Design and Computational Analysis of Protein Based Nanoscale Biomimetic Actuators

Dubey, A., Rutgers University, New Brunswick, NJ, USA
Mavroidis, C., Northeastern University, Boston, MA, USA
Tomassone M. S. Rutgers University, New Brunswick, NJ, USA

Abstract

A nanoscale biomolecular linear actuator taken from a virus is introduced, analyzed and methods to characterize and quantify its performance are discussed and applied. The HA2 domain of Influenza viral peptide is known to undergo a large conformational change in the cellular endosome upon a drop in pH. Targeted molecular dynamics techniques (TMD) are employed to study the opening/closing behavior of a hinge region that plays a critical role in the conformational change. Four different models of the peptide are subjected to TMD to trace a trajectory from closed (initial) to open (final) state and the differences between them are quantified using conformational energy and open state contacts to show the role of low pH as well as peptide mutations in order for such systems to act as nanoactuators. The results obtained are found to be in agreement with experimental findings that show that the protonated and mutated peptides are more likely to attain stable open state when compared to the wild type peptide.

Introduction

Development of discrete individual machine elements – from nature or synthetic – is a critical starting step towards the creation of fully functional complex nanomachines and nanorobots. Although recent advances in the nanorobotics and molecular biochemistry give us information on the possibilities of several molecules to be used as specific types of motors and other machines components such as one degree of freedom rotary ATP-based motors (1), linear actuators have not been thoroughly investigated. Even though various biomolecular elements have been suggested as possible parts in a nano-biomachinery, there have been few attempts to characterize them in such a perspective. In particular, it is still not clear the role of collective interactions with other actuators on the functional capabilities of a single one. In this context, we computationally characterize a novel biomolecule, the Viral Protein Linear actuator (VPL) and focus on the effect of trimerization on the final performance of a single actuator. The results are compared to the experimental findings in the literature. We study a small VPL actuator – about 6 nm when extended, and focus on the role of collective interactions by targeting monomers and trimers dynamics separately. The first part of this paper introduces the VPL NanoActuator and describes the techniques used to analyze its performance. The next part describes the results of performance parameters - pH variation and peptide mutation on the NanoActuator. In addition, we characterize different properties such as stability, energy changes, and the quality of the final structure. We perform all MD simulations in an implicit solvent environment.

Concept

The role of envelope glycoproteins (surface proteins) of various retroviruses in the process of membrane fusion is known to a great degree of accuracy. Such viruses infect their target cells by a process called Membrane Fusion (2). Membrane fusion is necessary for a large and diverse number of processes in biology such as protein trafficking, protein secretion, fertilization, viral invasion and neurotransmission. These processes require the fusion of distinct membranes to form a lipid bilayer. The mechanism is best understood among enveloped viruses such as the influenza virus (2-6). 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, the protein 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. Fig. 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 (7, 8). 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 sialic acid. HA1 binds to this receptor and hence brings the virus and cell together. Upon binding with sialic acid, the virus is endocytosed by the cell. HA remains dormant until the endosome begins to mature, and the pH in HA surroundings drops to 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 (see Fig. 2). 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.

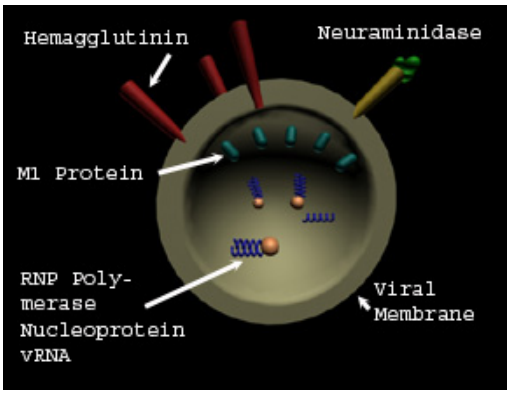


Figure 1 A schematic of influenza virus. Hemagglutinin (HA) polypeptides lies on the surface of the virus while viral RNA resides inside. HA has subunits that attach to cells and then facilitate membrane fusion in order to infect the cell.

