SOCIETY OF PHYSICS STUDENTS

An organization of the American Institute of Physics

SPS Chapter Research Award Proposal

Project Proposal Title	Tiny Dancers with Moves Like Jagger: Normal Mode Analysis of Anthropomorphic Organic Molecules
Name of School	Louisiana Tech University
SPS Chapter Number	3878
Total Amount Requested	$2000.00

Abstract

This project will provide Louisiana Tech's SPS members experience installing Linux, performing parallel and GPU-based quantum and classical simulations, and developing a web-app for normal mode analysis. It also serves as a pilot project for teaching computational methods to physics undergraduates and as a tool for education and outreach efforts.

Proposal Statement

Overview of Proposed Project

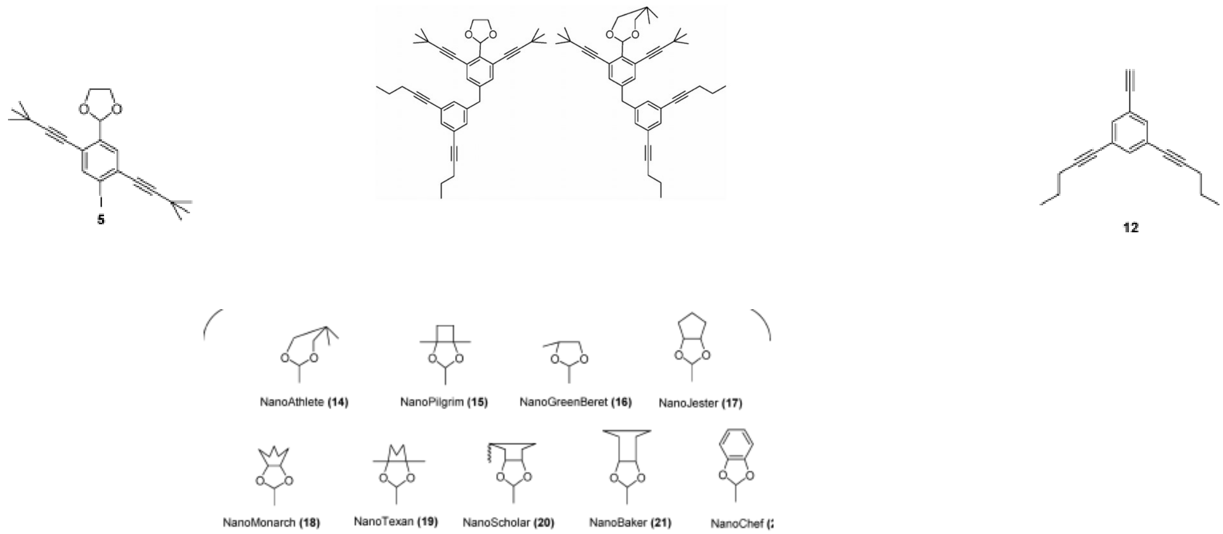

Figure 1: Images extracted from Chanteau and Tour, 2003, showing (clockwise from top left): upper body, two complete Dancing NanoPutians, lower body, and various head group substitutions.

Research Questions:
The specific research questions to be answered with this project are:

1) Are the dynamics of individual molecular structures (head groups, upper body, lower body) used in the synthesis of NanoPutians altered by the assembly of these groups into a full NanoPutian?
2) Does the addition of functional groups to the feet or hands alter the dynamics of a NanoPutian?
3) Does polymerization or cyclization of NanoPutians alter monomer dynamics?
4) Do any of the NanoPutians have moves like Jagger?

The education and outreach plan for this project is to have students configure a computer and develop a web application for demonstrating molecular simulations. The project also serves as a pilot project for training Louisiana Tech's physics students in computational methods and as the basis for a Louisiana Board of Regents "Speaking of Science" outreach activity targeting elementary and high school students.

Motivation:
The employment of NanoPutians in the chemical and industrial sector appears to be limited. However, we expect Dancing NanoPutians can be used for entertainment and educational outreach. Simply put, the presentation of complex organic molecules as personalities with unique shapes and dynamical characteristics is fun; also, there is an immediate appeal to audiences of all ages. The coupling of our Tiny Dancers web application with parallel and GPU-based quantum and classical dynamics compute engines, and the NIH's PubChem database provides educational opportunities for students at all levels of experience, from basic to advanced.

Background for Proposed Project

The organic molecules synthesized in 2003 by Chanteau & Tour are aptly labeled NanoPutians[1]. This entire class of organic molecules certainly exhibits anthropomorphic character, at least when represented in 2D. In addition to their common names and formal International Union of Pure and Applied Chemistry (IUPAC) nomenclature, one can describe each by a unique Simplified Molecular-Input Line-Entry System or SMILES string[2] . This string is ideal for informatics representations, including web/database searches, and a unique tag for indexing database entries. A properly constructed SMILES string includes all the information necessary to build 2D and 3D models, including stereoisomers. The NIH's PubChem is an example of a database that exploits these properties[3].

