UnityMol: Interactive Scientific Visualization for Integrative Biology

Sébastien Doutreligne, Tristan Cragnolini, Samuela Pasquali, Philippe Derreumaux, Marc Baaden

Laboratoire de Biochimie Théorique, CNRS UPR9080, Univ Paris Diderot, Sorbonne Paris Cité
13, rue Pierre et Marie Curie, 75005, Paris, France

Introduction

Nowadays, an increasing number of structures of molecular assemblies is available in a variety of databases. With the advances in experimental structure determinations, the size of these assemblies keeps expanding. Descriptive data from various measurements is abundant. Studying and exploring the high-dimensional set of properties of such huge structures is part of today's challenges to modern biology. We intend to transform UnityMol in a tool for both **visualization and analysis** of such data sets.

Visualization of Biological Data

UnityMol[1] primary use is to render **biological molecular systems**. UnityMol includes both original and more traditional implementations of **molecular graphics algorithms** based on polygonal static and ray-casted dynamic rendering techniques. The former provides common tools for a static view of the system, the latter elegantly renders dynamic biomolecular events such as the formation of bonds and interactions.

UnityMol makes use of polygons to render molecular features, including among others secondary structures based on splines, molecular isosurfaces determined by marching cubes and a novel polysaccharide representation for complex sugar molecules.

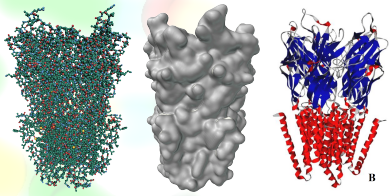


Figure 1: Protein composed of 5 identical sub-units forming an ion channel (PDB ID: 4HFI). The system has a total of 13195 atoms. Left: Hyperballs rendering. Middle: surface. Right: Secondary Structure.

UnityMol comes with an **HyperBalls** GPU implementation. The heart of this **ray-casting technique** is the use of **quadric surface equations**, perfectly compatible with a traditional rendering pipeline. Here, we successfully extend the representation to the **coarse-grained (CG) HiRE-RNA model**. Based on this model developed in the laboratory [2], nucleo-bases can be rendered as ellipsoids. Visual feedback of molecular properties can be enhanced using a lit-spheres lighting pass.

Complex and large-scale data may be rendered on advanced displays such as the 33 megapixels 4,3x2,4m **display wall** at Maison de la Simulation in Saclay (Figure 3).

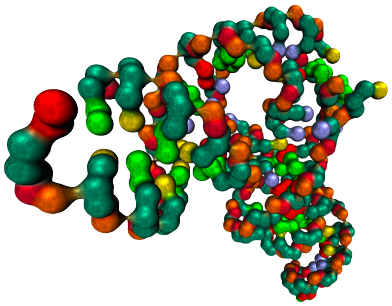


Figure 2: An RNA molecule of type Transfer RNA (tRNA) as renderer in UnityMol after the lit-spheres lighting pass (textured).
PDB ID: 3L0U, 73 nucleotides, 1564 atoms, 475 coarse-grained particles.

Figure 3: UnityMol on the display wall at Maison de la Simulation (Saclay)

Molecular Dynamics

We effectively coupled UnityMol with Molecular Dynamics (MD) Simulations through MDDriver [3] and the IMD protocol. Atoms positions are retrieved and displayed in UnityMol. IMD also enables us to apply **external forces** in simulation (Figure 4).

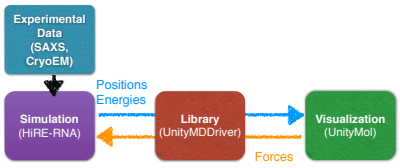


Figure 4: Overview of our architecture to bridge visualization and simulation

In the lab, we use the **HiRE-RNA** simulator for RNA molecules [4]. As a CG model, it provides an excellent compromise between simulation speed and biological fidelity, being in our experience more robust with respect to user interactions than computations carried out at an all-atom level. Enriched by special screen-space effects, the underlying dynamic molecular behavior is highlighted visually.

HiRE-RNA can use experimental heterogeneous raw data to guide simulations (SAXS, CryoEM). These data are a warrant for the **biological accuracy**. Simulations are also faster.

