Online Proceedings

SESSION 1D

FRONTIERS IN PEPTIDE AND PROTEIN SCIENCE

Session Moderator: Rachel Klevit, Biochemistry
MGH 228
12:30 PM to 2:15 PM

* Note: Titles in order of presentation.

VSEPR Encoding of Peptide Structures for Predicting Binding-Affinity
Jonathan Taylor (Jonathan) Francis Landau, Junior, Mathematics
Ximing Lu, Junior, Computer Science (Data Science), Statistics
Undergraduate Research Conference Travel Awardee
Mentor: Mehmet Sarikaya, Materials Science & Engineering
Mentor: Siddharth Rath, Materials science and engineering, Genetically Engineered Materials Science and Engineering Center

The goal of this project is to encode peptides, i.e., short amino acid sequences, in terms of smaller molecular components such as their VSEPR (Valence Shell Electron Pair Repulsion) features for training interpretable models with reasonable predictability of functionality. This enables us to go beyond the limitations imposed by treating peptides as sequences of letters, thereby enabling a generalized encoding that works for lipids and other biomolecules that are of interest in a comparable scenario. Biological processes are rarely disjoint and often complicated which lends justification to our approach. Current methods for binding affinity prediction, such as one-hot encoding, where letter-based sequences are converted to a binary representation, do not take into account molecular level features. Combined with a neural network, such a simple encoding is better at predicting affinities of short peptides, e.g., 5-9 Amino acids long, but with an increase in length from 9 to 10, the predictability suffers an exponential drop. Several alternatives have been employed in literature, but they also suffer from the negative impact of distal effects. In the VSEPR approach, encoding peptides in terms of their component functional-group geometries enables us to encode the actual physical length, rather than the number of amino acids. This leads to an overlap between peptides of different length, thereby reducing the fall in predictability. In this encoding, we create 5 channeled matrices with each channel corresponding to 'central-atom connectivity', 'bond-types', 'bond-lengths', 'bond-angles' and 'lone-pairs' that is then fed through a Deep Residual-Neural-Network. The metrics used to evaluate the models are Pearson-Correlation, Spearman-Rank-Correlation-Coefficient, and Area-under-Receiver-Operating-Curve. With this technique, we were able to consistently predict binding affinities of peptides without an appreciable loss between 9 or 10 length peptides. This method would allow one to create length invariant encodings, not limited to just peptides, significantly improving the practicality of using such a model. The research is supported by NSF/DMR-DMREF program under Materials Genome Initiative.

SESSION 1D

FRONTIERS IN PEPTIDE AND PROTEIN SCIENCE

Session Moderator: Rachel Klevit, Biochemistry
MGH 228
12:30 PM to 2:15 PM

* Note: Titles in order of presentation.

Learning the Ideal Similarity Matrix for Peptide Sequences with Given Functionality
Dylan Hylander, Sophomore, Engineering Undeclared
Mentor: Siddharth Rath, Materials science and engineering, Genetically Engineered Materials Science and Engineering Center
Mentor: Mehmet Sarikaya, Materials Science & Engineering

The Genetically Engineered Materials Science and Engineering Center (GEMSEC) labs revolve around designing and synthesizing genetically engineered peptides for inorganic materials (GEPIs). Experimentally characterizing GEPIs can be slow, and therefore a computational method that can predict functionalities would greatly accelerate the development of bio/inorganic interface design and implementations. The Pairwise Similarity Score is a proven predictor of relative binding affinity and has been used to predict GEPIs specific for quartz, gold, hydroxyapatite, and MoS2. In previous work, a similarity matrix was updated based on whether

a peptide (Strong or Weak binding) had higher similarity to strong peptides and less similarity with weak peptides. Our method instead obtains the most ideal similarity matrix via stochastic gradient descent to best predict the relative binding affinities. The values in an amino-acid similarity matrix are randomly initialized and subsequently updated until convergence by minimizing the errors in binding affinity prediction. 5-fold cross-validation is used as a metric to evaluate performance on test data. We expect to observe higher predictability with this learned similarity matrix than using a literature matrix. This would compound work done by the high throughput screening, confirming count numbers observed during phage display are correlated with their actual binding affinity, while using a novel large dataset to test known successful predictive models. All in all, the work carried out in this project accelerates the development pace of bio-nano-devices of the future. The research is supported by NSF/DMR-DMREF program under the Materials Genome Initiative.