In general, SMILES mapping from 1D to 3D is bidirectional. A 1D SMILES string can be converted to a 2D or 3D model, and a 2D or 3D model can be converted to a SMILES string without losing information. We will exploit this 1 to 1 mapping to associate 3D and 4D physical/material properties of the NanoPutations with 1D informatics data, e.g., searching chemical databases like PubChem. The SMILES string representation also provides a basis for ordering the normal mode coordinates that is unique and informative. The Nth character of a SMILES string provides information about the local chemical topology, thus using a SMILES string to also index the elements of a normal mode vector associates this informatics data with dynamical properties of the molecule.

Dr. Bishop's broader scientific goal is to develop a framework for modeling and analyzing multidimensional data that can be employed whenever there is a systematic mapping between a 1D representation, the informatics world view, and a 3D representation, the physical-material world view. Bishop's lab has developed such a framework for biophysical studies of chromatin[4]. When combined with artificial intelligence analysis, such frameworks provide a means of discovering new connections between informatics and physics data and interpreting those connections in a manner that is meaningful from a human point of view.

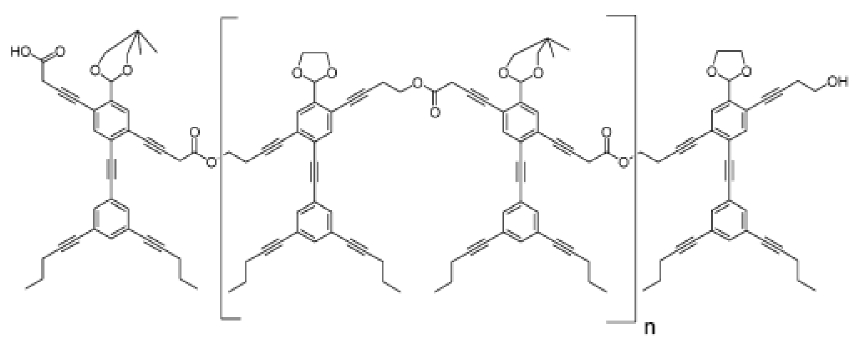

Figure 2: Polymeric assembly of Nanoputians. Image taken from Chanteau and Tour 2003. The outermost NanoPutians may be chemically bonded to form cyclic polymers (not shown).

Expected Results

Research Results:

We expect the individual functional groups (head, upper body, lower body) used in the synthesis of NanoPutians to have distinct, potentially unique dynamic signatures as determined by normal mode analysis. When these functional groups are combined to form a complete molecular structure, we expect specific modes will be selected and either amplified or eliminated. In a like manner, hetero- and homo- polymers and cyclic polymers might also exhibit such selectivity.

Education and Outreach Results:

We expect the SPS students participating in this project to learn to manage Linux-based computers, including operating system and software configuration, and to become comfortable using the command line, as well as, graphical user interfaces. They will also learn to use integrated development environments and version control software.

We expect SPS students to learn how to conduct CPU and GPU-based calculations and learn the methods of evaluating the performance of these calculations.

We expect this project to serve as a pilot for developing modules that match the research interests of other faculty and graduate students at Louisiana Tech. Many labs across campus utilize classical and quantum mechanics based simulation methods in their research. These modules will provide SPS students a pathway for contributing to research projects on campus and for research groups to train and identify undergraduates in computer maintenance.

We expect Dr. Bishop to incorporate the Dancing Nanoputians into a Louisiana Board of Regents "Speaking of Science" outreach presentation that targets elementary and high school students.

Description of Proposed Research - Methods, Design, and Procedures

Overview:

Laptops will be acquired and configured as Apache HTTP servers running the OpenSuse Leap operating system. The Amber suite of tools includes the necessary compute engines for conducting both quantum and classical modeling in both CPU and GPU modes and for conducting the normal mode analysis. Dr. Bishop has already developed web applications that integrate informatics and 3D modeling for genomics (https://dna.engr.latech.edu/~gdash/), and a NanoCar Racing web app (https://dna.engr.latech.edu/~nanocar/, pictured below). A NanoKid modeling web app is available (https://dna.engr.latech.edu/~nanokids/), but it does not have a molecular dynamics compute engine or normal mode analysis tool associated with it yet. There are no Tiny Dancers.

