**Phase I NIAC Grant**

**Final Report**

# PROTEIN BASED NANO-MACHINES FOR SPACE APPLICATIONS

Submitted to:


Dr. Robert A. Cassanova, Director
NASA Institute for Advanced Concepts (NIAC)
555A Fourteenth Street, N.W.
Atlanta, Georgia 30318
Tel: (404) 347-9633, FAX: (404) 347-9638
Email: bcass@niac.usra.edu




Principal Investigator:
Constantinos Mavroidis, Ph.D., Associate Professor
Department of Mechanical and Aerospace Engineering
Rutgers University, The State University of New Jersey
98 Brett Road, Piscataway New Jersey 08854-8058
TEL: (732) 445-0732, FAX: (732) 445-3124
EMAIL: mavro@jove.rutgers.edu, WEBPAGE: http://cronos.rutgers.edu/~mavro
PROJECT'S WEBPAGE: http://bionano.rutgers.edu


December 2, 2002

# PROTEIN BASED NANO-MACHINES FOR SPACE APPLICATIONS

**Constantinos Mavroidis[1], Martin L. Yarmush[2], Atul Dubey[1], Angela Thornton[2], Kevin Nikitczuk[2], Silvina Tomassone[3], Fotios Papadimitrakopoulos[4], Bernie Yurke[5]**

1 Department of Mechanical and Aerospace Engineering
Rutgers, The State University of New Jersey
98 Brett Road, Piscataway, New Jersey, 08854-8058
TEL: (732) 445-0732, FAX: (732) 445-3124
EMAIL: mavro@jove.rutgers.edu

2 Department of Biomedical Engineering
Rutgers, The State University of New Jersey
98 Brett Road, Piscataway, New Jersey, 08854-8058
TEL: (732) 445-4346
EMAIL: gmuniz@rci.rutgers.edu

3 Department of Chemical & Biochemical Engineering
Rutgers, The State University of New Jersey,
98 Brett Road, Piscataway, NJ 08854
TEL: (732) 445-2972; FAX: (732) 445-2581
EMAIL: silvina@sol.rutgers.edu

4 Institute of Materials Science (IMS)
University of Connecticut, Storrs, CT 06269
TEL: (860) 486-3447, FAX: (860) 486-4745
EMAIL: papadim@mail.ims.uconn.edu

5 Bell Laboratories, Lucent Technologies
Materials Research Department
600 Mountain Avenue, Murray Hill, NJ 07974
TEL: (908) 582-4961, FAX: (908) 582-4868
EMAIL: yurke@lucent.com

## TABLE OF CONTENTS

## 1.0 OVERVIEW

This six month, Phase I NIAC project was undertaken from May 1 to October 31, 2002. A multidisciplinary team from Rutgers University, The State University of New Jersey and Shriners Hospital for Children, a Harvard Medical School affiliated hospital, has been assembled to study the development of protein-based nano-motors and nano-robots. Consultants from the University of Connecticut and Lucent Technologies' Bell Labs provided additional technical support.

The long term goal of this project is to develop novel and revolutionary biomolecular machine components that can be assembled to form multi-degree of freedom nanodevices that will be able to apply forces and manipulate objects in the nanoworld, transfer information from the nano to the macro world and also be able to travel in the nanoenvironment. These machines are expected to be highly efficient, economical in mass production, work under little supervision and be controllable. The vision is that such ultra-miniature robotic systems and nano-mechanical devices will be the biomolecular electro-mechanical hardware of future planetary missions. Some proteins, due to their structural characteristics and physicochemical properties constitute potential candidates for this role.

The specific aims of this project are:

a) To identify proteins that can be used as motors in nano / micro machines and mechanisms. We are focusing our studies on the mechanical properties of viral proteins to open or close depending on the pH level of environment. Thus, a new, powerful, linear biomolecular actuator type is obtained that we call: Viral Protein Linear (VPL) motor. Various viral proteins will be studied and from them different VPL motors will be produced.

b) To develop concepts for bio-nano-sensors using natural sensor proteins called Heat Shock Factor (HSF), along with DNA and optical fibers. Being on such a small scale and with a possibility of self-replication *in situ*, these sensors would be very economical to propel through long distances and enter into tiniest of holes, cervices, or places where there is a possible chemical hazard to human beings. They would also be helpful in the

detection of life forms in other planets due to their ability to gauge humidity and other life-supporting chemicals. They can also be used to measure temperature.

c) To develop dynamic models and realistic simulations / animations to accurately predict the performance of the proposed VPL motors.

d) To perform a series of biomolecular experiments to demonstrate the validity of the proposed concept of VPL motors.

e) To study the interface of the proposed protein motors and sensors with other biomolecular components such as DNA joints and carbon-nanotube rigid links so that complex, multi-degree of freedom devices, machines and robots are created.

The activities of the Phase I focused on the development of concepts for bio-nano-machine components and their interface with each other so that complex nano-devices can be produced. We performed preliminary computational and biomolecular studies to demonstrate the concept of the VPL nano-motor. Finally, we developed a detailed plan for experiments that will be performed in a proposed Phase II.

## 2.0 THE TEAM AND MANAGEMENT

Table 1 summarizes the responsibilities and level of involvement of the team members.

Computational studies were performed at Rutgers University. Experimental studies were performed at the Shriners Hospital for Children in Boston and Rutgers University. Both Principal Investigators have appointments in both institutions and had weekly meetings with all researchers at both institutions. They also had daily email and telephone discussions with all researchers of the project. Weekly email and telephone communication has been maintained with all consultants. A bi-monthly meeting was convened at Rutgers University where all team members gathered to review the project's progress and discuss the results.

**TABLE 1**: Roles and Responsibilities and Level of Involvement of the Team Members.

| Researchers / Role | Title / Institution | Responsibility |
|---|---|---|
| **Principal Investigators** | | |
| C. Mavroidis, Ph.D.<br>Principal Investigator<br>50% | Associate Professor of Mechanical Engineering at Rutgers University and Visiting Associate Professor of Surgery at Harvard Medical School and Shriners Hospital | Program Management<br>Research Supervision on Design, Modeling and Simulations of Protein Based Nanomachines |
| M. Yarmush, M.D., Ph.D.,<br>Co-Principal Investigator<br>10% | Professor and Chair of Biomedical Engineering at Rutgers University and Director, Center for Engineering in Medicine, Massachusetts General Hospital and Shriners Hospital | Research Supervision on Biochemical and Biomedical Components of the Project Including Experiments on Protein Based Nanomachines |
| **Researchers For Computational Studies** | | |
| Atul Dubey<br>Research Assistant<br>100% | Graduate Fellow, Department of Mechanical and Aerospace Engineering, Rutgers University | Modeling and Performance Evaluation of Protein Based Nanomachines |
| Kevin Nikitczuk<br>Research Assistant<br>100% | Undergraduate Student, Department of Biomedical Engineering, Rutgers University | Graphic Simulations and Animations |
| Chris Kanik<br>Research Assistant<br>100% | Summer Undergraduate Research Fellow, Department of Mechanical & Aerospace Engineering, Rutgers University | Modeling of VPL Motors<br>New Concepts of Bio-Nano-Devices |
| **Researchers For Experimental Studies** | | |
| Angela Thornton<br>Research Assistant<br>100% | Graduate Fellow, Department of Biomedical Engineering, Rutgers University | Biochemical and Biomolecular Experiments |
| Scott Banta, Ph.D.<br>Post-Doc<br>20% | Post-Doctoral Associate, Shriners Hospital | Protein and Metabolic Engineering |
| Jennifer Cusick<br>Technician<br>10% | Technician, Shriners Hospital | Biochemistry and Molecular Biology |
| Jennifer Mercury<br>Technician<br>10% | Genomics Core Facility Research Technician, Shriners Hospital | Biochemistry and Molecular Biology |
| **Consultants** | | |
| S. Tomassone, Ph.D.<br>Consultant<br>10% | Assistant Professor of Chemical and Biochemical Engineering at Rutgers University | Consulting and Advising on Computational Molecular Dynamics |
| F. Papadimitrakopoulos, Ph.D.<br>Consultant<br>10% | Associate Professor of Chemistry and Associate Director of the Institute of Materials Science at the University of Connecticut | Consulting and Advising on Carbon-Nanotubes and their Interface With Proteins and DNA |
| B. Yurke, Ph.D.<br>Consultant<br>10% | Senior Research Scientist, Lucent Technologies, Bell Labs | Consulting and Advising on DNA Joints and their Interface With Carbon-Nanotubes |

## 3.0 SIGNIFICANCE / MOTIVATION

It is well known that outer space, planetary, military and in-body medical missions and interventions will benefit tremendously by decreasing considerably the size, weight and cost of hardware and payloads. The development of biomolecular nano-components and devices may be the technological solution in this problem. These devices will be lightweight and hence easy and cost-efficient to be launched or introduced into remote or difficult to reach worlds. They will be designed to be self-replicating, a property that will help create computing stations and manufacturing sites on remote and inaccessible environment and in turn, develop a whole nano-scale industry.

One way to fabricate such devices is by using the design and machinery from nature's biomolecular world. ATPase rotary motors [1] or our proposed concept of VPL motors are good examples of biomolecular machine components. The advantage of nature's machine components is that they are highly efficient, well optimized and as it is obvious they have been tested and their performance is excellent. Moreover, there is a provision of recycling and uninterrupted production depending upon the need. Since these machines will not actually be living organisms, they will be able to work in the unfavorable environment offered by remote worlds. They will also be able to grow, care and protect various forms of life transplanted from Earth to other planets. Figures 1 and 2 show an example of a biomolecular nanorobot repairing an infected cell in a blood vessel.

Our goal is that just as conventional macro-machines are used to develop forces and motions to accomplish specific tasks, bio-nano-machines can be used to manipulate nano-objects, to assemble and fabricate other machines or products, to perform maintenance, repair and inspection operations. The advantages in developing bio-nano-machines include: a) energy efficiency due to their extremely fast intermolecular and interatomic interactions; b) low maintenance needs and high reliability due to the lack of wear and also due to nature's homeostatic mechanisms (self-optimization and self-adaptation); c) low cost of production due their small size and natural existence.

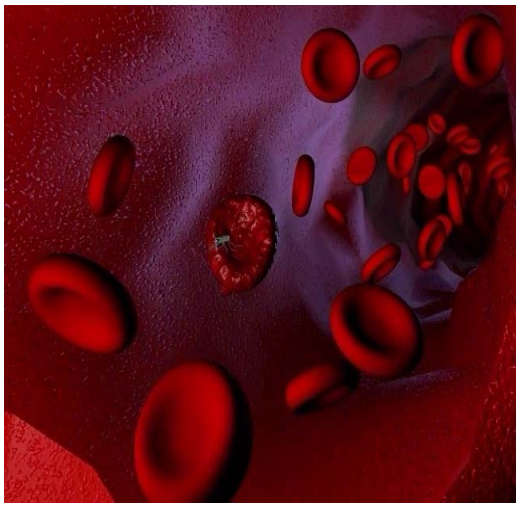

**FIGURE 1:** A "nano-robot" flowing inside a blood vessel, finds an infected cell.

**FIGURE 2:** The nanorobot attaches on the cell and projects a drug to repair or destroy the infected cell.

## 4.0 USING BIO-NANO-DEVICES FOR SPACE COLONIZATION

Peering into the future, we can envision a world where life does not take place before our eyes, yet at a level where the building blocks of life are interacting. This world of nanotechnology will enable us to explore, venture, and inhabit places beyond our current realms of reality. But to reach this state of technology, we must begin with the basics. We must understand the biological components that draw a parallel to current macro-robotics. With this knowledge in hand we can continue forward and join these components into assemblies. Some of these assemblies will execute specific tasks, while others perform a number of different operations. Eventually these bio-nano-robots will interact with one another, collaborating to build, repair, and manipulate other objects in the nano-world. Once these nano-robots are shown to sustain and create life, transporting them to far away planets will yield results not currently possible.

Development of robotic components from biological systems is the first step to future exploration (see Figure 3). We envision that this phase will last approximately 10 years from now. Since the planned systems and complex devices will be composed of these components, we must have an understanding of how they operate and how to control them. From the simple elements such as structural links to more advanced concepts as motors, each part must be carefully studied and possibly manipulated to understand the functions and limits of each

element. DNA and carbon nanotubes have already been manufactured into many shapes to build possible complex devices. These structures may house proteins such as rhodopsin and bacteriorhodopsin. Both are naturally found in biological systems as light sensors. These proteins can essentially be used as solar collectors to gather abundant energy from the sun and either harvest this energy for later use or have other systems use this energy immediately. Motors attached to the structure such as the ATPase protein will use this energy as a means of moving the device. Sensors, like Heat Shock Factor, attached to the structure will be activated when needed and used as planetary probes, gathering the required information.

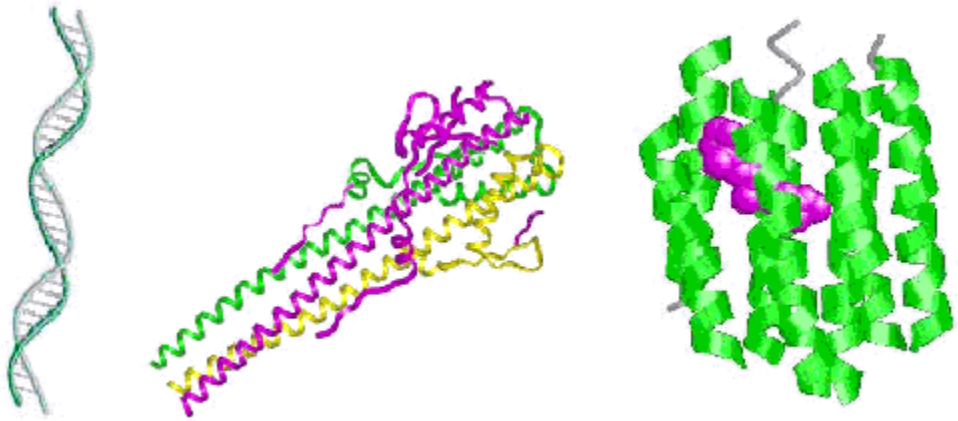

**FIGURE 3:** 0-10 Years: Understanding of basic biological components and controlling their functions as robotic components. From left we have: DNA which will be used in a variety of ways such as a structural element and a power source; hemagglutinin virus used as a VPL motor; bacteriorhodopsin as a sensor and power source.

The next step is to assemble the many components into robotic systems. Figure 4 shows one such concept of a nano-organism, with its 'feet' made of helical peptides and its body using carbon nanotubes while the power unit is a biomolecular motor. We believe that this phase will correspond to the next 10-year period, from 10-20 years from now. By the beginning of this phase we will have a library of biological elements of every category. We will then essentially be able to take macro-robotic systems and fabricate bio-nano-robots that function in the same manner. There will be systems that have mobile characteristics to transport themselves as well as other objects to desired locations. Some bio-nano-robots will manufacture more elements and various structures, like infrastructures from *in situ* resources to house manufacturing sites and storage areas, while others manipulate existing structures by repairing damaged walls or making other renovations. There will also be robots that not only perform physical labor, but also sense

the environment and react accordingly. There will be systems that will sense an oxygen deprivation and stimulate other components to generate oxygen, creating an environment with stable hemoeostasis.

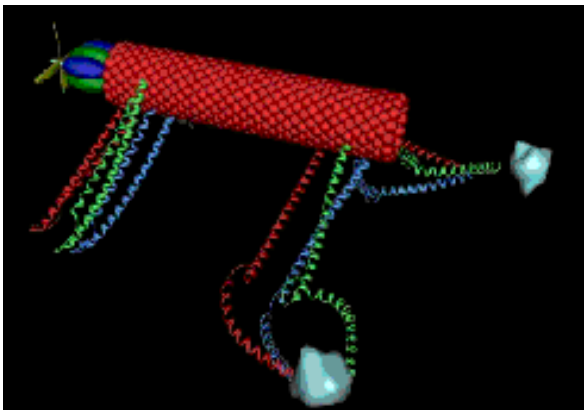

**FIGURE 4:** 10-20 Years: The biological elements once in study will now be used to fabricate robotic systems. A vision of a nano-organism: carbon nano-tubes form the main body; peptide limbs can be used for locomotion and object manipulation a biomolecular motor located at the head can propel the device in various environments.

With the nano-robots in full function, they will now need to collaborate with one another to further develop systems and "colonies" of similar and diverse nano-robots. We envision that this phase will take place between 20 to 30 years from now (see Figure 5). Since most of these robots will be composed of the basic fundamental elements of life, they will be able to self-replicate and multiply. They will produce identical copies of themselves and eventually variations of themselves to adapt to the environment and conditions. They will in effect create new life forms, perhaps biological systems, like plants, that will one day sustain human life on distant planets. The capabilities of robots to replicate and mutate into diverse structures would be valuable in discovering and designing stronger materials and exotic chemicals with unique properties.

Sending these robots on planetary exploration and colonization is the final step. This phase can take place in 30 to 50 years from now (see Figure 6). Due to the fact that these robots will be on the nanometer scale, the cost to transport them will be minimal. Along with being cost efficient, they will be able to operate under little supervision, with minimal human intervention. When the time arrives for these missions, we will have electronic, computational, and mechanical bio-nano-robots. They will initially explore the environments and provide feedback of the present conditions. These robots will take samples of the environment and analyze them, sending

information to the macro-world. There will be measurements taken of the planet for various factors as temperature and acidity, as well as detecting water and/or oxygen.

**FIGURE 5:** 20-30 Years: Robotic systems will be self-sustaining and replicate to produce colonies and new life. These robots will communicate with each other to sustain an environment in which to live.

