



ELSEVIER

Contents lists available at ScienceDirect

## Physics Reports

journal homepage: [www.elsevier.com/locate/physrep](http://www.elsevier.com/locate/physrep)

## Biomolecules under mechanical force

Sanjay Kumar<sup>a,\*</sup>, Mai Suan Li<sup>b,\*</sup><sup>a</sup> Department of Physics, Banaras Hindu University, Varanasi 221 005, India<sup>b</sup> Institute of Physics, Polish Academy of Sciences, Al. Lotników 32/46, 02-668 Warsaw, Poland

## ARTICLE INFO

## Article history:

Accepted 17 September 2009

Available online xxxx

editor: H. Orland

## Keywords:

Single-molecule force spectroscopy

DNA

Protein

Computer simulations

## ABSTRACT

Recent advances in single molecule experiments have raised many challenges. These challenges can be met by a proper understanding of the inter- and intra-molecular interactions in the framework of physics followed by suitable theoretical models substantiated by extensive numerical simulations. In this review, we briefly discuss experimental, theoretical and numerical techniques used to examine the dynamics of biomolecules under the application of external mechanical force. We focus on issues which require special attention: the relationship between the mechanical stability of a protein and the secondary structure of its native conformation, the dependence of the free energy landscape on the pulling direction and secondary structure content, unfolding of a protein through a pore, protein re-folding under quenched force etc. We pay special attention to a single stranded DNA, where the force-extension curve shows a multi-step plateau for the chain made up of adenine while poly-thymine exhibits entropic response only. There are many interesting predictions related to unzipping of double stranded DNA e.g. re-entrance in the force-temperature phase diagram, existence of an "Eye-phase", effects of random disorder etc. based on model studies which require further investigation. We will also discuss cases where the theoretical descriptions of the models fail to explain the experimentally observed behavior and when further refinement is needed in order to describe the outcomes of experiments. Finally we suggest certain experimental protocols to observe theoretical predictions *in vitro*.

© 2009 Elsevier B.V. All rights reserved.

## Contents

1. Introduction.....	3
2. Models .....	3
2.1. Continuum models of homopolymers.....	4
2.1.1. Excluded volume interaction: Self-avoiding chain model.....	6
2.1.2. Interactions that matter .....	7
2.1.3. Lattice model.....	8
2.2. Modeling of dsDNA.....	9
2.2.1. Poland–Scheraga (PS) model .....	9
2.2.2. Peyrard–Bishop model .....	11
2.2.3. Semi-microscopic model.....	12
2.2.4. Lattice models.....	13
2.3. Modeling of proteins .....	13
2.3.1. Lattice models.....	14
2.3.2. Off-lattice coarse-grained Go modeling.....	14

\* Corresponding author.

E-mail addresses: [yashankit@yahoo.com](mailto:yashankit@yahoo.com) (S. Kumar), [masli@ifpan.edu.pl](mailto:masli@ifpan.edu.pl) (M.S. Li).

2.3.3.	All-atom models .....	15
3.	Methods .....	15
3.1.	Generating function technique .....	15
3.2.	Exact enumeration technique .....	16
3.3.	Transfer integral technique .....	17
3.4.	Molecular dynamics .....	17
3.4.1.	Langevin dynamics simulation .....	18
3.4.2.	Brownian dynamics .....	18
3.5.	Kinetic theory .....	18
3.5.1.	Bell approximation: Constant force case .....	19
3.5.2.	Bell approximation: Force ramp case .....	19
3.5.3.	Beyond the Bell approximation .....	20
3.6.	Constant force and constant distance ensemble .....	21
4.	Experimental techniques .....	22
4.1.	Atomic force microscopy .....	22
4.2.	Laser optical tweezers .....	22
5.	Mechanical unfolding of homopolymers .....	23
5.1.	Reentrance transition .....	23
5.2.	Emergence of intermediates .....	25
5.3.	Role of stiffness .....	26
5.4.	Nature of transition .....	27
5.5.	Dependence of phase diagrams on pulling direction .....	28
6.	Mechanics of DNA .....	31
6.1.	DNA stretching: Double stranded DNA .....	31
6.2.	Single stranded DNA under tension .....	33
6.3.	Effects of temperature .....	35
6.4.	Evidence for structural transitions in ssDNA .....	35
6.5.	Unzipping of dsDNA .....	35
6.6.	Unzipping DNA from a globule state .....	40
6.7.	Rotational drag on DNA .....	40
6.8.	DNA twisting .....	42
7.	Protein unfolding .....	43
7.1.	Equilibrium force–temperature phase diagram of proteins .....	43
7.2.	Protein mechanical unfolding via intermediates .....	46
7.3.	Mechanical unfolding pathways are different from thermal and chemical ones .....	48
7.4.	Mechanical resistance of proteins: Correlation between unfolding force and secondary structure content .....	48
7.4.1.	Definition of mechanical stability of biomolecules .....	48
7.4.2.	Validity of Go models for estimating $f_u$ .....	48
7.4.3.	Unfolding force, secondary structure content, and contact order .....	49
7.5.	Free energy landscape in the Bell approximation .....	51
7.5.1.	Validity of Go models for estimation of $x_u$ .....	51
7.5.2.	Correlation between $x_u$ and helix content .....	52
7.5.3.	Correlation between $x_u$ and contact order .....	52
7.5.4.	Correlation between $f_u$ and $x_u$ .....	53
7.6.	Effect of pulling geometry .....	54
7.6.1.	Influence of pulling geometry on mechanical resistance .....	54
7.6.2.	Dependence of the free energy landscape on pulling positions .....	55
7.6.3.	Prediction of protein structures by mechanical pulling .....	56
7.7.	Free energy landscapes beyond the Bell approximation .....	56
7.8.	Other developments .....	58
7.8.1.	Construction of free energy landscape using Jarzynski's equality .....	58
7.8.2.	AC modulated force spectroscopy .....	58
7.8.3.	Unfolding of proteins pulled mechanically through a pore .....	58
7.8.4.	Crowding effect .....	59
7.9.	Summary of main factors governing mechanical stability of proteins .....	59
8.	Refolding under quenched force .....	59
8.1.	Proteins refold in stages .....	60
8.2.	Deciphering the free energy landscape using force-clamp measurements .....	61
8.3.	Can force-clamp spectroscopy probe folding pathways of a protein? .....	62
9.	What next? .....	63
10.	Conclusion .....	65
	Acknowledgements .....	65
	Appendix A. List of abbreviations .....	65
	Appendix B. List of symbols .....	66
	References .....	66

## 1. Introduction

Macromolecules from living organisms vary in their size, shape and function. Some of these are also called “biopolymers” because they are made up of long chains of “monomer” units. The monomers can be of different types. In DNA and RNA the constituent monomers are nucleotides while in proteins they are amino acids. DNA, RNA and proteins may be considered as fundamental building blocks of the cell. For example DNA stores the genetic code of a cell. The transfer of code describes the normal flow of biological information: DNA can be copied to DNA (DNA replication), DNA information can be copied into mRNA (transcription) and proteins can be synthesized using the information in mRNA as a template (translation) [1]. Once a protein is formed it acquires its native conformation and many biological functions depend on the specific properties of this native state. The process of transferring information from DNA to protein involves many complicated interactions and pathways. The last ten years have witnessed intense activity using Single Molecule Force Spectroscopy (SMFS) experiments in order to understand these pathways by manipulating single biomolecules. Such experiments have also resulted in measurements of inter- and intra-molecular forces in biological systems and hence enhanced our understanding of the role they play in determining the functions and structures of biomolecules. Following the seminal papers by Rief et al. [2] and Tskhovrebova et al. [3] much of the research has been focused on the elastic properties of proteins, DNA and RNA, i.e., their response to an external force.

The main advantage of SMFS is its ability to separate out the fluctuations of individual molecules from the ensemble average behavior observed in traditional bulk biochemical experiments. Using SMFS one can measure detailed distributions, describing certain molecular properties (for example, the distribution of unfolding forces of biomolecules [2]) and observe possible intermediates in chemical reactions. SMFS studies have provided unexpected insights into the strength of the forces driving biological processes as well as determined various biological interactions responsible for the mechanical stability of biological structures. Given the experimental setup of SMFS, it was originally thought that the interactions detected would be mostly of a mechanical nature. However, with the increasing number of experiments and insights gathered so far, it has become clear that SMFS allows one to detect molecular interactions, which arise from many different physical and chemical origins. A major concern of present efforts is to understand the role of these interactions in the stability of biomolecules. One would also like to know whether all interactions contribute at the same moment or do they have different life times (exchange rates)? How do these interactions follow each other thereby forming an interaction pathway?

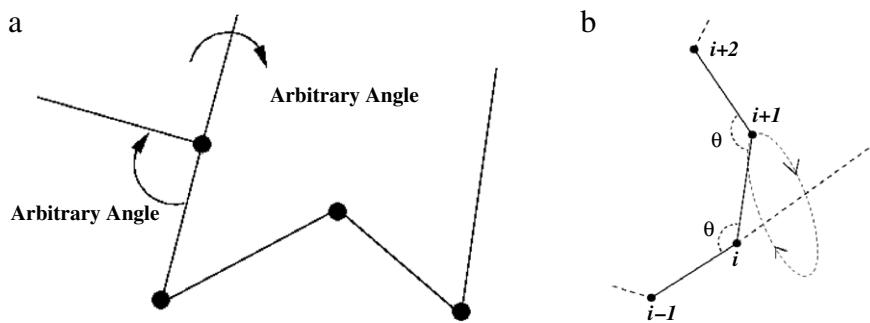
During recent years a number of excellent reviews on mechanical unfolding of biomolecules has appeared. However, some of them are relatively old [4–7] and do not reflect the latest developments in this fast growing field. Several surveys are only focused on specific problems like the free energy landscape (FEL) of biomolecules [8], the protein mechanical stability [9], experimental aspects of refolding under quenched force [10], advances in optical tweezers [11] etc. Moreover, most reviews are addressed to experts rather than to beginners. Therefore, the main goal of this review is to discuss the behavior of biomolecules subjected to an external mechanical force in the wider context compared to previous reviews. We do provide the essential polymer physics fundamentals, which have been used extensively by the scientific community and may be useful for beginners in this field.

The selection of topics for this review has been dictated by our own interests and expertise. We concentrate on mechanical properties of polymers, DNA and proteins. Those who are interested in RNA, one of the most important biomolecules, are recommended to read other surveys [7,12]. This review will be guided by the outcome of current experimental SMFS setups and the resulting issues which are of most concern from a physics understanding point of view. The physical background material will have a focus towards the understanding of molecular interactions contributing to an overall force and on current models describing these SMFS experiments. Moreover, we shall focus our discussions on those issues, which have not been discussed in previous reviews [6–10]. For example, the relationship between the FEL of proteins and their secondary structures, refolding under a weak quenched force, effects of the pulling direction on unfolding properties of biomolecules, the experimental construction of free energy landscape using Jarzynski’s equality [13], AC modulated force spectroscopy, unfolding of proteins pulled mechanically through a pore, re-entrance in force–temperature phase diagrams, structural transitions in ssDNA, effects of temperature and its consequences on the stretched state of biopolymers, and the role of bubbles (Eye phase) during unzipping of dsDNA etc. We shall also speculate force induced transitions *in vivo*, where molecular crowding and random force may influence these processes.

The article is organized as follows. In Section 2 we present the main models which have been used to study DNA and protein folding. Some exact techniques for studying lattice polymer models, molecular dynamics simulation methods and the kinetic theory for deciphering the free energy landscape of biomolecules are discussed in Section 3. Section 4 contains a brief discussion of experimental single molecule techniques. Section 5 is devoted to the mechanical unfolding of homopolymers. The mechanics of DNA and protein unfolding is reviewed in Sections 6 and 7, respectively. Section 8 describes recent progress in the study of protein refolding under a quenched force. Some questions which are interesting for future work are mentioned in Section 9.

## 2. Models

Biomolecules are polymers and quite frequently, physicists attempt to simplify their complex structures as much as possible and think of polymer chains merely as threads or necklaces made up of beads on a string. In some cases such simplified models can help to describe some essential properties of biopolymers but not always. This is because there are



**Fig. 1.** (a) Schematic of a freely jointed chain. (b) Freely rotating chain with fixed bond angle. These models do not incorporate excluded volume effects in their description.

big differences between them and polymers. Unlike heteropolymers, amino acid sequences of proteins are not random but rather designed by Mother Nature. Consequently, proteins are more structured having secondary structures such as  $\beta$ -strands and  $\alpha$ -helices. Chemically, DNA consists of two long polymers of simple units called nucleotides, with backbones made of sugars and phosphate groups joined by ester bonds. These two strands run in opposite directions to each other and are therefore anti-parallel. Careful modeling of biomolecules should take into account their specific features that make them different from polymers.

Despite plenty of existing models, they are divided into two broad categories: (i) Continuum models and (ii) Lattice models. In the following, we discuss some of them and illustrate when and how the theoretical descriptions of models fail to explain experimental observations and further refinements of the models are necessary. A few theoretical and numerical techniques will be reviewed briefly.

### 2.1. Continuum models of homopolymers

We briefly discuss three models namely the Gaussian Chain model, the Freely Jointed Chain (FJC) and the Worm Like Chain (WLC model) [14–16]. All these models ignore the excluded volume effect by allowing two monomers to overlap. The usefulness of these models comes from their simplicity allowing analytic expressions to be written in a very simple form. In these models a polymer chain consists of  $N$  beads (monomers). A point in  $d$ -dimensional space represents each monomer and the distance between two consecutive monomers is  $R_{i-1} - R_i$  (see Fig. 1). For the Gaussian chain (also known as an ideal chain) in the presence of an applied force  $f$ , the expression for the extension  $x$  can be written as

$$x = \frac{Nb}{3} \frac{fb}{k_B T}, \quad (1)$$

where  $N$  is the total number of monomers and  $b$  is the characteristic length scale known as the Kuhn length and defined as the mean square distance between monomers  $i$  and  $i + 1$ .  $k_B$  is the Boltzmann constant and  $T$  the temperature of the system. This is a linear force relation. The model describes the response of a single polymer chain under low force. The major limitation of the model is the property that the distance between two monomers can be extend without any limit. This limitation is removed in the FJC model, where the distance between two consecutive monomers (bond length) is kept fixed while the bond between three consecutive monomers can be oriented in any direction Fig. 1(b). The extension along the pulling direction is given by the relation

$$x_{FJC} = Nb \mathcal{L} \left( \frac{fb}{k_B T} \right). \quad (2)$$

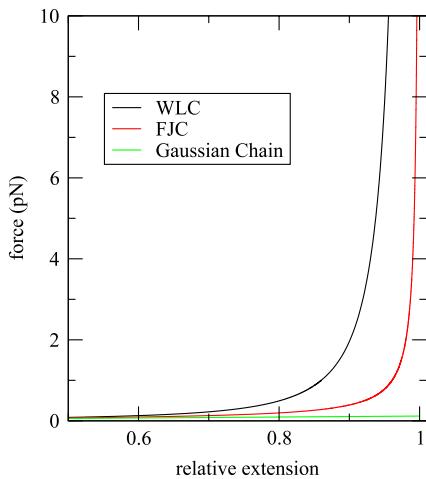
$\mathcal{L}$  is the Langevin function defined as

$$\mathcal{L} \left( \frac{fb}{k_B T} \right) = \frac{1}{\tanh \left( \frac{fb}{k_B T} \right)} - \frac{1}{\frac{fb}{k_B T}}. \quad (3)$$

The above expression is very useful in describing the properties of flexible polymers [17–21]. In the limit  $f \ll k_B T/b$  we get back the expression for the Gaussian chain model. However, in the large force limit, the expression reduces to

$$x = Nb \left( 1 - \frac{k_B T}{fb} \right), \quad (4)$$

which shows that the extension does not cross the contour length, and the polymer chain behaves like an entropic spring and the stretching force increases with temperature. Eq. (2), therefore, has been adopted in simulations to model bonding potentials.



**Fig. 2.** Force-extension curves for the Gaussian chain, FJC and WLC models.

At this point, it is important to introduce one more length scale, namely the persistence length ( $L_p$ ), which is frequently used to describe biopolymers. Any polymer segment less than a critical length  $L_c = R^2/L$  can be regarded as rigid, where  $R$  is the size of the chain and  $L$  is the contour length of the chain. The persistence length measures the distance along the chain over which the tangent vectors of the chain become uncorrelated [15]. The persistence length of the ideal chain is half the Kuhn length. For typical DNA one finds a statistical segment length of 80–250 bp and a persistence length of approximately  $L_p = 50$  nm (depends on the salt concentration) [17,22]. The size of the polymer can be related to the average end-to-end distance, which can be calculated from the following expression [15]

$$\langle \mathbf{R} \rangle_e = \sum_{i=1}^N \mathbf{R}_i. \quad (5)$$

As there is no preferred direction  $\langle \mathbf{R} \rangle_e = 0$ . Nevertheless,  $\langle \mathbf{R}^2 \rangle_e$  is nonzero and may be calculated from the following expression

$$\sum_{i=1}^N \sum_{j=1}^N \mathbf{R}_i \mathbf{R}_j = Nb^2 \quad (6)$$

and used to define the characteristic length of a polymer chain. It is pertinent to mention that the above quantity has been used extensively in polymer physics to define a critical exponent  $\nu$  of the chain as  $R_e^2 \sim N^{2\nu}$  in the limit  $N \rightarrow \infty$ . The FJC model can be regarded as a random walk model of polymers (Fig. 3(a)) [15], where the value of  $\nu$  is equal to 0.5 as seen from Eq. (6).

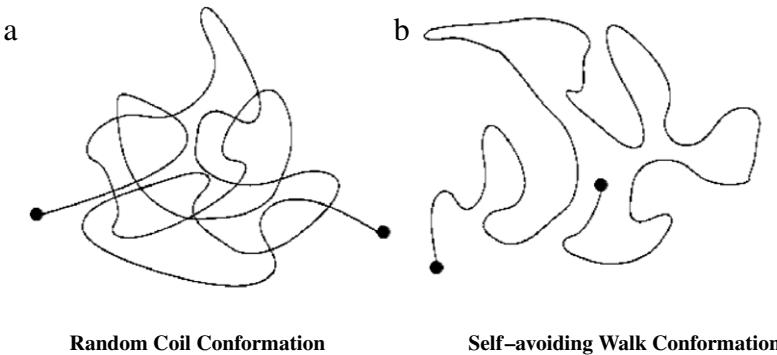
One of the important characteristics of biopolymers not described by the FJC model is the stiffness of the chain. The WLC model is the most commonly used model to describe the elastic response of biopolymers. The model builds an extension of the FJC with the assumption that the orientation of bonds is not free but costs additional energy. The extension within the WLC model can be calculated by transfer matrix methods, numerical simulations etc. One can find variational expressions for  $x$  as a function of  $f$  [23,24]. The one used extensively in the literature is [23]

$$f = \frac{k_B T}{2b} \left[ \frac{1}{(1-x/L)^2} - 1 + \frac{4x}{L} \right]. \quad (7)$$

This equation also reduces to the Gaussian chain result in the small force regime i.e.  $f < k_B T/b$  and for large force  $f > k_B T/b$ , it acquires the form

$$x = L \left( 1 - \sqrt{\frac{k_B T}{2fb}} \right). \quad (8)$$

One can see from this equation that the difference between the contour length  $L$  and the extension  $x$  varies as  $f^{-0.5}$ . The force-extension curves for the Gaussian chain, FJC and WLC are shown in Fig. 2. The behavior of a biopolymer subjected to a large force as described by the WLC model is different to the behavior as described by the FJC model, whereas at low force both models describe identical behavior. The Kuhn length of the WLC model corresponds to twice the persistence length.



**Fig. 3.** (a) Schematic of a random coil conformation in the continuum limit, where segments of the polymer are allowed to overlap. (b) A typical conformation of a self-avoiding walk (SAW) conformation in the continuum limit. Because of the non-crossing of the walk (excluded volume effect) the end-to-end distance is greater than for the random coil.

### 2.1.1. Excluded volume interaction: Self-avoiding chain model

Light-scattering and other experiments have shown that  $R_e$  grows more quickly than  $N^{0.5}$  as described by the WLC or FJC model [15,25]. Flory, who suggested that this discrepancy must be due to the excluded volume effect, discovered the answer to this puzzle [14]. In the WLC or FJC model monomers are allowed to overlap, while in reality the monomers try to bounce away from each other like the molecules of a gas.

Flory obtained a simple relation between the size exponent  $\nu$  and the spatial dimension  $d$  by postulating that the repulsive energy in the chain arises from excluded volume effects acting on the monomers balanced by the attractive elastic energy from their ideal chain. Minimizing the free energy, which is the sum of elastic energy and the monomer–monomer repulsive energy with respect to  $R_e$ , the expression for  $\nu$  at equilibrium can be written as [25]

$$\nu = \begin{cases} \frac{3}{d+2}, & d \leq 4 \\ \frac{1}{2}, & d \geq 4. \end{cases} \quad (9)$$

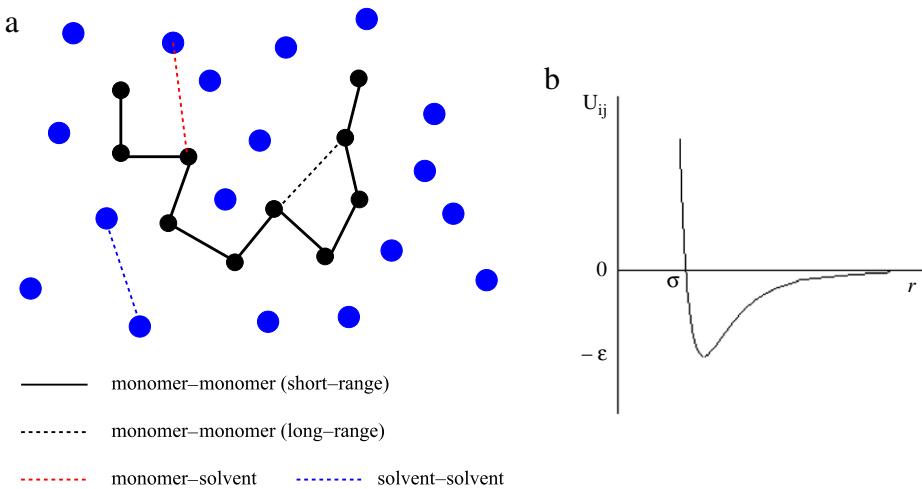
The repulsive energy is enormously overestimated in the Flory treatment when correlations are omitted, but this is effectively canceled by the elastic energy which is also overestimated. The cancelation of these two errors, results in a remarkably successful formula, which gives the exact value of  $\nu$  in one, two, and four dimensions while overestimating by merely 1% in three dimensions. Experimentalists have successfully tested Flory's theory. Ever since Flory presented his heuristic solution, physicists and mathematicians have tried to solve the problem of a random walk with excluded volume effects, also called the self-avoiding walk (SAW), which resembles a real polymer (Fig. 3(b)). From polymer statistics it is known that certain quantities associated with polymers, e.g., the number of distinct  $N$ -step SAWs,  $C_N$ , the number of closed  $N$ -step polygons,  $P_N$ , and the end-to-end distance,  $R_e$ , of SAWs of length  $N$  scales as [14]

$$\begin{aligned} C_N &\sim \mu^N N^{\gamma-1} \\ P_N &\sim \mu^N N^{\alpha-2} \\ R &= R_0 N^\nu, \end{aligned} \quad (10)$$

where  $\mu$  is the connectivity constant giving the average number of available steps for an infinitely long walk and  $R_0$  is a constant (lattice spacing). From the physics of phase transitions and critical phenomena, we know that certain physical quantities like the susceptibility  $\chi$ , specific heat  $C$  and the correlation length  $\xi$  near the transition point,  $T_c$ , scale as

$$\begin{aligned} \chi &\sim \chi_0 \left| \frac{T - T_c}{T_c} \right|^{-\gamma} \\ C &\sim C_0 \left| \frac{T - T_c}{T_c} \right|^{-\alpha} \\ \xi &\sim \xi_0 \left| \frac{T - T_c}{T_c} \right|^{-\nu}. \end{aligned} \quad (11)$$

A relation between polymer statistics and phase transitions was established by de Gennes [14] and des Cloiseaux [26] showing the correspondence between the polymer chain modeled by self-avoiding walks and the  $n$ -vector spin model of magnetization in the limit  $n \rightarrow 0$ . The correspondence between  $1/N$  and  $\frac{T-T_c}{T_c}$  is established as  $\frac{T-T_c}{T_c} \rightarrow 0$  and  $N \rightarrow \infty$ . This equivalence allows polymer science to benefit from the vast knowledge accumulated in the study of critical phenomena.



**Fig. 4.** (a) Diagram representing various kinds of interactions in a solution containing a polymer chain. (b) Diagrammatic representation of the monomer–monomer interaction in a polymer chain. The interaction is strongly repulsive at short-range ( $<\sigma$ ) and attractive at large distances. Units chosen here are arbitrary.

For example, the following scaling relations between critical exponents as established in statistical physics can also be used for polymers [14]

$$\begin{aligned} \alpha + 2\beta + \gamma &= 2 \\ 2 - \alpha &= d\nu, \\ \gamma &= \nu(2 - \eta). \end{aligned} \quad (12)$$

Here  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\nu$ , and  $\eta$  are the critical exponents and  $d$  is the dimensionality of the system. The above equivalence may serve an important role in describing the long-range behavior of biopolymers (Fig. 3(b)).

### 2.1.2. Interactions that matter

In a polymer solution consisting of polymer chains (as solute) and a solvent following three types of interaction must be considered (shown in Fig. 4(a))

- (i) Monomer–monomer interactions
- (ii) Monomer–solvent interactions
- (iii) Solvent–solvent interactions.

There are two kinds of monomer–monomer interaction: (a) Covalent-bonds (short range interactions) between consecutive monomers along the chain which are of the order of  $100\epsilon_H$ , where the hydrogen bond energy  $\epsilon_H \sim 1$  kcal/mol. Considering the typical distance between two monomers  $a \approx 0.4$  nm, the typical force is about  $100\epsilon_H/a \sim 1000$  pN. (b) Physical interactions between two monomers separated by some distance along the chain (long range interaction) which are about ten times weaker than covalent interactions that is a typical force of  $\sim 100$  pN.

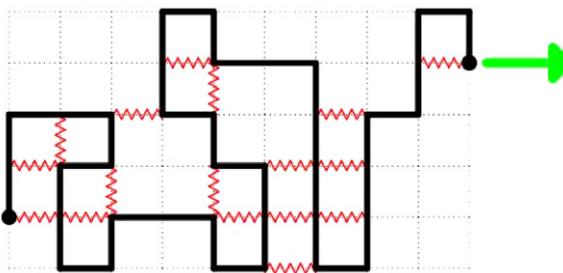
The long-range interaction has a form similar to the one between a pair of molecules e.g. a Lennard–Jones (L–J) potential

$$U(R_{ij}) = 4\epsilon \left[ \left( \frac{\sigma}{R_{ij}} \right)^{12} - \left( \frac{\sigma}{R_{ij}} \right)^6 \right], \quad (13)$$

where  $\epsilon$  is the depth of the potential well and  $\sigma$  is the distance at which the interparticle potential is zero (Fig. 4(b)). These parameters can be fitted to reproduce experimental data or accurate quantum chemistry calculations. The  $(1/r)^{12}$  term describes the Pauli repulsion at short range due to overlapping electron orbitals and the  $(1/r)^6$  term describes attraction at long range (van der Waals force, or dispersion force). Because of the repulsive core ( $R_{ij} < \sigma$ ), the polymer chain cannot cross itself or equivalently no two monomers of the chain can occupy the same space. This non-availability of space, which is already occupied by a monomer, takes care of the excluded volume effect discussed above.

The solvent–solvent interaction  $V(R)$  plays a crucial role in determining the structure of polymer chains in solution [27]. The competition between  $V(R)$  and  $U(R)$  results in many interesting features including the  $\Theta$ -point (coil–globule transition). We list some of them here [27]:

1. It may be that the net effect of the attractive and the repulsive interactions nearly cancel each other. In this case, the chain becomes an ideal or Gaussian chain and the solution is termed as an athermal solution.



**Fig. 5.** A self-attracting-self avoiding walk on the square lattice with one end attached to the surface and subject to a pulling force  $f$  on the other end. Interactions represented by wiggly lines are introduced between non-bonded monomers.

2. The net effect of the two interactions can be such that the monomers interact with each other only through a strong short-range repulsive interaction. A solvent in which such a situation occurs is called a good solvent.
3. The net effect of the competition between  $U(R)$  and  $V(R)$  may result in an interaction, which has both short-range repulsion and long-range attraction. In this situation the polymer chain may transform from an expanded coil to a compact globule phase by lowering the temperature. A solvent in which such a transition occurs is known as a poor solvent or bad solvent. The interaction between monomers in a polymer chain is often referred to as self-attraction [14].

#### 2.1.3. Lattice model

A linear polymer chain in a solvent can be modeled by a walk on a lattice in which each vertex of the walk represents a monomer (see Fig. 5) and the edge or step between vertices represents the covalent bond binding the chain together. In a random walk model of a polymer (Fig. 3(a)), the walk is allowed to cross itself without restriction and the results obtained from this model are same as those for the Gaussian chain as discussed above. The excluded volume effect can be incorporated in the walk with a constraint that a lattice site cannot be visited more than once [28–30]. This kind of walk is known as a self-avoiding walk (SAW), which simulates a linear polymer chain in a good solvent [14,16]. A polymer chain in a poor solvent is modeled by self-attracting self-avoiding walks (SASAWs) by including self-attraction among monomers [29,31] with  $-\epsilon > 0$  being an attractive monomer–monomer interaction energy.

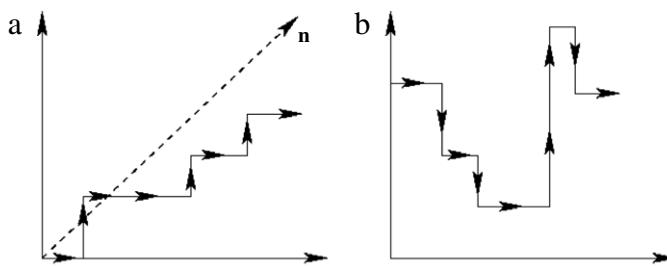
The polymer and solvent system is in one of three states: compact polymer or globule (poor or bad solvent) where  $\nu = 1/d$  for high values of  $\epsilon$ ; swollen or coil polymer (good solvent) with  $\nu = 0.588$  for low  $\epsilon$  and a  $\Theta$ -solvent at the transition temperature where the coil–globule transition takes place. In a  $\Theta$ -solvent the value of  $\nu$  is given by [16]

$$\nu_{\Theta} = \begin{cases} \frac{4}{7}, & \text{if } d = 2 \\ \frac{1}{2}, & d \geq 3. \end{cases} \quad (14)$$

Various kinds of underlying lattice have been used to study the conformational properties of linear polymer chains. The choice of the lattice depends on mathematical convenience, nature of the system and the interactions present in the system. As far as universality is concerned, the nature of the underlying lattice and the detail of the interactions do not matter much and many important properties associated with polymers can be derived. These properties are in qualitative agreement with experiments and sometimes quantitative agreement has also been achieved particularly in determining the critical exponents. Therefore, SAWs on a lattice have been studied extensively in attempts to describe the properties of polymers and biopolymers with suitable interactions. In spite of much effort, an exact solution of the problem in  $3d$  has not been possible so far. In  $2d$  the model has not been solved either but exact results have been obtained for some critical properties such as the critical exponents [32]. In  $2d$ , the exact value of the connectivity constant ( $\mu$ ) is known for the hexagonal lattice and conjectured for the square lattice [16].

In addition to self-avoidance, one can also introduce additional constraints on the walks e.g. certain direction(s) is (are) not accessible to the walker. Such walks are called directed walks (DWs). To define a directed walk, a preferred direction  $z$  on the lattice is assigned. Walkers are allowed to take steps only in the non-negative direction of  $z$ . The directed walk can be seen as a model of a polymer that is subject to some external force in the direction of  $z$  e.g. a flow in which the polymer is immersed or an electric field acting on electrically charged polymers. Depending on the choice of the direction  $z$ , directed walks may be defined in two ways as shown in Fig. 6:

1. Fully Directed walks: If the direction  $n$  is assigned as shown in Fig. 6(a) and walks in the non-negative projection of  $n$  are not allowed then such walks are called fully directed walks (FDWs).
2. Partial directed walks: If the direction  $n$  is assigned say along the  $x$ -axis Fig. 6(b), then the walks are said to be partial directed walks (PDSAWs). For example in  $2d$ , the walker can take steps in the  $\pm y$  directions but only in the  $+x$  direction. The model exhibits a collapse transition in the presence of an attractive interaction [16].



### TWO TYPES OF DIRECTED WALKS

**Fig. 6.** (a) Schematics of fully directed walk (FDW). Here the walker is allowed to take steps only in the  $+x$  and  $+y$  directions. (b) A diagrammatic representation of partial directed walks, where the walker is allowed to takes steps in  $\pm x$  direction but only in  $+x$  direction.

DWs can be solved analytically and exact results can be derived [33], however, because of the constraints stated above, its applicability is confined.

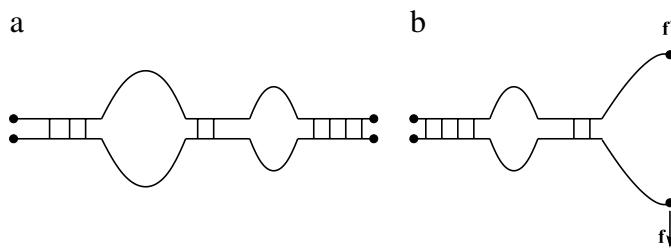
The kinds of macromolecules which can be described by a single polymer chain are proteins and RNA. The above models with suitable constraints can mimic the structural and functional properties of such biomolecules and have been used extensively to describe stretching experiments. The models discussed above can also describe the stretching of dsDNA and ssDNA.

## 2.2. Modeling of dsDNA

The DNA molecules in each cell of an organism contain all the genetic information necessary to insure the normal development and function of that organism [34,35]. It is a hetero-polymeric molecule consisting of residues (nucleotides) of four types, cytosine the shorthand is (C), guanine (G), adenine (A) and thymine (T). dsDNA has a helical structure *i.e.* it is made up of two strands paired to one another by hydrogen bonds. Bases A and G are classified as purine, C and T are as pyrimidines. The bases are capable of forming hydrogen bonds among themselves, but in a selective way. A can only couple with T and G with C. Therefore, in dsDNA there are two possible base pairs or base pairing interactions AT (or TA) and CG (or GC). A-T base pair involves the formation of two hydrogen bonds while the C-G base pair contains three hydrogen bonds. The strength of a hydrogen bond is typically 1 to 5 kcal/mol. The genetic information is encoded in the precise linear sequence of the nucleotide bases from which DNA is built. At room temperature, all base pairs are intact and the probability that a base pair breaks due to thermal fluctuation is of the order  $10^{-6}$ . However, with a rise in temperature or change of the pH of the solvent, the probability may increase and bubbles form. Ultimately, one reaches a value where the number of intact base pairs drops abruptly and the two strands are better represented as single stranded chains. This process is called thermal denaturation or melting of DNA. The temperature at which half the base pairs are open is the melting temperature, which lies in the range 80–100 °C. Experiments done in the 1960s on thermal denaturation revealed important information about the basic behavior of DNA. For example, it was observed that GC pair are relatively stronger than AT pair because of the triple hydrogen bond involved in C-G pairing. Another interesting observation, from these experiments was the multi-step transitions in some DNA. Differential melting curves (derivative of broken pairs with temperature) showed several peaks [36]. These peaks are the signature of sharp transitions of cooperatively melting regions, like, inner loop openings or the unbinding of double stranded regions at the edges of the chain [36]. For a very long chain ( $\sim 10^6$  bp) there is a nearly smooth transition that is supposedly a superposition of several such sub-transitions. The number of base pairs open in each step may be obtained by deconvoluting the differential melting curve into a superposition of Lorentzian curves [36]. Normalizing the areas of these curves to the length of the chain one finds that the cooperative melting regions are roughly 100–400 bps. Cooperatively melting regions have also been called thermalities. One of the interesting characteristics of biological macromolecules is the interplay between the cooperative interactions between different regions and the independent properties of these regions. Regions of DNA of the order of 100 base pairs unwind over temperature intervals of 0.3–0.7 K. When a specific region opens the stability of adjacent regions is decreased. The progress of unwinding depends on the cooperativity between regions and the internal stability of each region [37–39]. The stepwise melting has been structurally verified by electron microscopy. In the following, we are going to discuss some of the models, which have been developed in the context of DNA melting and widely applied in order to understand the mechanism of DNA unzipping.

### 2.2.1. Poland-Scheraga (PS) model

The Poland-Scheraga (PS) model [35,40] considers the DNA molecule as being composed of an alternating sequence of bound and denatured states as shown in Fig. 7. A bound state is energetically favored over an unbound state, while a denatured state is entropically favored over a bound one. In this simplified description, one ignores not only interactions between different parts of the chain, but also more generally other details regarding real DNA such as stiffness, helicity, chemical composition, and excluded volume effects of nucleotides. The model was later generalized by Fisher [41] to take



**Fig. 7.** (a) A schematic representation of the Poland–Scheraga model of the dsDNA unzipping. (b) A force  $f$  is applied at one end of the DNA strand. Segments of the chain are assumed to be non-interacting and the model therefore does not explicitly contain excluded volume effects.

into account the excluded volume effect within each loop. It was shown that the nature of the phase transition is determined by a critical exponent  $c$  of the underlying loop class. If  $c > 2$  the transition is first order while it is second order if  $1 < c < 2$ . For  $c \leq 1$ , no transition takes place and the strands remain bound.

For the sake of convenience, we impose the boundary condition where the nucleotides at the ends are always bound (Fig. 7(a)). However, when we apply force one of the ends will be denatured as shown in Fig. 7(b). Rest of the nucleotides of the chain will be either bound or unbound to a specific ( $i$ th nucleotide of one strand will be bound only with  $i$ th monomer of other strand) way called native interaction. The representation shown in Fig. 7 consists of sequence of bound segments separated by the denatured loops. The Boltzmann weight of the bound sequence of length  $i$  is  $s^i = \exp(-i\epsilon/k_B T)$ . The binding energy  $\epsilon$  is taken to be the same for all native nucleotides. The weight of the denatured loops is assigned as  $\Omega(2i) = k^i/i^c$  [42–44], where  $k$  is a non-universal constant. The model can be studied in the grand canonical ensemble, where the chain length is allowed to fluctuate. The grand partition function is given by [42–44]

$$Z = \sum_{N=0}^{\infty} Z(N)x^N = \frac{V_0(x)U_N(x)}{1 - U(x)V(x)} \quad (15)$$

where  $Z(N)$  is the canonical partition function of a chain length  $N$  and  $x$  is the fugacity. The function  $U(x)$ ,  $V(x)$  and  $U_N(x)$  are defined as

$$U(x) = \sum_{i=1}^{\infty} \frac{k^i}{i^c} x^i = \Phi_c(xk); \quad V(x) = \sum_{i=1}^{\infty} s^i x^i \quad (16)$$

and  $V_0(x) = 1 + V(x)$ ,  $U_N(x) = 1 + U(x)$ .  $\Phi_c(x)$  is the polylog function. The detail properties of this function may be found in Ref. [43]. In the thermodynamic limit,

$$\ln Z \approx N \ln x_1. \quad (17)$$

Here  $x_1$  is the value of the fugacity in the limit  $N \rightarrow \infty$ . This is the smallest value of the fugacity for which the partition function defined in Eq. (15) diverges. This can arise either from the divergence of the numerator or from the vanishing denominator. The latter is the situation relevant at lower temperature which corresponds to  $x_1$  satisfying,

$$U(x_1)V(x_1) = 1. \quad (18)$$

Since  $V(x) = sx/(1-sx)$ , the above equation can be put in the form

$$U(x_1) = 1/sx_1 - 1. \quad (19)$$

The order parameter, which may be defined as the fraction of bound pairs  $m$ ,

$$m = \frac{\partial \ln x_1}{\partial \ln s}. \quad (20)$$

The nature of the transition is determined by the dependence of  $x_1$  on  $s$ . The function  $U(x)$  does not depend on  $s$ . It remains finite for  $x < 1/k$  and diverges when  $x > 1/k$ . Moreover, the function  $1/V(x)$  in Eq. (15) increases continuously as the temperature increases. When the temperature increases from  $T = 0$ , the fraction of bound pairs decreases and  $x_1$  increases. However, if  $x_1$  reaches the value  $x_c = 1/k$ , it implies that  $1/V(1/k) \geq U(1/k)$ . An increase in temperature after that does not change  $x_1$  and thus  $m$  remains equal to 0. Therefore, the transition takes place at  $x_1 = x_c = 1/k$ . Its nature is determined by the behavior of  $U(x)$  in the vicinity of  $x_c$ . This is controlled in turn by the value of the exponent  $c$ . There are three regimes

1. For  $c \leq 1$ ,  $U(x_c)$  diverges and hence  $x_1$  will be an analytic function of  $s$  and therefore, there will be no phase transition.
2. For  $1 < c \leq 2$ ,  $U(x_c)$  converges but its derivative diverges at  $x_1 = x_c$ . Thus the transition is of second order.
3. For  $c > 2$ ,  $U(x_c)$  and its derivative converges at  $x_1 = x_c$  and hence the transition is of first order.

However, considering the case of force induced unzipping, a divergence in the partition function arising from the boundary condition (Fig. 7(b)) will play a crucial role, and in this case the grand canonical partition function takes the form [43]

$$\frac{V_0(x)O(x)}{1 - U(x)V(x)}. \quad (21)$$

The factor  $O(x)$  represents the grand partition function of the open tail under the force which is defined as

$$O(x) = 1 + \sum_{N=1}^{\infty} Z_{\text{end}}(N)x^N. \quad (22)$$

$Z_{\text{end}}$  is the canonical partition function of an open end composed of two strands each of length  $N$ . Kafri and Mukamel [45] studied the effect of force and found an expression for the partition function

$$O(x) = 1 + \sum_{N=1}^{\infty} \left[ x s \exp(A(fR_o/k_B T)^{1/\nu}) \right]^N. \quad (23)$$

Here  $A$  is a constant and  $R_o$  is defined by Eq. (10). According to Eq. (22) below the melting temperature  $T_M$  the end segment partition function diverges at a critical unzipping force  $f_c$  given by

$$\exp \left[ -A \left( \frac{f_c R}{k_B T} \right)^{\frac{1}{\nu}} \right] = kx^*(s) \quad (24)$$

where  $x^*(s)$  is the solution of Eq. (24). Since the average length of the loop in the polymer chain is finite the transition is of first order. Near the transition temperature, the length of the end segment diverges like  $|x^* - xU|^{-1}$  where  $xU = \exp(-kf_c/k_B T)^{1/\nu}/k$ . Since  $x^*$  is a regular function of  $f$ , one gets the same scaling as found by Bhattacharjee [46].

### 2.2.2. Peyrard–Bishop model

In 1989, Peyrard and Bishop [47] proposed a model of DNA by neglecting the asymmetry of the molecule and they assumed that each strand can be represented as a set of point masses (Fig. 8(a)), each corresponding to a nucleotide. The main characteristics of the model are:

- The longitudinal displacements of bases are neglected, as their amplitude of vibrations is much smaller than the amplitude of vibration of transverse motion. The displacement from the equilibrium of the  $n$ th nucleotide is denoted by  $u_n$  for one chain and  $v_n$  for the second chain.
- Two neighboring nucleotides on the same strand are connected by a harmonic potential. (Later on this term was modified by Dauxios et al. [48] to an anharmonic potential.)
- The hydrogen bonding between two base pairs on opposite strands is represented by a Morse potential. It represents not only the attraction between base pairs but also the repulsion between the two phosphate groups on the opposite strands (Fig. 8(b)).

The Hamiltonian of the model system can be written as [47],

$$H = \sum_n \left[ \frac{1}{2} m(u_n^2 + v_n^2) + \frac{1}{2} k \{(u_n - u_{n+1})^2 + (v_n - v_{n+1})^2\} + D(e^{-a(u_n - v_n)/\sqrt{2}} - 1)^2 \right]. \quad (25)$$

Here  $m$  is the reduced mass of the base pair and  $k$  is the force constant. The first term in the above equation represents the kinetic energy of the system and the second term is the stacking potential which represents the effect of stretching of a pair on nearest neighbor positions. The third term is the Morse Potential which represents the interaction between two bases. In reduced coordinates the motions of the two strands can be described as,

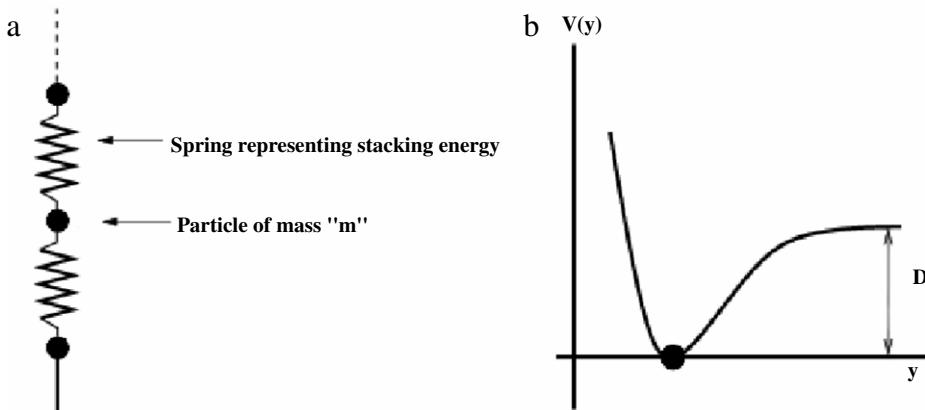
$$x_n = \frac{u_n + v_n}{\sqrt{2}}, \quad y_n = \frac{u_n - v_n}{\sqrt{2}}$$

which represents the in phase and out of phase motions, respectively. In terms of reduced coordinates, the Hamiltonian can be written as,

$$H = \sum_n \left\{ \left[ \frac{1}{2} m \dot{x}_n^2 + \frac{k}{2} (x_n - x_{n+1})^2 \right] + \left[ \frac{1}{2} m \dot{y}_n^2 + \frac{k}{2} (y_n - y_{n+1})^2 + D(e^{-ay_n} - 1)^2 \right] \right\}. \quad (26)$$

The part of the Hamiltonian which depends on the variable  $x_n$  is decoupled from the stretching part and corresponds merely to a harmonic chain without on-site potential. One can ignore this term in studying the opening of helix of DNA. The effective Hamiltonian in presence of force can be written as

$$H = \sum_n \left[ \frac{p_n^2}{2m} + \omega(y_n, y_{n+1}) + V(y_n) \right] - fy_1. \quad (27)$$



**Fig. 8.** (a) Pictorial representation of the Peyrard–Bishop model, which can be viewed as a model of a one-dimensional monotonic lattice with each atom having mass  $m$  and nearest neighbor interaction  $\omega$ . (b) Each atom of the lattice is subjected to an external potential  $V(y)$  whose effect is to confine the chain in a potential well.

Here  $y_n$  denotes the stretching from the equilibrium position of the hydrogen bonds connecting the two bases of the  $n$ th pair and  $m$  is the reduced mass of the base pair. The stacking interactions are contributed by dipole–dipole interactions,  $\pi$ -electron systems, London dispersion forces, and in water, the solvent–solvent induced hydrophobic interactions. The resultant of these is a complex interaction pattern between overlapping base pairs. The following anharmonic potential model mimics the essential features of the stacking energy

$$\omega(y_n, y_{n+1}) = \frac{1}{2}k(1 + \rho e^{-b(y_n + y_{n+1})})(y_n - y_{n+1})^2 \quad (28)$$

where the force constant  $k$  is related to the stiffness of a single strand and the second term in the bracket is the anharmonic term. The range of anharmonicity is defined by a parameter  $b$ . In the zipped state the force constant is equal to  $k(1 + \rho)$ . A decrease in the force constant in the unzipped state provides large entropy and hence favors unzipping either at high force or high temperature. The difference in the force constants between the zipped and the unzipped state of base pairs, create an energy barrier, which depends on  $\rho$  and  $b$ .

The Morse potential [49], which describes the interaction of two bases of the  $n$ th pair, defined as

$$V(y_n) = D_n(e^{-ay_n} - 1)^2. \quad (29)$$

In the model this term represents not only the hydrogen bond but also includes effects from the surrounding solvent and the repulsive interaction between phosphate strands. In Eq. (29),  $D_n$  measures the depth of the potential and,  $a y_n$  its range. In a homogeneous sequence of DNA,  $D_n$  is taken to be site independent, but in heterogeneous DNA the value of  $D_n$  depends on whether the  $n$ th base pair is AT or CG.

The fourth term in Eq. (27) is the contribution arising from the force as applied at one end of the strand. The Hamiltonian has five parameters  $D_n$ ,  $k$ ,  $\rho$ ,  $a$  and  $b$ . The model has been successfully applied to study the thermal denaturation and the unzipping of dsDNA and includes both heterogeneity and stacking. The model as presented here ignores the helicoidal structure which was later introduced by Barbi et al. [50]. The crucial shortcomings of the model are that it does not explicitly consider the embedding of the molecule in 3d space and the reselling long-range interactions.