Fig. 3 shows a cartoon drawing to illustrate the mechanism of viral infection by influenza virus. The green regions are the triple stranded coiled coils wound against each other. They are connected to smaller helices in the HA2 domain by loop regions shown in yellow. The smaller helices connect to the fusion peptide (residues 1-25). The globular HA1 domains are shown as spheres. They serve the purpose of attaching to the cellular receptor – sialic acid. After attachment and pH drop, the yellow loop regions convert into extended coiled coils effectively extending the green coils and moving the fusion peptide by 10 nm. The crystal structures of HA in both the native and the fusogenic conformations are known (9-12). As stated above, HA1 only serves the purpose of cellular recognition and binding to the cell surface receptor. HA2 subunit is then responsible for membrane fusion activity. In the native state, central region of HA2 folds as a helical hairpin structure (Fig. 3A).

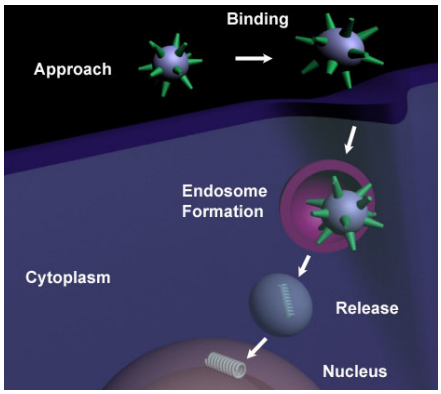


Figure 2 Various stages in influenza virus infection. The virus in endocytosed by the cell and an endosome is formed, wherein the conformational change occurs.

(A) **(B)** **(C)** **(D)**

Figure 3: A) Hairpin like structure of central HA2 non-fusogenic state monomer. It 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 brown). **B)** HA2 chains A and B in fusogenic states (complete structure equal to that in 3a is not available). The loop region shown in 3A (brown) converts into a relatively more rigid α-helix and forms a continuation of the long α-helix (blue) and the short α-helix is thus translated upwards along with the fusion peptide (not shown). **C)** The native state HA2 trimer. Small helices (yellow) are joined to the three coiled coils (blue helices) by loop regions (brown). The remaining structure (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-like regions now helical.

The hairpin shown in Fig. 3A is a monomer. 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 structural motif called the three stranded coiled coil. Coiled coils can be

visualized as an interwound rope of three interacting helices. The corresponding three short helices lie on the outside of the coiled coil (Fig. 3C). The 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.

The Viral Protein Linear Nano-Actuator

The conformational change mechanism in the Influenza virus membrane protein as described above has the potential to be used as nanoscale linear actuator. This mechanism converts chemical energy (pH) into mechanical energy, generating in turn a force and motion. Once the envelope protein is created separately, it can be made to undergo similar conformational change repetitively and reversibly to create controlled motion and the nano-scale. Furthermore, it can be interfaced with other elements such as nanotubes, DNA and other chemical molecules to create complex functional robotic assemblies at the nano scale. However, in order to accomplish this difficult task, step-by-step analysis and a careful design needs to be performed. Each individual element must be fully characterized and optimized on its own. In this study we study the performance of the VPL nanoactuator by performing dynamic simulations using molecular dynamics (MD) techniques.

Modeling

The Targeted Molecular Dynamics (15) approach allows elucidating forces and displacements generated by the acutator for different segments of the VPL protein. Since the entire influenza hemagglutinin sequence is large and has a complicated structure, we initially focus on a partial segment of the peptide, which is a 36 residue cut out of the VPL protein. The 36 residue segment forms a hinge region of the viral protein joining the two α-helical regions of each monomer (from now on, we will refer to this sequence as "*loop36")* (12, 13). An initial state as obtained from PDB (structure ID 1HGF) at pH of about 7.0, loop36 consists of a 15 residue long α-helical part with the remaining fragment in a random loop form. It is located in the segment B of the Influenza Hemagglutinin protein sequence from residues 54 to 89. The extended state of the peptide is obtained from crystal structure 1HTM in PDB. Loop36 has been recognized as being a critical segment for a pH dependant conformational change. The loop36 wild type sequence is:

```
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
```

