Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines

Atul Dubey, M. Silvina Tomassone
Department of Chemical and Biochemical Engineering, Rutgers, The State University of New Jersey, Piscataway, USA

Article Outline

Glossary

Nanodevice A Device with size dimensions of the order of the nanometer (10^{-9} m)

Actuator A device used to generate motion by converting some form of energy (such as chemical or light energy) into mechanical energy.

ATP Adenosine Triphosphate molecule used by living organisms to store and transport energy.

End-effector The last link in a robotic assembly that performs the desired tasks by the robot by interacting with or acting upon the elements outside the robot.

Retrovirus A type of virus that has a protein envelope around it, stores its genetic information in an RNA molecule instead of double stranded DNA and transcripts its genome into DNA which integrates into the chromosome of the infected cell thereby replicating the virus when the cell replicates.

Glycoprotein A protein molecule which has a carbohydrate attached to it.

Membrane fusion A process when viral and hose cellular membranes fuse together during infection.

Hemagglutinin A type of protein on the surface of the influenza virus.

Fusogenic The unit that facilitates or takes part in membrane fusion.

Disulfide bond A sulfur–sulfur atomic bond which typically joins two biological segments such as protein chains.

Proteolytic cleavage A process of breaking the peptide bond between amino acids in a protein chain.

Endosome Intracellular vesicle formed by the outer cell membrane used to transport material inside the animal cell.

Endocytosis The process of encapsulation of an external element by the cell membrane and subsequent formation of an endosome for transport within the cell.

Nucleocapsid The genome of a virus enclosed in a protein coat.

α-Helix A secondary structure of protein molecules characterized by a helical shape and a specific pitch and stabilized by hydrogen bonds along the length of the helix.

β-Sheet A secondary structure of protein molecules where two or more chains are interconnected by hydrogen bonds forming a sheet-like structure.

Titin A big filamentous protein used by muscles.

Holonomic constraint A constraint that puts a restriction on the motion of a system by the use of integrable system of differential equations.

Chymotrypsin An enzyme in the pancreas which is involved in the catalysis of hydrolysis of proteins in the small intestine.

PDB The protein data bank which contains structural and other information about a large number of protein molecules. http://www.rcsb.org/pdb/home/home.do

Implicit solvation A technique used in molecular simulation which allows for a potential function to be used in place of explicit solvent molecules surrounding the molecule of interest.

Protonation The process of accepting a positively charged hydrogen ion or a proton by an amino acid.

Helicity The measure of the helical content of a given peptide molecule.

Denaturation Process of loss of secondary structure of a peptide molecule, which can happen for example due to heating beyond a certain level.

Definition of the Subject

In this paper, we describe computational aspects of an initiative to create a nanodevice for chemo-mechanical energy conversion based on a biological system. We provide insights into the available methods and their applications to gain knowledge about whether a given molecule is suitable to function as a nanoactuator. The field of bio-nanotechnology is new and it is developing fast. Noticeably, its roots go back decades of studies in the fields of biophysics, biochemistry and engineering. These disciplines can now

come together to create useful and controllable devices at the nano-scale thereby adding another dimension to human scientific capabilities. Their applications in the fields of medicine, space exploration and colonization, and military are very promising.

Introduction

There is a persistent need of miniaturization of machines and energy conversion devices for various engineering applications. A great part of modern research is shifting from macro to nano-scale as consumer and scientific devices become smaller, faster and more energy efficient. In the field of robotics, there is a need to develop nano-scale actuators, joints, motors and other machine components. Since these nanodevices can be either single-molecule or macromolecular devices, the laws governing their behavior are different from the ones that govern macro-scale machines. Hence in order to study and develop them it is crucial to fully understand the material properties and the work environment and be able to create appropriate models that account for them.

Miniature molecular machines have been operating for centuries in living organisms. Nature has created and optimized these complex machines to carry out specific tasks, such as moving cellular cargo, and oxidizing high-energy molecules etc. With modern microscopic abilities, finer structural and functional details of these molecules are being discovered. Mechanical motions, structural features, operating principles, limitations, and design features of many natural molecular machines are being investigated. It is possible to use these machines or their individual parts under artificial stimuli in order to perform tasks not necessarily assigned to them by nature. Attempts at creating artificial molecular machines are also being made. Some of the principles and applications of molecular machines have been reviewed [4,5,7,26,32,38,47].

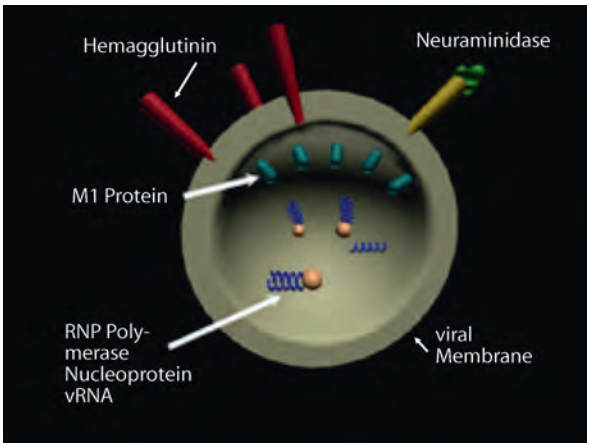
Molecular machines can be broadly classified into protein-based, DNA-based or chemical machines. From each of the three classes, completely different types of machines are obtained. Each of these machines has a different operating environment, force and displacement capabilities, and fuel requirements. Hence all of them figure in a molecular part-list and can be utilized when requirements are specific to them. For example protein-based ATP motors require protons or high energy molecules, whereas DNA-based machines require DNA strands as fuel. Both these machines can be used to produce rotary motion, although in different environmental conditions.

Molecular devices for use in robotics have not yet been thoroughly investigated. There is a need to find nanoscale equivalents of robot parts such as actuators, links, end-effectors, sensors etc. This study aims to develop one such protein based linear actuator, called the Viral Protein Linear (VPL) actuator. The inspiration for this type of actuator comes from nature – from the infection mechanism of some viruses. There is a particular segment of the surface protein of retroviruses such as influenza virus that undergoes a conformational change which can be used as a linear actuator producing about 10 nm of linear motion. In this work, we characterize the VPL protein actuator by computationally studying its behavior and predicting its performance.

The VPL Nanoactuator

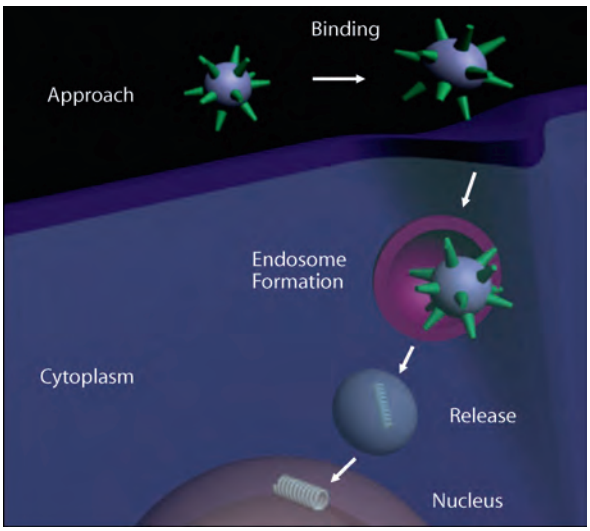
The role of envelope glycoproteins (surface proteins) of various retroviruses in the process of membrane fusion has been investigated and understood over the years. Such viruses infect their target cells by the process of membrane fusion. Membrane fusion is necessary for a large number of diverse processes in biology such as protein trafficking, protein secretion, fertilization, viral invasion and neurotransmission. The mechanism is best understood among enveloped viruses such as the influenza virus [31,35,43,44]. Specialized viral proteins are required to promote membrane fusion – a process which is otherwise very slow. In many cases, these membrane-fusion proteins also serve as agents that promote the binding of the virus onto the cell surface receptors. In the influenza virus, a protein segment called Hemagglutinin (HA) mediates both the binding of the virus to the cell surface and the subsequent fusion of viral and cellular membranes. The receptor binding subunit of HA is termed HA1, while the fusogenic subunit is denoted as HA2. Figure 1 shows a schematic of the influenza virus.