### 2.2.3. Semi-microscopic model

In this subsection we are going to present a model first proposed by Bhattacharjee [46] to describe DNA unzipping of the homo-sequence. Lubensky and Nelson [51,52] have studied the model for homo- and hetero-sequences. The Hamiltonian of the double stranded DNA of  $N$  base pair is

$$H_0 = \int_0^N \left\{ \frac{3kT}{4bd} \left( \frac{\partial \mathbf{R}}{\partial n} \right)^2 + V[\mathbf{R}(n)] \right\} dn, \quad (30)$$

where the first term is the elastic energy and the second term is the interaction energy between complementary bases. Here  $b$  is the Kuhn length, which can be interpreted as a distance between two elastic monomers and the term  $d$  has been introduced to distinguish between different bases. Base pairing has been taken in a similar form to the PS or PB model and the detailed form is not crucial. It has a different form for A-T and G-C pairs. The Eq. (30) with random interaction was studied by Bhattacharjee and Mukherjee [53,54]. The contribution of energy resulting from the applied external force is  $H_f = -fx$ . thus the total Hamiltonian of the system is

$$H = H_0 + H_f. \quad (31)$$

The model as defined above has a similar form to those of vortex lines in a mixed superconductor or directed polymers in  $d + 1$  dimensions [46,51,52]. The partition function  $Z$  of the system

$$Z = \int \mathcal{D}(\mathbf{R}(n)) \exp\left(\frac{-H(T)}{k_B T}\right) \quad (32)$$

is well studied and many results are known. Here  $\mathcal{D}(\mathbf{R}(n))$  means that integration is performed on all paths of  $\mathbf{R}(n)$ . If we treat  $n$  as a time like coordinate then the system may be described by the path integral formulation of a quantum particle in imaginary time. The next step is to introduce sequence dependence in order to study the random sequence of DNA. The form of the interaction potential will thus be

$$V[\mathbf{R}(n) = 0] = V_0 + \eta(n) \quad (33)$$

where  $\eta(n)$  is the random function. Since pairing is independent, the average of the disorder is

$$\overline{\eta(n)} = 0 \quad (34)$$

and

$$\overline{\eta(n)\eta(n')} = \Gamma(\delta(n - n')). \quad (35)$$

As stated above, heterogeneity in the sequence results in a different behavior when compared to a homo sequence. We shall discuss this later.

#### 2.2.4. Lattice models

The important shortcomings of the above models are that they are one-dimensional and ignore entropic contributions due to the orientation of the bases. Moreover they also ignore non-native contacts in the formation of base pairs. Therefore, the formation of hairpins, condensation of DNA and the role of intermediate states during melting cannot be studied in these models. As discussed above, these features can be taken into account in lattice models. One of the simplest models, which incorporates the self-avoidance and chain fluctuation was introduced by Causo et al. [55] and further analyzed and extended by Carlon et al. [56]. The model consists of two self-avoiding walks of length  $N$  on a cubic lattice which share a common origin. The  $i$ th monomer of one chain is allowed to overlap only with the  $i$ th monomer of the other chain. An overlap corresponds to a bound state of complementary DNA base pairs that gain energy  $\epsilon$ . The model includes bubbles of different lengths and exhibits a thermal melting transition. In this model mutual avoidance of two chains have been neglected and the orientation of bases has not been taken into account. Only native pairing has been allowed and it is therefore similar to the Poland–Scheraga (PS) model. The concept of native interactions has been extended to study the force induced transition using directed walk models and exact results have been derived which includes the presence of a novel re-entrant unzipping transition at low temperature [57,58]. However, because of the way the native interaction is imposed, these models have only one thermal phase transition and therefore cannot describe phenomena like DNA condensation [34]. In a system of two interacting linear polymer chains where any monomer of one chain can interact with any monomer of the other chain (non-native interaction allowed), there are two zero force thermal phase transitions [59,60]. The first transition is from the condensed globule state (in case of a square lattice it takes the shape of spiral state) to the zipped state. With a further increase in temperature a second transition occurs from the zipped state to the unzipped state. The presence of an entropy stabilized intermediate state (zipped state) gives rise to a force induced triple point in the force–temperature phase diagram [59,60]. This opens up the possibility of a much richer phase diagram and multicritical behavior in polymeric systems having potential applications in understanding the mechanism of unzipping.

Recently a more sophisticated model incorporating the directional nature of bases has been proposed to study the DNA unzipping [61,62]. This semi-microscopic model is rich enough to take care of the shortcomings of models similar to PS and also has some additional features like the existence of intermediate states, effects of stacking energy, excluded volume properties of nucleotides and the directional nature of the hydrogen bond. The role of the formation of hairpins can also be studied in this model. The results obtained from these models will be discussed later.

#### 2.3. Modeling of proteins

The linear amino acid sequence of a polypeptide chain encodes the information required to fold to a native state as first shown by Anfinsen [63]. The ability to predict not only the native structure from a linear amino acid sequence, but also the pathways along which the reaction proceeds is known as the protein folding problem. The basic question concerns the mechanism by which a polypeptide chain is able to fold rapidly to the native state, despite the very large number of conformations that exists for the chain. This is known as Levinthal's paradox [64] and it was solved by the realization that proteins fold through parallel processes leading to the idea that the free energy surface of a folding peptide is funnel shaped [65].

The ultimate goal of protein modeling is to construct a model the native conformation of which coincides to that deposited on the protein data bank (PDB) ([www.rcsb.org/pdb](http://www.rcsb.org/pdb)). Furthermore, it should correctly capture folding and unfolding mechanisms of proteins. In this section we briefly discuss the main models used to study protein dynamics.

### 2.3.1. Lattice models

In the last two decades considerable insight into the thermodynamics and kinetics of protein folding has been gained from simple lattice models [66,67]. Since a protein is a polypeptide chain, lattice models used to describe them are similar to those for polymers which have been discussed in the previous section (Fig. 5). Here amino acids are represented by single beads which are located at the vertices of a cubic lattice. However, the most important difference from homopolymer models is that amino acid sequences and the role of contacts must be taken into account. Due to the constraint that a contact is formed if two residues are nearest neighbors, but not successive in sequence, contact between residues  $i$  and  $j$  is allowed provided  $|i - j| \geq 3$ . In the simple Go modeling [68], the interaction between two beads which form a native contact is assumed to be attractive, while the non-native interaction is repulsive. This guarantees that the native conformation has the lowest energy. In more realistic models like the HP (H: hydrophobic, P: polar)- or other models [66], specific interactions between amino acids are taken into account. Several kinds of potential [69–71] are used to describe these interactions.

To mimic more realistic features of proteins such as a dense core packing [72,73] a natural step is to include the rotamer degrees of freedom [74]. One of the simplest models is a cubic lattice backbone (BB) sequence of  $N$  beads, to which side chain beads are attached [75,76]. The system has in total  $2N$  beads.

In order to monitor protein dynamics usually one uses the standard move set which includes the tail flip, corner flip, and crankshaft for BB beads. If a move is accepted then the side chain moves are determined. The Metropolis criteria is applied once the moves of both BB and side chain beads are allowed geometrically. If an attempt involving the move of BB monomers fails then one tries to move the corresponding side chain beads simultaneously. It should be noted that the standard tail and corner flips involve one bead while two beads are moved in the crankshaft attempt. If the length of a chain is long enough then this move set ceases to be ergodic [77], i.e., one cannot explore the whole phase space using it. This can be partially overcome using more involved move sets in which three or more beads are allowed to move simultaneously [78,79]. While lattice models have been widely used to study the protein folding problem, they attract little attention in mechanical unfolding simulations [80,81].

### 2.3.2. Off-lattice coarse-grained Go modeling

The major shortcoming of lattice models is that beads are confined to lattice vertices and they do not allow for an accurate description of the protein shape. This can be remedied with the help of off-lattice models in which beads representing amino acids can occupy any position. A number of off-lattice coarse-grained models with realistic interactions (not Go) between amino acids [82,83] has been developed to study the mechanical resistance of proteins [84,85]. However, it is not an easy task to construct such models for long proteins.

In the pioneering paper [68] Go introduced a very simple model in which non-native interactions are ignored. This native topology-based Go model turns out to be highly useful in predicting the folding mechanisms and deciphering the free energy landscapes of two-state proteins [86–89]. On the other hand, in mechanical unfolding one stretches a protein from its native conformation, and unfolding properties are mainly governed by the native topology [90]. Therefore, the native-topology-based or Go modeling is suitable for studying the mechanical unfolding. Various versions of Go models [87,90–94] have been applied to this problem. Since they give qualitatively similar results, in this review we will focus on the variant of Clementi et al. [87]. Here one uses a coarse-grained continuum representation for a protein in which only the positions of  $C_{\alpha}$ -carbons are retained. The interactions between residues are assumed to be Go-like and the energy of such a model is as follows [87]

$$E = \sum_{bonds} K_r (r_i - r_{0i})^2 + \sum_{angles} K_\theta (\theta_i - \theta_{0i})^2 + \sum_{dihedral} \{K_\phi^{(1)} [1 - \cos(\phi_i - \phi_{0i})] + K_\phi^{(3)} [1 - \cos 3(\phi_i - \phi_{0i})]\} \\ + \sum_{i>j-3}^{NC} \epsilon_H \left[ 5 \left( \frac{r_{0ij}}{r_{ij}} \right)^{12} - 6 \left( \frac{r_{0ij}}{r_{ij}} \right)^{10} \right] + \sum_{i>j-3}^{NNC} \epsilon_H \left( \frac{C}{r_{ij}} \right)^2 + E_f. \quad (36)$$

Here  $\Delta\phi_i = \phi_i - \phi_{0i}$ ,  $r_{i,i+1}$  is the distance between beads  $i$  and  $i + 1$ ,  $\theta_i$  is the bond angle between bonds  $(i - 1)$  and  $i$ , and  $\phi_i$  is the dihedral angle around the  $i$ th bond and  $r_{ij}$  is the distance between the  $i$ th and  $j$ th residues. Subscripts “0”, “NC” and “NNC” refer to the native conformation, native contacts and non-native contacts, respectively. Residues  $i$  and  $j$  are in native contact if  $r_{0ij}$  is less than a cutoff distance  $d_c$  taken to be  $d_c = 6.5 \text{ \AA}$ , where  $r_{0ij}$  is the distance between the residues in the native conformation.

The first harmonic term in Eq. (36) accounts for chain connectivity and the second term represents the bond angle potential. The potential for the dihedral angle degrees of freedom is given by the third term in Eq. (36). The interaction energy between residues that are separated by at least 3 beads is given by 10–12 Lennard–Jones potential. A soft sphere repulsive potential, the fifth term in Eq. (36), disfavors the formation of non-native contacts. The last term accounts for the force applied to C and N termini along the end-to-end vector  $\vec{R}$ . We choose  $K_r = 100\epsilon_H/\text{\AA}^2$ ,  $K_\theta = 20\epsilon_H/\text{rad}^2$ ,  $K_\phi^{(1)} = \epsilon_H$ , and  $K_\phi^{(3)} = 0.5\epsilon_H$ , where  $\epsilon_H$  is the characteristic hydrogen bond energy and  $C = 4 \text{ \AA}$ .

In the constant force simulations the last term in Eq. (36) is

$$E_f = -\vec{f} \cdot \vec{R}, \quad (37)$$

where  $\vec{R}$  is the end-to-end vector and  $\vec{f}$  is the force applied either to both termini or to one of them. In the constant velocity force simulation we fix the *N*-terminal and pull the *C*-terminal by force

$$f = K(vt - x), \quad (38)$$

where  $x$  is the displacement of the pulled atom from its original position [95], and the pulling direction was chosen along the vector from fixed atom to pulled atom. In order to mimic AFM experiments, one chooses the spring constant  $K = 10\text{--}1000 \text{ pN/nm}$ , while for description of Laser optical tweezers (LOT) experiments –  $K = 0.001\text{--}0.1 \text{ pN/nm}$  (see section Experimental technique).

It should be noted that a simplified Go model with binary variables [96] have recently been applied to study the unfolding of proteins [97,98]. This model is a generalization of a model developed by Wako and Saito [99] in a purely thermodynamic version and subsequently reconsidered by Munoz and Eaton [100], who used a kinetic version to analyze experimental data on protein folding.

### 2.3.3. All-atom models

The most intensive theoretical study of protein folding has been performed with the help of all-atom simulations. All-atom models, which provide the most detailed description on the atomistic level, include the local interaction and non-bonded terms. The later include the (6–12) Lenard–Jones potential [13], the electro-static interaction, and the interaction with environment. The all-atom model with the CHARMM (Chemistry at HARvard Molecular Mechanics) force field [101] and explicit TIP3P (transferable intermolecular 3 point potential) water [102] has been employed first by Grubmuller et al. [103] to compute the rupture force of the streptavidin–biotin complex. Two years later a similar model was successfully applied by Schulten and coworkers [95] to the titin domain I27. The NAMD (NANoscale Molecular Dynamics) software [104] developed by this group is now widely used for stretching biomolecules by the constant mechanical force and by the force with constant loading rate (see recent reviews [105–107] for more references). NAMD works with not only CHARMM but also with AMBER (Assisted Model Building with Energy Refinement) potential functions [108], parameters, and file formats. Recently, it has become possible to use the GROMACS (GROningen MAchine for Chemical Simulations) software [109] for all-atom simulations of mechanical unfolding in explicit water.

The CHARMM software with the implicit water [101,110], has been used by Paci et al. [111–115] to understand the response of proteins to mechanical perturbation. Irback et al. applied an all-atom model but without water [116] to study occurrence of unfolding intermediates of ubiquitin [117].

Using all-atom simulations with explicit water one can study the dynamics of hydrogen bond rupture in detail. One remarkable example is that Schulten and coworkers [95] predicted the existence of a “hump” on the force–extension profile for the titin domain I27 (see below). This theoretical prediction was confirmed experimentally one year later [118]. The disadvantage of all-atom simulations is that they are very CPU demanding. As a result, to stretch proteins within a reasonable amount of simulation time one has to apply forces of 1000 pN which are one to two orders of magnitude larger than those used in constant-force experiments. In NAMD simulations the force loading rate  $v \sim 10^9 \text{ nm/s}$  which is about six–seven orders of magnitude faster than in real experiments. If a protein is pulled with high  $f$  or  $v$ , then details of the unfolding dynamics may be lost. In order to overcome this difficulty one can employ coarse-grained models which allow for probing lower forces and lower ramp speeds. In this review we focus on results obtained by using simple off-lattice Go models.

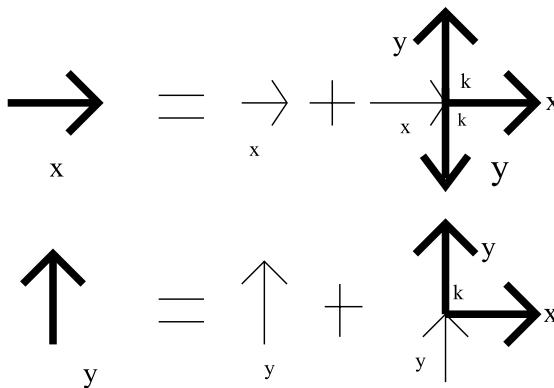
## 3. Methods

In this section we shall discuss basic analytical and numerical techniques used in describing the force induced transitions in biomolecules.

### 3.1. Generating function technique

The generating function technique is a very powerful technique, which may be adapted to study the conformational properties of biopolymers [33,119]. As discussed above biopolymers are in general semi-flexible. Here we use a simple example to illustrate how the generating function approach can be applied and exact results can be derived. Let us consider a directed walk model of a polymer chain [16,33]. For simplicity we consider partial directed walks (PDWs) in which the walker is allowed to move along  $x$ - and  $\pm y$ -axis only. Stiffness is introduced into the polymer chain by putting an energy cost  $\epsilon_b$  on every bend of the walk thus giving rise to an associated Boltzmann weight  $k = e^{-\frac{\epsilon_b}{k_B T}}$ . For  $k = 1$  ( $\epsilon_b = 0$ ) the chain is said to be flexible while for  $0 < k < 1$  ( $0 < \epsilon_b < \infty$ ) the chain is said to be semi-flexible. The grand canonical partition function of such a chain can be written as [33,120]

$$Z(x, k) = \sum_{N=0}^{\infty} \sum_{\text{all walks}} x^N k^{N_b}. \quad (39)$$



**Fig. 9.** The diagrammatic representation of the recursion relations (40) and (41) for PDWs. The thick arrows  $X$  and  $Y$  denote all possible walks with the initial step (fugacity) along  $+x$  and  $\pm y$  directions.

Here  $x$  is the fugacity of the walk and  $N_b$  is the number of bends in a given configuration. In 3d, the grand canonical partition function defined by Eq. (39) may be expressed as a sum of two components of PDWs. The recursion relations for the case of a semi-flexible polymer chain are

$$X = x + x(X + 2kY), \quad (40)$$

and

$$Y = y + x(kX + Y). \quad (41)$$

A schematic representation of the above equations (for 2D) has been shown in Fig. 9. It may be noted that the first term of Eqs. (40) and (41) is the fugacity of the walk and remain constant. Thus,  $x = y$ . Solving Eqs. (40)–(41) we get,

$$X = \frac{x + (2k - 1)x^2}{1 - 2x + x^2 - 2x^2k^2} \quad (42)$$

$$Y = \frac{x + (k - 1)x^2}{1 - 2x + x^2 - 2x^2k^2}. \quad (43)$$

The partition function of the system can, therefore, be written as

$$Z(x, k) = X + 2Y = \frac{(4k - 3)x^2 + 3x}{1 - 2x + x^2 - 2x^2k^2}. \quad (44)$$

The critical point for polymerization of an infinite chain is found from the relation

$$1 - 2x + x^2 - 2x^2k^2 = 0. \quad (45)$$

This leads to an expression for the critical value of the step fugacity as a function of  $k$ ,  $x_c = \frac{1}{(1+\sqrt{2k})}$ . In the limit  $k \rightarrow 1$ , it reduces to the well-known value for the flexible polymer chain in 2d [33]. The approach has been successfully applied to study DNA unzipping and the unfolding of a collapse polymer [24,57,58,121].

### 3.2. Exact enumeration technique

The partition function defined in Eq. (39) is a power series expansion about the origin  $x = 0$ . If

$$\lim_{N \rightarrow \infty} |C_N|^{-\frac{1}{N}} = x_c \quad (46)$$

then the series converges for  $|x| < x_c$  and diverges for  $|x| > x_c$ . Correspondingly, there must be at least one singularity (non-analytic point) on the circle of convergence  $|x| = |x_c|$ . If all the coefficients  $C_N$  are known exactly one can, in principle, analytically calculate the function across the complex  $x$ -plane at least up to any possible natural boundary of the function, beyond which it remains undefined [16,122]. The nature of the coefficients is determined by the singularities of the partition function. The singularity nearest to the origin will dominate the behavior at large  $N$ . If the dominant singularity is on the real positive axis, the coefficients will eventually all have the same sign. Conversely, if the dominant singularity is on the negative real axis, the coefficients alternate in sign. For our problems all coefficients are real and hence singularities must occur in complex conjugate pairs. For example, the partition function for semi-flexible polymer chain of length  $N$  can be enumerated exactly in  $d$  dimensions [120]. However, the number of distinct configurations  $C_N$  of length  $N$  and the time needed to enumerate all these conformations increases as  $\mu^N$ . So even with a rapid increase in computing power only

a few more terms can be obtained in each decade. Recently a finite lattice method has been used to greatly exceed the previous enumerations [123,124]. The basic idea of the finite lattice method is to count the number of SAWs on finite  $L \times W$  rectangles and then reconstruct the full series expansion by combining the results from the finite pieces. The number of distinct conformations in any finite rectangle is calculated using transfer matrix techniques [125]. Using transfer matrix calculations combined with parallel processing, the enumeration was extended to  $N = 45$  for the semi-flexible case. This extension of the chain length allowed for the prediction of intermediate states and their pathways in much greater detail.

Since a system of finite size does not show any phase transition, one has to approximate the partition function in the thermodynamic limit. Several methods like the ratio method, Padé approximation and differential approximation have been used [122]. Here we discuss briefly the ratio method which has been used quite extensively. One version consists of taking the logarithm of the ratio of consecutive terms of the series obtained from the partition function defined in Eq. (39). From Eq. (10) one sees that such a quantity should for large  $N$  (to leading order) be a linear function of  $1/N$

$$\log \frac{C_{N+1}}{C_N} \cong \log \mu + b \left( \frac{1}{N} \right). \quad (47)$$

A linear fit using Eq. (47) then determines the value of  $\mu$  and  $b$ . The analysis of  $R$  is even simpler as in this case, we can put  $\mu = 1$ . However, this method suffers from strong corrections to the form Eq. (10). There are many other more advanced methods which can be used to study the partition function.

### 3.3. Transfer integral technique

The configurational partition function for dsDNA consisting of  $N$  base pairs and at temperature  $T$  for the Peyrard–Bishop model is [47,48]

$$Z_N(\beta) = \int_{-\infty}^{\infty} \prod_{n=1}^N dy_n \exp[-\beta H(y_n, y_{n+1})] \quad (48)$$

where  $\beta = 1/k_B T$ .

To solve this equation the partition function is expressed in terms of a propagator or kernel  $K(y_n, y_{n+1}) = \exp[-\beta H(y_n, y_{n+1})]$ . For a chain with homogeneous base pairs the kernel can be expanded into a set of orthonormal base functions as,  $K(y_n, y_{n+1}) = \lambda_n \phi(y_n) \phi(y_{n+1})$ . Substituting this into the equation above and using the orthonormality conditions, the partition function is simply,  $Z_N(\beta) = \sum_{n=1}^N \lambda_n^N$ , which becomes  $Z_N(\beta) = \lambda_1^N$  for an infinite chain of base pairs. The ground state energy eigenvalue can be evaluated from the expression  $\epsilon_1 = -\frac{k_B T}{N} \ln(\lambda_1)$ . The critical force at temperature  $T$  can be calculated as suggested by [126],  $f_c = \sqrt{2k(D - \epsilon_1)}$ . The major problem with this technique is that it is only valid for a homogeneous base sequence. Once heterogeneity of the base sequence is considered the above expressions are no longer valid. In the case of thermal denaturation studies the method has been modified successfully, but for force induced unzipping the technique is no longer valid. One has to rely on the method of matrix multiplication for systems consisting of a random sequence of base pairs.

### 3.4. Molecular dynamics

One of the important tools that has been employed to study biological molecules is molecular dynamics (MD) simulation. It was first introduced by Alder and Wainwright in 1957 to study the interaction of hard spheres [127]. In 1977 the first biomolecule, the bovine pancreatic trypsin inhibitor (BPTI) protein, was simulated using this technique [128]. Nucleic acids have been considered a challenge for simulation because of the negative BB charge and the polyelectrolyte behavior [129]. Thus in oligonucleotide dynamics simulations, particular attention should be paid to the atomic charges. Nowadays the MD technique is quite common in the study of biomolecules such as solvated proteins, protein-DNA complexes as well as lipid systems addressing a variety of issues including the thermodynamics of ligand-binding, the folding and unfolding of proteins etc.

Before studying structural and dynamical properties of biomolecules it is important to note that these molecules exhibit a wide range of time scales over which specific processes take place. For example local motion, which involves atomic fluctuations, side chain motion, and loop motion occurs on the length scale of 0.01 to 5 Å and the time involved in such processes is of the order of  $10^{-15}$  to  $10^{-12}$  s. The motion of a helix, a protein domain or a sub-unit falls under the domain of rigid body motion with typical length scales between 1 and 10 Å and the time involved in such motion is between  $10^{-9}$  and  $10^{-6}$  s. Large-scale motion consists of helix-coil transitions or folding-unfolding transitions, which involve length scales of more than 10 Å and time scales from  $10^{-7}$  to  $10^1$  s. Typical time scales for protein folding are  $10^{-6}$  to  $10^1$  s [130]. In unfolding experiments, to stretch out a protein of length  $10^2$  nm, one needs a time of  $\sim 1$  s using a pulling speed of  $v \sim 10^2$  nm/s [2].

The method of steered MD (SMD) which combines the stretching condition with standard MD was initiated by Schulten and coworkers [105]. They simulated the force-unfolding of titin immunoglobulin domains, I27 and I11, showing atomic details of the molecular motion under force. The focus was on rupture events involving the hydrogen bonds that stabilize the structures. The structural and energetic analysis enabled them to identify the origin of the free energy barrier and

intermediates during mechanical unfolding. However, one has to notice that there is an enormous difference between the simulation conditions used in SMD and real experiments. In order to stretch proteins within a reasonable amount of CPU time, SMD simulations at constant pulling speed use speeds eight to ten orders higher and spring constants one to two orders larger than those of actual AFM experiments. Therefore, the effective force acting on the molecule is about seven to eight orders larger than in real experiments. It is unlikely, that the dynamics under such extreme conditions can mimic real experiments, and one has to be very careful about comparing simulation results with experimental results. In the literature the word "steered" therefore means MD at extreme conditions, where the constant force and pulling speed are chosen to be very high.

Excellent reviews on molecular dynamics and its use in biochemistry and biophysics are numerous (see, e.g., [131] and references therein). Below we focus only on Brownian dynamics as well as the second-order Verlet method for the Langevin dynamics simulation as they have been extensively used to obtain the main results presented in this review.

### 3.4.1. Langevin dynamics simulation

The Langevin equation is a stochastic differential equation which introduces friction and noise terms into Newton's second law to approximate effects of temperature and environment. This equation is

$$m \frac{d^2r}{dt^2} = F_c - \gamma \frac{dr}{dt} + \Gamma \equiv F \quad (49)$$

where  $\Gamma$  is a random force,  $m$  the mass of a bead,  $\gamma$  the friction coefficient and  $F_c = -dE/dr$ . The configuration energy  $E$  for the Go model, for example, is given by Eq. (36). The random force  $\Gamma$  is taken to be a Gaussian random variable with white noise spectrum and is related to the friction coefficient by the fluctuation-dissipation relation

$$\langle \Gamma(t)\Gamma(t') \rangle = 2\gamma k_B T \delta(t - t') \quad (50)$$

where  $\delta(t - t')$  is the Dirac  $\delta$ -function. The friction term only influences kinetic but not thermodynamic properties.

In the low friction regime, where  $\gamma < 25 \frac{m}{\tau_L}$  (the time unit  $\tau_L = (ma^2/\epsilon_H)^{1/2} \approx 3$  ps), Eq. (49) can be solved using the second-order Velocity Verlet algorithm [82,132]

$$x(t + \Delta t) = x(t) + \dot{x}(t)\Delta t + \frac{1}{2m}F(t)(\Delta t)^2, \quad (51)$$

$$\begin{aligned} \dot{x}(t + \Delta t) &= \left(1 - \frac{\gamma \Delta t}{2m}\right) \left(1 - \frac{\gamma \Delta t}{2m} + \left(\frac{\gamma \Delta t}{2m}\right)^2\right) \dot{x}(t) \\ &\quad + \left(1 - \frac{\gamma \Delta t}{2m} + \left(\frac{\gamma \Delta t}{2m}\right)^2\right) (F_c(t) + \Gamma(t) + F_c(t + \Delta t) + \Gamma(t + \Delta t)) \frac{\Delta t}{2m} + o(\Delta t^2), \end{aligned} \quad (52)$$

with time step  $\Delta t = 0.005\tau_L$ .

### 3.4.2. Brownian dynamics

In the over damped limit ( $\gamma > 25 \frac{m}{\tau_L}$ ) the inertial term can be neglected and we obtain a much simpler equation

$$\frac{dr}{dt} = \frac{1}{\gamma}(F_c + \Gamma). \quad (53)$$

This equation can be solved using the simple Euler method which gives the position of a biomolecule at the time  $t + \Delta t$  as follows

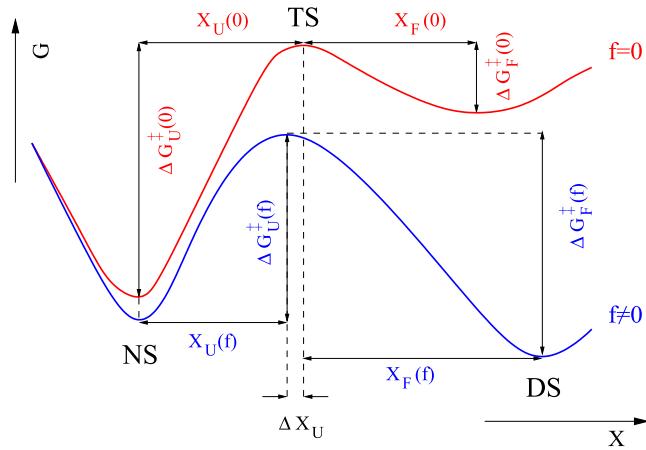
$$x(\Delta t + t) = x(t) + \frac{\Delta t}{\gamma}(F_c + \Gamma). \quad (54)$$

Due to the large value of  $\gamma$  we can choose the time step  $\Delta t = 0.1\tau_L$  which is 20-fold larger than the low viscosity case. Since the water has  $\gamma \approx 50 \frac{m}{\tau_L}$  [82] the Euler method is valid for studying protein unfolding.

### 3.5. Kinetic theory

One of the notable aspects of force experiments on single biomolecules is that the end-to-end distance  $R$  is directly measurable or controlled by instrumentation.  $R$  becomes a natural reaction coordinate for describing the mechanical processes. Hence, the mechanical unfolding theory is testable and has been developed in a more practical manner whereas many protein folding theories are rather elusive in practice because they use a difficult reaction coordinate like the fraction of native contacts.

The theoretical framework for understanding the effect of external constant force on rupture rates was first discussed in the context of cell-cell adhesion by Bell in 1978 [133]. Evans and Ritchie have extended the theory to the case where



**Fig. 10.** Conceptual plot for the FEL without (blue) and under (red) the external force.  $\Delta x_u$  is the shift of  $x_u$  in the presence of force.

the loading force increases linearly with time [134]. The phenomenological Bell theory is based on the assumption that the transition state (TS) does not move under stretching (Fig. 10). Since this assumption is not true Dudko et al. [135] developed a microscopic theory which is free from this shortcoming. In this section we discuss the phenomenological as well as microscopic kinetic theory.

### 3.5.1. Bell approximation: Constant force case

Suppose the external constant force,  $f$ , is applied to the termini of a biomolecule. The deformation of the free energy landscape under force is schematically shown in Fig. 10. Assuming that the force does not change the distance between the native state (NS) and TS ( $x_u(f) = x_u(0)$ ), Bell [133] stated that the activation energy is changed to  $\Delta G_u^\ddagger(f) = \Delta G_u^\ddagger(0) - fx_u$ , where  $x_u = x_u(0)$ . In general, the proportionality factor  $x_u$  has the dimension of length and may be viewed as the width of the potential. For example, if a ligand is pulled out of a binding pocket,  $x_u$  is the distance the ligand has to move before it traverses over the transition state. Using the Arrhenius law, Bell obtained the following formula for the unfolding/unbinding rate constant

$$k_u(f) = k_u(0) \exp(fx_u/k_B T), \quad (55)$$

where  $k_u(0)$  is the unfolding rate constant in the absence of a force. If a reaction takes place in the condensed phase, then according to Kramer's theory [136] the prefactor  $k_u(0)$  is

$$k_u(0) = \frac{\omega_0 \omega_{ts}}{2\pi\gamma} \exp(-\Delta G_u^\ddagger(0)/k_B T). \quad (56)$$

Here  $\gamma$  is a solvent viscosity,  $\omega_0$  the angular frequency (curvature) at the reactant bottom, and  $\omega_{ts}$  the curvature at the barrier top of the effective reaction coordinate [136,137]. For biological reactions, which belong to the Kramers category,  $\frac{\omega_0 \omega_{ts}}{2\pi\gamma} \sim 1 \mu\text{s}$  [64].

It is important to note that the unfolding rate grows exponentially with the force. This is the hallmark of the Bell model. Further this equation describes how the barrier to unfolding is lowered by an external force. In contrast, in the Bell approximation the external force increases the folding barrier by an amount  $\Delta \Delta G_f^\ddagger = fx_f$ , where  $x_f = x_f(0)$  is the distance between the unfolded state and the transition state. Then the refolding rate reads as

$$k_f(f) = k_f(0) \exp(-fx_f/k_B T). \quad (57)$$

Within Kramer's theory, the folding rate at zero force,  $k_f(0)$ , is given by Eq. (56) where  $\Delta G_u^\ddagger(0)$  is replaced by the folding barrier  $\Delta G_f^\ddagger(0)$ . Although Eqs. (55) and (57) are very simple, as we will see below, they fit most experimental data very well. Using these equations and the force dependence of  $k_u(f)$  and  $k_f(f)$  one can extract the distances  $x_u$  and  $x_f$ , i.e. locate the TS.

### 3.5.2. Bell approximation: Force ramp case

Assuming that the force increases linearly with a rate  $v$  (i.e.  $f = vt$ ), the probability  $n(t)$  that the system is in the bound state can be calculated by solving the following equation

$$\frac{dn(t)}{dt} = -k_u[vt]n(t). \quad (58)$$

One can relate the distribution of unbinding forces  $P(f)$  to  $n(t)$  by noting that  $P(f|v)df = \frac{n(t)}{dt}dt$ . Then from Eq. (58) we obtain

$$P(f|v) = \frac{-k_u(f)}{v} n\left(\frac{f}{v}\right). \quad (59)$$

Solving Eq. (58) with the boundary condition  $n(0) = 1$  and substituting the expression of  $n\left(\frac{f}{v}\right)$  into Eq. (59) we have

$$P(f) = \frac{k_u(f)}{v} \exp\left\{\frac{k_B T}{x_u v} [k_u(0) - k_u(f)]\right\}, \quad (60)$$

where  $k_u(f)$  is given by Eq. (55). The most probable unbinding force, the maximum of the distribution, is then given by

$$f^* = \frac{k_B T}{x_u} \ln\left(\frac{vx_u}{k_u(0)k_B T}\right). \quad (61)$$

This equation was first obtained by Evans and Ritchie in their seminal paper [134]. It is based on Bell's assumption that the rates of a reaction depend exponentially on the applied force. Eq. (61) shows that the dissociation kinetics under an applied force are governed by the thermal dissociation rate constant  $k_u(0)$  and the length  $x_u$ , interpreted to be the width of the potential bound state. The logarithmic dependence of  $f^*$  on the pulling speed  $v$  was confirmed by extensive experiments and simulations [138,139].

### 3.5.3. Beyond the Bell approximation

The major shortcoming of the Bell approximation is the assumption that  $x_u$  does not depend on the external force. Upon force application, the location of TS should move closer to the native state thus reducing  $x_u$  (see Fig. 10) as postulated by Hammond in the context of chemical reactions of small organic molecules [140,141]. The Hammond behavior has been observed in protein folding experiments [142–144] as well as in simulations of RNA [145] and protein refolding/unfolding [139,146]. If the end-to-end distance is a good reaction coordinate for mechanical unfolding then deviations from Hammond's postulate should be an exception rather than the rule. For example, Schlierf and Rief [147] have successfully interpreted their AFM data obtained for the protein ddFLN4 assuming that the anti-Hammond effect is relevant, i.e., the transition state moves toward the unfolded state with increasing denaturating conditions but not to the folded one. The necessity for a more sophisticated theory than the Bell formula was clearly demonstrated experimentally by these authors [147].

Using Kramer's one-dimensional theory of diffusive barrier crossing in the presence of force [134,148,149] alternative microscopic models have been formulated that remove the limitations of the phenomenological Bell relation [150,151]. These microscopic theories still give analytical expressions for the rate of rupture at constant force and for the rupture force distribution in the presence of time-varying external forces. Recently, Szabo and coworkers [135] have developed a unified approach that casts the phenomenological and microscopic theories into a unified framework. We briefly describe the main results followed from this theory. Assuming that the entire dynamics can be projected onto a single reaction coordinate,  $x$ , then under the external force  $f$ , a molecule moves along the reaction coordinate  $x$  on a free energy surface,  $V(x) = V_0(x) - fx$ . This assumption holds if the relaxation times of other degrees of freedom are much shorter than the characteristic time scale of  $x$  [152]. The bare free energy  $V_0(x)$  is assumed to have a single well, a barrier at a distance  $x_u$  from the well center, and an activation free energy  $\Delta G^\ddagger$ . Escape over the barrier involves unfolding/unzipping of molecules. Two types of  $V_0(x)$ , which allow for exact solution, have been considered. In the case of a linear-cubic potential we have

$$V_0(x) = \frac{3\Delta G^\ddagger}{2} \frac{x}{x_u} - 2\Delta G^\ddagger \left(\frac{x}{x_u}\right)^3, \quad (62)$$

while the cusp potential has the following form

$$V_0(x) = \begin{cases} \Delta G^\ddagger \left(\frac{x}{x_u}\right)^2, & \text{for } x < x_u, \\ -\infty, & \text{for } x \geq x_u. \end{cases} \quad (63)$$

We first consider the case of a constant force accelerating the rate of molecular rupture. The Dudko–Hummer–Szabo theory [135] predicts that for a single-well free energy surface the rupture rate under a constant force  $f$  is

$$k_u(f) = k_u(0) \left(1 - \frac{vx_u f}{\Delta G^\ddagger}\right)^{1/v-1} \exp\left\{\frac{\Delta G^\ddagger}{k_B T} [1 - (1 - vx_u f / \Delta G^\ddagger)^{1/v}] \right\}. \quad (64)$$

When the scaling parameter  $v = 1$ , Eq. (64) reduces to the Bell expression equation (55). In this case  $k_u(f)$  is independent of  $\Delta G^\ddagger$  or  $\Delta G^\ddagger \rightarrow \infty$  in the Bell approximation. The values  $v = 2/3$  and  $v = 1/2$  correspond to the linear-cubic potential equation (62) and cusp potential equation (63), respectively. For these  $v$  values, the new theory has one obvious advantage that it allows for estimating  $\Delta G^\ddagger$  from the force dependence of the rupture rate. For  $v \neq 1$ , Eq. (64) is no longer valid as

$f$  approaches the critical value  $f_c = \Delta G^\ddagger / (\nu x_u)$  at which the barrier to rupture vanishes. This problem is caused by the use of Kramer's high-barrier approximation and can be circumvented by using the mean first passage time (MFPT) formula [148,153,154] to determine the force-dependent rate of escape from the well at  $x_{\min}$  to a point  $x_M$  beyond the barrier,

$$k_{MFPT}(f) = D \left[ \int_{x_{\min}}^{x_M} dx e^{\beta[U_0(x)-fx]} \int_{-\infty}^x dy e^{-\beta[U_0(y)-fy]} \right]^{-1},$$

where  $D$  is diffusion constant. In the high-barrier regime probed by low-to-intermediate forces  $f$ , the full MFPT expression is practically identical to the approximation of Eq. (64). However, substantial improvements can be achieved at high ramp speeds where forces near  $f_c$  are probed.

When the force is ramped up linearly with time ( $f(t) = vt$ ), one can repeat the same procedure as in the Bell case to get the distribution of rupture forces for any values of  $\nu$ . Namely, one has to solve Eq. (58) with  $k_u(f)$  taken from Eq. (64). Then using Eq. (59) we obtain [135]

$$P(f|\nu) = \frac{k_u(f)}{\nu} \exp \left[ \frac{k_u(0)}{\beta x_u \nu} - \frac{k_u(f)}{\beta x_u \nu} \left( 1 - \frac{\nu f x_u}{\Delta G^\ddagger} \right)^{1-1/\nu} \right], \quad (65)$$

where  $k_u(f)$  is the force-dependent rupture rate in Eq. (64). From this equation one can obtain the approximate analytical expressions for the mean rupture force  $f^* = \langle f \rangle = \int f P(f|\nu) df$  and the variance  $\sigma_f^2 = \langle f^2 \rangle - \langle f \rangle^2$ :

$$f^* \simeq \frac{\Delta G^\ddagger}{\nu x_u} \left\{ 1 - \left[ \frac{1}{\beta \Delta G^\ddagger} \ln \frac{k_0 e^{\beta \Delta G^\ddagger + \gamma}}{\beta x_u \nu} \right]^\nu \right\}, \quad (66)$$

$$\sigma_f^2 \simeq \frac{\pi^2}{6(\beta x_u)^2} \left[ \frac{1}{\beta \Delta G^\ddagger} \ln \frac{k_0 e^{\beta \Delta G^\ddagger + \tilde{\gamma}}}{\beta x_u \nu} \right]^{2\nu-2}. \quad (67)$$

Here  $\tilde{\gamma} = \gamma^2 - 3\psi''(1)/\pi^2 \approx 1.064$ , where  $\gamma \approx 0.577$  is the Euler–Mascheroni constant and  $\psi''(1) \approx -2.404$  is a particular value of the tetragamma function [155]. If the free energy barrier  $\Delta G^\ddagger$  is large compared to  $k_B T$  ( $\gamma$  in Eq. (66) is negligible) then in the phenomenological limit  $\nu \rightarrow 1$  the Evans–Ritchie result for  $f^*$ , Eq. (61), is recovered. Furthermore, since the variance  $\sigma_f^2$  is independent of the loading rate for  $\nu = 1$  one may use this as a criterion for the validity of the phenomenological theory. For  $\nu = 2/3$  these expressions reduce to those of Dudko et al. [151]. It is appealing to analytically continue  $\nu$  in Eqs. (64)–(67) to all  $\nu$  and thus  $\nu$  can be used as an additional fitting parameter to find the best agreement with experiments.

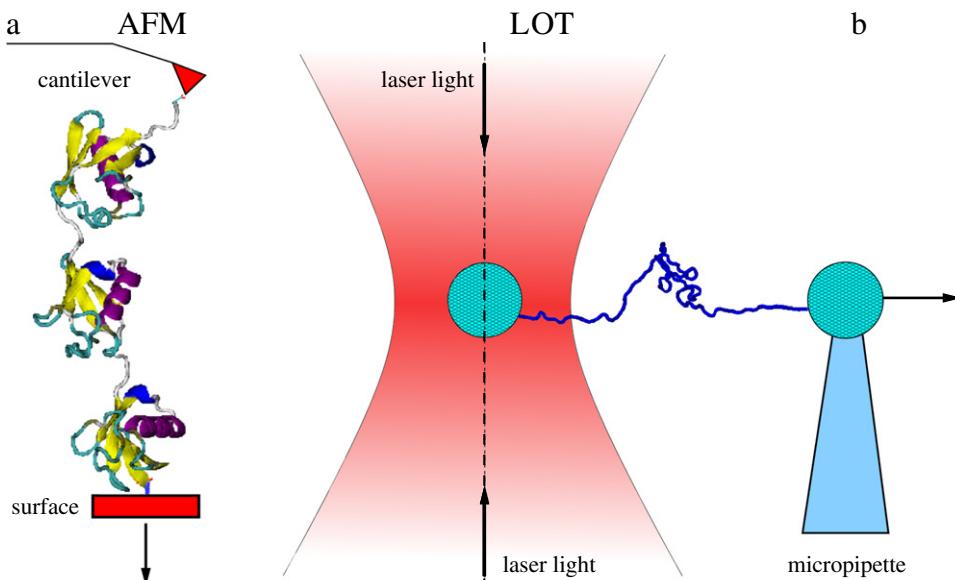
In the derivation of both phenomenological and microscopic models one has implicitly used the adiabatic assumption that the pulling speed is slow enough so that by the time the barrier is so low that Kramer's theory is invalid, and the survival probability,  $n(t)$ , is effectively zero [150]. For a single well potential  $n(t)$  then satisfies a first-order differential equation (58). If this adiabatic approximation is indeed valid and  $f(t) = vt$ , then the product  $v \ln n[t(f)]$  as a function of  $f$  is independent of  $v$  [156], where  $t(f) = f/v$ . In this case, the following relation between the constant-force experiments (measuring  $k(f)$ ) and constant speed experiments (measuring  $P(f|\nu)$ ) has been established [135]

$$k(f) = \frac{\nu P(f|\nu)}{1 - \int_0^f P(f'|\nu) df'} \quad \forall \nu. \quad (68)$$

This equation relates the two kinds of experiment in a model-free way. Finally, the theory, Eqs. (64)–(68), developed by Szabo et al. for molecular unfolding/unbinding can be applied to the reverse process, folding/binding where the force is reduced with time.

### 3.6. Constant force and constant distance ensemble

In any experiment on systems of small size the outcome depends crucially on whether the control parameter is the force or the extension. For example optical tweezers and AFM essentially control the position of the end monomer where a force is applied. On the other hand magnetic tweezer setups provide a constant force at the end monomer. In a constant force ensemble (CFE) the control parameter is the average extension. In most applications of AFM, the force is applied according to a linear ramp protocol  $x(t) = x(0) + vt$ . Here  $x(0)$  is the cantilever position at time  $t = 0$  and  $v$  is a constant velocity (typically very small). Hence such a system may be considered to be in quasi-static equilibrium and the appropriate ensemble is constant distance ensemble (CDE). In the thermodynamic limit both ensembles are expected to give the same results. However, single molecule experiments study systems of finite size and the results may depend on the ensemble. This has been reflected in experiments mentioned above where plateaus in the force–extension curve correspond to the stalling of fork at a GC-rich barrier region. In the constant force ensemble one must wait for a thermal fluctuation for unzipping to proceed. Lubensky and Nelson [52] have emphasized this effect in their theoretical work and the constant extension plateaus have been observed in experiments by Weeks et al. [157].



**Fig. 11.** Caricatures of two major techniques for manipulating single molecules: AFM (a) and LOT (b).

#### 4. Experimental techniques

There are a number of techniques for manipulating single molecules: atomic force microscopy (AFM) [158], laser optical tweezers (LOT) [159], magnetic tweezers [160], the bio-membrane force probe [161], etc. In this section we briefly discuss AFM and LOT which are often used to probe the mechanical response of proteins and DNA under external force. The description of other techniques may be found in reviews [7,162].

##### 4.1. Atomic force microscopy

In all techniques one terminal (end) of a biomolecule is anchored to a surface and the other terminal to a force sensor. The biomolecule is stretched by increasing the distance between the surface and the force sensor. In the AFM case the force sensor is a micron-sized cantilever (Fig. 11(a)). The force measured in experiments is proportional to the displacement of the cantilever.

If the stiffness  $k$  of the cantilever is known then a biomolecule experiences the force  $f = k\delta x$ , where  $\delta x$  is a cantilever bending which is detected by the laser. In general the resulting force versus extension curve is used in combination with theories for obtaining mechanical properties of biomolecules. The spring constant of the AFM cantilever tip is typically  $k = 10\text{--}1000 \text{ pN/nm}$ . The value of  $k$  and thermal fluctuations defines the spatial and force resolution of AFM experiments because when the cantilever is kept at a fixed position the force acting on the tip and the distance between the substrate and the tip fluctuate. The respective fluctuations are

$$\langle \delta x^2 \rangle = k_B T / k, \quad (69)$$

and

$$\langle \delta f^2 \rangle = k_B T. \quad (70)$$

For  $k = 10 \text{ pN/nm}$  and at the room temperature  $k_B T \approx 4 \text{ pN nm}$ , we have  $\sqrt{\langle \delta x^2 \rangle} \approx 0.6 \text{ nm}$  and  $\sqrt{\langle \delta f^2 \rangle} \approx 6 \text{ pN}$ . Thus AFM can probe forces larger than a few pN. For this reason it can serve as an ideal tool for studying relatively strong inter- and intra-molecular interactions, e.g., as in pulling nucleic acids and proteins.

##### 4.2. Laser optical tweezers

LOT operates on the same principle as AFM, but instead of the cantilever and the substrate, two micro-sized (typically  $1\text{--}3 \mu\text{m}$ ) polystyrene or silica beads are used to hold a molecule. One of them is optically trapped in the laser focus, and the other is sucked on to a micro pipette (Fig. 11(b)). The radiation pressure from a focused laser beam is able to trap neutral atoms. More details on the origin of optical trapping may be found in [163]. Using a piezoelectric actuator one can control the motion of the micro pipette and measure the deviation  $\delta x$  of its position from the trap center. To a very good approximation

the trapping potential is harmonic and the forces acting on the bead are  $f = k\delta x$ , where  $k$  is the stiffness constant of the trap. As follows from Stock force calibration or noise measurements [164], the typical values of  $k = 0.001\text{--}0.1 \text{ pN/nm}$ , which are about  $10^2\text{--}10^4$  times smaller than AFM tips. Since  $\sqrt{\langle \delta f^2 \rangle} \sim \sqrt{k}$  (Eq. (70)), the force resolution of LOT is at least 10 times better, that is on the order of 0.1 pN. However, as follows from Eq. (69),  $\sqrt{\langle \delta x^2 \rangle} \sim \frac{1}{\sqrt{k}}$  and LOT is thus less sensitive to the change in extension and this can mask minor unfolding events like short-lived intermediate states of a target biomolecule. One remarkable example of this is that the typical saw-tooth pattern of titin unfolding observed in AFM is not seen in LOT [165].

To summarize, AFM and LOT are practically identical except that the stiffness constants  $k$  are different in each setup. This difference affects the precision of the force and extension measurements. From Eqs. (69) and (70) we can see that the product

$$\sqrt{\langle \delta x^2 \rangle} \sqrt{\langle \delta f^2 \rangle} \approx k_B T \quad (71)$$

is independent of the spring constant. This relation is universal and implies that any experimental setup should obey the “uncertainty principle”: the more precise the determination is for the force the less precise the determination for location and vice versa. LOT is often used for studying nucleic acids and molecular motors.

