Undergraduate Research Symposium May 18, 2018 Mary Gates Hall

Online Proceedings

Poster Session 1

Commons East, Easel 68

11:00 AM to 1:00 PM

Image Analysis for Automation of Labeling Multivariate AFM Images of Coherent Bio-Nano Interfaces for Machine Learning Applications

Erik Matthew Johnson, Senior, Materials Science & Engineering
Jack Otto Ryan, Junior, Pre Engineering
Mentor: Siddharth Rath, Materials science and engineering, Genetically Engineered Materials Science and Engineering Center
Mentor: Mehmet Sarikaya, Materials Science & Engineering

Atomic Force Microscopy (AFM) images of peptide self-assembly on two dimensional atomically-thin inorganic solid substrates have features that are difficult to extract because of the diverse conformations these peptides have on the surfaces from nanoislands, no nanowires to confluent films and disorganized nanostructures. Analysis of these images can produce parameters such as thickness, aspect ratio, order/disorder ratio, and orientation distributions quickly and accurately. The first step is to input data in a table and convert it into an array. The images are then converted into a grayscale version, its background noise subtracted and subsequently renormalized. User determined edge detection techniques are then used to delineate the edges in the images. Appropriate segmentation methods are then used to separate different types of nanostructural textures and phases, and coupled with the edge information. Several further parameters such as percent area and volume fraction of each phase (surface coverage), percent-ordering in the long-range ordered phases, their shape and orientation distributions, as well as relative sizes are identified and quantified from the images. The sets of unique parameter are then used as the bases for assigning specific labels for the degree of molecular recognition of the substrate by the peptides and eventual orientation relationships between the peptide nanostructures and the crystalline lattices of the 2D atomically thin solid substrates. These labels are then used to train a machine learning algorithm to cluster these images and relate them to processing parameters associated with them via a relational database. The eventual goal here is to parameterize the images so that we can predict the ordering characteristics of any peptide-substrate pair to accelerate design of future technologies, e.g., bionanosensors, emanating from these hybrid material systems.

Poster Session 2

MGH 241, Easel 127

1:00 PM to 2:30 PM

Deciphering Interactors of the α1D Adrenergic Receptor

Diana Tram Anh Dinh, Senior, Biochemistry
Mentor: Chris Hague, Pharmacology, University of Washington School of Medicine
Mentor: Dorathy-Ann Harris, Pharmacology
Mentor: Eric Janezic, Pharmacology

G Protein-Coupled Receptors (GPCRs) are seven-transmembrane proteins present throughout the body that can be activated upon binding of drugs, hormones, or neurotransmitters. Acting like an inbox for messages, these multi-protein complexes play a significant role in the body by regulating cell expression and signaling, making them attractive targets for drug development. One class of GPCRs, known as the adrenergic receptors (ARs) are critical in modulating the function of numerous targets, including cardiac muscle, vascular smooth muscle, and bronchial smooth muscle. There are nine AR subtypes: three α1s, three α2s, and three βs. In my research, I focus on deciphering the interactions of proteins with the α1D-AR due to a recent publication of the Hague Lab, which discovered that α1D-AR forms a multi-PDZ protein complex. To determine the participants of this complex and identify the presence of associated proteins, I first purified the proteins that potentially bind to α1D-AR. In my purification process, I first conducted Polymerase Chain Reactions (PCRs), transformed my bacteria, grew my bacteria, and finally purified my proteins of interest. With this information, my lab and I intend to identify the complexes using mass spectrometry and Octet. Determining how α1D-ARs are organized can create a deeper understanding of how drugs work in the body and assist in new drug development.

