Characterization of Protein based Spring-like Elastic Joints for Biorobotic Applications

Mustapha Hamdi 1, Gaurav Sharma 2, Antoine Ferreira 1, Constantinos Mavroidis 2

1Laboratoire Vision et Robotique, Université d’Orléans, 18000, Bourges, France.
2Computational BioNano Robotics Laboratory, Northeastern University, Boston, MA-02120, USA

antoine.ferreira@ensi-bourges.fr, mavro@coe.neu.edu

Abstract—This paper presents a molecular mechanics study using a molecular dynamics software (NAMD) for characterization of molecular elastic joints for bio nanorobotic prototyping. Simple protein-like elastic joints elements for structural links have been carefully studied and simulated to understand the functional limits of each one of them: fibronectin, titin, deca-alanine, fibrillin, Rop and double-helical DNA molecules. We simulated the restoring forces involved under various external mechanical stress to predict the type of force spectra, reversibility, degrees of freedom and irreversible work that may be expected from single-molecule protein manipulation experiments. Their use as elementary molecular components for bio-nanorobotic structures are also simulated and the results discussed.

I. INTRODUCTION

Bio-nanorobotics is an emerging area of scientific and technological opportunity. It is a new and rapidly growing interdisciplinary field addressing the assembly, construction and utilization of biomolecular devices based on nanoscale principles and/or dimensions. Research and product development at the interface of physical sciences and biology as applied to this area require multi-skilled teams and often novel technical approaches for material synthesis, characterization and applications. In this way proteins and DNA could act as motors, mechanical joints, transmission elements, or sensors. If all of these different components were assembled together they can form bio-nanorobots with multiple degrees-of-freedom that are able to apply forces and manipulate objects in the nanoscale world. These motors, which are called biomolecular motors, have attracted a great deal of attention recently because they have high efficiency, they could be self-replicating, and hence cheaper in mass usage, and they are readily available in nature. A number of enzymes such as kinesin [1], RNA polymerase [2], myosin [3], dynein [4], adenosine triphosphate (ATP) synthase [5], viral protein linear (VPL) motor [6] and DNA can function as nanoscale biological motors. In addition, there are compliance devices such as spring-like proteins called fibronectin [7] and vorticellids [8], as well as synthetic contractile plant polymers [9] which can act as compliant joints in biomolecular robotic systems.

Contribution from molecular dynamic (MD) simulations is important in order to be able to understand the bi-nanomechanics of proteins and develop dynamic and kinematic models to study their performances. The ability to visualize the atom-to-atom interaction in real-time and see the results in a fully immersive 3-D environment is an additional feature of such simulations.

In this work, we consider a mechanical molecular study of unfolding proteins acting as passive joints in bio-nanorobotic systems using a molecular dynamics software (NAMD). We therefore decided to begin our investigations by simulating the forces involved under various external mechanical stress (stretching, contraction, shearing, bending) to predict the type of force spectra, reversibility, degrees of freedom and irreversible work that may be expected from single-molecule protein manipulation experiments.

The paper is organized as follows: Section II presents mechanical molecular simulations of various proteins: deca-alanine, Rop, fibronectin, immunoglobulin, fibrillin and double stranded DNA as passive links before to make a discussion about the current developments in bio-nanorobotics field in Section IV.

II. MECHANICAL SIMULATION OF BIOLOGICAL SPRINGS USING MOLECULAR DYNAMIC SIMULATION

A. Molecular Dynamic Simulations through NAMD

Biological springs can be classified as active mechano-chemical devices that store the energy of conformation from various bonded and non-bonded interactions as potential energy. To characterize the mechanochemical energy conversion and to study the elasto-mechanical properties of these biological springs external forces are applied using molecular mechanics technique. Figure 1 shows the structure of some of the molecules whose mechanical properties have been simulated through a molecular dynamic program NAMD [10] for protein structure prediction. These molecular structures are derived from X-ray diffraction and NMR data. Unfortunately, MD simulations are computer time- and memory-intensive and, currently, simulations of ten require...
several hours or days to complete. NAMD software is a parallel molecular dynamics program designed for high-performance simulations in structural biology and uses a spatial decomposition parallelization strategy coupled with a multi-threaded computation model. The results of NAMD simulations are displayed on a suitable molecular visualization program called VMD [11] for interactive display of molecular systems. A force feedback interface system has been developed [12] in order to allow the user to virtually steer the protein molecule in the space of the precomputed conformations for a mechanical characterization of the protein forces (stretching, shearing or bending).

