Bridging the Gap Between Computation and Experiment to Understand Structural Dynamics in the Hairpin Ribozyme

David A. Paul†, Bryan A. Leland†, and Dr. Brent P. Krueger
Department of Chemistry
Hope College
Holland, MI 49423
† authors contributed equally to this work

Funding provided by NSF-REU Chemistry, NSF-CIEG, NSF-MRI, ACS-PRF, and HHMI
We are developing a new method for studying the structural dynamics of biological systems that brings together fluorescence spectroscopy with computational modeling, providing a more complete understanding of the dynamics than is possible with either technique individually. Initial studies focus on a relatively simple soluble system—the hairpin ribozyme. Before molecular dynamics computer simulations can begin, force field parameters must be developed for the fluorescent probe molecules used in the experimental studies: fluorescein, tetramethylrhodamine, indiocarbocyanine-3, and indiocarbocyanine-5. Parameterization has been carried out using quantum mechanical calculations at the B3LYP/6-31G(d) level to determine low energy conformations and electrostatic potentials; atomic charges were derived using the RESP charge fitting procedure. Several RNA-fluorescent probe systems were constructed for use in the AMBER molecular dynamics package. In all cases the RNA sequence and structure were based on the hairpin ribozyme crystal structure (1X9C).
**Introduction**

One key advantage of FRET is that the experimental measurement is sensitive to structural dynamics in the system. However, assumptions that are usually made during FRET analysis necessarily remove those dynamic details. Several research groups have made the connection between calculation and experiment by comparing values obtained for the average Donor-Acceptor distance (green arrow, below). The goal of this work is to develop a new method that makes a more direct connection between fluorescence data and simulation (red arrow, below) and, therefore, avoids the usual assumptions and averaging that remove critical detail. Thus, by bridging the gap between computation and experiment more directly, we hope to improve the utility of FRET with soluble systems such as the hairpin ribozyme as well as with membrane-bound proteins, which play a key role in signal transduction.

The specific systems under study consist of four oligomers (right) taken from the hairpin ribozyme, each tagged with two of the four fluorescent probes (fluorescein, tetramethylrhodamine, indiocarbocyanine-3, and indiocarbocyanine-5).

\[
E = \sum \frac{k_{\text{bonds}}}{2} (r_{ij} - r_{ij})^2 + \sum \frac{k_{\text{angles}}}{2} (\theta_{ij} - \theta_{ij})^2 + \sum \frac{V_{\text{torsions}}}{2} (1 + \cos(\omega_{ij} - \gamma)) + \sum \sum 4\epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{6} \right] + \sum \sum \frac{q_i q_j}{r_{ij}}
\]

Molecular Dynamics uses the force field equation (above) to calculate the energy of the system as a function of its (constantly-changing) structure throughout the simulation. The force field equation expresses the energy as an aggregate of all forces imposed upon each atom in the system, and is calculated at each point of the MD simulation. The forces that are accounted for in the equation include: bonding interactions, bond angles, torsional forces, electrostatic interactions, and Van der Waals forces.
Methods

QM Structure Determination

Quantum Mechanical Calculations were carried out on all four probes and a linker molecule using Gaussian 03 Software at the B3LYP/6-31G(d) level. The WebMO user interface was used to submit Gaussian jobs to the MU3C and Curie Clusters at Hope College. These calculations were used to determine the lowest-energy configurations of the molecules. Care was taken to identify multiple low-energy configurations when present.

Charge Derivation

The electrostatic potentials of the dyes and linker were derived using Gaussian 03 at the HF/6-31G(d) level. This ESP was used to determine the atomic charges via the Multi-conformational RESP method described in Cornell, W.D., Cieplak, P., Bayly, C.I., Kollman, P.A., JACS, 1993, 115, 9620-9631. The caps used on the dyes for RESP charge determination were the traditional ACE-NME groups, while the linker was capped using ACE and a phosphate group. All charges used for the caps were obtained from Cieplak, P., Cornell, W.D., Bayly, C., Kollman, P.A., J.Comp.Chem., 1995, 16(11), 1357-1377.

Probe Attachment

The linker molecule was used to connect each probe to the phosphate backbone of an RNA oligomer. Coordinates for the RNA were taken from the PDB (1X9C) and converted to AMBER format by Xleap. The Sirius Visualization and Xleap programs provided for attachment, and the charges obtained from RESP were entered using Xleap.
Methods Continued

Identifying Low-Energy Conformations

*Geometry optimization at B3LYP/6-31G(d) with strict convergence criteria.
**Results/Discussion**

**Lowest Energy Structures**

Fluorescein had four lowest energy structures. The above graph illustrates the lowest energy level that each conformation attained, and is described in terms of C20-O16 Distance. Note that each of these four conformations has a ‘left’ and ‘right’ sub-conformation that are energetically identical.
The following charges for fluorescein were obtained from the multi-conformational RESP fit.

The graph below illustrates the atomic charges for fluorescein determined through several different RESP fitting procedures. RED charges were provided by F.-Y. Dupradaeu.

Raw charge data vrs atom (note: the NME cap is composed of the atoms on the LEFT)
Acknowledgements

Assistance provided by:
Dr. Ross Walker, San Diego Supercomputer Center
Mr. Paul VanAllsburg, Hope College
Dr. William Polik, Hope College Chemistry
Dr. Francois Dupradeau, Université de Picardie - Jules Verne, Amiens

Funding and computational resources provided by:
Howard Hughes Medical Institute
ACS-Petroleum Research Fund
National Science Foundation REU & CIEG programs
National Science Foundation MRI program
Hope College Computational Science and Modeling Laboratory
Midwest Undergraduate Computation Chemistry Consortium
San Diego Supercomputer Center (and UCSD)
TeraGrid Cyberinfrastructure