HA1/HA2 complex consists of the disulfide bonded HA1 and HA2 peptides. Each HA monomer is synthesized as a fusion-incompetent precursor polypeptide known as HA0 which undergoes proteolytic cleavage to give rise to the two chains [25,34]. The two subunits have different functions. For example, in the case of the Human Immunodeficiency Virus (HIV), HIV 1, the precursor glycoprotein is gp160, which is proteolytically cleaved into gp120 and gp41 subunits. The gp120 is the surface subunit and the gp41 is the transmembrane (TM) subunit. The surface subunit serves to recognize the cell to be infected when it comes in the vicinity of the virus with the help of receptors located on the cell surface. The gp41 mediates membrane fusion between the viral and cellular membranes. It has been found that gp41, and corresponding TM subunits in other retroviruses acquire an alpha-helical conforma-



Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 1

A schematic of the influenza virus. Hemagglutinin (HA) polypeptides lie on the surface of the virus while viral RNA resides inside. HA has subunits that attach to cells and then facilitate membrane fusion in order to infect the cell



Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 2

Various stages in influenza virus infection. The virus in endocytosed by the cell and an endosome is formed, wherein a conformational change of interest occurs

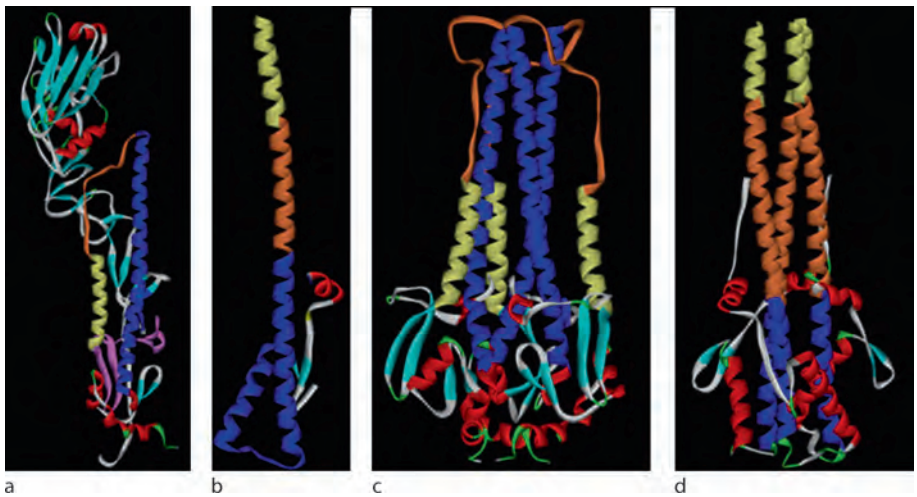
tion when the virus is in its active or fusogenic state. The structure is like a hairpin composed of three coils, having one C terminal (carboxy- end) and the other N terminal (amino-end). The carboxy regions pack in an anti-parallel manner around the three hydrophobic amino ends as shown in Fig. 3.

The native HA1/HA2 complex in the viral envelope is fusion-inactive. The cell-surface receptor needed for the virus to bind onto the cell is known as sialic acid. HA1 binds to this receptor and hence serves the purpose of bringing the virus and cell together. Upon binding with sialic acid, the virus is endocytosed by the cell (Fig. 2). HA remains dormant until the endosome begins to mature, and the pH in the HA surroundings drops to a value of about 5.

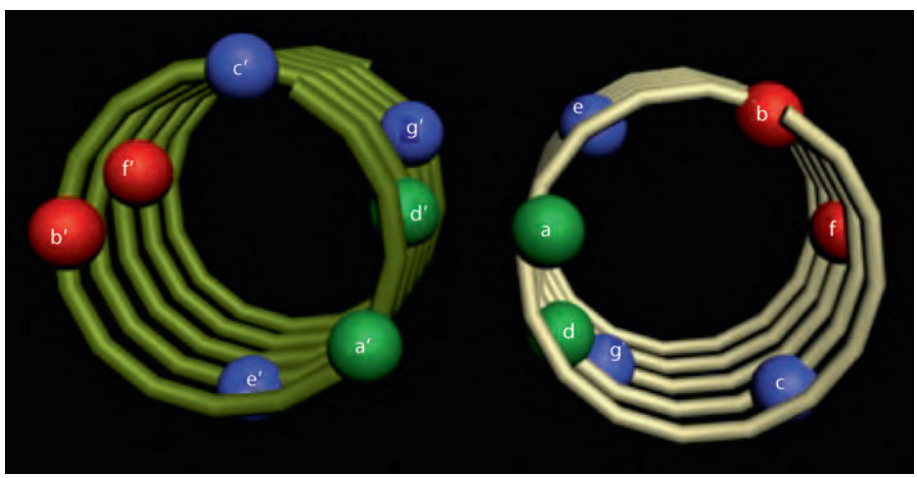
At this pH, there is a conformational change in HA2 domain of HA that induces the viral membrane to fuse with the cellular, endosomal membrane, thereby permitting the nucleocapsid of the virus to be deposited into the cytoplasm of the cell. Hence it can be inferred that the acidic pH acts as the physiological trigger for the HA conformational change. Since the low pH also activates membrane fusion, the low pH conformation of HA is also known as fusogenic conformation.

The crystal structures of HA in both the native and the fusogenic conformations are known [8,40,41,45]. As stated above, the HA1 only serves the purpose of cellular recognition and binding to the cell surface receptor and is not of much interest as a nanoactuator. The HA2 subunit, however, is then responsible for the membrane fusion activity. In the native state, the central region of HA2 folds as a helical hairpin-like monomeric structure (Fig. 3a). Two more such subunits (consisting of two segments each) are present to form a trimer together. The long α-helices from each subunit in the native state form a well known structure called the three stranded coiled coil system. Coiled coils can be visualized as an intertwined rope of three interacting helices. As is visible in Fig. 3c and 3d, in addition to the helical domains, there are other regions – small helices, β-sheets, turns and random loop regions. The blue regions in the native stage in Fig. 3c are triple stranded coiled coils wound against each other. They are connected to smaller helices (yellow) in the HA2 domain by loop regions shown in orange. The smaller helices connect to the fusion peptide (residue numbers 1–25). After attachment and pH drop, the orange loop regions convert into extended coiled coils effectively extending the blue coils and moving the fusion peptide by about 10 nm [11].

Coiled coil motifs contain hydrophobic and hydrophilic amino acids in a repeating heptad pattern (positions *a* through *g*) as shown in Figs. 4 and 5. Looking down the axis of the helix, hydrophobic residues tend to occur at the positions *a* and *d* of the heptad repeat and these residues form the interface between the helices. It is believed that HA folds into its thermodynamically most stable state at neutral pH, known as its native state. But at reduced pH, the fusogenic state becomes thermodynamically more feasible and hence the protein changes conformation in order to achieve it [10]. Another model



Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 3
a Hairpin like structure of central HA2 native state monomer. The monomer consists of two segments A and B. At the N-terminal is the fusion peptide contained in segment A (residues 1–25, shown in *pink*) followed by short anti-parallel β-sheet (residues 26–37, shown in *cyan*). The outer arm of the hairpin is a small α-helix (residues 38–53, shown *yellow*) which is connected to a long α-helix (residues 82–125, shown *blue*) by a loop region (residues 54–81, shown in *orange*); b HA2 chains A and B in fusogenic states. The loop region (*orange*) converts into a relatively more rigid α-helix and forms a continuation of the long α-helix (*blue*). The short α-helix is thus translated upwards along with the fusion peptide (not shown). c The native state HA2 trimer. The remaining structure other than the helices (*bottom*) is the fusion peptide from the N-terminal and the other smaller secondary structures following the long helices. d Fusogenic state of the HA2 trimer. The helical hairpins open up to form an extended three stranded coiled coil structure with the previously loop regions now helical. Such coiled coil motif is found in many other proteins such as the leucine zipper domain of some transcription factors. The α-helices are wrapped around each other with a left handed superhelical twist