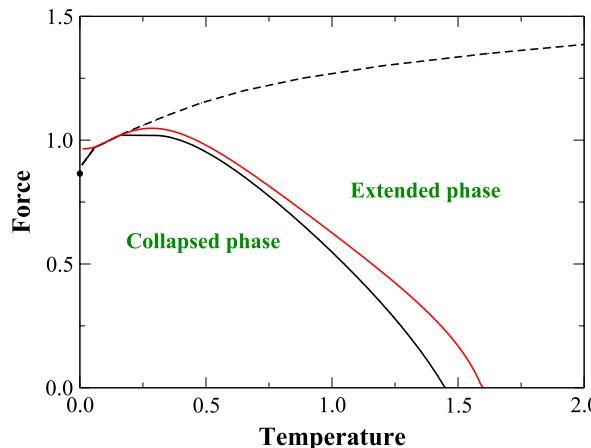
## 5. Mechanical unfolding of homopolymers

### 5.1. Reentrance transition

From polymer theory we know that a polymer chain will, depending on temperature, be in either a collapsed state or a swollen state [25]. The end-to-end distance  $R$  scales as  $N^\nu$ , where  $N$  is the chain length and  $\nu$  is the end-to-end distance exponent. In the collapsed state (low temperatures)  $\nu = 1/d$ , while at high temperatures  $\nu$  is given quite accurately by the Flory approximation equation (9) (actually this formula gives an exact value for  $d = 1, 2$  and  $4$ , whereas in  $d = 3$  it is only a good approximation). It should be noted that by varying temperature alone a polymer cannot acquire the conformation of a stretched state where  $\nu = 1$ . Hence, force not only “tilts” the free energy surface but also induces a new “stretched state” which is otherwise not accessible. In most experiments the reaction co-ordinate varies up to the contour length, hence it is more likely that the polymer may attain a stretched state before rupture takes place.

In some cases the force-extension curves obtained from experiments showed the existence of plateaus. For short polymer chains a stick release pattern with hysteresis has been found. The first observation is in good agreement with the mean-field theory proposed by Halperin and Zhulina [166]. The WLC model has been extensively used to understand the force-extension curves of biomolecules. However, the WLC model ignores excluded volume effects and attractive interactions between different chain segments and is thus only well suited for modeling the stretching of polymers in a good solvent [15]. Moreover, the mean-field analysis of Geissler and Shakhnowich suggested the existence of a re-entrant region for low temperatures similar to DNA unzipping [167]. The exact results derived by Marenduzzo et al. on fractal lattices showed that the mean-field theory is not valid [168].

Many efforts have been made [24,80,121–124,168–174] to obtain the phase-diagram of small chains on regular lattices using transfer matrix calculations, numerical simulations and exact enumeration technique on regular lattices. A self-attracting-self avoiding walks (SASAWs) on regular lattices models Fig. 5 a polymer chain in a poor solvent. Kumar et al. [123,124] used exact enumeration techniques to analyze the phase-diagram of a polymer in a poor solvent. The major advantage of this approach is that the complete finite-length partition function can be analyzed exactly. Furthermore scaling corrections can be taken into account by suitable extrapolation schemes enabling us to obtain accurate estimates in the thermodynamic (infinite length) limit [122]. To achieve a similar degree of accuracy using Monte Carlo simulations one typically has to use chains at least two orders of magnitude longer than in the exact enumerations [175]. Moreover, Monte Carlo studies were unable to access the low temperature regime (*i.e.* temperatures relevant in a biological context) corresponding to poor solvent. The greatest challenge facing exact enumerations is to increase the chain length. Using direct counting algorithms the time required to enumerate all the configurations increases as  $\mu^N$ , where  $\mu$  is the connective constant of the lattice ( $\mu \approx 2.638$  on the square lattice). So even with a rapid increase in computing power only a few more terms can be obtained each decade. Until now most of the exact results for models of small proteins were confined up to chain length 30 or so [169,171,172]. Combining parallel processing and transfer matrix calculations Kumar et al. [123,124] were able to extend the enumeration, not just by one or two steps, but to almost double the chain length namely up to 55 steps. In their study, one end of the polymer is attached to an impenetrable neutral surface (there are no interactions with this surface) while the polymer is being pulled from the other end with a force acting along the  $x$ -axis. Note that the SASAW does not extend beyond either end-point so the  $x$ -coordinate  $x_j$  of the  $j$ th monomer is restricted by  $0 = x_0 \leq x_j \leq xN = x$ . At first this restriction may appear artificial but it does in fact model the experimental setup. In typical experiments single proteins (or small pieces of DNA) are attached to the surface of beads (using ligand molecules). The beads are very large compared to the size of the proteins and the surface of the bead is thus well approximated by a flat surface. We introduce Boltzmann weights  $\omega = \exp(-\epsilon/k_B T)$  and  $u = \exp(-F/k_B T)$  conjugate to the nearest neighbor



**Fig. 12.** The ‘f-T diagram’ for flexible chains as obtained from the position of the peak in the contact fluctuation curves for  $N = 55$  by setting  $\epsilon = 1$  and  $k_B = 1$  [124]. The solid black curves are obtained by fixing the force and varying the temperature. The solid red curve is obtained by fixing the temperature and varying the force. Figure has been taken from Ref. [124].

interactions and force, respectively, where  $\epsilon$  is the interaction energy. The finite-length partition function of the flexible chain can be written as [123,124]

$$Z_N(F, T) = \sum_{\text{all walks}} \omega^m u^x = \sum_{m,x} C(N, m, x) \omega^m u^x, \quad (72)$$

where  $C(N, m, x)$  is the number of SASAWs of length  $N$  having  $m$  nearest neighbor contacts and whose end-points are a distance  $x = x_N - x_0$  apart. The partition functions of the CFE,  $Z_N(F, T)$ , and CDE,  $Z_N(x, T) = \sum_m C(N, m, x) \omega^m$ , are related by  $Z_N(F, T) = \sum_x Z_N(x, T) u^x$ . The free energies are evaluated from the partition functions

$$G(x) = -k_B T \log Z_N(x) \quad \text{and} \quad G(F) = -k_B T \log Z_N(F). \quad (73)$$

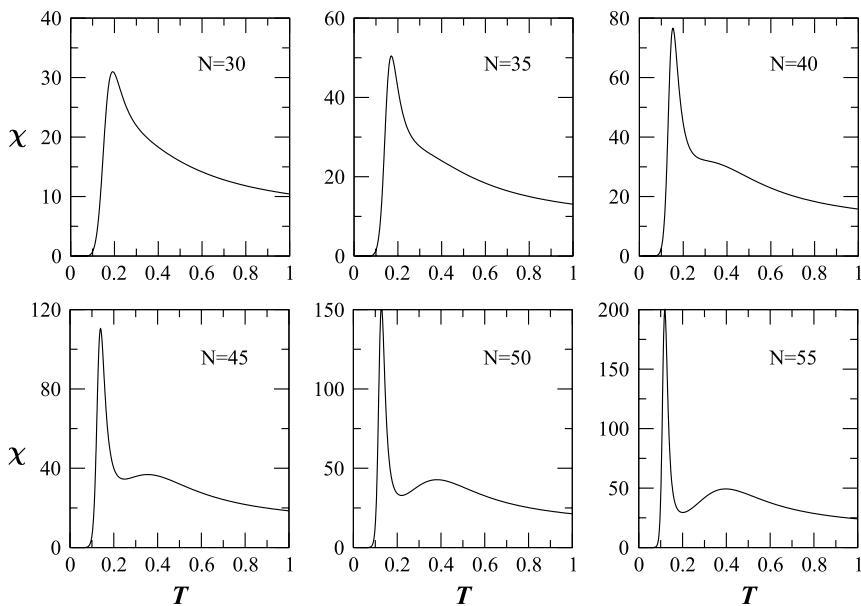
Here  $\langle x \rangle = \frac{\partial G(F)}{\partial F}$  and  $\langle F \rangle = \frac{\partial G(x)}{\partial x}$  are the control parameters of the CFE and CDE, respectively. At low temperature and low force the polymer chain is in the collapsed state and as the temperature is increased (at fixed force) the polymer chain undergoes a phase transition to a swollen state. The value of the transition temperature (for a fixed value of the force) can be obtained from the fluctuations of the number of non-bonded nearest neighbors (which can be calculated exactly for any finite  $N$  up to 55). The fluctuations are defined as  $\chi = \langle m^2 \rangle - \langle m \rangle^2$ , with the  $k$ th moment given by

$$\langle m^k \rangle = \sum_{m,x} m^k C(N, m, x) \omega^m u^x / \sum_{m,x} C(N, m, x) \omega^m u^x.$$

In Fig. 12, the force–temperature (f-T) ‘diagram’ for flexible chains as obtained from the peak positions (note that the true phase diagram should be obtained by extrapolating the data to the  $N \rightarrow \infty$  limit) is shown [124]. The qualitative features of the f-T diagram remains largely the same as those observed in previous studies [167,169]. In Fig. 12 we have shown the transitions as obtained by fixing the force (black curve) or fixing the temperature (red curve).

The transition (solid) line separates the collapsed state from the unfolded state. In the case of flexible polymer chain, they observed the phenomenon of re-entrance, *i.e.* the critical force goes through a maximum as the temperature is lowered. For example, one can see that the polymer chain at fixed force (say  $F = 0.9$ ) is in the extended state at low temperature. With an increase in temperature the chain is found to be in the collapsed state. After a further rise in temperature the chain again acquires conformations in the extended (swollen) state. Similar re-entrance behavior has also been found in other models of flexible polymer chains [91,171]. This indicates that the polymer goes from a high entropy state to a low entropy state under the application of force. In the panels of Fig. 13 the emergence of two peaks in the fluctuation curves with increasing  $N$  at fixed force  $F = 1.0$  can be seen. The twin-peaks reflect the fact that in the re-entrant region as we increase  $T$  (with  $F$  fixed) the polymer chain undergoes two phase transitions. Note that the twin-peaks are not apparent for small values of  $N$ . The other notable feature is that in the fixed force case a new line of transitions is observed from the extended state to the stretched state which is solely induced by the applied force. In contrast to the lower phase boundary (collapse transition), where the force decreases with temperature, the upper boundary (stretching) shows that the force increases with the temperature and is in fact a crossover line [176]. In the thermodynamic limit, this line does not exist [177].

The origin of re-entrance can be explained by using a phenomenological argument near  $T = 0$ . By equating the free energy of the folded state and the free energy of the stretched state due to the force, one can show that the critical force for the unfolding  $F_c(T) = -\epsilon + (1/\sqrt{N})\epsilon + TS$ . The dominant contribution to the free energy is due to the non-bonded nearest neighbor interactions. The second term is a surface correction term and it vanishes in the thermodynamic limit. It



**Fig. 13.** The fluctuations in the number of contacts as a function of temperature for fixed force  $F = 1.0$ . Figure has been taken from Ref. [124].

should be noted that for a flexible polymer chain the entropy associated with a globule is finite and hence a positive slope  $\frac{dF_c}{dT}$ , at  $T \rightarrow 0$ , confirms the existence of re-entrance. AFM experiments at constant force should be able to confirm this phenomenon.

### 5.2. Emergence of intermediates

Marenduzzo et al. showed the existence of intermediate states in a homo-polymer in 2 and 3 dimensions [171,172]. Their exact results for small chains showed that intermediates are a finite size and have a low temperature effect (poor solvent condition) and as the temperature increases the multi-step character washes out. The qualitative behavior remains the same for a continuum model in 3 dimensions. Kumar et al. [123,124] considered flexible chains of length up to 55 steps and calculated the average elongation using the following relation

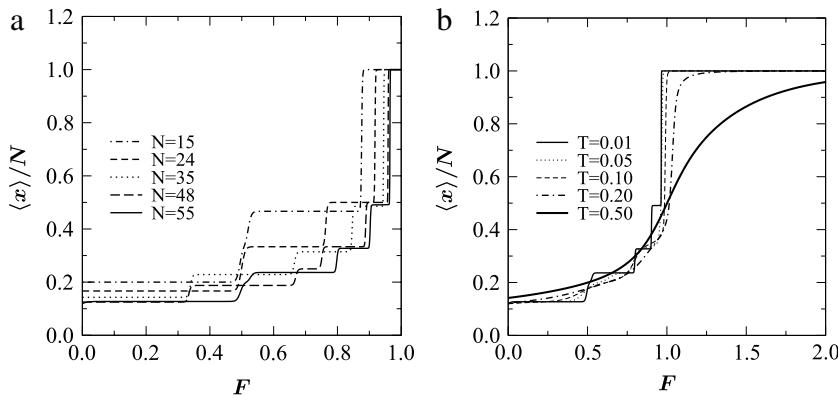
$$\langle x \rangle / N = \frac{1}{N} \frac{\sum_{m,x} x C(N, m, x) \omega^m u^x}{\sum_{m,x} C(N, m, x) \omega^m u^x}.$$

A simple theoretical argument for the observed behavior is that at low temperatures, where the entropy  $S$  (per monomer) of the chain is quite low the dominant contribution to the free energy

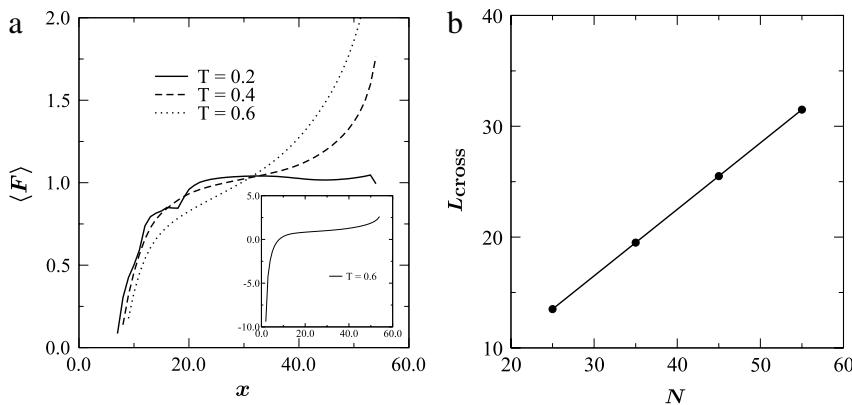
$$G = N\epsilon - \sigma(N, F)\epsilon - NTS \quad (74)$$

is from the non-bonded nearest neighbor interactions  $N\epsilon$ . The second term is the surface correction term and it vanishes in the thermodynamic limit. However, for finite  $N$  the system has many degenerate states depending on the shape of the globule. This leads to a surface correction term  $\sigma(N, F)$  which is a function of  $N$  and  $F$ . If  $F = 0$ , the shape of the globule is like a square (in 2d) and the surface correction term  $\sigma(N, 0)$  will be minimized and equal to  $2\sqrt{N}$ . In the constant force ensemble, there is a force induced additional contribution proportional to the extension of the globule that along with  $\sigma(N, F)$  stabilizes the intermediate states (Fig. 14). When the temperature increases the multi-step character of the force extension is washed out due to increased contributions from entropy.

Kumar et al. [123,124] extended their study to the CDE, best suited to the analysis of experiments performed using an apparatus like AFM. Enhanced knowledge of the density of states not only provided a unified treatment of both pulling and stretching, but also revealed the existence of a crossover length in the force-distance curve. It shows that when the distance between the first and last monomer (where force is applied) is less than the average size of the coil (without force), one needs a compressing force instead of a pulling force. The qualitative behavior of the curve remains the same as that found from experiments and computer simulations. The flat portion of the curve shows the average force needed to unfold the chain. Such plateaus have been seen in experiments. From Fig. 15(a) one can see that the force required to obtain a given extension initially decreases with the temperature. But beyond a certain extension (close to 30 in this case) the required force increases with temperature. They noted that the curves cross each other at a critical temperature. They identified this as a crossover point. The plot of the position of the crossover point  $L_{\text{cross}}$  as a function of length  $N$  of the polymer chain



**Fig. 14.** The average scaled elongation  $\langle x \rangle/N$  vs  $F$  at  $T = 0.01$  for various lengths (a) and temperatures at length  $N = 55$  (b). Figure has been taken from Ref. [123].



**Fig. 15.** Plot of the average force  $\langle F \rangle$  vs the elongation  $x$  at various temperatures  $T$  for  $N = 55$  (a) and the cross-over length vs  $N$  (b). Figure has been taken from Ref. [123].

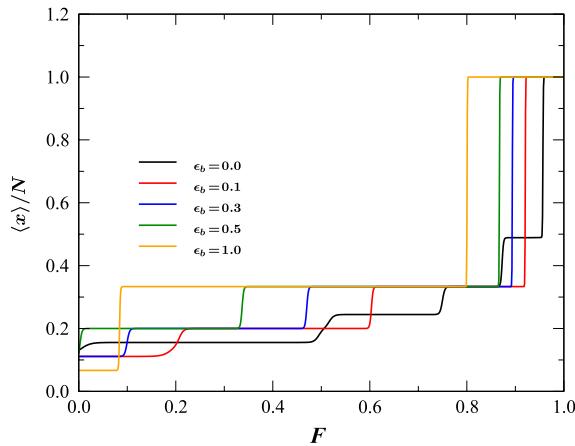
is shown in Fig. 15(b). The crossover extension increases linearly with length. This shows that above this point the chain acquires the conformation of the stretched state. The increase in force with temperature generates a tension in the chain sufficient to overcome the entropic effect. Since the contribution of this term to the free energy is  $TS$ , more force is needed. The exact analysis for finite chains shows that applying a force at first favors taking the polymer from the folded state to the unfolded state. However, a second unfolding occurs when the unfolded chain attains the stretched state and one requires more force.

### 5.3. Role of stiffness

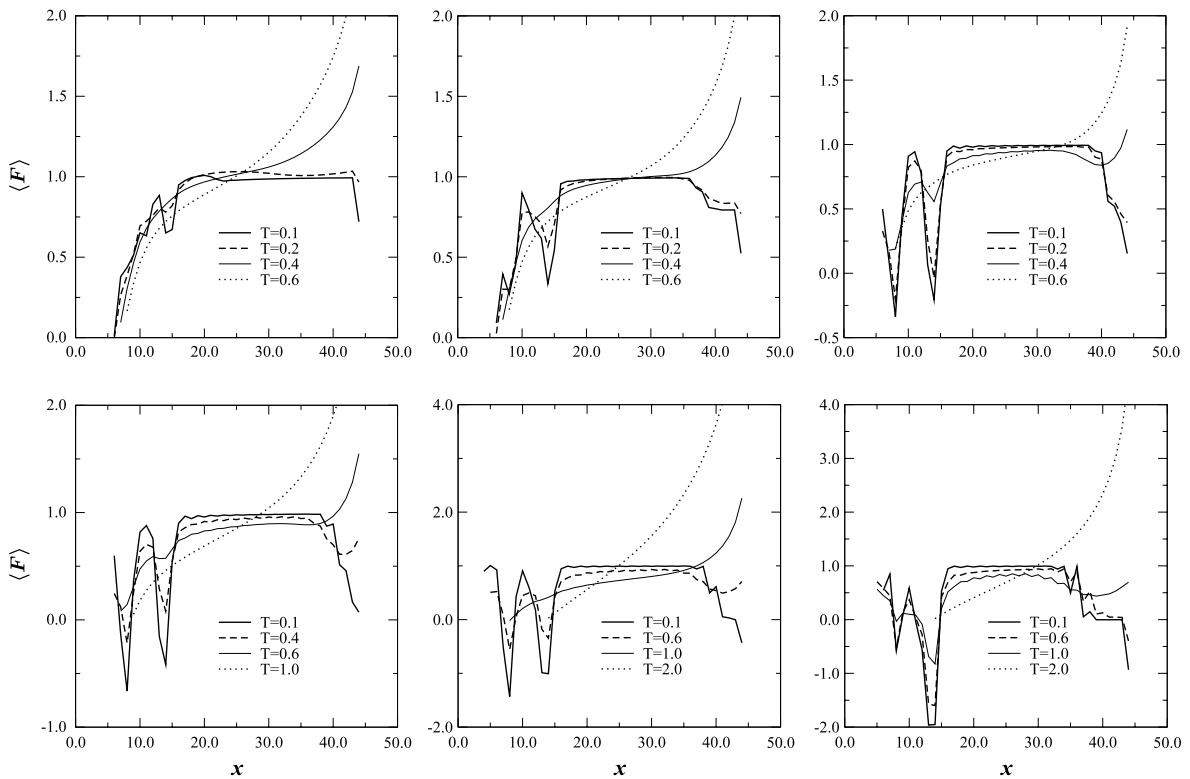
Semi-flexible polymers can be modeled by associating a positive energy  $\epsilon_b$  with each turn or bend of the walk [123, 124, 178, 179]. The corresponding Boltzmann weight is  $\omega_b = \exp(-b\epsilon_b)$ , where  $b$  is the number of bends in the SASAW. Kumar et al. enumerated all walks, but because of the additional parameter  $\omega_b$ , they were restricted to 45 steps. For a semi-flexible polymer chain, a stretched state may be favored by increasing the stiffness. The phase diagram for semi-flexible chains is now well established. It has three states namely (i) an open coil state at high temperature, (ii) a molten globule at a low temperature and low stiffness and (iii) a ‘frozen’ or ‘folded’ state at a low temperature and large stiffness [169, 178, 179]. We note that while the flexible and semi-flexible  $F-T$  phase-diagrams are qualitatively similar [169], the re-entrant behavior is suppressed because of stiffness and becomes less pronounced with increasing bending energy.

In Fig. 16, we show the force-extension curves for semi-flexible chains with bending energy ranging from  $\epsilon_b = 0.0$  (flexible chains) to  $\epsilon_b = 1.0$  at low temperature  $T = 0.01$  for chain length  $N = 45$ . One can observe that as the bending energy is increased the chain undergoes fewer intermediate transitions between the compact state (low force) and the fully stretched state (high force). In particular the flexible chain ( $\epsilon_b = 0.0$ ) has five plateaus while the chain with  $\epsilon_b = 1.0$  has only three.

In the CFE the probability distribution of the end-to-end distance has a “saw-tooth” like behavior corresponding to the intermediate states during unfolding [169]. Therefore, it is important to study the effect of stiffness on force-extension curves in the CDE. The force extension curves shown in Fig. 17 have striking differences to the flexible ones. At low temperatures we see strong oscillations which vanish as the temperature is increased. Since the polymer chain has “frozen conformations” like  $\beta$ -sheets (the zero-force limit of which describes zig-zag configurations inscribed in a square), it takes



**Fig. 16.** Plot of the average scaled elongation  $\langle x \rangle / N$  vs  $F$  for semi-flexible chains with bending energy ranging from  $\epsilon_b = 0.0$  (flexible chains) to  $\epsilon_b = 1.0$  at low temperature  $T = 0.01$  for chain length  $N = 45$ . Figure has been taken from Ref. [124].

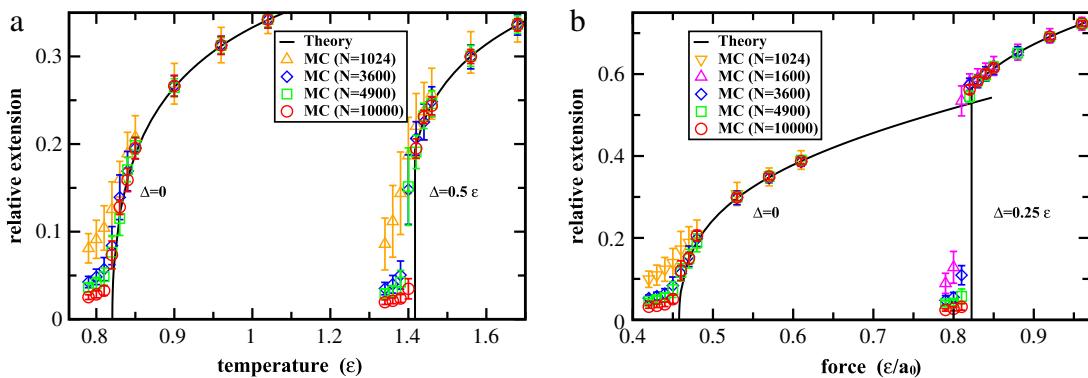


**Fig. 17.** Plot of  $\langle F \rangle$  vs  $x$  for a semi-flexible chain with bending energy (from left to right and top to bottom)  $\epsilon_b = 0.0, 0.1, 0.3, 0.5, 1.0$  and 2.0 at different temperatures  $T$  for  $N = 45$ . Figure has been taken from Ref. [124].

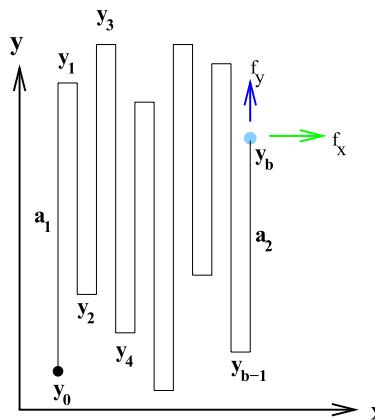
more force to unfold a layer. When about half a layer has been opened, the bending energy favors a complete stretching of the layers and hence the force decreases (see Fig. 17). This phenomenon allows us to probe a molecule like Titin which has a similar  $\beta$ -sheet structure [2].

#### 5.4. Nature of transition

We discussed the collapse transition that occurs at the so-called  $\theta$ -temperature in Section 2. The exact nature of the collapse transition, however, is not yet completely settled, despite decades of extensive effort [16, 180, 181]. This transition can be studied by mapping to the tricritical point of the  $\phi^4 - \phi^6 O(n)$  field theory in the  $n \rightarrow 0$  limit [182–184] and is expected to be second-order in two dimensions and beyond [185]. Furthermore, the exponents of the temperature-driven collapse and adsorption of a 2D polymer grafted on a linear boundary has been obtained by analytical calculations [186, 187].



**Fig. 18.** (a) Temperature-extension curve at  $f = 0$  (b) force extension curve at  $T = 0.590928\epsilon$ . Solid lines are analytical predictions, and symbols are MC simulation results for different system sizes. Figure has been taken from Ref. [170].



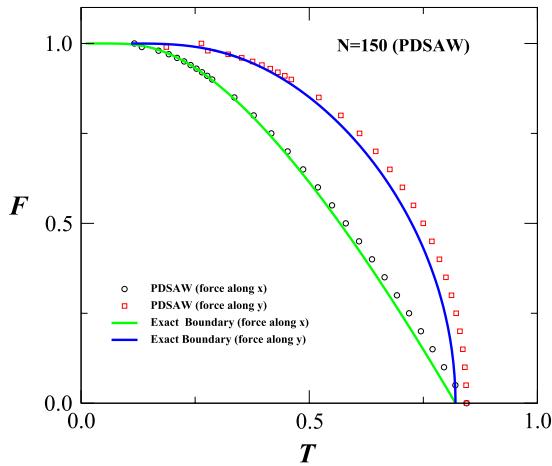
**Fig. 19.** Schematic representation of a PDSAW on the square lattice with one end kept fixed and the other end subjected to a pulling force along ( $f_x$ ) and perpendicular ( $f_y$ ) to the preferred direction. At low temperature the conformation mimics the  $\beta$ -sheet.

However, experimentally observed collapse transitions of both relatively flexible [188] and semi-flexible [189] 3D polymers are more like first-order transitions. Recent Monte Carlo (MC) simulations [190] are also in favor of the latter interpretation. According to mean-field theory [166] the force-induced collapse transition is first order in all dimensions. This claim is confirmed by MC simulation in 3D [191], while in 2D the transition is argued to be second order by MC simulation [191] and scaling analysis [121,171]. In order to understand the collapse transition more deeply Zhou et al. [170] investigated a 2D partially directed lattice model of a polymer chain that is exactly solvable (Fig. 6). As the free energy density of this model system can be calculated exactly, they could draw definite conclusions concerning the nature of the collapse transition and showed that the collapse transition is second order for a polymer with zero bending stiffness, while it changed to first order when the bending stiffness is non-zero (Fig. 18). This is also in agreement with earlier mean-field calculations of Doniach et al. [178]. Therefore, bending energies have a dramatic effect on the cooperativity of the 2D globule-coil transition. They further substantiated their predictions by performing extensive MC simulations. Now it is experimentally feasible to confine polymers to 2D mobile surfaces (see [192]). Therefore, future experiments will be able to verify the theoretical predictions.

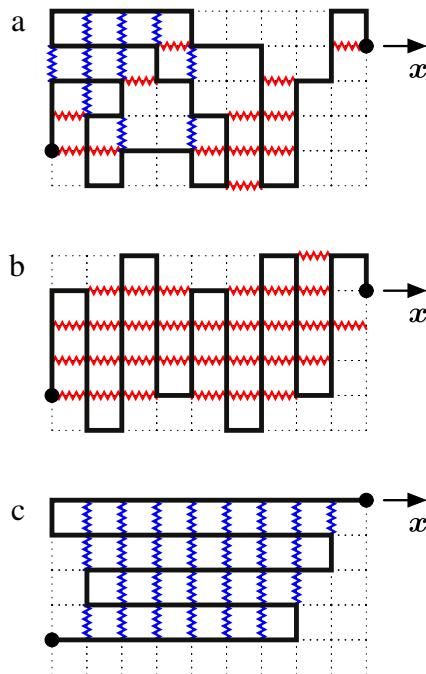
### 5.5. Dependence of phase diagrams on pulling direction

Theoretical modeling have reproduced experimentally observed force-extension curve quite accurately using, say, the FJC or WLC models. However, the difference in the resistance of a protein to unfolding solely due to a change in the pulling geometry [193–196] cannot be understood by these theoretical models and they thus provide only limited insight into biomolecules. Here, it will suffice to point out that these models are isotropic in nature and ignore crucial excluded volume effects in their description. They are thus only well suited to modeling the stretching of polymers (isotropic) in a good solvent. Notably all proteins studied so far are highly anisotropic in shape and interactions [197]. One therefore expects that their mechanical properties, e.g., compressibility and Young's modulus etc. depend on the direction and orientation [198].

The molecular origin of protein resistance to unfolding has recently been numerically investigated by Kumar and coworkers [173,174] by considering a partial directed walk model of a polymer. Their results based on considerable chain



**Fig. 20.** The globule-coil phase boundary in the force-temperature plane. The response to the force when applied along the  $y$ -direction (squares) is distinctly different to the case where the force is applied along the  $x$ -direction (circles). The continuous lines show the exact phase boundaries which are in excellent agreement with our numerical results for finite  $N = 150$ . Figure has been taken from Ref. [173].

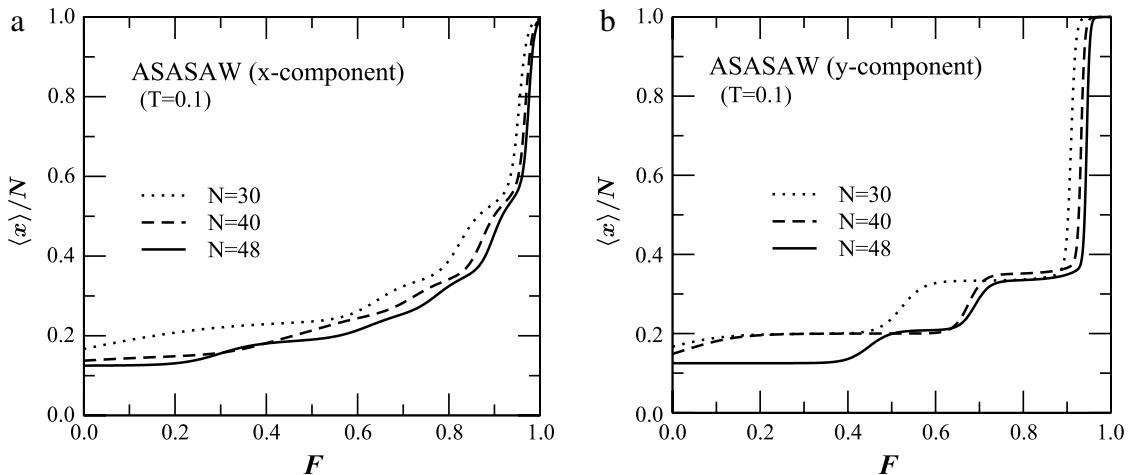


**Fig. 21.** Schematic representation of ASASAWs on the square lattice. (a) Different nearest neighbor interactions along the  $x$ - and  $y$ -directions introduce anisotropy in the model polymer. (b) represents ground state conformations which dominate the partition function similar to Fig. 19 when subjected to force along the  $x$ -direction. (c) Same as (b) but resembles Fig. 19 when subjected to force along the  $y$ -direction.

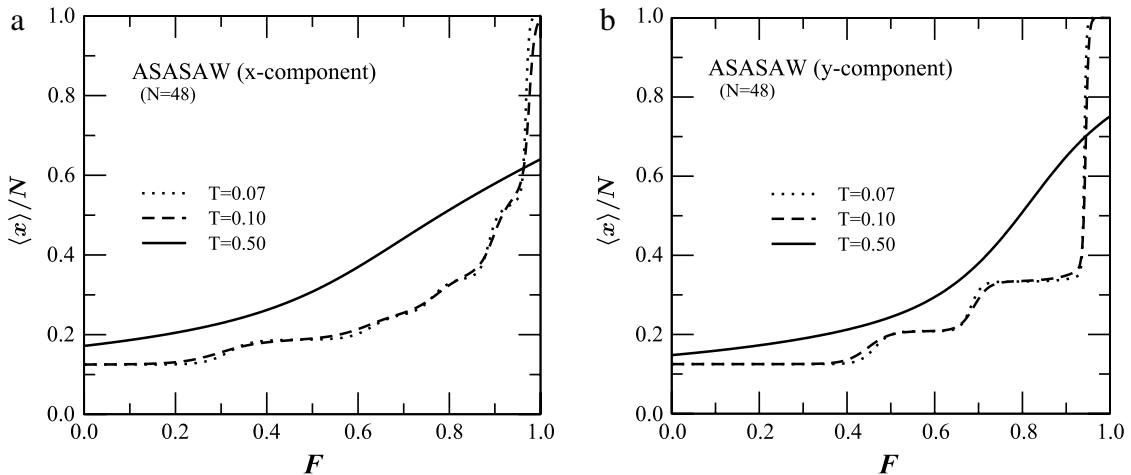
size of length up to  $N = 150$  indicate that when a chain is being pulled along the  $x$ -direction the nature of the transition is akin to unzipping, but when force is applied along the  $y$ -direction the transition is akin to shearing (Fig. 18).

They provided for the first time the analytical solution of the model and showed that a change in the pulling direction gives rise to different phase-diagrams even in the thermodynamic (infinite chain length) limit. In order to do so, they calculated the phase boundary between the collapsed and the extended phase which can be obtained by calculating the macroscopic shape of the collapsed phase at low temperature. It is convenient to work with the reduced variables  $\omega_x = \exp(\beta F_x)$ ,  $\omega_y = \exp(\beta F_y)$  and  $u = \exp(\beta \epsilon_u)$ , where  $\beta$  is the inverse temperature. The energy of the configuration, which mimics the  $\beta$ -sheet, is (see Fig. 19)

$$E = -u\epsilon + \frac{\epsilon}{2} + (a_1 + a_2 + 2b) + \frac{\epsilon_u}{2} \sum_{j=0}^{b-2} |y_{j+2} - y_j|. \quad (75)$$



**Fig. 22.** (a) The average scaled elongation  $\langle x \rangle / N$  as a function of the force with interactions along the  $x$ -direction at fixed low temperature  $T = 0.01$  for different chain lengths; (b) Same as (a) but for interactions along the  $y$ -direction. Figure has been taken from Ref. [173].



**Fig. 23.** (a) The average scaled extension  $\langle x \rangle / N$  along the  $x$ -direction as a function of the force at different temperatures with chain length  $N = 49$ ; (b) As above but the interactions are along the  $y$ -direction. Figure has been taken from Ref. [173].

They found an expression for  $\omega_x$  in terms of  $\omega_y$  and  $u$  as

$$\omega_x^c = \frac{u}{(u-1)\sqrt{\omega_y}} [(\sqrt{u} - \sqrt{\omega_y})(\sqrt{u}\sqrt{\omega_y} - 1)]. \quad (76)$$

From this one can obtain the complete three-dimensional phase boundary. For  $F_y = 0$ , Eq. (76) reduces to the following simple expression for the critical force when pulling along the  $x$ -direction

$$\omega_x^c = \frac{u}{(\sqrt{u} + 1)} (\sqrt{u} - 1). \quad (77)$$

This is similar to the equation found by using transfer matrix methods, but the method does not work when a force is applied along the arbitrary direction. However, Eq. (76) derived by Rajesh et al. [173] is sufficiently general to yield an expression for the critical force when pulling in the  $y$ -direction

$$\omega_y^c = \left[ \frac{1 + u^2 + \sqrt{1 + 2u^2 - 4u^3 + u^4}}{2u^{\frac{3}{2}}} \right]^2. \quad (78)$$

The force-temperature phase diagrams obtained from these expressions are shown in Fig. 20. Their numerical analysis based on a 150 steps walk showed the existence of intermediate states. Moreover, in the CDE the force-extension curves showed unzipping like transitions characterized by smooth plateaus when a chain is being pulled along the  $x$ -direction. When the chain is being pulled along the  $y$ -direction the force-extension curve showed saw-tooth like behavior indicating that the transitions are akin to shearing.

A more realistic model of polymers is SASAWs. However, this model is isotropic in nature and hence cannot be used to describe the effect of a change in pulling direction. Rajesh et al. [173] introduced anisotropy into the model by assigning different nearest neighbor interactions along  $x$ - and  $y$ -directions as shown in Fig. 21. This is in accordance with real proteins where the interactions along different directions can differ by orders of magnitude. Hence by varying the strength of the interactions one can change the anisotropy of the system. The force extension curves plotted in Fig. 22 not only show the existence of intermediate states but also the emergence of new states when the pulling direction is changed. As the temperature is increased, they observed that the intermediate state washed out due to the resulting increase in entropy. It was further seen that system attains a higher stability against the force along the  $y$ -direction and that intermediate states survive even at high temperatures Fig. 23.

## 6. Mechanics of DNA

The intra- and inter-molecular forces of DNA are central to understanding its structure and functional behavior. The understanding of these molecular forces was until recently based on indirect physical and thermodynamic measurements such as crystallography, light scattering and nuclear magnetic resonance spectroscopy [36]. For the direct measurement of interaction forces it is essential that the state of the system be monitored while an independent force is applied [17,199–202]. For example, osmotic pressure techniques have been applied to measure the nonspecific inter-molecular forces between DNA helices. For a process such as molecular recognition a direct measurement between individual molecules is required. With the advent of techniques such as optical tweezers, magnetic tweezers, the atomic force microscope etc. discussed in Section 4, it is now possible to measure these interactions. In this section we are going to discuss the forces associated with the intra-chain interactions that give rise to the elasticity of DNA and inter-chain interactions associated with Watson–Crick base pairing between complementary strands of DNA. So far there is no clear understanding about the origin of the stiffness in DNA. According to one view it is the tendency to maximize the base stacking which is the dominant contribution resisting DNA deformations such as lateral bending, torsion and stretching. In another view mutual inter-phosphate repulsion may give rise to DNA rigidity by resisting deformed conformations. It is also thought that there is a possibility that DNA stiffness reflects the equilibrium between forces trying to compress the DNA and inter-phosphate repulsion [203]. Here our discussions will focus mainly on the response of force on DNA namely DNA stretching and slippage, twisting of DNA and unzipping of DNA.

### 6.1. DNA stretching: Double stranded DNA

Experimental studies conducted by Cluzel et al. [204] and Smith et al. [17] have revealed unusual elastic properties of DNA. It was found that dsDNA is a semi-flexible macromolecule while single stranded (ssDNA) behaves like a flexible polymer chain. Due to intra-strand electrostatic repulsions, the chain is stiff over short length scales (the persistence length). In ssDNA the persistence length consists of about 7 to 8 bases while in dsDNA it consists of typically 50 to 100 base pairs. In the low force regime ( $< 10$  pN) the elasticity of dsDNA is entropy dominated. At small forces (10–60 pN) dsDNA obeys Hooke's law and the WLC model can describe the experimentally observed force extension curves (Fig. 24). In the high force regime ( $\sim 65$  pN) it was found that the dsDNA molecule can be overstretched about 1.7 times the B-form contour length and a phase transition occurs from the B-form to a stretched or S-form [17,22,204–207]. An explanation of this regime is attributed to the short range nature of base pair stacking interactions. At high forces the stacking potential can no longer stabilize the B-form configuration of dsDNA and the stacked helical pattern becomes distorted. Moreover, the slope of the increase in force after the over-stretching transition has been shown to depend on the rate at which the DNA molecule is stretched. At a low pulling rate the force-extension curve matches that of ssDNA in this regime. At high pulling rates the slope at this point increases, but at forces greater than 120 pN it again matches the ssDNA curve. This description applies only to the case in which one strand of DNA is allowed to rotate freely so that DNA can untwist while being stretched. If the DNA strand is torsionally constrained a much broader over-stretching transition occurs at 110 pN.

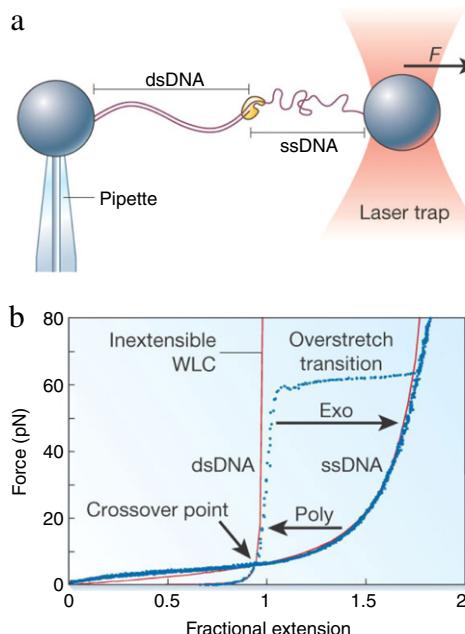
A two state model proposed by Cizeau and Viovy [208], inspired by models introduced earlier in the context of DNA melting [35,181], is capable of reproducing the B- to S-transition. In this approach the dsDNA is described as a chain of  $N$  monomers (base-pairs) which can be in the states B (energy  $E_B$ , length  $l_B$ ) or S (energy  $E_S$ , length  $l_S > l_B$ ) [209]. Each base pair is associated with a spin +1 or -1 for the B and S state respectively. The energy of the chain can then be written as

$$H = \sum_{i=1}^N (E_{Si} - fl_{Si}) + \frac{\omega}{2} \sum_{i=1}^N (1 - S_i S_{i+1}), \quad (79)$$

where  $\omega$  represents a “domain wall” energetic cost of a B-S frontier. Up to an additive constant,

$$H = -\frac{\omega}{2} \sum_{i=1}^{N-1} S_i S_{Si+1} - \frac{1}{2} (\Delta E - f \Delta l) \sum S_i \quad (80)$$

where  $\Delta E = E_S - E_B$  and  $\Delta l = l_S - l_B$ . Eq. (80) resembles an Ising model with magnetic field  $h = \Delta E - f \Delta l$ . Ashan et al. proposed an Ising type model of denaturation and coupled it to the mesoscopic elastic deformations [210].



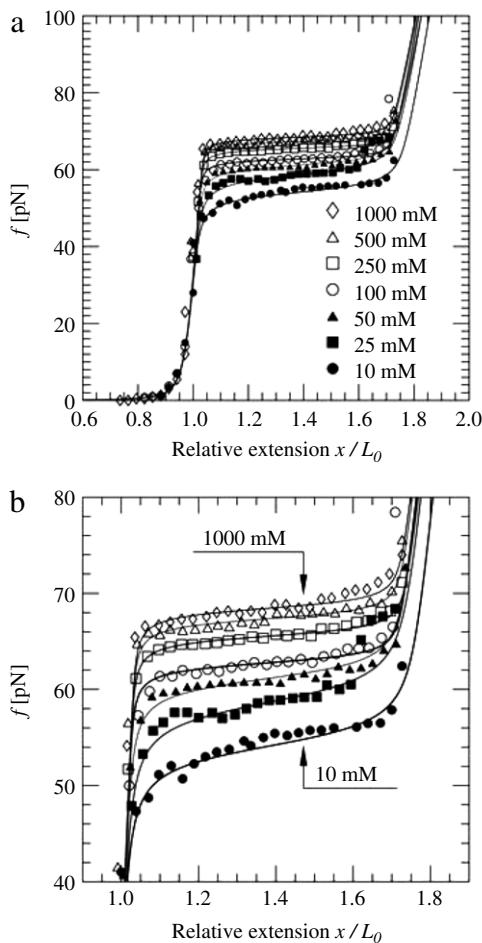
**Fig. 24.** Stretching a dsDNA and ssDNA by using an optical tweezer, where the external force is exerted by the optical trap on the trapped bead. Force extension curves have been taken from [205].

Source: By permission from Ref. [205].

However the elasticity as defined in this model is too approximate since it assumes that the elastic parameters of BDNA and SDNA are similar. Podgornic et al. [211] studied the effect of inter-segment interactions on the effective bending and stretching moduli of a semi-flexible polymer chain with a finite stretching modulus. For an interaction potential of a screened Debye–Hückel type, renormalization of the stretching modulus is derived on the same level of approximation as the celebrated Odijk–Skolnick–Fixman result for the bending modulus [212,213]. The presence of mesoscopic inter-segment interaction potentials couples the bending and stretching moduli in a manner different from that predicted by macroscopic elasticity theory. This leads to a dependence of the elastic module of a flexible polyelectrolyte on the ionic conditions: stretchability. The theory proposed by them and its consequences compare favorably with experiments on DNA bending and stretching at not too low salt conditions. In another approach [214] a hybrid model combined the Ising model approach of Ashan et al. [210] and the elasticity theory by Podgornic et al. [211]. It was found that this model fit the experimental data very well [214].

Rouzina and Bloomfield have developed a theory that predicts that DNA melting occurs during the over-stretching transition [215,216]. The proposed transition takes place in two stages. First it undergoes a cooperative transition and an equilibrium melting process occurs. At this stage a non-zero fraction of bases remains bound. In the second stage the remaining base pairs are broken. However, the removal of the bonds separating the melted domains represents an irreversible process in which the two strands completely separate. Because the process is irreversible the two strands are unable to sample bound and unbound states so the entropy gained by the strands upon melting is not meaningful. Under these conditions it is the enthalpy of  $15k_B T/\text{bp}$  rather than the free energy of  $2k_B T/\text{bp}$  that determines the stretching force. Thus the force required to completely separate the two strands is much greater than the over-stretching force. A series of experiments by Wenner et al. [217] showed that changes in solution conditions, which favor the melting of the double helix, reduce the force needed for over-stretching. Wenner et al. interpret their results as indicating that S-DNA is tension induced melted DNA. However, it is also a 1d model, which cannot describe the spatial effect of the double helix structure of dsDNA on the state transitions or how internal forces in the DNA molecule work during this transition. Recently, Shokri et al. [218] have studied the overstretching transition in presence of the solvent glyoxal. It is pertinent to mention here that the solvent glyoxal used in the above experiments reacts with DNA. It introduces an additional ring to the G-base (to form a tricyclic compound, i.e. glyoxaldG), thereby sterically preventing the G–C base pair re-annealing [219,220]. They held DNA overstretched at different lengths during the transition in presence of glyoxal. If stretching involves strand separation, then force melted bases will be exposed to the solvent and thus may become single stranded. It was found that in subsequent stretching that a significant fraction of the DNA remain permanently melted. Their results thus demonstrates that DNA overstretching is accompanied by disruption of the DNA helical structure, including the loss of hydrogen bonding [218].

The tension induced melting transition proposed by Rouzina and Bloomfield was recently examined by Cocco et al. [221] in the light of data obtained from a number of experiments. Using a simple thermodynamical model for tension melted dsDNA it was shown that the over-stretching transition near 65 pN cannot be explained in terms of the conversion of the



**Fig. 25.** Force-extension curve for a ssDNA in different salt concentrations. Solid lines are from [214] and dots are from the experiments by Wenner et al. [217]. (a) Data for entire regimes. (b) the same data for the over-stretching regime.  
Source: By permission from Ref. [214].

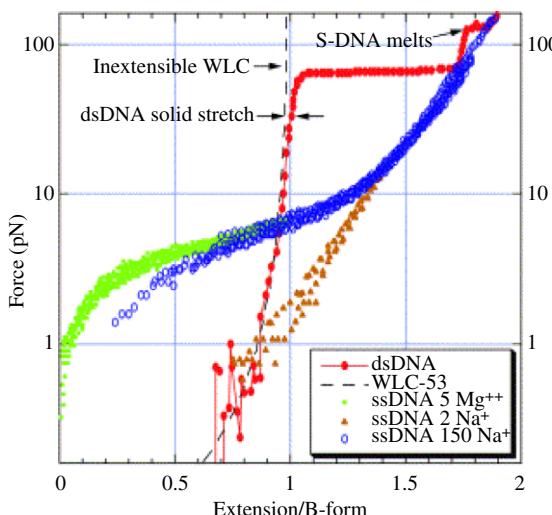
double helix into non-interacting poly-nucleotide strands. It was found that force response of SDNA neither matches with a single ssDNA nor to that of two parallel ssDNAs. This indicates that the mechanical properties of SDNA are distinct from those of B-DNA.

Zhou et al. [222,223] proposed a 3d model based on bending and base stacking in which the double helix nature of the DNA structure was considered via the introduction of a structural parameter, namely the folding angle. They provided a unified framework using Green function methods to understand systematically and qualitatively all aspects of the mechanical properties of DNA [222,223]. Excellent agreement between theoretical predictions obtained using this model and experimental results have been reported by several groups.