Effect of Protonation and Mutation

Since the VPL actuator works in an environment where a drop in pH is the primary reason for the conformational change, it is important to incorporate it in the simulations. Certain amino acids take different states – protonated or unprotonated – depending on their pKa values and the pH of the environment. At a pH lower then their pKa, the HIS, ASP and GLU residues accept protons and exist in their protonated states. In case of VPL, the conformation change is known to occur in acidic pH of about 5.0. In order to simulate the actual environment all the titrable groups in the sequence, should hence be protonated. From the available software and hardware capabilities at present, it is not possible to perform protonation in real time; i.e. to protonate the amino acids after building the protein and then observe the conformation change. We generate separate protonated structures of the residues GLU, HIS and ASP and incorporate them into the sequence, replacing their corresponding un-protonated counterparts. In the loop36 peptide, there are eight GLU, two ASP and one HIS residues making a total

of eleven "protonatable" residues out of 36. A noticeable difference in the structure and energetics of loop36 can be observed after such protonation.

The structural and chemical properties of proteins are a function of their amino-acid sequence. Any change in this sequence leads to a change in these properties. We utilize this property by changing the sequence of the proteins (i.e. mutating). Our aim is to find the optimal sequence with an aim of designing the devices that meet specific requirements. These requirements can be for example, maximum force, maximum displacement, structural rigidity, desired α-helical content, etc. In the case of VPL, a mutation G22A (Glycine at position 22 replaced by Alanine) is known to improve α-helical content in the final state, hence improve the actuator function.

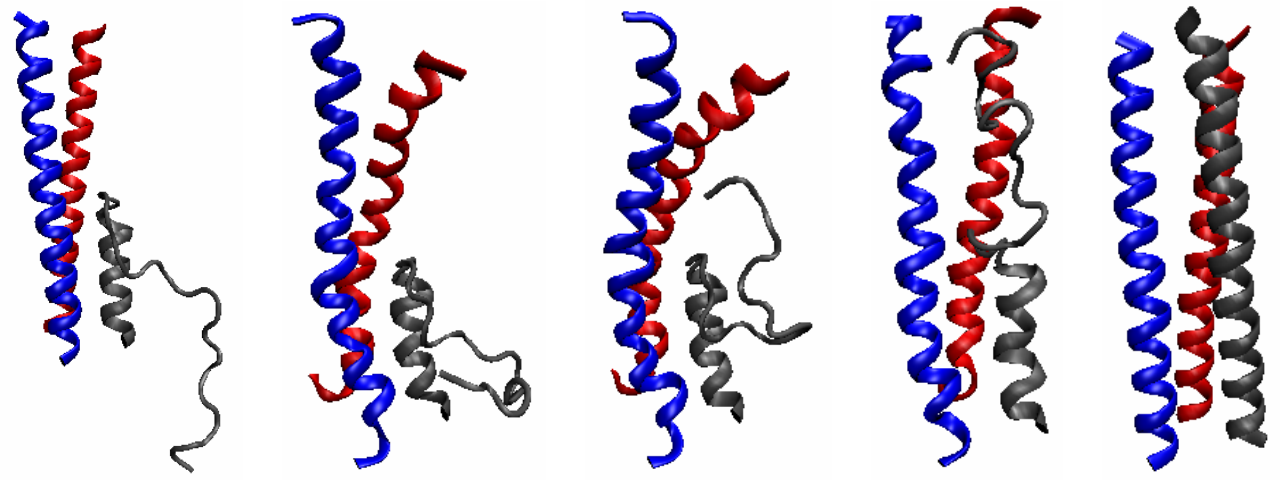
Results

The VPL peptide in the Influenza virus is naturally found in its trimer form; i.e. three identical peptides are located very close to each other due to the attractive hydrophobic interactions. The conformational change occurs in all three monomers leading to an open state of three longer α-helical coiled coil peptides. However, it is not clear whether the conformational changes in all three peptides occur at the same time. Simulations were performed using CHARMM force field (14) in order to study the behavior of one monomer (in the closed state) in company of the other two members of the trimer system in the open state as shown in Figure 5.