The nano-robots will begin to colonize and build a new world, a world where we may one day thrive. Carrying out construction projects in hostile environments is a prime example of the benefits of these machines. Self-replicating robots, utilizing local materials and local energy will assemble space habitats that can be completely constructed by remote control. Other robots will be used not to colonize, but to maintain life by providing proper energy needed to other systems or organic organisms. Nanometer size computers, actuators, and sensors will allow robots to rebuild damaged parts of existing structures, such as walls, transportation vessels, and even space suits. There will be a new world built from bio-nano-robots that function on the nanometer scale, building and maintaining an environment where they, and we, may one day live. Figure 6 shows a vision of an established nano-industry on a planetary surface, composed of nano-computing stations, nano-manipulators and other mechanical systems, and thus, providing the foundations towards establishing colonies on uninhibited planets as well as searching for life forms on other planets.

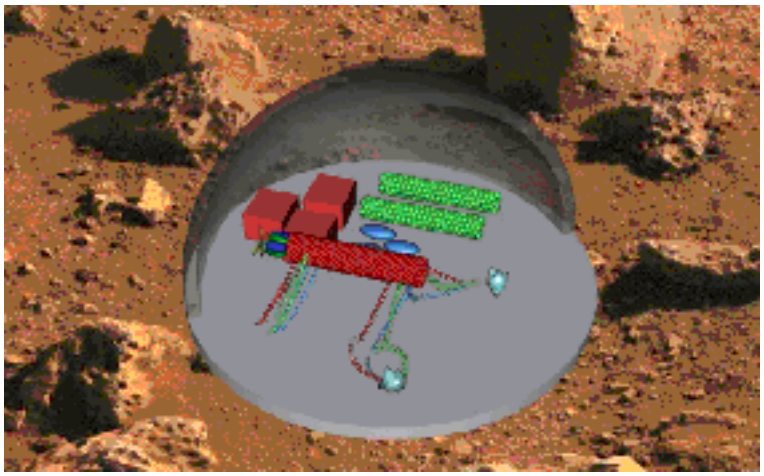

**FIGURE 6:** 30-50 Years: Future missions to planets will be undertaken by bio-nano-robots, in which they will survey and colonize the planets. Robots will be designed to build and maintain the environment that they have created, which will one day be stable for human colonization.

## 5.0 BACKGROUND STUDY ON BIOMOLECULAR NANODEVICES

In this section we present briefly today's state-of-the-art on the development of bio-molecular machine components and systems. While, the majority of the prior research in this field has been focused on biomolecular motors, several other nano-components such as sensors and even assemblies of components in the form of mechanisms have been studied.

In the macroscopic world, what we understand by a 'motor' is *a machine capable of imparting motion associated with the conversion of energy.* While a number of such devices, like electrical motors or gasoline motors have proven indispensable to mankind, there is an interest in utilizing motors, which have always been an integral part of all life. These motors, called *Biomolecular Motors* have attracted a great deal of attention recently due to the fact that they are highly efficient, they could be self-replicating, hence (cheaper in mass usage), and they are readily available in nature. A number of enzymes such as kinesin [2,3], RNA polymerase [4], myosin [5], and adenosine triphosphate (ATP) synthase function as nanoscale linear or rotary biological motors [6-11].

### *The ATPase Motor*

One of the most abundant rotary motors found in life forms is $F_0F_1$ ATP synthase, commonly known as the ATPase Motor. A vital cellular process called Oxidative Phosphorylation was

demonstrated over 50 years ago as an important process by which our bodies capture energy from the food we eat. How this process actually occurs was not known until 1997, when Nobel Laureates Paul D. Boyer and John Walker demonstrated that Adenosine Triphosphate (ATP) plays a key role in the process [1, 11]. The adenosine part acts as a handle that attaches ATP to enzymes, whereas the three phosphates take part in the energy capture. When the energy stored in ATP is used, the terminal anhydride bond is split, and ATP hydrolyses into adenosine diphosphate (ADP) and inorganic phosphate (Pi). An enzyme called ATP synthase, which is present in abundance in intracellular membranes of animal mitochondria, plant chloroplasts, bacteria, and other organisms, catalyzes the re-synthesis of ATP, coupled to energy input. The ATP made by ATP synthase is transported out of the mitochondria and used for the function of muscle, brain, nerve, kidney, liver, and other tissues, for transport and for making a host of compounds that the cell requires. The ADP and phosphate formed when ATP is used return to the mitochondria, and ATP is synthesized again using the energy from the oxidations.

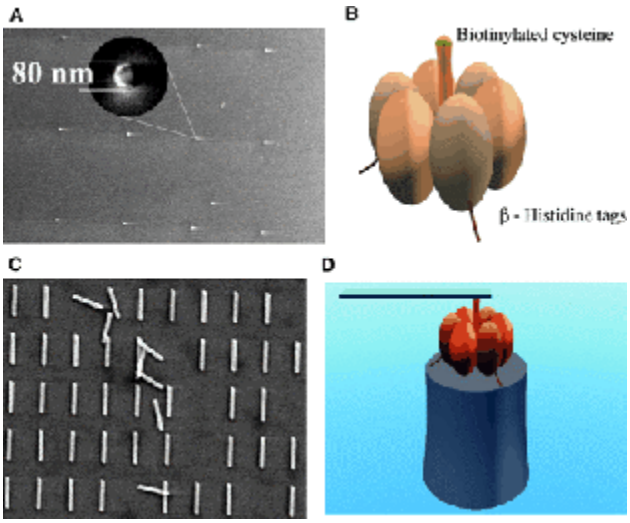


**FIGURE 7:** Schematic diagram of the $F_1$-ATPase biomolecular motor-powered nano mechanical device. The device consisted of (A) a Ni post (height 200 nm, diameter 80 nm), (B) the $F_1$-ATPase biomolecular motor, and (C) a nanopropeller (length 750 to 1400 nm, diameter 150 nm). The device (D) was assembled using sequential additions of individual components and differential attachment chemistries [6].

Noji et al published the structural and performance data of the ATPase motor in 1997 [9,10]. According to this study, the γ subunit, which is about 1 nm in diameter, rotates inside the F1 subunit, which is about 5 nm in diameter, to produce approximately 40 pN-nm of rotary torque. Carlo Montemagno and his group were the first to indicate that the rotation of the γ subunit of

the ATPase motor could be mechanically useful as they have been able to fabricate nanomechanical inorganic devices – which could be compatible with the force production and dimensions of the molecular motors [6-8]. They further showed this integration by making (i) engineered, nanofabricated substrates of nickel (Ni) posts; (ii) recombinant F1-ATPase biomolecular motors specifically engineered to selectively interface with nanofabricated structures; and (iii) engineered nanopropellers [7, 8] (see Figure 7). Frasch's group at Boston College is studying the binding of metals to amino acids of the motor protein. These experiments are providing new insights into the means by which the energy obtained from the hydrolysis of ATP can be converted into the physical action of pumping a proton in a unilateral direction [12].

***Kinesin and Myosin***

Motor proteins are tiny vehicles that transport molecular cargoes around inside cells. These minute cellular machines exist in three families- the kinesins, the myosins and the dyneins. There are over 250 kinesin-like proteins, and they are involved in processes as diverse as the movement of chromosomes and the dynamics of cell membranes. They all have a similar catalytic portion, known as the motor domain, but beyond this they are astonishingly varied - in their location within cells, their structural organization, and the movement they generate [13]. Modern microscopy has transformed our view of the cell interior from a relatively static environment to one that is churning with moving components. Kinesin and myosin are two kinds of motor proteins that utilize energy released by ATP and are responsible for motility. Molecular motors that move unidirectionally along protein polymers (actin or microtubules) drive the motions of muscles as well as much smaller intracellular cargoes. Myosin is an actin-based motor, whereas conventional kinesin transports membrane organelles along microtubules. Muscle myosin, whose study dates back to 1864, has served as a model system for understanding motility for decades. Kinesin, discovered using in vitro motility assays in 1985, is a relative newcomer by comparison. Conventional kinesin was found to be a highly processive motor that could take several hundred steps on a microtubule without detaching [14,15], whereas muscle myosin was shown to execute a single "stroke" and then dissociate [16]. A detailed analysis and modeling of these motors has been done [17].

Prof. Hackney's group at Carnegie Mellon University has concentrated upon the usage of ATP energy by motors like kinesin, myosin, dynenin and related motor families [18]. Unger's group in

IMB Jena is working towards developing a microtubule-kinesin system as a biological linearmotoric actuator. They are currently attempting to produce force multiplication by parallel action of numerous single driving units as well as a more efficient means for system control [19]. Researchers at the University of North Carolina have discovered a new member of the myosin-V family (Myo5c) and have implicated this myosin in the transport of a specific membrane compartment [20]. The role of ATP hydrolysis in kinesin motility has been explained in ref. [13].

### *The Flagella Motors*

*Escherichia coli* is a single-celled organism that lives in the human gut. It is equipped with a set of rotary motors only 45 nm in diameter. Each motor drives a long, thin, helical filament that extends several cell body lengths out into the external medium. The assemblage of motor and filament is called a *flagellum*. The concerted motion of several flagella enables a cell to swim. A cell can move toward regions that it deems more favorable by measuring changes in the concentrations of certain chemicals in its environment (mostly nutrients), deciding whether life is enhancing or debilitating, and then modulating the direction of rotation of its flagella. Thus, in addition to rotary engines and propellers, *E. coli's* standard accessories include particle counters, rate meters, and gearboxes. This microorganism is a nanotechnologist's dream [21].

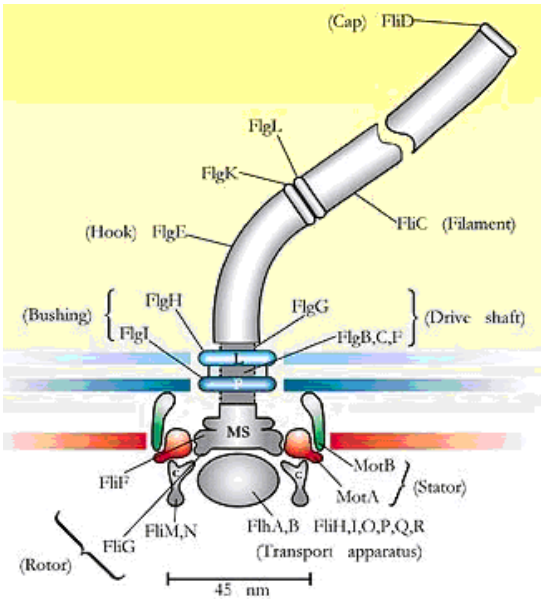


**FIGURE 8:** Flagella Motor and Drive Train. Shown are the L ring, the P ring, and MS ring and the C ring. Also shown are the stator, the rotor, the drive shaft, the bushing and the filament that is driven to achieve motion [21].

The flagellum is an organelle that has three parts (see Figure 8). There is a basal body consisting of a reversible rotary motor embedded in the cell wall, beginning within the cytoplasm and ending at the outer membrane. There is a short proximal hook, which is a flexible coupling or universal joint. There is also a long helical filament, which is a propeller. Torque is generated between a stator connected to the rigid framework of the cell wall (to the peptidoglycan) and a rotor connected to the flagellar filament. The proteins MotA and MotB are thought to constitute the elements of the stator- FliF, G, M, and N (the MS and C rings), those of the rotor- FlgB, C, F, and G, those of the drive shaft- and FlgH and I (the L and P rings), and those of the bushing that guides the driveshaft out through the outer layers of the cell wall.

The motor is driven by protons flowing from the outside to the inside of the cell (except for marine bacteria and bacteria that live at high pH, where sodium ions are used instead). There have been several models proposed for the performance evaluation and emulation of flagella motors. However, a bulk of our knowledge about flagella motor is due to the research of Howard C. Berg, who started working on the motile behavior of bacteria since 1971. His work was initially focused on the bacterial motility (1971-1975) and then he broadened his scope to motility in microorganisms in general. He gave the one of the earliest models for the rotary motor [22]. Improved models came in 1992 [23, 24]. Flagella motor analysis coupled to real-time computer assisted analysis of motion has been performed at The Albert Einstein College of Medicine [25]. Researchers in Japan have applied crystallographic studies in order to understand the molecular structure of flagella motors as well as that of kinesin [26]. Henry Hess' group is trying to build a nanoscale train system, complete with tracks, loading docks and a control system. This effort is inspired by nature, since cells have evolved a complex transport system. In this system specialized motor proteins connect to small containers filled with proteins and transport them along the skeleton of the cell. Since motor proteins are a thousand times smaller than any man-made motor, they try to utilize them in a synthetic environment as the engines powering the nano-trains [27].

***Other Motors and Mechanisms***

Other than working with the naturally existing motors, attempts are being made to create synthetic molecular motors. The structure of the ATP synthase – a rod rotating inside a static wheel – suggests Rotaxanes as potential artificial models for natural motors [28]. Rotaxanes are

organic compounds consisting of a dumbbell-shaped component – which incorporates one or more recognition sites in its rod section and is terminated by bulky 'stoppers' – encircled by one or more ring components. The possibility of manufacturing specific forms of rotaxane and creating molecular motors capable of guided, unidirectional rotary motion and the possibility of fueling such a motor by light, electrons and chemical energy has been proposed [29].

In addition to linear and rotary motors, there are quite a few biological mechanisms that have recently been recognized as candidates for the famous 'bottom up' manufacturing. Schemes for the use of Pseudorotaxanes, Rotaxanes and Catenanes as molecular switches and perform chemical, electrochemical and photochemical switching, and controllable molecular shuttles have recently been proposed [30]. Molecular shuttles have been reported using $\alpha$-cyclodextrin – a parent of rotaxanes and catenanes [28]. A light-driven monodirectional rotor made of helical alkene, with rotation around a central Carbon-Carbon covalent bond due to chirality has been reported [31]. Another simple way to convert chemical energy into mechanical motion in a controlled fashion has been shown in [32], wherein a metal ion can be translocated reversibly between two organic compartments with the change of its ionization state, controllable by redox reaction or pH change. A significant way of motility of unicellular organisms like vortecellids reminds us of energy storage and release by mechanical springs on a macromolecular scale. Spring-like action has been observed in sperm cells of certain marine invertebrates during fertilization. Springs and supramolecular ratchets by actin polymerization are yet to be built in vitro, but they theoretically contain good promiscuity, as [33].

***DNA-based Molecular Nanomachines, Joints and Actuators***

Several researchers are exploring the use of DNA in nanoscale mechanisms. Compared with protein structures, DNA is small, simple and homogeneous, and its structure and function is well understood. The generally predictable self-assembling nature of the double helix makes it an attractive candidate for engineered nanostructures. This property was exploited to build several complex geometric structures, including knots, cubes and various polyhedra [34]. Mathematical analyses of the elastic structure of DNA using energy minimization methods were performed to examine its molecular stability, wherein short DNA strands were treated as an elastic rod [35]. Initial physical experiments on DNA visualization and manipulation using mechanical, electrical, and chemical means have been underway for a decade [36, 37]. A dynamic device providing

atomic displacements of 2-6 nm was proposed in [38], wherein the chemically induced transition between the B and Z DNA morphologies acts as a moving nanoscale device. A method for localized element-specific motion control was seen in the reversible transition between four stranded topoisomeric DNA motifs (PX and $JX_2$) thereby producing rotary motion [39]. A very important, though simple DNA machine that resembles a pair of tweezers was successfully created, whose actuation, opening and closing, is also fueled by adding additional DNA fuel strands [40] (see Figure 9).

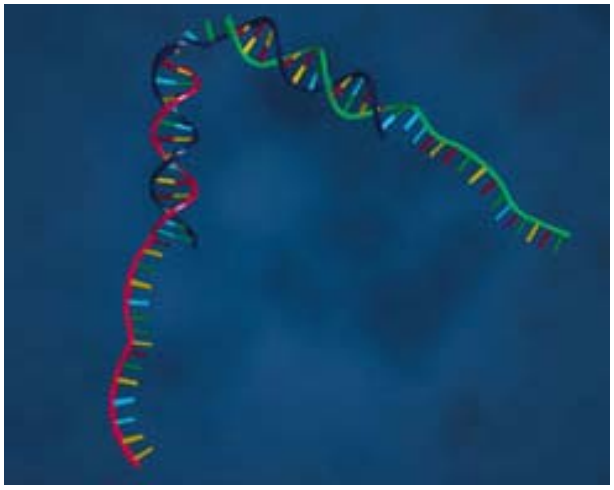

**FIGURE 9:** DNA hinge/joint with three strands A, B and C (red, black and green respectively). Introduction and subsequent removal of a fuel strand F facilitates the closing and opening of the tweezers [40].

***Nanosensors***

The technology of nanosensing can be broadly divided into two classes – a) developing nanoscale devices to sense environment/processes at this scale, and, b) developing sensors which are at the nanoscale or have nanoscaled parts and are capable of performing at a remote site. For the former problem silicon probes with single walled carbon nanotube tips are being developed (MIT media library) [41]. However for the development of nanodevices capable of remote sensing, responding and transmitting signals, a great deal of work needs to be done. The parameters that have to be measured are, but not limited to, temperature, pH, humidity, electronic charge, chemical concentration, radiation etc. With the threat of biological weapons increasing and becoming a reality, there is a need to have miniature devices that respond to such chemicals and send warning signals in advance. For sensing certain analytes, genetically engineered versions of pore-forming proteins like *Staphylococcus aureus* alpha-hemolysin are

being studied at NIST [42]. Efforts to detect biological warfare agents like cholera toxins by utilizing their ability to bind to a bilayer membrane in the presence of gangliosides are being made [43]. Light sensors could be made using certain photoreceptive polypeptides containing azobenzene or spyropyran units as they respond to light or dark environmental conditions by undergoing conformational change, e.g. transition from random to a $\alpha$-helix [44]. Such peptides could also serve as the 'secondary' sensors in remote sensing as in we have proposed heat shock proteins that respond to environmental changes by binding to DNA and producing a color change in the solution which could be picked up by these to produce a 'mechanical' effect as an output, instead of light. An optical DNA biosensor platform was developed using etched optical fiber bundles filled with oligonucleotide-functionalized microsphere probes at Tufts University [45]. The DNA biosensor array is capable of gene expression analysis, Single Nucleotide Polymorphism (SNPs) discrimination, or detection of DNA linked to disease states. Sensors with important future medical applications are being developed at the University of Newcastle in UK. Work is in progress to develop implantable sensors inside human brain, which would foretell a possibility of a stroke and perioperative on-line monitoring during coronary by-pass surgery [46].