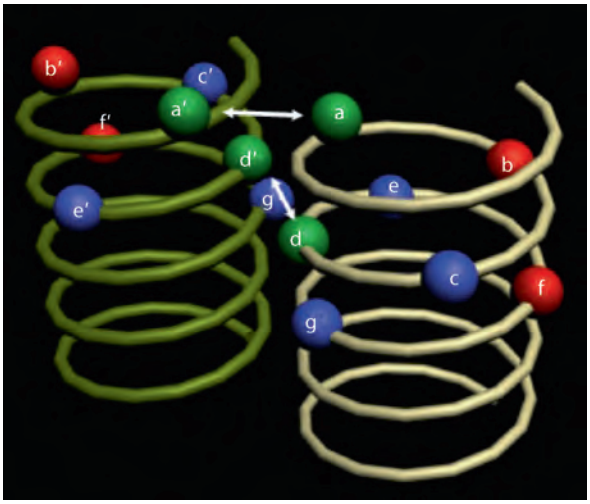


Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 4
Two helices of a coiled coil system. The seven positions (*a*–*e*) denote locations of hydrophobic/hydrophilic residues. Hydrophobic residues (green balls) tend to occur on the inside positions *a* and *d* whereas the hydrophilic residues (red balls) reside in other positions

that is widely accepted suggests that the native state is a metastable state that can be disturbed by any agent, not only pH. It has been observed that heat can also produce this conformational change at neutral pH and so does a denaturant such as urea [10].

In addition to Influenza, the following other viruses show similar infection mechanism and can be candidates for a VPL-like nanoactuator:

(i) The Human Immunodeficiency Virus type 1 (HIV 1) peptide gp41 [12]



Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 5

Packing of two helices of a coiled coil system. The hydrophobic residues in positions *a* — *a′* and *d* — *d′* of each chain interact with each other to stabilize the coiled coil system

(ii) The Human Respiratory Syncytial Virus (HRSV) protein subunit F1 [33]
(iii) The Simian Immunodeficiency Virus (SIV) protein gp41 [9]
(iv) The Human T cell Leukemia virus type 1, protein gp21 [23]
(v) The Simian Parainfluenza Virus peptide unit SV5 [3]
(vi) Ebola virus protein gp2 [42]

Each one of these peptides can result in a different actuator that can have different properties such as weight, volume, range of motion, force and speed capabilities. However, their principle of actuation is the same. Studies have shown [13,39,43] that the common characteristic in these viruses is the structure of a portion of the envelope glycoprotein and the mode of infection.

Projected Robotic Applications of VPL Nanoactuators

Fig. 6 shows a schematic of the VPL motor supporting a moving platform. The motor is shown in its initial, "contracted" phase in the first image that corresponds to the virus' native state and at its extended, fusogenic state in the second image of the figure. The total outward protrusion is measured to be 10 nanometers. The estimated time for filling up of the hydrophobic membrane defect by lipids is 2–4 ns, which gives an idea that the peptide unfolding process takes about the same time. To augment the force capabilities of the VPL motors, several VPL actuating elements could be attached in parallel as it is shown in Fig. 7. Such parallel attachment of multiple VPL motors could result extremely powerful, micro, meso or even macro actuators that will be able to apply ultra large forces while their dimensions are extremely small.

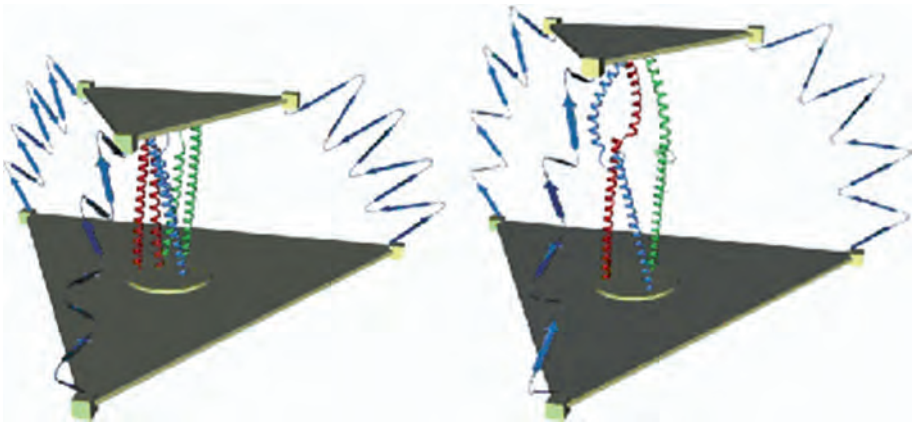
In a similar way, to increase the displacement capability of the VPL motors, several VPL elements could be connected in series (see Fig. 7). For attaching the titin fibers to the platforms DNA joints could be manufactured. By biochemical matching, the platforms could be designed to have proteins that bind to DNA on their corners. DNA strands could then bind onto the corners from one end and to the fibers on the other end to create a flexible joint. The titin fibers will have to be functionalized in such a way that their ends bind to the DNA as well.

Methods

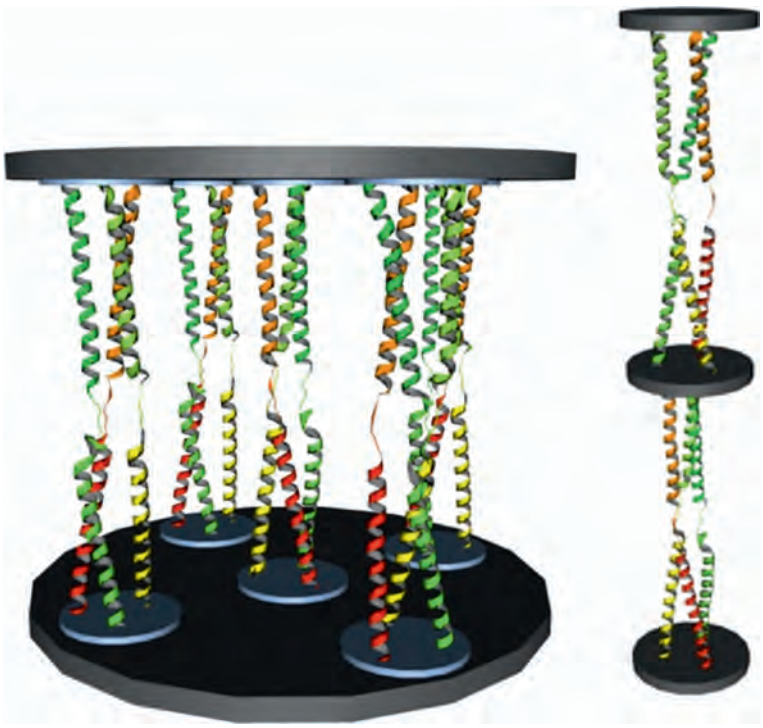
Targeted Molecular Dynamics (TMD) is a technique used to simulate conformational pathways between two known states by resorting to Newtonian molecular dynamics calculations with a physical potential and a holonomic constraint [29,30]. It provides an approximate modeling of processes that would otherwise span long time-scales (up to milliseconds) in realistic time scales for simulations. Protein folding is a microsecond scale phenomenon [46] and cannot be captured by unbiased MD simulations in realistic time frames. TMD is an example of a biasing technique that drives the trajectory from one known state to another. A number of such techniques have been investigated, for example, internal coordinate constraints [36], minimum biasing methods [20], biased molecular dynamics [27], essential dynamics sampling [14], steered molecular dynamics [22], modified targeted molecular dynamics [2], accelerated molecular dynamics [19] and more recently, the restricted perturbation targeted molecular dynamics (RP-TMD) [37] method.