**B. α–Helix Deca-Alanine Protein**

We applied the techniques described above to an exemplary system: helix-to-coil transition of deca-alanine in vacuum and in solvent. The well-known deca-alanine is an oligopeptide composed of ten alanine residues. In vacuum at room temperature ($T=300^\circ$K), the stable configuration of deca-alanine is a α-helix. Stretching the molecule by an external force can induce its transition to an extended (coiled) form. This helix-to-coil transition represents a simple but basic folding system acting as a passive like spring. In the simulation, we fix one end of the molecule (the main chain nitrogen atom of the first residue) at the origin and constrain the other end (the capping nitrogen atom of the C-terminus) to move only along the $z$ axis, thereby removing the irrelevant degrees of freedom, i.e., removing the overall translation and rotation. For this, a guiding potential [13]

$$u(r; \lambda) = \left( k / 2 \right) \left( \xi(r) - \lambda \right)^2$$

is added to control the end-to-end distance $\xi$ which is a function of the $3N$-dimensional position $r$ of the system. The moving guiding potential used in the pulling simulations is represented by a spring which is connected to the C-terminus and pulled with a constant velocity $v$. The parameter $\lambda$ is varied from 15 to 35 Å and $v$ is constant with a speed of 0.1 Å/ns. A force constant of $k = 500$ pN/Å is used in order to allow end-to-end distance $\xi$ closely follow the constraint center $\lambda$. Figure 2a shows the maximal force (before breaking) as a function of stretching distance. Depending on the sign of $v$, the external work can be defined for either stretching or contracting motion of the protein. The results shows two distinct conformational transitions that provoke the conversion of α-helix to an extended form approaching the coil conformation. The force increases rather smoothly to almost 10 pN up to a relative length of the protein. Once this transition is completed, the hydrogen bonds start to break and the structure further extends toward the coil conformation (intrinsic elastic regime), as shown in Part II, of the force curve (~700 pN). At the lowest forces, the molecule behaves as a Hookian spring and its extension is proportional to the force applied at the end with a reversible motion. A useful approximation for spring constant $k_{stret}$ is given by the inextensible worm-like chain (WLC) model [14]. The WLC model of entropic elasticity predicts the relationship between the relative extension of a polymer ($z/L$) and the entropic restoring force ($f$) through

$$f = k_{stret} z/L + (1/4)(1-z/L)^2 - 1/4.$$  (2)

where $k_B$ is the Boltzmann constant, $T$ is absolute temperature, $A$ is the persistence length, $z$ is the end-to-end length, and $L$ is the length. Then, if the stretching force increases gradually until it reaches its stretching limit (~3600 pN) with an irreversible motion in Part III. Fig.2b shows the nonlinear reversibility of the protein when relaxed. The previous results showed that the native structure is not destroyed under normal physiological conditions. In some unnatural conformations, lateral shearing and bending forces applied on a protein molecule has been simulated: fixed-free boundary conditions and fixed-fixed boundary conditions. These tests simulate disturbances of the bio-nanorobotic component under various operating conditions. For these conformational, Fig.3 presents the lateral forces obtained. As shown in Fig.3a, the lateral Hookian spring $k_{shear}$ has high stiff spring value that is able to counteract microenvironment variations and mechanical disturbances of a bio-nanorobotic platform. The extension of the Hookian spring is proportional to the force applied at its end such as:
\[ F = \frac{k_B T}{A} \left( \frac{\chi}{L} \right) \]  

In this case, in contrast to stretching described above, the force variation is roughly monotonic with different plateaus leading to different Hookian spring values. Shearing is largely limited to breaking hydrogen bonds as there is little conformational change in the extended peptide backbones. Conversely, Fig. 3b shows a linear variation of force behavior when considering pure bending deformation of the protein. It shows that for little length deformation, it requires low constant force. Finally, these results indicate that it may be possible to obtain uncoupled mechanical spring behavior of the protein: stiff in lateral and compliant in longitudinal directions. Furthermore, under normal operating conditions, deca-alanine protein shows a reversibility of displacement-force characteristics which allows its use as spring-like joint in bio-nanorobotic platforms. Simulations also showed that the serial/parallel connection of deca-alanine proteins permits to augment considerably the resulting spring-force and to decrease the overall displacement.

**C. ROP Protein**

The repressor of primer protein (ROP) is a small, dimeric molecule consisting of two identical chains of 63 amino acids. Each monomer consists of two \( \alpha \)-helices connected by a short turn and a seven-residue C-terminal tail. The two monomers pack together as a fully antiparallel four helix-bundle. The bend region of Rop has attracted considerable interest as a parallel molecular spring due to its stability and elasticity properties [15]. In the simulation, we constrain both ends of two \( \alpha \)-helices of the molecule to move only along the \( z \) axis for stretching simulations and fix the short turn.

