–743, November 2013.

[35] F. Eghiaian, F. Rico, A. Colom, I. Casuso, and S. Scheuring. High-speed atomic force microscopy: Imaging and force spectroscopy. *FEBS Letters*, 588(19):3631–3638, October 2014.

[36] H. Takahashi, F. Rico, C. Chipot, and S. Scheuring. α-Helix Unwinding as Force Buffer in Spectrins. *ACS Nano*, 12(3):2719–2727, March 2018.

[37] C. Valotteau, F. Sumbul, and F. Rico. High-speed force spectroscopy: Microsecond force measurements using ultrashort cantilevers. *Biophysical Reviews*, 11(5):689–699, October 2019.

[38] F. Rico, A. Russek, L. González, H. Grubmüller, and S. Scheuring. Heterogeneous and rate-dependent streptavidin–biotin unbinding revealed by high-speed force spectroscopy and atomistic simulations. *Proceedings of the National Academy of Sciences of the United States of America*, 116(14):6594–6601, April 2019.

[39] F. Gräter, J. Shen, H. Jiang, M. Gautel, and H. Grubmüller. Mechanically induced titin kinase activation studied by force-probe molecular dynamics simulations. *Biophysical Journal*, 88(2):790–804, February 2005.

[40] E. H. Lee, J. Hsin, M. Sotomayor, G. Comellas, and K. Schulten. Discovery through the computational microscope. *Structure (London, England: 1993)*, 17(10):1295–1306, October 2009.

[41] T. R. Strick, J.-F. Allemand, D. Bensimon, A. Bensimon, and V. Croquette. The Elasticity of a Single Supercoiled DNA Molecule. *Science*, 271(5257):1835–1837, March 1996.

[42] J. K. Fisher, J. Cribb, K. V. Desai, L. Vicci, B. Wilde, K. Keller, R. M. Taylor, J. Haase, K. Bloom, E. T. O'Brien, and R. Superfine. Thin-foil magnetic force system for high-numerical-aperture microscopy. *The Review of scientific instruments*, 77(2):023702–1–023702–9, February 2006.

[43] C. Gosse and V. Croquette. Magnetic tweezers: Micromanipulation and force measurement at the molecular level. *Biophysical Journal*, 82(6):3314–3329, June 2002.

[44] C. Haber and D. Wirtz. Magnetic tweezers for DNA micromanipulation. *Review of Scientific Instruments*, 71(12):4561–4570, November 2000.

[45] J. Yan, D. Skoko, and J. F. Marko. Near-field-magnetic-tweezer manipulation of single DNA molecules. *Physical Review E*, 70(1):011905, July 2004.

[46] I. De Vlaminck and C. Dekker. Recent Advances in Magnetic Tweezers. *Annual Review of Biophysics*, 41(1):453–472, 2012.

[47] A. Ashkin. History of Optical Trapping and Manipulation of Small Neutral Particles, Atoms, and Molecules. In Rudolf Rigler, Michel Orrit, and Thomas Basché, editors, *Single Molecule Spectroscopy: Nobel Conference Lectures*, Springer Series in Chemical Physics, pages 1–31. Springer, Berlin, Heidelberg, 2001.

[48] A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm, and S. Chu. Observation of a single-beam gradient force optical trap for dielectric particles. *Optics Letters*, 11(5):288–290, May 1986.

[49] M. J. Lang and S. M. Block. Resource Letter: LBOT-1: Laser-based optical tweezers. *American Journal of Physics*, 71(3):201–215, February 2003.

[50] J.-P. Galaup. La pince optique. *Photoniques*, (66):45–49, July 2013.

[51] M. Ribezzi-Crivellari and F. Ritort. Force Spectroscopy with Dual-Trap Optical Tweezers: Molecular Stiffness Measurements and Coupled Fluctuations Analysis. *Biophysical Journal*, 103(9):1919–1928, November 2012.

[52] J. Camunas-Soler, M. Ribezzi-Crivellari, and F. Ritort. Elastic Properties of Nucleic Acids by Single-Molecule Force Spectroscopy. *Annual Review of Biophysics*, 45(1):65–84, 2016.

[53] A. M. Monge, M. Manosas, and F. Ritort. Experimental test of ensemble inequivalence and the fluctuation theorem in the force ensemble in DNA pulling experiments. *Physical Review E*, 98(3):032146, September 2018.

[54] F. Landuzzi, X. Viader-Godoy, F. Cleri, I. Pastor, and F. Ritort. Detection of single DNA mismatches by force spectroscopy in short DNA hairpins. *The Journal of Chemical Physics*, 152(7):074204, February 2020.

[55] C.-H. Chiou and G.-B. Lee. A micromachined DNA manipulation platform for the stretching and rotation of a single DNA molecule. *Journal of Micromechanics and Microengineering*, 15:109–117, January 2005.

[56] F. Ren, Y. Zu, K. Kumar Rajagopalan, and S. Wang. Regulation of DNA conformations and dynamics in flows with hybrid field microfluidics. *Biomicrofluidics*, 6(4):044103, October 2012.

[57] M. C. Tarhan, N. Lafitte, Y. Tauran, L. Jalabert, M. Kumemura, G. Perret, B. Kim, A. W. Coleman, H. Fujita, and D. Collard. A rapid and practical technique for real-time monitoring of biomolecular interactions using mechanical responses of macromolecules. *Scientific Reports*, 6(1):28001, June 2016.

[58] M. S. Hanay, S. Kelber, A. K. Naik, D. Chi, S. Hentz, E. C. Bullard, E. Colinet, L. Duraffourg, and M. L. Roukes. Single-protein nanomechanical mass spectrometry in real time. *Nature Nanotechnology*, 7(9):602–608, September 2012.

[59] C. Yamahata, D. Collard, B. Legrand, T. Takekawa, M. Kumemura, G. Hashiguchi, and H. Fujita. Silicon Nanotweezers With Subnanometer Resolution for the Micromanipulation of Biomolecules. *Journal of Microelectromechanical Systems*, 17(3):623–631, June 2008.