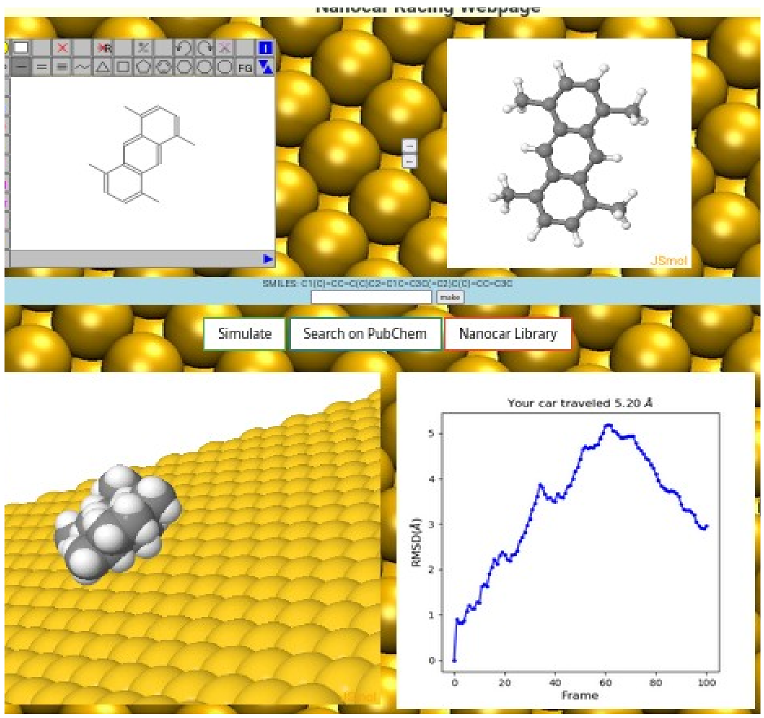

Figure 3: The NanoCar Racing web app https://dna.engr.latech.edu/~nanocar that will be modified to create the Tiny Dancers web app. Top: 1D SMILES string, 2D Chemical Sketch and 3D representations are used to “Simulate” a race, “Search PubChem” or add an entry to the “Nanocar Library”. The race is computed as diffusion on a gold surface and appears as molecular animation in JSmol. Race results are evaluated based on root mean square displacement of the molecule as a function of time.

Methods:

After acquiring the laptops, the students will install the latest version of OpenSuse Leap and configure it as an Apache HTTP server. Molecular modeling and analysis tools (Gaussian, Amber, NAMD, GROMOS, VMD) will be compiled (if necessary) and configured. Note that only Amber is required to achieve the specific goals of this project[5; 6] , but students will be encouraged to explore how the same or similar computations can be achieved with other computational tools. The Tiny Dancer web application will be created by copying the existing Nanocar Racing web app to any user's “public_html” directory on a suitably configured Apache HTTP server and making necessary modifications to the computational tasks associated with the Linux shell script invoked by the “Simulate” button.

For the NanoCar Racing web app OpenBabel[7] is used to convert a SMILES string or 2D sketch into a 3D model that is displayed with Jsmol(http://www.jmol.org/). The Simulate button in this web app initiates a workflow that optimizes the initial 3D conformation and assigns partial charges to each atom in the structure using quantum mechanics based techniques. The molecule is then docked onto a gold surface and the entire system is parameterized and subjected to a short molecular dynamics simulation using classical mechanics. The root mean square displacement of the NanoCar’s center of mass motion is determined and reported to assess the NanoCar’s racing ability. All of these computations are combined into a single unix shell script that is initiated by simply pressing Simulate.

For the Tiny Dancers web app the Simulate button will be modified to analyze the normal modes of motion of the molecule observed during a short molecular dynamics simulation. This is a standard calculation provided with the Amber suite of tools. We expect that even the largest of the tiny dancers can be modeled with our existing workflow since the NanoCar Racer has been applied to molecules that are larger than any of the NanoPutians.

Plan for Carrying Out Proposed Project

Personnel:
The primary students involved are Mac Sandel (Undergraduate Physics and SPS member) and Caden Edwards (Undergraduate Physics and SPS member). While Mac and Caden will lead the experiment, the research project is open to all SPS members.

Expertise:
Dr. Bishop earned his PhD from Klaus Schulten's Theoretical and Computational Biophysics Group in 1996. Since then he has continued to utilize and develop computational methods for the simulation and analysis of biomolecular complexes. He acquired a Little Fe computer in 2013 and with support from the Louisiana Board of Regents has used it to demonstrate real time molecular dynamics simulations and "The ABC's of DNA" to high school and elementary school students all across Louisiana. He is eager to add Nanocars and Tiny Dancers to his list of outreach activities.

Research Space:
The project will primarily be conducted in Bishop's Theoretical Molecular Biology Lab at Louisiana Tech. This location provides a central location for all participants to meet, provides desk space for the students, and houses related computer resources. Working in Bishop's lab also provides opportunities for SPS students to interact with graduate students.