**D. Giant Titin Protein**

Titin, a 1-\( \mu \)m-long protein found in striated muscle myofibrils, possesses unique elastic and extensibility properties. Titin is composed of \( \sim300 \) repeats of two types of domains, fibronectin type III-like (Fn-III) domains and immunoglobulin-like (Ig) domains.

- **Immunoglobulin-like (Ig) domains.** The behavior of titin as a serial link entropic spring depends on the reversible unfolding of individual Ig domains. Currently, I27 is the only I-band Ig with an experimentally solved structure [16] and hence has been selected for investigation. The Ig domain was placed in the center of the water box and equilibrated with a thermal bath of 300°C. Stretching simulations were carried out by fixing one terminus of the domain and applying external forces to the other terminus. The simulation began...
with an equilibrated folded structure and was stopped when a fully extended polypeptide was obtained. The extension of I27 was performed with a pulling speed \( v = 0.5 \, \text{Å/ns} \) and was stopped when the extension reached 33nm. The extension domain (Fig.7a) is divided into four sections: I. preburst at extension of 4nm during which the protein maintains \( \beta \)-sheet structure and the external force remains smaller than 1500 pN; II. major burst immediately after the preburst burst at extension of 8nm; III. post-burst at extension of 27 nm during which the protein unravels; IV. pulling of fully extended chain up-to an extension of 33nm.

Other simulations showed similar features of the unfolding process and force profiles with only small variations in force peak value and degree of extension at the force peak. Results of stretching-relaxation curve (Fig.7b) show a good reversibility of the protein motion when completely relaxed. Fig.7c suggests a mechanical portrait model. Its behavior during extension might be modeled as series of elastic spring with a viscous element corresponding to the unfolding of the individual I27 domain. Stretch would result first in straightening of the Ig domain chain (corresponding to the preburst-part I) as an entropic spring. The tightly folded Ig domain might function as a “shock absorber” (parts II: major burst and III: post-burst) by reversible unfolding only in the case of extremely high stretching forces. This structure allows avoiding the complete rupture of the protein due to overstretching. When subjected to bending or shearing forces, the Ig27 domain behaves as a stiff hookian spring. As example of serial Ig configuration, Fig.8a shows the typical force-extension curve by stretching two immunoglobulin-like proteins (I1 and I27-Ig modules). The simulation parameters are \( v = 10 \, \text{Å/ns}, T=300^\circ \) and a force constant \( k = 500 \, \text{pN/Å} \).

\[ \begin{align*}
\text{Fig.7: Ig 27 stretching simulations. (a) Force-distance curve; (b) stretching-relaxation curve for reversibility; (c) mechanical model of the molecular spring I27-Ig module.}
\end{align*} \]

The initial part of the force-extension curve is fitted with the WLC model to obtain the entropic spring of I27-Ig module. After the unfolding of the I27-Ig module, the second I1-Ig module unfolds and can be also fitted by a WLC model. The extension of (I27-I1) domains can be represented by an elastic spring. When considering several Ig modules connected in series, the force-distance curve behaves as a sequential unfolding of the Ig modules. The number of force peaks arising in the profiles is equal to the number of Ig domains involved in the stretched protein. A portrait of the molecular springs is shown in Fig.8b.

- **Fibronectin-like domains**: It is noteworthy that Ig and Fn-III constitute \( \beta \) sandwich domains with N-terminal and C-terminal strands parallel to each other, but pointing in opposite direction [17]. FnIII10 was stretched from its initially compact and folded structure to its fully elongated configuration at a pulling speed of 1 Å/ns, \( T=300^\circ \) and a force constant \( k = 700 \, \text{pN/Å} \). A detailed analysis of the force-extension curve show that the first force peak corresponds to the main energy barriers which separate folded and the unfolded states of the domain (2177pN) at an extension of 5.8nm. After the barrier is overcome, the force decreases quickly to average values of 750pN until the second force peak appears corresponding to the second Fn-III10 module. Then, the force averages 750nm until the second module is completely folded. The comparison between Ig/Fn-III modules shows similar force-extension curves and reversibility behavior. The elastic mechanical model is similar to the one proposed for Ig domains.

\[ \begin{align*}
\text{E. Fibrillin protein}
\end{align*} \]

Fibrillins form the structural framework of an essential class of extracellular microfibrils that endow dynamic
connective tissues with long-range elasticity (skin, lung, eye) [18]. Mechanical extension induces a conformational change with an apparent decrease in randomly coiled regions of the protein and a relative increase in α-helical regions. As we can see in Fig.9, a series of three unfolding events that results ultimately in the generation of a stabilized tensioned polymer.