## 6.0 SYSTEM CONCEPTS AND ARCHITECTURES

In Phase I we developed concepts for various bio-nano-machine components that we present in this section.

### 6.1 The VPL Motor and Its Types

In this project, we are focusing on the mechanical properties of viral proteins to transit between two conformations depending on the pH level of environment. Thus, a new linear biomolecular actuator type is obtained that we call: Viral Protein Linear (VPL) motor.

So far, we know, to a great degree of accuracy, the role of envelope glycoproteins of various retroviruses for the process of membrane fusion, which is a process necessary for the virus to be able to infect a cell. During the process of membrane fusion, there is a distinct conformational change in the peptide on the viral surface as it 'readies' itself for infecting the cell. This change is due to the pH change associated with the vicinity of the cell. Given similar conditions, it is proposed to use this conformational change to produce VPL motors.

We have selected the following viral peptides for further consideration in this project:

1) The Influenza virus protein Hemagglutinin (HA) peptide HA2 [47]
2) The Human Immunodeficiency Virus type 1 (HIV 1) peptide gp41 [48]
3) The Human Respiratory Syncytial Virus (HRSV) protein subunit F1 [49]
4) The Simian Immunodeficiency Virus (SIV) protein gp41 [50]
5) The Human T cell Leukemia virus type 1, protein gp21 [51]
6) The Simian Parainfluenza Virus peptide unit SV5 [52]
7) Ebola virus protein gp2 [53]

Each one of these peptides can result in a different VPL motor that can have different properties such as different weight, volume, range of motion, force and speed capabilities. However, the principle of actuation is the same. Studies have shown that the common characteristic in these viruses is the structure of a portion of the surface protein (envelope glycoprotein) and the mode of infection. The envelope glycoproteins of these viruses can be divided into two subunits, which are a result of proteolytic cleavage of a common precursor protein. The two subunits have different functions. For example, in the case of 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 (above listed) viruses acquire an alpha-helical conformation 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 Figure 10. This coiled coil structure undergoes a conformational change induced by mildly acidic conditions (i.e. pH around 5). This change is required for the process of membrane fusion, i.e. the fusion of viral and cellular membranes essential for infection of the cell. With the change in pH, the N-terminals pop out of the inner side and the peptide acquires a straightened position or the fusogenic state.

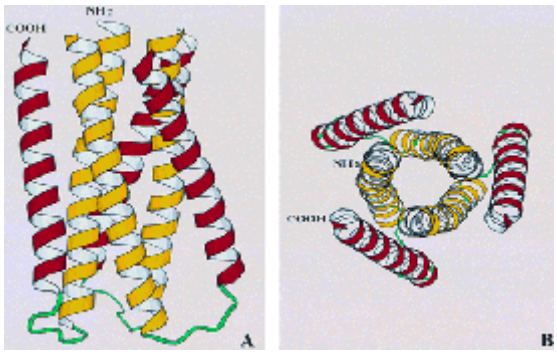


**FIGURE 10:** Ribbon drawing of the trimer with the N and the C terminus of one molecule labeled. Also shown is the end view of the trimer, looking down the 3-fold axis [54].

Figures 11 and 12 show a schematic of the VPL motor supporting a moving platform. The motor is shown in its initial, "contracted" phase that corresponds to the virus' native state (Figure 11) and at its extended, fusogenic state (Figure 12). 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 opening process takes about the same time. A rough estimate of the velocity of the actuator end can be made by this data as 2.5-5 m/s. By estimating the energy release during the transition from native to fusiogenic state, we have estimated that the actuator force of the proposed motor will be approximately equal to 21 pN. Compared to other bio-molecular nano-motors such as the muscle myosin, which can develop a force of 5 pN, the VPL motor will be a very powerful actuator. To augment the force capabilities of the VPL motors, several VPL actuating elements could be attached in parallel as it is shown in Figure 13. 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 Figure 13).

In this project, computational and experimental studies will be performed using the Influenza virus protein Hemagglutinin (HA) as the basis for forming a VPL motor. The reason for making this peptide selection, is that, based on current literature, this peptide seems to be able to perform repeatable motion controlled by variation of the pH. In the case of the influenza virus, membrane fusion is mediated by the protein hemagglutinin (HA). HA is composed of two subunits, HA1 and HA2. The sialic acid binding domain is located on the HA1 subunit, while the fusion

peptide region and the transmembrane segment are located on the HA2 subunit (shown in Figure 14).

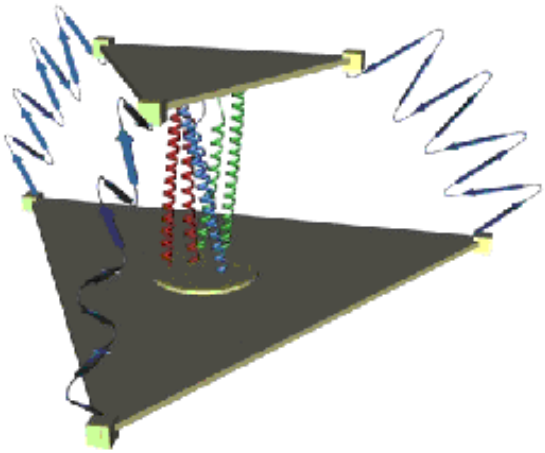

**FIGURE 11:** The titin fibers can be used as passive spring elements to join two platforms and form a single degree of freedom parallel platform that is actuated by a viral protein linear (VPL) actuator (center).

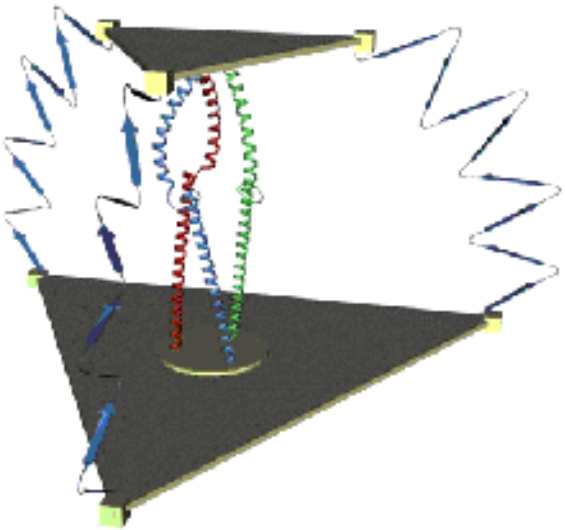

**FIGURE 12:** The VPL actuator has stretched out resulting in the upward linear motion of the platform. The three titin fibers acting as springs are also stretched out. Their elastic behavior can be used as a passive control element or as the restitution force that will bring the platform back to its original position.

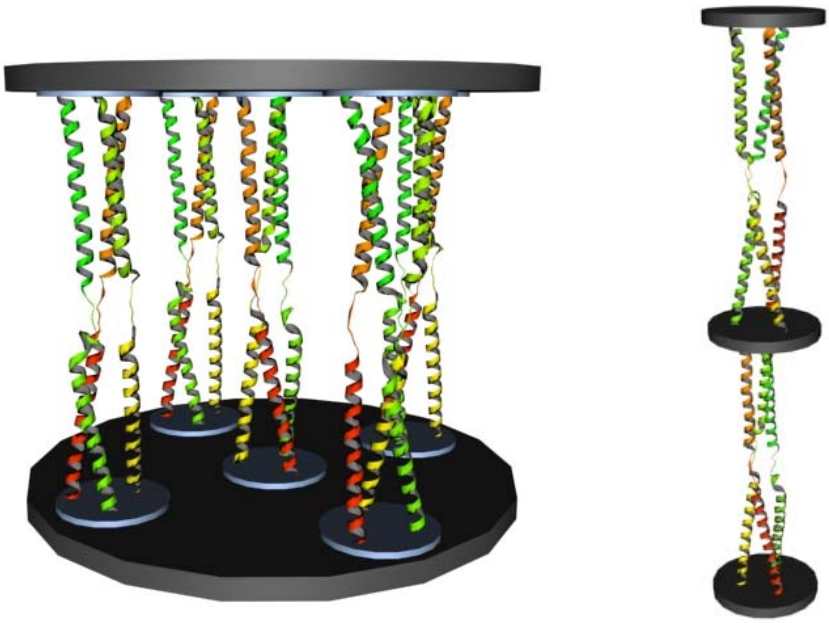

**FIGURE 13:** Several VPL motors placed in parallel (left) and series (right) to multiply force and displacement respectively.

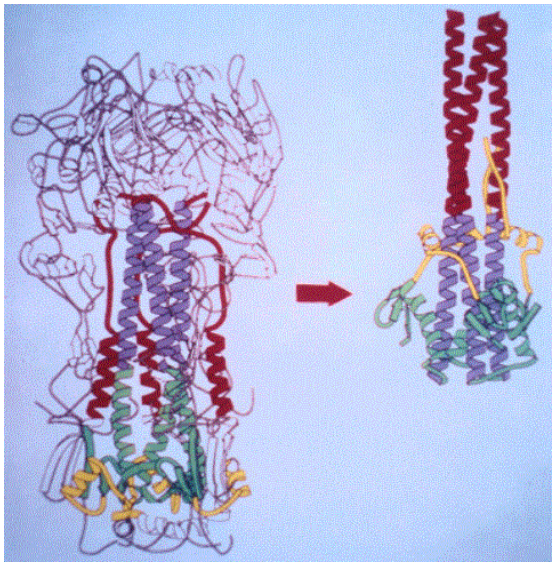

**FIGURE 14:** Influenza Hemagglutinin HA2 peptide shown in contracted, native state and in extended, fusogenic state [55].

When HA binds to the sialic acid on the surface of the target cell, the virus is internalized by a cellular endosome. The low pH of the endosome - between 5.0 and 6.5 - activates a conformational change in HA. The fusion peptide is projected 10 nm, acquires a straightened conformation, and inserts into the target cell, bringing the target cell close enough to the transmembrane region to cause fusion [56, 57].

### 6.2 The Heat Shock Factor / DNA Based Optical Bio-Nano-Sensor

As a mechanism against heat shock and other environmental stresses, virtually all organisms respond by a dramatic and rapid increase in the synthesis rates of a small number of protein chaperones. These proteins work to combat the deleterious effects of heat shock by binding to partially unfolded proteins and act to prevent their aggregation and to facilitate their refolding [58]. The promoters of the heat shock protein genes in eukaryotes contain a heat shock element (HSE). The protein that binds to this HSE in response to stress is termed heat shock factor (HSF) [59]. Upon activation, these heat shock factors trimerize and function as transcriptional activators that bind with similar specificities to the HSE, thus regulating heat shock gene transcription [60] (see Figures 15 and 16). This cellular defense system acts to protect cells from heat shock as well as from heavy metals, ethanol, developmental programs, and many diseases. However, the most common stresses that induce the heat shock response are elevated temperature and oxidative stresses. As observed through studies [59], maximal activation of purified, deactivated

HSF occurs at the heat shock temperature of 43 ºC, with less activation observed at temperature above or below 43 ºC. Through other experiments [58], it is found that HSF responds directly to $O_2^-$ but not to hydrogen peroxide or the hydroxyl radical. It was also found that $Mg^{+2}$ acts indirectly on HSF by stimulating the production of $O_2^-$. Superoxide anion ($O_2^-$) is produced during partial oxygen deprivation or during the recovery from anoxia that occurs upon reperfusion after ischemia.

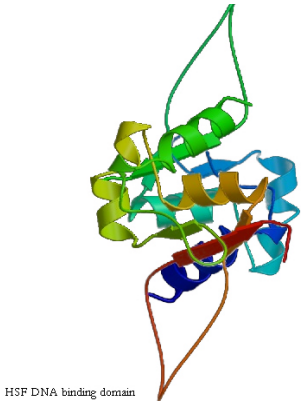


**FIGURE 15:** A HSF trimer ribbon-diagram – the coils are individual helices of the trimer [75].

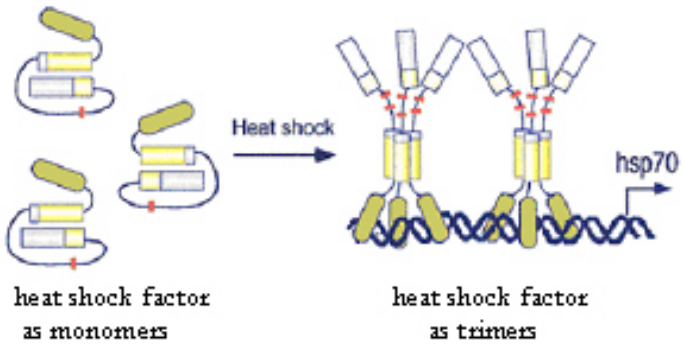


**FIGURE 16:** HSF monomers trimerise into a coiled coil structure upon experiencing heat shock and other chemical stimuli [63].

After HSF monomers are stimulated and undergo a molecular rearrangement, these trimers bind DNA. However, transcriptional activation is not a necessary consequence of DNA binding. Treatment of human cells with salicylate induces HSF to trimerize and bind DNA but not to activate transcription [61]. If the activation and binding of HSF to DNA can be controlled as previously shown, coupling of HSF with fiber optic DNA microarrays can provide a sensor that is activated by heat, or other previously mentioned stimuli, and transmit the signal into light pathways. The proposed DNA biosensor platform would contain optical fiber bundles filled with oligonucleotide-functionalized microsphere probes. The microsphere terminal end is a well covered with single stranded DNA probes. After hybridization with a fluorescently labeled target solution, the complementary fluorescently labeled target is hybridized to the surface of the well, making the microsphere surface covered with the fluorescent target. The optical fiber transmits this fluorescent response and the intensity of the signal can be measured [62]. Since the oligonucleotide probe sequence of each microsphere can be specified, a sequence complementary to the DNA binding domain of HSF will create a link to join HSF and the

microsphere array. If the specific complement targets of these DNA microarrays are the exact DNA binding domains of HSF, a heat to light transducer can be fabricated. When activated by heat, or other stresses, the fluorescent-tagged HSF will trimerize and bind to the DNA microarrays, transmitting the fluorescent signal by means of fiber optics. Hence, we can sense specific heat and oxidative stresses manually.

### 6.3 DNA based Rotary Actuator

A rotary actuator based on double-cross-over DNA molecules was conceived by our team in this project. Our concept is based on the work presented in [39]. The authors in that work developed the device for use in DNA-based computations. In this work, we want to use it as a component of complex molecular machines for space applications. In such machines of the future, one of the technical challenges to address will be the control of individual machine components independently of each other. The DNA based rotary actuator proposed here is based on, and is controlled by DNA hybridization, which will not affect components other than itself, allowing the actuator to be controlled independent of the other devices. This device works in a four-step cycle and is fuelled by induction of interconversion between two topological motifs – the paranemic crossover PX DNA and its topoisomer $JX_2$ DNA. PX DNA, shown in Figure 17, is a four-stranded molecule in which two parallel double helices are joined by reciprocal exchange (crossing over) of strands at every point where the strands come together [64]. $JX_2$ DNA is a topoisomer of PX DNA that contains two adjacent sites where backbones juxtapose without crossing over. The color coding of the strands and labels in Figure 17 indicates that the top ends, A and B, are the same in both molecules, but the bottom ends, C and D, are rotated 180°. This rotation is the basis for the operation of the device. Strand replacement [40] is used to interconvert the PX and $JX_2$ motifs.

One strand of each of the blue and red strand pairs is broken into three strands. The principles of operation are illustrated in Figure 17b, where the red and blue strands of opposite polarity are shown connected by hairpin loops. Thus, the PX molecule shown there consists of one red strand, one blue strand, and two green strands, termed the 'set' strands, because they set the state of the device to be in the PX conformation; similarly, the $JX_2$ molecule has purple set strands. The set strand associated with the red strand has a 5' single-stranded extension, and the set strand associated with the blue strand has a similar 3' extension.

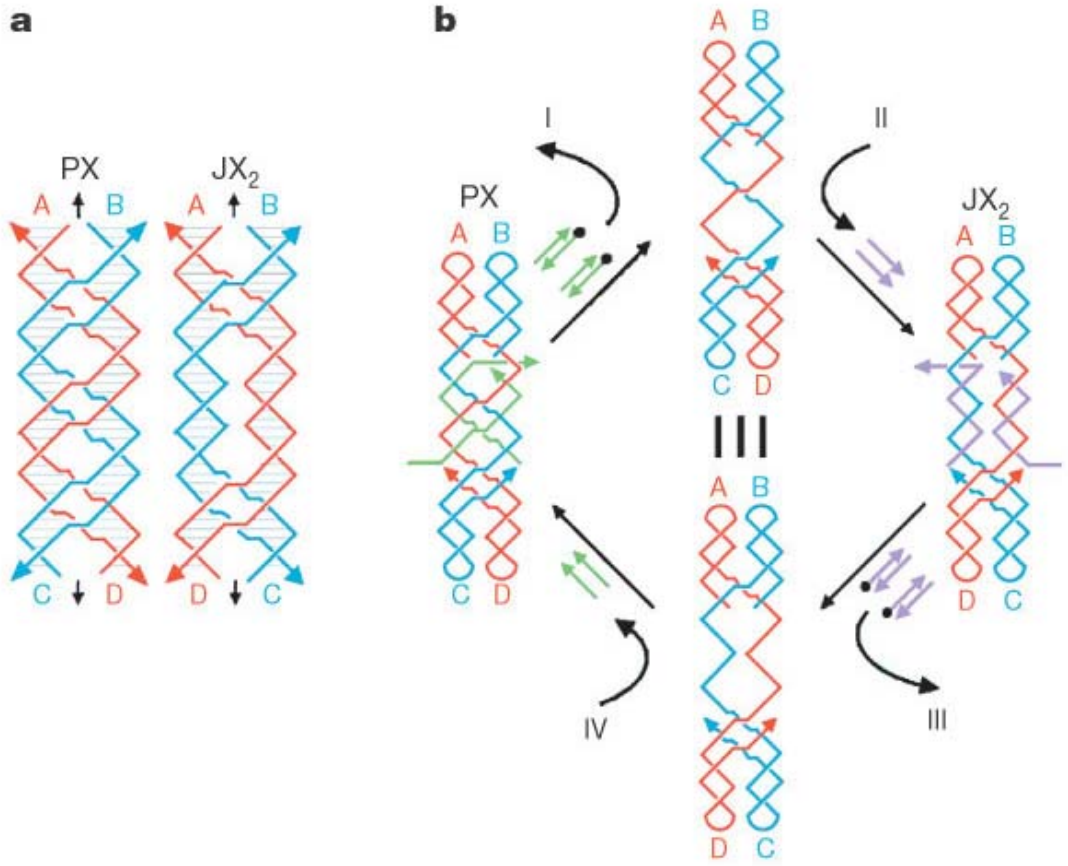


**FIGURE 17:** (a) PX and $JX_2$ motifs; (b) Inter-conversion between the two motifs with the rotation of lower end [39].