[60] G. Perret, T. Lacornerie, F. Manca, S. Giordano, M. Kumemura, N. Lafitte, L. Jalabert, M. C. Tarhan, E. F. Lartigau, F. Cleri, H. Fujita, and D. Collard. Real-time mechanical characterization of DNA degradation under therapeutic X-rays and its theoretical modeling. *Microsystems & Nanoengineering*, 2(1):1–9, December 2016.

[61] DNA. https://en.wikipedia.org/wiki/DNA.

[62] Types of RNA. https://www.news-medical.net/life-sciences/-Types-of-RNA-mRNA-rRNA-and-tRNA-(French).aspx.

[63] Chromatin. https://www.genome.gov/genetics-glossary/Chromatin.

[64] H. Meng, K. Andresen, and J. van Noort. Quantitative analysis of single-molecule force spectroscopy on folded chromatin fibers. *Nucleic Acids Research*, 43(7):3578–3590, April 2015.

[65] M. Kruithof, F.-T. Chien, A. Routh, C. Logie, Daniela D., and J. van Noort. Single-molecule force spectroscopy reveals a highly compliant helical folding for the 30-nm chromatin fiber. *Nature Structural & Molecular Biology*, 16(5):534–540, May 2009.

[66] F.-T. Chien and J. Noort. 10 Years of Tension on Chromatin: Results from Single Molecule Force Spectroscopy. *Current pharmaceutical biotechnology*, 10:474–85, September 2009.

[67] Y. Cui and C. Bustamante. Pulling a single chromatin fiber reveals the forces that maintain its higher-order structure. *Proceedings of the National Academy of Sciences of the United States of America*, 97(1):127–132, January 2000.

[68] S. Cocco, J. F. Marko, R. Monasson, A. Sarkar, and J. Yan. Force-extension behavior of folding polymers. *The European Physical Journal E*, 10(3):249–263, March 2003.

[69] J. D. Watson and F. H. C. Crick. Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*, 171(4356):737–738, April 1953.

[70] S. B. Smith, L. Finzi, and C. Bustamante. Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads. *Science*, 258(5085):1122–1126, November 1992.

[71] J. F. Marko and E. D. Siggia. Stretching DNA. *Macromolecules*, 28(26):8759–8770, December 1995.

[72] C. Bustamante, J. F. Marko, E. D. Siggia, and S. Smith. Entropic elasticity of lambda-phage DNA. *Science*, 265(5178):1599–1600, September 1994.

[73] S. B. Smith, Y. Cui, and C. Bustamante. Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules. *Science (New York, N.Y.)*, 271(5250):795–799, February 1996.

[74] I. Rouzina and V. A. Bloomfield. Force-induced melting of the DNA double helix 1. Thermodynamic analysis. *Biophysical Journal*, 80(2):882–893, February 2001.

[75] I. Rouzina and V. A. Bloomfield. Force-induced melting of the DNA double helix. 2. Effect of solution conditions. *Biophysical Journal*, 80(2):894–900, February 2001.

[76] S. Cocco, J. Yan, J.-F. Léger, D. Chatenay, and J. F. Marko. Overstretching and force-driven strand separation of double-helix DNA. *Physical Review E*, 70(1):011910, July 2004.

[77] N. Bosaeus, A. H. El-Sagheer, T. Brown, S. B. Smith, B. Åkerman, C. Bustamante, and B. Nordén. Tension induces a base-paired overstretched DNA conformation. *Proceedings of the National Academy of Sciences*, 109(38):15179–15184, September 2012.

[78] M. Manghi, N. Destainville, and J. Palmeri. Mesoscopic models for DNA stretching under force: New results and comparison with experiments. *The European Physical Journal E*, 35(10):110, October 2012.

[79] C. Bustamante, Z. Bryant, and S. B. Smith. Ten years of tension: Single-molecule DNA mechanics. *Nature*, 421(6921):423–427, January 2003.

[80] O. Kratky and G. Porod. Röntgenuntersuchung gelöster Fadenmoleküle. *Recueil des Travaux Chimiques des Pays-Bas*, 68(12):1106–1122, 1949.

[81] A. Rosa, T. X. Hoang, D. Marenduzzo, and A. Maritan. Elasticity of Semiflexible Polymers with and without Self-Interactions. *Macromolecules*, 36(26):10095–10102, December 2003.

[82] A. Rosa, T. X. Hoang, D. Marenduzzo, and A. Maritan. A new interpolation formula for semiflexible polymers. *Biophysical Chemistry*, 115(2):251–254, April 2005.

[83] S. Doniach. Biological Physics: Energy, Information, Life. *Physics Today*, 57(11):63–64, November 2004.

[84] J. Lipfert, J. W. J. Kerssemakers, T. Jager, and N. H. Dekker. Magnetic torque tweezers: Measuring torsional stiffness in DNA and RecA-DNA filaments. *Nature Methods*, 7(12):977–980, December 2010.

[85] J. Lipfert, M. Wiggin, J.W. J. Kerssemakers, F. Pedaci, and N. H. Dekker. Freely orbiting magnetic tweezers to directly monitor changes in the twist of nucleic acids. *Nature Communications*, 2:439, August 2011.

[86] S. K. Nomidis, F. Kriegel, W. Vanderlinden, J. Lipfert, and E. Carlon. Twist-Bend Coupling and the Torsional Response of Double-Stranded DNA. *Physical Review Letters*, 118(21):217801, May 2017.

[87] S. K. Nomidis, E. Skoruppa, E. Carlon, and J. F. Marko. Twist-bend coupling and the statistical mechanics of the twistable wormlike-chain model of DNA: Perturbation theory and beyond. *Physical Review E*, 99(3):032414, March 2019.

[88] Genetic code. https://www.genome.gov/genetics-glossary/Genetic-Code.

[89] L. Pauling, R. B. Corey, and H. R. Branson. The structure of proteins: Two hydrogen-bonded helical configurations of the polypeptide chain. *Proceedings of the National Academy of Sciences*, 37(4):205–211, April 1951.

[90] Protein. https://www.genome.gov/genetics-glossary/Protein.

[91] A. E. Mirsky and L. Pauling. On the Structure of Native, Denatured, and Coagulated Proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 22(7):439–447, July 1936.