These simulations were performed with the four types of peptides –Wild Type (WT), Protonated, Mutated and Protonated-Mutated. For each case, the behavior of loop36 due to trimerization at 273K was analyzed in the following manner:

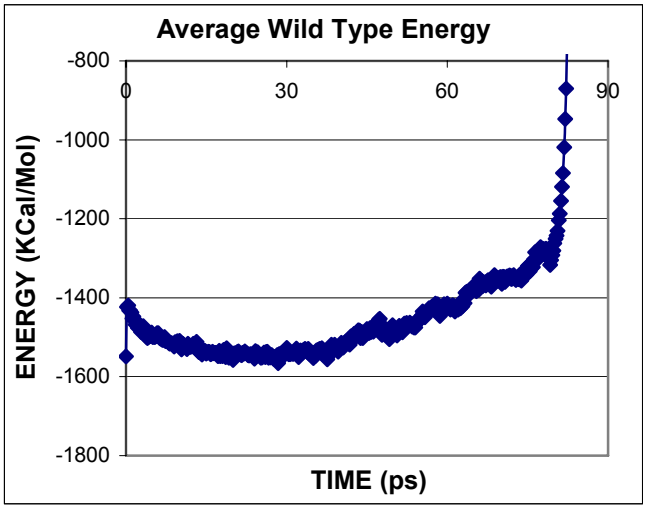
Let each of the monomers in loop36 be called L, M and N respectively. When monomer L is in closed state and M and N are in the open state, we call this setup as –COO (closed-open-open). Likewise in the cases when M and N are in closed states respectively, the setups are termed –OCO and –OOC. The conformational energy was averaged over the three setups and a single average curve was generated for each of the four peptide types. This averaging was done to prevent any local fluctuations from affecting the conclusions. The physical significance of this is that in a solution where there are monomers in both the closed state and open state, any closed monomer can come together with two other open monomers to form a trimer. The energy plots (Fig. 6) show that for the wild type monomer Fig. 6A) there is a continuous increase in the average conformational energy making the overall open state conformations are energetically unfavorable. The mutation M1 (GLY to ALA) appears to remove this extreme infeasibility of the open state (Fig 6B), however there needs to be a large energy barrier jump to achieve the open state. The protonated and protonated—mutated peptides show much smaller energy barriers and are the most likely candidates to achieve a stable open state (Figs. 6C-D). However the protonated (Fig. 6C) has a higher barrier as compared to the protonated-mutated (Fig. 6D). Since targeted molecular dynamics technique applies an external constraint on the peptide, it does achieve an open conformation in each of the four cases; but the energy profiles tell us which of the cases is the most feasible. Such characterization allows for the design of an optimal actuator based on its amino-acid sequence and the corresponding energy profile.

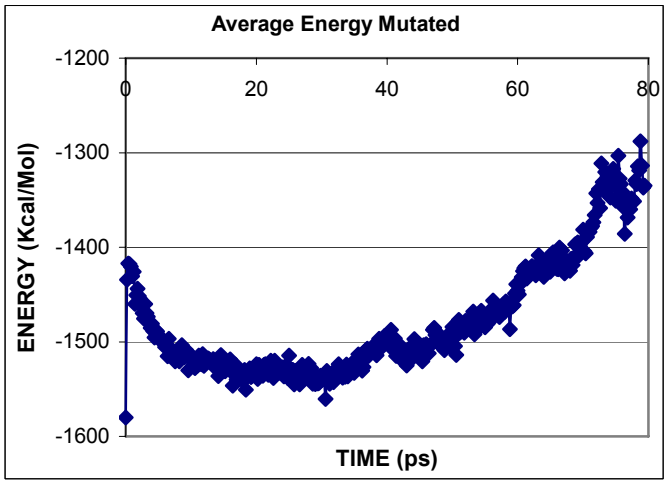
We hence deduce that this mutation in the presence of low pH at 273K temperature is a good candidate peptide for the creation of a VPL nanoactuator. Our simulations show that monomers of all four types can undergo a conformational change - the WT, protonated and G22A mutated will need some excitation while the PM will be spontaneous. However the open states might not be stable for the wild type and unprotonated mutated peptides.