The role of salt on this type of transition raised a number of questions (Fig. 25). First of all one may wonder if the electrostatic component of the B-to-S transition is a manifestation of effects already included in the mesoscopic elasticity or whether other effects are also involved. Based on an idea proposed by Manning [224] the data fitting analysis of Wenner et al. [217] supports the latter view. Yet the mesoscopic elasticity model proposed by Podgornik [211] does not rule out the possibility of merely including the contribution to denaturation of electrostatics through the elastic parameters. Secondly, one may ask how well the data conform to the much-invoked Manning condensation theory [225]. Rouzina and Bloomfield have discussed that approach in the context of the thermotropic properties of DNA, and irrespective of the detailed correctness of the Manning theory. The model proposed by Zhou et al. has been extended by Dong et al. [226] and Fu et al. [227] to study the effect of salt. It was shown that the over-stretching force is linear with the natural logarithm of the ionic strength, which is in good agreement with experiments.

## 6.2. Single stranded DNA under tension

The force response curve of dsDNA is quite well understood by now and much attention has recently turned to ssDNA or RNA molecules. ssDNA has twice the contour length per base compared to dsDNA because the helical BB of the double helix



Current Opinion in Structural Biology

**Fig. 26.** Force-extension response of dsDNA and ssDNA. Pulling and relaxation curves are shown. The effect of ion concentration is evident from the plot. See [22,228] for details.

Source: By permission from Ref. [228].

extends about 0.8 nm per base compared to 0.3 nm per base of dsDNA. As a linear chain of nucleotides with thin diameter and large flexibility ssDNA is more contractile than dsDNA at low forces, but it can be stretched to greater length at high forces because of the absence of the helix. The most striking difference between dsDNA and ssDNA is that ssDNA can stick (form hairpins) to itself by base pairing and stacking interactions between bases along the same molecule.

The counterion atmosphere of DNA neutralizes the charges of anionic phosphates and affects the electrostatic interactions of the system. The DNA structure is very sensitive to the composition and concentration of ions. This was clearly seen in recent force-extension curves for ssDNA (Fig. 26) [228]. Using the approach developed for worm like polyelectrolytes, the FJC and WLC models were used to explain the force-extension curves under different ionic strengths. It was found that in a 150 mM NaCl solution the force-extension curve of a ssDNA melted from  $\lambda$ -phage DNA can be fitted with a FJC model of Kuhn length 1.5 nm with an additional stretch modulus [22]. However recent measurements have shown that the theoretical description of the FJC model fails in low (2 mM NaCl) and high (5 mM MgCl<sub>2</sub>) ionic solutions [229]. To explain the high concentration data it was argued that secondary structures (hairpins) may be formed when ssDNA sticks to itself and its complementary bases form base pairs and consequently a larger force is required to pull ssDNA having such hairpins than expected from the FJC model [230].

Furthermore the measurements of Rief et al. shows that the force-extension characteristic of ssDNA is sequence dependent [231]. When a designed poly(dA-dT) or poly(dG-dC) strand is pulled with an atomic force microscope they found that at a particular stretching force (9 pN for poly(dA-dT) and 20 pN for poly(dG-dC)) the end-to-end distance of the molecules suddenly change nearly to zero, a value comparable to its total contour length. This is drastically different from the gradual elongation of a ssDNA chain with a relatively random base composition [232]. Zhou et al. observed that there is a higher probability for a paired base to be neighbors in the designed ssDNA than in a random sequence and hence a strong base pair stacking interaction exists in the designed ssDNA chain [233].

Montaniri and Mezard have developed a theoretical analysis of the elasticity of a polymer chain with hairpins as secondary structures [230]. The formation of such structures is a crucial step in the folding of single stranded nucleic acids polymers. Its importance stems from the rather large values of the binding energy involved in their formation as compared to the much smaller energy scale of the interaction between secondary structures that governs the final three-dimensional shape of the molecules. The model developed by Montaniri and Mezard reproduces the experimental force extension curve measured on ssDNA chains whose nucleotide bases are arranged in relatively random order. The force-induced transition in ssDNA is found to be second order characterized by a gradual decrease in the number of base pairs as the external force is increased.

The question as to why a force induced transition is second order in one case and highly cooperative first order in another case was addressed by Zhou et al. [233] who noticed that there may be strong base pair stacking interactions in the designed ssDNA chain while a stacking energy is absent or negligible in a relatively random sequence. By including the effects of base stacking in the model developed by Montaniri and Mezard they found that when the base stacking interaction is small the ssDNA chain is in a hairpin state with base pairs only weakly stacked in a so-called hairpin-I phase. When the base stacking interactions are large the ssDNA is in a hair-pinned state with almost competing pairing and stacking energies which they called hairpin-II. They showed that the force induced a transition between the hairpin-II state and the coil state is first order while the transition between the hairpin-I state and the coil state is continuous. A continuous transition between hairpin-I

and hairpin-II states may be induced by changing the nucleotide sequence or by changing the solvent conditions such as temperature or ionic strength [234].

The comparison of force-extension curves with theoretical predictions over the entire force range including the entropic regime (low force) and the energetic regime (high force) is still a challenge. The most popular WLC model, could not even reproduce the entire range with three parameters. It was found that fit parameters depend on the range of forces used for the fit [235,236]. Hugel et al. [237] studied three different chains namely ssDNA, poly vinylamine and peptide at very high force (2 nN). At such a high force conformational entropy does not play a significant role, therefore, zero temperature *ab-initio* calculation can be compared with the experimental results. They showed that with a single parameter the contour length  $L$ , one can obtain different elastic constants which describes force-extension curve quite well at high force.

### 6.3. Effects of temperature

Danilowicz et al. [238] studied experimentally the elastic properties of ssDNA and showed that temperature has a significant effect on the force-extension behavior. It was found that below a certain force the extension increases with temperature, while above that same force the extension decreases with temperature. Known theoretical models like the FJC, WLC or PS model could not provide the explanation to this abrupt decrease. Kumar and Mishra [176] developed a simple model based on SAWs and showed that this decrease is an entropic effect. Their results indicated that in the constant force ensemble the observed decrease in extension with temperature may be observed in other SMFS experiments such as protein unfolding, RNA unfolding and DNA unzipping.

### 6.4. Evidence for structural transitions in ssDNA

Early stretching experiments related to ssDNA showed that the entropic response and the force-extension curves could be reproduced by the FJC model [17,205,239,240]. Recently it was found that semi-microscopic changes in the monomer strongly influence the elastic properties of ssDNA. Attempts have been made to monitor the force-extension curve [241,242] of ssDNA consisting of only one type of nucleotide. It was found that the elastic properties of ssDNA made of adenine (poly(A)) are significantly different from ssDNA made of thymine (poly(T)) or uracil (poly(U)). It is known that base stacking is strongest among adenine (poly(A)) and weakest among thymine (poly(T)) and uracil (poly(U)) [243].

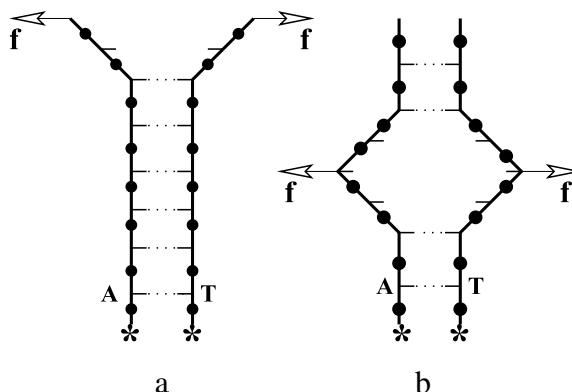
The force-extension curve for poly(T) and poly(U) shows the entropic elasticity. Whereas poly(A) exhibits the existence of a plateau in the force-extension curve [241]. Ke et al. [242] found the existence of multi-step plateaus in the case of poly(A). The first plateau occurs at a force  $23 \pm 1$  pN and over-stretches the nucleotide by  $\sim 74\%$ , which has been predicted by the model proposed by Buhot and Halperin [244]. This prediction has also been observed in the case of homopolymeric RNA [241]. An attempt has been made to explain a plateau observed in the case of ssRNA (poly(A)) on the basis of seven parameters [241] in Zimm-Bragg model [245]. A qualitative understanding of the first plateau has been achieved in terms of the unwinding of helical structure of poly(A) arising due to base stacking. This transition appears to be weakly cooperative, but needs further attention from the statistical mechanics point of view [241,244]. Moreover, the second plateau which occurs at a force  $113 \pm 1$  pN and over-stretches ssDNA by an additional 16% [242] was not predicted by the model proposed by Buhot and Halperin [244]. Though it was argued that it is the mechanical signature of the base stacking interactions it appears that there is still no clear understanding of this transition [242].

Mishra et al. [246] developed a model in which the flipping of bases has been introduced for the first time. This allowed them to study the effect of stacking interactions in the case of ssDNA consisting of only one type of nucleotide. It was found that the stacking interaction is not sufficient to describe the force-extension curve of ssDNA. They showed that when bases are stacked a helix is formed and the effective end-to-end distance is reduced for poly(A) but remains the same for poly(T). The first plateau is the result of this effect. When ssDNA acquires a stretched state and force is further increased there is an additional increase (about 18% of length) due to a structural transition which gives rise to the second plateau in the force-extension curve in accordance with experiments.

Physically this may be understood in terms of the structurally different conformations arising due to the rotation of furanose rings in the nucleic acids [247]. The two basic forms C3'-endo (5.9 Å distance between neighboring phosphates) and C2'-endo (7.0 Å distance between neighboring phosphates) of furanose rings characterizes two distinctly different families of nucleic acid structures [247]. The reorientation of bases at high force may lead to a conformational transition in the deoxyribonucleic acid rings from C3'-endo pucker to C2'-endo pucker in a step like jump [248]. If so then it gives an additional extension of about 18% in the chain [242]. It may be noted that in the case of ssDNA the potential energy associated with a ring is much less than for RNA. This is the reason why RNA molecules are stiffer than ssDNA [248]. Inclusion of this extension in the partition functions qualitatively reproduces the experimental force-extension curve.

### 6.5. Unzipping of dsDNA

The realization that there can be a force-induced unzipping transition of DNA with a force applied solely at one end raised a lot of interest among experimentalists and theorists [43,46,51,52,249–251]. In both the thermal and the forced cases the



**Fig. 27.** Schematics of DNA unzipping by force  $f$  applied at (a) one end of the two strands (Y-model); (b) at the interior of the strands, which gives rise to the “bubble or eye-type” conformation.

phase transition takes a double stranded form to two single strands. Since the elucidation of the structure of DNA it has become clear that DNA replication and DNA transcription into messenger RNA necessitate the unwinding or unzipping of the two paired strands. Indeed it is now clear that DNA replication is a correlated process involving many proteins and other molecules working at different points in space and time. Understanding the nature and origin of these correlations is in fact a major motivation for statistical mechanics models of DNA unzipping. It has been demonstrated that the force induced unzipping of DNA is a genuine phase transition, different from its thermal denaturation. It was proposed that the initiation of replication at the origins along the DNA, e.g. DNA for *Escherichia Coli*, or by the “origin of recognition complex” in eukaryotes is like the unzipping near the phase transition point and the resulting correlations during unzipping leads to the cooperativity required for replication. Furthermore, a sound investigation of DNA replication *in vitro* requires an understanding of the coupling between the opening of the strands and subsequent events during the process.

The mechanical separation of DNA using force was first observed by Bockelmann and his coworkers [249]. They used  $\lambda$ -phage DNA to study the unzipping phenomenon. One strand of this  $\lambda$ -phage DNA was attached to a glass slide and the other strand to a micro-needle. The other end of the  $\lambda$ -phage DNA is capped with a hairpin molecule to avoid complete separation. The deflection of the tip was measured on a video image as a function of displacement. The mechanical forces involved in the strand separation was found to be of an irregular nature, which is due to the sequence of AT/GC pairs along the chain and the elastic nature of the mechanical configuration. They performed experiments in the range of 20–40 nm/s which corresponds to a rate of opening of around 20–40 base pairs per second.

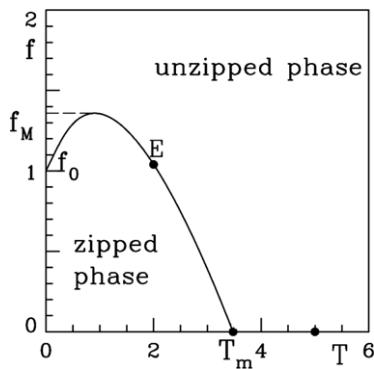
A number of theoretical efforts have been made to understand various aspects of dsDNA unzipping and use the technique to gain further insights into DNA. As an example we mention that sequencing genomes by unzipping the strands of a single dsDNA molecule and following the force fluctuations was considered by Viovy et al. [252]. The difficulties associated with single strand separation due to the elasticity were emphasized. Later on predictions of an unzipping transition based on an interacting Gaussian chain were made by Bhattacharjee [46] and almost immediately confirmed by a dynamic approach [253]. It was shown that while a homogeneous dsDNA molecule gains considerable entropy by opening in response to the external force, and therefore the unzipping is entropy driven, a heterogeneous dsDNA molecule is believed to unzip primarily for energetic reasons [51]. Lubensky and Nelson [52] have studied the force induced unzipping of a randomly disordered dsDNA molecule using a Hamiltonian which was coarse grained over many bases with an unknown number thereof. By mapping it onto a non-Hermitian quantum mechanics problem it was shown that there is a critical force  $f_c$ , required to unzip the dsDNA and the phenomenon is different than the thermal melting of dsDNA. For the homogeneous sequence the number of open base pairs scales as  $|f - f_c|^{-1}$  [46]. Nelson and Lubensky and Nelson studied both homogeneous and heterogeneous sequences. For a heterogeneous sequence they found that the number of open base pairs scales as  $|f - f_c|^{-2}$  [51] using a model similar to the Poland–Scheraga model discussed earlier. In both cases the critical force is  $f_c$  is given by

$$f_c = 3f_0k_B T/b^2. \quad (81)$$

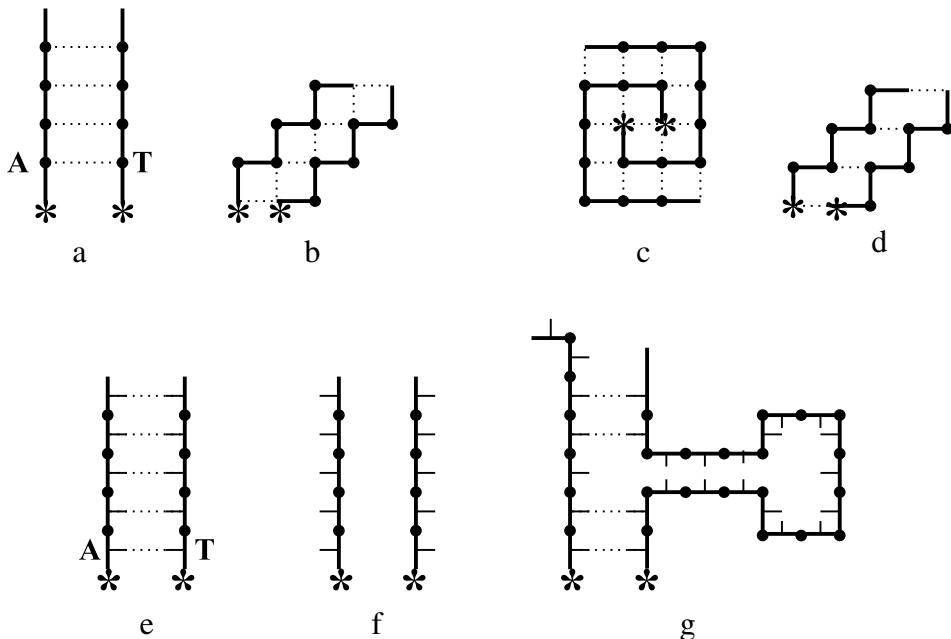
Here  $b^2 = 2a^2$  with a Kuhn length of ssDNA and  $f_0$  is the average energy per base pair in the absence of a pulling force. Recently Lam et al. studied the excluded volume effect and found that the force vs. temperature diagram depends on only two parameters namely the persistence length and the denaturation temperature [254,255]. The scaling form is parameter independent and depends only on the spatial dimension.

Molecular information on the unzipping transition is now available in both the constant distance and constant force ensembles from extensive exact solutions of lattice models. Using the directed walk model of dsDNA, Marenduzzo et al. studied two different models [58].

- (i) the “Y model” in which the two strands are zipped together up to a bifurcation point, i.e., allowed configurations are those which have the first  $N - m$  bases bound and the remaining  $m$  bases separated as in a Y-shape (Fig. 27(a)).
- (ii) The “b-model or eye model” in which configurations with “bubbles or eyes” are also allowed (Fig. 27(b)).



**Fig. 28.** Force-temperature diagram for the model which exhibits reentrance at low temperature [58].  
Source: By permission from Ref. [58].

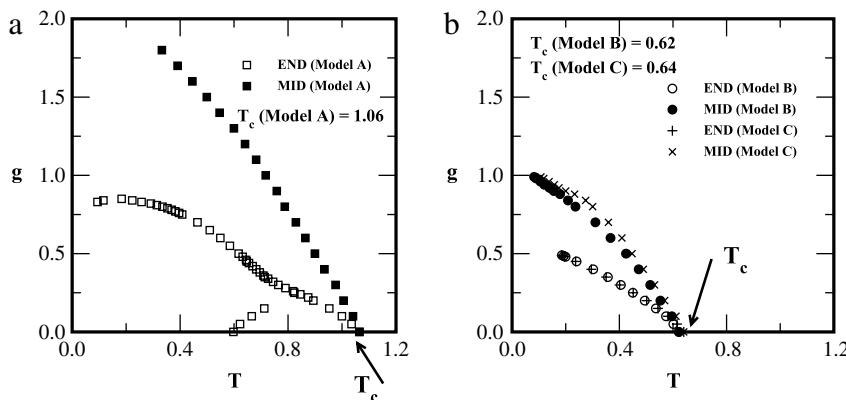


**Fig. 29.** Figures (a–c) represent the possible conformations in models where any monomer from one chain can interact with any monomer from the other (Model A). Figures (a) and (d) are possible conformations of a model where the  $i$ th monomer of one strand can interact with the  $i$ th monomer of the other strand (Model B). Figure (c) corresponds to the ground state of model A, while Figure (a) corresponds to the ground state of model B. Figures (e, f, g) represent model C where bases are on the links of the strands with short stubs representing the orientation of the possible interactions. (e) is the completely zipped state, (f) is a non-pairing configuration, and (g) is a partially bound configuration with a hairpin loop. In all cases the dotted lines show the attractive interactions.

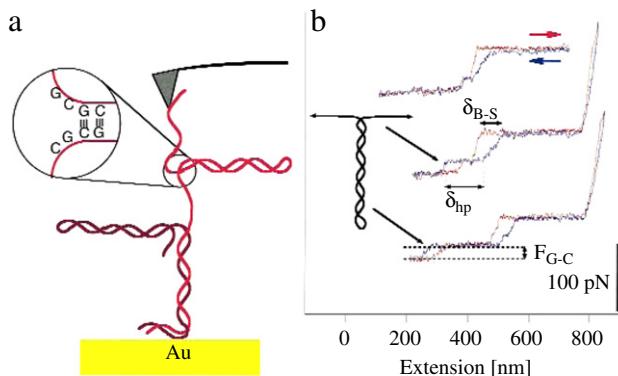
The schematics of these models are shown in Fig. 27. Following the generating function technique they obtained the exact phase diagram shown in Fig. 28. A major result from the exact solution was the occurrence of re-entrance in the low temperature region. Monte Carlo simulations of a three-dimensional model with self- and mutual-avoiding walks representing strands of dsDNA also showed this re-entrance [57].

Kumar and coworkers [59–62] considered two linear polymer chains made up of say Adenine (A) and its complementary Thymine (T) which are mutually-attracting-self-avoiding walks (MASAWs) on a square lattice as shown in Fig. 29(a–d). There is another more realistic model of dsDNA where the directional nature of the interactions has been taken into account [60–62]. It was called model C and some details are shown in Fig. 29(e–g). They considered two cases: (I) a force is applied at the end (Fig. 27(a)) of the chain (analogous to replication) or (II) a force is applied at the middle (Fig. 27(b)) of the chain (analogous to transcription). The model may be generalized to any dimension [59–62], but for computational ease, they considered it on the square lattice only. The phase diagram for all these models is shown in Fig. 30.

In another experiment, Rief et al., studied the effect of longitudinal stress on dsDNA [231]. They observed that with a loading rate of 200–6500 pN per second dsDNA changes into ssDNA around a force of 65 pN. In [231] the progressive opening was monitored for different opening velocities. It was found that the average value and amplitude of the force signal are almost independent of the opening velocity in the interval of 20 nm/s to 800 nm/s. Rief et al. have also observed



**Fig. 30.** Variation of critical force ( $g_c(T)$ ) (END and MID case) as a function of temperature ( $T$ ) for (a) model A, (b) model B and C. In figure (b) the phase boundary separates the zipped state from the unzipped state for both the END and MID case. Whereas Figure (a) (Model A) shows (i) the spiral to zipped transition, (ii) zipped to unzipped transition and (iii) spiral to unzipped transition for the END case. All three lines meet at a triple point. For the MID case one gets only a zipped to unzipped transition. Figure has been taken from Ref. [59].

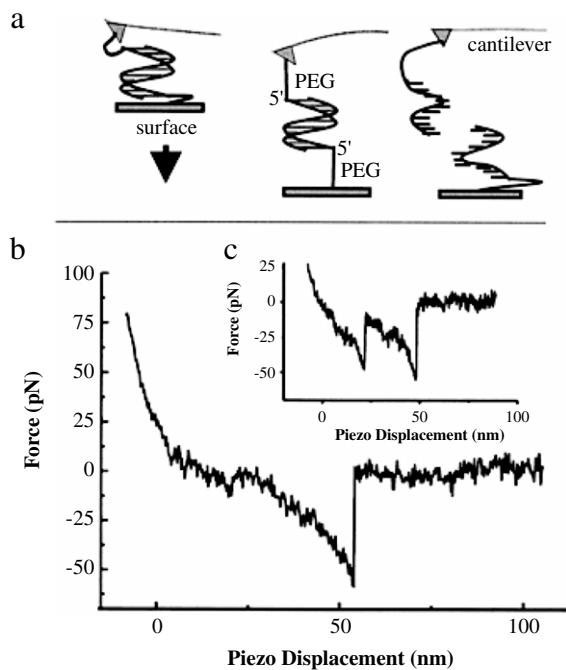


**Fig. 31.** (a) The self-complementarity of poly(dG-dC) and poly(dA-dT) allows for hairpin formation within the two DNA strands. After force-induced melting of the double helix base-pairing forces can be directly measured in the strand attached to the tip. (b) Three repeated extension (red) and relaxation (blue) curves of one poly(dG-dC) molecule. For interpretation of the references to colour in this figure legend, the reader is referred to the web version of Rief et al., Nature Structure Biology, Vol. 6, 346 (1999).  
Source: By permission from Ref. [231].

the formation of hairpin like structures in DNA. When the tension on the molecule is relaxed the single strand of DNA form hairpins (Fig. 31). This hairpin was found to unzip at a force close to 20 pN for a chain containing only GC pairs and close to 9 pN for a chain containing AT pairs only. The force required to unzip dsDNA by over-stretching was found to depend on the velocity of pulling, the ionic strength, the temperature and the sequence of the DNA molecule (Figs. 31–33).

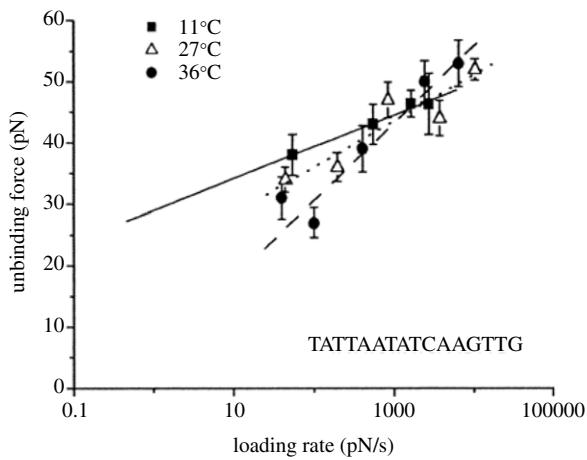
In some cases the term “unzipping” is not appropriate because the inter-chain interactions were carried out in a different way (Fig. 32). In one experiment the complementary segments were covalently attached to opposite surfaces. For example the force versus displacement was measured between  $(ACTG)_5$  and  $(CAGT)_5$  fictionalized surfaces [199]. The highest forces for 20, 16 and 12 bps were 1520, 1110 and 830 pN, respectively. It appears that the highest force was 76 pN/bp. Another group of researchers studied the mechanical separation of short DNA sequences [256,257]. They considered the DNA oligomer  $a = 5'-G-G-C-T-C-C-T-C-T-A-C-C-A-C-T-G-A-C-A-T-C-G-C-A-C-G-G-3'$  and tested it against its complement  $b = 3'-C-C-G-A-G-G-G-A-A-G-A-T-G-G-T-G-A-C-T-G-T-A-G-C-G-T-T-G-C-C-5'$  and the truncated complements  $c = 3'-A-T-G-G-T-G-A-C-T-G-T-A-G-C-G-T-T-G-C-C-5'$  and  $d = 3'-T-A-G-C-G-T-T-G-C-C-5'$ . In these cases the lengths of the sequences  $c$  and  $d$  are 20 and 10 bps while the complimentary oligomers  $a$  and  $b$  are 30 bps long. They applied force not on the opposite strands (5' and 3') but rather on the opposite ends of two strands at 5' and 5'. This experiment helps to elucidate the mechanical and elastic properties of dsDNA. They found that the cooperative unbinding of base pairs in the duplex leads to a scaling of the unbinding forces with the logarithmic of the loading rate. Depending on the loading rate, which they varied from 16–4000 pN per second, and sequence length the unbinding force of a single duplex is found to lie between 20–50 pN (Fig. 33). They also showed that the distance of the energy barrier to the minimum energy along the separation path, and the logarithmic thermal dissociation rate are proportional to the number of base pairs of the DNA duplex.

Porschke proposed a mechanism for DNA slippage [258]. A formation of bulge of a few bases at one end may diffuse along the molecule and anneal at the other end. This means that two strands may slide on each other with only a small energetic



**Fig. 32.** Schematic diagrams for the unzipping of dsDNA when a force was applied along 5'-ends of the DNA. Force–displacement curves for different conditions. B and C show how a kink in the force data can actually identify the presence of a mechanical intermediate along a single reaction coordinate. For a detailed discussion see [256].

Source: By permission from Ref. [256].



**Fig. 33.** Unbinding force with loading rate. Different slopes at different temperatures is evident from the plot [262].  
Source: By permission from Ref. [257].

barrier compared to the large barrier of complete unbinding. Neher and Gerland studied the base pairing dynamics of dsDNA with repetitive sequences where the local strand slippage can create, annihilate and move bulge loops [259]. Using analytical methods and Monte Carlo simulations they found that the response of periodic dsDNA to a shear force is distinctly different to that of non-periodic sequences. Above a critical force  $f_c$ , but below a dynamical force  $f^*$ , bulge loop diffusion allows periodic dsDNA to open by sliding. Predictions of the viscoelastic behavior for periodic DNA with time and force may be used with micro structure devices similar to programmable DNA based force sensors [260]. It was found that repetitive sequences dissociate at lower forces and elongate above a certain threshold force. This yield force was found to be rate dependent [261]. In the case of rapid stretching of the DNA duplex the applied force relaxes by a stepwise elongation of the duplex.

In the experiments by Bockelmann and coworkers [249,251], dsDNA was unzipped with a constant velocity and a fluctuating force applied at one end of a strand. In an interesting experiment by Danilowicz et al. dsDNA was unzipped using a constant force applied at one end of a strand [263,264]. They observed the unzipping of the first 1500 base pairs

of  $\lambda$ -phage DNA. They performed the experiments at different temperatures and proposed an experimental phase-diagram of DNA unzipping. One of the most interesting features of these experiments are the observation of certain “pauses” and “jumps” in the opening of dsDNA. They did the experiments on 50 beads in which 10 beads were found to unzip slowly over the course of the experiment pausing at several points. An individual bead paused at a fixed extension until through thermal fluctuations the unzipping proceeded. The number of unzipped base pairs was calculated using the following arguments. The distance between beads, attached to fully zipped strands of  $\lambda$ -phage DNA, and round capillary is around  $77.4 \mu\text{m}$  at  $15 \text{ pN}$  when the DNA is fully unzipped. Thus DNA strands stretched to a length of  $60.9 \mu\text{m}$ . Since  $\lambda$ -phage DNA have 48502 base pairs, there are on average  $48502/60.9 \approx 800 \text{ bp per } \mu\text{m}$ . Since there are two strands of DNA the monomer spacing along a ssDNA is nearly  $6 \text{ \AA}$  which was also found by standard experiments [34].

Pauses and jumps observed in these experiments are due to many reasons. Near the melting transition AT rich regions unbind and form a transient “bubble”. The probable reason for these bubbles is the differences in bond energies of AT and GC pairs, the twist of the double helix etc. However, in these experiments the location of pauses is found to be highly conserved from one strand to the other. Thus one can say that transient bubbles and twists play only minor roles in determining the jumps and pauses. Weeks et al. [157] used the model of Lubensky and Nelson [52] and calculated the pause point spectrum in the constant force ensemble of  $\lambda$ -phage DNA. For a brief review of these experiments see the articles by Bockelmann [265] and Weeks [157].

There is another class of DNA unzipping experiments in which the effects of solution conditions, like pH, on unzipping phenomenon have been studied. A simple model proposed by Rouzina and Bloomfield predicts that the DNA over-stretching force is a decreasing function of temperature and conversely, that the melting temperature is a decreasing function of the applied force [215,216,266]. Their model also quantitatively explains the observation that cross linking the dsDNA makes little difference to its stretching behavior in high salt concentrations, but significantly raises the over-stretching force in lower salt concentrations compared to non-cross linked DNA.

### 6.6. Unzipping DNA from a globule state

In the systems studied so far there is only one zero force thermal phase transition. In such a case, with two intensive variables, temperature  $T$  and force  $f$ , there is an unzipping transition line  $f = f_c(T)$  in the  $f-T$  plane, demarcating the zipped state from the unzipped phase. If there is more than one transition then the phase diagram will be influenced by the intermediate phase. Kumar et al. [59] considered two models which differ in the nature of the mutual interaction between the two strands. These models are general enough to be defined in any dimension  $d$ . In one of the models (model A) any monomer from one chain can interact with any monomer from the other chain (Fig. 29). In this model there are two phase transitions. The first phase transition is from a compact globule (spiral) state to the zipped state. With a further increase in temperature a second transition occurs from the zipped state to the unzipped state (Fig. 30). They showed that an entropy stabilized force gives rise to a force induced triple point in the force–temperature phase-diagram. This opens up the possibility of much richer phase diagrams and multicritical behavior in polymeric systems having potential applications to our understanding of DNA unzipping.

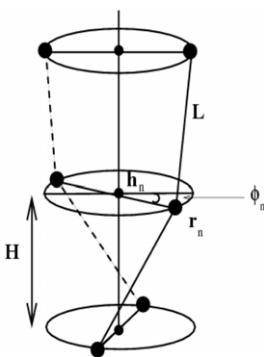
Lam and Levy [267] studied the phase transition from a compact globule state to an unzipped state at constant force. They found that the number of unzipped monomers scales as  $|f - f_c|^{-3}$  for both homogeneous and heterogeneous sequences. This is different from the case of the unzipping of DNA from an extended coil state. In the constant distance case the number of unzipped base pairs at the unraveling transition scales as  $N^{4/3}$  in accordance with theoretical predictions about the unraveling transition of polymers stretched by an external force. Further experimental, numerical and theoretical work is required to fully understand the complete phase diagram.

An advantage of force-induced melting over thermal melting is that it is isothermal, thereby avoiding poorly characterized thermal contributions to the transition enthalpy and entropy. The resulting data can be used for independent verification of current ideas about DNA melting thermodynamics, including the issue of the heat capacity increment in DNA melting.

### 6.7. Rotational drag on DNA

Torsional friction drag on DNA is a problem of long standing interest since the discovery of the structure of dsDNA. Levinthal and Crane in 1956 proposed a simple model by considering dsDNA as a cylinder rotating about its axis analogous to a speedometer cable and calculated the friction torque required to open the strands [268]. Friction torque has also been invoked in kinetics of *in vitro* thermal denaturation and renaturation. Nelson [269] developed a model to evaluate the frictional torque and proposed that the torque may actually be orders of magnitude higher than estimated by Levinthal and Crane. Nelson argued that DNA does not act like a speedometer cable as assumed in the model of Levinthal and Crane. Instead the presence of natural bends in the DNA should induce a sideways motion of the molecule when rotating, an effect expected to be associated with much higher friction.

Nelson [269] found the theoretical expression of torque by considering DNA as a semi-flexible chain. The effective rotating chain length  $L_{\text{eff}}$  of dsDNA is usually less than the actual length because of presence of randomly distributed nick. For  $L_{\text{eff}} = 12 \mu\text{m}$  Nelson estimates for the torque is about  $15\,000 k_B T$  for an angular velocity of  $12\,000 \text{ rad/s}$ . This value



**Fig. 34.** The helicoidal DNA model: Base pairs are modeled through their radius  $r_n$  and angle  $\phi_n$ . The axial distance between successive base pair planes varies while the BB length along the strand is fixed to be  $L$  [273].

is much larger than the value reported by Levinthal and Crane [268] and the value obtained by Thomen et al. [270] which suggests that dsDNA is a rigid rod. It has been reported both experimentally [271] and through theoretical modeling [23,272] that a tethered DNA in a uniform flow is not free draining, up to a ratio  $r$  between the extension and length of the molecule of 0.8 which is close to the value 0.4 estimated by Thomen et al. [270]. Hence hydrodynamic coupling is such that non-free draining conditions occur for translation and there should be no extreme effect of the flow on rotational friction. Moreover, Thomen et al. also studied the rotational drag on dsDNA by measuring the force during mechanical opening and closing of the double helix at different rates [270]. They used a microscope set up using an optical trapping interferometer as a force-measuring device. The molecule is cranked at one end by the effect of unzipping and is free to rotate at the other end. They found that at  $\sim 14$  pN of applied force the molecule starts opening. The rotational drag applied on one end of the molecule, while keeping the other end fixed, gives rise to a torque which was found to be of the order of  $1k_B T$  per 10 000 base pairs of DNA when cranked at 2000 turns per second. It was also shown that the effect of rotational drag increases with the length of the molecule and is approximately proportional to the angular velocity.

Cocco and Monasson developed a unified theory to describe the denaturation transition of DNA, driven by temperature or induced by torque [50,273,274]. The proposed model is an extension of the PB model which couples the hydrogen bond opening and the untwisting of the helicoidal molecular structure. The model is shown schematically in Fig. 34.

Each base pair ( $n = 1, \dots, N$ ) is described by its radius  $r_n$ , the angle  $\phi_n$  in the plane perpendicular to the helical axis and the height  $h_n$  along the ladder. Since the sugar phosphate BB is comparatively rigid the distance between two consecutive bases,  $L$ , along the same strand is assumed to be constant. The BB distance introduces a geometrical constraint between two successive base pairs

$$L = \sqrt{(h_n - h_{n+1})^2 + r_n^2 + r_{n+1}^2 - 2r_n r_{n+1} \cos(\phi_n - \phi_{n+1})}. \quad (82)$$

This couples the degrees of freedom  $r_n, r_{n+1}, \phi_n, \phi_{n+1}, h_n$  and  $h_{n+1}$ . From this equation one can express  $l_n = h_n - h_{n+1}$  as a function of  $r_n, r_{n+1}$  and of the twist angle  $\theta_n = \phi_n - \phi_{n+1}$ ,

$$l_n(r_n, r_{n+1}, \theta_n) = \sqrt{L^2 - r_n^2 - r_{n+1}^2 + 2r_n r_{n+1} \cos \theta_n}. \quad (83)$$

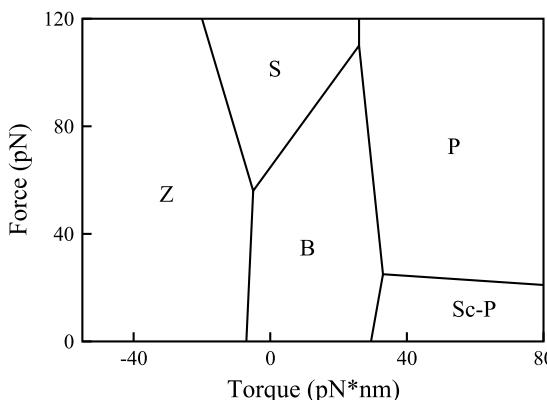
We can then choose the radii  $r_n$  and angles  $\phi_n$  as independent degrees of freedom from which the heights,  $h_n$ , may be unambiguously obtained. The Hamiltonian  $H_c$  associated with a configuration of independent degrees of freedom  $\{r_n, \phi_n\}$  is the sum of three different contributions,

$$H_c(r_n, \phi_n) = \sum_{n=1}^N V(r_n) + \sum_{n=2}^N [W(r_n, r_{n+1}) + B(r_n, r_{n+1}, \phi_n, \phi_{n+1})]. \quad (84)$$

Here the first two terms are the same as in the planar PB model while the third term is an angle dependent term, which takes into consideration the helicity of DNA

$$\begin{aligned} B(r_n, r_{n+1}, \theta_n) &= k(l_n - H)^2 \\ &= k \left[ \sqrt{L^2 - r_n^2 - r_{n+1}^2 + 2r_n r_{n+1} \cos \theta_n} - H \right]^2. \end{aligned} \quad (85)$$

The helicoidal structure arises when  $H < L$ . The model predicts that the denaturation transition is of first order and is in good agreement with experimental results.



**Fig. 35.** Phase-diagram for DNA under torque. The lines correspond to phase coexistence. Z, S, B, P and Sc-P are the left hand twisted phase, over-stretched phase, right handed phase, Paulling phase and supercoiled phase, respectively. The detail may be found in [279].  
Source: By permission from Ref. [279].

### 6.8. DNA twisting

So far most single molecule experiments have focused on stretching properties, but recently the response of a DNA molecule to an external torsional stress has been studied. From a biological point of view torsional stress is quite important in the context of living cells and may strongly influence DNA functionality.

As mentioned above, experiments by Cluzel et al. [204] and Smith et al. [22] showed that there is a sharp transition in dsDNA around 65 pN of tension where the length of dsDNA is about 1.7 times the length of the double helix. The transition is possibly due to the fact that the covalently bonded sugar phosphate BBs along each strand are helically coiled inside B-DNA. Other DNA stretching experiments showed that dsDNA can be stretched almost to double the length of B-DNA. This state has been termed as S-DNA [275,276]. The transition has been referred as a B-S transition. The 65 pN B-S transition occurs on dsDNA with single strand attachments at their ends or on a molecule where single strands breaks (nicks) somewhere inside them. Such molecules support torsional stress while coiled inside B-DNA. In order to extend a molecule up to 1.7 times of its B-form length it is more likely that the two strands must untwist. Two groups [204,206,277] have developed experimental techniques and studied the force-extension curves subject to the constraint that the double linking number is fixed. Early experiments [277] showed that at small forces, 5 pN, effects on the entropic elasticity of DNA could be observed and studied as a function of twisting. Studies of this low force regime verified theoretical predictions of the coexistence of super coiled and extended domains and denaturation of under-twisted DNA at forces of a few pN.

In other experiments Strick et al. [278] studied the behavior of twisted DNA at higher forces and showed that dsDNA could be denatured by untwisting and very heavily over-twisted. Legner et al. [206] made a global study of the force-extension behaviour over a wide range of twisting and for forces in the range 1–100 pN. These experiments suggested that there are four phases in addition to B-DNA, which can be accessed by varying the force and amount of twisting.

For a straight dsDNA molecule with both strand affixed to surfaces, the linking number  $Lk$  of the two strands is a topological invariant. This is defined as the number of times two strands of the DNA double helix are intertwined.  $Lk$  is the sum of two geometrical characteristics of the double strand the twist  $Tw$  and the writhe  $Wr$ :

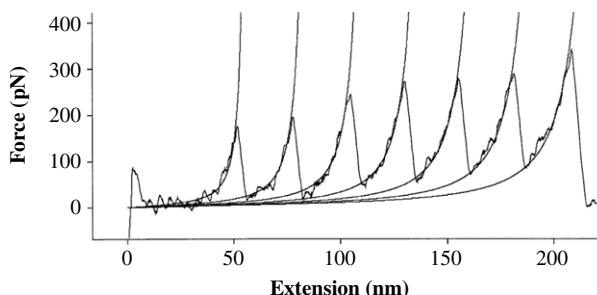
$$Lk = Tw + Wr. \quad (86)$$

Here  $Tw$  reflects the helical winding of the two strands around each other and  $Wr$  is a measure of the coiling of the axis of the DNA about itself, like a twisted cord forming interwound structures to relieve its torque. For a long unconstrained DNA,  $Lk = Lk_0$ , the number of turns of the double helix. Since there is one turn after every 10.5 bp for the unstressed DNA,  $Lk = N_{bp}/10.5$ , where  $N_{bp}$  is the total number of base pairs. It is convenient to describe the relative change in linking number or the degree of supercoiling

$$\sigma = (Lk - Lk_0)/Lk_0. \quad (87)$$

For either a ladder like DNA or for separated single strands  $\sigma = -1$ . By comparison DNA *in vivo* is usually constrained to have  $\sigma = -0.05$ .

Sarakar et al. [279] studied a simple discretized model, which is the extension of a continuum model proposed earlier by Lenger et al. [206]. The simple model predicts five structural states separated by first order phase transition lines and involves three triple point like regions where three phases can coexist (Fig. 35). The structural transitions between these five states are induced by force and torque and have been supported by good global fits to the experimental force-extension curves [205,280]. For forces in the 50 pN range the double helix is stable roughly over the torque range  $-2k_B T$  to  $10k_B T$ . For a force  $f < 10$  pN, a transition to Z-DNA occurs for unwinding torque  $\approx -2k_B T$  and a transition to sc-P occurs for overwinding with  $\approx 7k_B T$  torque. These are in good agreement with experiment [277]. There is Sc-P DNA to P-DNA transition which takes place at force  $> 20$  pN where torque remains almost constant. At a force  $\approx 50$  pN, there is a transition to S-DNA for unwinding torque  $\approx -3k_B T$ . Since S-DNA remains stable at zero torque for high forces, micromanipulation is possible in this regime.



**Fig. 36.** Shown is a force–extension curve obtained by the stretching of a Ig8 titin fragment. Each peak corresponds to the unfolding of a single domain. Smooth curves are fits to the worm-like model.

Source: By permission from Ref. [2].

## 7. Protein unfolding

As mentioned above, in deference to polymer, proteins consists of secondary structures. This difference gives rise to distinct mechanical properties. Due to the unfolding of individual secondary structures, the force–extension profile of a protein, for example, displays local minima, while the corresponding behavior of a homopolymer may be described by the simple WLC model. Proteins are finite-size objects to which statistical approaches developed for infinite systems can hardly be applied. Therefore, the main results on their mechanical properties have been obtained by numerical simulation. This chapter is mostly based on the results following from simulations and experiments. The kinetics theory is used for their interpretation.

In order to gain a better understanding of the energy landscape and mechanical stability of proteins attempts have now been made to apply a force in the pN range, most commonly using AFM. After the pioneering AFM experiment of Gaub et al. [281] a lot of experimental as well as theoretical work has been done on various proteins (see Table 1 and references therein). Proteins are pulled either with a constant force or by a force ramped with a constant loading rate. The breaking of hydrogen bonds requires a constant force of  $\sim 10$  pN or a pulling rate  $v \sim 100$  nm/s. The force–extension curves from constant rate pulling has a saw-tooth shape due to domain by domain unfolding (see Fig. 36). Grubmuller et al. [103] and Schulten et al. [148] were the first to reproduce this remarkable result by SMD simulations.

The single molecule force techniques have the following advantages that make them more appropriate for comparison to theory or simulations than conventional folding studies: (a) Unlike ensemble measurements it is possible to observe differences in the nature of the unfolding event of single molecules; (b) the reaction co-ordinate is well defined which makes comparisons with theory and numerical simulations easier, and (c) the techniques have the advantage of probing the free energy landscape in great detail.

Most of the proteins studied so far display varying degrees of mechanical resistance, i.e., the unfolding force,  $f_u$ , above which unfolding occurs are very different.  $f_u$  depends on the pulling speed [134].  $\beta$ -rich proteins are believed to be mechanically more resistant compared to  $\alpha$ -rich ones. The 27th Ig module of cardiac titin I-band (I27) which contains only  $\beta$ -sheets has  $f_u \approx 200$  pN at  $v = 600$  nm/s [2] while  $f_u \approx 30$  pN for  $\alpha$ -spectrin [282]. For helix-rich proteins the unfolding force was shown to decay linearly with the helix content [283].

What one can learn from mechanical unfolding experiments? First, from the force dependence of the unfolding time one can extract the distance between the native state and the transition state, this being one of the most important parameters of the free energy landscape. Second, by fitting the force–extension curve to the WLC model the persistence and contour lengths can be estimated. Third, the unfolding force which characterizes the mechanical resistance of a protein can be defined as the maximum of the force–extension profile. Finally, Rief and coworkers [195] have shown that one can use so-called mechanical triangulation to determine protein structures. This technique may prove to be as useful as standard NMR or X-ray crystallography.

Recent reviews [6,107] discuss issues related to elastic, structural conformational and functional properties of proteins using SMFS techniques. We will focus on the relationship between mechanical stability and the parameters characterizing the native topology [283] such as the secondary structure content and the contact order [284]. The dependence of the FEL shape on these parameters will be presented. We will mention recent interesting results obtained by using a microscopic theory [135,285], which is more accurate than the Bell approximation. The effect of pulling directions on unfolding properties of proteins will be also reviewed.

### 7.1. Equilibrium force–temperature phase diagram of proteins

Klimov and Thirumalai were the first to study the  $f$ – $T$ -phase diagram of model proteins [81]. For off-lattice models we define the fraction of native contacts or the overlap function [286]

$$\chi = \frac{1}{Q_{total}} \sum_{i < j+1}^N \theta(1.2r_{0ij} - r_{ij})\Delta_{ij} \quad (88)$$

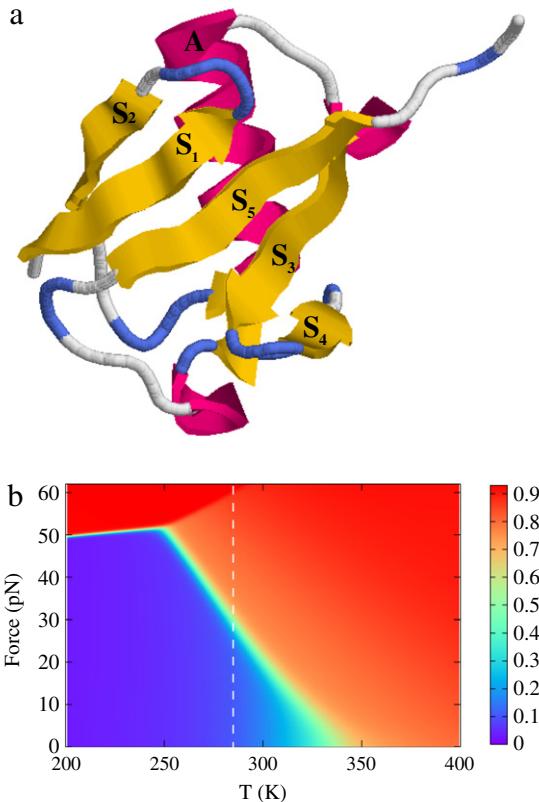
**Table 1**Survey of experimental and Go-simulation results for  $f_u$ .