[92] J. Sabelko, J. Ervin, and M. Gruebele. Observation of strange kinetics in protein folding. *Proceedings of the National Academy of Sciences of the United States of America*, 96(11):6031–6036, May 1999.

[93] Y. Fichou. Hydration water dynamics of the tau protein in its native and amyloid states. March 2015.

[94] M. Caruel and L. Truskinovsky. Physics of muscle contraction. *Reports on Progress in Physics*, 81(3):036602, February 2018.

[95] M. Caruel, P. Moireau, and D. Chapelle. Stochastic modeling of chemical-mechanical coupling in striated muscles. *Biomechanics and Modeling in Mechanobiology*, 18(3):563–587, June 2019.

[96] H. A. Wu. Molecular dynamics study on mechanics of metal nanowire. *Mechanics Research Communications*, 33(1):9–16, January 2006.

[97] W. Liang, D. J. Srolovitz, and M. Zhou. A micromechanical continuum model for the tensile behavior of shape memory metal nanowires. *Journal of the Mechanics and Physics of Solids*, 55(8):1729–1761, August 2007.

[98] X. Guo, W. Liang, and M. Zhou. Mechanism for the Pseudoelastic Behavior of FCC Shape Memory Nanowires. *Experimental Mechanics*, 49(2):183–190, April 2009.

[99] A. Rafsanjani and D. Pasini. Bistable Auxetic Mechanical Metamaterials Inspired by Ancient Geometric Motifs. *Extreme Mechanics Letters*, 9:291–296, 2016.

[100] S. Katz and S. Givli. Solitary waves in a bistable lattice. *Extreme Mechanics Letters*, 22:106–111, July 2018.

[101] I. Müller and P. Villaggio. A model for an elastic-plastic body. *Archive for Rational Mechanics and Analysis*, 65(1):25–46, March 1977.

[102] B. Fedelich and G. Zanzotto. Hysteresis in discrete systems of possibly interacting elements with a double-well energy. *Journal of Nonlinear Science*, 2(3):319–342, September 1992.

[103] G. Puglisi and L. Truskinovsky. Thermodynamics of rate-independent plasticity. *Journal of the Mechanics and Physics of Solids*, 53:655–679, March 2005.

[104] M. Caruel, J. M. Allain, and L. Truskinovsky. Mechanics of collective unfolding. *Journal of the Mechanics and Physics of Solids*, 76:237–259, March 2015.

[105] Y. R. Efendiev and L. Truskinovsky. Thermalization of a driven bi-stable FPU chain. *Continuum Mechanics and Thermodynamics*, 22(6):679–698, September 2010.

[106] A. Mielke and L. Truskinovsky. From Discrete Visco-Elasticity to Continuum Rate-Independent Plasticity: Rigorous Results. *Archive for Rational Mechanics and Analysis*, 203:577–619, February 2011.

[107] I. Benichou and S. Givli. Structures undergoing discrete phase transformation. *Journal of the Mechanics and Physics of Solids*, 61(1):94–113, January 2013.

[108] S. Liu, A. I. Azad, and R. Burgueño. Architected materials for tailorable shear behavior with energy dissipation. *Extreme Mechanics Letters*, 28:1–7, April 2019.

[109] A. F. Huxley and R. M. Simmons. Proposed Mechanism of Force Generation in Striated Muscle. *Nature*, 233(5321):533–538, October 1971.

[110] T. L. Hill. Theory of muscular contraction extended to groups of actin sites. *Proceedings of the National Academy of Sciences of the United States of America*, 70(10):2732–2736, October 1973.

[111] M. Caruel, J.-M. Allain, and L. Truskinovsky. Muscle as a Metamaterial Operating Near a Critical Point. *Physical Review Letters*, 110:248103, June 2013.

[112] M. Caruel and L. Truskinovsky. Statistical mechanics of the Huxley-Simmons model. *Physical Review E*, 93(6):062407, June 2016.

[113] S. Giordano. Spin variable approach for the statistical mechanics of folding and unfolding chains. *Soft Matter*, 13(38):6877–6893, October 2017.

[114] R. Law, P. Carl, S. Harper, P Dalhaimer, D.W. Speicher, and D. E. Discher. Cooperativity in forced unfolding of tandem spectrin repeats. *Biophysical Journal*, 84(1):533–544, January 2003.

[115] W. Bialek. Perspectives on theory at the interface of physics and biology. *Reports on Progress in Physics. Physical Society (Great Britain)*, 81(1):012601, January 2018.

[116] M. V. Vol'kenshteĭn. *Configurational Statistics of Polymeric Chains*. Interscience Publishers, 1963.

[117] H. Kleinert. *Path Integrals in Quantum Mechanics, Statistics, Polymer Physics, and Financial Markets*. World scientific, fifth edition, May 2009.

[118] R. H. Boyd and P. J. Phillips. *The Science of Polymer Molecules*. Cambridge University Press, October 1996.

[119] M. Doi. *Introduction to Polymer Physics*. Clarendon Press, 1996.

[120] J. H. Weiner. *Statistical Mechanics of Elasticity*. Courier Corporation, February 2012.

[121] M. Rubinstein and R. H. Colby. *Polymer Physics*. OUP Oxford, June 2003.

[122] R. G. Winkler. Deformation of semiflexible chains. *The Journal of Chemical Physics*, 118(6):2919–2928, January 2003.

[123] G. Glatting, R. G. Winkler, and P. Reineker. Partition function and force extension relation for a generalized freely jointed chain, May 2002.

[124] F. Manca, S. Giordano, P. L. Palla, R. Zucca, F. Cleri, and L. Colombo. Elasticity of flexible and semiflexible polymers with extensible bonds in the Gibbs and Helmholtz ensembles. *The Journal of Chemical Physics*, 136(15):154906, April 2012.

[125] F. Manca, S. Giordano, P. L. Palla, F. Cleri, and L. Colombo. Theory and Monte Carlo simulations for the stretching of flexible and semiflexible single polymer chains under external fields. *The Journal of Chemical Physics*, 137(24):244907, December 2012.