TMD applies a holonomic constraint that reduces the root mean square deviation (RMSD) of the protein from a predefined target structure by a preset value at each dynamics step [29]. It has been shown that TMD can produce results equivalent to those from traditional MD simulations techniques in a much shorter time by using and lower temperatures. Ferrara et al. [17] have shown using the example of *chymotrypsin inhibitor 2* that the energies of conformation sampled using TMD are in agreement with those obtained from higher temperature unrestrained MD simulations and also with simulations performed using different force fields.

Because the distance to be traveled by the N-terminal of the VPL peptide is relatively very large, unbiased MD techniques cannot be applied to simulate the entire con-



Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 6
A single degree of freedom VPL actuated parallel platform VPL actuator



Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 7
Several VPL motors placed in parallel and series

formational change. High resolution crystal structures of initial and final states of VPL peptides are available from PDB. The closed state was termed as the initial state *I* and the open state as the final state *F*. TMD applies the holonomic constraint to reduce the root mean square deviation (RMSD) from *I* to *F* by a finite value at each MD step. The 3N position coordinates corresponding to N atoms in the molecule can be defined as:

$$\mathbf{X} = (X_1, X_2, \ldots, X_{3N})^T , \tag{1}$$

where 3N are the Cartesian coordinates of the position vectors $\mathbf{r_1}, \mathbf{r_2}, \ldots, \mathbf{r_N}$ of each individual atom.

The position of an individual atom is time dependent, and is described at time t as $X_i(t)$. Let X_{Fi} be the final (target) position of an atom i. The holonomic constraint is of

the form:

$$\Phi[\underline{X}(t)] = \sum \left[\underline{X}_i(t) - \underline{X}_{Fi}\right]^2 - \underline{\rho}^2(t) = 0\,. \tag{2}$$

The underlined terms indicate that they have been scaled by a factor of $\sqrt{m_i/\langle m\rangle}$ where m_i is the mass of atom i and $\langle m\rangle$ is the average atomic mass. This mass scaling helps avoid the net translation of the system as a whole [30]. The holonomic constraint in Eq. (2) gives rise to a constraint force of the form:

$$F_i^c = \lambda \nabla_i \Phi = \frac{2\lambda m_i}{\langle m\rangle}(X_i - X_{Fi})\,, \tag{3}$$

where λ is a Lagrange multiplier. It is determined by the following method. The Leap-frog [1] propagator is a common method used to update variables along a trajectory of positions. According to this algorithm, the position of an atom i at time t is given by:

$$\begin{aligned} X_i(t) &= \left[X_i(t-\Delta t) + v_i\left(t - \frac{3}{2}\Delta t\right)\Delta t \right. \\ &\quad \left. + \frac{F_i(t-\Delta t)}{m_i}(\Delta t)^2\right] + \frac{F_i^c(t-\Delta t)}{m_i}(\Delta t)^2 \\ &= x_i(t) + p_i(t) \end{aligned} \tag{4}$$

In the above equation, Δt is the time increment, F_i is the force on atom i due to unbiased Newtonian molecular dynamics, and v_i is the velocity of the atom. The variable $x_i(t)$ is the position of the atom in the absence of the constraint and $p_i(t)$ is the perturbation due to the holonomic constraint [37]. Noting that:

$$p_i(t) = \frac{F_i^c(t-\Delta t)}{m_i}(\Delta t)^2$$

from Eq. (4), and

$$F_i^c(t-\Delta t) = \frac{2\lambda m_i}{\langle m\rangle}\{X_i(t-\Delta t) - X_{Fi}\}$$

from Eq. (3), and defining

$$\gamma = \frac{2(\Delta t)^2\lambda}{\langle m\rangle}$$

we get:

$$p_i(t) = \gamma[X_i(t-\Delta t) - X_{Fi}]\,. \tag{5}$$

Hence the perturbation on each atom is now simply the product of a Δt dependent scaling factor γ and the difference in position between the current and previous time step. In order to obtain γ (and hence λ), substitute Eq. (4) into the original constraint Eq. (2) to get:

$$\Phi[\underline{X}(t)] = \sum \left[\underline{X}_i(t) - \underline{X}_{Fi}\right]^2 - \underline{\rho}^2(t) = 0\,.$$

Which can be written as:

$$\sum \left[\left\{\underline{x}_i(t) + \underline{p}_i(t)\right\} - \underline{X}_{Fi}\right]^2 - \underline{\rho}^2(t) = 0\,,$$

and further simplified as:

$$\begin{aligned} &\sum \left[\left\{\underline{x}_i(t) + \underline{p}_i(t)\right\}^2 \right. \\ &\qquad \left. - 2\left\{\underline{x}_i(t) + \underline{p}_i(t)\right\}\underline{X}_{Fi} - \underline{X}_{Fi}^2\right] \\ &\qquad - \underline{\rho}^2(t) = 0 \\ &\sum \left[\left\{\underline{x}_i^2(t) + \underline{p}_i^2(t) + 2\underline{x}_i(t)\underline{p}_i(t)\right\} \right. \\ &\qquad \left. - 2\underline{x}_i(t)\underline{X}_{Fi} - 2\underline{p}_i(t)\underline{X}_{Fi} - \underline{X}_{Fi}^2\right] \\ &\qquad - \underline{\rho}^2(t) = 0 \\ &\sum \left[\left\{\underline{p}_i^2(t) + 2\underline{x}_i(t)\underline{p}_i(t) - 2\underline{p}_i(t)\underline{X}_{Fi}\right\} \right. \\ &\qquad \left. + \underline{x}_i^2(t) - 2\underline{x}_i(t)\underline{X}_{Fi} + \underline{X}_{Fi}^2\right] \\ &\qquad - \underline{\rho}^2(t) = 0 \\ &\sum \left[\underline{p}_i^2(t) + 2\underline{p}_i(t)\{\underline{x}_i(t) - \underline{X}_{Fi}\}\right] \\ &\qquad + \sum \left[\underline{x}_i(t) - \underline{X}_{Fi}\right]^2 - \underline{\rho}^2(t) = 0\,. \end{aligned} \tag{6}$$

Using Eq. (5) into (6) yields:

$$\begin{aligned} &\sum [\gamma^2\{\underline{X}_i(t-\Delta t) - \underline{X}_{Fi}\}^2 + 2\gamma\{\underline{X}_i(t-\Delta t) - \underline{X}_{Fi}\} \\ &\cdot \{\underline{x}_i(t) - \underline{X}_{Fi}\}] + \sum [\underline{x}_i(t) - \underline{X}_{Fi}]^2 - \underline{\rho}^2(t) = 0\,, \end{aligned}$$

which can be rearranged as –

$$\begin{aligned} &\gamma^2 \sum \left[\{\underline{X}_i(t-\Delta t) - \underline{X}_{Fi}\}^2\right] \\ &+ \gamma \sum 2[\{\underline{X}_i(t-\Delta t) - \underline{X}_{Fi}\}\{\underline{x}_i(t) - \underline{X}_{Fi}\}] \\ &\qquad + \sum [\underline{x}_i(t) - \underline{X}_{Fi}]^2 - \underline{\rho}^2(t) = 0\,. \end{aligned} \tag{7}$$

Defining,

$$\begin{aligned} a &= \sum \left[\{\underline{X}_i(t-\Delta t) - \underline{X}_{Fi}\}^2\right], \\ b &= 2\sum [\{\underline{X}_i(t-\Delta t) - \underline{X}_{Fi}\}\{\underline{x}_i(t) - \underline{X}_{Fi}\}]\,. \end{aligned}$$

And noting that:

$$\Phi\left[\underline{x}(t)\right] = \sum \left[\underline{x}_i(t) - \underline{X}_{Fi}\right]^2 - \underline{\rho}^2(t)\,.$$

Equation (7) can be simply written as –

$$a\gamma^2 + b\gamma + \Phi[\underline{x}(t)] = 0\,. \tag{8}$$

The Eq. (8) yields two roots for γ. In order to minimize the total perturbation $\sum |p_i|$ (from Eq. (5)), the root with lower absolute value is chosen.