Protein	PDB ID	N	SCOP Class	alpha (%)	beta (%)	CO	$f_u$ (pN) Sim.	$f_u$ (pN) Exp.	v (nm/s)	Ref.
Protein L	1HZ6	62	$\alpha/\beta$	23.8	44.4	0.161	254.3	136	400	[305]
Ubiquitin (Ub) <sub>9</sub>	1UBQ	76	$\alpha/\beta$	21.1	43.4	0.150	163.9	203	400	[194]
Ubiquitin (Ub) <sub>8</sub>	1UBQ	76	$\alpha/\beta$	21.1	43.4	0.150	163.9	230	1000	[302]
Ubiquitin(48-C)	1UBQ	28	$\alpha/\beta$				84	$85 \pm 20$	300	[194,302]
Barnase	1BNI	108	$\alpha/\beta$	23.2	21.3	0.109	44.2	70	300	[308]
Human DHFR	1HFR	186	$\alpha/\beta$	26.3	24.7	0.131	50.3	$72 \pm 15$	400	[369]
GFP	1B9C	224	$\alpha/\beta$	9.8	54.9	0.127	203.7	104	1000	[291]
GFP(3–212)	1EMB	210	$\alpha/\beta$				152	$117 \pm 19$	600	[196]
GFP(132–212)	1EMB	81	$\alpha/\beta$				193	$127 \pm 23$	600	[196]
GFP(3–132)	1EMB	130	$\alpha/\beta$				323	$366 \pm 45$	600	[196]
GFP(117–182)	1EMB	129	$\alpha/\beta$				282	$548 \pm 57$	600	[196]
GFP(182–212)	1EMB	129	$\alpha/\beta$				135	$356 \pm 61$	600	[196]
Bacterio (Memb)	1AT9	230	$\alpha/\beta$	72.6	6.1	0.065	87	350		[376,377]
T4-Lysozyme	1B6I	162	$\alpha/\beta$	35.2	3.7	0.067	70	64		[378]
Ribonuclease H	1RNH	151	$\alpha/\beta$	35.1	28.5	0.127	142.8	19		[300]
Bovine	1V9E	259	$\alpha/\beta$	20.8	40.2	0.132	165	1100		[379]
E2lip3	1QJO	80	all $\beta$	0	53.8	0.211	57.8	$15 \pm 10$	700	[193]
E2lip3(N-41)	1QJO	40	all $\beta$				151	$177 \pm 3$	700	[193]
I27	1TIT	89	all $\beta$	0	59.6	0.178	248.2	$204 \pm 30$	200–1500	[2,3,113,138]
Titin I1	1G1C	100	all $\beta$	0	64.3	0.182	272.7	127	600	[380]
Titin I1	1G1C	100	all $\beta$	0	64.3	0.182	272.7	127	600	[380]
Titin I28	1WIT	93	all $\beta$	0	59.1	0.203	248.9	230	600	[381]
Tenascin (TNfn3)	1TEN	89	all $\beta$	0	80.9	0.171	80.2	113	200–600	[382]
Fibronectin M10	1FNF	94	all $\beta$	0	44.7	0.174	76.2	$74 \pm 20$	600	[383,384]
<sup>12</sup> FNIII <sub>13</sub> FNIII	1FNH	92	all $\beta$	0	43.5	0.176	127.1	$124 \pm 18$	600	[384]
ddFLN4	1KSR	100	all $\beta$	0	39.0	0.152	121.0	$45 \pm 20$	200–400	[147,294,385]
Domain C2A	1RSY	126	all $\beta$	0	43.7	0.159	59.8	60	600	[309]
Domain I54-59	1NCT	98	all $\beta$	0		0.186	231	210	500–1000	[386,387]
<sup>1</sup> FNIIIi27	1OWW	97	all $\beta$	0		0.164	86.5	120	600	[384]
$\alpha$ -spectrin R16	1AJ3	98	all $\alpha$	87.8	0	0.080	29.9	30, 60–80	800, 3000	[282,296]
$\alpha$ -spectrin13–18	1U4Q	106	all $\alpha$	87.7	0	0.095	36.7	$26 \pm 15$	300	[282,388,389]
$\alpha$ -spectrin8–9	1S35	106	all $\alpha$	91.5	0	0.093	32.6	$27 \pm 13$	300	[282,388,389]
$\alpha$ -actin1–4	1HCl	119	all $\alpha$	73.1	0	0.089	39	38	300	[282,389]
Calmodulin	1CFC	148	all $\alpha$	56.8	0	0.055	28	15	600	[309]
Ankyrin 1	1N11	33	all $\alpha$	63.6	0	0.044	40.0	37		[318,368]
Ankyrin 6-repeats	1N11	198	all $\alpha$	63.6	0	0.044	56	$50 \pm 20$	600	[318]
Ankyrin 24-repeats	1N11	792	all $\alpha$	63.6	0	0.044		450		[368]
VACC1 <sup>1+DTT</sup>	1VCS	96	all $\alpha$	0		0.084	53.7	40	1000	[374]
b	2PDD	43	all $\alpha$	48.8	0	0.110	48.2			[390] <sup>a</sup>
IM9	1IMQ	86	all $\alpha$	53.5	0	0.118	33.8			[391] <sup>a</sup>
Cytochrome C	1YCC	103	all $\alpha$	51.5	0	0.115	55.8			[392] <sup>a</sup>
Cytochrome C	1HRC	104	all $\alpha$	44.2	0	0.111	53.7			[393] <sup>a</sup>
Acyl-coenzyme A	2ABD	86	all $\alpha$	68.6	0	0.137	31.1			[394] <sup>a</sup>
Cytochrome B562	256B	106	all $\alpha$	78.3	0	0.073	40.1			[395] <sup>a</sup>
$\lambda$ -repressor	1LMB	80	all $\alpha$	66.3	0	0.080	37.9			[396] <sup>a</sup>
Myoglobin	1F63	154	all $\alpha$	87.0	0	0.082	35.8			[397] <sup>a</sup>
DNA								14	40	[249,379]
RNA (P5ab hairpin)								15		[398]

<sup>a</sup> References, taken from PDB, refer to structures of proteins whose mechanical properties have not been studied yet.<sup>b</sup> The molecule full name is Dihydrolipoamide Acetyltransferase. The simulation results for  $f_u$  obtained at  $v = 3.6 \times 10^7$  nm/s, are mainly taken from Ref. [283].

where  $\Delta_{ij}$  is equal to 1 if residues  $i$  and  $j$  form a native contact and 0 otherwise, and  $\theta(x)$  is the Heaviside function. The argument of this function guarantees that a native contact between  $i$  and  $j$  is classified as formed when  $r_{ij}$  is shorter than  $1.2r_{0ij}$ . Here  $r_{0ij}$  is the distance between beads  $i$  and  $j$  in the native state. The probability of being in the native state,  $f_N$ , which can be measured by various experimental techniques, is defined as  $f_N = \langle \chi \rangle$ , where  $\langle \dots \rangle$  stands for a thermal average. The  $f$ – $T$  phase-diagram (a plot of  $1 - f_N$  as a function of  $f$  and  $T$ ) and thermodynamic quantities were obtained by the multiple histogram method [287] extended to the case when the external force is applied to the termini [81,288]. In this case the reweighting is carried out not only for temperature but also for force. Histograms are collected for several values of  $T$  at  $f = 0$  and for a few values of  $f$  at a fixed value of  $T$ .

For illustration, following Li et al. [94], we consider the protein ubiquitin (Ub) (the PDB structure is given in Fig. 37(a)) using the Go model (Eq. (36)). With the choice  $d_c = 6.5$  Å the total number of native contacts  $Q_{\max} = 99$ . We assume



**Fig. 37.** (a) Native state conformation of ubiquitin taken from the PDB (PDB ID: 1ubq). There are five  $\beta$ -strands: S1 (2–6), S2 (12–16), S3 (41–45), S4 (48–49) and S5 (65–71) and one helix A (23–34). (b) The  $f$ – $T$  phase-diagram obtained by the extended replica exchange and histogram method (taken from Ref. [94]). The color code for  $1 - f_N$  is given on the right, where  $f_N$  is the population of the native state. The blue color corresponds to the state in the folded state, while the red color indicates the unfolded state. The vertical dashed line refers to  $T = 0.85T_F \approx 285$  K at which most simulations were performed.

Source: By permission from Ref. [94].

that the dynamics of the polypeptide chain obeys the Langevin equation. The equations of motion (see Ref. [289] for details) were integrated using the velocity form of the Verlet algorithm [132] with a time step  $\Delta t = 0.005\tau_L$ , where  $\tau_L = (ma^2/\epsilon_H)^{1/2} \approx 3$  ps. The  $f$ – $T$  phase-diagram of Ub is shown in Fig. 37(b). The dimensionless melting temperature is  $T_F = 0.675\epsilon_H/k_B$ . Using the experimental value  $T_F = 332.5$  K [290] we obtain  $\epsilon_H = 4.1$  kJ/mol = 0.98 kcal/mol. Then the force unit [ $f$ ] =  $\epsilon_H/\text{\AA}$  = 68.0 pN, which is widely used in this review.

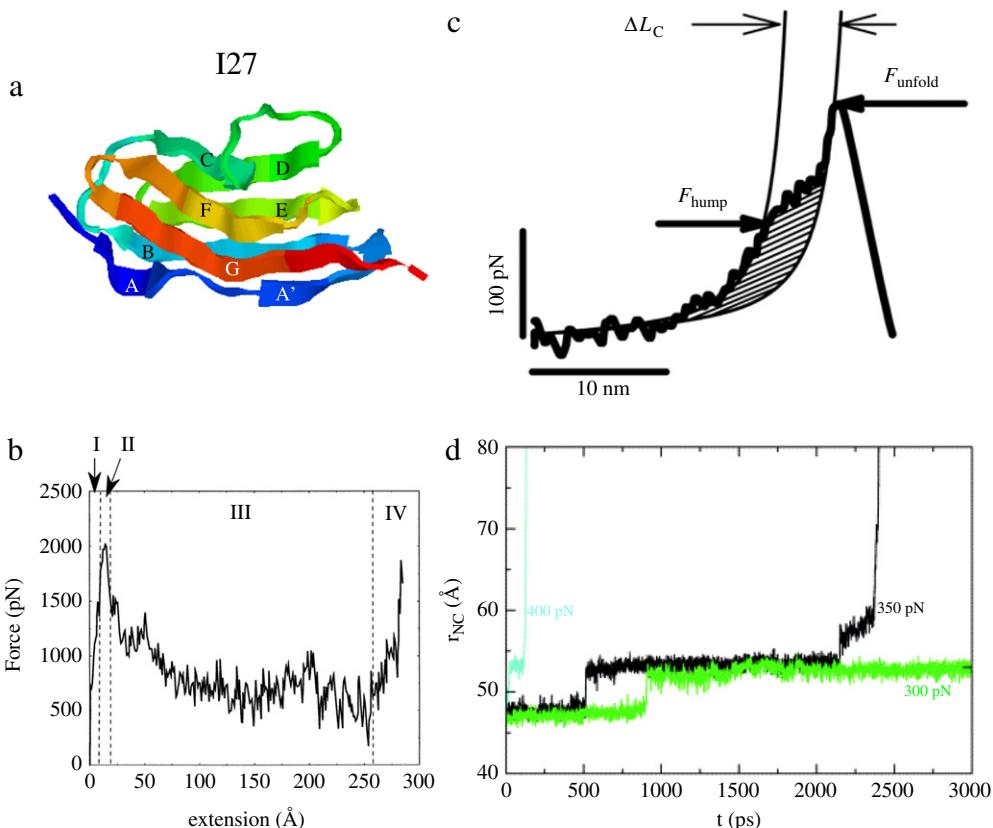
The folding–unfolding transition, defined by the yellow region, is sharp in the low temperature region, but it becomes less cooperative (the fuzzy transition region is wider) as  $T$  increases. One can show that, as in the homopolymer case, the transition is first order. The weak re-entrancy (the critical force slightly increases with  $T$ ) occurs at low temperatures. This seemingly strange phenomenon occurs as a result of a competition between the energy gain and the entropy loss upon stretching. A similar cold unzipping transition was also observed in a number of models for heteropolymers [167] and proteins [81] including the C <sub>$\alpha$</sub> –Go model for I27 (MS Li, unpublished results). As follows from the phase diagram, at  $T = 285$  K the equilibrium critical unfolding force  $f_c^{eq} \approx 30$  pN.

In order to estimate  $f_c^{eq}$  from experimental data, we recast Eq. (61) in the following form

$$f^* = f_c^{eq} \ln \frac{v}{v_{eq}}, \quad (89a)$$

$$f_c^{eq} = \frac{k_B T}{x_u}, \quad v_{eq} = \frac{k_u(0)k_B T}{K_r x_u} \approx k_u(0)L, \quad (89b)$$

where  $L$  is the contour length. The physical meaning of  $f_c^{eq}$  in the last equation is clear: in equilibrium a mechanical energy  $f_c^{eq}x_u$  of the same order as the thermal energy is required to unfold a protein. For Ub, using the typical value  $x_u \approx 2$  Å, for  $T = 285$  K, we obtain  $f_c^{eq} \approx 20$  pN, which is close to the theoretical value in Fig. 37. Given the simplicity of the Go model this agreement can be considered satisfactory. For I27 which has  $x_u \approx 2.5$  Å we have  $f_c^{eq} \approx 17$  pN for  $T = 300$  K. If  $x_u$  is not known one can use the logarithmic dependence of  $f_c$  on  $v$  (Eq. (89a)) and experimental data to extract  $f_c^{eq}$ . Using raw pulling data from Refs. [2, 194] we obtain  $f_c^{eq} \approx 18$  pN and 25 pN for I27 and Ub, respectively. These estimates are close to those obtained by using  $x_u$ . The green fluorescent protein (GFP) starts to fracture at  $f_c \approx 35$  pN in a loading-rate-independent



**Fig. 38.** (a) Native state conformation of Ig27 domain of titin (PDB ID: 1tit). There are 8  $\beta$ -strands: A (4–7), A' (11–15), B (18–25), C (32–36), D (47–52), E (55–61), F(69–75) and G (78–88). (b) Force extension profile of SMD all-atom simulations with a pulling speed of  $5 \times 10^{10}$  nm/s. The extension domain is divided into four regions: I, preburst; II, major burst; III, postburst; IV, pulling of fully extended chain [95].

Source: By permission from Ref. [95].

(c) Plot of the first force peak of a sawtooth pattern. The hump begins at a force,  $F_{\text{hump}}$ , that is smaller than the force required to completely unfold the module,  $F_u$ . The thin line is the fit of the WLC model to the data before the hump. The contour length of the second fit is  $\Delta L_c = 46.5$  Å longer than the one before the hump [118].

Source: By permission from Ref. [118].

(d) The end-to-end distance  $r_{NC}$  of the I27 domain as a function of time for three simulations. In all three simulations a metastable intermediate (I) is observed. At 300 pN (green) I is stable for  $> 3$  ns. At an applied force of 350 pN (black) I is seen to unfold further after  $\sim 2.7$  ns, while at 400 pN (cyan) I is short-lived and the transition state is crossed at  $r_{NC} > 59$  Å.

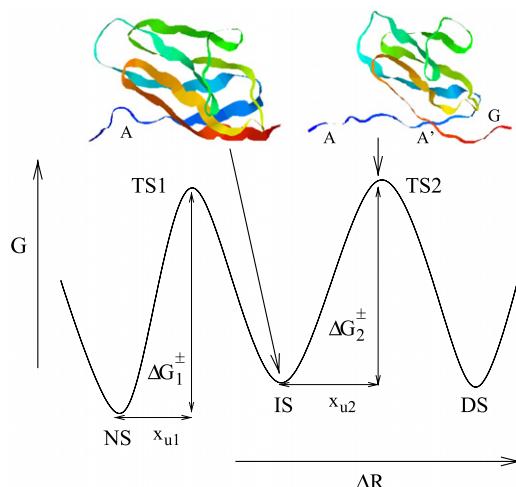
Source: By permission from Ref. [113].

equilibrium process [291]. Thus, the typical equilibrium unfolding force of the proteins is of the order of tens of pN. For helix-rich proteins which have high values of  $x_u$  (see below)  $f_c^{\text{eq}}$  may drop to a few pN.

## 7.2. Protein mechanical unfolding via intermediates

Many proteins mechanically unfold in an all-or-none fashion. In this case the force-extension curves show a characteristic saw-tooth like pattern, where each peak is attributed to the breakage of a folded structure (see, for example, Fig. 36). However, many proteins mechanically unfold via intermediates. For illustration we consider the titin domain I27 (see Fig. 38(a)), which has been studied in great detail. Fig. 38(b) shows the force-extension curve obtained by all-atom simulations [95]. The “hump” is seen in the first preburst region before the unfolding peak at about 200 pN. Interestingly, its existence was predicted first by simulation [95] and then confirmed by the experiment of Marszalek et al. [118] (Fig. 38(c)). This “hump”, centered at  $\Delta R \approx 6$  Å, is attributed to as the obligatory intermediate on mechanical unfolding pathways of I27. It occurs due to the breaking of six hydrogen bonds between A and G strands [95]. *In vitro* at a pulling speed  $v = 200$  nm/s the intermediate is populated at forces of  $\approx 150$  pN, which is lower than the unfolding force  $f_u \approx 200$  pN [118].

The occurrence of unfolding intermediates can be detected by experiments or simulations where a constant force is applied. In this case, instead of the force-extension curve, one studies the time dependence of the end-to-end distance (Fig. 38(d)). The existence of two plateaus is a clear signature of the population of intermediate states since unfolding without intermediates would proceed with only one plateau. The stepwise behavior of the force-extension curve is also observed for homopolymers which unfold via intermediates (see Section 6).



**Fig. 39.** Schematic plot of the free energy landscape for a three-state protein as a function of the end-to-end distance.  $x_{u1}$  and  $x_{u2}$  refer to the distance between the NS and the first transition state (TS1) and the distance between the IS and the second transition state (TS2).  $\Delta G_1^\ddagger = G_{TS1} - G_{NS}$  and  $\Delta G_2^\ddagger = G_{TS2} - G_{IS}$  are unfolding barriers. Shown are typical snapshots for the intermediate and second transition state of I27.

If a protein unravels in the all-or-none fashion then its free energy landscape has two local minima as shown in Fig. 10. In the case where unfolding proceeds via intermediates one observes an additional local minimum corresponding to the intermediate state (Fig. 39). A typical snapshot for intermediates of I27 (Fig. 39) corresponds to the hump in Fig. 38(c) or to the first plateau in Fig. 38(d). Strand A is fully detached from the core, while the remaining seven  $\beta$ -strands remain ordered. Thus, the intermediate state of I27 is very close to the native state. The second transition state occurs at the end-to-end extension  $\Delta R = 12\text{--}15 \text{ \AA}$  (second plateau in Fig. 38(c)). As seen from the snapshot in Fig. 39, strands A and A' have lost all native contacts whereas strand G almost unfolds. From the dependence of the hump and unfolding forces on the pulling speeds Williams et al. [292] obtained the distance between NS and the first transition state (TS1),  $x_{u1} = 2.2 \text{ \AA}$  and the distance between the intermediate state (IS) and second transition state (TS2),  $x_{u2} = 3.0 \text{ \AA}$ . They also estimated the unfolding barriers  $\Delta G_1^\ddagger \approx 21k_B T$  and  $\Delta G_2^\ddagger \approx 24k_B T$ .

The fourth domain of *Distyostelium discoideum* filamin (DDFLN4) was found to unfold via an intermediate which contains a core of the 60 carboxy-terminal residues [293,294]. Thus, even though the native topology of this protein is similar to that of I27 the mechanical intermediate is much less native-like. It occurs at  $\Delta R \approx 120 \text{ \AA}$ , while the intermediate of I27 is at  $\Delta R \approx 6 \text{ \AA}$ . Schwaiger et al. obtained  $x_{u1} = 4 \text{ \AA}$ ,  $x_{u2} = 5.3 \text{ \AA}$ , and  $\Delta G_1^\ddagger \approx \Delta G_2^\ddagger \approx 17k_B T$  [294]. These results agree with the theoretical estimates [295] using Go modeling [87].

A study by Rief et al. [282] of  $\alpha$ -spectrin unfolding showed that its domain unravel at much lower force (25–35 pN) than those of either titin or tenascin. Spectrin unfolds in a cooperative way and no intermediate was found. Lenne et al. [296], using engineered protein constructed with 16 spectrin domains, showed that unfolding occurs in a stepwise fashion during stretching. The force extension pattern gives a clear signature of the existence of an intermediate state between the folded and unfolded state. In another experiment Altmann et al. [297] clearly showed that the unfolding of spectrin is different from Ig-like domains. It was also shown that several unfolding pathways exist. Paci and Karplus [112] did a molecular dynamics study of the unfolding of spectrin. Their simulations also gave a clear signature for an intermediate state but could not reproduce the observation of Altmann et al. [297]. In another effort Klimov and Thirumalai [84] obtained predictions from a simple model based on comparative interaction between complete secondary structure block, which they assumed to unfold in a all-or-none fashion. Based on this model of spectrin they assumed that it is mainly the native topology that determines a proteins resistance to an external force in accordance with simulation performed by Lu and Schulten [95].

The T4-lysozyme (PDB ID: 1L63), which mainly consists of helices, unfolds via multiple distinct unfolding pathways. The majority of T4-lysozymes unfolds in all-or-none fashion involving a dominant unfolding kinetic barrier [298]. A small fraction of T4-lysozymes unfolds in a three-state manner involving unfolding intermediate states. The three-state unfolding pathways are not following well-defined routes, instead they display great variability and diversity [298]. These results provide direct evidence that the mechanical unfolding of T4-lysozyme is governed by the same general kinetic partitioning mechanism as in the case of folding [299].

Cecconi et al. [300] studied the force induced unfolding and refolding trajectory of a single molecule of the *Escherichia coli* protein ribonuclease H (Rnase H) using optical tweezers. They found that the protein unfolds in a two-state manner, but refolds through an intermediate state that correlates with the transient molten globule like intermediate. This intermediate displays unusual mechanical compliance but unfolds at substantial lower forces than the native state.

Dietz and Rief [291] used AFM to drive single GFP molecules from their native state through their complex energy landscape into the completely unfolded state. They found that GFP unfolding proceeds by means of two subsequent intermediate states. They could construct complex mechanical unfolding pathways through a rough energy landscape. The

existence of two intermediates has also been seen in Go simulations [93]. In another interesting paper Dietz et al. [196] have examined the occurrence of mechanically unfolding intermediates due to pulling at different positions. They found that a single intermediate appears when force is applied to residue 3 and 212. In other four cases ((117, 182), (3, 132), (132, 212) and (182, 212)) unfolding occurs in a two-state manner.

Fernandez and coworkers [301] and Chyan et al. [302] investigated the mechanical properties of polyubiquitin by SMFS. It was found that the unfolding may proceed through rare intermediates but the overall behavior remains two-state like [301] although the three-state scenario has also been reported [303]. So, from the experimental point of view the question about the existence of intermediates on unfolding pathways of ubiquitin remains a puzzle. Recently Irback et al. [117] performed Monte Carlo simulations on the mechanical unfolding of ubiquitin and found that unfolding can occur either in a single step or through intermediates. The frequency of the occurrence of intermediates showed a clear systematic dependence on the strength of the applied force. Using the Go model [87] Li et al. [94] did not observe intermediates. However, with the help of an even simpler spin Go-like model [99], one was able to reproduce them [304].

### 7.3. Mechanical unfolding pathways are different from thermal and chemical ones

In mechanical experiments the force replaces the role of a denaturant by pulling biomolecules apart. One of the issues debated by the community is the similarity of free energy landscape parameters when using different techniques to probe the landscape. Thermodynamic properties such as the free energy of unfolding are state functions and they therefore depend only on the initial conditions and final state of the process. However, the kinetics of a reaction is pathway dependent. In single molecule experiments there is a reaction co-ordinate which is different from the bulk. Although it was initially believed that mechanical unfolding probes the same unfolding pathways [138], later experimental work clarified that it was not the case [113,193,305]. One of the most well known examples is that under thermal fluctuations strand C of I27 unfolds first [112] while a mechanical force detaches strand A from the core first [95]. In the case of ubiquitin strand S5 is weakest in thermal unfolding [306] but S1 is predicted [94,117] to unzip first under mechanical force.

The difference between mechanical and thermal or chemical unfolding is manifested by the difference in unfolding rates. For example, Brockwell et al. [307] found a five times difference in the unfolding rate constant at zero force in comparison to the chemical denaturant rate constant  $1.1 \times 10^{-2} \text{ s}^{-1}$  for I27. Studies by Best et al. [308] on Barnase revealed similar results with an unfolding rate constant an order of magnitude slower than I27 WT at zero denaturant but two orders of magnitude faster at zero force. The pronounced difference between mechanical and thermal/chemical unfolding comes from the fact that the force is applied locally to the termini while thermal fluctuations or chemical denaturants have a global effect on an entire protein. In the force case unzipping should propagate from the termini whereas under thermal fluctuations the most unstable part of a polypeptide chain unfolds first.

### 7.4. Mechanical resistance of proteins: Correlation between unfolding force and secondary structure content

#### 7.4.1. Definition of mechanical stability of biomolecules

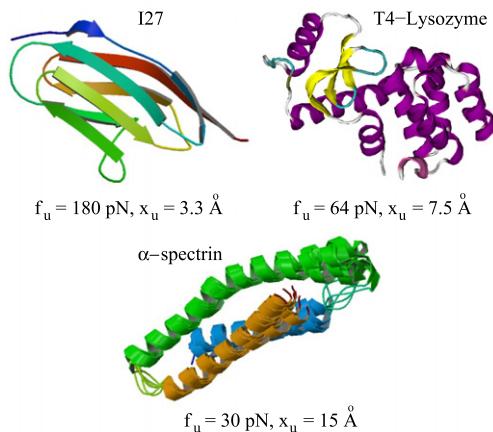
The question we now ask is what quantity can be used to measure the mechanical stability of a biomolecule at the quantitative level? Clearly the more stable it is the larger a force is needed to stretch it out. Therefore, one can use the critical equilibrium force  $f_c^{eq}$  given by Eq. (89b) as a measure of mechanical resistance. However, pulling experiments do not probe  $f_c^{eq}$  directly because they are carried out in non-equilibrium conditions with finite pulling speeds. For this reason, to characterize protein mechanical stability, we use the unfolding force  $f_u$ , which is identified as the maximum force,  $f_{max}$ , in the force-extension profile, i.e.  $f_u \equiv f_{max}$ . If this profile has several local maxima then we choose the largest one. Since  $f_u$  depends on the pulling speed (in the Bell-Evans-Richtie approximation this dependence is logarithmic as shown in Eq. (89a)), in order to compare the mechanical stability of different biomolecules, one must use data obtained for the same  $v$ . Note that  $f_c^{eq}$  and  $f_u$  are mutually related and by using Eq. (89a) one can extract  $f_c^{eq}$  if  $f_u$  is known for at least two pulling speeds.

Experimental and theoretical results, accumulated over more than one decade, allow one to ask: What are the main factors that govern mechanical strength? As a consequence of the local nature of the applied force the type of the secondary structural motif is thought to be important, with  $\beta$ -sheet structures being more mechanically resistant than all  $\alpha$ -helix ones [112,282,305,309–312]. For example, the  $\beta$ -protein I27 and the  $\alpha/\beta$ -protein ubiquitin have  $f_u \approx 200 \text{ pN}$  which is considerably higher than  $f_u \approx 30 \text{ pN}$  for purely  $\alpha$ -spectrin (Fig. 40).

#### 7.4.2. Validity of Go models for estimating $f_u$

A summary of the values for  $f_u$  obtained from constant speed pulling experiments is listed in Table 1.  $f_u$  was computed using either all-atom or coarse-grained models. Simple Go models allow one to estimate  $f_u$  for a huge number of proteins [9,283]. The results [283] obtained by using the Go model [87] for pulling rate  $v_p = 3.6 \times 10^7 \text{ nm/s}$ , are presented in Table 1. Note that the values of  $f_u$  for  $\alpha$ -proteins in this table are higher than those from Ref. [283]. This is because they were obtained at  $T = 0.3\epsilon_H/k_B \approx 285 \text{ K}$ , while the older results were obtained at  $T = 0.53\epsilon_H/k_B$  which is too high for this class of proteins.

The pulling speed used in Go simulations [283] is about five orders of magnitude faster than typical experimental values but is about two to three orders of magnitude slower than those used in all-atom SMD simulations [106]. Despite very



**Fig. 40.** Typical examples for three classes of proteins. Experimental values of  $f_u$  and  $x_u$  are given.

**Table 2**  
Survey of SMD results for  $f_u$ .

Protein	PDB ID	N	SCOP Class	$f_u$ (pN) SMD	$v$ (Å/ps)	Ref.
Ubiquitin(N-C)	1UBQ	76	$\alpha/\beta$	2000	0.1	[194,315]
Ubiquitin(48-C)	1UBQ	28	$\alpha/\beta$	1200	0.1	[194,302]
Barnase	1BNI	108	$\alpha/\beta$	500	0.01	[308]
bovine	1V9C	259	$\alpha/\beta$	3000	0.5	[379]
T4 Lysozyme	1BGI	164	$\alpha/\beta$	73		[84] <sup>a</sup>
cad1	1EDH	211	all $\beta$	1850	0.5	[399]
cad2	1EDH	211	all $\beta$	1970	0.5	[399]
VCAM1	1VSC	89	all $\beta$	2050	0.5	[399]
VCAM2	1VSC	108	all $\beta$	1620	0.5	[399]
I1 oxidized	1GCG	97	all $\beta$	2397	0.5	[400]
I1 reduced	1GCG	97	all $\beta$	2090	0.5	[400]
I27	1TIT	89	all $\beta$	2479	0.5	[302,400]
I27	1TIT	89	all $\beta$	2040	0.5	[95,399,401–403]
I28	1TIT	93	all $\beta$	2082	0.5	[95]
<sup>1</sup> FNIII	1OWW	97	all $\beta$	1500	0.01	[404]
<sup>2</sup> FNIII	1OWW	91	all $\beta$	1600	0.01	[404]
<sup>7</sup> FNIII	1FNF	93	all $\beta$	1638	0.5	[399,405]
<sup>9</sup> FNIII	1FNF	91	all $\beta$	2000	0.1	[403,405]
<sup>10</sup> FNIII	1FNF	94	all $\beta$	1580	0.5	[399,403,405]
DDFLN4	1KSR	100	all $\beta$	700	0.01	[406]
cytochrome C6	1CYI	89	all $\alpha$	no peak	0.5	[399]
binding protein igb	1BDD	60	all $\alpha$	no peak	0.5	[399]
synaptotagmin(c2)	1RSY	125	all $\alpha$	no peak	0.5	[399]
Ankyrin 4-repeats	1N11	132	all $\alpha$	$\approx 210$	0.01	[407]

The values of unfolding force obtained by all-atom SMD simulations.

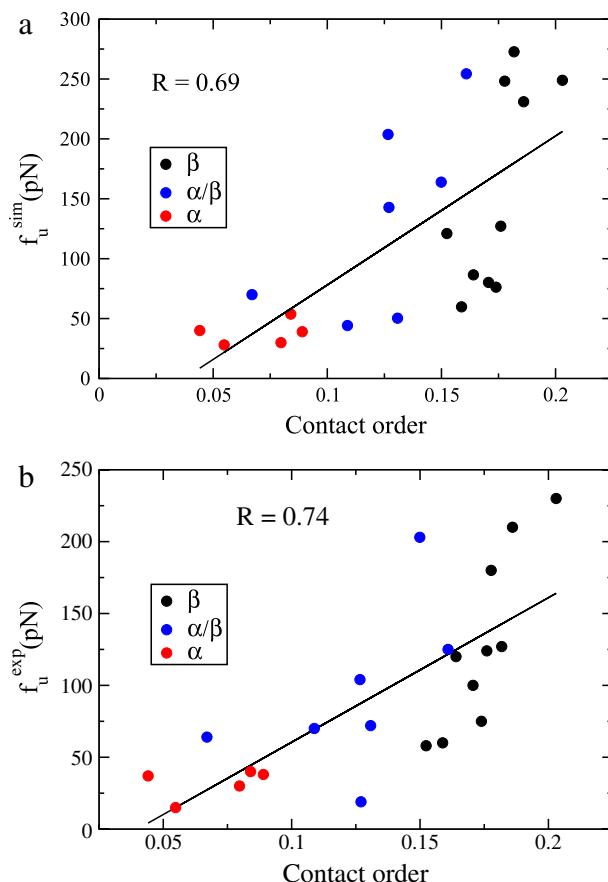
<sup>a</sup> In Ref. [84] the unfolding force was estimated by the protocol based on the secondary structures of the native state but not by all-atom MD.

different loading rates, the values of  $f_u$  obtained from Go modeling and experiments are of the same order of magnitude (Table 1). This is probably an artifact of the simple Go modeling where non-native interactions and the effect of the environment are not taken into account. For spectrin Rief et al. [282], for example, reported  $f_u \approx 30$  pN at the pulling speed  $v = 800$  nm/s. Applying the formula  $f_u \sim \ln v$  [134] we obtain  $f_u \approx 321$  pN for the speed used in Go simulations [283]. This value is much higher than the Go estimate  $f_u \approx 30$  pN (Table 1). Thus, Go modeling is not able to reproduce experimental values for  $f_u$  for individual proteins. However, it remains a useful model to predict their relative values because the correlation level ( $R \approx 0.80$ ) between the Go and experimental results is very high [9,283]. Therefore, one can use the Go model to study the mechanical stability of proteins at a qualitative level.

Table 2 shows values of  $f_u$  obtained by all-atom simulations for various proteins. Typically  $f_u \sim 1000$  pN at a pulling speed  $v \sim 10^{10}$  nm/s. From the logarithmic dependence  $f_u \sim \ln v$  one can obtain  $f_u \sim 100$  pN at experimental speed  $v \sim 100$  nm/s. So all-atom simulations give reasonable estimates for  $f_u$  for proteins *in vitro*.

#### 7.4.3. Unfolding force, secondary structure content, and contact order

As mentioned above, the mechanical stability of proteins depends on the content of their secondary structures. This is also clearly seen from Table 1. Using this table one can show [283] that  $f_u$  decreases with the helix content linearly (the



**Fig. 41.** (a) Dependence of theoretical values of  $f_u$  on CO (linear fit  $y = -46.287 + 1243.7x$ ). The data set is of 22 proteins: 10  $\beta$ - (PDB ID: 1TIT, 1G1C, 1WIT, 1TEN, 1FNF, 1FNH, 1KSR, 1RSY, 1NCT, and 10WW), 7  $\alpha/\beta$ - (1HZ6, 1UBQ, 1BNI, 1HFR, 1B9C, 1B6I, and 1RNH), and 5  $\alpha$ -proteins (1AJ3, 1HCl, 1CFC, 1VCS, and 1N11). (b) The same as in (a) but for the experimental results (linear fit  $y = -40.131 + 1005.8x$ ). Figure has been taken from Ref. [283].

helix content is defined as  $N_\alpha/N$ , where  $N$  and  $N_\alpha$  are the total numbers of amino acids and the numbers of amino acids from  $\alpha$ -fragments). It is valid for both experimental and theoretical sets. A very poor correlation between the mechanical stability and the beta content [283] may be due to a limited number of collected proteins. This calls for further theoretical and experimental studies.

As far as the correlation between  $f_u$  and the beta content remains low, it is highly desirable to find out if there is any correlation between  $f_u$  and the so called contact order (CO) [284]. Intuitively, such a correlation should exist because this parameter reflects the native topology which plays an important role in the mechanical unfolding of proteins. Moreover, the beta-rich proteins have a higher CO compared to helix-rich ones [283].

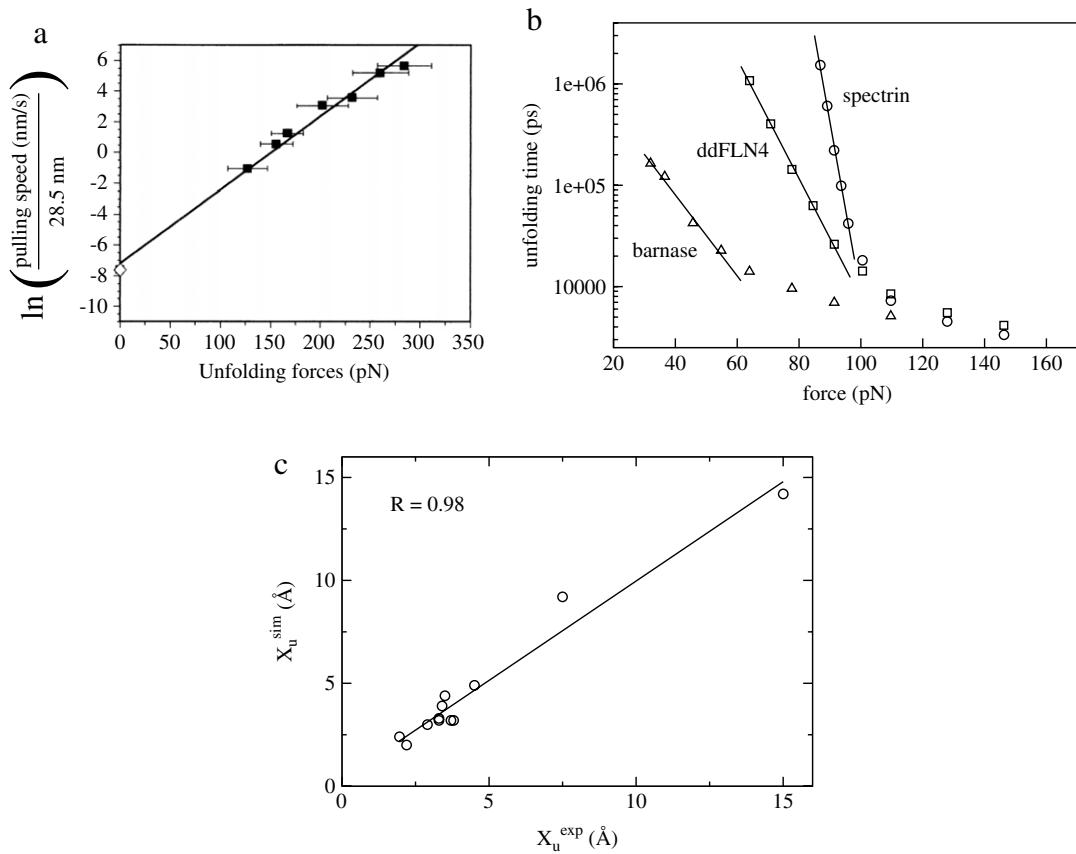
The CO is defined as follows [284]

$$\text{CO} = \frac{\sum_{ij} \Delta_{ij} |i - j|}{N \sum_{ij} \Delta_{ij}}, \quad (90)$$

where  $N$  is the number of residues,  $\Delta_{ij} = 1$  if amino acids form a contact and  $\Delta_{ij} = 0$  otherwise. The value of CO depends on how a contact is defined. Here, we adopt a definition based on the position of side chains [313]. A contact between any two amino acids ( $|i - j| \geq 1$ ) is said to be formed if the distance between the centers of mass of side chains  $d_{ij} \leq 6.0 \text{ \AA}$  (see [http://depts.washington.edu/bakerpg/contact\\_order/](http://depts.washington.edu/bakerpg/contact_order/) for the code for computing the CO by this definition). The CO's of various proteins are shown on Table 1.

Fig. 41 shows the dependence of  $f_u$  on the CO for 22 proteins. Since helix-rich proteins are less mechanically stable compared to helix-poor proteins  $f_u$  grows with the CO. The linear fit (Fig. 41) gives

$$f_u = \begin{cases} -46.287 + 1243.7 \times \text{CO}, & \text{from simulations,} \\ -40.131 + 1005.8 \times \text{CO}, & \text{from experiments,} \end{cases} \quad (91)$$



**Fig. 42.** (a) The dependence of the unfolding force on pulling speeds for the domain 27 (the result was obtained by AFM experiments). This dependence yields  $x_u = 2.5 \text{ \AA}$  [138].

Source: By permission from Ref. [138].

(b) Dependence of the unfolding times on  $f$  from Go simulations [283]. Fitting to the Bell Eq. (55) gives  $x_u$  as shown in Table 3. (c) The simulation values of  $x_u$  are plotted versus the experimental ones. In the case when the experiments provide different values of  $x_u$  (Table 3), we took their average. The linear fit ( $y = 0.31 + 0.966x$ ) has a high correlation level  $R = 0.98$ . Figure has been taken from Ref. [283].

where  $f_u$  is measured in pN. The correlation level is  $R = 0.69$  and  $0.74$  for the simulation and experimental sets, respectively. The correlation for the experimental set is much improved compared to the smaller set of  $14$  proteins studied in Ref. [283]. Thus, one can expect that the correlation between the mechanical stability and CO is robust. The hallmark of Eq. (91) is that it can be used to estimate the mechanical stability using the native topology only.

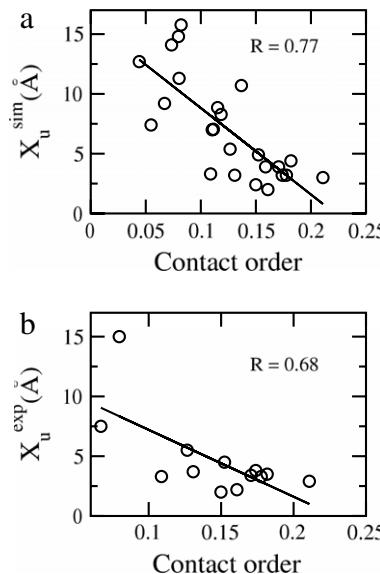
### 7.5. Free energy landscape in the Bell approximation

Nowadays AFM or optical tweezers [2] have proved to be a very useful tool to probe the free energy landscape of biomolecules. Assuming that  $x_u$  does not depend on the external force one can use the Bell formula to extract it from dependencies of unfolding times on the external force. In the last decade the Bell approximation, Eq. (57), has been repeatedly refined using various approximations [135,147,150,314]. However, this approximation works pretty well for many experiments [301,307] and we will use it to extract the distance between NS and TS,  $x_u$ . The main goal of this chapter is to show that  $x_u$  is strongly correlated with the secondary structure content and CO. Beyond the Bell approximation, using the microscopic kinetic theory [135] one can calculate not only  $x_u$  but also unfolding barriers. This problem will be discussed in the next section.

#### 7.5.1. Validity of Go models for estimation of $x_u$

In the Bell approximation one can estimate  $x_u$  either from the dependence of the unfolding force on the pulling speed, see Eq. (61), or from the dependence of the unfolding times on the external forces as per Eq. (57). The first approach is often used in pulling experiments, while the second one is widely adopted in simulations (Fig. 42(a) and b).

To the best of our knowledge  $x_u$  was experimentally determined for  $13$  proteins as listed in Table 3. With the help of all-atom simulations Li et al. [315] have shown that at low forces, where the Bell approximation is valid,  $x_u = 10 \text{ \AA}$  for Ub. This is noticeably higher than the experimental value  $x_u = 2.4 \text{ \AA}$  (Table 3). Presumably, this is due to the fact that these



**Fig. 43.** (a) Dependence of theoretical values of  $x_u$  on CO (linear fit  $y = 16.049 - 72.312x$ ). (b) The same as in (a) but for the experimental set (linear fit  $y = 12.8 - 55.953x$ ). The correlation level of the linear fitting is shown on the plot. Figure has been taken from Ref. [283].

authors computed  $x_u$  from equilibrium data, but their sampling was not good enough for such a long protein as ubiquitin. Here we focus on theoretical estimates of  $x_u$  [283] obtained by Go modeling [87] for many proteins (Table 3). Using a more sophisticated version of the Go model [92] West et al. [316] have obtained almost the same value for protein L. Applying the so called Wako-Saito-Munoz-Eaton model [99,100,317] to I27 Imparato et al. [97] got  $x_u \approx 3.1 \text{ \AA}$ , which is consistent with the Go result listed in Table 3. Furthermore, it was shown that  $x_u$  depends very weakly on temperature [97].

As is evident from Table 3 and Fig. 42 Go modeling [87] gives an acceptable agreement with the experiments. The correlation level between theoretical and experimental values is very high ( $R = 0.98$ ) [283] and this justifies the use of Go models for the computation of  $x_u$ . Our result is very appealing because this simple modeling provides even a quantitative agreement with experiments. Again this is because unfolding is mainly defined by the native topology captured by Go models.

### 7.5.2. Correlation between $x_u$ and helix content

As follows from Fig. 40, the  $\alpha$ -protein spectrin has  $x_u \approx 15 \text{ \AA}$  [282] which is much larger than that for  $\beta$ -I27 and  $\alpha/\beta$ -ubiquitin. Therefore  $\alpha$ -rich proteins seem to have large values of  $x_u$ . Since  $x_u$  was experimentally determined only for spectrin (Table 3), to ascertain this we have calculated  $x_u$  for another 8  $\alpha$ -proteins [283] ( $\alpha$ -protein ankyrin and calmodulin, have been studied experimentally [309,318] but their values of  $x_u$  were not reported). Based on the results from Table 3, one can divide proteins into two main classes. The first class consists of  $\alpha$ - and  $\alpha/\beta$ -proteins which have  $x_u \approx 2-5 \text{ \AA}$ . The second class consists of purely  $\alpha$ -proteins with markedly higher values of  $x_u \approx 7-15 \text{ \AA}$ . In terms of  $x_u$  neither our simulations nor the experiments can distinguish clearly between  $\beta$ -proteins and the mixed ones. Better statistics is required to resolve this delicate issue. Nevertheless, it is obvious that the secondary structure of the native conformation play a decisive role in determination of the distance between the TS and NS.

One can show [283] that the linear regression for the experimental and simulation data gives

$$x_u = \begin{cases} -0.431 + 0.174\alpha, & \text{simulations} \\ 0.159 + 0.16\alpha, & \text{experiments,} \end{cases} \quad (92)$$

where  $x_u$  is measured in  $\text{\AA}$  and the helix content  $\alpha$  is in percent. The correlation level between  $x_u$  and the helix content is equal to 0.91 and 0.94 for the experimental and the simulation sets, respectively. Such a high quality of fitting unambiguously shows the strong correlation between these two quantities: the higher the helix content, the larger is  $x_u$ . It should be stressed that Eq. (92) is useful for estimating  $x_u$  based solely on the topology of the native state.

The correlation level between  $x_u$  and the beta content is very low for both the experimental ( $R = 0.02$ ) and simulation ( $R = 0.25$ ) sets [283]. However, this result was obtained using rather small data sets. Hence, in order to establish if there is any pronounced correlation between  $x_u$  and the beta content one has to generate better statistics.

### 7.5.3. Correlation between $x_u$ and contact order

The dependence of  $x_u$  on the CO is shown in Fig. 43 where the correlation levels are relatively high for both theoretical ( $R = 0.77$ ) and simulation sets ( $R = 0.68$ ). The quality of the linear fitting to the experimental data is lower, but one can

**Table 3**Survey of experimental and theoretical results for  $x_u$ .

Protein	PDB ID	N	SCOP Class	alpha (%)	beta (%)	CO	$x_u$ (Å) Sim.	$x_u$ (Å) Exp.	Ref.
Protein L	1HZ6	62	$\alpha/\beta$	23.8	44.4	0.161	2.0	2.2	[305]
Ub (N-C)	1UBQ	76	$\alpha/\beta$	21.1	43.4	0.150	2.4	1.4–2.5	[194,301,302]
Ub (N-Lys48)	1UBQ	76	$\alpha/\beta$				6.1	6.3	[194]
Barnase	1BN1	108	$\alpha/\beta$	23.2	21.3	0.109	3.3	3.3	[308]
Human DHFR	1HFR	186	$\alpha/\beta$	26.3	24.7	0.131	3.2	3.7	[369]
GFP	1B9C	224	$\alpha/\beta$	9.8	54.9	0.127	5.38	5.5	[291]
GFP (3-132)	1EMB	228	$\alpha/\beta$					1.3	[195]
GFP (3-212)	1EMB	228	$\alpha/\beta$					4.5	[195]
GFP (117-182)	1EMB	228	$\alpha/\beta$					1.2	[195]
GFP (132-212)	1EMB	228	$\alpha/\beta$					3.2	[195]
GFP (182-212)	1EMB	228	$\alpha/\beta$					1.4	[195]
T4-Lysozyme	1B6I	162	$\alpha/\beta$	35.2	3.7	0.067	9.2	7.5	[298]
E2lip3	1QJO	80	all $\beta$	0	53.8	0.211	3.0	2.9	[193]
I27	1TIT	89	all $\beta$	0	59.6	0.178	3.2	3.3	[114]
Titin I1	1G1C	100	all $\beta$	0	64.3	0.182	4.4	3.5	[380]
Tenascin (TNfn3)	1TEN	89	all $\beta$	0	80.9	0.171	3.9	2.0–4.8	[115,408]
Fibronectin M10	1FNF	94	all $\beta$	0	44.7	0.174	3.2	3.8	[384,409]
ddFLN4	1KSR	100	all $\beta$	0	39.0	0.152	4.9	4.0–5.0	[147,294,385]
Domain C2A	1RSY	126	all $\beta$	0	43.7	0.159	3.9		[309]
spectrin	1A3J	98	all $\alpha$	87.8	0	0.080	14.2	15.0	[282]
Calmodulin	1CFC	148	all $\alpha$	56.8	0	0.055	7.4		[309]
Ankyrin 6-repeats	1N11	198	all $\alpha$	63.6	0	0.044	12.6		[318]
b	2PPD	43	all $\alpha$	48.8	0	0.110	7.00		[390] <sup>a</sup>
IM9	1IMQ	86	all $\alpha$	53.5	0	0.118	8.3		[391] <sup>a</sup>
Cytochrome C	1YCC	103	all $\alpha$	51.5	0	0.115	8.85		[392] <sup>a</sup>
Cytochrome C	1HRC	104	all $\alpha$	44.2	0	0.111	7.04		[393] <sup>a</sup>
Acyl-coenzyme A	2ABD	86	all $\alpha$	68.6	0	0.137	10.70		[394] <sup>a</sup>
Cytochrome B562	256B	106	all $\alpha$	78.3	0	0.073	14.10		[395] <sup>a</sup>
$\lambda$ -repressor	1LMB	80	all $\alpha$	66.3	0	0.080	11.30		[396] <sup>a</sup>
Myoglobin	1F63	154	all $\alpha$	87.0	0	0.082	15.77		[397] <sup>a</sup>

The dependence of simulated (23 proteins) and experimental (12 proteins) values of  $x_u$  on the secondary structure content and CO. Theoretical estimates of  $x_u$  were obtained [283] using the Go model [87].

<sup>a</sup> References, taken from PDB, refer to structures of proteins whose mechanical properties have not been studied yet.

b The molecule full name is Dihydrolipoamide Acetyltransferase.

expect it to improve as more data becomes available. So there is a strong correlation between  $x_u$  and the CO. The decrease of  $x_u$  with the CO is consistent with the fact that  $x_u$  grows with the helix content (Eq. (93)), which is anti-correlated with the CO [283]. The linear fit (Fig. 43) gives:

$$x_u = \begin{cases} 16.49 - 75.11 \times CO, & \text{simulations} \\ 14.43 - 66.03 \times CO, & \text{experiments.} \end{cases} \quad (93)$$

Thus, we obtain the very important result that the CO might be used to estimate the distance between the TS and NS of globular proteins regardless of whether they are helix- or beta-rich. CO is more universal than the helix content because the latter can be applied to the  $\alpha$ - and  $\alpha/\beta$ -proteins only. It should be noted that the CO correlates with the folding rates of small two-state proteins but not with three-state ones [313]. On the other hand, this parameter, as shown above, can be used to estimate not only  $f_u$  but also  $x_u$  for two-state proteins as well as multi-state ones. In this sense, CO is more useful for studying unfolding than folding. This is because the energetic frustration plays a much more important role for folding of biomolecules compared to their mechanical unfolding.