[126] T. Su and P. K. Purohit. Thermomechanics of a heterogeneous fluctuating chain. *Journal of the Mechanics and Physics of Solids*, 58(2):164–186, February 2010.

[127] J. Kierfeld, O. Niamploy, V. Sa-yakanit, and R. Lipowsky. Stretching of semiflexible polymers with elastic bonds. *The European Physical Journal E*, 14(1):17–34, May 2004.

[128] M. Rief, J. M. Fernandez, and H. E. Gaub. Elastically Coupled Two-Level Systems as a Model for Biopolymer Extensibility. *Physical Review Letters*, 81(21):4764–4767, November 1998.

[129] H. J. Kreuzer, S. H. Payne, and L. Livadaru. Stretching a Macromolecule in an Atomic Force Microscope: Statistical Mechanical Analysis. *Biophysical Journal*, 80(6):2505–2514, June 2001.

[130] F. Manca, S. Giordano, P. L. Palla, F. Cleri, and L. Colombo. Two-state theory of single-molecule stretching experiments. *Physical Review E*, 87(3):032705, March 2013.

[131] S. Giordano. Helmholtz and Gibbs ensembles, thermodynamic limit and bistability in polymer lattice models. *Continuum Mechanics and Thermodynamics*, 30(3):459–483, May 2018.

[132] M. S. Z. Kellermayer, S. B. Smith, H. L. Granzier, and C. Bustamante. Folding-Unfolding Transitions in Single Titin Molecules Characterized with Laser Tweezers. *Science*, 276(5315):1112–1116, May 1997.

[133] J. Wang, T. B. Kouznetsova, R. Boulatov, and S. L. Craig. Mechanical gating of a mechanochemical reaction cascade. *Nature Communications*, 7(1):13433, November 2016.

[134] A. Imparato, F. Sbrana, and M. Vassalli. Reconstructing the free-energy landscape of a polyprotein by single-molecule experiments. *EPL (Europhysics Letters)*, 82(5):58006, May 2008.

[135] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub. Reversible Unfolding of Individual Titin Immunoglobulin Domains by AFM. *Science*, 276(5315):1109–1112, May 1997.

[136] M. Rief, J. Pascual, M. Saraste, and H. E. Gaub. Single molecule force spectroscopy of spectrin repeats: Low unfolding forces in helix bundles. *Journal of Molecular Biology*, 286(2):553–561, February 1999.

[137] F. Manca, S. Giordano, P. L. Palla, F. Cleri, and L. Colombo. Monte Carlo simulations of single polymer force-extension relations. *Journal of Physics: Conference Series*, 383:012016, October 2012.

[138] F. Manca, S. Giordano, P. L. Palla, F. Cleri, and L. Colombo. Response to "Comment on 'Elasticity of flexible and semiflexible polymers with extensible bonds in the Gibbs and Helmholtz ensembles"' [J. Chem. Phys. 138, 157101 (2013)]. *The Journal of Chemical Physics*, 138(15):157102, April 2013.

[139] R. G. Winkler. Equivalence of statistical ensembles in stretching single flexible polymers. *Soft Matter*, 6(24):6183–6191, November 2010.

[140] A. E. B. Pupo, F. Falo, and A. Fiasconaro. DNA overstretching transition induced by melting in a dynamical mesoscopic model. *The Journal of Chemical Physics*, 139(9):095101, September 2013.

[141] F. Hanke and H. J. Kreuzer. Conformational transitions in single polymer molecules modeled with a complete energy landscape: Continuous two-state model. *The European Physical Journal E*, 22(2):163–169, February 2007.

[142] D. B. Staple, S. H. Payne, A. L. C. Reddin, and H. J. Kreuzer. Stretching and unfolding of multidomain biopolymers: A statistical mechanics theory of titin. *Physical Biology*, 6(2):025005, July 2009.

[143] A. Prados, A. Carpio, and L. L. Bonilla. Sawtooth patterns in force-extension curves of biomolecules: An equilibrium-statistical-mechanics theory. *Physical Review E*, 88(1):012704, July 2013.

[144] L. L. Bonilla, A. Carpio, and A. Prados. Theory of force-extension curves for modular proteins and DNA hairpins. *Physical Review E*, 91(5):052712, May 2015.

[145] D. De Tommasi, N. Millardi, G. Puglisi, and G. Saccomandi. An energetic model for macromolecules unfolding in stretching experiments. *Journal of The Royal Society Interface*, 10(88):20130651, November 2013.

[146] D. E. Makarov. A Theoretical Model for the Mechanical Unfolding of Repeat Proteins. *Biophysical Journal*, 96(6):2160–2167, March 2009.

[147] C. Guardiani and F. Bagnoli. A toy model of polymer stretching. *The Journal of Chemical Physics*, 125(8):084908, August 2006.

[148] G. Puglisi and L. Truskinovsky. Mechanics of a discrete chain with bi-stable elements. *Journal of the Mechanics and Physics of Solids*, 48(1):1–27, January 2000.

[149] I. Benichou and S. Givli. Rate Dependent Response of Nanoscale Structures Having a Multiwell Energy Landscape. *Physical Review Letters*, 114(9):095504, March 2015.

[150] I. Benichou, Y. Zhang, O. K. Dudko, and S. Givli. The rate dependent response of a bistable chain at finite temperature. *Journal of the Mechanics and Physics of Solids*, 95:44–63, October 2016.

[151] I. D. Vladescu, M. J. McCauley, I. Rouzina, and M. C. Williams. Mapping the Phase Diagram of Single DNA Molecule Force-Induced Melting in the Presence of Ethidium. *Physical Review Letters*, 95(15):158102, October 2005.

[152] C. A. Plata, F. Cecconi, M. Chinappi, and A. Prados. Understanding the dependence on the pulling speed of the unfolding pathway of proteins. *Journal of Statistical Mechanics: Theory and Experiment*, 2015(8):P08003, August 2015.

[153] P. Malhotra and J. B. Udgaonkar. How cooperative are protein folding and unfolding transitions? *Protein Science: A Publication of the Protein Society*, 25(11):1924–1941, November 2016.