Poster Session 2

MGH 241, Easel 129

1:00 PM to 2:30 PM

Characterizing the Interaction of the hDLG1 PDZ Domains to the Alpha-1D Adrenergic Receptor
George Williams, Sophomore, Pre-Sciences
Mentor: Dorathy-Ann Harris, Pharmacology
Mentor: Chris Hague, Pharmacology, University of Washington School of Medicine

The alpha-1D adrenergic receptor (A1DAR) is a G-protein coupled receptor (GPCR) and is known for its function as a regulator for cardiovascular, urinary, and central nervous system function. This unique GPCR has a PDZ ligand at the C-terminus of this receptor. A PDZ ligand is a region on a protein that can bind to a PDZ domain on another protein and is named based on the three proteins that it was observed in. Previously, it has been discovered that this ligand interacts with the syntrophin family of PDZ domain proteins. The hDLG1 protein is a human homologue of the Drosophila disc large (DLG1) tumor suppressor protein and contains three PDZ domains that have been implicated in interacting with A1DARs. By determining how/if the PDZ domains on the hDLG1 protein bind to the A1DAR, a more accurate depiction of this macromolecular complex can eventually be determined. Several methods were used in order to determine this. First, I cloned hDLG1 PDZ domains separately as well as varying combinations of the three domains. Next, I purified the protein constructs so I could determine the binding efficiency of the hDLG1 proteins to the A1DAR via Bio-Layer Interferometry (Octet). SNAP gel assays can also be used to determine the interaction with the A1DAR. Upon reviewing the results of these techniques, the nature of the interaction between the PDZ domains of the hDLG1 protein to the A1DAR can be determined. Understanding the architecture of this structure has many implications in pharmaceuticals as this protein can be specifically targeted by medications to treat a multitude of diseases. Side effects can be minimized by targeting specific parts of this large protein complex with medications than produce more specific responses thus limiting unwanted side effects. Medications that target the proteins involved in this complex can affect the function of the cardiovascular, urinary, and central nervous system.

POSTER SESSION 4

MGH 206, Easel 173
4:00 PM to 6:00 PM

Effective Utilization of Experimental and Modeling Data in Innovation via Machine Learning, Data Analytics, and AI: Looking inside the Black Box
John Taylor (John) Hamann, Senior, Mechanical Engineering
Jack Otto Ryan, Junior, Pre Engineering
Benjamin (Ben) Mac Millan, Sophomore, Pre-Sciences
Antonio R. Crowe, Junior, Chemistry, Materials Science & Engineering
Mentor: Mehmet Sarikaya, Materials Science & Engineering
Mentor: Siddharth Rath, Materials science and engineering, Genetically Engineered Materials Science and Engineering Center
Mentor: Burak Berk Ustundag, Materials Science and Engineering
Mentor: David Starkebaum, MSE

In scientific research labs, in general, experiments are generally treated as a black box: a prepared sample goes in, something happens, and one gets results that are then obtained via elaborate characterization steps. Several important dependent or correlated parameters are either discarded or ignored because of a lack of coherent dependency analyses that require critical thinking, linking, and pattern recognition. In this research we are working to stop treating experiments and computational simulations as black boxes, and create a cohesive platform where materials used, processes and parameters utilized and results achieved can be brought together as separate but related sets of databases. In the next step, the relationships between all the different parameters can then be connected, analyzed and visualized. Machine learning and AI techniques can then be used to predict results using these databases, thereby reducing experiment time, and taking away the traditional 'trial and error' method of experimentation. The research involves creation of a software interface, with numerous image and signal processing tools and applications running on libraries made customizable to research fields, types of experiments, etc. Assorted variety of services such as parallelization, compression, data analysis, and visualization, caching (among others) are also provided. We are improving the accuracy of time series data analysis and using fingerprinting to depict all parameters for improved predictability, flexibility and accuracy. When fully developed, we anticipate that the program will enable experimental and computational researchers to extensively use, customize and apply data analytics, machine learning and AI even in niche research in the hard sciences at the intersection of biology and genetics, materials science (physics, chemistry) and engineering, and computational modeling and informatics, enabling faster and accurate cross disciplinary innovation in technology and medicine. The research is supported by NSF-DMREF (DMR-1629071) program at GEMSEC-MSE, as part of National Materials Genome Initiative.