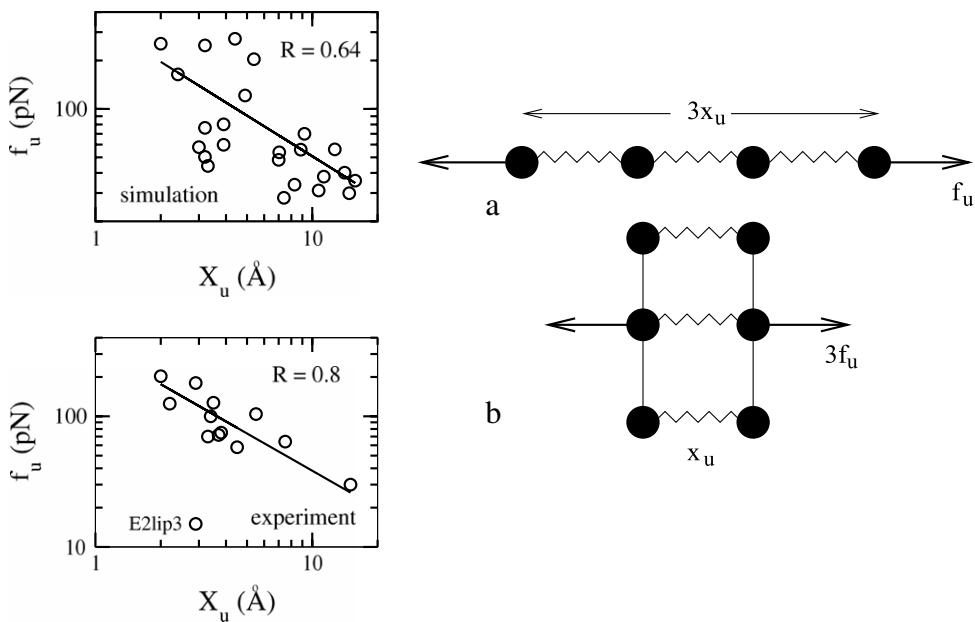
#### 7.5.4. Correlation between $f_u$ and $x_u$

Using data from Tables 1 and 3, one can obtain a relationship between  $f_u$  and  $x_u$  (Fig. 44). The best fit gives the following power law [283]

$$x_u = cf_u^{-\mu}, \quad (94)$$

where the exponent  $\mu = 0.70 \pm 0.15$  and  $0.92 \pm 0.14$  for the simulation and experimental sets, respectively. The high correlation level for the experimental set ( $R = 0.92$ ) implies that there is a strong correlation between these two quantities. A poorer correlation ( $R = 0.65$ ) was obtained for the theoretical data. More data points are expected to improve the correlation. The present estimate of  $\mu$  is much higher than the old value  $\mu \approx 0.4$  [283]. This is because we used improved theoretical values of  $f_u$  for  $\alpha$ -proteins (see Section 7.4.2) and the data for protein E2lip3 was excluded from the experimental set (Fig. 44).

The “scaling” law (Eq. (94)) shows that mechanically more stable biomolecules should have lower values of  $x_u$  because they are less sensitive to external perturbations. The “fractal” nature of the exponent  $\mu$  is physically unclear. However, with



**Fig. 44.** Left: The dependence of theoretical (upper panel) and experimental (lower panel) values of  $f_u$  on  $x_u$ . The straight lines are nonlinear fits to Eq. (94) (for the experimental set, protein E2lip3 is not taken into account). For the experimental set exponent  $\mu = 0.92 \pm 0.14$ , constant  $c = 339.8$ , and the correlation level  $R = 0.92$ . In the simulation case  $\mu = 0.75 \pm 0.15$ , constant  $c = 351$ , and  $R = 0.65$ . (a) Schematic plot for the case where a force is applied parallel with the peptide bonds. (b) The same as in (a) but a force is applied perpendicular to the peptide bonds. Figure has been taken from Ref. [283].

error bars one can expect that  $\mu = 1$  and Eq. (94) becomes

$$f_u x_u = \text{constant}. \quad (95)$$

Recently, a simple network model has been proposed to support the validity of Eq. (95) [319], which can be understood using the following simple argument. Consider two situations with different combinations of peptide bonds (Fig. 44). In case (a) amino acids are connected in serial while in case (b) peptide bonds are parallel. In the former case, the unfolding force is three times smaller compared to the latter one, but the characteristic distance  $x_u$  is effectively three times larger. Therefore, the product  $f_u x_u$  remains constant. The averaged value of  $f_u x_u \approx 468 \text{ pN } \text{\AA}$  and  $\approx 378 \text{ pN } \text{\AA}$  for the simulation and experimental sets, respectively. These values are not far from those quoted in [319].

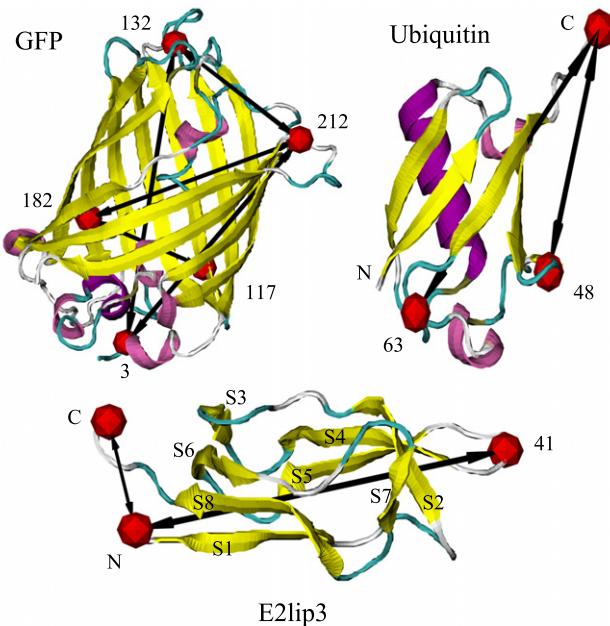
Finally, using the data from Table 3, one can show that  $x_u$  has little correlation with the number of amino acids (the correlation level is below 0.5). Probably, the size of proteins affects the prefactor in the Bell equation but not the exponent itself.

## 7.6. Effect of pulling geometry

### 7.6.1. Influence of pulling geometry on mechanical resistance

Until recently it was believed that not all residues contribute equally to the thermodynamic stability of globular proteins and that its hydrophobic core provides for the most part the stability of the protein structure [320,321]. It was therefore expected that the mechanical unfolding of a protein should be insensitive to the direction of the applied force. Surprisingly, recent experimental investigations [193,194,196] showed that the mechanical stability of ubiquitin, E2lip3 protein, and GFP (Fig. 45) depends notably on the linkage through which the mechanical force is applied to. The protein resistance depends not only on linkage, but also on pulling directions [193,195]. In some sense these two effects are not independent as linkage changes would lead to a change in the pulling direction. The pronounced effect of pulling direction on protein stability may be explained as follows. If a chain is pulled along the direction of hydrogen bonds, then the unfolding is akin to shearing, but when force is applied perpendicular to this direction, the unfolding is akin to unzipping. The force needed to break hydrogen bonds in the former case should be larger than in the latter one. This has been demonstrated in a convincing manner by Kumar and coworkers [123] using a homopolymer lattice model. The effect of pulling direction on  $f_u$  has also been explored using Go modeling [90,139], and all-atom SMD simulations [322].

Pulling Ub at Lys48 and N-terminal decreases the unfolding force almost three-fold compared to the case where the force is applied to both termini (Table 1). For GFP the linkage 3–132 changes the mechanical resistance substantially ( $f_u \approx 350$  at pulling rate  $v = 1000 \text{ nm/s}$  [195]) but pulling at 132–212 and 3–212 positions leaves it almost unchanged. The most drastic linkage effect has been observed for E2lip3 [193] when the protein is stretched at N and position 41. In this case  $f_u$



**Fig. 45.** Native conformations of GFP, Ub and E2lip3 for which the effect of pulling positions on the mechanical resistance was studied experimentally. The external force is applied at amino acids denoted by red circles. Arrows refer to force directions. For GFP the force is applied at (3, 132), (3, 212), (117, 182), (132, 212), and (182, 212) [196]. In the Ub case [194] the force is directed long (N, 48) and (N, 63), while for E2lip3: (N, 41).

increases by one order of magnitude, *i.e.* the unfolding force is found to be 177 pN (at  $v = 700$  nm/s) which is compatible with the mechanical resistance of I27 at the same pulling speed.

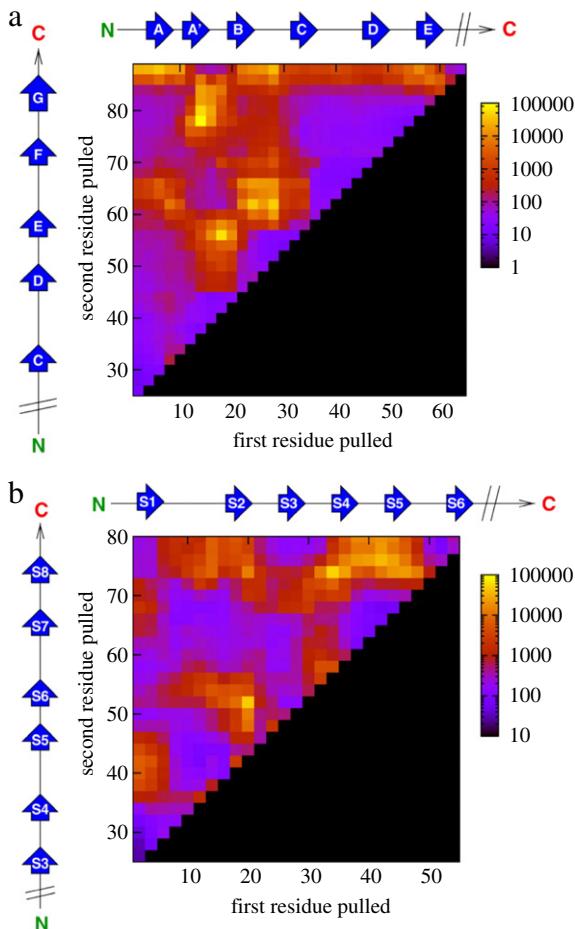
Characterizing mechanical stability by unfolding times (the larger the stability the longer  $\tau_u$ ) and pulling in all possible directions West et al. [90] have constructed the complete unfolding landscape for I27 and E2lip3 (Fig. 46). The unfolding landscape of I27 is simple: High mechanical resistance occurs when the force induces the longitudinal shearing of  $\beta$ -strands. This longitudinal shearing can appear between points that are on the same sheet (the loops between A'B and DE), on different sheets (the C-terminus of G-strand and the DE loop) or on the loops linking each sheet (BC and EF). By contrast it is soft when force is applied orthogonally to  $\beta$ -strands (*e.g.* between BC and DE loops and between CD and FG loops or between the N-terminus and the FG loop) [90]. Interestingly these soft directions are orthogonal to one end of a set of  $\beta$ -strands supporting the hypothesis that mechanical resistance is topologically determined.

E2lip3, by contrast, has a more complex mechanical unfolding landscape. A weak response is found when a force is applied orthogonally to  $\beta$ -strands (for example between the N- and C-terminus) or when loops are peeled apart. This is presumably because each element of mechanical resistance is loaded and fails sequentially. On the other hand, E2lip3 is highly resistant to extension under application of force between pairs of residues (such as 20 and 50) which are in loops connecting both sheets together [90]. Extension of these strands, therefore, implies the shearing apart of a large number of hydrogen bonds.

As shown by experiments [193] and simulations [90], E2lip3 unfolds at significant forces when extended between the N-terminus and residues in the loop connecting stands 4 and 5 (residues 39–43). It shows even greater stability when extended between the C-terminus and residue 41 than when extended by the N-terminus and residue 41 despite their close proximity. This observation may be of functional importance because in the pyruvate dehydrogenase complex of *Escherichia coli*, E2lip3 shuttles an acetyl group (attached via a lipoyl moiety to residue 41) between active sites in the inner icosahedral core and the outer spherical shell [323]. E2lip3 is connected to the outer shell by a C-terminal linker. Thus, any force applied onto this domain during its functional cycle will be through an extension geometry in which this domain is mechanically robust. Some of these highly force-resistant geometries shown by this analysis may be a result of their function, while more mechanically labile geometries may be utilized to allow rapid protein turnover *in vivo*.

#### 7.6.2. Dependence of the free energy landscape on pulling positions

Suppose a protein is pulled at two arbitrary points (not necessary at termini), then the distance between them should be chosen as the reaction coordinate. In other words one has to project the free energy landscape onto this pulling direction. One of the most interesting experimental results of Carrion-Vazquez et al. [194] and Dietz and Rief [196] is that the distance between the NS and TS  $x_u$  of a projected free energy landscape changes drastically. If the force is directed along C-terminal and Lys48 of Ub, then in the Bell approximation  $x_u \approx 0.63$  nm, which is about two and half times larger than the case when the termini N and C are pulled (see Table 3). Using the Go model [87] at low forces, one obtained  $x_u \approx 0.61$  nm [139], which is in good agreement with the experiment [194]. Again the success of the Go model in estimating  $x_u$  while not only pulling at termini but also at other positions is presumably due to the fact that unfolding is mainly governed by the native topology



**Fig. 46.** Go models reveal anisotropy in the mechanical unfolding landscape of (a) I27 (constant force = 150 pN) and (b) E2lip3 (constant force = 100 pN). Yellow to blue colors denote geometries of high to weak mechanical resistance, black denotes regions not pulled. Scale is in picoseconds. Source: By permission from Ref. [90].

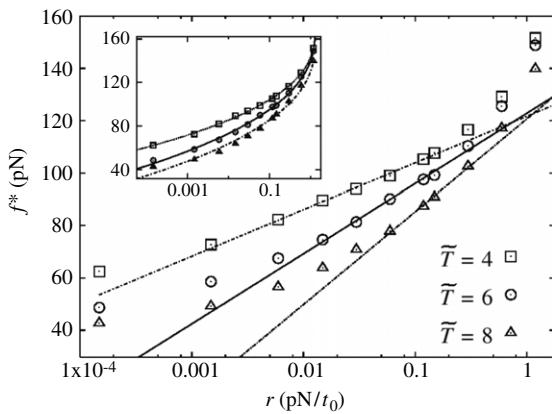
of the part of the protein between the two points subject to the external force. In the GFP case [196] experimental values of  $x_u$  also vary between 1.2 Å and 4.5 Å, depending on pulling directions.

#### 7.6.3. Prediction of protein structures by mechanical pulling

Another interesting development in the use of single molecule experiments has been made by Dietz and Rief [195] who proposed to use pulling in different directions as a way to probe the structure of a protein provided its sequence is known. Their idea is very simple: the distance between any amino acids  $i$  and  $j$ ,  $d_{ij}$ , can be determined as  $d_{ij} = d_{ij}^{full} - \Delta L_{ij}$ , where  $d_{ij}^{full} = |j - i|a$ ,  $\Delta L_{ij}$  is the length gain due to stretching and  $a$  is the distance between two neighboring amino acids. Having all possible  $d_{ij}$  one can obtain the protein structure with angstrom precision. This so-called mechanical triangulation method may be competitive with the standard techniques of X-ray crystallography and NMR.

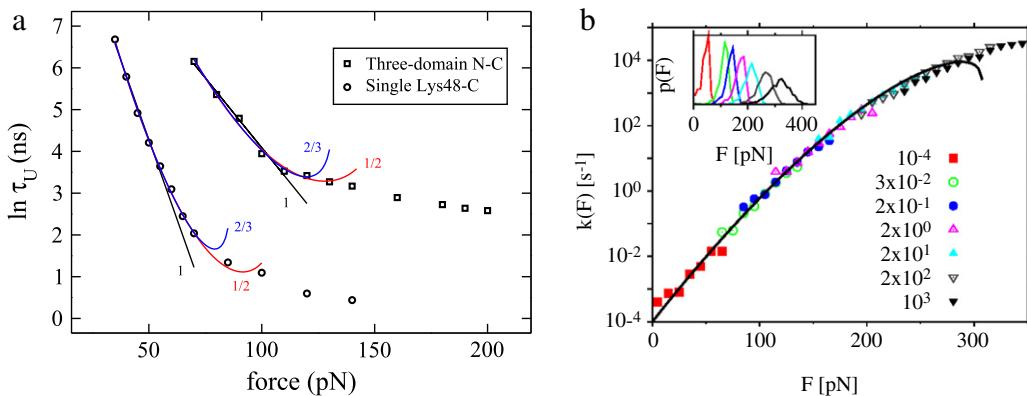
#### 7.7. Free energy landscapes beyond the Bell approximation

In the previous sections we have seen that if one uses the Bell formula, Eq. (55), to compute  $x_u$  then Go modeling provides a reasonable agreement with experiments. Here, we apply a more general microscopic theory [135] (Section 3.5. *Kinetic theory*), which is expected to be more accurate than the Bell approximation, to estimate  $x_u$ . This theory also allows one to estimate the unfolding barrier  $\Delta G^\ddagger$ . From the conceptual point of view, it is worth checking if general formulas such as Eqs. (66) and (68) are really essential for the interpretation of unfolding data. Fig. 47 shows the dependence of  $f^*$  on the pulling speed for I27. The results have been obtained by Imparato et al. [97] using the simple Wako–Saito–Munoz–Eaton model. Clearly, the microscopic theory works for a wider interval of pulling speeds compared to the phenomenological one. This is also evident from the dependence of the unfolding times of Ub on a constant force [139] (Fig. 48(a)). For both cases, when the protein is extended between termini and between residue Lys48 and the C-terminus the Dudko–Hummer–Szabo theory



**Fig. 47.** Dependence of the unfolding (rupture) force  $f^*$  on the loading rate  $r = Kv$  for protein domain I27 at three values of  $T$ . Straight lines are linear fits using Eq. (61). From such fits we obtain nearly temperature-independent  $x_u \approx 3 \text{ \AA}$ . Inset: the lines are the fits of the data to the equation given by Eq. (66) with  $\nu = 2/3$ .

Source: By permission from Ref. [97].



**Fig. 48.** (a) The semi-log plot for the force dependence of unfolding times at  $T = 285 \text{ K}$  for ubiquitin. Results were obtained by using the Go model [87]. Squares and circles refer to the cases when the force is applied to N- and C-terminal of three-domain ubiquitin and to Lys48 and C-terminal of single ubiquitin, respectively. In the Bell approximation if the N- and C-terminal of the trimer are pulled then we have the linear fit  $y = 10.448 - 0.066x$  (black line) and the distance between the native and transition states,  $x_u \approx 0.24 \text{ nm}$ . In the case when we pull at Lys48 and C-terminal of single Ub the linear fit (black line) at low forces is  $y = 11.963 - 0.168x$  and  $x_u = 0.61 \text{ nm}$ . The correlation level of fitting is about 0.99. The red and blue curves correspond to the fits with  $\nu = 1/2$  and  $2/3$ , respectively to Eq. (64). Results are taken from [139].

Source: By permission from Ref. [87].

(b) Constant-force rupture rate  $k(F)$  from the collapse of a constant-speed rupture-force histogram for I27, using Eq. (68). Simulation data (symbols) cover pulling speeds  $v$  from  $10^{-4}$  to  $10^3 \text{ pN/ms}$ , with each  $f$  probing different range of  $F$ . The solid line is the analytical prediction for  $k(F)$  from Eq. (64) with  $\nu = 2/3$ . Deviations at the highest pulling speeds (and rupture forces) are caused by a breakdown of the Kramer's high-barrier approximation. The inset shows the histograms that were used in Eq. (68).

Source: By permission from Ref. [135].

is more reliable. However the difference is not so drastic and it may well be masked in experiments. Presumably, this is a reason of why most experimental data are fitted to the Bell scheme pretty well.

According to the microscopic theory the rupture force distribution  $P(f|v)$  at different ramp speeds can be collapsed onto a single curve for the force-dependent rate of molecular rupture,  $k(f)$  as per Eq. (68). This prediction is supported by the experimental results shown in Fig. 48, where  $v$  spans over seven orders of magnitude. Kinetics parameters, obtained by various fitting schemes for proteins, RNA and DNA, are shown on Table 4. The values of  $x_u$  obtained from the Bell approximation are always lower than microscopic estimates [98,135,139]. Much more experimental as well as theoretical work has to be done to see to what extent the microscopic theory is superior compared to the Bell theory.

To estimate the unfolding barrier from experimental data we use the following formula

$$\Delta G^\ddagger = -k_B T \ln(\tau_A/\tau_0) \quad (96)$$

where  $\tau_0$  denotes the unfolding time in the absence of force and  $\tau_A$  is a typical unfolding prefactor. Since  $\tau_A$  for unfolding is not known, we use the typical value for folding  $\tau_A = 1 \mu\text{s}$  [324–326]. For ubiquitin, e.g., using  $\tau_0 = 10^4/4 \text{ s}$  [327] we obtain  $\Delta G^\ddagger = 21.6k_B T$ . This value is also consistent with  $\Delta G^\ddagger \approx 17.2k_B T$  obtained from Go modeling with  $\nu = 1/2$  (Table 4). For titin, the value of  $\Delta G^\ddagger$ , obtained with the help of experimental kinetic data and Eq. (64) [135] is in reasonable agreement

**Table 4**

Kinetic parameters from different fitting procedures.

	$\nu = 1$		$\nu = 1/2$			$\nu = 2/3$		
	$k_0$ ( $s^{-1}$ )	$x_u$ ( $\text{\AA}$ )	$k_0$ ( $s^{-1}$ )	$x_u$ ( $\text{\AA}$ )	$\Delta G^\ddagger$ ( $k_B T$ )	$k_0$ ( $s^{-1}$ )	$x_u$ ( $\text{\AA}$ )	$\Delta G^\ddagger$ ( $k_B T$ )
I27	$38 \times 10^{-4}$	2.1	$10^{-4}$	4	20	$2/3 \times 10^{-4}$	3.4	17.6
Ub (N-C)	$10^5$	2.4	114	8.43	17.15	942	6.09	14.75
Ub (48-C)	6369	6.1	216	12.35	15.90	434	10.59	13.94
RNA	$34 \times 10^{-4}$	97	$10^{-4}$	120	40	$1.1 \times 10^{-4}$	118	31.3
DNA	0.61		0.05		11.9	0.12		10.5

Model parameters for unfolding of titin domain I27 [135], Ub [139], RNA [135] and DNA [285] obtained by fitting data to the phenomenological ( $\nu = 1$ ) and microscopic theory ( $\nu = 1/2$  and  $2/3$ ).

with other independent estimates [138,328,329]. So it seems that the microscopic theories predict reasonable values for unfolding barriers for proteins, RNA and DNA [135,285,330].

## 7.8. Other developments

### 7.8.1. Construction of free energy landscape using Jarzynski's equality

In a pioneering paper [13] Jarzynski has shown that the difference in equilibrium free energy  $\Delta G$  is related to the fluctuations of work performed during a non-equilibrium process  $W_\lambda$  by

$$\langle e^{-\beta W_\lambda} \rangle_M = \int dW_\lambda \rho(W_\lambda) e^{-\beta W_\lambda} = e^{-\beta \Delta G}. \quad (97)$$

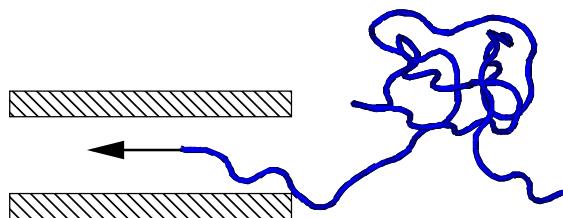
Here  $\langle \dots \rangle_M$  denotes an average over  $N$  realizations of the process and the equality is exact in the limit  $M \rightarrow \infty$ . The work  $W_\lambda = \int F d\lambda$  and its non-equilibrium distribution  $\rho(W_\lambda)$  depends on the schedule for varying work parameter  $\lambda$ . The Jarzynski's equality and a closely related fluctuation theorem by Crooks [331] inspired a lot of work because they allow for the recovery of the free energy landscape of biomolecules from single molecule experiments [332] and simulations [333]. Since most of these works have been reviewed by Ritort [7] and Hummer and Szabo [330], we restrict ourselves to recent developments for proteins. Harris et al. [329] have experimentally reconstructed the free energy surface for the protein I27 using Eq. (97) and the AFM technique. The consistency of their estimate of the unfolding barrier with other results [135], obtained by independent approaches, partially validates the Jarzynski's equality. With the help of this equality and the Go model [92] Paci and coworkers [316] have obtained the free energy landscape for the protein E2lip3. Using an extended form [334] of the Jarzynski equality and the Wako-Saito-Munoz-Eaton model, Imparato et al. [97] computed the free energy landscape of the 39-residue protein PIN1 (PDB ID: 1I6C) as a function of the molecule length. However, they failed to get a reliable estimate for the free energy profile for the longer protein I27 with 89 amino acids. This is probably related to a sampling problem because the average of exponential term in Eq. (97) is dominated by trajectories corresponding to small amounts of work. The longer a protein the less infrequent are these trajectories in the simulations. Currently, practical applications are limited to slow processes where the fluctuations in the amount of work are comparable to the thermal energy [333].

### 7.8.2. AC modulated force spectroscopy

Most of experiments have been performed with a constant force loading rate. The AFM force clamp needs  $\sim 1$  ms feedback to work, while it can provide forces up to nN at a resolution of 5–10 pN, thereby limiting its applicability to the low force regime [301,335] (a laser tweezer would provide more stringent constraints but offers a resolution up to 0.1 pN). This limitation may be overcome by using AC force spectroscopy [336–339] in which the force amplitude and frequency depend on time ( $f(t) = f_0(t) \cos \omega(t)t$ ). This technique can improve the resolution of AFM force spectroscopy by one order of magnitude ( $\approx 0.4$  pN [340]). Schlierf et al. [340] have applied lock-in force spectroscopy to study refolding of the protein ddFLN4 which unfolds [147] and refolds via intermediates. Individual intermediates associated with unfolded I27 titin domains have been revealed by frequency modulation atomic force spectroscopy [341]. The dynamic response of single bacteriorhodopsins has also been studied by a molecular force modulation technique [342]. Braun et al. [343] have shown that periodic loading can offer a more accurate way to probe the free energy landscape of biomolecules compared to other approaches.

### 7.8.3. Unfolding of proteins pulled mechanically through a pore

Although biologically active in a folded state, many proteins, such as those translocated through membranes or those degraded by ATP-dependent proteases must unfold [344,345]. A schematic description of this process is shown in Fig. 49. One important example is the case when precursor proteins, which synthesized and folded in the cytosol, are imported into mitochondria. The precursor proteins usually carry an N-terminal positively charged targeting sequence which is called a presequence. This presequence directs the protein to specific receptors on the mitochondrial outer surface and then through a proteinaceous pore that spans the outer and inner membranes into the innermost mitochondrial compartment



**Fig. 49.** Schematic representation of unfolding of a molecule pulled through a pore. The pore diameter should be much smaller than the molecule length.

called the matrix [346–348]. Because the diameter of the importing pore is smaller than the size of the folded precursor proteins [345,349] (the width of the degradation channel is 10–15 Å at its narrowest point), they must unfold to pass through the pore while unraveling from their N-termini. The unfolding via translocation is a few orders of magnitude faster than chemical/thermal denaturation [344,350–352], suggesting that unfolding in cell may occur via different pathways. Until recently very little was known about the molecular details of this process. One hypothesis is that the cell machinery accomplishes unfolding by mechanically pulling at the end of the polypeptide chain that is labeled for degradation or translocation [198,344,353]. Most insights into the unfolding through a pore mechanism were inferred from single-molecule pulling experiments [6] on polypeptide chains which unfold in free space. However, recent simulations of unfolding of Ub [354], barnase [355] and domain I27 [90] dragged into a pore show significant differences between these two cases. Barnase, a kinetically stable protein, is rapidly imported into a mitochondrion via a pathway different from that found under chemical denaturation [351,355]. The unfolding mechanism depends on the pore diameter, the magnitude of the pulling force and on whether the force is applied at the N- or the C-terminus of the chain. The difference between mechanical unfolding through a pore and in free space is expected because the direction and the geometry of the applied force may dramatically affect the unfolding mechanism [193,194,196].

#### 7.8.4. Crowding effect

So far most of SMFS experiments have been carried out *in vitro* [2,165]. However, *in vivo* folding and unfolding take place in cells which contain different kinds of biomolecules like sugar, nucleic acids, lipids etc. Hence, it is important to take into account these crowders which occupy about 40% of the total volume with steric repulsion among themselves [344,356,357]. The confined environment induces phenomena like “molecular confinement” and “molecular crowding” and has major thermodynamic and kinetic consequences on the cellular processes [358–361]. In the presence of crowders, biomolecules become thermodynamically more stable [360,362]. Consequently, the mechanical stability of proteins gets enhanced [363,364]. The nature of force induced transitions and unfolding pathways may become strongly modified in the presence of molecular crowding [363–367].

#### 7.9. Summary of main factors governing mechanical stability of proteins

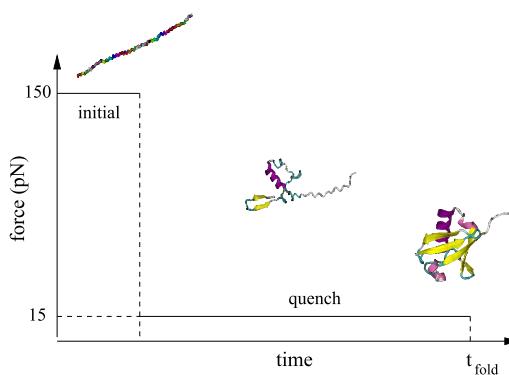
So far about 50 proteins have been studied by SMFS with differing degrees of detail (most of them are shown in Table 1). The unfolding force varies between 15 (Calmodulin) and 1100 pN (Bovine). Although the full molecular basis underlying the mechanical resistance of proteins is still lacking, several main determinants have been identified:

1. The secondary structure plays a crucial role as  $\beta$ -proteins tend to unfold at higher forces than  $\alpha$ -helical ones. As a result  $f_u$  may grow linearly with the CO.
2. In addition to the secondary structure the tertiary structure may influence the mechanical resistance. The 24-domain ankyrin is mechanically more stable than the single- or six-domain ones (see Table 1 and [368]).
3. The mechanical stability depends on the pulling geometry. The points of application of the force to a protein matter, as they can substantially alter its unfolding force. The pulling direction is also relevant. If a force is applied parallel with hydrogen bonds (unzipping), then  $\beta$ -proteins are less stable than the case where the force direction is orthogonal to them (shearing).
4. The mechanical resistance of most proteins tends to be determined by a mechanical pile usually formed by a patch of highly localized hydrogen bonds [95], but in some cases the hydrophobic core gives a contribution to it as well [115].
5. The mechanical stability can be affected by ligand binding [309,369,370] and disulphide bond formation [371–375].
6. Crowders can increase the unfolding force.

Contrary to proteins the structure of DNA is simple and it does not play a role in the mechanical stability. The resistance of DNA to an external force is mainly dictated by the base sequence as well as by the pulling geometry (unzipping or shearing).

#### 8. Refolding under quenched force

An ongoing challenge in molecular biology is to decipher the folding routes of biomolecules so that the underlying energy landscape can be quantitatively mapped. Major advances in theory, simulations, and ensemble experiments have resulted



**Fig. 50.** Schematic plot of the force-clamp experimental setup. Initially a force  $f_i$  of several hundred pN is applied to stretch out a protein. Then the force is reduced to a much lower value  $f_q$  ( $\approx 10$  pN) and it is kept quenched at this value during the folding process. Typical snapshots are shown for different times.

in a better understanding of how proteins and RNA fold. In conventional experiments folding (unfolding) is triggered by decreasing (increasing) the concentration of denaturants or temperature. Under folding conditions biomolecules traverse the energy landscape from high entropy unfolded states to a low entropy native basin of attraction (NBA). Because the initial conformations in the unfolded basin of attraction (UBA) are difficult to characterize, and their fluctuations are averaged out in ensemble experiments, it is difficult to unambiguously decipher the folding mechanisms.

Recently, in an impressive series of experiments, mechanical force  $f$  has been used as a new variable to prepare well defined initial states of proteins [410]. The basic idea of this approach (force-clamp technique) is illustrated in Fig. 50. Initially one applies a high force  $f_i$  of a few hundred pN to generate initial stretched conformations. Then the force is reduced to a lower value  $f_q$  and it is kept quenched at this value during the refolding process. As shown in an important paper by Fernandez and Li [410] the NBA remains preferably populated because  $f_q < f_c^{eq}$ , where  $f_c^{eq}$  is the equilibrium critical force separating the folded region from the unfolded one (see Section 7.1).

In contrast to ensemble averaged experiments, mechanical folding experiments probe at the single molecule level the dynamics from a stretched low entropy state to the low entropy NBA. The refolding under quenched force has a number of advantages compared to the case where mechanical force is not applied. First, one can control the initial conformations by monitoring the end-to-end distance. Second, the quenched force slows down the folding process considerably and this makes a study of the details of chain collapse and folding pathways much easier. Third, measurements of folding times at various forces allows one to obtain the distance between the transition state and the native state  $x_f$ , using the Bell equation (57). In other words, force-clamp spectroscopy can serve as a tool to decipher the free energy landscape of biomolecules. Finally, the direct observation of molecule-dependent collapse offers an unprecedented glimpse not only into the energy landscape of proteins, but also the relationship between folding and collapse of polypeptide chains.

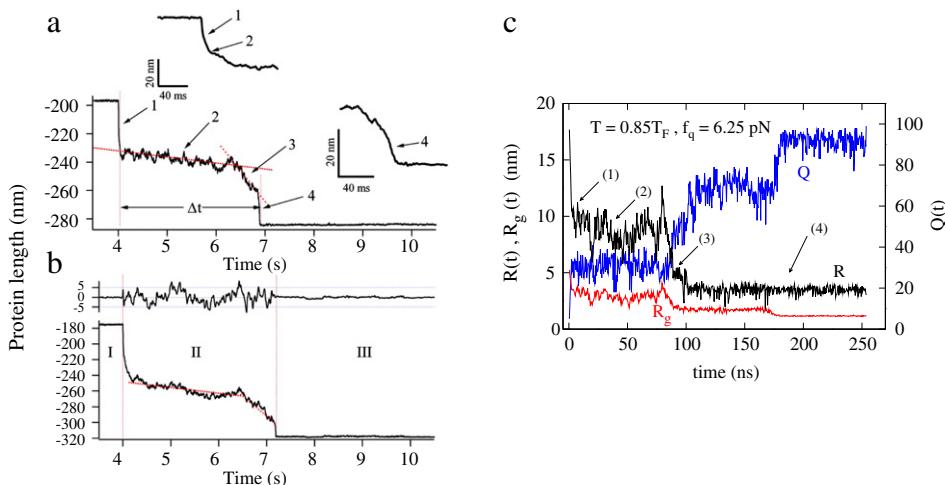
A novel force-clamp method which produces a constant force has been applied to obtain mechanical folding trajectories for tandem constructs of ubiquitin (Ub) [410]. However, the poly-Ub construct makes it difficult to infer the nature of refolding of a monomeric protein. The dynamics of a particular Ub module may be influenced by covalent linkage to other Ub domains. In order to assess these effects and establish the generality of the results several groups have undertaken force-quench simulations of single Ub [94,304,411] and I27 [412] using Go models. Since the simulation results are qualitatively similar for these two proteins, we will mainly focus on Ub to make direct comparisons with the experiments [410].

### 8.1. Proteins refold in stages

One of the most remarkable results of Fernandez and Li [410] is that the refolding from an ensemble of force-induced denatured (FDE) conformations is not an all-or-none process, but rather one sees a continuous collapse in the length of a protein (Fig. 51(a) and (b)). The first stage is fast and one can interpret it as the elastic recoil of an ideal polymer. The next three steps are marked by abrupt changes in the slope and correspond to the folding trajectory of Ub. The individual stages can be distinguished by different slopes. Stages 2 and 3 show fluctuations in the length of several nanometers. The rapid final contraction of stage 4 is the end of the folding event. As seen from the inset of Fig. 51(a) even this stage is not instantaneous.

To monitor force-quench refolding coarse-grained protein models were used, because the refolding simulations of all-atom polypeptide models are computationally prohibitive. Go modeling can capture the step-wise nature of force-induced refolding [94,412]. Typically, to follow the dynamics of the approach to the NBA, one initially prepares stretched structures by applying a constant  $f_i = 100$  pN. A molecule is stretched, if the end-to-end distance  $R \geq 0.85L$ , where  $L = (N - 1)a$  is the contour length. Starting from such a FDE the force is quenched to  $f_q < f_c^{eq}$  to study the refolding kinetics. The obvious advantage of simulations is that one can monitor the time dependence not only of  $R$  but also of  $Q$  and  $R_g$ .

Fig. 51(c) shows a typical trajectory from Go simulations. In qualitative agreement with the experiments [410] the decrease in  $R$  occurs in four stages marked by arrows. The acquisition of the native structure, as measured by  $Q$ , also occurs in steps from the initial value  $Q \approx 0$  until the NBA is reached. The increase in  $Q$  mirrors the reduction in  $R$ , which shows



**Fig. 51.** (a) Four distinct stages can be identified for the refolding of a poly-Ub from a stretched conformation. (b) The folding collapse is marked by large fluctuations in the length of the protein. These fluctuations greatly diminish after folding is complete. Results shown in (a) and (b) are taken from the work of Fernandez and Li [410]. The quench force  $f_q = 15$  pN.

Source: By permission from Ref. [410].

(c) The  $T =$  typical four stages for the total number of native contacts,  $Q$ , the gyration radius,  $R_g$ , and the end-to-end distance,  $R$ , obtained from Go simulations of a single Ub [94].  $T = 0.85T_F$  and  $f_q = 6.25$  pN. Figure has been taken from Ref. [94].

that the compaction of the chain and the acquisition of the native structure are correlated. It is remarkable that even though  $R(t)$  and  $Q(t)$  reflect the stepwise acquisition of structure,  $R_g(t)$  in all cases decreases more continuously, thus masking the presence of any barriers to folding. Overall, refolding trajectories upon force quench show considerable heterogeneity [93,94,412]. Different molecules traverse entirely different routes in the transition from FDE to the NBA.

The qualitative difference between FDE refolding and standard folding commencing from a temperature-denatured ensemble (TDE) may be revealed by monitoring the evolution of the averaged end-to-end distance  $\langle R(t) \rangle$ , gyration radius  $\langle R_g(t) \rangle$ , and the total number of native contacts  $\langle Q(t) \rangle$ , where the angular brackets indicate averages over individual trajectories [139,412]. In the FDE case the evolution of these quantities is described by two time scales  $\tau_{\text{col}}$  and  $\tau_{\text{fold}}$ , where the collapse time  $\tau_{\text{col}} \ll \tau_{\text{fold}}$ . Within the Go model  $\tau_{\text{col}}$  is a few nanoseconds and  $\tau_{\text{fold}}$ , which has the same order of magnitude as the folding time, is about two orders of magnitude larger than  $\tau_{\text{col}}$  [94,412].

After a temperature quench in the TDE the folding transition starts from a high-entropy state and proceeds to a low-entropy NBA (Fig. 52). In contrast, in FDE refolding, the low-entropy NBA is reached from a low-entropy stretched state with the initial rapid kinetics being a FDE  $\rightarrow$  random coil-like state ( $RC_f$ ) transition. If the relaxation of tension along the chain is rapid compared with protein conformation changes then the ensemble of transiently populated coil conformations upon force quench  $RC_f$  must be distinct from the TDE (Fig. 52). These physical arguments suggest that the FDE refolding pathways and time scales must be different from those observed for TDE refolding. It was shown [412] that  $\langle R(t) \rangle$  of TDE folding decreases exponentially with a single characteristic time, while in the FDE case the evolution of  $\langle R(t) \rangle$  is characterized by two time scales. The simulation data [412] show that collapse and folding starting from the TDE conformations are effectively synchronous. In contrast FDE folding becomes decoupled from the collapse upon the change in initial conformations.

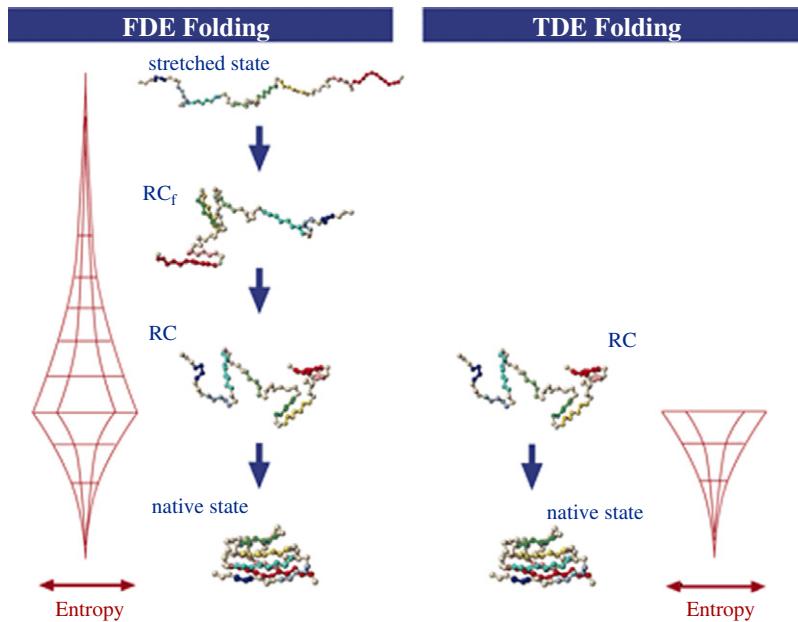
Thus, the folding initiated with the FDE differs from TDE folding in two aspects. First, the stretched polypeptide chain in FDE must initially contract to the ensemble of conformations (under tension) with the dimensions similar to random coil TDE conformations. The rapid initial contraction introduces an additional time scale. Second, FDE results in a delay in the formation of native interactions, because the chain must not only contract but also equilibrate within a random coil ensemble. Therefore, when folding begins from a stretched state the collapse and folding are decoupled. In contrast when TDE is used the collapse and folding can occur in concert for two-state folders.

## 8.2. Deciphering the free energy landscape using force-clamp measurements

In the Bell approximation the dependence of refolding  $\tau_F(f_q)$  on the quenched force  $f_q$  obeys Eq. (57). We rewrite this equation in the following form

$$\tau_F(f_q) \simeq \tau_F(0) \exp \left( \frac{f_q x_f}{k_B T} \right), \quad (98)$$

where  $x_f$  is the distance from the unfolded state to the TS. Thus, force-clamp measurements allow one to obtain this parameter of the free energy landscape. The typical dependence of  $\tau_F(f_q)$  on  $f_q$  for Ub is shown in Fig. 53. Both simulations and experiments support the validity of the Bell approximation, but the refolding in the Go models is about seven orders of



**Fig. 52.** Conceptual differences in FDE and TDE refolding pathways. The mechanically stretched state in the FDE has a low conformational entropy.

magnitude faster than in the experiments. However this simple modeling still gives a reasonable estimate for  $x_f$ . Namely, from Go simulations of Ub (Fig. 53(a)) we obtain  $x_f \approx 0.96 \text{ nm}$  which is close to the experimental estimate of Fernandez and Li [410]  $x_f \approx 0.8 \text{ nm}$  (Fig. 53(b)). Given the simplicity of the model the agreement in  $x_f$  should be considered remarkable. In the case of I27 Go modeling gives [412]  $x_f \approx 0.6 \text{ nm}$ .

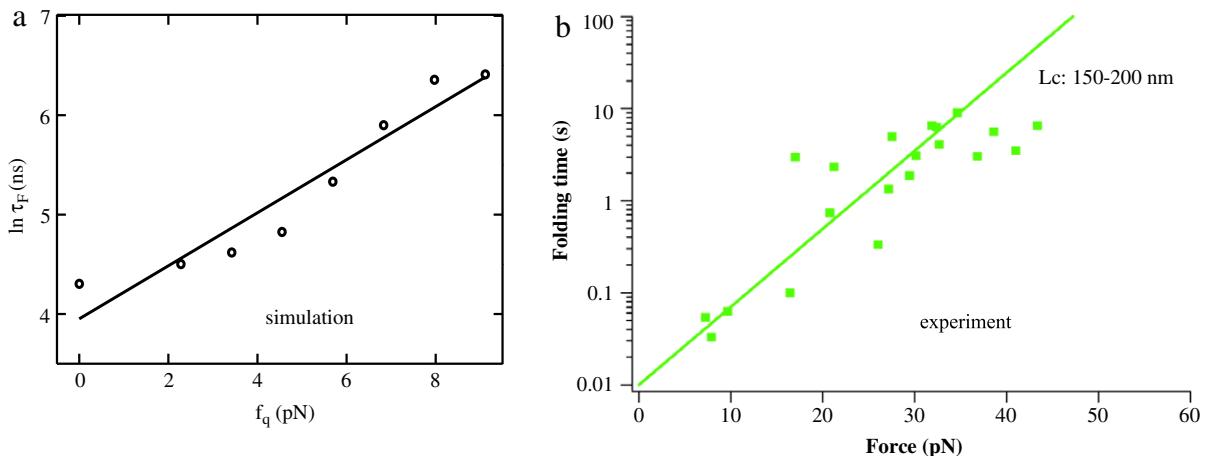
In force quench experiments  $R$  should be treated as a measure of the progress in the stretched-unfolded basin of attraction  $\rightarrow$  native basin of attraction (NBA) transition [410]. The dependence of  $\tau_F$  on  $f_q$  can be written as  $\tau_F \approx \exp(\Delta G_{uf} x_f f_q / x_{uf} k_B T f_c)$ , where  $\Delta G_{uf}$  is the equilibrium free energy of stability of the NBA with respect to the unfolded states and  $x_{uf}$  is the distance between the NBA and the unfolded states. In force-quench experiments the search for the transition ensemble occurs among the ensemble of unfolded states. Within this interpretation a Tanford-like parameter [413] can be written as  $\beta \approx x_f / x_{uf}$ , where  $x_{uf} \approx a(N^{3/5} - N^{1/3})$ . Using the estimates of  $x_f$  from Go simulations we have  $\beta \approx 0.15$  for I27. In the case of Ub Go modeling and experiments give  $\beta \approx 0.27$  and 0.23, respectively. Given the small value of  $\beta$  the native states of the  $\beta$ -sandwich protein I27 and  $\alpha/\beta$  protein Ub may be “brittle”, i.e., once the extension exceeds a small value the structure unravels. A similar conclusion follows from the SMD simulations of  $\beta$ -sandwich proteins [402]. Finally, estimates of the position of the average transition state using the end-to-end distance as a reaction coordinate suggests that for force quench refolding the transition state moves closer to the NBA [94]. This result is consistent with the Hammond’s postulate [140].

Force-quench experiments on poly Ub [410] and the theoretical study of monomeric proteins [94,412] and multi-domain proteins [139] show that the  $f_q$ -dependent refolding time obeys Eq. (98). However, the value of  $\tau_F$  extrapolated to  $f_q$  is larger than that obtained by conventional experiments. Fernandez and Li obtained  $\tau_F(0) \approx 0.01 \text{ s}$  for Ub, whereas the folding time from the ensemble experiments  $\tau_F(0) \approx 3 \text{ ms}$  [327]. The difference between the two experimental values is caused by the vastly different initial conditions from which refolding was initiated. In the transition from the stretched state to the NBA a large free energy cost is involved in burying a substantial surface area  $\Delta A \approx (R_l^2 - a^2 N^{2/3})$ . In contrast the NBA is accessed from TDE by much smaller conformational fluctuations. The discrepancy between  $\tau_F^{FDE}$  and  $\tau_F^{TDE}$  should decrease, if  $f_l$  does not fully stretch the polymer chain [412].

In the Fernandez and Li experiments [410]  $f_l$  was fixed ( $f_l = 150 \text{ pN}$ ) and  $f_q$  was varied. This affects the free energies of the folded and unfolded states. The linear increase in the free energy barrier to folding, which is consistent with the Bell-type model of Eq. (98)), is caused by the stabilization of the  $RC_f$  states with increasing  $f_q$ . The presence of a free energy barrier to folding is consistent with theoretical predictions [166,414] and simulation results [81] showing that equilibrium-force unfolding is a first order transition. On the other hand, if  $f_q$  is fixed and  $f_l$  is varied, then we expect  $\tau_F \approx \exp(\gamma R_l^2)$ , where  $\gamma$  is an effective surface tension [412]. This finding suggests that the dependence of  $\tau_F$  on  $f_l$  should be different from the dependence of  $\tau_F$  on  $f_q$ . Thus, force-quench refolding with fixed  $f_l$  and varying  $f_q$  is qualitatively different from folding with  $f_q$  fixed and differing  $f_l$  values. This inherent asymmetry can be exploited to map the free energy landscape of proteins.

### 8.3. Can force-clamp spectroscopy probe folding pathways of a protein?

Since one terminal of a protein is anchored in the force-clamp technique, it is not clear if it probes the same folding pathways as the free-end case. Fig. 54 shows [139] the dependence the fraction of native contacts of different structures of



**Fig. 53.** The dependence of refolding times on the quench force for Ub. (a) The simulation results at  $T = 0.85T_F$  from [94]. (b) The experimental results from Fernandez and Li [410].

Source: By permission from Ref. [410].