SESSION 1I

ROBOTS HUMAN SYSTEMS

Session Moderator: Santosh Devasia, Mechanical Engineering

MGH 248

12:30 PM to 2:15 PM

* Note: Titles in order of presentation.

Assistive Feeding Using Fingertip Tactile Sensors with Rich Haptic Feedback

Connor Geiman, Senior, Mechanical Engineering
Mentor: Siddhartha Srinivasa, CSE
Mentor: Hanjun Song, Computer Science and Engineering, Paul G. Allen Center

The ability to eat our meals how and when we want is often taken for granted, but is not a reality for some with limited upper body mobility. Current haptic sensory inputs for autonomous robotic feeding systems are inconvenient and expensive. We propose using GelSight, a fingertip tactile sensor capable of providing rich haptic feedback. Built with silicone and 3D printed parts, GelSight uses a camera to track a grid of dots embedded in the silicone and estimate forces. GelSight is accurate and inexpensive to manufacture and overcomes many of the limitations of currently available and previously tested options. Based on laboratory tests and feedback from potential users we continue to improve our hardware and software, seeking to generalize our solution to any food.

POSTER SESSION 2

Balcony, Easel 107

1:00 PM to 2:30 PM

Analysis of Epithelial-Mesenchymal Transition in HEK 293 Cells

Kevin Ngoc Nguyen, Senior, Anthropology: Medical Anth & Global Hlth
Mentor: Chris Hague, Pharmacology, University of Washington School of Medicine
Mentor: Dorathy-Ann Harris, Pharmacology

Epithelial-mesenchymal transition (EMT) refers to a biologic process that allows a polarized epithelial cell, which normally functions in the basement membrane of a cell, to undergo biochemical changes that makes it express as a mesenchymal cell phenotype. This mesenchymal phenotype allows the cell to have enhanced migratory capacity, invasiveness, elevated resistance to apoptosis, and increased production of extracellular matrix (ECM) elements. The process of EMT is considered completed once the underlying basement membrane breaks down, and the mesenchymal cell becomes migratory. Another component that proves EMT is the loss of e-cadherin. E-cadherin refers to cell-to-cell adhesion and the degradation of e-cadherin levels are a hallmark of EMT happening. There are three distinct types of EMTs; I will be focusing on type II EMT. Type II EMTs are associated with inflammation/wound repair but usually stops once inflammation subsides. However, in the context of organ fibrosis, type II EMTs can continue to over-respond to a persisting inflammation and can lead to organ death. In my experiment, I hypothesize that in HEK 293 human cells, SNAP -Δ1-91 alpha-1D adrenergic receptors undergo type II EMT. SNAP -Δ1-91 alpha-1D adrenergic receptors are a truncation of the extracellular portion of the receptor. Certain receptors undergo this truncation to increase its expression. It is shown that in SNAP – Full Length alpha-1D adrenergic receptors (wild type receptors) do not undergo EMT. I will be able to observe the process of type II EMT through imaging the breakdown of the cell membrane in SNAP- Δ1-91 alpha-1D adrenergic receptors and the measuring of e-cadherin levels. The purpose of this research would be to potentially influence future therapeutic interventions that target wild type receptors to induce would repair.

POSTER SESSION 3

MGH 241, Easel 136

2:30 PM to 4:00 PM

A Neural Network for Predicting Peptide Binding Affinity

Francesca Caroline Green, Senior, Materials Science & Engineering

Louis Stokes Alliance for Minority Participation, NASA Space Grant Scholar

Mentor: Mehmet Sarikaya, Materials Science & Engineering

Mentor: Siddharth Rath, Materials science and engineering, Genetically Engineered Materials Science and Engineering Center

Our Lab, GEMSEC, uses molecular biology, bioinformatics, genome sciences, and engineering for de novo design of short amino acid sequences for various applications such as tooth remineralization strategies in dentistry, biosensing in cancer diagnostics, and bioelectronics in single-molecule detection. Designing and constructing peptides for a desired function begins with selecting the appropriate sequence of amino acids with the predictive conformation that affects the function. In this project we use latent-space representation (matrix factorization) in conjunction with a simple neural network to create a model that is able to predict peptide binding affinity to several alleles of MHC-I protein. Python was used to encode amino acids by creating data frames defining the functional groups within them, differing by n-terminus, intermediate, and c-terminus of each amino acid and their placement along the backbone of each structure. A tensor was created using the data frames describing each amino acid to encode the 9- and 10-length sequences of thousands of unique peptides from the Immune Epitope Database. Each chemical structure and peptide sequence can be described by k attributes, or latent features. Matrix factorization was used to discover the latent features and send this feature encoding to a neural network (NN) to determine binding affinity. The goal is to minimize the mean-squared-error by stochastic gradient descent in a supervised learning protocol. The two modules of matrix factorization and NN provide an optimum between interpretability and predictability simultaneously.The successful prediction of peptide binding affinity towards nanoscale targets provides novel opportunities for drug design towards targeted public health initiatives and in technology applications such as bio/nano hybrid devices. This research is supported by NSF/DMR-DMREF program under the Materials Genome Initiative.