Extensions like these can be used to initiate branch migration that leads to removal of the strand from the branched motif, because it is paired with a complementary strand along its entire length. Thus, a complement to the entire length of the set strand (termed a 'fuel' strand) will pair with it in preference to the partially paired set strand in the PX (or $JX_2$) motif. Process I, shown in Figure 17b shows the addition of fuel strand complements to the two green set strands of the PX device, producing the unstructured intermediate at the top of the drawing. Process II shows the addition of pale-purple set strands that convert the intermediate to the $JX_2$ conformation. Process III shows the addition of fuel strands that convert the $JX_2$ molecule to the unstructured intermediate, and process IV shows the addition of the green set strands to produce the PX conformation again. Alternation between a paired structure and a partially unpaired structure analogous to this intermediate characterized the action of the system in [40]. In the device described here, the four-step cycle leads to two robust end points, the PX state and the $JX_2$ state.

We plan to utilize this motion to rotate, for example, a simple disc which could be a part of a membrane or a protein structure, in application. This is shown in Figure 18.

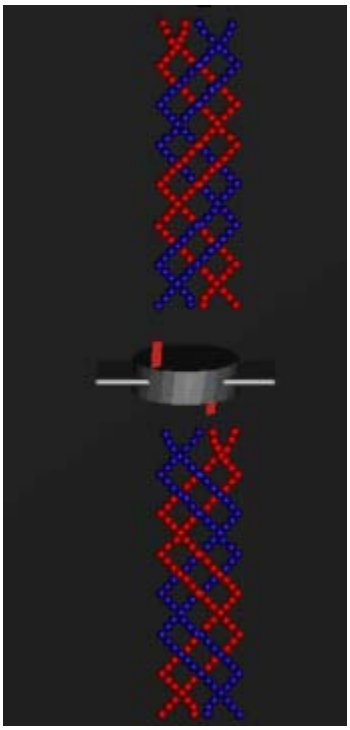

**FIGURE 18:** Two sets of DNA double-crossover molecules can rotate an object by $180^0$ each, producing a combined rotation of $360^0$. The linear motion can be achieved by B-Z transition of DNA.

The red tags on the disc in Figure 18 are pushed in a rotary motion alternately by the upper and lower motifs, each spanning half a circle. The motifs have to undergo a linear motion in order to touch the red tags and then translate back after their rotation. For this simple motion we suggest the use of B-Z transition of DNA strands [38], which produces a small linear extension in the device length. The DNA undergoing B-Z transition can be easily linked to these motifs and depending upon the selective fuelling, the linear and rotary motions will work in tandem.

### 6.4 Complex Devices

The proposed bio-nano machine-elements will be coupled with other elements and form multi-degree-of-freedom bio-nano-devices. These bio-nano-devices will be used as positioning instruments for the HSF bio-nano-sensors. The idea is to use the VPL motors as the actuators of such nano-devices where the structural elements are carbon nanotubes while the joints are formed by appropriately designed DNA elements. An example of such a bio-nano-machine is

shown in Figure 19. This is a 3-degree of freedom, 3-legged parallel manipulator. The top and bottom platforms are made from carbon nanotubes. They are connected to each other using 3 legs made from the VPL motors. The interface of the VPL motors to each one of the carbon-nanotube platforms is made using DNA based universal joints. Due to the motions of the 3 VPL motors the top platform presents a 3-degree of freedom motion. In Phase II we will study the interface of the VPL motors to both the DNA joints and carbon-nanotube structural elements.

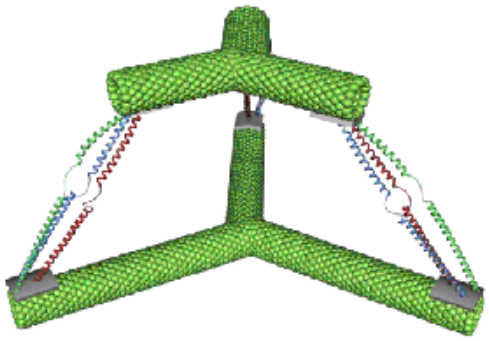

**FIGURE 19:** Three degree of freedom parallel platforms can be formed using controllable VPL actuators attached at the three legs of the platform.

DNA is a versatile nanoassembly material from which a variety of nanostructures and nanodevices have been constructed. Oligomers shorter than 100 nucleotides are easily synthesized with end modifications, such as thiol or amine groups, that aid in the construction of DNA protein conjugates or conjugates with other molecules. Single stranded DNA has a persistence length of about 1 nm. It is thus a flexible molecule suitable for linking members that must be joined in a hinge like manner. Double stranded DNA has a persistence length of 50 nm and is thus suitable for linking members for which greater flexural rigidity is desired. Thiol or amine modified DNA strands allow the linkage of DNA with VPL motors or other proteins using standard linkage chemistries [65] such as those we intend to use for the attachment of dyes to HSF. Carbon nanotubes can be prepared with carboxyl groups at the ends [66]. These can serve as reactive groups allowing the attachment of DNA to the nanotubes. As an alternative to covalent bonding for attachment to proteins, DNA with suitable base sequences will bind specifically to proteins through supramolecular interactions. DNA strands with such base sequences (aptimers) can be obtained through a suitably constructed selection process on a pool of random DNA strands [67]. Similarly, peptide sequences that exhibit specific binding to substrates can be obtained by selection experiments using phage display [68]. In this manner it

should be possible to find peptide sequences that specifically bind to the ends of carbon nanotubes. Linking DNA between such peptides could then form DNA joints. The binding specificity that can be achieved by DNA and protein aptimers would facilitate the construction of complex nanomachines through self-assembly. Each part can be given a unique recognition group allowing the structure to self-assemble only one way. We will explore the suitability of these approaches for constructing DNA based hinges.

Purification, de-aggregation and solubilization of single wall carbon nanotubes is the key to the fabrication of advanced nano-sized structures and scaffolds. Our team has demonstrated considerable expertise in this direction, by not only purifying [69] and solubilizing SWNTs, [70] but also length-fractionating these nano-rods in sizes from 10-400 nm. [70]. In addition, our ability to assemble them as nano-forest arrays, [71] enables us to functionalize each of their ends (composed of terminal carboxylic groups) with a variety of chemical functionalities where complex structures and assemblies such as those of Figure 19 can be realized.

So far we have developed preliminary concepts for DNA joints. Several researchers believe that the sticky ends of DNA oligonucleotides can join to the carboxyl groups at the end of acid-soaked nanotubes. We propose two types of DNA joints shown in Figure 20: double stranded (Type 1), and double stranded with a single strand gap (Type 2).

To create a joint of Type 1, no gap, we need to synthesize two oligonucleotides with complementary sequences. However, one strand should have additional bases on both ends. These strands will fold into a long double strand with "sticky ends". The double strand part should be long enough to provide a useful amount of angular flexibility, about 1-2 persistence lengths (P). Assuming the sticky ends bond two nanotubes, the joint should act roughly as a 2-DOF universal joint, providing freedom in two spatial components but little or no rotation. Double strand DNA keeps its traditional form for longitudinal forces less than about 50 pN. This type of joint will be stronger than the type with the gap.

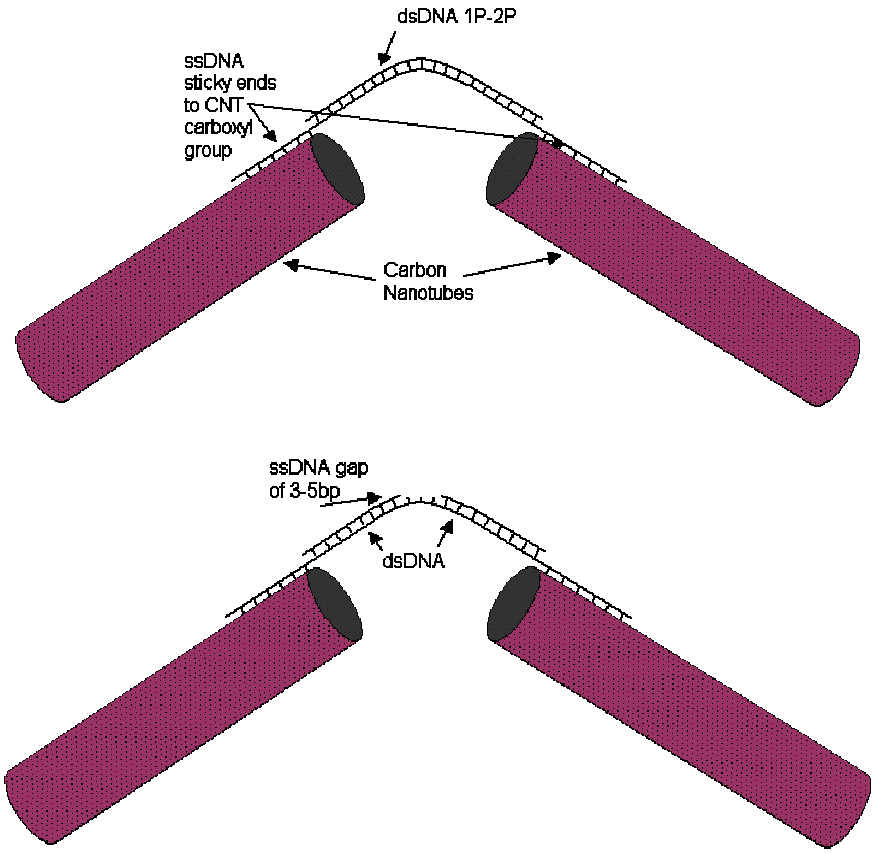


**FIGURE 20:** Type 1 (Top) and Type 2 (Bottom) DNA Joints.

Type 2, having a gap, can be smaller than Type 1. Type 1 requires 1-2P for angular freedom, but in Type 2 the gap provides the freedom of motion. To create these joints, begin with one oligonucleotide and add to it two short strands with complimentary sequences. Leave sticky ends and a gap of 3-5 bp (base pairs) in the middle. The sticky ends will connect to the carbon nanotubes and the joint will act roughly as a universal joint. This type may provide rotational motion as well since the gap has only single strand DNA. Therefore it may behave more like a 3-DOF joint. Also, the ssDNA (single strandes DNA) gap causes a weakness in the joint. SsDNA can only withstand longitudinal forces on the order of 1pN before changing in length. Therefore, a force analysis should be considered before choosing the type of DNA joint.

## 7.0 COMPUTATIONAL STUDIES

In order to predict the performance of the proposed bio-nano-machines, and more specifically of the VPL motors, mathematical models and 3D graphic animations are currently being developed. As it has been stated in Section 6.1, we are focusing our studies on the Influenza Haemagglutinin

(HA2) peptide. The aim is to build an appropriate computational model for the opening and closing of the HA2. In Phase I the goal was to identify the methods and software packages that will be used in the model development and obtain preliminary computational results. The modeling method is based on the calculation of the free energy that is released during the transition from native to fusogenic state using Molecular Dynamics (MD) Simulations.

### 7.1 Molecular Dynamics Simulations

The idea to study the behavior of VPL motors is based on the creation of a statistical description of the process by taking a large number of VPL actuators in an ensemble and then perform Molecular Dynamics Simulations. Software packages like CHARMM [72, 73] and AMBER [74] make it possible to model peptides and calculate energy interactions for their ensembles. In this work the software package CHARMM is used (see Section 7.2). The geometrical and structural data for the HA2 protein of Influenza virus are obtained from the Protein Data Bank (PDB) [75] at Rutgers University (http://rutgers.rcsb.org/) that provides models of specific protein chains that can be analyzed and manipulated as desired. These PDB files contain the precise molecular make up of the proteins, including the size, shape, angle between bonds, and a variety of other aspects. The file format provided by the Protein Data Bank can be exported, using a wide variety of available software, such as RasMol, to the common VRML format.

The structure file can be read by CHARMM to define the static structure of the protein. These all-atom models take into account very minute details of the molecular structure. A large number of intra and intermolecular forces have to be considered. In addition, with bond bending and torsion being taken into account, there are numerous degrees of freedom for the macromolecule. Such simulations have been done for extremely small peptides, and on a very small time scale (a few nanoseconds) [76, 77]. Also, the reported simulations have been done with an aim of understanding protein folding, which is not of interest in this project. Therefore, there are several modifications in the source code of CHARMM that need to be made to simulate efficiently the behavior of the VPL motors. These special considerations are described below:

a) Due to the large ranges of motions (10 nanometers) and large time ranges (in the order of microseconds) needed to study the behavior of VPL motors, the all-atom Molecular Dynamics (MD) methods become computationally very intensive. To simplify the computational burden and reduce computational time, we are making several

approximations that are based on the fact that our main interest in this work is the calculation of the free energy in the initial and final configurations and not in the intermediate states.

b) Due to the presence of the VPL motor in a solution, the protein has to be modeled in the presence of a solvent, which is usually taken to be water. The solvent can be considered explicitly by adding a number of water molecules interacting with the protein. However, this increases the computational work. Another way is to include the effect of the solvent implicitly by adding solvent correction terms to the CHARMM force field [78].

c) To further reduce computational intensity, the use of the Targeted Molecular Dynamics (TMD) toolbox, included in CHARMM will be used [79]. This toolbox is used for approximate modeling of processes spanning long time-scales, and is definitely within the time scales required by the VPL motors. In the method involving targeted molecular dynamics, the peptide is modeled and instead of letting it unfold at high temperature, it is 'forced' to achieve a desired final configuration by applying a time dependant, purely geometrical constraint to find the end structure [80]. This approach has successfully been applied to model a number of systems for time periods as large as 4 microseconds [81].

d) To reduce computational time, the Rutgers University parallel processing computing machines will be used. More specifically, The Center for Advanced Information Processing (CAIP) at Rutgers University has a Sun Microsystems Enterprise10000 computer which is, Sun's highest-end multi-processor server composed of 128 CPUs, 512 MB per processor, and approximately 3/4 TB of disk storage. This system will be used in our studies, to run the software package CHARMM.

## 7.2 Software Description: CHARMM 28

***Software Structure***

The structure and capabilities of CHARMM are shown in Figure 21.a while a number of support subroutines are shown in Figure 21.b. A brief description of these capabilities is given in this section.

The Input and Output to and from CHARMM can be: a) *Parameters* – which themselves can be of many types (commonly structural) and can be a part of the input as well as output. These can be Force Constants and Equilibrium Geometries useful for computing the energy. b) *Coordinates* – CHARMM can assign internal coordinates to the protein. It can also modify and convert from one coordinate system to another. c) *Data* Structures – I/O facilities are available for all the major data structures, which are described below. CHARMM can accept the commonly used *.pdb (from Protein Data Bank) structure format for an input structure. The structure file will contain the static structure of the protein molecule as obtained from the Protein Data Bank. d) *Results* – printing the results during the course of computation.

The Data Structures essentially contain all information about the static structure and inherent forces within the protein in vacuum. a) *Topology files* contain information about the structure and bonding of each monomer unit, which is used for constructing the system. The initial and/or final structure can be input into the program and the resulting structure is obtained as output topology and structure files. b) *Structure file* is the main repository of structural information. It is built using entries from the topology file. c) *H-bond list* – that contains information of all hydrogen bonds within the macromolecule, which is crucial for the dynamics calculations. d) *Nonbond list* – it contains a listing of all nonbonded interactions which affect the molecular dynamics, most prominent being the Van der Waal's interactions. e) *Symmetric Images* – the program can build a symmetric image of the protein in case the problem requires multiple proteins, which have similar structure and properties. Specifically, for the VPL motor, we have three similar alpha-helical structures out of which we can model just one (monomer) and we will try to use the symmetric images for building the trimer. This will save computational effort. This facility will also be useful when we will go on to use VPL actuators in series and parallel for displacement and force manipulation respectively.

Coordinate Manipulation (building, orienting and transforming) – a) *Internal Coordinate files* - As the protein is modeled, it is given internal coordinates which help to specify the position of each atom/molecule with respect to a datum. These coordinates can be manipulated and transformed from one coordinate system to another, if needed. b) *Build* – facilities to build coordinates using internal coordinate information. c) *Orient* – there are subroutines to orient molecules with respect to their principal axis and with respect to other molecules.

d) *Transformation* – these are general-purpose routines to translate, rotate and transform the non-orthogonal coordinate systems.