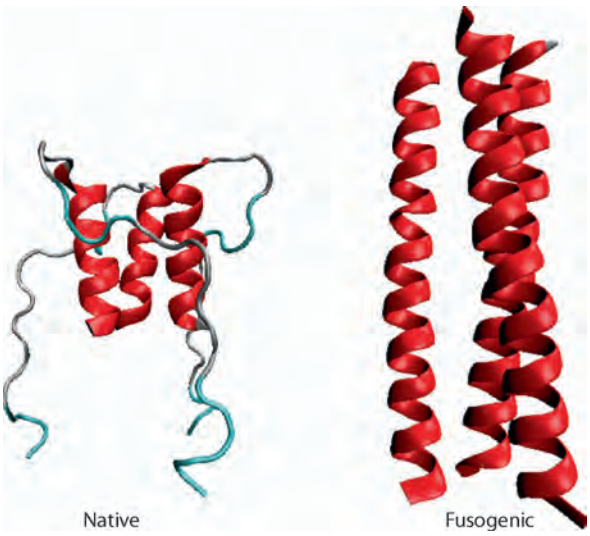
The RMSD is decreased by a predetermined (and user defined) value at each time step by a value of $\Delta\rho$ as:

$$\rho(t) = \rho(t - \Delta t) - \Delta\rho\,. \tag{9}$$

The TMD algorithm steps are:

(i) Set $\rho = \rho_0 = |\mathbf{x}_\mathrm{I} - \mathbf{x}_\mathrm{F}|$ where I is the initial and F is the final conformation.
(ii) Numerically solve the equations of motion with the additional constraint force F^c.
(iii) After each time step Δt, diminish ρ as in Eq. (9).

At the end of the simulation, the final distance ρ_f is reached. In this way, a monotonous reduction of ρ forces the system to find a pathway from $\mathbf{x}_\mathrm{I}$ to a final configuration $\mathbf{x}_\mathrm{F}$.



Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 8
The peptide loop36 in the initial and final states of the conformational change. Coiled coils are shown in red and the random loop regions in white and cyan. The random loop regions rise and convert into α-helical coiled coils upon activation

Simulation Results and Discussion

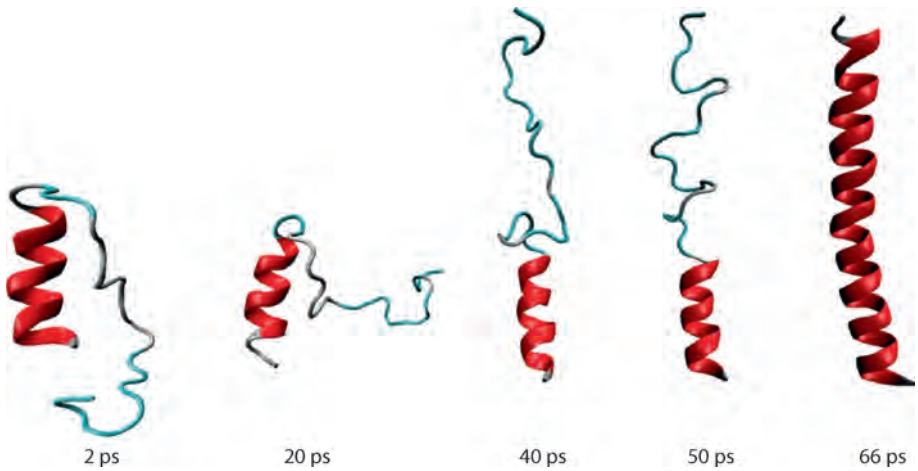
Molecular Dynamics Simulations were performed on VPL peptides in order to study the performance of the nanoactuator. Specifically, a fragment from the VPL segment, the loop36 was chosen for these studies due to its established role in the conformational change [11]. The large conformational change was simulated using the Targeted Molecular Dynamics (TMD) technique. The TMD analyzes were performed using implicit solvent conditions. In order to simulate the effect of environmental pH change on the VPL motor protein, acidic amino acids were protonated. In the TMD simulations the protein was allowed to undergo a transition from one known state to another. This is a useful approach to ensure the feasibility of the conformational change and to get an idea of the magnitude of the energy change involved. The VPL nanoactuator works in a solution where there are protons bombarding the protein due to a reduction in the solution pH. This allows acidic amino acids in the protein to accept the protons on their negatively charged sites, and hence get protonated. Protonation changes the energy balance of the solvent-protein system and can create conditions that promote a conformational change. Low pH simulations on the small fragment loop36 were performed to study the effect on the protein conformation and the way it interacts energetically with its surroundings.

In addition to the TMD analyzes, classical molecular dynamics approaches were performed to quantify some other aspects such as helicity and free energy of the peptides. In order to be able to perform classical unbiased MD simulations in a realistic time scale, it is important to energize the system by using high temperatures. The open and closed structures were taken from the protein data bank (PDB) with the open state as the target state and the closed state as the starting point [16]. The simulations were performed using an implicit solvation model known as Effective Energy Function for proteins EEF1 [24].

Structural Details of Loop36

Loop36 is a 36 residue long peptide of the VPL protein that forms a hinge region of the viral protein joining the two α-helical regions of each monomer [8,11]. An initial state as obtained from PDB [6] file 1HGF at pH of about 7.0, loop36 consists of a 15 residue long α-helical part with the remaining fragment in a random and flabby form (Fig. 8). It is located in the segment B of the Influenza Hemagglutinin protein sequence from residues 54 to 89. It has been recognized as being critical for a pH dependent conformational change [10,11]. The loop36 wild type (naturally occurring) sequence is as follows:

ARG VAL ILE GLU LYS THR ASN GLU LYS PHE HIS GLN ILE GLU LYS GLU PHE SER GLU VAL GLU GLY ARG ILE GLN ASP LEU GLU LYS TYR VAL GLU ASP THR LYS ILE



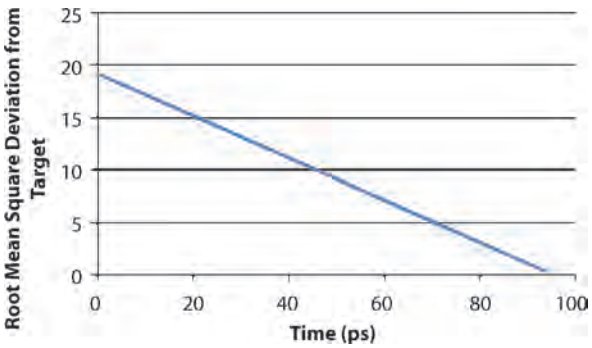
Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 9
Transition of a loop36 monomer from native to fusogenic state during a TMD simulation

The initial and final states of *loop36* as obtained by the crystal structures 1HGF (pH ∼ 7.0) and 1HTM (pH ∼ 5.0) from PDB are shown in Fig. 8.

In order to simulate low pH, the biochemical response of certain amino acids to the acidic conditions was considered. Some amino acids are acidic in nature, some basic and some neutral; and they get protonated at different pH values. Out of the 20 constituent types of amino acids, Glutamic Acid (GLU), Aspartic Acid (ASP), and Histidine (HIS) are in their unprotonated states at neutral pH (7.0). There is only one histidine residue in loop36, whereas there are eight glutamic acid residues and two aspartic acid residues which make a large part of loop36. Hence, overall, there were eleven residues out of the 36 that were protonated.

Figure 9 shows the Visual Molecular Dynamics (VMD) [21] rendered transition of a loop36 monomer from initial (closed) state to an open (fusogenic) state. The entire transition is achieved in 66 picoseconds when using TMD, while it is not achievable in 30 nanoseconds using traditional MD techniques. As seen in the figure, after 20 ps the loop regions of the peptide are halfway open and after 50 ps they are completely opened but have not attained an α-helical conformation.

The deviation of the final state from the initial state of the monomer in Fig. 9 can be seen in the root mean square deviation (RMSD) from the open state. Results are depicted in Fig. 10 and values of RMSD indicate that the open and close conformation differ by approximately 19 Å. The rate of opening of the monomer can be estimated by the slope of this curve which in this case is 0.2 Å/ps. It must be kept in mind that to have an estimation of the real rate of opening, this value should be rescaled to account for the increase in speed of the process given by TMD.