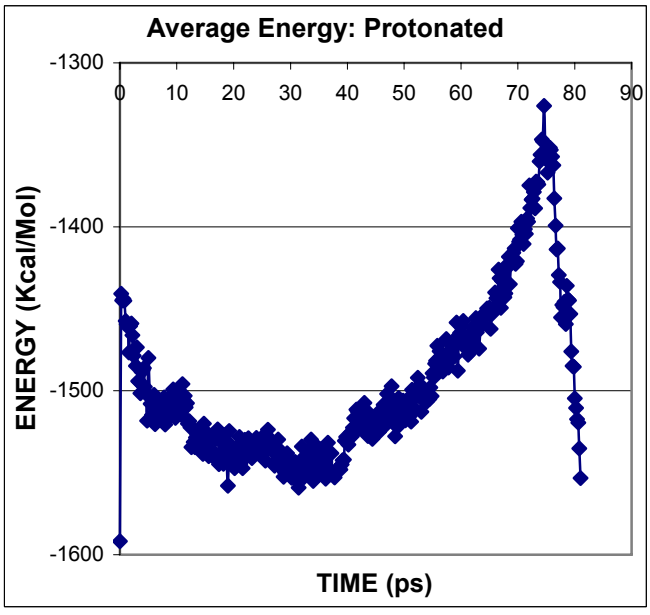


(A) (B) (C) (D) (E)

Figure 5: TMD progression of Loop36 trimer with one monomer in the closed state and other two in the open state.







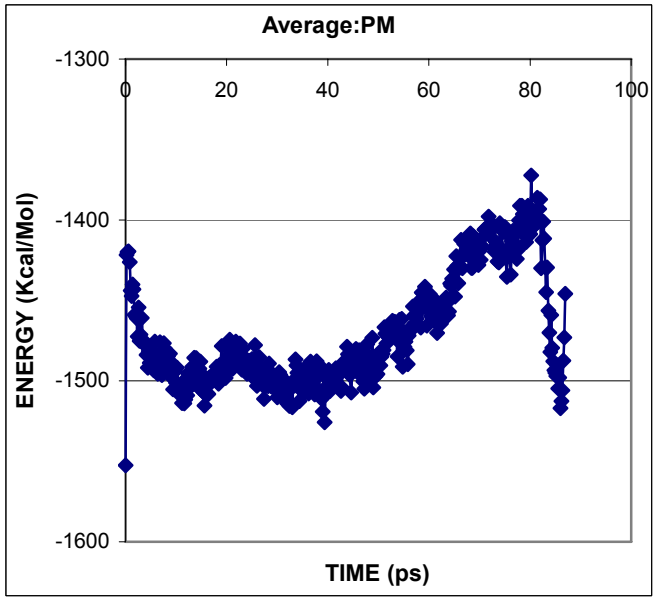


Figure 6 A) Average conformational energy of wild type loop36 trimer shows an unfavorable open state. **B)** The mutated peptides show less unfavorability even though the open states may be unstable. **C)** The protonated peptide shows stable open states but there is a high energy barrier. **D)** The energy barrier is reduced in the protonated-mutated system.

Discussion

Characterization of molecular scale devices with the aim of application in nano-machinery is one of the challenges of present times. In order for nanoscale engineering to be successful, simple and individually functional machine elements must be developed, characterized, and optimized before they are put together to form complex assemblies. We introduce one of the possible elements of such machine, the viral protein linear nanoactuator and describe how its inspiration has been taken from nature. We further apply molecular dynamics simulation approaches to study the actuator computationally. Due to the limitations imposed by present day computational power, we incorporate targeted molecular dynamics as a means to approximate the conditions that the system encounters. We tweak the system based on a previously known mutation and protonate with an aim to not only validate our computational setup but also to quantify the observed phenomena. We perform quantification by using energy profiles for monomers and trimers, which ultimately defines the feasibility of conformational change and stability of the target state.

All molecular dynamics simulations are performed in an implicit solvent environment to save computational time and resources while keeping the approximation within an acceptable limit of accuracy.