Contributions:
In addition to the proposed laptops to be purchased, Bishop's Lab also has a Little Fe and a few decommissioned computers that can be used for configuration testing.

Project Timeline

Note: We have set the date for the configuration of the laptop computers to be February 2022; however, given recent delays in shipment of computational equipment we have designed the project timeline such that this activity can be moved to other dates without significantly affecting overall progress.

January 2022: Initiate grant accounting. Update and finalize purchase order. Analyze the design and implementation of the existing NanoCar web application.

February 2022: Receive laptop computers. Install OpenSuse and configure laptops as an Apache HTTP servers with full KDE desktop user environment. Compile and install molecular modeling software and explore existing modeling tutorials as means of both testing the installation and learning the basics of molecular modeling.

March 2022: Plan modifications of NanoCar web application and design workflow to be initiated by the Simulate button.

April 2022: Develop SMILES strings representations of the various head groups, upper and lower body parts. Beta testing of the Simulate button.

May 2022: Extend SMILE strings library to include whole NanoPutians and finalize basic Simulate button computations.

Summer Break: Work on developing advanced models, e.g. hetero- , homo- and cyclic polymers and extend functionality of Simulate button. Compare GPU vs. CPU implementations. Note the laptops will also be configured to support remote access so that students can continue to work on this project regardless of whether or not they are physically located on campus.

September 2022: Evaluate CPU and GPU performance of the Simulate button and begin to categorize normal mode analysis as a personality inventory for each NanoPutian. Update Bishop's list of outreach activities for the Louisiana Board of Regents Speaking of Science series to include "Tiny Dancers with Moves like Jagger. Simulating Anthropomorphic Organic Molecules on Parallel Computers" and "A Day at the Races: NanoCars at the Golden Brickyard"

October 2022: Finalize results, including outreach materials and demos for the Speaking of Science seminars. Explore alternate compute engines for use in the Simulate button workflow as time permits.

November 2022: Advertise website to public and present Tiny Dancers during a College of Engineering and Science Thursday Science seminar. Evaluate project as an undergraduate experience in computational modeling.

December 2022: Submit final report to SPS national and submit abstract for Louisiana Tech University's undergraduate research symposium in 2023.

Budget Justification

The budget includes only two items: two multicore, CUDA-capable laptop computers. The operating system (OpenSUSE) and webserver (Apache) are open-source. The modeling and analysis tools include a combination of public (namd, vmd, ambertools, jsmol, nglview) and licensed software tools (Gaussian and Amber Suite). The university maintains a license for Gaussian, and Bishop's Lab holds a license for the Amber Suite.

Given the volatility in the supply of computer parts, especially CUDA-capable graphics cards, we decided against building our own computer or updating Bishop's Little Fe.

The quote provided is a representative computer configuration intended to show that we can acquire two laptops with the necessary hardware specifications. The idea of two laptops is, in part, forward-looking. We wish to acquire a variety of computing resources for use by the SPS students. It also allows students to keep one "properly" configured and running while exploring alternate configurations and installations on the other..

Bibliography

[1] **Chanteau, S. and Tour, J.** (**2003**). *Synthesis of Anthropomorphic Molecules: The NanoPutians*, Journal of Organic Chemistry 68 : 8750-8766.

[2] **Weininger, D.** (**1988**). *SMILES, a Chemical Language and Information System: 1: Introduction to Methodology and Encoding Rules*, Journal of Chemical Information and Computer Sciences 28 : 31-36.

[3] **Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B.; Zaslavsky, L.; Zhang, J. and Bolton, E. E.** (**2021**). *PubChem in 2021: new data content and improved web interfaces.*, Nucleic Acids Research 49 : D1388-95.

[4] **Li, Z.; Sun, R. and Bishop, T. C.** (**2020**). *Genome Dashboards: Framework and Examples.*, Biophysical Journal 118 : 2077-85.

[5] **Case, D.; Cheatham, T. E.; Darden, T.; Gohlke, H..; Luo, R.; Merz, K..; Onufriev, A.; Simmerling, C.; Wang, B. and Woods, R. J.** (**2005**). *The Amber Biomolecular Simulation Programs.*, Journal of computational chemistry 26 : 1668-88.

[6] **Salomon-Ferrer, R.; Case, D. and Walker, R.** (**2013**). *An overview of the Amber biomolecular simulation package*, Wiley Interdisciplinary Reviews: Computational Molecular Science 3 : 198-210.

[7] **O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T. and Hutchison, G. R.** (**2011**). *Open Babel: An open chemical toolbox.*, Journal of Cheminformatics 3 : 33.