Mechanics and Energetics – a) *Minimization* - *e*nergy minimization is crucial for any molecular dynamics program to perform. It is done to make the molecules stable inside the solution, i.e. to relax and cool the structures. Energy minimization corresponds to a state of equilibrium. Initial velocities are assigned and the protein is allowed to open towards the predetermined target conformation. b) *Dynamics* – refers to the integration of Newton's equations of motion with Gear [82] or Verlet algorithms [83] and the option of internal coordinate constraints (SHAKE) [84]. c) *Normal Modes* – diagonalization of the second derivative matrix and analysis of normal modes.

Analysis Facility – Various analysis facilities are available, as shown in Figure 21.a. These are a) *Tables* – these are general-purpose data structures used to store and manipulate analysis information. b) *Static Properties* – many properties of the protein at a fixed point of time may be entered into a table. c) *Dynamic Properties* - averages, fluctuations and other dynamic properties from a MD simulation in tabular form. d) *Comparison* – structural and energetic comparison of structures. e) *Homology* – homologies in residue sequences of two related systems can be computed for comparison purposes. f) *Statistics* – applicable to all table data. g) *Correlation* – generation of auto and cross correlation for analysis of trajectories; f) *Spectral Analysis* – frequency analysis of dynamical properties; and, g) *Graphics Interfaces* – available interface routines and programs to display results on an Evans and Sutherland picture system and hard copy plots.

A number of Support Subroutines are necessary, shown in Figure 21.b. These are: a) *Energy Evaluation* – The CHARMM energy function will be described later. It is available as a subroutine called in many different parts of the program for tasks such as total energy evaluation, first and second derivative evaluation, and term-by-term analysis of the potential energy. b) *Dynamic Energy Evaluation* – DMA subroutines written in FORTRAN are used in many parts of the program to maximize the flexibility of the program and minimize the demands on the host system. c) *Free Format String Processing* – a free format string manipulation library used for input and command parsing and for adaptable formats for outputs; d) *Sorting, Searching, Linked*

*Lists, Array Math, I/O etc.* – a variety of subroutines exists for these tasks as well as miscellaneous functions.

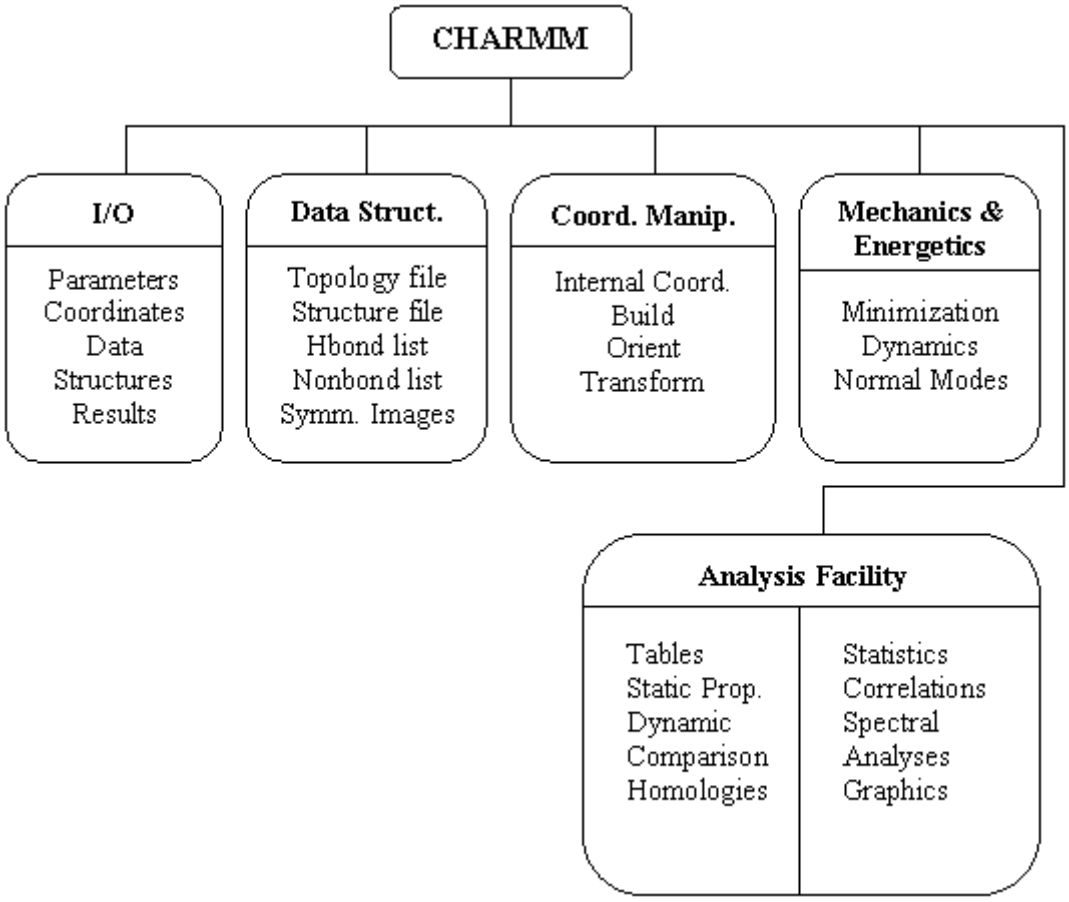


**FIGURE 21.a:** Structure and capabilities of CHARMM.

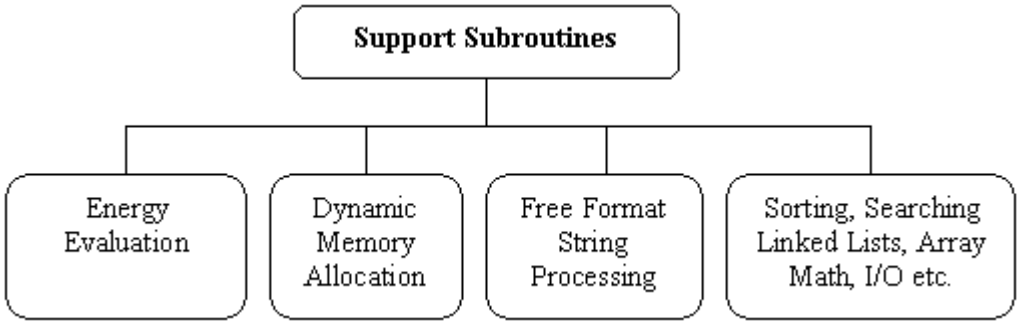


**FIGURE 21.b:** Support subroutines of CHARMM.

***Energy Calculation***

The CHARMM energy function E, which is the main computational aim of CHARMM, consists of a sum of the following energy terms such as the Bond Potential $E_b$, the Bond Angle Potential $E_\theta$, the Dihedral Angle Potential $E_\phi$, the Improper Torsion Potential $E_\omega$, the Van der Waals $E_{vdW}$, the Dielectric Constant $E_{el}$, the Hydrogen Bonds energy $E_{hb}$, and the Constraint Terms $E_{cr}$ and $E_{cf}$:

$$E = E_b + E_\theta + E_\phi + E_\omega + E_{vdW} + E_{el} + E_{hb} + E_{cr} + E_{cf} \quad (1)$$

The individual energy terms are defined in Equations (2)-(10).

The Bond Potential $E_b$ is equal to:

$$E_b = \Sigma\, k_b\, (r - r_0)^2 \quad (2)$$

where: $r_0$ is the initial position for a given atom; r is its instantaneous position; $K_b$ is a constant.

The Bond Angle Potential $E_\theta$ is equal to:

$$E_\theta = \Sigma\, k_\theta\, (\theta - \theta_0)^2 \quad (3)$$

where: $\theta_0$ is the initial angle for a particular given bond; $\theta$ is its instantaneous value; $K_\theta$ is a constant.

The Dihedral Angle Potential $E_\phi$ is equal to:

$$E_\phi = \Sigma\, |k_\phi| - k_\phi \cos(n\phi) \quad (4)$$

This torsion energy term is a four-atom term based on the dihedral angle $\phi$ about an axis defined by a pair of atoms. The energy constant, $k_\phi$, in this expression depends on the conformation of the molecule (cis or trans, pertaining to the amino acids). Integer n can take the values from 1 to 6.

The Improper Torsion Potential, $E_\omega$, is equal to:

$$E_\omega = \Sigma\, k_\omega(\omega - \omega_0)^2 \quad (5)$$

This improper torsion term maintains chirality about a tetrahedral extended heavy atom and planarity about certain polar atoms. The force constants $k_\omega$ and geometric constant $\omega_0$ are selected from the parameter table read into CHARMM.

The Van der Waals $E_{vdW}$ is equal to

$$E_{vdW} = \sum_{excl(i,j)=1} \left( \frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) sw\left(r_{ij}^2 , r_{on}^2 , r_{off}^2\right) \tag{6}$$

where: $r_{ij}$ is the distance between the center of gravity from one atom ($i^{th}$) group to another ($j^{th}$); index excl(i,j) is equal to 0 if the atoms are connected by angles or bonds and equal to 1 otherwise (unless explicitly specified to be zero); $r_{off}$ and $r_{on}$ are two extreme values of the interatomic distance; 'sw' is a switching function that is defined as :

$$sw(x , x_{on}, x_{off}) = 1 \quad \text{when} \quad x \le x_{on} \tag{6a}$$

$$sw(x , x_{on}, x_{off}) = \frac{(x_{off} + 2x - 3x_{on})}{(x_{off} - x_{on})^3}, \text{ when } x_{on} < x \le x_{off} \tag{6b}$$

$$sw(x , x_{on}, x_{off}) = 0 , \text{ when } x > x_{off} \tag{6c}$$

The electrostatic function $E_{el}$ can have different forms depending on the problem at hand. One such form is a constant dielectric, where the symbols q define atomic charges:

$$E_{el} = \sum_{excl(i,j)=1} \frac{q_i q_j}{4\pi\varepsilon_0 r_{ij}} \tag{7}$$

The Hydrogen Bonds energy $E_{hb}$ is equal to:

$$\begin{aligned} E_{hb} = & \sum \left( \frac{A}{r_{AD}^{i}} - \frac{B}{r_{AD}^{j}} \right) \cos^m(\theta_{A\text{-}H\text{-}D}) \times \cos^n(\theta_{AA\text{-}A\text{-}H}) \times sw\left( r_{AD}^2, r_{h_{on}}^2, r_{h_{off}}^2 \right) \\ & \times sw\left[ \cos^2(\theta_{H\text{-}A\text{-}D}), \cos^2(\theta_{h_{on}}), \cos^2(\theta_{h_{off}}) \right] \end{aligned} \tag{8}$$

where: i and j are positive integers; m can take the values 0,2 or 4; n can take the values 0 or 2. Subscripts 'A', 'D', 'H' and 'AA' represent the acceptor, donor hydrogen and antecedent atoms respectively. The selection of hydrogen bonds is based on the A-D distance and the A-H-D bond angle. The total term is zeroed if $\theta_{A\text{-}H\text{-}D}$ is less than $90^0$ or if $\theta_{AA\text{-}A\text{-}H}$ is less than $90^0$ along with the value of n equal to 0. The exponents i and j are determined when the parameters are read and

the exponents n and m are determined in the topology file when donor and acceptor atoms are listed. The exponent n is determined by the type of acceptor while m is determined by the donor type. The constants A and B are determined experimentally and the subroutine associated with hydrogen bonding automatically takes those values.

In order to restrict the changes that occur in the structure, certain constraints have to be placed. These constraints are placed in energy. One option is to maintain rigidly the position of certain atoms and to delete energy terms involving only these atoms. Atom harmonic constraints are used to avoid large displacements of atoms when minimizing energy, while still allowing the structure to relax. The atom harmonic constraint term $E_{cr}$ is equal to:

$$E_{cr} = K_i( r_i - r_{io} )^2 \tag{9}$$

where: $K_i$ is a mass-weighted constant, available in the subroutine. Dihedral constraints are used to maintain certain local conformation when a series of different conformations need to be examined in order to make potential energy maps.

The dihedral constraint is equal to:

$$E_{c\varphi} = \sum K_i(\varphi_i - \varphi_{io})^2 \tag{10}$$

where: $\phi$ is the dihedral angle and $\phi_i$ and $\phi_{io}$ are its instantaneous and initial values respectively, while the Rigid Distance Constraint from SHAKE [84] is $\delta_{rij}=0$. The rigid distance constraint puts a limit to the bond length and avoids large bond extensions. The SHAKE subroutine is standard and is used in many MD applications.

Once the energy function E is calculated then the following Dynamics - Newtonian Equation can be formed:

$$\frac{\partial^2 x_i}{\partial t^2} = -\frac{\nabla E(x_i)}{m_i} \tag{11}$$

where: index ‘i’ pertains to each atom, x is the position coordinate, ‘m’ is the atomic mass and ‘*t*’ is the time.

Equation (11) will be solved for every atom (or every significant atom, as an approximation) by the program after the construction of the energy function E.

***Targeted Molecular Dynamics***

Targeted Molecular Dynamics (TMD) is a toolbox of CHARMM 28 that is used for approximate modeling of processes spanning long time-scales, and is definitely within the time scales required by the VPL motors. Because the distance to be traveled by the N-terminal of the viral protein is relatively very large (10 nm), we cannot let the protein unfold by itself. Instead of 'unfolding' we want it to undergo a large conformational change and 'open' up. To achieve this, the macromolecule will be 'forced' towards a final configuration 'F' from an initial configuration 'I' by applying constraints. The constraint is in the form of a bias in the force field. If we define the 3N position coordinates corresponding to N atoms in the molecule as $\mathbf{x} = (x_1,\ldots..x_{3N})^T$ where 3N are the Cartesian coordinates of the position vectors $\mathbf{r}_1$, $\mathbf{r}_2$, …$\mathbf{r}_N$ of each individual atom, then for each configuration x, its distance, ρ, to the target configuration 'F' is defines as:

$$\rho = |\mathbf{x}-\mathbf{x}_F| = [\Sigma(x_i-x_{Fi})^2]^{1/2} \tag{12}$$

The distance ρ is a purely geometric control parameter here, which will be used to force the macromolecule to undergo the desired transformation. The constraint applied for this is equal to:

$$f(\mathbf{x}) = |\mathbf{x}-\mathbf{x}_F|^2-\rho^2 = 0 \tag{13}$$

which results in an additional constraint force:

$$\mathbf{F}_C = \lambda\, df/dx = 2\lambda\, (\mathbf{x}-\mathbf{x}_F), \tag{14}$$

where λ is the Lagrange parameter.

The TMD algorithm steps are:

1) Set $\rho = \rho_0 = |\mathbf{x}_I-\mathbf{x}_F|$, where I is the initial and F is the final conformation.

2) Choose initial coordinates $x_i(0) = x_{Ii}$ and appropriate initial velocities.

3) Solve, numerically, the equations of motion with the additional constraint force $\mathbf{F}_C$.

4) After each time step Δt, diminish ρ by $\Delta\rho = (\rho_0 - \rho_f)\, \Delta t/t_S$, where $t_S$ is the total simulation time.

At the end of the simulation, the final distance $\rho_f$ is reached. In this way, a monotonous decrease of $\rho$ forces the system to find a pathway from $x_I$ to a final configuration $x_F$. The total time for the simulations will be 2 microseconds.

### 7.3 Computational Results

In Phase I, preliminary computational results have been obtained and are presented in this section. In line with our experimental work, the computational studies are initially focused on the group of 36 amino acids that forms the hinge region of the influenza virus protein Haemagglutinin (HA), joining the two alpha helical regions of each monomer [55, 56]. All calculations are carried out using CHARMM. The native "closed" conformation structure for this peptide is obtained from the crystal structure in the PDB (PDB identifier code: 2VIU). The "open" structure is generated arbitrarily by forcing the structure away from the native conformation with constrained high-temperature molecular dynamics. After a short equilibration, these two "closed" and "open" structures are then used as reference end-point states to study the transformation between the open and closed conformations. The transformation is enforced through a root mean square difference (RMSD) harmonic constraint in conjunction with molecular dynamics simulations. Both the forward (closed to open) and the reverse (open to closed) transformations are carried out. The RMSD between the two end-point structures is about 9Å, therefore the transformation is carried out in 91 intermediate steps or windows with a 0.1 Å RMSD spacing between each intermediate window. At each intermediate window, the structure is constrained to be at the required RMSD value away from the starting structure, it is minimized using 100 steps of Steepest Descent minimization, and then equilibrated with 0.5 picoseconds of Langevin dynamics using SHAKE [84] with a friction coefficient of 25 ps on the non-hydrogen atoms. The harmonic RMSD constraint is mass-weighted and has a force constant of 500 kcal/mol/ $Å^2$ applied only to the non-hydrogen atoms. The decision to calculate the RMSD only for non-hydrogen atoms is usually done by convention since the conformation of the protein is more directly dependent on the heavy atoms than on the hydrogen atoms. Spontaneous transformation between the two conformations using unconstrained molecular dynamics may occur in the microsecond timescale. In the present case, however, the transformation is accelerated by using the artificial RMSD constraint such that a conformation close to the final state is approached successfully. Figure 22 shows the RMSD values of the structures at the end

of each intermediate window from both the native closed (non-fusogenic) and the open (fusogenic) conformations.

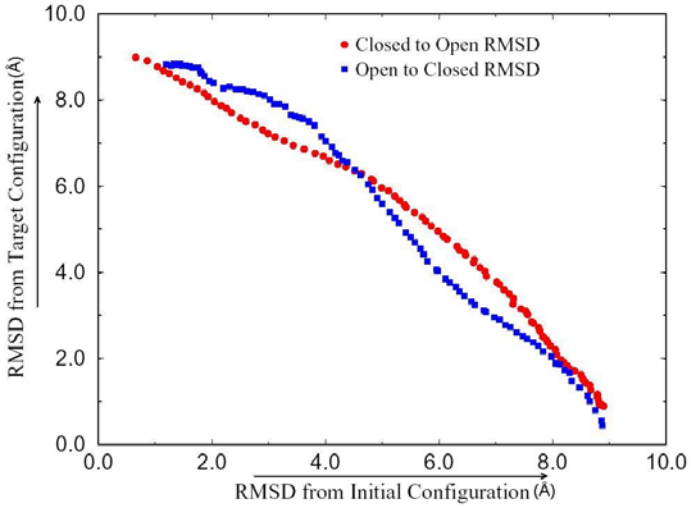


**FIGURE 22:** The non-hydrogen RMSD values of the structures at the end of each window in the constrained molecular dynamics transformation between the closed (native, non-fusogenic) and the open (fusogenic) states of the 36 amino acid peptide test system. The RMSD values (in Å) in comparison to the closed state are on the x-axis, those in comparison to the open state are on the y-axis. The two curves correspond to the forward (closed to open) and the reverse (open to closed) transformations, respectively.