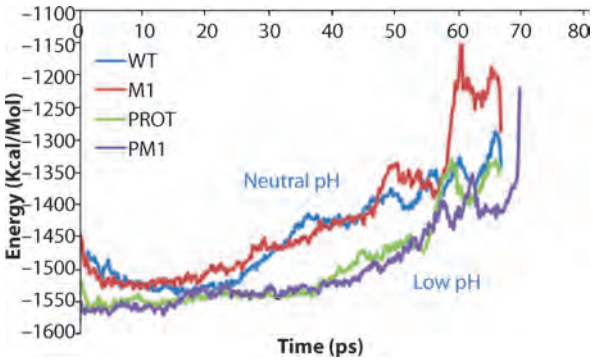


Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 10
In the beginning the RMSD from the final state has a maximum value of about 19 Å and it decreases linearly with time as the open state is achieved due to the holonomic constraint applied by TMD

Mutational Analysis

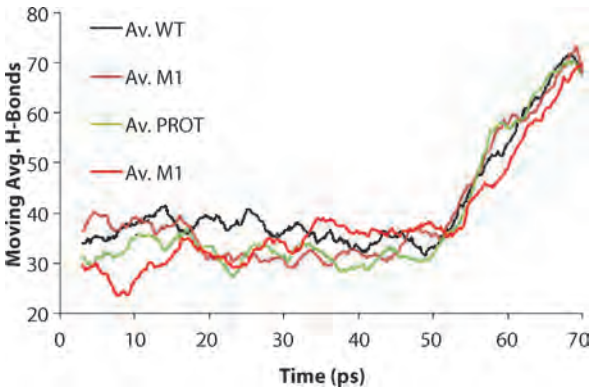
Mutations in a protein can be made by replacing one or more constituent amino acids with different amino acids. The structure and properties of a protein change when they undergo a mutation. It is therefore possible to design a protein according to specific requirements; for example, if a high hydrophobicity is desired, hydrophobic amino acids should replace some of the residues.

The loop36 peptide contains one Glycine residue which is known as a helix-breaker amino acid because its side chain is simply one hydrogen atom. In experimental



Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 11

Comparison of potential energy paths of four types of loop36 monomers. (i) wild type WT (the protein in this state does not have any mutation or protonation), (ii) mutated type M1 (in this state, GLY (glycine) is replaced by ALA (alanine), (iii) protonated type PROT (GLU (glutamic acid)) is replaced by protonated GLU, ASP (aspartic acid) by protonated ASP, HIS (histidine) replaced by doubly protonated histidine and (iv) protonated and mutated type PM1 (in this state, the protein has both protonation and mutation). The low pH structures follow lower energy pathways than the neutral pH structures although all structures eventually achieve similar endstate energy values



Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 12

The number of H-bonds in each of the monomer WT, M1, PROT and PM1 modified structures varies as the open state is achieved. A moving average of 15 points (3 ps) is shown in each curve. The H-bonds oscillate around a constant value up to about 50 ps in each of the cases. A steep rise in the number is observed beyond this point. This region corresponds to the helix formation which results in intra-helical H-bond formation

studies, it is observed that replacing this residue by an Alanine results in better α-helix formation [11]. In line with the experiments, in our simulations the Glycine is replaced with the Alanine. The glycine (GLY) residue is the 22nd residue in loop36, and a mutated version of the loop36 with GLY replaced by alanine (ALA) is called mutation M1. The unmutated version of the peptide is simply called the wild type (WT). GLY contains no side chain and it has no atom attached to the α-carbon (CA). On the other hand, ALA contains a simple side chain (CH3) attached to its CA atom.

The four different cases compared in this study are as follows:

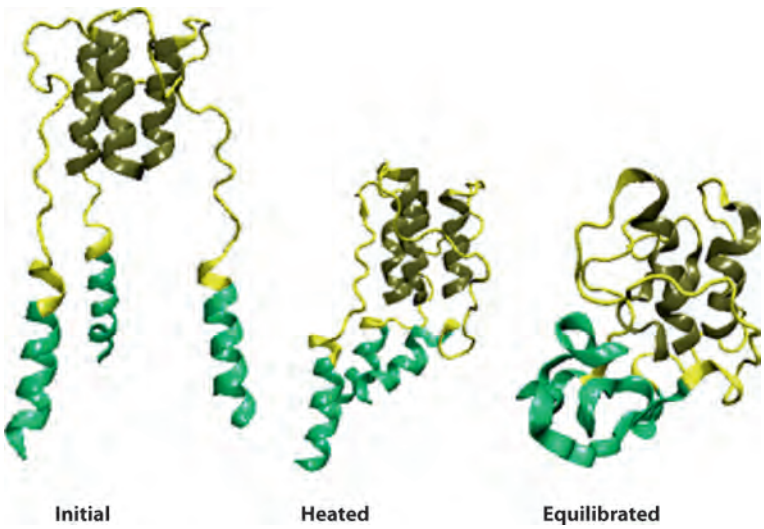
(i) the wild type (no mutation or protonation) guided to an open state
(ii) one mutation M1
(iii) protonation of amino acids (with no mutations), and
(iv) mutation and protonation (PM1). Conformational energies of these simulations are compared in Fig. 11.

H-Bond Analysis

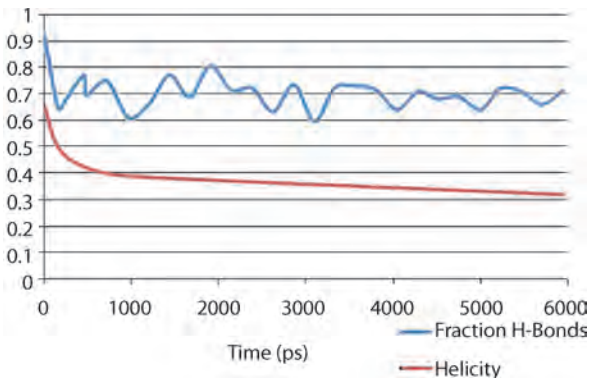
The intra-helical H-bonds stabilize α-helical structures. The hydrogen bonds for the loop36 conformational transition (Fig. 12) were found to be oscillatory in nature and hence a moving average with a period of 15 time units was chosen. Each observation was made every 0.2 ps hence 15 points correspond to 3 ps time interval. In the initial stages of the simulation (time $<$ 50 ps) the partial helix of loop36 is the only region with a significant number of H-bonds. The random loop region is converted into α-helix beyond 50 ps and hence a steep increase in the numbers is observed. The variation in the open state contacts, namely the H-bonds quantitatively describe the achievement of the open state.

Classical Molecular Dynamics Studies

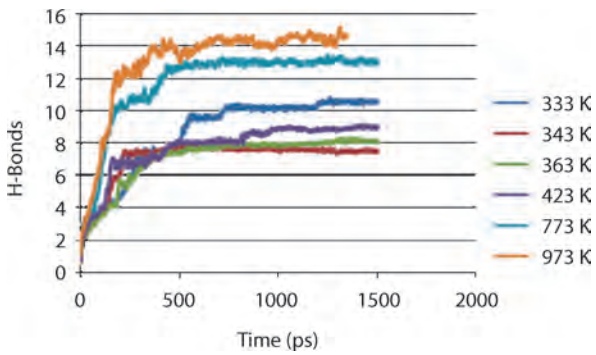
Even though classical methods cannot capture a large conformational change such as that exhibited in the VPL nanoactuator they provide insight into important structural details and properties. Typically higher temperatures are used to accelerate structural events in macromolecule simulations. When the protein is exposed to high temperatures, not only can it fold/unfold but it may also become denatured. Heating a protein corresponds to increasing the kinetic energy of its atoms, and this, in turn causes the molecules to vibrate more, possibly disrupting many of the hydrogen bonds and non-polar interactions and hence the secondary structure of the protein. The classical simulations on VPL protein were performed to understand the behavior of the VPL trimer at elevated temperatures. Knowledge of this will also help to ascertain whether there is a range of temperatures in which changes of structure such as folding or α-helix formation were initiated, be-



Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 13
Effect of temperature on VPL peptide. The heated state corresponds to 500 °C temperature achieved in 150 ps and the equilibrated state corresponds to a constant temperature equilibration for up to 6 ns. The peptide gets denatured at this point



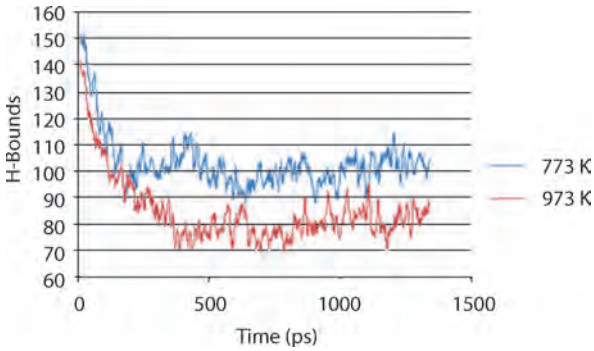
Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 14
Helicity and fraction of open state contacts for a 6 ns classical high temperature molecular dynamics simulation



Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 15
RMSD of the loop50 trimer from its initial state for each of the simulations ranging from 333 K to 973 K. Higher temperatures take the peptide further away from the initial state which indicates a larger conformational change. However, there is significant denaturation of the peptide at 773 K and 973 K

fore the protein denatures [15]. Figure 13 shows temperature effects on the VPL trimer loop50, which is an extended loop36 trimer. In contrast to the TMD simulations in which the open state was achievable in about 60 ps, the classical simulations need an approximate time of the order of few microseconds to achieve the same results [18].

It has been shown experimentally [28] that viral membrane fusion in the hemagglutinin of influenza virus occurs at 335 K. The loop50 trimer was subjected to a range of simulated temperatures ranging from 333 K to 973 K with an attempt to capture the conformational changes required for fusion and hence the performance of the VPL nanoactuator. The RMSD from the initial equilibrated state at various temperatures is shown in Fig. 15. The expected trend would be that increasing temperature increases the RMSD from the initial state as it allows the peptide to jump energy barriers and denatures it. This was observed in simulations at 773 K and 973 K. However, the 333 K simulation indicates that the peptide has a larger difference with the initial state (larger RMSD values) in the same time frame than those at 343 K, 363 K and 423 K. This result is in agreement with the experimental observa-



Viral Protein Nano-Actuators: Computational Studies of Bio-Nanomachines, Figure 16

H-bonding data for the conformational changes achieved at 773 K and 973 K. The higher temperature results in loss of H-bonds

tion and suggests that larger conformational changes may take place in the temperature range of 333–343 K [11]. The extremely high temperatures 773 K and 973 K result in denaturation of the peptides as seen by the H-bond data in Fig. 16.

Conclusions

The recent explosion in nanotechnology has given place to the creation of biomolecular machines and motors such as viral protein linear actuators (VPL). The idea of the VPL motors, whose constituent proteins belong to a family of retroviruses such as HIV and influenza virus, originates from the mechanism they utilize to infect a cell. When these types of viruses are endocytosed by the cell, they experience a drop in pH and change their structure from a coil strand to an alpha helix conformation. In order to exploit the functional mechanism of viral proteins, it is important to characterize and understand the sensing capabilities of the peptides under different conditions of the environment. In this study, some aspects of the response of the VPL motor to different stimuli, using molecular dynamics have been illustrated. For the study of nanoactuators, in particular, molecular simulations provide a very valuable tool, especially since sometimes; experiments at the nanoscale may not be simple to perform. Alteration to the protein sequence (mutations) is a good example of a process in which computational studies are much easier to perform than experimental ones due to the complexities involved. One of the challenges is to learn how to control structural mechanisms, behavior and properties of the basic nanocomponents involved in molecular motors. In this study we learn about changes in the structure experienced by the viral protein actuator monomer as it interacts with other monomers in conditions of trimerization, mutation of its sequence, and changes in temperature.

The VPL undergoes a remarkably large conformational change upon a drop in pH. In order to simulate the motion of the protein, targeted and traditional MD techniques are employed. Structural changes induced by pH variations are studied by protonating a number of amino acids in the peptide sequence. The TMD results showed a clear preference to conformational transition paths at low pH values. Our findings confirm that the protein forms a helical coil at acidic pH values but it shows a coiled structure at neutral pH. More studies are needed to establish conclusive results on the effect of pH and mutations.

The experimental observation that influenza hemagglutinin peptides showed a conformational change with a significant formation of alpha helix occurs in the wild type (unmutated) at 65°C at neutral pH [10]. In our simulations of a reduced loop50 trimer, this conformational change was not captured within 6 ns. Given the remarkably short time frames of our simulations and the size of the peptides, we are unable to observe regeneration of alpha helix. Even more, in order to be able to observe visible conformational changes, we had to work at very high temperatures. Even at that value of the temperature we do not see complete denaturation, again due to the short time frames under study. In an experimental situation the protein denatures at a much lower value of the temperature (at approximately 71°C). At these relatively low temperatures, a much longer time scale (micro- or even millisecond is required to fully capture alpha helix regeneration.

It is important to mention that high temperature approaches remain a valuable tool for conformational exploration; one must understand, however, that torsional transition kinetics are not correct at higher temperatures. Furthermore, rescaling to lower temperatures yields only approximate results. In spite of this, our goal is to present a methodology for studying large conformational changes in nanoscale molecular systems using both targeted and traditional molecular dynamics techniques that can ultimately be extended to much longer time frames using parallel computing.

We are at the dawn of a new era in the development of molecular machinery, and we are just starting to elucidate some of the challenges that these nanoactuators represent. Future studies will involve investigating interfacing with other molecular components such as carbon nanotubes, biological membranes, inorganic substrates, ions etc. As mentioned earlier, the end-effector of VPL can be designed according to the requirement. Moreover, the composition of the peptide can be varied according to the environ-

ment, as long as the rules of having hydrophobic and polar residues to ensure stability of the coiled-coil system are adhered to. The length of the VPL peptides, the role of the solvent (and solvent composition), binding energies to various target objects, are other variables that need to be investigated in order to assess conditions for optimum performance, (i. e. for example maximum force and stability, a given velocity etc.)

The VPL peptides are known to work better with specific mutations [11]. This opens up doors for further mutational analysis of the peptide, again directed towards achieving optimal performance parameters. The goal will be to determine the optimal sequence of the VPL peptide that results in the most stable, quickest, powerful and robust nanoactuator.

Future Directions

The future of this study involves the use of larger peptides, employment of better computational algorithms and interfacing of the VPL nanoactuator to other nanodevices to produce complex nanosystems. This work has shown great potential for further research. Better algorithms need to be developed for the study of the forces exerted by the nanoactuator. These studies have concentrated on developing a suitable approach to characterize and study conformational changes of loop36 and loop50. Preliminary studies have shown that longer peptides, such as loop114 (an extended version of loop50) are promising nanoactuator candidates as they provide a larger displacement ($\sim$ 9 nm) and they have a larger helical structure. The added secondary structure may provide more stability and larger forces to the actuator.

Mutations M1 to M7 were applied to loop36. Future studies will involve mutations to loop50 and larger peptides, which provide more mutation points and can significantly alter the performance of the actuator. The VPL nanoactuator can also be interfaced with other protein/lipid based molecular platform/devices such as lipid bilayer membranes. A possible scenario is to attach a cysteine tag to the end of VPL peptide and a lipid bilayer via a disulfide bond. TMD and Steered Molecular Dynamics (SMD) [22] can be used to study the performance of the nanoactuator while bound to a platform.

The TMD algorithm that was used to simulate conformational changes by applying a bias in the force field in this study has recently been succeeded by an updated restricted perturbation TMD (RP-TMD) [37] algorithm. The RP-TMD method depends not only on reduction of RMSD at each time step, but also on the minimal energy path. The two-way check yields lower energy pathways for the trajectories. The TMD algorithm has a few disadvantages:

(i) Transition is based solely on RMSD, which is a geometric parameter. No attempt is made to ensure that the conformational transition follows the minimum possible energy pathway.
(ii) Due to the above, large energy barriers are crossed which may not be realistic.
(iii) Some trajectories obtained are not reversible.