[154] T. Xu, H. Lannon, S. Wolf, F. Nakamura, and J. Brujic. Domain-Domain Interactions in Filamin A (16–23) Impose a Hierarchy of Unfolding Forces. *Biophysical Journal*, 104(9):2022–2030, May 2013.

[155] D. Barrick, D. U. Ferreiro, and E. A. Komives. Folding landscapes of ankyrin repeat proteins: Experiments meet theory. *Current Opinion in Structural Biology*, 18(1):27–34, February 2008.

[156] P. Bechtluft, R. G. H. van Leeuwen, M. Tyreman, D. Tomkiewicz, N. Nouwen, H. L Tepper, A. J. M. Driessen, and S. J. Tans. Direct Observation of Chaperone-Induced Changes in a Protein Folding Pathway. *Science*, 318(5855):1458–1461, November 2007.

[157] A. Mashaghi, G. Kramer, P. Bechtluft, B. Zachmann-Brand, A. J. M. Driessen, B. Bukau, and S. J. Tans. Reshaping of the conformational search of a protein by the chaperone trigger factor. *Nature*, 500(7460):98–101, August 2013.

[158] J. W. Gibbs. *Elementary Principles in Statistical Mechanics Developed with Especial Reference to the Rational Foundation of Thermodynamics.* C. Scribner, 1902.

[159] G. I. Bell. Models for the specific adhesion of cells to cells. *Science*, 200(4342):618–627, May 1978.

[160] Lord Rayleigh O. M. F.R.S. XXXI. On the problem of random vibrations, and of random flights in one, two, or three dimensions. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 37(220):321–347, April 1919.

[161] G. Polya. Berechnung eines bestimmten Integrals. *Mathematische Annalen*, 74(2):204–212, June 1913.

[162] L. R. G. Treloar. The statistical length of long-chain molecules. *Transactions of the Faraday Society*, 42(0):77–82, January 1946.

[163] M. C. Wang and E. Guth. Statistical Theory of Networks of Non-Gaussian Flexible Chains. *The Journal of Chemical Physics*, 20(7):1144–1157, July 1952.

[164] M. Radiom, P. Maroni, and M. Borkovec. Influence of Solvent Quality on the Force Response of Individual Poly(styrene) Polymer Chains. *ACS Macro Letters*, 6(10):1052–1055, October 2017.

[165] K. Wang, X. Pang, and S. Cui. Inherent Stretching Elasticity of a Single Polymer Chain with a Carbon–Carbon Backbone. *Langmuir*, 29(13):4315–4319, April 2013.

[166] C. Friedsam, H. E. Gaub, and R. R. Netz. Probing surfaces with single-polymer atomic force microscope experiments. *Biointerphases*, 1(1):MR1–MR21, March 2006.

[167] 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*, 94(4):048301, January 2005.

[168] M. Radiom and M. Borkovec. Influence of ligand-receptor interactions on force-extension behavior within the freely jointed chain model. *Physical Review E*, 96(6):062501, December 2017.

[169] N. K. Balabaev and T. N. Khazanovich. Extension of chains composed of freely joined elastic segments. *Russian Journal of Physical Chemistry B*, 3(2):242–246, April 2009.

[170] I. S. Gradshteyn and I. M. Ryzhik. *Table of Integrals, Series, and Products*. Academic Press, New York, 1965.

[171] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. U.S. Government Printing Office, 1948.

[172] Y. Hu, T. Lindstrøm, B. Øksendal, J. Ubøe, and T. Zhang. Inverse Powers of White Noise. 1993.

[173] A. A. Zadpoor. Mechanical meta-materials. *Materials Horizons*, 3(5):371–381, 2016.

[174] A. Taroni. 90 years of the Ising model. *Nature Physics*, 11(12):997–997, December 2015.

[175] R. J. Baxter. The inversion relation method for some two-dimensional exactly solved models in lattice statistics. *Journal of Statistical Physics*, 28(1):1–41, May 1982.

[176] A. Perez-Riba, M. Synakewicz, and L. S. Itzhaki. Folding cooperativity and allosteric function in the tandem-repeat protein class. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 373(1749), June 2018.

[177] M. Caruel and L. Truskinovsky. Bi-stability resistant to fluctuations. *Journal of the Mechanics and Physics of Solids*, 109:117–141, December 2017.

[178] L. Truskinovsky and A. Vainchtein. The origin of nucleation peak in transformational plasticity. *Journal of the Mechanics and Physics of Solids*, 52(6):1421–1446, June 2004.

[179] G. Puglisi. Hysteresis in multi-stable lattices with non-local interactions. *Journal of the Mechanics and Physics of Solids*, 54(10):2060–2088, October 2006.

[180] M. Mickler, R. I. Dima, H. Dietz, C. Hyeon, D. Thirumalai, and M. Rief. Revealing the bifurcation in the unfolding pathways of GFP by using single-molecule experiments and simulations. *Proceedings of the National Academy of Sciences*, 104(51):20268–20273, December 2007.

[181] E. Paci and M. Karplus. Unfolding proteins by external forces and temperature: The importance of topology and energetics. *Proceedings of the National Academy of Sciences*, 97(12):6521–6526, June 2000.

[182] K.R. Chaurasiya, T. Paramanathan, M.J. McCauley, and M.C. Williams. Biophysical characterization of DNA binding from single molecule force measurements. *Physics of Life Reviews*, 7(3):299–341, June 2010.

[183] J. Zakrisson, K. Wiklund, M. Servin, O. Axner, C. Lacoursière, and M. Andersson. Rigid multibody simulation of a helix-like structure: The dynamics of bacterial adhesion pili. *European biophysics journal: EBJ*, 44(5):291–300, July 2015.

[184] H. W. Turnbull. *Theory of Equations*. Oliver and Boyd, 4d ed., rev edition edition, January 1947.

[185] H. W. Gould. *The Girard-Waring Power Sum Formulas for Symmetric Functions and Fibonacci Series*. May, 1997.

[186] D. G. Mead. Newton's Identities. *The American Mathematical Monthly*, 99(8):749–751, October 1992.

[187] D. Kalman. A Matrix Proof of Newton's Identities. *Mathematics Magazine*, 73(4):313–315, 2000.

