High-Performance and Quantum Computing for Solving Biological Problems

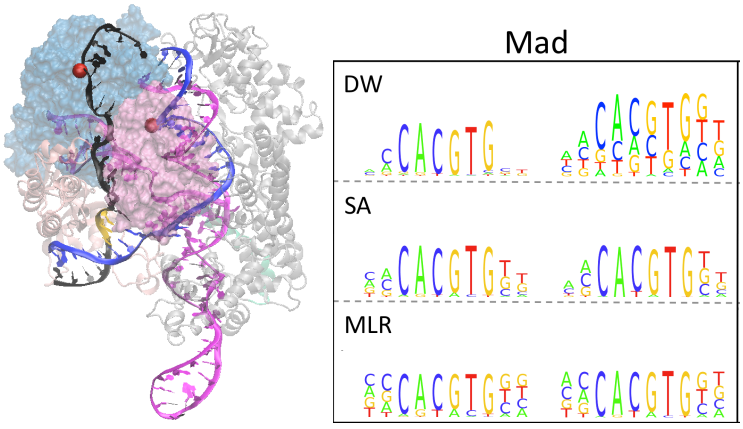
Rosa Di Felice and Remo Rohs
Department of Biological Sciences, Computational Biology and Bioinformatics, and Department of Physics and Astronomy, University of Southern California, Los Angeles, CA 90089

Ph: (213) 740-0555 Email: difelice@usc.edu

Quantitative approaches have become increasingly decisive to solve biological problems. In particular, two classes of computational methods are at the core of biological research: (i) methods that are based on the physicochemical properties of molecules, cells and systems and (ii) computer science methods that can navigate big data, which characterize genomics, proteomics, interactomics, etc. In this presentation we illustrate two examples of such research, applied to protein-DNA complexes.

The first example is an atomistic-level molecular dynamics (MD) simulation of CRISPR-Cas9 complexes on the microsecond time scale [1]. Our findings support a model in which the unwound non-target DNA strand is stabilized by a positively-charged patch located between the two nuclease domains of Cas9, and its flexibility increases in an uneven fashion upon scissions of the DNA backbone. The computed intra-DNA distances compare very well to electron paramagnetic resonance (EPR) measurements using spin labels and establish a joint experimental-computational approach to monitor conformational changes in DNA upon interacting with Cas9.

The second example is a quantum machine learning application to simplified transcription factor (TF)-DNA binding data [2], using the commercially available D-Wave Two X (DW2X) processor [3]. The results obtained on the quantum annealer DW2X are compared to state-of-the-art classical approaches for the same simplified datasets, including multiple linear regression (MLR), simulated annealing (SA), and extreme gradient boosting. Despite technological limitations, we find an advantage in classification performance and nearly equal ranking performance using the quantum annealer for these fairly small training datasets. Thus, we propose that quantum annealing might be an effective method to implement machine learning for computational biology problems with scarce training data.



References

[1] N. Tangprasertchai, N., Di Felice, R., Zhang, X., Slaymaker, I.M., Vazquez Reyes, C., Jiang, W., Rohs, R. & Qin, P.Z. CRISPR-Cas9 mediated DNA unwinding detected using site-directed spin labeling. *Manuscript under review*.

[2] R. Li, R., Di Felice, R., Rohs, R. & Lidar, D. Quantum vs classical machine learning applied to a simplified computational biology problem. *Manuscript in preparation*.

[3] http://www.dwavesys.com/resources/publications.