Ub on the progressive variable  $\delta$  for three scenarios (free end, N-fixed and C-fixed). Clearly, for the free-end and C-fixed case the refolding from stretched conformations follows the same pathways. We have the following dominant pathway

$$S2 \rightarrow S4 \rightarrow A \rightarrow S1 \rightarrow (S3, S5). \quad (99)$$

As evident from Fig. 54 another pathway  $S2 \rightarrow S4 \rightarrow A \rightarrow S3 \rightarrow S1 \rightarrow S5$  also has a high probability. These two pathways differ only in the sequencing of  $S1$  and  $S3$ . One can check the validity of the folding sequencing given by Eq. (99) with available experimental results on  $\phi$ -values [415] and  $\psi$ -values [416,417]. In addition, if one assumes that unfolding is the reverse of the refolding process (it is true for Go models [94]) then one can infer information about the folding pathways from experimental results on thermal unfolding of Cordier and Grzesiek [306] and by Chung et al. [418]. A detailed comparison with these experimental results as well as with all-atom simulations [419,420] and simulations using different coarse-grained models [83,421–424] have been carried out recently [94,425]. Overall, there is partial agreement between the sequencing given by Eq. (99) and experimental results.

However, if the N-terminus is fixed, then the pathways change notably. The probability of observing the main pathway

$$S4 \rightarrow S2 \rightarrow A \rightarrow S1 \rightarrow (S3, S5) \quad (100)$$

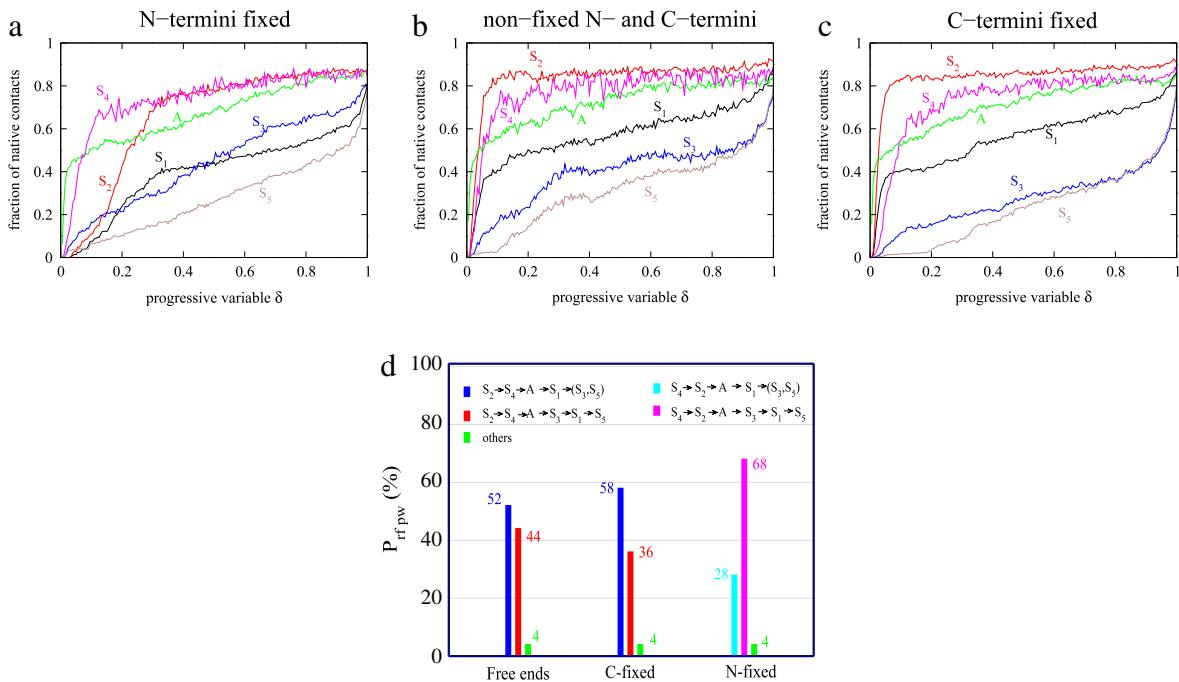
is very high (68%) compared to others. Thus, the force-clamp technique may give different pathways for Ub if the N-terminus is anchored [139,411].

Using the Go model [87] we have shown that fixing one end of the three-domain Ub does not affect its refolding pathways [139]. To understand this interesting result we recall that applying a low quenched force to both termini does not change the folding pathways of single Ub [94]. So in the three-domain case with the N-end of the first domain fixed, both termini of the first and second domains might be considered as effectively subjected to an external force and their pathways should remain the same as in the free-end case. The N-terminal of the third domain is tethered to the second domain but this should have a much weaker effect compared to the case where it is anchored to a surface. Thus this unit has almost free ends and its pathways remain unchanged.

So one can predict that force-clamp spectroscopy can probe the folding pathways of a free end Ub if one deals with either a multi-domain construction or a single unit with the C-terminal fixed. This conclusion follows from Go simulations [139]. It would be interesting to check if it remains valid when non-native interactions are taken into account. Using Go modeling we have also examined the effect of anchoring one terminal on the refolding pathways of the  $\beta$ -domain I27 [139]. It turns out that, contrary to the Ub case, the folding sequencing is not sensitive to which end is fixed. If this is true then the force-clamp technique always probes the folding pathways for I27. The results, obtained for Ub and I27, suggest that the effect of fixing one end on refolding pathways is not universal for all proteins.

## 9. What next?

We are not going to speculate what will be the trend of SMFS experiments in the next decade, but we are confident that the next decade will witness further advancements in our understanding of various molecular interactions and their contributions to biological processes. The progress will come from the interplay between theories, experiments and numerical simulations. In this process we expect that there will very soon be more experiments to verify the predictions based on model studies or theoretical models to explain the experimental observations.



**Fig. 54.** (a) The dependence of native contacts of  $\beta$ -strands and helix A of single Ub on the progressive variable  $\delta$  when the N-terminal is fixed. (b) As in (a) but both ends are free. (c) As in (a) but C-terminal is anchored. The results are averaged over 200 trajectories. (d) The probability of observing different pathways is shown above histograms. Figure has been taken from Ref. [139].

Among these the issue related to the re-entrance of the force-temperature phase-diagram is of a particular interest. Although for more than a decade exact results supported by extensive numerical calculations have confirmed this phenomenon it has not yet been observed in any experiment. A recent proposal by Kumar [426] shows that a Molecular Beacon in formamide may be an ideal candidate which will exhibit reentrance *in vitro*.

Early experiments on stretching of ssDNA revealed an entropic response which can be described by the FJC model. Recently there has been considerable interest in pulling ssDNA consisting of only one type of nucleotide (poly(A) or poly(T)). It was found that poly(T) shows an entropic response while poly(A) shows two plateaus in the force-extension curve. These two plateaus have been related to base stacking associated with structural transitions in poly(A). It appears that there is no better understanding and additional NMR and quantum calculation are required to understand the structural transitions in ssDNA which is solely induced by a force.

Understanding the role of bubbles in the transcription process is another area where SMFS has a vital role to play. So far force has been applied only at the end of the chains. The phase diagram obtained by Bhattacharjee and others resulted in important information about the “existence of an Eye-phase”. In this context the role of the ssb (single-strand binding) protein becomes important which may apply force in the interior of the chain. One of the ways to elucidate such a problem is to apply random forces at different sites and study the phase diagram.

Understanding of the structure and function of proteins *in vivo* by analyzing it *in vitro* is one approach [2, 165], but another route is to perform the analysis / experiments in presence of the environment similar to *in vivo* [344, 356]. Interestingly in most of the SMFS efforts (experimental, theoretical and numerical simulation) so far, the effect of the cellular environment has been ignored. A real challenge will be to perform further experiments *in vivo* and particularly in cells. This will permit us to understand the functioning of proteins and their relation to diseases.

Recently, it was found that the folding of two-state proteins is surprisingly simple: their folding rates are just determined by the contact order [427]. Whether this is true for unfolding rates remains to be elucidated. We have discussed the application of SMFS to decipher the free energy landscape of biomolecules that unfold in the two-state manner. However, many proteins like I27 [292] and DDFLN4 [294, 428] mechanically unfold via intermediates. It would be desirable to study the free energy landscape to locate intermediate state for this class of biomolecule [295].

It is worth mentioning some important topics that have not been discussed in this review. Mechanical force has been shown to accelerate rearrangement reactions and bias reaction pathways to yield products not obtainable from purely thermal or light-induced reactions [429]. Therefore, by using SMFS one can control chemical reactions in which reactants are sensitive to mechanical perturbation. Recent studies [430] showed that the response of living cells to stretching is governed by simple phenomenological universal laws, despite the fact that the effects of biochemistry, molecular crowding and physical forces are complex and inseparable. More simulations have to be performed to support this fascinating finding.

It is now well established that many protein sequences have a propensity to assemble into amyloid fibrils under the appropriate external conditions [431]. Solid state NMR experiments have revealed a parallel in-registry arrangement

of Alzheimer's A $\beta$  peptides in amyloid fibrils [432]. From the perspective of biology and biotechnology an important characteristic of amyloid assemblies is their remarkable stability against denaturation. As a result their formation is essentially irreversible under physiological conditions. It is important to investigate the stability of amyloid fibril against a variety of external factors. The first AFM experimental [433–439] and theoretical [440] studies probing their mechanical properties have already been reported. It was found that A $\beta$  fibrils are mechanically stronger than most protein domains. Their mechanical dissociation is highly anisotropic and proceeds via different pathways when the force is applied either in a direction parallel with or perpendicular to the fibril axis. Since understanding of structural diseases is very important the study of mechanical stability of amyloid fibrils is expected to be an active research area in the near future. Another interesting research area is to perform exhaustive SMD study of stability of receptor-ligand complex by pulling a ligand. This would help to screen out leads in the drug design problem.

## 10. Conclusion

In this review we have concentrated on recent experimental and theoretical work on the mechanical properties and structural transitions of biopolymers. In order to do so we briefly reviewed some of the models and methods which have been used in describing results from SMFS. From these works we have tried to show the importance of molecular interactions in the stability of biomolecules in terms of structure and function. We expect future research in this direction will elucidate the relationship between the function and structure of biomolecules. We have also highlighted some of the problems which may be of interest for the single molecule community.

## Acknowledgements

We have benefited greatly from discussions with numerous collaborators and colleagues. We would particularly like to thank D. Giri, S. M. Bhattacharjee, I. Jensen, N. Singh, P. Mishra, A. R. Singh, G. Mishra, C. K. Hu, D. K. Klimov, M. Kouza, M. Rief, and D. Thirumalai. Financial supports from the Department of Science and Technology, India and the Ministry of Science and Informatics in Poland (grant No 202-204-234) are greatly acknowledged. We thank all the publishers for kind permission to use their figures in this review.

## Appendix A. List of abbreviations

A	Adenine
AFM	Atomic Force Microscopy
BB	Backbone
CO	Contact order
C	Cytosine
DDFLN4	Fourth domain of <i>Dictyostelium discoideum</i> filamin
DS	Denatured State
DW	Directed walk
dsDNA	Double stranded DNA
FDE	Force denatured ensemble
FEL	Free Energy Landscape
FJC	Freely Jointed Chain
FDW	Fully directed walk
G	Guanine
HBs	Hydrogen bonds
IS	Intermediate State
LOT	Laser optical tweezer
MD	Molecular Dynamics
NBA	Native basin of attraction
NS	Native State
PDSAW	Partial directed walk
PS	Poland–Scheraga
SAW	Self-avoiding walk
SASAW	Self-attracting self-avoiding walk
SMD	Steered Molecular Dynamics
SMFS	Single Molecular Force Spectroscopy
ssDNA	Single stranded DNA
T	Thymine
TDE	Thermal denatured ensemble
T-f	Temperature-force

TS	Transition State
Ub	Ubiquitin
WLC	Worm Like Chain

## Appendix B. List of symbols

$L$	Contour length
$\xi$	Correlation length
$\alpha, \beta, \gamma, \eta, \nu$	Critical exponents
$T_c$	Critical temperature
$x_f$	Distance between transition state and denatured
$x_u$	Distance between native state and transition state
$\Delta R$	End-to-end extension
$\epsilon_H$	Hydrogen bond energy
$b$	Kuhn length
$N$	Number of monomers (beads)
$Z$	Partition function
$L_p$	Persistence length
$C$	Specific heat
$\chi$	Susceptibility
$f_u$	Unfolding force

## References

- [1] F. Crick, Central dogma of molecular biology, *Nature* 227 (1970) 561.
- [2] M. Rief, M. Gautel, F. Oesterhelt, J.M. Fernandez, H.E. Gaub, Reversible unfolding of individual titin immunoglobulin domains by AFM, *Science* 276 (1997) 1109–1112.
- [3] L. Tskhovrebova, K. Trinick, J.A. Sleep, R.M. Simons, Elasticity and unfolding of single molecules of the giant muscle protein titin, *Nature* 387 (1997) 308–312.
- [4] T.R. Strick, J.F. Allemand, D. Bensimon, V. Croquette, Behavior of supercoiled DNA, *Biophys. J.* 74 (1998) 2016–2028.
- [5] R. Lavery, A. Lebrun, J.F. Allemand, D. Bensimon, V. Croquette, Structure and mechanics of single biomolecules: Experiment and simulation, *J. Phys.: Condens. Matter.* 14 (2002) R383–R414.
- [6] C. Bustamante, Y.R. Chemla, N.R. Forde, D. Izhaky, Mechanical processes in biochemistry, *Annu. Rev. Biochem.* 73 (2004) 705–748.
- [7] F. Ritort, Single-molecule experiments in biological physics: Methods and applications, *J. Phys.: Condens. Matter.* 18 (2006) R531–R583.
- [8] C. Hyeon, D. Thirumalai, Measuring the energy landscape roughness and the transition state location of biomolecules using single molecule mechanical unfolding experiments, *J. Phys.: Condens. Matter.* 19 (2007) 113101.
- [9] J.I. Sulkowska, M. Cieplak, Mechanical stretching of proteins – A theoretical survey of the protein data bank, *J. Phys.: Condens. Matter.* 19 (2007) 283201.
- [10] A. Borgia, P.M. Williams, J. Clarke, Single-molecule studies of protein folding, *Annual Rev. Biochem.* 77 (2008) 101–125.
- [11] J.R. Moffitt, Y.R. Chemla, S.B. Smith, C. Bustamante, Recent advances in optical tweezers, *Annual Rev. Biochem.* 77 (2008) 205–228.
- [12] P.T.X. Li, J. Vieregg, I.T. Jr., How RNA unfolds and refolds, *Annual Rev. Biochem.* 77 (2008) 77–100.
- [13] C. Jarzynski, Non-equilibrium equality for free energy differences, *Phys. Rev. Lett.* 78 (1997) 2690–2693.
- [14] P.G. de Gennes, *Scaling Concepts in Polymer Physics*, Cornell University Press, Ithaca, London, 1979.
- [15] M. Doi, S.F. Edwards, *The Theory of Polymer Dynamics*, Oxford University Press, Oxford, 1986.
- [16] C. Vanderzande, *Lattice Models of Polymers*, Cambridge University Press, Cambridge, England, 1998.
- [17] S.B. Smith, L. Finzi, C. Bustamante, Direct mechanical measurements of the elasticity of single DNA molecules by using magnetic beads, *Science* 258 (1992) 1122–1126.
- [18] B.C. vander Aa, R.M. Michel, M. Asther, M.T. Zamora, P.G. Rouxhet, Y.F. Dufrêne, Stretching cell surface macromolecules by atomic force microscopy, *Langmuir* 17 (2001) 3116–3119.
- [19] A. Vologodskii, DNA extension under the action of an external force, *Macromolecules* 27 (1994) 5623–5625.
- [20] C. Ortiz, G. Hadzioannou, Entropic elasticity of single polymer chains of poly(methacrylic acid) measured by atomic force microscopy, *Macromolecules* 32 (1999) 780–787.
- [21] W.K. Zhang, S. Zou, C. Wang, X. Zhang, Single polymer chain elongation of poly(*n*-isopropylacrylamide) and poly(acrylamide) by atomic force microscopy, *J. Phys. Chem. B* 104 (2000) 10258–10264.
- [22] S.B. Smith, Y. Cui, C. Bustamante, Overstretching B-DNA: The elastic response of individual double-stranded and single-stranded DNA molecules, *Science* 271 (1996) 795–799.
- [23] J. Marko, E. Siggia, Stretching DNA, *Macromolecules* 28 (1995) 8759–8770.
- [24] A. Rosa, T.X. Hoang, D. Marenduzzo, A. Maritan, Elasticity of semiflexible polymers with and without self-interactions, *Macromolecules* 36 (2003) 10095–10102.
- [25] P.J. Flory, *Statistical Mechanics of Chain Molecules*, Hanser, Munich, 1989.
- [26] J.D. Cloizeaux, Langrangian theory for a self-avoiding random chain, *Phys. Rev. A* 10 (1974) 1665–1669.
- [27] Y. Singh, Molecular theory of solvent induced self-interactions in a polymer, *J. Phys. A* 20 (1987) 3949–3954.
- [28] S. Kumar, Y. Singh, D. Dhar, Surface adsorption of a self avoiding polymer chain on a family of finitely ramified fractals, *J. Phys. A: Math. Gen.* 26 (1993) 4835.
- [29] S. Kumar, Y. Singh, Collapse transition of linear polymers on a family of truncated *n*-simplex lattices, *Phys. Rev. A* 42 (1990) 7151–7154.
- [30] S. Kumar, Y. Singh, Critical behaviour of 2 interacting linear polymer chains: Exact results for a state of interpenetration of chains on a fractal lattice, *J. Phys. A: Math. Gen.* 26 (1993) L987.
- [31] D. Dhar, J. Vannimeten, The collapse transition of linear-polymers on fractal lattices, *J. Phys. A: Math. Gen.* 20 (1987) 199–213.

- [32] B. Duplantier, H. Saleur, Exact critical properties of two-dimensional dense self-avoiding walks, *Nuclear Phys. B* 290 (1987) 291–326.
- [33] V. Privman, N.M. Svarkic, *Directed Models of Polymers, Interfaces, and Clusters*, Springer-Verlag, Berlin, 1989.
- [34] V.A. Bloomfield, D.M. Crothers, I. Tinoco Jr., *Nucleic Acids: Structure, Properties, and Functions*, University Science Books, Sausalito, California, 2000.
- [35] D. Poland, H.A. Scheraga, *Theory of Helix-coil Transitions in Biopolymers: Statistical Mechanical Theory of Order and Disorder Transitions in Biological Macromolecules*, Academic Press, New York, 1970.
- [36] R.M. Wartell, A.S. Benight, Thermal-denaturation of DNA molecules – A comparison of theory with experiment, *Phys. Rep.* 126 (1985) 67–107.
- [37] Y. Zeng, A. Montrichok, G. Zocchi, Bubble nucleation and cooperativity in DNA melting, *J. Mol. Biol.* 339 (2004) 67–75.
- [38] A. Montrichok, G. Gruner, Zocchi, Trapping intermediates in the melting transition of DNA oligomers, *Europhys. Lett.* 62 (2003) 425–458.
- [39] Y. Zeng, A. Montrichok, G. Zocchi, Length and statistical weight of bubbles in DNA melting, *Phys. Rev. Lett.* 91 (2003) 148101.
- [40] D. Poland, H.A. Scheraga, Phase transitions in 1 dimension and helix-coil transition in polyamino acids, *J. Chem. Phys.* 45 (1966) 1456.
- [41] M.E. Fisher, Effect of excluded volume on phase transitions in biopolymers, *J. Chem. Phys.* 45 (1966) 1469.
- [42] Y. Kafri, D. Mukamel, L. Peliti, Why is the DNA denaturation transition first order? *Phys. Rev. Lett.* 85 (2000) 4988–4991.
- [43] Y. Kafri, D. Mukamel, L. Peliti, Melting and unzipping of DNA, *Eur. Phys. J. B* 27 (2002) 135–146.
- [44] A. Bar, Y. Kafri, D. Mukamel, Loop dynamics in DNA denaturation, *Phys. Rev. Lett.* 98 (2007) 038103.
- [45] Y. Kafri, D. Mukamel, L. Peliti, Denaturation and unzipping of DNA: Statistical mechanics of interacting loops, *Physica A* 306 (2002) 39–50.
- [46] S.M. Bhattacharjee, Unzipping DNAs: Towards the first step of replication, *J. Phys. A* 33 (2000) L423–L428.
- [47] M. Peyrard, A.R. Bishop, Statistical mechanics of a nonlinear model for DNA denaturation, *Phys. Rev. Lett.* 62 (1989) 2755–2758.
- [48] T. Dauxois, M. Peyrard, A. Bishop, Entropy-driven DNA denaturation, *Phys. Rev. E* 47 (1993) R44–R47.
- [49] P.M. Morse, Diatomic molecules according to the wave mechanics. ii. vibrational levels, *Phys. Rev.* 34 (1929) 57.
- [50] M. Barbi, S. Cocco, M. Peyrard, Helicoidal model for DNA opening, *Phys. Lett. A* 253 (1999) 358.
- [51] D.K. Lubensky, D.R. Nelson, Pulling pinned polymers and unzipping DNA, *Phys. Rev. Lett.* 85 (2000) 1572–1575.
- [52] D.K. Lubensky, D.R. Nelson, Single molecule statistics and the polynucleotide unzipping transition, *Phys. Rev. E* 65 (2002) 031917.
- [53] S.M. Bhattacharjee, S. Mukherjee, Directed polymers with random interaction: Marginal relevance and novel criticality, *Phys. Rev. Lett.* 70 (1993) 49.
- [54] S. Mukherjee, S.M. Bhattacharjee, Directed polymers with random interaction—An exactly solvable case, *Phys. Rev. E* 48 (1993) 3483–3496.
- [55] M.S. Causo, B. Coluzzi, P. Grassberger, Simple model for the DNA denaturation transition, *Phys. Rev. E* 62 (2000) 3958–3973.
- [56] E. Carlon, E. Orlandini, A. Stella, Roles of stiffness and excluded volume in DNA denaturation, *Phys. Rev. Lett.* 88 (2002) 198101.
- [57] D. Marenduzzo, A. Trovato, A. Maritan, Phase diagram of force-induced DNA unzipping in exactly solvable models, *Phys. Rev. E* 64 (2001) 031901.
- [58] D. Marenduzzo, S.M. Bhattacharjee, A. Maritan, E. Orlandini, F. Seno, Dynamical scaling of the DNA unzipping transition, *Phys. Rev. Lett.* 88 (2001) 028102.
- [59] S. Kumar, D. Giri, S.M. Bhattacharjee, Force induced triple point for interacting polymers, *Phys. Rev. E* 71 (2005) 051804.
- [60] S. Kumar, D. Giri, Probability distribution analysis of force induced unzipping of DNA, *J. Chem. Phys.* 125 (2006) 044905.
- [61] S. Kumar, D. Giri, Y. Singh, Statistical mechanics of coil-hairpin transition in a single-stranded DNA oligomer, *Europhys. Lett.* 70 (2005) 15–21.
- [62] D. Giri, S. Kumar, Effects of the eye phase in DNA unzipping, *Phys. Rev. E* 73 (2006) 050903(R).
- [63] C.B. Anfinsen, Principles that govern the folding of protein chains, *Science* 181 (1973) 223–230.
- [64] C. Levinthal, Are there pathways for protein folding, *J. Chim. Physique* 65 (1968) 44–45.
- [65] P. Leopold, M. Montal, J. Onuchic, Protein folding funnels: A kinetic approach to the sequence–structure relationship, *Proc. Natl. Acad. Sci. USA* 89 (1992) 8721–8725.
- [66] K.A. Dill, S. Bromberg, K.Z. Yue, K.M. Fiebig, D.P. Yee, P.D. Thomas, H.S. Chan, Principles of protein-folding – A perspective from simple exact models, *Protein Sci.* 4 (1995) 561–602.
- [67] A. Sali, E. Shakhnovich, M. Karplus, Kinetics of protein folding – A lattice model study of the requirements for folding to the native state, *J. Mol. Biol.* 235 (1994) 1614–1636.
- [68] N. Go, Theoretical studies of protein folding, *Ann. Rev. Biophys. Bioeng.* 12 (1983) 183–210.
- [69] S. Miyazawa, R.L. Jernigan, Estimation of effective interresidue contact energies from protein crystal structures: Quasi-chemical approximation, *Macromolecules* 18 (1985) 534–562.
- [70] A. Kolinski, A. Godzik, J. Skolnick, A general method for the prediction of the three dimensional structure and folding pathway of globular proteins: Application to designed helical proteins, *J. Chem. Phys.* 98 (1993) 7420–7433.
- [71] M.R. Betancourt, D. Thirumalai, Pair potentials for protein folding: Choice of reference states and sensitivity of predicted native states to variations in the interaction schemes, *Protein Sci.* 8 (1999) 361–369.
- [72] F.M. Richards, Areas, volumes, packing, and protein structure, *Ann. Rev. Biophys. Bioeng.* 6 (1977) 151–176.
- [73] F.M. Richards, W.A. Lim, An analysis of packing in the protein folding problem, *Q. Rev. Biophys.* 26 (1993) 423–498.
- [74] A. Kolinski, W. Galazka, J. Skolnick, On the origin of the cooperativity of protein folding: Implications from model simulations, *Proteins: Struct. Funct. Gen.* 26 (1996) 271–287.
- [75] S. Bromberg, K.A. Dill, Side chain entropy and packing in proteins, *Protein Sci.* 3 (1994) 997–1009.
- [76] D.K. Klimov, D. Thirumalai, Cooperativity in protein folding: From lattice models with side chains to real proteins, *Fold. Design* 3 (1998) 127–139.
- [77] H.J. Hilhorst, J.M. Deutch, Analysis of Monte Carlo results on the kinetics of lattice polymer chains with excluded volume, *J. Chem. Phys.* 63 (1975) 5153–5161.
- [78] M.R. Betancourt, Smoothing the landscapes of protein folding: Insights from a minimal model, *J. Chem. Phys.* 109 (1998) 1545–1554.
- [79] M.S. Li, D.K. Klimov, D. Thirumalai, Dependence of folding rates on protein length, *J. Phys. Chem. B* 106 (2002) 8302–8305.
- [80] N.D. Soccia, J.N. Onuchic, P.G. Wolynes, Stretching lattice models of protein folding, *Proc. Natl. Acad. Sci. USA* 96 (1999) 2031–2035.
- [81] D.K. Klimov, D. Thirumalai, Stretching single-domain proteins: Phase diagram and kinetics of force-induced unfolding, *Proc. Natl. Acad. Sci. USA* 96 (1999) 6166–6171.
- [82] T. Veitshans, D.K. Klimov, D. Thirumalai, Protein folding kinetics: Timescales, pathways and energy landscapes in terms of sequence-dependent properties, *Fold. Design* 2 (1997) 1–22.
- [83] J.M. Sorenson, T. Head-Gordon, Toward minimalist models of larger proteins: A ubiquitin-like protein, *Proteins: Struc. Funct. Bioinform.* 46 (2002) 368–379.
- [84] D.K. Klimov, D. Thirumalai, Native topology determines force-induced unfolding pathways in globular proteins, *Proc. Natl. Acad. Sci. USA* 97 (2000) 7254–7259.
- [85] S. Kirmizialtin, L. Huang, D.E. Makarov, Topography of the free energy landscape probed via mechanical unfolding of proteins, *J. Chem. Phys.* 122 (2005) 234915–234926.
- [86] S. Takada, Gō-ing for the prediction of protein folding mechanisms, *Proc. Natl. Acad. Sci. USA* 96 (1999) 11698–11700.
- [87] C. Clementi, H. Nymeyer, J.N. Onuchic, Topological and energetic factors: What determines the structural detail of the transition state ensemble and en-route intermediates for protein folding? An investigation for small globular proteins, *J. Mol. Biol.* 298 (2000) 937–953.
- [88] N. Koga, S. Takada, Roles of native topology and chain-length scaling in protein folding: A simulation study with a Go-like model, *J. Mol. Biol.* 313 (2001) 171–180.
- [89] J.N. Onuchic, P.G. Wolynes, Theory of protein folding, *Curr. Opin. Struct. Biol.* 14 (2004) 70–75.
- [90] D.K. West, D.J. Brockwell, P.D. Olmsted, S.E. Radford, E. Paci, Mechanical resistance of proteins explained using simple molecular models, *Biophys. J.* 90 (2006) 287–297.
- [91] M. Cieplak, T.X. Hoang, M.O. Robbins, Thermal folding and mechanical unfolding pathways of protein secondary structures, *Proteins: Struct. Funct. Bioinform.* 49 (2002) 104–113.
- [92] J. Karanicolas, C.L. Brooks, The origins of asymmetry in the folding transition states of protein L and protein G, *Protein Sci.* 11 (2002) 2351–2361.

- [93] C. Hyeon, R.I. Dima, D. Thirumalai, Pathways and kinetic barriers in mechanical unfolding and refolding of RNA and proteins, *Structure* 14 (2006) 1633–1645.
- [94] M.S. Li, M. Kouza, C.K. Hu, Refolding upon force quench and pathways of mechanical and thermal unfolding of ubiquitin, *Biophys. J.* 92 (2007) 547–567.
- [95] H. Lu, B. Isralewitz, A. Krammer, V. Vogel, K. Schulten, Unfolding of titin immunoglobulin domains by steered molecular dynamics simulation, *Biophys. J.* 75 (1998) 662–671.
- [96] P. Bruscolini, A. Pelizzola, Exact solution of the munoz-eaton model for protein folding, *Phys. Rev. Lett.* 88 (2002) 258101–258104.
- [97] A. Imparato, A. Pelizzola, M. Zamparo, Ising-like model for protein mechanical unfolding, *Phys. Rev. Lett.* 98 (2007) 148102–148105.
- [98] A. Imparato, A. Pelizzola, M. Zamparo, Protein mechanical unfolding: A model with binary variables, *J. Chem. Phys.* 127 (2007) 145105.
- [99] H. Wako, N. Saito, Statistical mechanical theory of the protein conformation. I. General considerations and the application to homopolymers, *J. Phys. Soc. Japan* 44 (1978) 1931–1938.
- [100] V. Munoz, W. Eaton, A simple model for calculating the kinetics of protein folding from three-dimensional structures, *Proc. Natl. Acad. Sci. USA* 96 (1999) 11311–11316.
- [101] B.R. Brooks, R.E. Brucoleri, B.D. Olafson, D.J. States, S. Swaminathan, M. Karplus, A program for macromolecular energy, minimization, and dynamics calculations, *J. Comp. Chem.* 4 (1983) 187–217.
- [102] W.L. Jorgenson, J. Chandrasekhar, J.D. Madura, R.W. Impey, M.L. Klein, Comparison of simple potential functions for simulating liquid water, *J. Chem. Phys.* 79 (1983) 926–935.
- [103] H. Grubmuller, B. Heymann, P. Tavan, Ligand binding: Molecular mechanics calculation of the streptavidin–biotin rupture force, *Science* 271 (1996) 997–999.
- [104] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skele, L. Kale, K. Schulten, Scalable molecular dynamics with NAMD, *J. Comp. Chem.* 26 (2005) 1781–1802.
- [105] B. Isralewitz, M. Gao, K. Schulten, Steered molecular dynamics and mechanical functions of proteins, *Curr. Opin. Struct. Biol.* 11 (2001) 224–230.
- [106] M. Gao, M. Sotomayor, E. Villa, E.H. Lee, K. Schulten, Molecular mechanisms of cellular mechanics, *Phys. Chem. Chem. Phys.* 8 (2006) 3692–3706.
- [107] M. Sotomayor, K. Schulten, Single-molecule experiments in vitro and in silico, *Science* 316 (2007) 1144–1148.
- [108] P.K. Weiner, P.A. Kollman, AMBER: Assisted model building with energy refinement. A general program for modeling molecules and their interactions, *J. Comp. Chem.* 2 (1981) 287–303.
- [109] W.F. van Gunsteren, S.R. Billeter, A.A. Eising, P.H. Hünenberger, P. Krüger, A.E. Mark, W.R.P. Scott, I.G. Tironi, Biomolecular Simulation: The Gromos96 Manual and User Guide, vdf Hochschulverlag AG an der ETH, Zurich, 1996.
- [110] A.D. MacKerell, J.B. Brooks, C.L.B. III, L.N.B. Roux, Y. Won, M. Karplus, in: R. Schleyer, et al. (Eds.), *The Encyclopedia of Computational Chemistry*, John Wiley and Sons, Chichester, 1998.
- [111] E. Paci, M. Karplus, Forced unfolding of fibronectin type 3 modules: An analysis by biased molecular dynamics simulations., *J. Mol. Biol.* 288 (1999) 441–459.
- [112] E. Paci, M. Karplus, Unfolding proteins by external forces and temperature: The importance of topology and energetics, *Proc. Natl. Acad. Sci. USA* 97 (2000) 6521–6526.
- [113] S. Fowler, R.R. Best, J.L. Toca-Herrera, T. Rutherford, A. Steward, E. Paci, M. Karplus, J. Clarke, Mechanical unfolding of a titin domain Ig: Structure of unfolding intermediate revealed by combining AFM, molecular dynamics simulations, NMR and protein engineering, *J. Mol. Biol.* 322 (2002) 841–849.
- [114] R. Best, S. Fowler, J. Herrera, A. Steward, E. Paci, J. Clarke, Mechanical unfolding of a titin Ig domain: Structure of transition state revealed by combining atomic force microscopy, protein engineering and molecular dynamics simulations, *J. Mol. Biol.* 330 (2003) 867–877.
- [115] S. Ng, R. Rounsevell, A. Steward, C. Geierhass, P. Williams, E. Paci, J. Clarke, Mechanical unfolding of TNfn3: The unfolding pathway of a fnIII domain probed by protein engineering, AFM and MD simulations, *J. Mol. Biol.* 350 (2005) 776–789.
- [116] A. Irback, B. Samuelsson, F. Sjunnesson, S. Wallin, Thermodynamics of  $\alpha$ - and  $\beta$ -structure formation in proteins, *Biophys. J.* 85 (2003) 1466–1473.
- [117] A. Irback, S. Mittetnacht, S. Mohanty, Dissecting the mechanical unfolding of ubiquitin, *Proc. Natl. Acad. Sci. USA* 102 (2005) 13427–13432.
- [118] P.E. Marszalek, H. Lu, H. Li, M. Carrion-Vazquez, A.F. Oberhauser, K. Schulten, J.M. Fernandez, Mechanical unfolding intermediates in titin modules, *Nature* 402 (1999) 100–103.
- [119] G. Forgacs, R. Lipowsky, T.M. Nieuwenhuizen, in: C. Domb, J.L. Lebowitz (Eds.), *Phase Transition and Critical Phenomena*, Clarendon Press, Oxford, 1995.
- [120] P.K. Mishra, S. Kumar, Y. Singh, A simple and exactly solvable model for a semi flexible polymer chain interacting with a surface, *Physica A* 323 (2003) 453–465.
- [121] A. Rosa, D. Marenduzzo, A. Maritan, F. Seno, Mechanical unfolding of directed polymers in a poor solvent: Critical exponents, *Phys. Rev. E* 67 (2003) 041802.
- [122] C. Domb, J.L. Lebowitz, *Phase Transition and Critical Phenomena*, vol. 13, Academic Press, New York, 1989.
- [123] S. Kumar, I. Jensen, J.L. Jaconson, A.J. Guttmann, Role of conformational entropy in force induced biopolymer unfolding, *Phys. Rev. Lett.* 98 (2007) 128101–128104.
- [124] A.J. Guttmann, J.L. Jacobsen, I. Jensen, S. Kumar, Modeling force induced bio-polymer unfolding, *J. Math. Chem.* 45 (2009) 223–237.
- [125] I. Jensen, Enumeration of self-avoiding walks on the square lattice, *J. Phys. A: Math. Gen.* 37 (2004) 5503.
- [126] N. Singh, Y. Singh, Statistical theory of force-induced unzipping of DNA, *Eur. Phys. J. E* 17 (2005) 7–19.
- [127] B.J. Alder, T.E. Wainwright, Phase transition for a hard sphere system, *J. Chem. Phys.* 27 (1957) 1208–1209.
- [128] J.A. McCammon, B.R. Gelin, M. Karplus, Dynamics of folded proteins, *Nature* 267 (1977) 585–590.
- [129] V. Daggett, M. Levitt, Molecular dynamics simulations of helix denaturation, *J. Mol. Biol.* 223 (1992) 1121–1138.
- [130] J. Kubelka, J. Hofrichter, W.A. Eaton, The protein folding ‘speed limit’, *Curr. Opin. Struct. Biol.* 14 (2004) 76–88.
- [131] S.A. Adcock, J.A. McCammon, Molecular Dynamics: Survey of methods for simulating the activity of proteins, *Chem. Rev.* 106 (2006) 1589–1615.
- [132] W.C. Swope, H.C. Andersen, P.H. Berens, K.R. Wilson, A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters and molecules: Application to small water clusters, *J. Chem. Phys.* 76 (1982) 637–649.
- [133] G.I. Bell, Models for the specific adhesion of cells to cells, *Science* 200 (1978) 618–627.
- [134] E. Evans, K. Ritchie, Dynamic strength of molecular adhesion bonds, *Biophys. J.* 72 (1997) 1541–1555.
- [135] O.K. Dudko, G. Hummer, A. Szabo, Intrinsic rates and activation free energies from single-molecule pulling experiments, *Phys. Rev. Lett.* 96 (2006) 108101–108104.
- [136] H.A. Kramers, Brownian motion in a field of force and the diffusion model of chemical reactions, *Physica (Utrecht)* 7 (1940) 284–304.
- [137] P. Hanggi, P. Talkner, M. Borkovec, Reaction-rate theory: Fifty years after Kramers, *Rev. Modern Phys.* 62 (1990) 251–341.
- [138] M. Carrion-Vasquez, A.F. Oberhauser, S.B. Fowler, P.E. Marszalek, S.E. Broedel, J. Clarke, J.M. Fernandez, Mechanical and chemical unfolding of a single protein: A comparison, *Proc. Natl. Acad. Sci. USA* 96 (1999) 3694–3699.
- [139] M. Kouza, C. Hu, M.S. Li, New force replica exchange method and protein folding pathways probed by force-clamp technique, *J. Chem. Phys.* 128 (2008) 045103.
- [140] G.S. Hammond, A correlation of reaction rates, *J. Am. Chem. Soc.* 77 (1953) 334–338.
- [141] J.E. Leffler, Parameters for the description of transition states, *Science* 117 (1953) 340–341.
- [142] A. Matouschek, D.E. Otzen, L. Izaki, S.E. Jackson, A.R. Fersht, Movement of the position of the transition state in protein folding, *Biochemistry* 34 (1995) 13656–13662.
- [143] P.A. Dalby, M. Oliveberg, A.R. Fersht, Movement of the intermediate and rate determining transition state of barnase on the energy landscape with changing temperature, *Biochemistry* 37 (1998) 4674–4679.

- [144] G. Pappenberger, C. Saudan, M. Becker, A.E. Merbach, T. Kiehaber, Denaturant-induced movement of the transition state of protein folding revealed by high-pressure stopped-flow measurements, *Proc. Natl. Acad. Sci. USA* 97 (2000) 17–22.
- [145] C. Hyeon, D. Thirumalai, Forced-unfolding and force-quench refolding of RNA hairpins, *Biophys. J.* 90 (2006) 3410–3427.
- [146] D.J. Lacks, Energy landscape distortions and the mechanical unfolding of proteins, *Biophys. J.* 88 (2005) 3494–3501.
- [147] M. Schlierf, M. Rief, Single-molecule unfolding force distributions reveal a funnel-shaped energy landscape, *Biophys. J.* 90 (2006) L33–L35.
- [148] S. Izrailev, S. Stepaniants, M. Balsera, Y. Oono, K. Schulten, Molecular dynamics study of unbinding of the avidin–biotin complex, *Biophys. J.* 72 (1997) 1568–1581.
- [149] B.E. Shapiro, H. Qian, A quantitative analysis of single protein-ligand complex separation with the atomic force microscope, *Biophys. Chem.* 67 (1997) 211–219.
- [150] G. Hummer, A. Szabo, Kinetics from nonequilibrium single-molecule pulling experiments, *Biophys. J.* 85 (2003) 5–15.
- [151] O.K. Dudko, A.E. Filippov, J. Klafter, M. Urbakh, Beyond the conventional description of dynamic force spectroscopy of adhesion bonds, *Proc. Natl. Acad. Sci. USA* 100 (2003) 11378–11381.
- [152] R. Zwanzig, *Nonequilibrium Statistical Mechanics*, Oxford University Press, New York, 2001.
- [153] A. Szabo, K. Schulten, First passage time approach to diffusion controlled reactions, *J. Chem. Phys.* 72 (1980) 4350–4357.
- [154] J.C. Shillcock, U. Seifert, Escape from a metastable well under a time-ramped force, *Phys. Rev. E* 57 (1998) 7301–7304.
- [155] M. Abramowitz, I.A. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1972.
- [156] M. Raible, M. Evstigneev, P. Reimann, F.W. Bartels, P. Ros, Theoretical analysis of dynamic force spectroscopy experiments on ligand-receptor complexes, *J. Biotechnol.* 112 (2004) 13–23.
- [157] J.D. Weeks, J.B. Lucks, Y. Kafri, C. Danilowicz, D.R. Nelson, M. Prentiss, Pause point spectra in DNA constant-force unzipping, *Biophys. J.* 88 (2005) 2752–2765.
- [158] G. Binnig, C.F. Quate, C.H. Berger, Atomic Force Microscope, *Phys. Rev. Lett.* 56 (1986) 930–933.
- [159] R. Simmons, J. Finer, S. Chu, J. Spudich, Quantitative measurements of force and displacement using an optical trap, *Biophys. J.* 70 (1996) 1813–1822.
- [160] F. Amblard, B. Yurke, A. Pargellis, S. Leibler, A magnetic manipulator for studying local rheology and micromechanical properties of biological systems, *Rev. Sci. Instrum.* 67 (1996) 818–827.
- [161] E. Evans, K. Ritchie, R. Merkel, Sensitive force technique to probe molecular adhesion and structural linkages at biological interfaces, *Biophys. J.* 68 (1995) 2580–2587.
- [162] T.R. Strick, M.-N. Dessinges, G. Charvin, N.H. Dekker, J.-F. Allemand, D. Benmison, V. Croquette, Stretching of macromolecules and proteins, *Rep. Prog. Phys.* 66 (2003) 1–45.
- [163] A. Ashkin, Optical trapping and manipulation of neutral particles using lasers, *Proc. Natl. Acad. Sci. USA* 94 (1997) 4853–4860.
- [164] K.C. Neuman, S.M. Block, Optical trapping, *Rev. Sci. Instrum.* 75 (2004) 2787–2809.
- [165] M.S.Z. Kellermayer, S.B. Smith, H.L. Granzier, C. Bustamante, Folding-unfolding transitions in single titin molecules characterized with laser tweezers, *Science* 276 (1997) 1112–1116.
- [166] A. Halperin, E.B. Zhulina, On the deformation behaviour of collapsed polymers, *Europhys. Lett.* 15 (2003) 417–421.
- [167] P.L. Geissler, E.I. Shakhnovich, Reversible stretching of random heteropolymers, *Phys. Rev. E* 65 (2002) 056110.
- [168] D. Marenduzzo, A. Maritan, F. Seno, Force-induced unfolding of a homopolymer on a fractal lattice: Exact results versus mean-field predictions, *J. Phys. A: Math. Gen.* 35 (2002) L233–L240.
- [169] S. Kumar, D. Giri, Force-induced conformational transition in a system of interacting stiff polymers: Application to unfolding, *Phys. Rev. E* 72 (2005) 052901.
- [170] H.J. Zhou, J. Zhou, Z.C. Ou-Yang, S. Kumar, Collapse transition of two-dimensional flexible and semiflexible polymers, *Phys. Rev. Lett.* 97 (2006) 158302.
- [171] D. Marenduzzo, A. Maritan, A. Rosa, F. Seno, Stretching of a polymer below the  $\theta$  point, *Phys. Rev. Lett.* 90 (2003) 088301.
- [172] D. Marenduzzo, A. Maritan, A. Rosa, F. Seno, Stepwise unfolding of collapsed polymers, *Eur. Phys. J. E* 15 (2004) 83–93.
- [173] R. Rajesh, D. Giri, I. Jensen, S. Kumar, Role of pulling direction in understanding the energy landscape of proteins, *Phys. Rev. E* 78 (2008) 021905.
- [174] S. Kumar, D. Giri, Does changing the pulling direction give better insight into biomolecules? *Phys. Rev. Lett.* 98 (2007) 048101.
- [175] Y. Singh, D. Giri, S. Kumar, Crossover of a polymer chain from bulk to surface states, *J. Phys. A: Math. Gen.* 34 (2001) L67–L74.
- [176] S. Kumar, G. Mishra, Force-induced stretched state: Effects of temperature, *Phys. Rev. E* 78 (2008) 011907.
- [177] J. Krawczyk, I. Jensen, A.L. Owczarek, S. Kumar, Pulling self-interacting polymers in two dimensions, *Phys. Rev. E* 79 (2009) 031912.
- [178] S. Doniach, T. Garel, H. Orland, Phase diagram of a semiflexible polymer chain in a  $\theta$  solvent: Application to protein folding, *J. Chem. Phys.* 105 (1996) 1601–1608.
- [179] U. Bastolla, P. Grassberger, Phase transitions of single semistiff polymer chains, *J. Chem. Phys.* 89 (1997) 1061–1078.
- [180] J. des Cloizeaux, G. Jannink, *Polymers in Solution*, Clarendon Press, Oxford, UK, 1990.
- [181] A.Y. Grosberg, A.R. Khokhlov, *Statistical Physics of Macromolecules*, American Institute of Physics, New York, 1994.
- [182] P.G. deGennes, Collapse of a polymer-chain in poor solvents, *J. Phys. (Paris)* 36 (1975) L55.
- [183] M.J. Stephen, Collapse of a polymer-chain, *Phys. Lett.* 53 (1975) 363.
- [184] B. Duplantier, Lagrangian tricritical theory of polymer-chain solutions near the theta-point, *J. Phys. (Paris)* 43 (1982) 991.
- [185] A.L. Owczarek, T. Prellberg, First-order scaling near a second-order phase transition: Tricritical polymer collapse, *Europhys. Lett.* 51 (2000) 602–607.
- [186] C. Vanderzande, A.L. Stella, F. Seno, Percolation, the special  $\Theta'$  point, and the  $\Theta-\Theta'$  universality puzzle, *Phys. Rev. Lett.* 67 (1991) 2757.
- [187] B. Duplantier, Higher conformal multifractality, *J. Stat. Phys.* 110 (2003) 691–738.
- [188] X. Wang, X. Qiu, C. Wu, Comparison of the coil-to-globule and the globule-to-coil transitions of a single poly(*n*-isopropylacrylamide) homopolymer chain in water, *Macromolecules* 31 (1998) 2792–2796.
- [189] K. Yoshikawa, M. Takahashi, V.V. Vasilevskaya, A.R. Khokhlov, Large discrete transition in a single DNA molecule appears continuous in the ensemble, *Phys. Rev. Lett.* 76 (1996) 3029–3031.
- [190] F. Rampf, W. Paul, K. Binder, On the first-order collapse transition of a three-dimensional, flexible homopolymer chain model, *Europhys. Lett.* 70 (2005) 628–634.
- [191] P. Grassberger, H.P. Hsu, Stretched polymers in a poor solvent, *Phys. Rev. E* 65 (2002) 031807.
- [192] B. Maier, J.O. Radler, Conformation and self-diffusion of single DNA molecules confined to two dimensions, *Phys. Rev. Lett.* 82 (1999) 1911–1914.
- [193] D.J. Brockwell, E. Paci, R.C. Zinober, G.S. Beddard, P.D. Olmsted, D.A. Smith, R.N. Perham, S.E. Radford, Pulling geometry defines the mechanical resistance of a beta-sheet protein, *Nature Struct. Biol.* 10 (2003) 731–737.
- [194] M. Carrion-Vazquez, H.B. Li, H. Lu, P.E. Marszalek, A.F. Oberhauser, J.M. Fernandez, The mechanical stability of ubiquitin is linkage dependent, *Nature Struct. Biol.* 10 (2003) 738–743.
- [195] H. Dietz, M. Rief, Protein structure by mechanical triangulation, *Proc. Natl. Acad. Sci. USA* 103 (2006) 1244–1247.
- [196] H. Dietz, F. Berkemeier, M. Bertz, M. Rief, Anisotropic deformation response of single protein molecules, *Proc. Natl. Acad. Sci. USA* 103 (2006) 12724–12728.
- [197] T.E. Creighton, *Protein Folding*, Freeman, New York, 1992.
- [198] A. Matouschek, C. Bustamante, Finding a protein's achilles heel, *Nature Struct. Biol.* 10 (2003) 674–676.
- [199] G.U. Lee, L.A. Chrisey, R.J. Colton, Direct measurement of the forces between complementary strands of DNA, *Science* 266 (1994) 771–773.
- [200] D.C. Rau, V.A. Parsegian, Direct measurement of the intermolecular forces between counterion-condensed DNA double helices – Evidence for long-range attractive hydration forces, *Biophys. J.* 61 (1992) 246–259.
- [201] T. Boland, B.D. Ratner, Direct measurement of hydrogen-bonding in DNA nucleotide bases by atomic force microscopy, *Proc. Natl. Acad. Sci. USA* 92 (1995) 5297–5301.