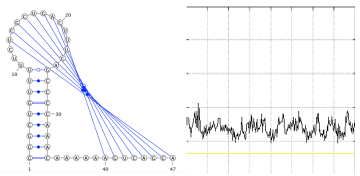
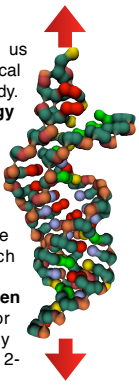


Figure 5: VARNA graph of the triple helix (Left) and Energy (kcal/mol) over time (200 timesteps displayed - Right)

Data Analysis

The HiRE-RNA simulator provides us some information about biophysical properties of the system under study. Among them figures the **total energy** of the molecule. In UnityMol, we provide the user with this information by the mean of plots. This helps the user to figure out what is happening and how the molecular structure relates to the resilience and function of such system.

We also incorporated an **hydrogen bonds detection algorithm** for interactive analysis. We ultimately want to display the secondary 2-



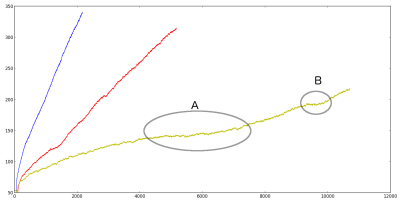


Figure 6: Distance (Angstrom) over time with various forces. yellow: 350 pico Newtons (pN), red: 700pN, blue:1400pN

dimensional structure of molecules to show how it relates to the biophysical function and properties [5].

To demonstrate the new abilities of UnityMol, we setup an experiment in which an RNA triple helix (PDB ID: 2K96) is stretched from both ends. We measured the **distance between atoms** at each side over time. This is analog to what is done experimentally with Atomic Force Microscopy. We performed the test with a constant force of 350pN and stopped once the helix was completely straight. In such simulation and throughout our tests, we used an integration time of 4 femtoseconds. The experiment last about 11,000 time steps, providing us with 44 picoseconds of simulated time.

We can map the distance stagnation with the resilience of bonds between base pairs. These bonds, called hydrogen bonds, are known to favor the **stability** of such molecular assembly.

Navigation guided by data and Human/Machine interaction

Exploring these molecules is hard because structures are usually complex. This also reduces our ability to interact with the simulation. We implemented many solutions to this problem. One is the **guided navigation based on data**. The goal is to set a target for the camera to track. The camera is set to a position at which their is the less possible obstacles. To achieve this goal, spherical coordinates of atoms are projected on a 2D map. A **Voronoï diagram** is generated from this map. The algorithm finds the biggest empty circle from Voronoï vertices. The selected vertex gives the position the camera will travel to.

UnityMol supports **special devices** such as an haptic arm and a 3D mouse via VRPN. The interactive approach opens up perspectives to guide simula- tions via user input, within a dedicated graphical environment. Peripherals such as a 6 d.o.f. 3D mouse ease the navigation in the virtual space. The user can change the point of view intuitively, and focus on the relevant part. Haptic devices help to render ac- curate molecular models more real and tangible to the scientists. Hence, the user feels an immediate force feedback by a straightforward combination of molecular modeling and virtual reality leading to an intuitive understanding of the causal relationship between the theoretical model and its biologically relevant properties.

References

(1)Z. Lv et al. Game on, science - how video game technology may help biologists tackle visualization challenges. PLoS ONE, 8(3):e57990, Mar. 2013.

(2)M. Chavent et al. GPU-accelerated atom and dynamic bond visualization using hyperballs: a unified algorithm for balls, sticks,and hyperboloids. J. Comput. Chem., 32(13):2924–2935, Oct. 2011.

(3)O. Delalande et al. Complex molecular assemblies at hand via interactive simulations. J. Comput. Chem., 30(15):2375–2387, Nov. 2009.

(4)S. Pasquali and P. Derreumaux. HiRE-RNA: a high resolution coarse-grained energy model for RNA. J. Phys. Chem. B, 114(37):11957–11966, Sept. 2010.

(5)K. Darty et al. VARNA: Interactive drawing and editing of the RNA secondary structure. Bioinformatics, 25(15):1974–1975, Aug. 2009.