[188] S.W. Englander and L. Mayne. The case for defined protein folding pathways. *Proceedings of the National Academy of Sciences of the United States of America*, 114(31):8253–8258, June 2017.

[189] M. Yang, B. Yordanov, Y. Levy, R. Brüschweiler, and S. Huo. The sequence-dependent unfolding pathway plays a critical role in the amyloidogenicity of transthyretin. *Biochemistry*, 45(39):11992–12002, October 2006.

[190] A. Yadav, S. Paul, R. Venkatramani, and S. R. K. Ainavarapu. Differences in the mechanical unfolding pathways of apo- and copper-bound azurins. *Scientific Reports*, 8(1):1989, January 2018.

[191] H. Li, A. F. Oberhauser, S. B. Fowler, J. Clarke, and J. M. Fernandez. Atomic force microscopy reveals the mechanical design of a modular protein. *Proceedings of the National Academy of Sciences*, 97(12):6527–6531, June 2000.

[192] M. Bonin, R. Zhu, Y. Klaue, J. Oberstrass, E. Oesterschulze, and W. Nellen. Analysis of RNA flexibility by scanning force spectroscopy. *Nucleic Acids Research*, 30(16):e81, August 2002.

[193] J. Lipfert, G. M. Skinner, J. M. Keegstra, T. Hensgens, T. Jager, D. Dulin, M. Köber, Z. Yu, S. P. Donkers, F.-C. Chou, R. Das, and N. H. Dekker. Double-stranded RNA under force and torque: Similarities to and striking differences from double-stranded DNA. *Proceedings of the National Academy of Sciences*, 111(43):15408–15413, October 2014.

[194] H. Li, M. Rief, P. Oesterhelt, H. E. Gaub, X. Zhang, and J. Shen. Single-molecule force spectroscopy on polysaccharides by AFM – nanomechanical fingerprint of α-(1,4)-linked polysaccharides. *Chemical Physics Letters*, 305(3):197–201, May 1999.

[195] Grégory Francius, David Alsteens, Vincent Dupres, Sarah Lebeer, Sigrid De Keersmaecker, Jos Vanderleyden, Hermann J. Gruber, and Yves F. Dufrêne. Stretching polysaccharides on live cells using single molecule force spectroscopy. *Nature Protocols*, 4(6):939–946, 2009.

[196] M. Schlierf and M. Rief. Single-Molecule Unfolding Force Distributions Reveal a Funnel-Shaped Energy Landscape. *Biophysical Journal*, 90(4):L33–L35, February 2006.

[197] M. Schlierf and M. Rief. Temperature Softening of a Protein in Single-molecule Experiments. *Journal of Molecular Biology*, 354(2):497–503, November 2005.

[198] M. Carrion-Vazquez, A. F. Oberhauser, S. B. Fowler, P. E. Marszalek, S. E. Broedel, J. Clarke, and J. M. Fernandez. Mechanical and chemical unfolding of a single protein: A comparison. *Proceedings of the National Academy of Sciences of the United States of America*, 96(7):3694–3699, March 1999.

[199] H. Lei, Y. Guo, X. Hu, C. Hu, X. Hu, and H. Li. Reversible Unfolding and Folding of the Metalloprotein Ferredoxin Revealed by Single-Molecule Atomic Force Microscopy. *Journal of the American Chemical Society*, 139(4):1538–1544, January 2017.

[200] J. Xia, J. Zuo, and H. Li. Single molecule force spectroscopy reveals that the oxidation state of cobalt ions plays an important role in enhancing the mechanical stability of proteins. *Nanoscale*, 11(42):19791–19796, November 2019.

[201] K. Klamecka, P. M. Severin, L. F. Milles, H. E. Gaub, and H. Leonhardt. Energy profile of nanobody-GFP complex under force. *Physical Biology*, 12(5):056009, September 2015.

[202] W. T. King, M. Su, and G. Yang. Monte Carlo simulation of mechanical unfolding of proteins based on a simple two-state model. *International Journal of Biological Macromolecules*, 46(2):159–166, March 2010.

[203] G. Hummer and A. Szabo. Kinetics from Nonequilibrium Single-Molecule Pulling Experiments. *Biophysical Journal*, 85(1):5–15, July 2003.

[204] E. Evans and K. Ritchie. Dynamic strength of molecular adhesion bonds. *Biophysical Journal*, 72(4):1541–1555, April 1997.

[205] O. K. Dudko. Decoding the mechanical fingerprints of biomolecules. *Quarterly Reviews of Biophysics*, 49:e3, January 2016.

[206] O. K. Dudko, A. E. Filippov, J. Klafter, and M. Urbakh. Beyond the conventional description of dynamic force spectroscopy of adhesion bonds. *Proceedings of the National Academy of Sciences*, 100(20):11378–11381, September 2003.

[207] O. K. Dudko, G. Hummer, and A. Szabo. Intrinsic Rates and Activation Free Energies from Single-Molecule Pulling Experiments. *Physical Review Letters*, 96(10):108101, March 2006.

[208] O. K. Dudko, G. Hummer, and A. Szabo. Theory, analysis, and interpretation of single-molecule force spectroscopy experiments. *Proceedings of the National Academy of Sciences of the United States of America*, 105(41):15755–15760, October 2008.

[209] C. A. Pierse and O. K. Dudko. Kinetics and Energetics of Biomolecular Folding and Binding. *Biophysical Journal*, 105(9):L19–L22, November 2013.

[210] F. Manca, P.-M. Déjardin, and S. Giordano. Statistical mechanics of holonomic systems as a Brownian motion on smooth manifolds. *Annalen der Physik*, 528(5):381–393, May 2016.

[211] S. Giordano. Stochastic thermodynamics of holonomic systems. *The European Physical Journal B*, 92(8):174, August 2019.

[212] H. Risken. *The Fokker-Planck Equation: Methods of Solution and Applications*. Springer Verlag, Berlin ; New York, April 1989.

[213] W. T. Coffey, Y. P. Kalmykov, and J. T. Waldron. *The Langevin Equation: With Applications to Stochastic Problems in Physics, Chemistry and Electrical Engineering*, volume 14 of *World Scientific Series in Contemporary Chemical Physics*. World scientific, second edition, March 2004.