- [202] A. Noy, D.V. Vezenov, J.F. Kayyem, T.J. Maade, C.M. Lieber, Stretching and breaking duplex DNA by chemical force microscopy, *Chem. Biol.* 4 (1997) 519–527.
- [203] L.D. William, L.J.M. III, Electrostatic mechanism of DNA deformation, *Annu. Rev. Biophys. Biomol. Struct.* 29 (2000) 497–521.
- [204] P. Cluzel, A. Lebrun, C. Heller, R. Lavery, J.L. Viovy, D. Chatenay, F. Caron, DNA: An extensible molecule, *Science* 271 (1992) 792–794.
- [205] C. Bustamante, Z. Bryant, S.B. Smith, Ten years of tension: Single-molecule DNA mechanics, *Nature* 421 (2003) 423–427.
- [206] J.F. Leger, G. Romano, A. Sarkar, J. Robert, L. Bourdieu, D. Chatenay, J.F. Marko, Structural transitions of a twisted and stretched DNA molecule, *Phys. Rev. Lett.* 83 (1999) 1066–1069.
- [207] J.F. Marko, DNA under high tension: Overstretching, undertwisting, and relaxation dynamics, *Phys. Rev. E* 57 (1998) 2134–2149.
- [208] P. Cizeau, J.L. Viovy, Modeling extreme extension of DNA, *Biopolymers* 42 (1997) 383–385.
- [209] S. Cocco, J.F. Marko, R. Monasson, Theoretical models for single molecule DNA and RNA experiments: From elasticity to unzipping, *C. R. Physique* 3 (2002) 569–584.
- [210] A. Ahsan, J. Rudnick, R. Bruinsma, Elasticity theory of the B-DNA to S-DNA transition, *Biophys. J.* 74 (1998) 132–137.
- [211] R. Podgornik, P.L. Hansen, V.A. Parsegian, Elastic moduli renormalization in self-interacting stretchable polyelectrolytes, *J. Chem. Phys.* 113 (2000) 9343–9350.
- [212] T.J. Odijk, Polyelectrolytes near rod limit, *J. Polym. Sci. Part B-Polymer Physics* 15 (1977) 477–483.
- [213] J. Skolnick, M. Fixman, Electrostatic persistence length of a worm-like polyelectrolyte, *Macromolecules* 10 (1977) 944–948.
- [214] O. Punkkinen, P.L. Hansen, L. Miao, I. Vattulainen, DNA overstretching transition: Ionic strength effects, *Biophys. J.* 89 (2005) 967–978.
- [215] I. Rouzina, V.A. Bloomfield, Force-induced melting of the DNA double helix 1: Thermodynamic analysis, *Biophys. J.* 80 (2001) 882–893.
- [216] I. Rouzina, V.A. Bloomfield, Force-induced melting of the DNA double helix 2: Effect of solution conditions, *Biophys. J.* 80 (2001) 894–900.
- [217] J.R. Wenner, M.C. Williams, I. Rouzina, V. Bloomfield, Salt dependence of the elasticity and overstretching transition of single DNA molecules, *Biophys. J.* 82 (2002) 3160–3169.
- [218] L. Shokri, M.J. McCauley, I. Rouzina, M.C. Williams, DNA overstretching in the presence of glyoxal: Structural evidence of force induced DNA melting, *Biophys. J.* 95 (2008) 1248.
- [219] N.E. Broude, E.I. Budowsky, Reaction of glyoxal with nucleic-acid components. 5. Denaturation of DNA under action of glyoxal, *Biochem. Biophys. Acta* 294 (1973) 378–384.
- [220] H. Kasai, N.I. Tanaka, S. Fukada, DNA modifications by the mutagen glyoxal: Adduction to G and C, deamination of C and GC and GA cross-linking, *Carcinogenesis* 19 (1998) 1459.
- [221] S. Cocco, J. Yan, J.F. Leger, D. Chatenay, J.F. Marko, Overstretching and force-driven strand separation of double-helix DNA, *Phys. Rev. E* 70 (2004) 011910.
- [222] H.J. Zhou, Y. Zhang, Z.C. Ou-Yang, Bending and base-stacking interactions in double-stranded DNA, *Phys. Rev. Lett.* 82 (1999) 4560–4563.
- [223] H.J. Zhou, Y. Zhang, Z.C. Ou-Yang, Elastic property of single double-stranded DNA molecules: Theoretical study and comparison with experiments, *Phys. Rev. E* 62 (2000) 1045–1058.
- [224] G.S. Manning, Limiting laws and counterion condensation in polyelectrolyte solutions i. Colligative properties, *J. Chem. Phys.* 51 (1969) 924–933.
- [225] M.O. Fenley, G.S. Manning, W.K. Olson, A numerical counterion condensation analysis of the B-Z transition of DNA, *Biopolymers* 30 (1990) 13–14.
- [226] R.X. Dong, X.L. Yan, G.F. Yu, S.G. Liu, B- to S-form transition of double-stranded DNA in solutions of various salt concentrations, *Phys. Lett. A* 318 (2003) 600–606.
- [227] H.X. Fu, C.G. Koh, H. Chen, Ionic effects on overstretching transition of B-DNA, *Eur. Phys. J. E* 17 (2005) 231–235.
- [228] C. Bustamante, S.B. Smith, J. Liphardt, D. Smith, Single-molecule studies of DNA mechanics, *Current Opinion Struct. Biol.* 10 (2000) 279–285.
- [229] G.J. Wuite, S.B. Smith, M. Young, D. Keller, C. Bustamante, Single molecule studies of the effect of template tension on T7 DNA polymerase activity, *Nature* 404 (2000) 103–106.
- [230] A. Montanari, M. Mezard, Hairpin formation and elongation of biomolecules, *Phys. Rev. Lett.* 86 (2000) 2178–2181.
- [231] M. Rief, H. Clausen-Schaumann, H.E. Gaub, Sequence-dependent mechanics of single DNA molecules, *Nat. Struct. Biol.* 6 (1999) 346–349.
- [232] B. Maier, D. Bensimon, V. Croquette, Replication by a single DNA polymerase of a stretched single-stranded DNA, *Proc. Natl. Acad. Sci. USA* 97 (2000) 12002–12007.
- [233] H.J. Zhou, Y. Zhang, Z.C. Ou-Yang, Stretch-induced hairpin-coil transitions in designed polynucleotide chains, *Phys. Rev. Lett.* 86 (2001) 356–359.
- [234] Y. Zhang, H. Zhou, Z.-C. Ou-Yang, Stretching single-stranded DNA: Interplay of electrostatic, base-pairing, and base-pair stacking interactions, *Biophysics J.* 81 (2001) 1133–1143.
- [235] T. Hugel, M. Grosholz, H. Clausen-Schaumann, A. Pfau, H. Gaub, M. Sei, Elasticity of single polyelectrolyte chains and their desorption from solid supports studied by AFM based single molecule force spectroscopy, *Macromolecules* 34 (2001) 1039.
- [236] D.W. Urry, T. Hugel, M. Seitz, H.E. Gaub, L. Sheiba, J. Dea, J. Xu, T. Parker, Elastin: A representative ideal protein elastomer, *Philos. Trans. R. Soc. B* 357 (2002) 169.
- [237] T. Hugel, M. Rief, M. Seitz, H.E. Gaub, R.R. Netz, Highly stretched state single polymers: Atomic force microscope experiments versus *ab-initio* theory, *Phys. Rev. Lett.* 94 (2005) 48301.
- [238] C. Danilowicz, C.H. Lee, V.W. Coljee, M. Prentiss, Effects of temperature on the mechanical properties of single stranded DNA, *Phys. Rev. E* 75 (2007) 030902(R).
- [239] M.N. Dessinges, B. Maier, Y. Zhang, M. Peliti, D. Bensimon, V. Croquette, Stretching single stranded DNA, a model polyelectrolyte, *Phys. Rev. Lett.* 89 (2002) 248102.
- [240] Y. Seol, G.M. Skinner, K. Visscher, Elastic properties of a single-stranded charged homopolymeric ribonucleotide, *Phys. Rev. Lett.* 93 (2004) 118102.
- [241] Y. Seol, G.M. Skinner, K. Visscher, Stretching of homopolymeric RNA reveals single-stranded helices and base-stacking, *Phys. Rev. Lett.* 98 (2007) 158103.
- [242] C. Ke, M. Humeniuk, H. S-Gracz, P.E. Marszalek, Direct measurements of base stacking interactions in DNA by single-molecule atomic-force spectroscopy, *Phys. Rev. Lett.* 99 (2007) 018302.
- [243] N.L. Goddard, G. Bonnet, O. Krichevsky, A. Libchaber, Sequence dependent rigidity of single stranded DNA, *Phys. Rev. Lett.* 85 (2000) 2400.
- [244] A. Buhot, A. Halperin, Effects of stacking on the configurations and elasticity of single-stranded nucleic acids, *Phys. Rev. E* 70 (2004) 020902.
- [245] B.H. Zimm, J.K. Bragg, Theory of phase transition between helix and random coil in polypeptide chains, *J. Chem. Phys.* 31 (1959) 526.
- [246] G. Mishra, D. Giri, S. Kumar, Stretching of a single stranded DNA: Evidence for structural transition, *Phys. Rev. E* 79 (2009) 031930.
- [247] M. Daune, *Molecular Biophysics: Structure in Motion*, Oxford, New York, 1999.
- [248] W.K. Olson, J.L. Sussman, How flexible is the furanose ring? 1. A comparison of experimental and theoretical studies, *J. Am. Chem. Soc.* 104 (1982) 270.
- [249] U. Bockelmann, B. Essevaz-Roulet, F. Heslot, Molecular stick-slip motion revealed by opening DNA with piconewton forces, *Phys. Rev. Lett.* 79 (1997) 4489–4492.
- [250] B. Essevaz-Roulet, U. Bockelmann, F. Heslot, Mechanical separation of the complementary strands of DNA, *Proc. Natl. Acad. Sci. USA* 94 (1997) 11935–11940.
- [251] U. Bockelmann, B. Essevaz-Roulet, F. Heslot, DNA strand separation studied by single molecule force measurements, *Phys. Rev. E* 58 (1998) 2386–2394.
- [252] J.L. Viovy, C. Heller, F. Caron, P. Cluzel, D. Chatenay, Ultrafast sequencing of DNA by mechanical opening of the double helix – A theoretical investigation, *Comptes Rendus De L'Academie Des Sciences Serie III-Sciences De La Vie-Life Sciences* 317 (1994) 795–800.
- [253] K.L. Sebastian, Pulling a polymer out of a potential well and the mechanical unzipping of DNA, *Phys. Rev. E* 62 (2000) 1128–1132.
- [254] P. Lam, Excluded volume effects in gene stretching, *Biopolymers* 64 (2002) 57–62.
- [255] P. Lam, J.C.S. Levy, H.C. Huang, Excluded volume effect in unzipping DNA with a force, *Biopolymers* 73 (2004) 293–300.

- [256] T. Strunz, K. Oroszlan, R. Schafer, H.-J. Guntherodt, Dynamic force spectroscopy of single DNA molecules, *Proc. Natl. Acad. Sci. USA* 96 (1999) 11277–11282.
- [257] I. Schumakovitch, W. Grange, T. Strunz, P. Bertoncini, Hans-Joachim Guntherodt, M. Hegner, Temperature dependence of unbinding forces between complementary DNA strands, *Biophys. J.* 82 (2002) 517–521.
- [258] D. Porschke, Model calculations on kinetics of oligonucleotide double helix coil transitions – Evidence for a fast chain sliding reaction, *Biophysical Chem.* 2 (1974) 83–96.
- [259] R.A. Neher, U. Gerland, Dynamics of force-induced DNA slippage, *Phys. Rev. Lett.* 93 (2004) 198102.
- [260] R.A. Neher, U. Gerland, DNA as a programmable viscoelastic nanoelement, *Biophys. J.* 89 (2005) 3846–3855.
- [261] F. Kuhner, J. Morfill, R.A. Neher, K. Blank, H.E. Gaub, Force-induced DNA slippage, *Biophys. J.* 92 (2007) 2491–2497.
- [262] R. Nevo, V. Brumfeld, R. Kapon, P. Hinterdorfer, Z. Reich, Direct measurement of protein energy landscape roughness, *EMBO Reports* 6 (2005) 482.
- [263] C. Danilowicz, V.W. Coljee, C. Bouzigues, D.K. Lubensky, D.R. Nelson, M. Prentiss, DNA unzipped under a constant force exhibits multiple metastable intermediates, *Proc. Natl. Acad. Sci. USA* 100 (2003) 1694–1699.
- [264] C. Danilowicz, Y. Kafri, R.S. Conroy, V.W. Coljee, J. Weeks, M. Prentiss, Measurement of the phase diagram of DNA unzipping in the temperature-force plane, *Phys. Rev. Lett.* 93 (2004) 078101.
- [265] U. Bockelmann, Single-molecule manipulation of nucleic acids, *Current Opinion Struct. Biol.* 14 (2004) 368–373.
- [266] M.C. Williams, I. Rouzina, V.A. Bloomfield, Thermodynamics of DNA interactions from single molecule stretching experiments, *Account of Chemical Research* 35 (2002) 159–166.
- [267] P.M. Lam, J.C. Levy, Unzipping DNA from the condensed globule state – Effects of unraveling, *Biopolymers* 79 (2005) 287–291.
- [268] C. Levinthal, R. Crane, On the unwinding of DNA, *Proc. Natl. Acad. Sci. USA* 42 (1956) 436–438.
- [269] P. Nelson, Transport of torsional stress in DNA, *Proc. Natl. Acad. Sci. USA* 96 (1999) 14342–14347.
- [270] P. Thomen, U. Bockelmann, F. Heslot, Rotational drag on DNA: A single molecule experiment, *Phys. Rev. Lett.* 88 (2002) 248102.
- [271] T.T. Perkins, D.E. Smith, R.G. Larsen, S. Chu, Stretching of a single tethered polymer in a uniform flow, *Science* 268 (1995) 83–87.
- [272] B.H. Zimm, Extension in flow of a DNA molecule tethered at one end, *Macromolecules* 31 (1998) 6089–6098.
- [273] S. Cocco, R. Monasson, Statistical mechanics of torque induced denaturation of DNA, *Phys. Rev. Lett.* 83 (1999) 5178–5181.
- [274] M. Barbi, S. Cocco, M. Peyrard, S. Ruffo, A twist opening model for DNA, *J. Biol. Phys.* 24 (1999) 97–114.
- [275] D. Bensimon, A.J. Simon, V. Croquette, A. Bensimon, Stretching DNA with a receding meniscus: Experiments and models, *Phys. Rev. Lett.* 74 (1995) 4754–4757.
- [276] H. Clausen-Schaumann, M. Rief, C. Tolksdorf, H.E. Gaub, Mechanical stability of single DNA molecules, *Biophys. J.* 78 (2000) 1997.
- [277] T.R. Strick, J. Allemand, D. Bensimon, A. Bensimon, V. Croquette, The elasticity of a single supercoiled DNA molecule, *Science* 271 (1996) 1835.
- [278] T.R. Strick, J.F. Allemand, D. Bensimon, V. Croquette, Stress-induced structural transitions in DNA and proteins, *Annu. Rev. of Biophys. and Biomol. Struc.* 29 (2000) 523–543.
- [279] A. Sarkar, J.F. Leger, D. Chatenay, J.F. Marko, Structural transitions in DNA driven by external force and torque, *Phys. Rev. E* 63 (2001) 051903.
- [280] Z. Bryant, M.D. Stone, J. Gore, S.B. Smith, N.R. Cozzarelli, C. Bustamante, Structural transitions and elasticity from torque measurements on DNA, *Nature* 424 (2003) 338–341.
- [281] E.-L. Florin, V. Moy, H. Gaub, Adhesion forces between individual ligand-receptor pairs, *Science* 264 (1994) 415–417.
- [282] M. Rief, J. Pascual, M. Saraste, H. Gaub, Single molecule force spectroscopy of spectrin repeats: Low unfolding forces in helix bundles, *J. Mol. Biol.* 286 (1999) 553–561.
- [283] M.S. Li, Secondary structure, mechanical stability and location of transition state of proteins, *Biophys. J.* 93 (2007) 2644–2654.
- [284] K.W. Plaxco, K.T. Simon, D. Baker, Contact order, transition state placement and the refolding rates of single domain proteins, *J. Mol. Biol.* 277 (1998) 985–994.
- [285] O.K. Dudko, J. Mathe, A. Szabo, A. Meller, G.H. Hummer, Extracting kinetics from single-molecule force spectroscopy: Nanopore unzipping of DNA hairpins, *Biophys. J.* 92 (2007) 4188–4195.
- [286] C.J. Camacho, D. Thirumalai, Kinetics and thermodynamics of folding in model proteins, *Proc. Natl. Acad. Sci. USA* 90 (1993) 6369–6372.
- [287] A.M. Ferrenberg, R.H. Swendsen, Optimized Monte-Carlo data analysis, *Phys. Rev. Lett.* 63 (1989) 1195–1198.
- [288] D.K. Klimov, D. Thirumalai, Lattice model studies of force-induced unfolding of proteins, *J. Phys. Chem. B* 105 (2001) 6648–6654.
- [289] M. Kouza, C.F. Chang, S. Hayryan, T.H. Yu, M.S. Li, T.H. Huang, C.K. Hu, Folding of the protein domain hbSBD, *Biophys. J.* 89 (2005) 3353–3361.
- [290] S.T. Thomas, V.V. Loladze, G.I. Makhatadze, Hydration of the peptide backbone largely defines the thermodynamic propensity scale of residues at the C' position of the C-capping box of  $\alpha$ -helices, *Proc. Natl. Acad. Sci. USA* 98 (2001) 10670–10675.
- [291] H. Dietz, M. Rief, Exploring the energy landscape of GFP by single-molecule mechanical experiments, *Proc. Natl. Acad. Sci. USA* 101 (2004) 16192–16197.
- [292] P.M. Williams, S.B. Fowler, R.B. Best, J. Luis, T. Herrera, K.A. Scott, A. Steward, J. Clarke, Hidden complexity in the mechanical properties of titin, *Nature* 422 (2003) 446–449.
- [293] I. Schweiger, A. Kardinal, M. Schleicher, A.A. Noegel, M. Rief, A mechanical unfolding intermediate in an actin-crosslinking protein, *Nat. Struct. Mol. Biol.* 11 (2003) 81–85.
- [294] I. Schweiger, M. Schlierf, A. Noegel, M. Rief, The folding pathways of a fast-folding immunoglobulin domain revealed by single-molecule mechanical experiments, *EMBO Reports* 6 (2005) 46–51.
- [295] M.S. Li, A.M. Gabovich, A.I. Voitenko, New method for deciphering free energy landscape of three-state proteins, *J. Chem. Phys.* 129 (2008) 105102–105107.
- [296] P.F. Lenne, A.J. Raue, S.M. Altmann, M. Saraste, J.K.H. Horber, States and transitions during forced unfolding of a single spectrin repeat, *FEBS Lett.* 476 (2000) 124–128.
- [297] S.M. Altmann, R.G. Grunberg, P.F. Lenne, J. Ylanne, A. Raue, K. Herbert, M. Saraste, M. Nilges, J.K.H. Horber, Pathways and intermediates in forced unfolding of spectrin repeats, *Structure* 10 (2002) 1085–1096.
- [298] Q. Peng, H.B. Li, Atomic force microscopy reveals parallel mechanical unfolding pathways of T4 lysozyme: Evidence for a kinetic partitioning mechanism, *Proc. Natl. Acad. Sci. USA* 105 (2008) 1885–1890.
- [299] D. Thirumalai, D.K. Klimov, S.A. Woodson, Kinetic partitioning mechanism: A unifying theme in the folding of biomolecules, *Theor. Chem. Accounts* 96 (1997) 14–22.
- [300] C. Cecconi, E.A. Shank, C. Bustamante, S. Marqusee, Direct observation of the three-state folding of a single protein molecule, *Science* 309 (2005) 2057–2060.
- [301] M. Schlierf, H. Li, J.M. Fernandez, The unfolding kinetics of ubiquitin captured with single-molecule force-clamp techniques, *Proc. Natl. Acad. Sci. USA* 101 (2004) 7299–7304.
- [302] C. Chyan, F. Lin, H. Peng, J. Yuan, C. Chang, S. Lin, G. Yang, Reversible mechanical unfolding of single ubiquitin molecules, *Biophys. J.* 87 (2004) 3995–4006.
- [303] J. Brujic, R.I. Hermans, K.A. Walther, J.M. Fernandez, Single-molecule force spectroscopy reveals signatures of glassy dynamics in the energy landscape of ubiquitin, *Nature Phys.* 2 (2006) 282–286.
- [304] A. Imparato, A. Pelizzola, Mechanical unfolding and refolding pathways of ubiquitin, *Phys. Rev. Lett.* 100 (2008) 158104.
- [305] D.J. Brockwell, G. Beddard, E. Paci, D. West, P. Olmsted, D. Smith, S. Radford, Mechanically unfolding the small, topologically simple protein L, *Biophys. J.* 89 (2005) 506–519.
- [306] F. Cordier, S. Grzesiek, Temperature-dependence of protein hydrogen bond properties as studied by high-resolution NMR, *J. Mol. Biol.* 317 (2002) 739–752.

- [307] D.J. Brockwell, G.S. Beddar, J. Clarkson, R.C. Zinober, A.W. Blake, J. Trinick, P.D. Omsted, D.A. Smith, S.E. Radford, The effect of core destabilization on the mechanical resistance of I27, *Biophys. J.* 83 (2002) 458–472.
- [308] R. Best, B. Li, A. Steward, V. Daggett, J. Clarke, Can non-mechanical proteins withstand force? stretching barnase by atomic force microscopy and molecular dynamics simulations, *Biophys. J.* 81 (2001) 2344–2356.
- [309] M. Carrión-Vázquez, A.F. Oberhauser, T.E. Fisher, P.E. Marszalek, H. Li, J.M. Fernandez, Mechanical design of proteins studied by single-molecule force spectroscopy and protein engineering, *Prog. Biophys. Mol. Biol.* 74 (2000) 63–91.
- [310] R. Law, G. Liao, S. Harper, G. Yang, D.W. Speicher, D.E. Discher, Pathway shifts and thermal softening in temperature-coupled forced unfolding of spectrin domains, *Biophys. J.* 85 (2003) 3286–3293.
- [311] V. Ortiz, S.O. Nielsen, M.L. Klein, D.E. Discher, Unfolding a linker between helical repeats, *J. Mol. Biol.* 349 (2005) 638–647.
- [312] A.M. Gabovich, M.S. Li, Mechanical stability of proteins, *J. Chem. Phys.* 131 (2009) 024121.
- [313] D.N. Ivankov, S.O. Garbuzynskiy, E. Alm, K.W. Plaxco, D. Baker, A.V. Finkelstein, Contact order revisited: Influence of protein size on the folding rate, *Protein Sci.* 19 (2003) 2057–2062.
- [314] D. Bartolo, I. Derenyi, A. Ajdari, Dynamic response of adhesion complexes: Beyond the single-path picture, *Phys. Rev. E* 65 (2002) 051910–051913.
- [315] P.-C. Li, D.E. Makarov, Simulation of the mechanical unfolding of ubiquitin: Probing different unfolding reaction coordinates by changing the pulling geometry, *J. Chem. Phys.* 121 (2004) 4826–4832.
- [316] D. West, P. Olmsted, E. Paci, Mechanical unfolding revisited through a simple but realistic model, *J. Chem. Phys.* 124 (2006) 154909–154917.
- [317] H. Wako, N. Saito, Statistical mechanical theory of the protein conformation. II. folding pathway for protein, *J. Phys. Soc. Japan* 44 (1978) 1939–1945.
- [318] L. Li, S. Wetzel, A. Pluckthun, J.M. Fernandez, Stepwise unfolding of ankyrin repeats in a single protein revealed by atomic force microscopy, *Biophys. J.* 90 (2006) L30–L32.
- [319] H. Dietz, M. Rief, Elastic bond network model for protein unfolding mechanics, *Phys. Rev. Lett.* 100 (2008) 098101.
- [320] L. Serrano, J.T. Kellis, P. Cann, P. Matouschek, A.R. Fersht, The folding of an enzyme II. Substructure of barnase and the contribution of different interactions to protein stability, *J. Mol. Biol.* 224 (1992) 783–804.
- [321] A.E. Eriksson, W.A. Baase, X.J. Zhang, D.W. Heinz, M. Blaber, E.P. Baldwin, B.W. Matthews, Response of a protein structure to cavity-creating mutations and its relation to the hydrophobic effect, *Science* 255 (1992) 178–183.
- [322] R.D. Toochnik, P.M. Williams, Simulations of multi-directional forced unfolding of titin I27, *J. Mol. Graphics Model.* 24 (2006) 396–403.
- [323] J.L. Milne, D. Shi, P.B. Rosenthal, J.S. Sunshine, G.J. Domingo, X. Wu, B.R. Brooks, R.N. Perham, R. Henderson, S. Subramanian, Molecular architecture and mechanism of an icosahedral pyruvate dehydrogenase complex: A multifunctional catalytic machine, *Embo J.* 21 (2002) 5587–5598.
- [324] M.S. Li, D.K. Klimov, D. Thirumalai, Thermal denaturation and folding rates of single domain proteins: Size matters, *Polymer* 45 (2004) 573–579.
- [325] B. Schuler, E.A. Lipman, W.A. Eaton, Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy, *Nature* 419 (2002) 743–747.
- [326] W.Y. Yang, M. Gruebele, Folding at the speed limit, *Nature* 423 (2003) 193–197.
- [327] S. Khorasanizadeh, L.D. Peters, T.R. Butt, H. Roder, Folding and stability of a tryptophan-containing mutant of ubiquitin, *Biochemistry* 32 (1993) 7054–7063.
- [328] C.F. Wright, A. Steward, J. Clarke, Thermodynamic characterisation of two transition states along parallel protein folding pathways, *J. Mol. Biol.* 338 (2004) 445–451.
- [329] N.C. Harris, Y. Song, C.-H. Kiang, Experimental free energy surface reconstruction from single molecule force spectroscopy using Jarzynski's equality, *Phys. Rev. Lett.* 99 (2007) 068101–068104.
- [330] G. Hummer, A. Szabo, Free energy surfaces from single-molecule force spectroscopy, *Acc. Chem. Res.* 38 (2005) 504–513.
- [331] G.E. Crooks, Entropy production fluctuation theorem and the nonequilibrium work relation for free energy differences, *Phys. Rev. E* 60 (1999) 2721–2726.
- [332] J. Liphardt, S. Dumont, S.B. Smith, L. Tinoco, C. Bustamante, Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality, *Science* 296 (2002) 1833–1835.
- [333] S. Park, F. Khalili-Araghi, E. Tajkhorshid, K. Schulten, Free energy calculation from steered molecular dynamics simulations using Jarzynski's equality, *J. Chem. Phys.* 119 (2003) 3559–3566.
- [334] G. Hummer, A. Szabo, Free energy reconstruction from nonequilibrium single molecule experiments, *Proc. Natl. Acad. Sci. USA* 98 (2001) 3658–3661.
- [335] M.C. Williams, I. Rouzina, Force spectroscopy of single DNA and RNA molecules, *Curr. Opin. Struct. Biol.* 12 (2002) 330–336.
- [336] Y. Sakai, T. Ikehara, T. Nishi, K. Nakajima, M. Hara, Nanorheology measurement on a single polymer chain, *Appl. Phys. Lett.* 81 (2002) 724–726.
- [337] E. Thormann, D.R. Evans, V.S.J. Craig, Experimental studies of the dynamic mechanical response of a single polymer chain, *Macromolecules* 39 (2006) 6180–6185.
- [338] M.A. Lantz, S.P. Jarvis, H. Tokumoto, T. Martynski, T. Kusumi, C. Nakarura, J. Miyake, Stretching the  $\alpha$ -helix: A direct measurement of the hydrogen-bond energy of a single-peptide molecule, *Chem. Phys. Lett.* 315 (1999) 61–68.
- [339] M. Kawakami, K. Byrne, B.S. Khatri, T.C.B. McLeish, S.E. Radford, D.A. Smith, Viscoelastic measurements of single molecules on a millisecond time scale by magnetically driven oscillation of an atomic force microscope cantilever, *Langmuir* 21 (2005) 4765–4772.
- [340] M. Schlierf, F. Berkemeier, M. Rief, Direct observation of active protein folding using lock-in force spectroscopy, *Biophys. J.* 93 (2007) 3989–3998.
- [341] M.J. Higgins, J.E. Sader, S.P. Jarvis, Frequency modulation atomic force spectroscopy reveals individual intermediates associated with each unfolded I27 titin domain, *Biophys. J.* 90 (2006) 640–647.
- [342] H. Janovjak, D.J. Muller, A.D.L. Humphris, Molecular force modulation spectroscopy revealing the dynamic response of single bacteriorhodopsins, *Biophys. J.* 88 (2005) 1423–1431.
- [343] O. Braun, A. Hanke, U. Seifert, Probing molecular free energy landscapes by periodic loading, *Phys. Rev. Lett.* 93 (2004) 158105–158108.
- [344] A. Matouschek, Protein unfolding – An important process *in vitro*, *Curr. Opin. Struct. Biol.* 13 (2003) 98–109.
- [345] S. Prakash, A. Matouschek, Protein unfolding in the cell, *Trends Biochem. Sci.* 29 (2004) 593–600.
- [346] N. Pfanner, E.A. Craig, A. Hoolinger, Mitochondrial preprotein translocase, *Annu. Rev. Cell. Dev. Biol.* 13 (1997) 25–51.
- [347] W. Voos, H. Martin, T. Krimmer, N. Pfanner, Mechanisms of protein translocation into mitochondria, *Biochem. Biophys. Acta* 1442 (1999) 235–254.
- [348] N. Pfanner, A. Geissler, Versatility of the mitochondrial protein import machinery, *Nature Rev. Mol. Cell Biol.* 2 (2001) 339–349.
- [349] M.P. Schwartz, A. Matouschek, The dimensions of the protein import channels in the outer and inner mitochondrial membranes, *Proc. Natl. Acad. Sci. USA* 96 (1999) 13086–13090.
- [350] C. Lee, M.P. Schwartz, S. Prakash, M. Iwakura, A. Matouschek, ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal, *Mol. Cell* 7 (2001) 627–637.
- [351] S. Huang, S. Ratliff, M.P. Schwartz, J.M. Spenner, A. Matouschek, Mitochondria unfold precursor proteins by unraveling them from their N-termini, *Nat. Struct. Biol.* 6 (1999) 1132–1138.
- [352] J.A. Kenniston, T.A. Baker, J.M. Fernandez, R.T. Sauer, Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA(+) degradation machine, *Cell* 114 (2003) 511–520.
- [353] K. Shariff, S. Ghosal, A. Matouschek, The force exerted by the membrane potential during protein import into the mitochondrial matrix, *Biophys. J.* 86 (2004) 3647–3652.
- [354] L. Huang, S. Kirmizitdin, D.E. Makarov, Computer simulations of the translocation and unfolding of a protein pulled mechanically through a pore, *J. Chem. Phys.* 123 (2005) 124903.
- [355] P. Tian, I. Andrićioaei, Repetitive pulling catalyzes co-translational unfolding of barnase during import through a mitochondrial pore, *J. Mol. Biol.* 350 (2005) 1017–1043.
- [356] B.B. van den, R.J. Ellis, C.M. Dobson, Effects of macromolecular crowding on protein folding and aggregation, *Embo. J.* 18 (1999) 6927–6933.
- [357] D.S. Goodsell, Inside a living cell, *Trends Biochem. Sci.* 16 (1991) 203–206.

- [358] A.P. Minton, The influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media, *J. Biol. Chem.* 276 (2001) 10577–10580.
- [359] A.P. Minton, Implications of macromolecular crowding for protein assembly, *Curr. Opin. Struct. Biol.* 10 (2000) 34–39.
- [360] M.S. Cheung, D. Klimov, D. Thirumalai, Molecular crowding enhances native state stability and refolding rates of globular proteins, *Curr. Opin. Struct. Biol.* 102 (2005) 4753–4758.
- [361] G.-W. Li, O.G. Berg, J. Elf, Effects of macromolecular crowding and DNA looping on gene regulation kinetics, *Nature Phys.* 5 (2009) 294–297.
- [362] H.X. Zhou, G. Rivas, A.P. Minton, Macromolecular crowding and confinement: Biochemical, biophysical, and potential physiological consequences, *Ann. Rev. Biophys.* 37 (2008) 375–397.
- [363] J.M. Yuan, C.L. Chyan, H.X. Zhou, T.Y. Chung, H.B. Peng, G.H. Ping, G.L. Yang, The effects of macromolecular crowding on the mechanical stability of protein molecules, *Protein Sci.* 17 (2008) 2156–2166.
- [364] D.L. Pincus, D. Thirumalai, Crowding effects on the mechanical stability and unfolding pathways of ubiquitin, *J. Phys. Chem. B* 113 (2009) 359–368.
- [365] A.R. Singh, D. Giri, S. Kumar, Force induced unfolding of bio-polymers in a cellular environment: A model study, *J. Chem. Phys.* 131 (2009) 065103.
- [366] A.R. Singh, D. Giri, S. Kumar, Effects of molecular crowding on stretching of polymers in poor solvent, *Phys. Rev. E* 79 (2009) 051801.
- [367] V. Blavatska, W. Janke, Polymers in crowded environment under stretching force: Globule-coil transitions, *Phys. Rev. E* 80 (2009) 051805.
- [368] G. Lee, K. Abdi, Y. Jiang, P. Michael, V. Bennett, P.E. Marszalek, Nanospring behavior of ankyrin repeats, *Nature* 440 (2006) 246–249.
- [369] S.R.K. Ainavarapu, L. Li, C.L. Badilla, J.M. Fernandez, Ligand binding modulates the mechanical stability of dihydrofolate reductase, *Biophys. J.* 89 (2005) 3337–3344.
- [370] Y. Cao, M.M. Balamurali, D. Sharma, H. Li, A functional single-molecule binding assay via force spectroscopy, *Proc. Natl. Acad. Sci. USA* 104 (2007) 15677–15681.
- [371] A.P. Wiita, R. Perez-Jimenez, K.A. Walther, F. Gräter, B.J. Berne, A. Holmgren, J.M. Sanchez-Ruiz, J.M. Fernandez, Probing the chemistry of thioredoxin catalysis with force, *Nature* 450 (2007) 124–127.
- [372] S.R.K. Ainavarapu, J. Brujic, H.H. Huang, H.L.A.P. Wiita, L. Li, K.A. Walther, M. Carrion-Vazquez, H. Li, J.M. Fernandez, Contour length and refolding rate of a small protein controlled by engineered disulfide bonds, *Biophys. J.* 92 (2007) 225–233.
- [373] A.P. Wiita, S.R.K. Ainavarapu, H.H. Huang, J.M. Fernandez, Force-dependent chemical kinetics of disulfide bond reduction observed with single-molecule techniques, *Proc. Natl. Acad. Sci. USA* 103 (2006) 7222–7227.
- [374] N. Bhasin, P. Carl, S. Harper, G. Feng, H. Lu, D.W. Speicher, D.E. Discher, Chemistry on a single protein, vascular cell adhesion molecule-1 during forced unfolding, *J. Biol. Chem.* 279 (2004) 45865–45874.
- [375] P. Carl, C.H. Kwok, G. Manderson, D.W. Speicher, D.E. Discher, Forced unfolding modulated by disulfide bonds in the Ig domains of a cell adhesion molecule, *Proc. Natl. Acad. Sci. USA* 98 (2001) 1565–1570.
- [376] F. Oesterhelt, D. Oesterhelt, M. Pfeiffer, A. Engel, H.E. Gaub, D.J. Muller, Unfolding pathways of individual bacteriorhodopsins, *Science* 288 (2000) 143–146.
- [377] H. Janovjak, M. Kessler, D. Oesterhelt, H.E. Gaub, D.J. Muller, Unfolding pathways of native bacteriorhodopsin depend on temperature, *Embo J.* 22 (2003) 5220–5229.
- [378] G. Yang, C. Cecconi, W.A. Baase, I.R. Vetter, W.A. Breyer, J.A. Haack, B.W. Matthews, F.W. Dahlquist, C. Bustamante, Solid-state synthesis and mechanical unfolding of polymers of T4 lysozyme, *Proc. Natl. Acad. Sci. USA* 97 (2000) 139–144.
- [379] C.G. Baumann, V.A. Bloomfield, S.B. Smith, C. Bustamante, M.D. Wang, S.M. Block, Stretching of single collapsed DNA molecules, *Biophys. J.* 78 (2000) 1965–1978.
- [380] H. Li, J.M. Fernandez, Mechanical design of the first promixal Ig domain of human cardiac titin revealed by single molecules force spectroscopy, *J. Mol. Biol.* 334 (2003) 75–86.
- [381] H. Li, A.F. Obserhauser, S.B. Fowler, J. Clarke, J.M. Fernandez, Atomic force microscopy reveals the mechanical design of a modular protein, *Proc. Natl. Acad. Sci. USA* 97 (2000) 6527–6531.
- [382] M. Rief, M. Gautel, A. Schemmel, H.E. Gaub, The mechanical stability of immunoglobulin and fibronectin III domains in the muscle protein titin measured by atomic force microscopy, *Biophys. J.* 75 (1998) 3008–3014.
- [383] Y. Oberdorfer, H. Fuchs, A. Janshoff, Conformational analysis of native fibronectin by means of force spectroscopy, *Langmuir* 16 (2000) 9955–9958.
- [384] A. Oberhauser, C. Badilla-Fernandez, M. Carrion-Vazquez, J.M. Fernandez, The mechanical hierarchies of fibronectin observed with single-molecule AFM, *J. Mol. Biol.* 319 (2002) 433–447.
- [385] M. Schlierf, M. Rief, Temperature softening of a protein in single-molecules experiments, *J. Mol. Biol.* 354 (2005) 497–503.
- [386] K. Watanabe, C. Mühle, M.S.Z. Kellermayer, S. Labeit, H. Granzier, Different molecular mechanics displayed by titin's constitutively and differentially expressed tandem Ig segment, *Struct. Biol. J.* 137 (2002) 248–258.
- [387] K. Watanabe, P. Nair, D. Labeit, M.S.Z. Kellermayer, M. Greaser, S. Labeit, H.L. Granzier, Molecular mechanics of cardiac titins PEVK and n2b spring elements, *J. Biol. Chem.* 277 (2002) 11549–11558.
- [388] R. Law, P. Carl, S. Harper, P. Dalhaimer, D.W. Speicher, D.E. Discher, Cooperativity in force unfolding of tandem spectrin repeats, *Biophys. J.* 84 (2003) 533–544.
- [389] R. Law, P. Carl, S. Harper, D.W. Speicher, D.E. Discher, Influence of lateral association on forced unfolding of antiparallel spectrin heterodimers, *J. Biol. Chem.* 279 (2004) 16410–16416.
- [390] Y.N. Kalia, S.M. Brocklehurst, D.S. Hippis, E. Appella, K. Sakaguchi, R.N. Perham, The high-resolution structure of the peripheral subunit-binding domain of dihydrolipoamide acetyltransferase from the pyruvate dehydrogenase multienzyme complex of bacillus stearothermophilus, *J. Mol. Biol.* 230 (1993) 323–341.
- [391] M.J. Osborne, A.L. Breeze, L.Y. Lian, A. Reilly, R. James, C. Kleanthous, G.R. Moore, Three-dimensional solution structure and <sup>13</sup>C nuclear magnetic resonance assignments of the colicin E9 immunity protein Im9, *Biochemistry* 35 (1996) 9505–9512.
- [392] G.V. Louie, G.D. Brayer, High-resolution refinement of yeast iso-1-cytochrome c and comparisons with other eukaryotic cytochromes c, *J. Mol. Biol.* 214 (1990) 527–555.
- [393] G. Bushnell, G.V. Louie, G.D. Brayer, High-resolution three-dimensional structure of horse heart cytochrome c, *J. Mol. Biol.* 214 (1990) 585–595.
- [394] K. Andersen, F. Poulsen, The three-dimensional structure of acyl-coenzyme a binding protein from bovine liver: Structural refinement using heteronuclear multidimensional NMR spectroscopy, *J. Biomol. NMR* 3 (1993) 271–284.
- [395] F. Lederer, A. Glatigny, P. Bethge, H. Bellamy, F. Matthew, Improvement of the 2.5 Å resolution model of cytochrome b\_562 by redetermining the primary structure and using molecular graphics, *J. Mol. Biol.* 148 (1981) 427–448.
- [396] L. Beamer, C. Pabo, Refined 1.8 Å crystal structure of the lambda repressor-operator complex, *J. Mol. Biol.* 227 (1992) 177–196.
- [397] M. Brunori, F. Cutruzzola, C. Savino, C. Travagliani-Allocatelli, B. Vallone, Q. Gibson, Structural dynamics of ligand diffusion in the protein matrix: A study on a new myoglobin mutant Y(B10)Q(E7)R(E10), *Biophys. J.* 76 (1999) 1259–1269.
- [398] J. Liphardt, B. Onoa, S.B. Smith, L. Tinoco, C. Bustamante, Reversible unfolding of single RNA molecules by mechanical force., *Science* 292 (2001) 733–737.
- [399] H. Lu, K. Schulten, Steered molecular dynamics simulations of force-induced protein domain unfolding, *Proteins: Struc. Fun. Bioinform.* 35 (1999) 453–463.
- [400] M. Gao, M. Wilmanns, K. Schulten, Steered molecular dynamics studied of titin i1 domain unfolding, *Biophys. J.* 83 (2002) 3435–3445.
- [401] H. Lu, K. Schulten, Steered molecular dynamics simulations of conformational changes of immunoglobulin domain I27 interpret atomic force microscopy observation, *Chem. Phys.* 247 (1999) 141–153.
- [402] H. Lu, K. Schulten, The key event in force-induced unfolding of titin's immunoglobulin domain., *Biophys. J.* 79 (2000) 51–65.
- [403] A. Krammer, H. Lu, B. Isralewitz, K. Schulten, V. Vogel, Force unfolding of the fibronectin type iii module reveals a tensile molecular recognition switch, *Proc. Natl. Acad. Sci. USA* 96 (1999) 1351–1356.

- [404] M. Gao, D. Craig, O. Lequin, I.D. Campbell, V. Vogel, K. Schulten, Structure and functional significance of mechanically unfolded fibronectin type iii, Proc. Natl. Acad. Sci. USA 100 (2003) 14784–14789.
- [405] D. Craig, A. Krammer, K. Schulten, V. Vogel, Comparison of the early stages of forced unfolding for fibronectin type iii modules, Proc. Natl. Acad. Sci. USA 98 (2001) 5590–5595.
- [406] M. Kouza, C.K. Hu, H. Zung, M.S. Li, Protein mechanical unfolding: Importance of non-native interactions, J. Chem. Phys. 131 (2009) 215103.
- [407] M. Sotomayor, D.P. Corey, K. Schulten, In search of the hair-cell gating spring: Elastic properties of ankyrin and cadherin repeats, Structure 13 (2005) 669–682.
- [408] M. Wang, Y. Cao, H. Li, The unfolding and folding of TNfnALL probebd by single molecule force-ramp spectroscopy, Polymer 47 (2006) 2548–2554.
- [409] L. Li, H. Huang, C.L. Badilla, J.M. Fernandez, Mechanical unfolding intermediates observed by single-molecule force spectroscopy in a fibronectin type iii module, J. Mol. Biol. 345 (2005) 817–826.
- [410] J.M. Fernandez, H. Li, Force-clamp spectroscopy monitors the folding trajectory of a single protein, Science 303 (2004) 1674–1678.
- [411] M. Cieplak, P. Szemczak, Protein folding in a force clamp, J. Chem. Phys. 124 (2006) 194901–194904.
- [412] M.S. Li, C.K. Hu, D.K. Klimov, D. Thirumalai, Multiple stepwise refolding of immunoglobulin domain I27 upon force quench depends on initial conditions, Proc. Natl. Acad. Sci. USA 103 (2006) 93–98.
- [413] C. Tanford, Chemical potential of bound ligand, an important parameter for free energy transduction, Proc. Natl. Acad. Sci. USA 78 (1981) 270–273.
- [414] P. Geissler, E. Shakhnovich, Mechanical response of random heteropolymers, Macromolecules 35 (2002) 4429–4436.
- [415] H.M. Went, S.E. Jackson, Ubiquitin folds through a highly polarized transition state, Protein Eng. Des. Sel. 18 (2005) 229–237.
- [416] B.A. Krantz, R.S. Dothager, T.R. Sosnick, Discerning the structure and energy of multiple transition states in protein folding using  $\psi$ -analysis, J. Mol. Biol. 337 (2004) 463–475.
- [417] T.R. Sosnick, B.A. Kratz, R.S. Dothager, M. Baxa, Characterizing the protein folding transition state using  $\psi$  analysis, Chem. Rev. 106 (2006) 1862–1876.
- [418] H.S. Chung, M. Khalil, A.W. Smith, Z. Ganim, A. Tomakoff, Conformational changes during the nanosecond-to-millisecond unfolding of ubiquitin, Proc. Natl. Acad. Sci. USA 102 (2005) 612–617.
- [419] D. Alonso, V. Daggett, Molecular dynamics simulations of hydrophobic collapse of ubiquitin, Protein Sci. 7 (1998) 860–874.
- [420] N.J. Marianayagam, S.E. Jackson, The folding pathway of ubiquitin from all-atom molecular dynamics simulations, Biophys. Chem. 111 (2004) 159–171.
- [421] A. Fernandez, Conformation-dependent environments in folding proteins, J. Phys. Chem. 114 (2001) 2489–2502.
- [422] A. Fernandez, Time-resolved backbone desolvation and mutational hot spots in folding proteins, Proteins 47 (2002) 447–457.
- [423] A. Fernandez, A. Colubri, R.S. Berry, Three-body correlations in protein folding: The origin of cooperativity, Physica A 307 (2002) 235–259.
- [424] D. Gilis, M. Roonan, Identification and ab initio simulations of early folding units in proteins, Proteins: Struct. Funct. Bioinform. 42 (2001) 164–176.
- [425] A. Irbac, S. Mitternacht, Thermal versus mechanical unfolding of ubiquitin, Proteins: Struct. Func. Bioinform. 65 (2006) 759–766.
- [426] S. Kumar, Can reentrance be observed in force induced transitions, Europhys. Lett. 85 (2009) 38003.
- [427] D. Baker, A surprising simplicity to protein folding, Nature 405 (2000) 39–42.
- [428] M.S. Li, M. Kouza, Dependence of protein mechanical unfolding pathways on pulling speeds, J. Chem. Phys. 130 (2009) 145102.
- [429] C.R. Hickenboth, J.S. Moore, S.R. White, N.R. Sottos, J. Baudry, S.R. Wilson, Biasing reaction pathways with mechanical force, Nature 446 (2007) 423–427.
- [430] X. Trepat, L. Deng, S.S. An, D. Navajas, D.J. Tschumperlin, W.T. Gerthoffer, J.P. Butler, J.J. Fredberg, Universal physical responses to stretch in the living cell, Nature 447 (2007) 592–596.
- [431] F. Chiti, C.M. Dobson, Protein misfolding, functional amyloid and human disease, Annu. Rev. Biochem. 75 (2006) 333–366.
- [432] O.N. Antzutkin, J.J. Balbach, R.D. Leapman, N.W. Rizzo, J. Reed, R. Tycko, Multiple quantum solid-state NMR indicates a parallel, not antiparallel, organization of  $\beta$ -sheets in alzheimer's  $\beta$ -amyloid fibrils, Proc. Natl. Acad. Sci. USA 97 (2000) 13045–13050.
- [433] C. Ray, B.B. Akhremitchev, Conformational heterogeneity of surface-grafted amyloidogenic fragments of  $\alpha$ -synuclein dimers detected by atomic force microscopy, J. Am. Chem. Soc. 127 (2005) 14739–14744.
- [434] S. Guo, B.B. Akhremitchev, Packing density and structural heterogeneity of insulin amyloid fibrils measured by AFM nanoindentation, Biomacromolecules 7 (2006) 1630–1636.
- [435] M.S.Z. Kellermayer, L. Grama, A. Karsai, A. Nagy, A. Kaln, Z.L. Datki, B. Penke, Reversible mechanical unzipping of amyloid  $\beta$ -fibrils, J. Biol. Chem. 280 (2005) 8464–8470.
- [436] C. McAllister, M.A. Karymov, Y. Kawano, A.Y. Lushnikov, A. Mikheikin, V.N. Uversky, Y.L. Lyubchenko, Protein interactions and misfolding analysed by AFM, J. Mol. Biol. 354 (2005) 1028–1042.
- [437] A.V. Kranoleslobodtsev, L.S. Shlyakhtenko, E. Ukrainstsev, T.O. Zaikova, J.F.W. Keana, Y.L. Lyubchenko, Nanomedicine and protein misfolding diseases, Nanomedicine 1 (2005) 300–305.
- [438] J.F. Smith, T.P.J. Knowles, C.M. Dobson, C.E. MacPhee, M.E. Welland, Characterization of the nanoscale properties of individual amyloid fibrils, Proc. Natl. Acad. Sci. USA 103 (2006) 15806–15811.
- [439] A. Karsai, Z. Martonfalvi, A. Nagy, L. Grama, B. Penke, M.S.Z. Kellermayer, Mechanical manipulation of alzheimer's amyloid  $\beta$ 1 – 42 fibrils, J. Struct. Biol. 155 (2006) 316–326.
- [440] E.P. Raman, T. Takeda, V. Barsegov, D.K. Klimov, Mechanical unbinding of A  $\beta$  peptides from amyloid fibrils, J. Mol. Biol. 373 (2007) 785–800.