In the RP-TMD method, the above deficiencies are considerably reduced.

Acknowledgments

The authors gratefully acknowledge support of the National Science Foundation (Grant NSF NIRT #0303950).

Bibliography

1. Allen M, Tildesley D (1987) Computer Simulation of Liquids. Oxford University Press, New York
2. Apostolakis J, Ferrara P, Caflisch A (1999) Calculation of conformational transitions and barriers in solvated systems: application to the alanine dipeptide in water. J Chem Phys 110(4):2099–2108
3. Baker KA et al (1999) Structural basis for paramyxovirus-mediated membrane fusion. Mol Cell 3(3):309–19
4. Balzani VV et al (2000) Artificial Molecular Machines. Angew Chem Int Edn Engl 39(19):3348–3391
5. Berg HC (2003) The Rotary Motor of Bacterial Flagella. Annu Rev Biochem 72:19–54
6. Berman HM et al (2000) The Protein Data Bank. Nucl Acids Res 28(1):235–42
7. Boyer PD (1998) Energy, life, and ATP. Biosci Rep 18(3):97–117
8. Bullough PA et al (1994) Structure of influenza haemagglutinin at the pH of membrane fusion. Nature 371(6492):37–43
9. Caffrey M et al (1998) Three-dimensional solution structure of the 44 kDa ectodomain of SIV gp41. EMBO J 17(16):4572–84
10. Carr CM, Chaudhry C, Kim PS (1997) Influenza hemagglutinin is spring-loaded by a metastable native conformation. Proc Natl Acad Sci USA 94(26):14306–13
11. Carr CM, Kim PS (1993) A spring-loaded mechanism for the conformational change of influenza hemagglutinin. Cell 73(4):823–32
12. Chan DC et al (1997) Core structure of gp41 from the HIV envelope glycoprotein. Cell 89(2):263–73
13. Colman PM, Lawrence MC (2003) The structural biology of type I viral membrane fusion. Nat Rev Mol Cell Biol 4(4):309–19
14. Daidone I et al (2003) Molecular Dynamics Simulation of Protein Folding by Essential Dynamics Sampling: Folding Landscape of Horse Heart Cytochrome c. Biophys J 85(5):2865–2871
15. Dubey A, Mavroidis C, Tomassone MS (2006) Molecular Dynamic Studies of Viral-Protein Based Nano-Actuators. J Comput Theor Nanosci 3:885
16. Dubey A, Sharma G, Mavroidis C, Tomassone MS, Nikitczuk K, Yarmush ML (2004) Computational Studies of Viral Protein Nano-Actuator. J Comput Theor Nanosci 1(1):1–11

17. Ferrara P, Apostolakis J, Caflisch A (2000) Targeted Molecular Dynamics Simulations of Protein Folding. J Phys Chem B 104:4511–4518
18. Ferrara P, Apostolakis J, Caflisch A (2000) Thermodynamics and Kinetics of Folding of Two Model Peptides Investigated by Molecular Dynamics Simulations. J Phys Chem B 104(20):5000–5010
19. Hamelberg D, Mongan J, McCammon JA (2004) Accelerated molecular dynamics: a promising and efficient simulation method for biomolecules. J Chem Phys 120(24):11919–29
20. Harvey SC, Gabb HA (1993) Conformational transitions using molecular dynamics with minimum biasing. Biopolymers 33(8):1167–72
21. Humphrey W, Dalke A, Schulten K (1996) VMD: visual molecular dynamics. J Mol Graph 14(1):27–8, 33–8
22. Isralewitz B, Gao M, Schulten K (2001) Steered molecular dynamics and mechanical functions of proteins. Curr Opin Struct Biol 11(2):224–230
23. Kobe B et al (1999) Crystal structure of human T cell leukemia virus type 1 gp21 ectodomain crystallized as a maltose-binding protein chimera reveals structural evolution of retroviral transmembrane proteins. Proc Natl Acad Sci USA 96(8): 4319–24
24. Lazaridis T, Karplus M (1999) Effective energy function for proteins in solution. Proteins 35(2):133–52
25. Lazarowitz SG, Compans RW, Choppin PW (1971) Influenza virus structural and nonstructural proteins in infected cells and their plasma membranes. Virology 46(3):830–43
26. Oster G, Wang H (2003) Rotary protein motors. Trends Cell Biol 13(3):114–21
27. Paci E, Karplus M (1999) Forced unfolding of fibronectin type 3 modules: an analysis by biased molecular dynamics simulations. J Mol Biol 288(3):441–59
28. Ruigrok RW et al (1986) Conformational changes in the hemagglutinin of influenza virus which accompany heat-induced fusion of virus with liposomes. Virology 155(2):484–97
29. Schlitter J (1994) Targeted Molecular Dynamics: A New Approach for Searching Pathways of Conformational Transitions. J Mol Graph 12:84–89
30. Schlitter JEM, Krüger P, Jacoby E, Wollmer A (1993) Targeted molecular dynamics simulation of conformational change: application to the T?R transition in insulin. Mol Simul 10(2–6):291–308
31. Schoch C, Blumenthal R, Clague MJ (1992) A long-lived state for influenza virus-erythrocyte complexes committed to fusion at neutral pH. FEBS Lett 311(3):221–5
32. Seeman NC (2003) DNA in a material world. Nature 421(6921):427–31
33. Singh M, Berger B, Kim PS (1999) LearnCoil-VMF: computational evidence for coiled-coil-like motifs in many viral membrane-fusion proteins. J Mol Biol 290(5):1031–41
34. Skehel JJ, Waterfield MD (1975) Studies on the primary structure of the influenza virus hemagglutinin. Proc Natl Acad Sci USA 72(1):93–7
35. Skehel JJ, Wiley DC (2000) Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin. Annu Rev Biochem 69:531–69
36. Tobias DJ, Brooks CL III (1988) Molecular dynamics with internal coordinate constraints. J Chem Phys 89:5115–5127
37. van der Vaart A, Karplus M (2005) Simulation of conformational transitions by the restricted perturbation–targeted molecular dynamics method. J Chem Phys 122:114903
38. Vale RD, Milligan RA (2000) The way things move: looking under the hood of molecular motor proteins. Science 288(5463):88–95
39. Walker JE (1998) ATP Synthesis by Rotary Catalysis (Nobel Lecture). Angew Chem Int Edn 37:2308–2319
40. Weis WI et al (1988) Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. Nature 333(6172):426–31
41. Weis WI et al (1990) Refinement of the influenza virus hemagglutinin by simulated annealing. J Mol Biol 212(4):737–61
42. Weissenhorn W et al (1998) The central structural feature of the membrane fusion protein subunit from the Ebola virus glycoprotein is a long triple-stranded coiled coil. Proc Natl Acad Sci USA 95(11):6032–6
43. Weissenhorn W et al (1999) Structural basis for membrane fusion by enveloped viruses. Mol Membr Biol 16(1):3–9
44. Wiley DC, Skehel JJ (1987) The structure and function of the hemagglutinin membrane glycoprotein of influenza virus. Annu Rev Biochem 56:365–94
45. Wilson IA, Skehel JJ, Wiley DC (1981) Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution. Nature 289(5796):366–73
46. Yang W et al (2003) The missing link between thermodynamics and structure in F1-ATPase. Proc Natl Acad Sci USA 100(3): 874–9
47. Yurke B, Turberfield AJ, Mills AP, Simmel FC, Neumann JL (2000) A DNA-Fuelled Molecular Machine Made of DNA. Nature 415:62–65

http://www.springer.com/978-0-387-75888-6

Encyclopedia of Complexity and Systems Science
Editor-in-chief: Meyers, R.A.
2009, MCXX, 10398 p. In 14 volumes, not available separately., Hardcover
ISBN: 978-0-387-75888-6