The two curves shown correspond to the forward and reverse transformations, respectively, and the difference between the two curves is due to hysteresis. A small RMSD value on the x-axis indicates that the structure is close to the native conformation while a small RMSD value on the y-axis indicates that the structure is close to the closed conformation. In both curves, it can be seen that structures close to the end-point states (RMSD < 1 Å) are obtained. Increase in the amount of sampling for each window and decrease in the value of the force constant used in this transformation can lead to a progressively better description of the transformation.

### 7.4 Graphic Representations

In Phase II, 3D simulations will be performed and give us a visual feedback on the geometric and dynamic motion of the protein-based machine elements and their assemblies. The powerful and complex package of 3D Studio Max provides a wide range of capabilities and applications for molecular modeling. Accurate and proportionally correct models can be designed within the software console, as well as rendered to provide a rational look into the nano-world. Once the

proper representation of machine components have been designed and modeled, position script controllers within 3D Studio Max can be implemented to provide real-time animation and movement according to derived data from the simulation software. Preliminary animations of the studied peptides have been performed using 3D Studio Max. Using the PDB files, protein structures are directly inputted into 3D Studio Max using VRML file format, allowing structures to be rendered and modified to provide accurate graphical representations.

### 7.5 Design Analysis of VPL Motors

Parametric analysis of the design of the VPL motor will be performed in Phase II, once valid models and simulations is completed. The goal of the analysis will be to study the changes in the performance of VPL motor such as changes in forces, speeds and ranges of motion when certain structural, operational or environmental conditions are changed. Such design studies could be used in the future to optimize the performance of the VPL motors for certain tasks and also to create a database of all possible VPL motor types and sizes.

The structure of the VPL actuators can be varied to alter its performance characteristics to cater to the varying demands of the different composite machinery it is interfaced with. These demands could be operation under unfavorable environmental conditions, need for increased/reduced speed of actuation, or force and displacement multiplication. One way to alter the structure of VPL motor protein is to change the sequence of amino acids. A different sequence results in different alpha helical content, different hydrogen bonding and different energetics. Since the secondary structure is different, steric constraints also dictate the speed, force and displacement offered by the actuator. Addition of molecules like cysteines could dramatically change the energetics due to the additions of disulfide bonds to the structure. Disulfide bonds, formed by oxidation of cysteine residues, can cross-link between the chains of the hinge and can manipulate the protein structure.

Temperature plays a crucial role in the performance of any protein-based system. Proteins tend to denature at heightened temperatures. Maintenance of the operating temperature range is crucial in order to extract optimum performance from a system of protein actuators. The solvent molecules interact with the protein molecules and atoms and greatly affect the dynamics of the protein. The performance of a VPL actuator will, therefore, depend on the type of solvent the motor is operating in. The variation in the ionic concentration of a given solvent will similarly

affect the performance as the protein is formed by different charged, uncharged and polar or non-polar amino acids which will interact with the ions in the solution. We will study the variation of the performance of the VPL motor when the temperature is changing.

Several parameters which are uncontrollable and often act as disturbances during the operation of VPL motors such as presence of impurities, dust particles in the environment will affect the operation and performance of the VPL actuators. We will simulate the presence of such parameters and study their effect on the performance of the VPL motors.

## 8.0 EXPERIMENTAL STUDIES

In Phase I of this project we focused on the demonstration of the VPL motor concepts and performed a series of experiments. The goals of the project's experimental tasks in Phase I included protein expression and purification, and protein conformation characterization as a function of pH. Experiments need to be continued in the proposed Phase II and a detailed plan has been outlined.

### 8.1 Materials and Methods

We constructed a synthetic gene sequence (Figure 23) corresponding to the loop region between the short and long α-helices of the influenza haemagglutinin HA2 peptide described in Section 6.1 and shown in Figure 14. The HA2 subunit is one of the VPL motors proposed types. Sedimentation equilibrium and circular dichroism (CD) experiments have shown that this loop section of the peptide folds into a trimeric α-helical coiled-coil at the pH of membrane fusion, but it is monomeric at neutral pH [56]. The nucleotide sequence (Figure 24) for the protein was chosen to minimize occurrence of rare codons, or those codons not highly expressed in *E. coli* cells.

**R V I E K T N E K F H Q I E K E F S E V E G R I Q D L E K Y V E D T K**

**FIGURE 23:** Amino acid sequence for protein studied.

**AAC CGTGTTATCGAAAAAACCAACGAAAAATTCCACCAGATCGAA**
**AAAGAATTCTCTGAAGTTGAAGGTCGTATCCAGGACCTGGAAAAA**
**TACGTTGAAGACACCAAAATC TAATCC**

**FIGURE 24:** The base sequence for the synthetic protein, including the TAAT "stop" codon and the cohesive ends to anneal to the complementary plasmid overhangs.

We inserted the DNA into the purified genome of a self-replicating genetic element – a plasmid. The DNA is joined to the chromosome of the plasmid in a test tube, and the new recombinant DNA is then introduced into an *E. coli* bacterial cell. The normal replication mechanisms can produce millions of identical cells containing the recombinant plasmid construct within a day.

The plasmids generally used are small circular molecules of double-stranded DNA derived from larger, naturally occurring plasmids in bacterial cells. Upon centrifugation, the larger chromosomal DNA molecules precipitate into pellet, allowing the smaller plasmids to be easily separated and purified. The purified plasmids are then digested, or cut, with a restriction endonuclease, creating a linear DNA molecule. The target DNA is digested with the same nuclease, yielding restriction fragments with cohesive ends. The target DNA can then be added to the linearized plasmid DNA and their cohesive ends annealed to form circular recombinant plasmid DNA.

For our purposes, the plasmid was double digested, that is it was cut by two endonucleases. Nucleases SapI and XhoI were used for the digestion of the plasmid. Rather that cutting the target sequence with these enzymes to yield the cohesive ends, the cohesive ends were included in the construction of the target (Figure 24). The ends were allowed to anneal to generate the circular recombinant plasmid.

The recombinant plasmids were introduced, or transformed, into bacterial cells made competent – that is, transiently permeable to DNA. The cells became transfected, meaning they took up the recombinant plasmid. These cells grew and divided, creating multiple copies of cells containing the recombinant DNA. In order to select the cells that were successfully transfected, we grew the cells in the presence of the antibiotic ampicillin. As the plasmids carried genes for ampicillin resistance, only those cells containing the recombinant DNA survived.

Before purifying the proteins to begin the characterization portion of the work, a sample of the recombinant plasmid was purified and sent for sequencing. The results showed that the base sequence for the amino acids was missing one base pair. In other words, the entire reading frame for the protein was out of frame by one base pair (Figure 25).

AAC[ ]GTGTTATCGAAAAAACCAACGAAA…
[ ]CACAATAGCTTTTTTGGTTGCTTT…

**FIGURE 25:** Base sequence of double-stranded target protein, showing the location of the missing base pair at the beginning of the sequence.

PCR mutagenesis was used to insert the missing base pair. In this procedure, four oligonucleotide primers were used to PCR-amplify two segments of the gene. These segments were then fused together. The mutation, or correction, was contained in the sequences of the two inside mutagenizing primers, while the two outside primers did not contain the mutation. The primers were designed so that their sequences contained the desired correction to be incorporated into the plasmid construct and so their annealing temperatures were the same.

The PCR procedure was optimized for the amplification of the two segments of the gene construct. After the two PCR products were obtained, gel electrophoresis was utilized for verification. These products were then extracted from the gel and used as templates in the PCR-amplification to obtain the final product having the correct reading frame for the target protein. The two outside primers were utilized as well. This product was also verified by gel electrophoresis.

The plasmid was again double digested, but with nucleases SacII and XhoI, to linearize the DNA. The target insert was also digested with these enzymes to generate the cohesive ends, which were allowed to anneal to yield the recombinant plasmid. The transformation process was then repeated with the corrected target base sequence.

Current work focuses on the purification and characterization of the single synthetic protein as an actuator. Before the protein is purified, the recombinant plasmid construct will be extracted from the cells, and a sample will be sent for sequencing for verification. The bacterial cells will be harvested, lysed, and then clarified by centrifugation. Protein purification will be done by a novel system in which an internal protein segment, a self-cleavable protein-splicing element, is used to separate the target protein from an affinity tag. Protein splicing is a post-translational event in which the internal protein segment, or intein, catalyzes its own excision from the precursor protein and ligates the two flanking regions, exteins, together [85]. In our system, the intein undergoes specific self-cleavage, in the presence of thiols such as DTT or cysteine, and releases the target protein from the chitin-bound affinity tag. In this way, the target is purified in a single step (Figure 26).

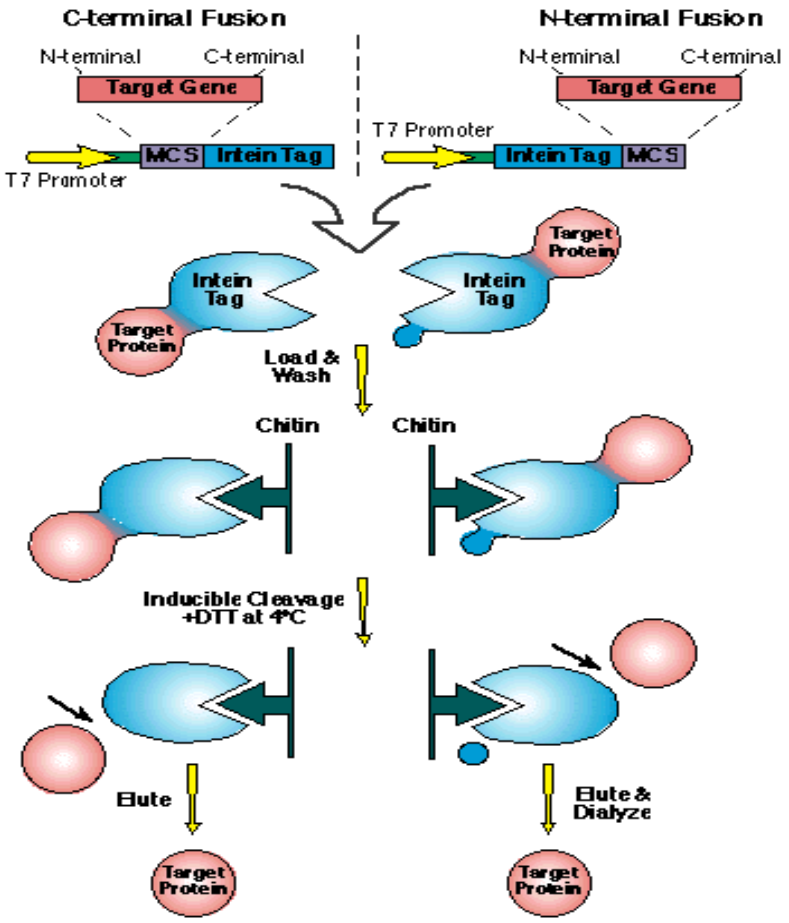


**FIGURE 26**: Schematic of the IMPACT™-CN protein purification system. The protein is fused to a tag consisting of the intein and its chitin-binding domain. The intein undergoes specific self-cleavage and releases the target from the intein tag, allowing affinity purification of the protein [86].

## 8.2 Future Experiments

### *Peptide Characterization*

The ability of the VPL Motor proteins to change conformation upon pH change will be measured using several different methods. CD spectroscopy will be used to monitor the secondary structure of the purified peptides. Gel filtration chromatography will be used to determine if the peptides form trimers.

Once the initial characterization is complete, fluorescence resonance energy transfer (FRET) will be used to study in real-time the kinetics of the motor protein. In order to use FRET, the N-terminus of the peptide will be labeled with tetramethylrhodamine by conjugation with a tetramethylrhodamine succimidyl ester. The C-terminus of the peptide will be labeled with

fluorescein using a fluoresceinyl glycine amide and the coupling agent 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDAC). Because the absorption spectrum of tetramethylrhodamine overlaps with the fluorescence emission spectrum of fluorescein, close proximity of the N- and C-terminus will cause FRET. Under such conditions, excitation of the fluorescently labeled peptides with blue light (absorbed by fluorescein) is expected to yield red fluorescence (emitted by tetramethylrhodamine). When the peptide is in the fully extended α-helix conformation, the two fluorescent labels will be too far apart for FRET to occur, and the blue excitation will yield green fluorescence from fluorescein only.

One potential problem with the FRET technique is the fluorescent labels are fairly large and may influence the physicochemical properties of the peptide. Therefore, other methods for performing FRET will be investigated. Recently, a synthetic fluorescent amino acid called aladan has been reported [87]. This novel fluorophore, along with the naturally fluorescent amino acid tryptophan may be better suited for the monitoring of the conformational change of the peptide.

The secondary tertiary structures and kinetics of folding will also be studied using nuclear magnetic resonance (NMR) spectroscopy. *E. coli* cells harboring the plasmid DNA that encodes for the fusion protein will be grown in minimal media with an $^{15}N$ labeled nitrogen source. As the cells express the fusion protein, all Nitrogen atoms in the peptide backbone will be labeled with the isotope, and this atom can be used to obtain an NMR signal. This can then be used to determine the 3-dimensional solution structure of the peptide at any pH. In addition, it will also be possible with NMR to actually observe the conformational change of the peptide, which is not possible with any other spectroscopy method.

***Protein Engineering***

In order to make the LOOP-36 peptide useful as a VPL motor, several features must be engineered into the peptide. According to previous work with the LOOP-36 peptide, the transition from α-helix to random coil required low temperatures, which suggests that the α-helix is not very stable. When additional α-helix forming amino acids were added to the end of the peptide, it no longer exhibited the pH-dependent conformational change. Therefore, we will need to stabilize the peptide without losing its hinge-like properties.

Our first approach to this problem will be to add elements to the ends of the peptide that will stabilize it without introducing more α-helical content. For example, we have already ordered oligonucleotide primers so that we can use PCR to add codons for cysteines to the ends of the LOOP-36 peptide. The cysteines will be able to form disulfide bonds that will hold the peptide in a folded conformation. Upon addition of a reducing agent and a pH change, the disulfide bond will break and allow the peptide to straighten into its linear conformation (Figure 27). We will also be able to use these flanking cysteines to attach other functional groups to the ends of the peptide.

Our second approach to the improvement of the LOOP-36 peptide will be to mutagenize the loop by both a directed and random approach to find changes that can improve the properties of the peptide. Our first step will be to perform an "alanine-scan" of the peptide [88]. This will involve replacing each amino acid in the peptide with an alanine, and then assaying the effect of this mutation using CD spectroscopy. This will allow us to determine which amino acids have the greatest effect on the pH-dependent conformational change. From this data we can then make rationally chosen mutations at the "hot-spots" in order to gain further improvements.

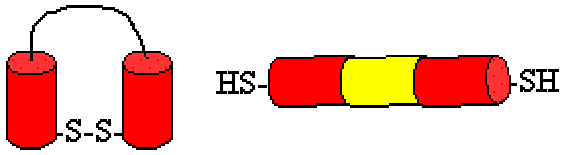


**FIGURE 27:** Predicted conformational change of LOOP-36 with cysteins attached to either end of the peptide.

We will also try a random mutagenesis approach. Oligonucleotide primers will be purchased so that chosen amino acids of the LOOP-36 peptide will be randomized. This can be used to create a library of different possible peptides. The expression and purification system that we have chosen is very well suited to this approach. A selection of the mutants from the peptide library will be expressed and screened using CD spectroscopy until peptides are found that exhibit superior properties. The amino acid sequence of the peptides can then be determined by sequential digestion.

## 9.0 INTERFACE OF VPL MOTORS WITH OTHER ELEMENTS

In Phase I, we developed concepts and made plans for experimental demonstration of the interface of the VPL motors with DNA joints and carbon-nanotube structural elements. In this section we present the concepts that we developed and the experimental plan to be performed in the proposed Phase II.

***DNA Joints***

DNA joints consisting of single-stranded DNA, 4nm long, have been employed in several DNA based motors devised by Yurke [40, 89, 90]. The length was chosen to be of order the persistence length of single-stranded DNA. The effect of hinge length on hinge performance has not been systematically investigated. We propose to carry out a series of experiments employing DNA-based nanostructures similar to the DNA-based motors devised by Yurke to evaluate hinge performance as a function of length. The hinges that Yurke has employed were single-stranded and, hence, the motion of these hinges is not well constrained mechanically. We propose to construct hinges that employ double or multiple crossover structures to limit the degrees of motion the hinges can have. The use of double or multiple crossover structures will also result in stiffer structures, allowing the structures to be larger than the 50 nm-persistence length of duplex DNA. Figure 28 shows a triple crossover hinge. The black bands represent single stranded DNA. The cylinders represent double stranded DNA. The cylinders are held together by virtue of the fact that a given strand of DNA from one duplex will cross over to another duplex. The construction of double and triple crossover structures is already a well-developed technology [40, 91, 92]. In the study of these hinges, DNA based motors will be used to flex the hinges. FRET and atomic force microscopy will be used to study their motion. We will also explore ways to bind joints to proteins and carbon nanotubes.