The model used in this study (loop36) describes a small system ideally suited to computational studies while maintaining the essential properties of the peptide. Larger peptides than loop36 can also be chosen, such as loop52 which has been studied experimentally (16) which will result in a larger actuator. The extended state loop36 actuator is about 6nm long.

The simulations on loop36 show that while the mutation G22A may not have an appreciable effect on the system performance in the absence of low pH, protonation alone without mutation will not help either. The WT and mutated peptides show similar energy profiles to the mutated protein profiles suggesting some probability of surviving in an open state. This is expected because the glycine residue was replaced with an intention of facilitating helix formation. A further analysis of more mutations in the loop36 as well as larger peptides will help to design peptides specific applications with desired performance. After protonation, the G22A mutated peptide shows a clear feasible pathway with a low energy barrier towards the open state indicating the success of this particular mutation.

References

1. C. D. Montemagno and G. D. Bachand, 1999, Constructing Nanomechanical Devices Powered by Biomolecular Motors, *Nanotechnology*, 10, 225-331.
2. J. Bentz, 2000, Membrane fusion mediated by coiled coils: a hypothesis, *Biophysical Journal*, **78:886-900**.
3. J. J. Skehel and D. C. Wiley, 2000, Receptor binding and membrane fusion in virus entry: the influenza hemagglutinin, *Annual Review of Biochemistry*, 69, 531-569.
4. D. C. Wiley and J. J. Skehel, 1987, The structure and function of the hemagglutinin membrane glycoprotein of influenza virus, *Annual Review of Biochemistry*, 56, 365-394.
5. C. Schoch, R. Blumenthal and M. J. Clague, 1992, A long-lived state for influenza virus-erythrocyte complexes committed to fusion at neutral pH, *FEBS Letters*, 311(3), 221-225.

6. W. Weissenhorn, A. Dessen, L. J. Calder, S. C. Harrison, J. J. Skehel and D. C. Wiley, 1999, Structural basis for membrane fusion by enveloped viruses, *Molecular Membrane Biology*, 16(1), 3-9.
7. S. G. Lazarowitz, R. W. Compans and P. W. Choppin, 1971, Influenza virus structural and nonstructural proteins in infected cells and their plasma membranes, *Virology*, 46(3), 830-843.
8. J. J. Skehel and M. D. Waterfield, 1975, Studies on the primary structure of the influenza virus hemagglutinin, *Proceedings of the National Academy of Sciences, U.S.A.*, 72(1), 93-97.
9. I. A. Wilson, J. J. Skehel and D. C. Wiley, 1981, Structure of the haemagglutinin membrane glycoprotein of influenza virus at 3 A resolution, *Nature*, 289(5796), 366-373.
10. W. Weis, J. H. Brown, S. Cusack, J. C. Paulson, J. J. Skehel and D. C. Wiley, 1988, Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid, *Nature*, 333(6172), 426-431.
11. W. I. Weis, A. T. Brunger, J. J. Skehel and D. C. Wiley, 1990, Refinement of the influenza virus hemagglutinin by simulated annealing, *Journal of Molecular Biology*, 212(4), 737-761.
12. P. A. Bullough, F. M. Hughson, J. J. Skehel and D. C. Wiley, 1994, Structure of influenza haemagglutinin at the pH of membrane fusion, *Nature*, 371(6492), 37-43.
13. C. M. Carr and P. S. Kim, 1993, A spring-loaded mechanism for the conformational change of influenza hemagglutinin, *Cell*, 73(4), 823-832.
14. R. Brooks, Bruccoleri R.E., Olafson B.D., States D.J., Swaminathan S., Karplus M., 1983, CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations, *Journal of Computational Chemistry*, 4, 187-217.
15. J. Schlitter, M. Engels and P. Kruger, 1994, Targeted molecular dynamics: a new approach for searching pathways of conformational transitions, *J Mol Graph*, 12(2), 84-89.
16. C. M. Carr, C. Chaudhry and P. S. Kim, 1997, Influenza hemagglutinin is spring-loaded by a metastable native conformation, *Proceedings of the National Academy of Sciences, U.S.A.*, 94(26), 14306-14313.