developed DeepSPT a machine learning framework for single particle tracking data analysis, processing, and classification. DeepSPT, exclusively using diffusional behavior and within milliseconds instead of days, maps with 95% accuracy the key timepoints of rotavirus internalization, predicts cellular localization of clathrin-coated pits and differentiates early from late endosomes. DeepSPT leverages masked diffusional variations outputting biological information, usually requiring multi-colour labeling, and provide results indicating that besides structure, motion encodes biomolecular identity and functionality.

208-Plat

Optimization and historical contingency in protein sequences

Anne-Florence Bitbol.
EPFL, Lausanne, Switzerland.

Protein sequences are shaped by functional optimization on the one hand and by evolutionary history, i.e. phylogeny, on the other hand. A multiple sequence alignment of homologous proteins contains sequences which evolved from the same ancestral sequence and have similar structure and function. In such an alignment, correlations in amino acid usage at different sites can arise from structural and functional constraints due to coevolution, but also from historical contingency. Correlations arising from phylogeny often confound coevolution signal from functional or structural optimization, impairing the inference of structural contacts from sequences. However, inferred Potts models are more robust than local statistics to these effects, which may explain their success. Dedicated corrections can further increase this robustness. Moreover, phylogenetic correlations can in fact provide useful information for some inference tasks, especially to infer interaction partners from sequences among the paralogs of two protein families. In this case, signal from phylogeny and signal from constraints combine constructively, and explicitly exploiting both further improves inference performance. Protein language models have recently been applied to sequence data, greatly advancing structure, function and mutational effect prediction. Language models trained on multiple sequence alignments capture coevolution and structural contacts, but also phylogenetic relationships. They are able to disentangle signal from structural constraints and from phylogeny more efficiently than Potts models, and they have promising generative properties. Furthermore, they allow predicting interacting partners from protein sequences, outperforming traditional coevolution methods on difficult datasets.

209-Plat

Capturing the biomechanics of SARS-COV-2/antibody complexes by GōMartini simulation

Luis F. Cofas-Vargas, Adolfo Poma.
Department of Biosystems and Soft Matter, Institute of Fundamental Technological Research, Polish Academy of Sciences, Warsaw, Poland.

Molecular dynamics (MD) simulation is a powerful tool for revealing the underlying mechanisms governing protein mechanostability. A typical disadvantage of the all-atom representation is the use of pulling speeds several orders of magnitude higher than those employed in single-molecule force spectroscopy (SMFS). In contrast, coarse-grained (CG) representation has the advantage of reducing the computational cost at the cost of losing information on the interaction strength at protein interfaces. This effect is more pronounced in protein complexes. The GōMartini approach is an alternative tool to circumvent this limitation, and in its recent implementation, it employs virtual sites near the C-alpha atom positions in the Martini 3 force field. This approach requires the determination of a contact map that includes the most relevant interactions between residues (i.e., native contacts). Large-scale applications, including mechanical stability and conformational changes, can be studied using the GōMartini. In this work, we have applied this approach to study the mechanostability associated with the immune response. Through refinement of the interaction potential between residues at the interface of the protein complex, we reproduced the results of all-atom MD and contrasted them with reported experimental values. GōMartini approach allows us to approach the speeds of atomic force microscopy (AFM) cantilevers in SMFS while preserving crucial information about the interaction between residues. This method is extremely useful in identifying the most crucial interactions that are responsible for the enhanced mechanostability in SARS-CoV-2 variants, information that can be used to develop antibodies with greater affinity.

210-Plat

Deep learning convolutional neural networks for the quantification of glycogen in electron micrographs of calcium-stressed human muscle

Eduardo Rios[1], Montserrat Samso[2], Eshwar Reddy Tammineni[1], Lourdes C. Figueroa[3], Carlo D. Manno[1], Sheila Riazi[4].
[1]Department of Physiology and Biophysics, Rush University, Chicago, IL, USA, [2]Department of Physiology and Biophysics, Virginia Commonwealth University School of Medicine, Richmond, VA, USA, [3]Department of Molecular Biophysics and Physiology, Rush University Medical Center, Chicago, IL, USA, [4]Department of Anesthesiology and Pain Management, University of Toronto, Toronto, ON, Canada.

42% of subjects diagnosed with malignant hyperthermia susceptibility (MHS) in Toronto's MHIU later developed hyperglycemia; many transitioned to diabetes 2 (Altamirano, BJA, 2020). Muscle glucose processing was altered due to the patients' elevated $[Ca^{2+}]$cyto, thereby lowering their glycogen content. Glycogen exists as 20-40 nm granules, at locations dictated by ATP needs, glycogen synthesis, or space availability. Systematic quantification of granules in cell loci should clarify their distribution mechanism. Changes under disease conditions could reveal pathogenesis. Given their large numbers, quantifying granules in EM images should be done automatically. We developed artificial intelligence (AI) neural networks (''models'') for two tasks: quantifying granules and locating them by cell region. The models were applied on images (4096 × 4096 pixels, 0.4 or 0.8 nm/pixel) of ultrathin muscle slices from normal and MHS patients. Combined, their results yielded regional content and density of granules. One approach developed categorical classification models. Images were divided into sub-images of 80 nm × 80 nm; a ''locations classifier'' found the location within cells of individual sub-images (A-band, I-band, Z-disk, SR, near-SR, mitochondria), and a ''granules classifier'' quantified their granules content. Another approach developed semantic segmentation models, whereby every pixel of the images is classified. A ''locations segmenter'' assigned pixels to locations; a ''granules segmenter'' classified pixels as within or outside granules. Both classifiers and segmenters reached ∼82% accuracy but segmenters were more robust vs changed magnification or image quality. In preliminary results, granule density was greatest in inter-myofibrillar regions, greater in I than A bands, and greater in normal than in MHS images. We will detail model training and testing, compare the relative virtues of the categorical and segmentation approaches, provide codes and discuss other possible applications of AI to image analysis.

211-Plat

Are phases an appropriate description for cells? Fluctuation-dominated regime in finite systems with many components

Martin Girard.
Max-Planck-Institut für Polymerforschung, Mainz, Germany.

Phase separation has emerged as an important topic for cellular function. From lipid rafts to liquid-liquid phase separation, our current understanding is that it is crucial for organization. We putatively expect that rules extracted from simple systems, two component mixtures, extend to multicomponent systems. While this is true in the thermodynamic limit, I will discuss here the thermodynamic limit for multicomponent systems. Using a toy model, I will show that what we consider ''large systems'' is largely subjective and dependent on details in multicomponent systems. For ''small'' systems, rules are very different, and the system is dominated by fluctuations. Usual assumptions, such as equivalence of thermodynamic ensembles, are broken. Still, the system can be driven to exhibit behavior that is similar to a phase transition, for instance by changing the statistical ensemble. Practically, this means that observed phase behavior may be largely dependent on system preparation. The typical signature of this regime is eerily similar to many observations in cells. This naturally leads to a fundamental question: is the traditional phase behavior an appropriate description for cellular behavior?

212-FlashTalk

Quantifying biological variability using appropriate data representations of microscopy images

Ritvik Vasan[1], Julie Theriot[2], Matheus Palhares Viana[1], Susanne M. Rafelski[1].
[1]Allen Institute for Cell Science, Seattle, WA, USA, [2]Department of Biology, University of Washington, Seattle, WA, USA.

The Allen Institute for Cell Science aims to understand the principles by which human induced pluripotent stem cells (hiPSCs) establish and maintain