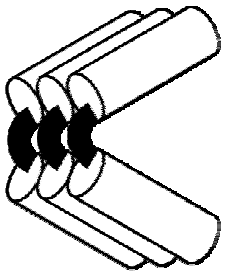

**FIGURE 28:** Triple crossover hinge made from DNA.

***Surfactant Organized System of VPL Motors***

Using the Langmuir-Blodget film technique, the viral protein monomers and (or) trimer will be organized with the help of surfactant bilayers on a variety of solid (*i.e.* mica, Si, etc) and liquid (i.e. Hg, $H_2O$, etc.) supports as shown in Figure 29. In this configuration, the viral protein monomers are stabilized by surfactant bilayers and can be organized side by side in trimers as they exist in their natural environment. This configuration is particularly amenable to Atomic Force Microscopy (AFM) characterization, where both VPL motor displacement and force parameters can be determined. Examining the actuation stroke in solid versus liquid interfaces (where positive pressure can be applied to the surfactant) we will establish the magnitude of cooperativity from the bilayer membrane to the VPL motor function.

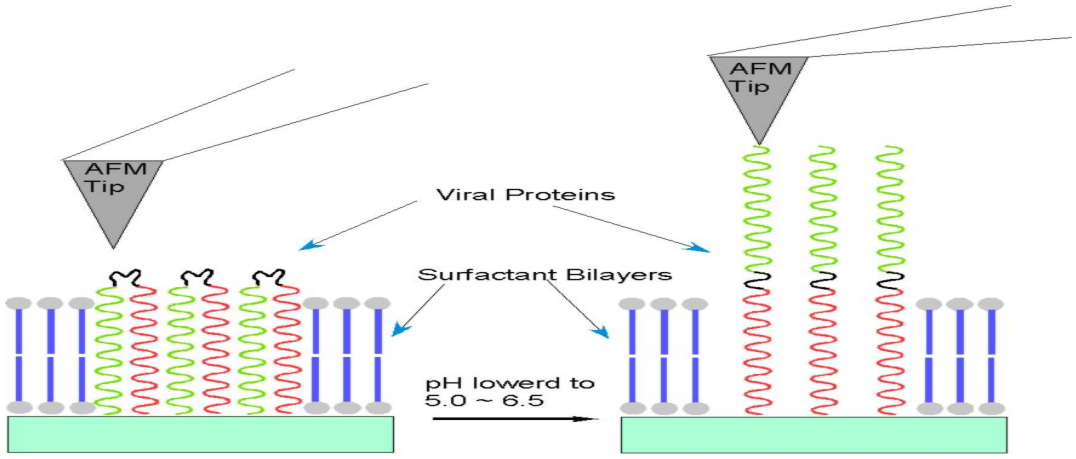


**FIGURE 29:** VP trimers organized with the help of surfactant bilayers on a variety of substrates.

***Nanotube Outfitted VPL Motors***

Assuming that adequate actuation is realized from individual VPL motors in the absence of surrounding surfactants (*i.e.* the VPL motor in its extended form is stiff enough to carry some load), then such linear motors can be equipped with rigid arms to extend actuation to greater lengths. This is schematically illustrated in Figure 30.

The fabrication of these nanotube-VPL motor assemblies relies on the non-covalent organization of rigid rod moieties that Papadimitrakopoulos' laboratory has pioneered for a variety of systems such as small chelates [93] and single wall carbon nanotubes [94]. Our demonstrated ability of fractionating SWNT by length [95] comes hand-in-hand with the capacity to non-covalently

assemble monolayers consisting of dense arrays of nanotubes oriented normal to the Fe(OH)x coated substrates (see Figure 31). Assuming that we can preferentially link through covalent functionalization either the C- or N-terminus of VPL motors or nearby pendant groups, we can envision building in a layer-by-layer fashion extended SWNT/VP/SWNT adducts, that can be released from their growing substrate by lowering the pH below 3.2 [96, 97] and solubilizing away the immobilizing $Fe(OH)_x$ layer (see Figure 32).

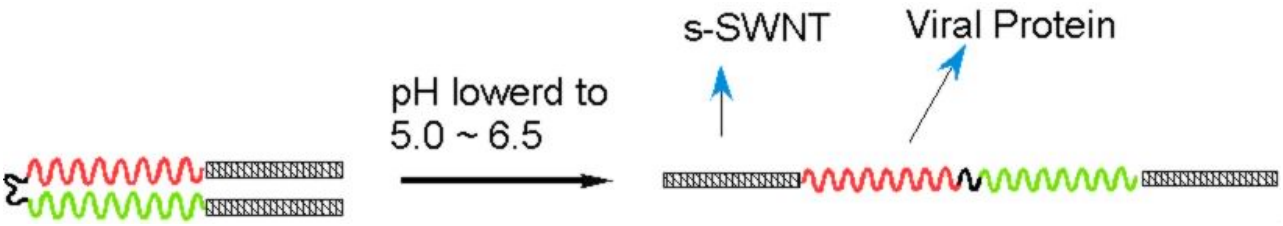


**FIGURE 30:** Single wall carbon nanotube (SWNT) outfitted VPL motors.

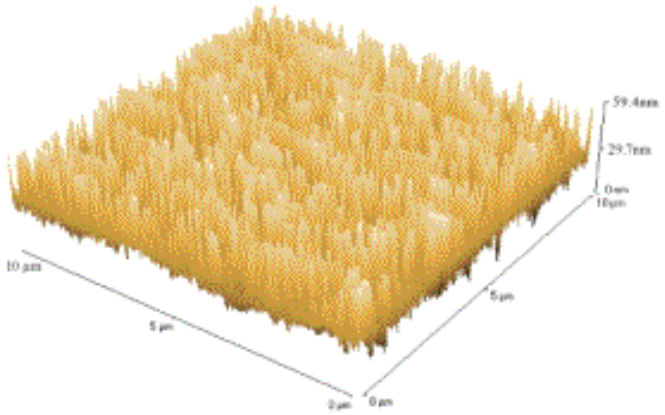

**FIGURE 31:** Typical non-covalent shortened SWNT forest assembly on a Si Substrate.

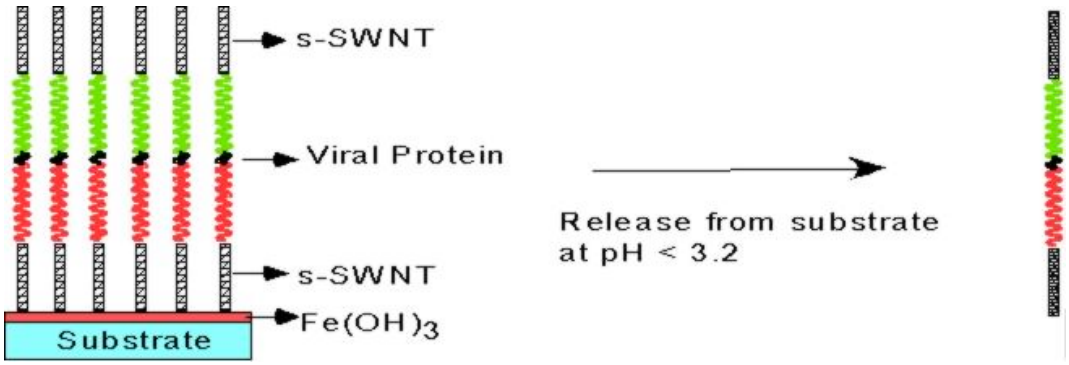


**FIGURE 32:** Layer-by-layer growth of SWNT/VP/SWNT adducts followed by their substrate release at pH below 3.2.

The action of such SWNT/VP/SWNT adducts can be monitored by various characterization methods such as dynamic light scattering and viscometry. These methods can readily allow us to investigate the effect of SWNT interaction in the collapsed state and up to what SWNT length and surface treatment of SWNTs can VP extend this adduct [98, 99].

***DNA / Nanotube Outfitted VPL Motors***

We will then proceed, using the similar concept and outfit the other two ends of the SWNT arms with different DNA oligomers (*i.e.* with sequences B and C). These DNA(B)/SWNT/VP/SWNT/DNA(C) adducts can be particularly useful in careful examination of the strength of complementary B'-X-C' linker to prevent extension of the VP through DNA slippage (see Figure 33). Here, the length of X can be varied as well as the temperature, ionic strength, mismatch density of the complementary B' and C' units, etc. to provide a detailed understanding of the strength of DNA linkage in light of its use for fabricating universal joints with molecular recognition.

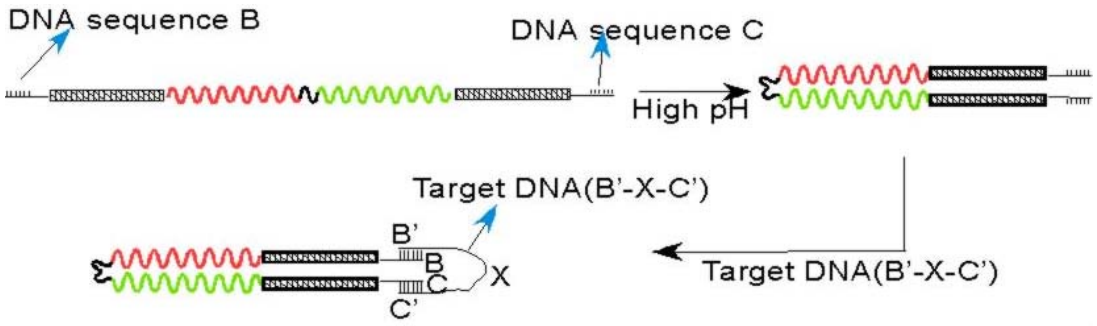


**FIGURE 33:** DNA(B)/SWNT/VP/SWNT/DNA(C) adducts and its coupling with the complementary B'-X-C' linker.

Finally, such DNA(B)/SWNT/VP/SWNT/DNA(C) adducts can be used to link DNA(B')-functionalized nano or submicron sized particles to DNA(C')-functionalized substrates to induce lifting of such particles from the surface at distances proportional to the SWNT/VP/SWNT length at its extended state (see Figure 34). Similar immobilization of DNA(B')-functionalized submicron polystyrene latex spheres to DNA(C')-functionalized substrates (see Figure 35) have been successfully demonstrated at the Papadimitrakopoulos' group as part of an AFOSR grant on DNA-assisted fabrication of opaline based photonic crystals.

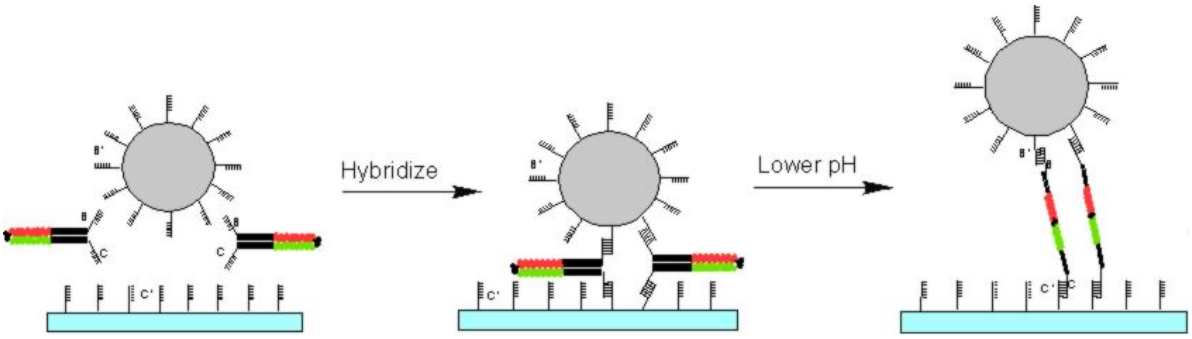


**FIGURE 34:** VP/SWNT/DNA/Latex-Sphere based linear motor (see text for details).

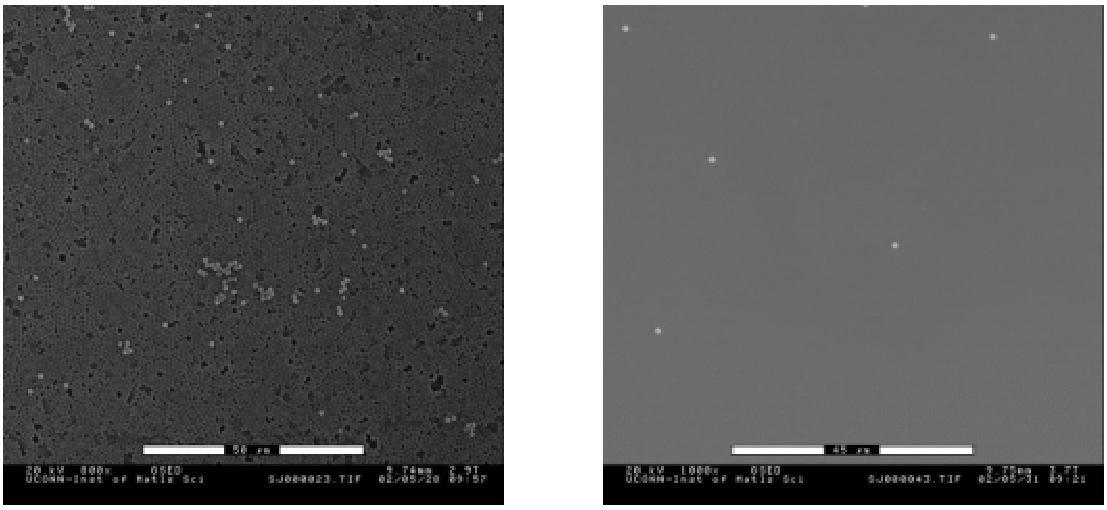

**FIGURE 35:** DNA immobilized monolayer of submicron spheres (A) below and (B) above the DNA melting point ($T_m$). At temperatures above $T_m$, nearly all the spheres can be washed away.

## 10.0 OUTREACH ACTIVITIES

The following outreach activities were performed during Phase I:

1. The Computational Bio-Nano-Robotics Laboratory was created by Prof. Mavroidis on July 1, 2002. Its goal is to perform high-level computational research to calculate the performance of novel bio-nano-robotic systems. The lab is fully equipped with state of the art computers and software and has access to CAIP's parallel processing capabilities.
2. An invention disclosure was submitted at Rutgers University and a provisional patent application was filed by Rutgers [100].

3. A webpage was created at Rutgers University in order to describe the state of the art on bio-nano-robotic systems, disseminate the research results of the proposed project and serve as a forum of discussions for this subject. Our project webpage can be accessed at http://bionano.rutgers.edu.

4. High school students participated in the research activities of this project.

5. Two presentations were given at the Industrial Advisory Board (IAB) meeting (September 2002) and the Annual Research Review (November 2002) of CAIP: Center for Advanced Information Processing.

## 11.0 REFERENCES


1. Boyer P.D., 1998, "Energy, Life and ATP", (Nobel Lecture) *Angewandte Chemie International Edition*, **37**, 2296-2307.

2. Block S.M., 1998, "Kinesin, What Gives?" *Cell*, **93**, 5–8.

3. Schnitzer M.J. and Block S.M., 1997, "Kinesin Hydrolyses One ATP per 8-nm Step", *Nature*, **388**, 386-390.

4. Wang M.D., Schnitzer M.J., Yin H., Landick R., Gelles J., Block S. M., 1998, "Force and Velocity Measured for Single Molecules of RNA Polymerase", *Science*, **282**, 902-907.

5. Kitamura K., Tokunaga M., Iwane A.H., Yanagida T., 1999, "A Single Myosin Head Moves Along an Actin Filament With Regular Steps of ~5.3nm", *Nature*, **397**, 129.

6. Montemagno C.D., and Bachand G.D., 1999, "Constructing Nanomechanical Devices Powered by Biomolecular Motors", *Nanotechnology*, **10**, 225-331.

7. Bachand G.D. and Montemagno C.D., 2000, "Constructing Organic / Inorganic NEMS Devices Powered by Biomolecular Motors", *Biomedical Microdevices*, **2**, 179-184.

8. Bachand G.D., Neves H.P., Olkhovets A.G., Craighead H.G., and Montemagno C.D., 2000, "Powering Inorganic Nanodevices With a Biomolecular Motor", *Science*, **290**, 1555-1558.

9. Noji H., Yasuda R., Yoshida M., and Kinosita K, Jr., 1997, "Direct Observation of the Rotation of F1-ATPase," *Nature*, **386**, 299-302.

10. Yasuda R., Noji H., Kinosita K. Jr., and Yoshida M., 1998, "F1-ATPase is a Highly Efficient Motor that Rotates with Discrete 120-Degree Steps", *Cell*, **93**, 1117-1124.

11. Walker J.E., 1998, "ATP Synthesis by Rotary Catalysis", (Nobel lecture), *Angewandte Chemie Intternational Edition*, **37**, 2308-2319.

12. Frasch W.D., 2000, "Vanadyl as a Probe of the Function of the F1-ATPase-Mg2+ Cofactor", *Journal of Bioenergetics and Biomembranes*, **32**, 539-546.

13. Farrell C.M., Mackey, A.M., Klumpp, L.M. and Gilbert, S.P., 2002, "Role of ATP Hydrolysis for Kinesin Processivity", *Journal of Biological Chemistry*, **277**, 17079-17087.

14. Block S.M., Goldstein L.S., Schnapp B.J., 1990, "Bead Movement by Single Kinesin Molecules Studied with Optical Tweezers", *Nature*, **348**(6299), 348-52.

15. Howard J., Hudspeth A.J., Vale R.D., 1989, "Movement of Microtubules by Single Kinesin Molecules", *Nature*, **342**(6246), 154-158.

16. Finer J.T., Simmons R.M., Spudich J.A., 1994, "Single Myosin Molecule Mechanics: Piconewton Forces and Nanometre Steps", *Nature*, **368**, 113-119.

17. Vale R.D. and Milligan R.A., 2000, "The Way Things Move: Looking Under the Hood of Molecular Motor Proteins", *Science*, **288**, 88-95.