POSTER SESSION 3

Balcony, Easel 104

2:30 PM to 4:00 PM

Differences Between 2D and 3D Cell Modeling

Jessica Giang, Senior, Public Health-Global Health, Linguistics

Mentor: Dorathy-Ann Harris, Pharmacology

Mentor: Chris Hague, Pharmacology, University of Washington School of Medicine

Mentor: Eric Janezic, Pharmacology

2D cell models have traditionally been used in labs to test the effects of new drugs on certain cell types due to the ease and convenience of use. While 2D methods are great, they often simplify the cell-to-cell interactions and may not accurately represent cell systems in humans. 3D methods show the complex cell communication systems and better simulate actual organ systems. Research comparing these two methods can inform scientists on the benefits of 3D models which can help efficiency in creating new drugs. Our lab looked into various 3D models to determine their effectiveness and reliability and looked into the differences in perceived cell mechanics and functionality between 2D and 3D methods. We tried Corning Matrigel and Corning 3D Spheroid microplates for 3D cell modeling using HEK293 cells, which are human embryonic kidney cells that were grown in lab. They are known for being easy to grow and transfect. We used SNAP-Gels, which are protein assays that show the protein levels in the cells, to ensure that the protein levels were similar between the 2D and 3D systems. We then did florescent imaging to determine cell localization and EPIC dynamic mass redistribution (DMR) to determine cell functionality. We found Matrigel to have inconsistent results, so we focused on using the spheroid microplates. Based on our initial results, we saw increased functionality and expression levels for full-length protein cells compared to cells with a truncated N-terminal protein in the 3D method. This increase in functionality and expression levels was not seen in the 2D method. Our results show that 3D modeling methods can be reliable, and do show results that differ from 2D models. This is important for future studies that require cell modeling because 3D models can provide a more accurate and reliable modeling system to create novel therapeutics.

POSTER SESSION 4

Commons West, Easel 10

4:00 PM to 6:00 PM

Differential N-Glycosylation Controls Function and Expression of α1D-Adrenergic Receptors

George Williams, Senior, Neurobiology

UW Honors Program

Mentor: Eric Janezic, Pharmacology

Mentor: Chris Hague, Pharmacology, University of Washington School of Medicine

Mentor: Dorathy-Ann Harris, Pharmacology

G-protein coupled receptors (GPCRs) - characterized by seven transmembrane alpha helical domains - are the largest family of membrane proteins, constituting ˜1% of the human genome. The α1D-adrenergic receptor (A1DAR) is a GPCR that regulates function of the cardiovascular, urinary, and central nervous systems. Dysfunction of this receptor can lead to various diseases including schizophrenia, benign prostate hypertrophy, hypertension, and PTSD. Prazosin, a non-specific α1-antagonist is the first line treatment for PTSD, however, chronic use has deleterious side effects including orthostatic hypotension and potentially fatal reflex tachycardia due to interactions with off-target related receptors. Thus, understanding how A1DARs are regulated will allow for the development of targeted therapeutics. To this end, the Hague Lab has previously discovered that A1DAR undergoes an endogenous cleavage of its extracellular N-terminal domain, affecting its membrane localization and response to agonist stimulation. Located within the N-terminal domain of A1DAR are two glycosylation sites at amino acids 65 and 82. Currently, how glycosylation of these sites regulates the cleavage event remains unknown. To characterize this phenomena, I used molecular cloning to mutate the glycosylation sites of A1DAR in the pSNAP vector for expression in Human Embryonic Kidney 293 (HEK293) cells. Near Infrared PAGE analysis revealed that glycosylation of both amino acids is required for cleavage and proper expression of A1DAR. Sucrose density gradient and dynamic mass redistribution further showed that glycosylation controls function and trafficking of A1DAR to the membrane. These results allow for the development of targeted medications specific to the N-terminal glycosylation sites of A1DAR, further reducing the potential side effects experienced by patients.