[214] C. A. Plata and A. Prados. Modelling the unfolding pathway of biomolecules: Theoretical approach and experimental prospect. *arXiv:1706.02475 [cond-mat, physics:physics, q-bio]*, 232:137–151, 2018.

[215] G. E. P. Box and M. E. Muller. A Note on the Generation of Random Normal Deviates. *Annals of Mathematical Statistics*, 29(2):610–611, June 1958.

[216] Z. Schuss. *Theory and Applications of Stochastic Processes: An Analytical Approach*. Applied Mathematical Sciences. Springer-Verlag, New York, 2010.

[217] G. Florio and G. Puglisi. Unveiling the influence of device stiffness in single macromolecule unfolding. *Scientific Reports*, 9(1):4997, March 2019.

[218] L. Bellino, G. Florio, and G. Puglisi. The influence of device handles in single-molecule experiments. *Soft Matter*, 15(43):8680–8690, November 2019.

[219] S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, and K. Schulten. Molecular dynamics study of unbinding of the avidin-biotin complex. *Biophysical Journal*, 72(4):1568–1581, April 1997.

[220] P. Fucini, C. Renner, C. Herberhold, A. A. Noegel, and T. A. Holak. The repeating segments of the F-actin cross-linking gelation factor (ABP-120) have an immunoglobulin-like fold. *Nature Structural Biology*, 4(3):223–230, March 1997.

[221] G. M. Popowicz, R. Müller, A. A. Noegel, M. Schleicher, R. Huber, and T. A. Holak. Molecular structure of the rod domain of dictyostelium filamin. *Journal of Molecular Biology*, 342(5):1637–1646, October 2004.

[222] I. Schwaiger, M. Schleicher, A. A. Noegel, and M. Rief. The folding pathway of a fast-folding immunoglobulin domain revealed by single-molecule mechanical experiments. *EMBO Reports*, 6(1):46–51, January 2005.

[223] H. Chen, X. Zhu, P. Cong, M. P. Sheetz, F. Nakamura, and J. Yan. Differential Mechanical Stability of Filamin A Rod Segments. *Biophysical Journal*, 101(5):1231–1237, September 2011.

[224] H. Chen, S. Chandrasekar, M. P. Sheetz, T. P. Stossel, F. Nakamura, and J. Yan. Mechanical perturbation of filamin A immunoglobulin repeats 20-21 reveals potential non-equilibrium mechanochemical partner binding function. *Scientific Reports*, 3, April 2013.

[225] S. Improta, A. S. Politou, and A. Pastore. Immunoglobulin-like modules from titin I-band: Extensible components of muscle elasticity. *Structure*, 4(3):323–337, March 1996.

[226] E. C. Eckels, R. Tapia-Rojo, J. A. Rivas-Pardo, and J. M. Fernández. The work of titin protein folding as a major driver in muscle contraction. *Annual review of physiology*, 80:327–351, February 2018.

[227] H. Chen, G. Yuan, R. S. Winardhi, M. Yao, I. Popa, J. M. Fernandez, and J. Yan. Dynamics of Equilibrium Folding and Unfolding Transitions of Titin Immunoglobulin Domain under Constant Forces. *Journal of the American Chemical Society*, 137(10):3540–3546, March 2015.

[228] G. Yuan, S. Le, M. Yao, H. Qian, X. Zhou, J. Yan, and H. Chen. Elasticity of the Transition State Leading to an Unexpected Mechanical Stabilization of Titin Immunoglobulin Domains. *Angewandte Chemie (International ed. in English)*, 56(20):5490–5493, April 2017.

[229] M. S. Li and M. Kouza. Dependence of protein mechanical unfolding pathways on pulling speeds. *The Journal of Chemical Physics*, 130(14):145102, April 2009. Comment: 26 pages, 9 figures, 1 table.

[230] C. Guardiani, D. Di Marino, A. Tramontano, M. Chinappi, and F. Cecconi. Exploring the Unfolding Pathway of Maltose Binding Proteins: An Integrated Computational Approach. *Journal of Chemical Theory and Computation*, 10(9):3589–3597, September 2014.

[231] C. A. Plata, Z. N. Scholl, P. E. Marszalek, and A. Prados. Relevance of the Speed and Direction of Pulling in Simple Modular Proteins. *Journal of Chemical Theory and Computation*, 14(6):2910–2918, June 2018.

[232] E. L. Florin, V. T. Moy, and H. E. Gaub. Adhesion forces between individual ligand-receptor pairs. *Science*, 264(5157):415–417, April 1994.

[233] V. Barsegov, D. K. Klimov, and D. Thirumalai. Mapping the Energy Landscape of Biomolecules Using Single Molecule Force Correlation Spectroscopy: Theory and Applications. *Biophysical Journal*, 90(11):3827–3841, June 2006.

[234] J. W. Weisel, H. Shuman, and R. Litvinov. Protein-protein unbinding induced by force: Single-molecule studies. *Current Opinion in Structural Biology*, 13(2):227–235, April 2003.

[235] A. Fuhrmann and A. J. Engler. The Cytoskeleton Regulates Cell Attachment Strength. *Biophysical Journal*, 109(1):57, July 2015.

[236] J. Helenius, C.-P. Heisenberg, H. E. Gaub, and D. J. Muller. Single-cell force spectroscopy. *Journal of Cell Science*, 121(11):1785–1791, June 2008.

[237] A. Beaussart, S. El-Kirat-Chatel, R. M. A. Sullan, D. Alsteens, P. Herman, S. Derclaye, and Y. F. Dufrêne. Quantifying the forces guiding microbial cell adhesion using single-cell force spectroscopy. *Nature Protocols*, 9(5):1049–1055, May 2014.

[238] X. Cao, E. Ban, B. M. Baker, Y. Lin, J. A. Burdick, C. S. Chen, and V. B. Shenoy. Multiscale model predicts increasing focal adhesion size with decreasing stiffness in fibrous matrices, June 2017.

[239] W. Mason. The Lüders' Lines on Mild Steel. *Proceedings of the Physical Society of London*, 23(1):305–333, December 1910.