18. Hackney D.D., 1996, "The Kinetic Cycles of Myosin, Kinesin, and Dynein", *Annual Review of Physiology* **58**, 732-750.

19. Böhm K.J., Steinmetzer P., Daniel A., Baum M., Vater W., and Unger E., 1997, "Kinesin Driven Microtubule Motility in the Presence of Alkaline-Earth Metal Ions: Indication for a Calcium Ion-Dependent Motility", *Cell Motility and the Cytoskeleton*, **37**, 226-231.

20. Mehta A.D., Rock R.S., Rief M., Spudich J.A., Mooseker M.S., and Cheney R.E., 1999, "Myosin-V is a Processive Actin-Based Motor", *Nature*, **400**, 590-593.

21. Berg H.C., "Motile Behavior of Bacteria", *Physics Today*, **53** (1), 24-29.

22. Berg H.C., 1974, "Dynamic Properties of Bacterial Flagellar Motors", *Nature*, **249**, 77-79.

23. Ueino T., Oosawa K., and Aizawa S-I., 1992, "M ring, S ring and Proximal Rod of the Flagellar Basal Body of Salmonella Typhimurium are Composed of Subunits of a Single Protein, FliF", *Journal of Molecular Biology*, **227**, 672-677.

24. Ueino T., Oosawa K., and Aizawa S-I., 1994, "Domain Structures of the MS Ring Component Protein (FliF) of the Flagellar Basal Body of Salmonella Typhimurium", *Journal of Molecular Biology* **236**, 546-555.

25. Khan S., Zhao R., and Reese T.S., 1998, "Architectural Features of the Salmonella Typhimurium Flagellar Motor Switch Revealed by Disrupted C-Rings", *Journal of Structural Biology*, **122**, 311-319.

26. Namba K, Vonderveczt F., 1997 "Molecular Structure of Bacterial Flagellum", *Quarterly Review of Biophysics*, **30**(1), 1-65.

27. Hess, Henry and Vogel, Viola, 2001, "Molecular Shuttles Based on Motor Proteins: Active Transport in Synthetic Environments", *Reviews in Molecular Biotechnology*, **82**, 67-85.

28. Harada Akira, 2001, "Cyclodextrin-Based Molecular Machines", *Accounts of Chemical Research* **34**, 456-464.

29. Schalley C.A., Beizai K. and Vogtle F.K., 2001, "On the Way to Rotaxane-Based Molecular Motors: Studies in Molecular Mobility and Topological Chirality", *Accounts of Chemical Research*, **34**, 465-476.

30. Balzani V., Lopez M.G. and Stoddart J.F., 1998, "Molecular Machines" *Accounts of Chemical Research*, **31**, 405-414.

31. Koumura N., Zijlstra R.W.J., van Delden, R.A., Harada N., Feringa B.N., 1999, "A Light-driven Monodirectional Molecular Rotor", *Nature*, **401**, 152-155.

32. Amendola V., Fabbrizzi L., Mangano C., and Pallavicini P., 2001, "Molecular Machines Based on Metal Ion Translocation", *Accounts of Chemical Research*, **34**, 488-493.

33. Mahadevan L. and Matsudaira P., 2000, "Motility Powered by Supramolecular Springs and Ratchets", *Science*, **288**, 95-99.

34. Seeman Nadrian C., 1998, "DNA Nanotechnology: Novel DNA Constructions", *Annual Review of Biophysics and Biomolecular Structure,* **27**, 225-48.

35. Tobias I., Swigon D., Coleman B.D., 2000, "Elastic Stability of DNA Configurations I - General Theory", *Physical Review E*, **61**(1), 747-58.

36. Yuqiu J., Juang C–B, Keller D., Bustamante C., Beach D., Houseal T., and Builes E., 1992, "Mechanical, Electrical, and Chemical Manipulation of Single DNA Molecules", *Nanotechnology*, **3**, 16-20.

37. Hu J., Zhang Y., Gao H., Li M., Hartmann U., 2002, "Artificial DNA Patterns by Mechanical Nanomanipulation", *Nanoletters*, **2**, No. 1, 55-57.

38. Mao C., Sun W., Shen Z., and Seeman N., 1999, "A Nanomechanical Device Based on the B-Z Transition of DNA", *Nature*, **397**, 144-146.

39. Yan H., Zhang X., Shen Z, and Seeman N., 2002, "A Robust DNA Mechanical Device Controlled by Hybridization Topology", *Nature*, **415**, 62-65.

40. Yurke B., Turberfield A.J., Mills A.P., Simmel F.C., and Neumann J.L., 2000, "A DNA-Fuelled Molecular Machine Made of DNA" *Nature*, **406**, 605-608.

41. http://www.media.mit.edu/nanoscale/

42. Kasianowicz J.J., Bayley H., NIST, http://www.cstl.nist.gov/biotech/biomat/Projects/metal.html

43. Plant A.L., and Silin V., NIST, http://www.cstl.nist.gov/biotech/biomat/Projects/warfare.html

44. Pieroni O., Fissi A., Angelini N., Lensi F., 2001, "Photoreceptive Polypeptides", *Accounts of Chemical Research*, **34**(1), 9-17.

45. Ferguson J.A., Boles T.C., Adams C.P., Walt D.R., 1996, "A Fiber-Optic DNA Biosensor Microarray for the Analysis of Gene Expression", *Nature Biotechnology*, **14**, 1681-1684.

46. Manning P., and McNeil C., ongoing research work, "Microfabricated Multi-Analyte Amperometric Sensors" http://nanocentre.ncl.ac.uk/

47. Wilson I.A., Skehel J.J., and Wiley D.C., 1981, "Structure of the Hemagglutinin Membrane Glycoprotein of Influenza Virus at 3 Å Resolution", *Nature*, **289**, 366-373.

48. Chan D.C., Fass D., Berger J.M. and Kim P.S., 1997, "Core Structure of gp41 from HIV Envelope Glycoprotein," *Cell*, **89**, 263-273.

49. Singh M., Berger B., Kim P., 1999, "Learncoil VMF: Computational Evidence for Coiled-Coil Like Motifs in Many Viral Membrane Fusion-Proteins," *Journal of Molecular Biology,* **290**, 1031-1041.

50. Caffrey M., Cai M., Kaufman J., Stahl S. J., Wingfield P.T., Covell D.G., Gronenborn, A. M., Clore G. M., 1998, "Three-Dimensional Structure of 44kD Ectodomain of SIV gp41," *EMBO Journal,* **17**, 4572-4584.

51. Kobe B., Center R.J., Kemp B.E., Poumbourios P., 1999, "Crystal Structure of Human T-Cell Leukemia Virus Type 1 gp21 Ectodomain Crystallized as a Malsote-Binding Protein Chimera Reveals Spectral Evolution of Retroviral Transmembrane Proteins," *Proceedings of the National Acdemy Sciences USA,* **96**, 4319-4324.

52. Baker K., Dutch R. E., Lamb R. A., Jardetzky T. S., 1999, "Structural Basis for Paramyxovirus-Mediated Membrane Fusion," *Molecular Cell,* **3**, 309-319.

53. Weissenhorn W., Calder L.J., Wharton S.A., Skehel J.J. and Wiley D.C., 1998, "The Central Structural Feature of Membrane Fusion Protein Subunit from the Ebola Virus Glycoprotein is a Long Triple Stranded Coiled Coil," *Proceedingsof the National Academy of Sciences USA,* **95**, 6032-6036.

54. Tan K., Liu J., Wang J., Shen S., Lu M., 1997, "Atomic Structure of a Thermostable Domain of HIV-1 gp41," *Proceedings of the National Academy of Sciences USA,* **94**,12303-12308.

55. Bullough P.A., Hughson F.M., Skehel J.J., and Wiley D.C., 1994, "Structure of Influenza Haemagglutinin at the pH of Membrane Fusion", *Nature*, **371**, 37-43.

56. Carr C. M. and Kim P., 1993, "A Spring-Loaded Mechanism for the Conformational Change of Influenza Hemagglutinin," *Cell*, **73**, 823-832.

57. Bentz J., 2000, "Membrane Fusion Mediated by Coiled Coils: A Hypothesis", *Biophysical Journal*, **78**, 886-900.

58. Lee et al, 2000, "The Yeast Heat Shock Transcription Factor Changes Conformation in Response to Superoxide and Temperature." *Molecular Biology of the Cell,* **11,**1752-1764.

59. Larson J.S., Schuetz T.J., Kingston R.E., 1995, "In Vitro Activation of Purified Human Heat Shock Factor by Heat," *Biochemistry*, **34**,1902-1911.

60. Mathew A., Mathur S.K., Jolly C., Fox S.G., Kim S. and Morimoto R.I., 2001, "Stress-Specific Activation and Repression of Heat Shock Factors 1 and 2," *Molecular and Cellular Biology*, **21**,7163-7171.

61. Jurivich et al, 1992, "Effects of Sodium Salicylate on the Human Heat Shock Response," *Science*, **255**,1243-1254.

62. Ferguson J.A., Steemers F.J., Walt D.R., 2000, "High-Density Fiber-Optic DNA Random Microsphere Array," *Analytical Chemistry*, **72**,5618-5624.

63. Morimoto R.I., 1998, " Regulation of Heat Shock Transcriptional Response: Cross Talk Between a Family of Heat Shock Factors, Molecular Chaperones, and Negative Regulators", Genes and Development, 12, 3788-3796.

64. Seeman N. C., 2001, "DNA Nicks and Nodes and Nanotechnology", *Nano Lett.* **1**, 22-26.

65. Hermanson, G. T., Bioconjugate Techniques*, Academic Press*, 1996.

66. Wong S.S., Joselevich E., Woolley A.T., Cheung C.L., Lieber C.M., 1998, "Covalently Functionalized Nanotubes as Nanometre-sized Probes in Chemistry and Biology", *Nature*, **394**(6688), 52-55.

67. Ellington A.D., and Szostak J.W., 1992, "Selection *in vitro* of Single-Stranded-DNA Molecules that Fold into Specific Ligand-binding Structures," *Nature*, **382**,818-822.

68. Whaley S.R., English D.S., Hu E.L., Barbara P.F., Belcher A.M., 2000, "Selection of Peptides with Semiconductor Binding Specificity for Directed Nanocrystal Assembly," *Nature*, **405**: 665-668.

69. Chattopadhyay D., Galeska I., Papadimitrakopoulos F., 2002, "Complete Elimination of Metal Catalyst from Single Wall Carbon Nanotubes", *Carbon*, **40**(7), 985-988.

70. Chattopadhyay D., Lastella S., Kim S., Papadimitrakopoulos F., 2002 "Length Separation of Zwitterion-Functionalized Single Wall Carbon Nanotubes by GPC", *Journal of American Chemical Society*, **124**(5), 728-729.

71. Chattopadhyay D., Galeska I., Papadimitrakopoulos F., 2001, "Metal-Assisted Organization of Shortened Carbon Nanotubes in Monolayer and Multilayer Forest Assemblies", *Journal of American Chemical Society*, **123**, 9451-9452.

72. Brooks R., Bruccoleri R.E., Olafson B.D., States D.J., Swaminathan S., and Karplus M., 1983, "CHARMM: A Program for Macromolecular Energy, Minimization, and Dynamics Calculations", *Journal of Computational Chemistry* **4**, 187-217.

73. MacKerell Jr. A.D., Brooks B., Brooks C.L., Nilsson L., Roux B., Won Y. and Karplus M., 1998, "CHARMM: The Energy Function and Its Parameterization with an Overview of the Program", in *The Encyclopedia of Computational Chemistry*, **1**, 271-277, P. v. R. Schleyer *et al.*, Editors, John Wiley & Sons: Chichester.

74. Kollman P.A., Pearlman D.A., Case D.A., Caldwell J.W., Ross W.S., Cheatham T.E., III, DeBolt S., Ferguson D.M. and Seibel G., 1998, "*AMBER* - *Encyclopedia of Computational Chemistry*, Wiley-Interscience: N.Y.

75. Berman H.M, Westbrook J., Feng Z., Gilliland G., Bhat T.N., Weissig H., Shindyalov I.N., and Bourne P.E., 2000, "The Protein Data Bank," *Nucleic Acid Research,* **28**, 235-242.

76. Bryant Z., Pande V.S., and Rokhsar D.S., 1999, "Mechanical Unfolding of a Beta Hairpin Using Molecular Dynamics," *Biophysical Journal,* 2000 **78**, 584-589.

77. Dinner A.R., Lazaridis T., and Karplus M., 1999, "Understanding the Beta Hairpin Formation" *Proceedings of the National Academy of Sciences*, **96**(16), 9068-9073.

78. Lazaridis T., and Karplus M., 1999, "Effective Energy Function for Proteins in Solution," *Proteins: Structure, Function and Genetics*, **35**,133-152.

79. Schlitter J., 1994, "Targeted Molecular Dynamics: A New Approach for Searching Pathways of Conformational Transitions," *Journal of Molecular Graphics*, **12**,84-89.

80. Schlitter J., "Targeted Molecular Dynamics Simulation of Conformational Change – Application to the T-R Transition in Insulin," 1993, *Molecular Simulation,***10**(2-6), 291-308.

81. Ferrara P., Apostolakis J., and Caflisch, A., 2000, "Thermodynamics and Kinetics of Folding of two Model Peptides Investigated by Molecular Dynamic Simulations" *Journal of Physical Chemistry B*, **104**(20), 5000-5010.

82. Gear C.W., 1966, "Numerical Integration of Ordinary Differential Equations of Various Orders", Technical Report 7126, Argonne National Laboratory.

83. Verlet L., 1967, "Computer 'experiments' on Classical Fluids I: Thermodynamical Properties of Lennard Jones Molecules", *Physical Review*, **159**, 98-103.

84. Ryckaert J.-P., Ciccotti G., and Berendsen H.J.C., 1977, "Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-alkanes", *Journal of Computational Physics*, **23**, 327-341.

85. Mathys S., Evans T.C., Chute I.C., Wu H., Chong S., Benner J., Liu X.Q., Xu M.Q., 1999, "Characterization of a Self-Splicing Mini-Intein and its Conversion into Autocatalytic N- and

C-terminal Cleavage Elements: Facile Production of Protein Building Blocks for Protein Ligation," *Gene*, **231**, 1-13.

86. New England Biolabs, Inc. IMPACT™-CN Protein Purification System Instruction Manual, Version 1.9 July 2001.

87. Cohen, B. E., McAnaney, T. B., Park, E. S., Jan, Y. N., Boxer, S. G. and Jan, L. Y., 2002, "Probing Protein Electrostatics With a Synthetic Fluorescent Amino Acid", *Science,* **296,** 1700-3.

88. Cunningham, B. C. and Wells, J. A., 1989, "High-Resolution Epitope Mapping of hGH-Receptor Interactions by Alanine- Scanning Mutagenesis", *Science,* **244,** 1081-5.

89. Simmel, F.C., Yurke, B, "A DNA-Based Molecular Device Switchable Between Three Distinct Mechanical States", *Appl. Phys. Lett.,* **80:** 883-885, 2002.

90. Simmel, F. C., Yurke, B., "Using DNA to Construct and Power a Nanoactuator", *Phys. Rev. E,* 63: 041913-[1-5], 2001.

91. Winfree, E., Lui, F., Wenzier, L.A., Seeman, N.C., "Design and Self-Assembly of Two-Dimensional DNA Crystals", *Nature*, **394**: 539-544, 1998.

92. Mao, C., LaBean, T.H., Reif, J.H., Seeman, N.C., "Logical Computations Using Algorithmic Self-Assembly of DNA Triple-Crossover Molecules", *Nature*, **407**: 493-496, 2000.

93. Thomsen D. L. III, Phely-Bobin T., Papadimitrakopoulos F.,1998, "Zinc-Bisquinoline Coordination Assemblies of High Refractive Index and Film Uniformity", *Journal of the American Chemical Society* , **120**(24), 6177-6178.

94. Chattopadhyay D., Galeska I., Papadimitrakopoulos F., 2001, "Metal-Assisted Organization of Shortened Carbon Nanotubes in Monolayer and Multilayer Forest Assemblies", *Journal of the American Chemical Society,* **123**(38), 9451-9452.

95. Chattopadhyay D., Lastella S., Kim S., Papadimitrakopoulos F., 2002, "Length separation of zwitterion-functionalized single wall carbon nanotubes by GPC", *Journal of the American Chemical Society*,**124**(5), 728-729.

96. Galeska I., Chattopadhyay D., Moussy F., Papadimitrakopoulos F., 2000, "Calcification-Resistant Nafion/Fe3+ Assemblies for Implantable Biosensors", *Biomacromolecules* , **1**(2), 202-207.

97. Galeska I., Hickey T., Moussy F., Kreutzer D., Papadimitrakopoulos F., 2001 "Characterization and Biocompatibility Studies of Novel Humic Acids Based Films as Membrane Material for an Implantable Glucose Sensor", *Biomacromolecules*, **2**(4), 1249-1255.

98. O'Connell M. J., Bachilo S. M., Huffman C. B., Moore V. C., Strano M. S., Haroz E. H., Rialon K. L., Boul P. J., Noon W. H., Kittrell C., Ma J., Hauge R. H., Weisman R. B., and Smalley R. E., "Band Gap Fluorescence from Individual Single-Walled Carbon Nanotubes", *Science* 2002; **297**: 593-596.

99. Zhao W., Song C., Pehrsson P. E., 2002, "Water-Soluble and Optically pH-Sensitive Single-Walled Carbon Nanotubes from Surface Modification", *Journal of the American Chemical Society (Communication),* **124**(42), 12418-12419.

100. Mavroidis C., Yarmush M., Dubey A. and Nikitczuk K., "Viral Protein Linear (VPL) Nano-Motors and Nano-Machines", Disclosure Invention submitted May 2002, Rutgers Docket Number 02-136. Provisional patent application was submitted by Rutgers University in June 2002.