Upon relaxation it appears that the protein spontaneously refolds to the original state, indicated by a return to the native state. WLC model fails in this case for small amount of stretching forces. It can be represented by a spring-stiff pair. For lengths below the length \( l_1 \), spring constant of the protein is \( k_1 \), whereas for lengths greater than \( l_1 \), spring constant is \( k_2 \). As the tension increases, the protein extends according to the unfolding events until to reach the length \( l_1 \) (as a spring component), and then the folded protein extends according to its stiffness (as a stiff component). When subjected to external mechanical bending force, the fibrillin protein behaves as a linearly elastic stiff spring. Such proteins are very interesting in compliant molecular links constituted by fibrillin-rich microfibrills (Fig.10a). In this structure, the individual microfibrills acts as relatively stiff elastic polymers. Based on simulation studies (with \( v=1 \) Å/ns and \( k=10 \) pN/Å), it was concluded that individual isolated microfibrills were reversibly extensible (see Fig.10b) although the mechanism of this elasticity is unknown. The mechanical elastic model can be represented as a series model of spring-string components (Fig.10c).

**F. Double Strand DNA Protein**

Chemically, DNA is a long polymer made up of a linear series of subunits known as nucleotides. Structurally, DNA is usually found as a double helix, with two strands wrapped around one another. However, DNA can adopt other configurations and can also exist in single-stranded forms. Double-stranded DNA (dsDNA) has sparked the renewed interest in the force versus extension of polymers for biomolecular springs. The DNA is solvated in water with 30 Na\(^+\) ions added to neutralize the charge. The water-DNA system was gradually heated over 7ps to 300K, and then equilibrated with a thermal bath at 300K for another 7ps. The dynamics of the DNA were performed with the double-stranded terminations stretched and the other end fixed. This elastic behavior is thus purely entropic [19].

For very low tension \( f \leq 1 \) pN, the restoring force is provided by "entropic elasticity". In the absence of any force applied to its ends, the DNA's RMS end-to-end distance (chain length, \( L \)) is small compared to its contour length defined as the maximum end-to-end distance (maximum length, \( L_0 \)) and the chain enjoys a large degree of conformational disorder. Stretching DNA reduces its entropy and increases the free energy. The corresponding force \( f \) increases linearly as a Hooke’s law with the extension \( L \):

\[
f \equiv \frac{3k_BT}{A_{\text{DNA}}} \frac{L}{L_0}, \quad L \ll L_0 \quad (4)
\]

The length \( A_{\text{DNA}} \) is known as the "thermal persistence length" of DNA and is of the order 50nm. For higher forces \( (f \geq 10 \) pN), the end-to-end distance \( L \) is close to \( L_0 \) and the elastic restoring force is due to distortion of the internal structure of DNA. In this regime, the force extension curve can be approximated by two models that are often used to describe the entropic elasticity of DNA. In the freely jointed chain (FJC) model, the molecule is made up of rigid, orientationally independent Kuhn segments whose length, \( b \), is a measure of chain stiffness. The resulting entropic elastic behavior can be summarized in the force-extension relation:

\[
\frac{z}{F_{\text{tot}}} = \coth \left( \frac{fb}{k_BT} \right) - \frac{f_b}{fb} \quad (5)
\]
defining the well-known Langevin function. Expanding Eq.(5) gives the effective spring constant for the FJC as $k^{FJC} = 3k_BT/b^4$. The terms $L_{tot}$ represent the total length of the protein and $f$ the stretching force. The alignment of segments by tension is described by Boltzmann distribution. In the inextensible worm-like chain model (WLC) model, the molecule is treated as a flexible rod of length $L$ that curves smoothly as a result of thermal fluctuation. Results, shown in Fig.11a, indicate that, even though the FJC model can describe the behavior of dsDNA in the limit of low and intermediate forces, it fails at high forces. The WLC model, on the other hand, provides an excellent description of molecular elasticity at intermediate and high forces. Both models behave as a Hooke’s law for low stretching forces. The dsDNA is fully reversible as shown in Fig.11b. When considering an anti-parallel shearing of dsDNA (Fig.11c), a mechanical unzipping of the 12 hairpin DNA occurs sequentially (Part I) after a phase of extension of dsDNA (Part II) before the rupture point (Part III). As it is shown, the strand unzipping occurred abruptly at 500pN and displayed a reproducible “saw-tooth” force variation pattern with an amplitude of +/-10pN along the DNA.

III. CONCLUSION

Development of bio-nano components from biological systems is the first step towards the design and development of an advanced bio-nanorobot. Simple protein-like elastic joints elements for structural links have be carefully studied and simulated to understand the functional limits of each one of them. This paper presents a molecular mechanics study carried out using molecular dynamics simulation techniques for the characterization of various protein based spring-like elastic joints. Our future work will be oriented to the experimental characterization through an AFM-based nanomanipulation system of the proposed elastic proteins.

REFERENCES