[240] A. Yilmaz. The Portevin–Le Chatelier effect: A review of experimental findings. *Science and Technology of Advanced Materials*, 12(6):063001, December 2011.

[241] A. A. Shibkov, A. E. Zolotov, M. A. Zheltov, and A. A. Denisov. Morphological diagram of Savart-Masson bands of macrolocalized deformation. *Crystallography Reports*, 57(1):105–111, January 2012.

[242] F. Chmelík, A. Ziegenbein, H. Neuhäuser, and P. Lukac. Investigating the Portevin–Le Châtelier effect by the acoustic emission and laser extensometry techniques. *Materials Science and Engineering: A*, 324:200–207, February 2002.

[243] P.-P. Cortet, M. Ciccotti, and L. Vanel. Imaging the stick–slip peeling of an adhesive tape under a constant load. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(03):P03005, March 2007.

[244] V. Barreau, D. Yu, R. Hensel, and E. Arzt. Elevated temperature adhesion of bioinspired polymeric micropatterns to glass. *Journal of the Mechanical Behavior of Biomedical Materials*, 76:110–118, December 2017.

[245] G. Schoeck. Thermally activated crack-propagation in brittle materials. *International Journal of Fracture*, 44(1):1–14, July 1990.

[246] A. Cochard, O. Lengliné, K. J. Måløy, and R. Toussaint. Thermally activated crack fronts propagating in pinning disorder: Simultaneous brittle/creep behaviour depending on scale. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 377(2136):20170399, January 2019.

[247] M. Peyrard and A. R. Bishop. Statistical mechanics of a nonlinear model for DNA denaturation. *Physical Review Letters*, 62(23):2755–2758, June 1989.

[248] N. Theodorakopoulos, T. Dauxois, and M. Peyrard. Order of the Phase Transition in Models of DNA Thermal Denaturation. *Physical Review Letters*, 85(1):6–9, July 2000.

[249] J. Palmeri, M. Manghi, and N. Destainville. Thermal denaturation of fluctuating DNA driven by bending entropy. *Physical Review Letters*, 99(8):088103, August 2007.

[250] J. Palmeri, M. Manghi, and N. Destainville. Thermal denaturation of fluctuating finite DNA chains: The role of bending rigidity in bubble nucleation. *Physical Review. E, Statistical, Nonlinear, and Soft Matter Physics*, 77(1 Pt 1):011913, January 2008.

[251] G. E. Fantner, E. Oroudjev, G. Schitter, L. S. Golde, P. Thurner, M. M. Finch, P. Turner, T. Gutsmann, D. E. Morse, H. Hansma, and P. K. Hansma. Sacrificial Bonds and Hidden Length: Unraveling Molecular Mesostructures in Tough Materials. *Biophysical Journal*, 90(4):1411–1418, February 2006.

[252] S. S. Brenner. Mechanical Behavior of Sapphire Whiskers at Elevated Temperatures. *Journal of Applied Physics*, 33(1):33–39, January 1962.

Statistical mechanics and thermodynamics of systems with conformational transitions: applications to biological macromolecules

The thermo-elastic behaviour of biological macromolecules, such as DNA and proteins, is addressed in a wide field of studies, with plenty of theoretical and practical applications. Mathematical modelling of the macromolecules' response to deformation and to thermal fluctuations allows the validation of the statistical mechanics of small systems, through the comparison with experimental results obtained by force spectroscopy, thus providing information on the static and dynamic responses induced by applied forces. These analyses are even more important for bistable macromolecules with conformational transitions, corresponding to folding/unfolding processes between two stable positions. In order to obtain analytically the force-deformation response of a chain composed of bistable units, it is necessary to calculate the partition functions, which are essential objects of the statistical mechanics. Thus, the bistable potential energy is decomposed into two parabolas, both corresponding to the folded and unfolded states and identified using the spin variables technique, largely used throughout the manuscript. The first part of this thesis concerns the extensibility of the bonds between the bistable units of a chain. The second part allows us to consider the interactions between the bistable units of a chain, thanks to the Ising model. The third part deals with heterogeneity, an important parameter to determine the unfolding sequence in proteins unfolding. Finally, in the last part, the dynamics of deformation is described.

Keywords thermomechanics of macromolecules, statistical mechanics, chains of bistable units, statistical ensembles of Gibbs and Helmholtz, models for single-molecule force spectroscopy

Mécanique statistique et thermodynamique des systèmes avec transitions conformationnelles : applications aux macromolécules biologiques

Le comportement thermoélastique des macromolécules biologiques, telles que l'ADN et les protéines, est abordé dans un vaste champ d'études, aux nombreuses applications théoriques et pratiques. La modélisation mathématique de la réponse des macromolécules à la déformation et aux fluctuations thermiques permet de tester la validité de la mécanique statistique des petits systèmes, grâce à la comparaison avec les résultats expérimentaux obtenus par la spectroscopie de force, qui fournit des informations sur les réponses statiques et dynamiques induites par les forces appliquées. Ces analyses sont encore plus importantes pour les macromolécules bistables présentant des transitions conformationnelles, correspondant à des processus de pliage/dépliage entre deux positions stables. Pour obtenir analytiquement la réponse force-déformation d'une chaîne composée d'unités bistables, il est nécessaire de calculer les fonctions de partition, essentielles en mécanique statistique. Ainsi, l'énergie potentielle bistable est décomposée en deux paraboles, correspondant l'une et l'autre aux états plié et déplié et identifiées grâce à la technique des variables de spin, utilisée tout au long du manuscrit. La première partie de cette thèse concerne l'extensibilité des liens entre les unités bistables d'une chaîne. La seconde partie permet, quant à elle, de considérer les interactions entre les unités bistables d'une chaîne, grâce au modèle d'Ising. Dans la troisième partie, il est question de l'hétérogénéité, paramètre important pour déterminer la séquence de dépliage dans le repliement des protéines. Enfin, dans la dernière partie, la dynamique de la déformation est décrite.

Mots-clés thermomécanique des macromolécules, mécanique statistique, chaînes d'unités bistables, ensembles statistiques de Gibbs et Helmholtz, modèles pour la spectroscopie de force à molécule unique