5TH INTERNATIONAL SYMPOSIUM ON CRYO-EM 3D IMAGE ANALYSIS 2024

March 13 –16, 2024
Granlibakken Conference Center & Lodge
Lake Tahoe, CA USA

The organizing committee acknowledges the generous support of our sponsors:

Special thanks are extended to Dr. Sneka Raveendran for the dedicated efforts in maintaining the website.

Program

WEDNESDAY, MARCH 13TH

3:00 – 6:00 PM	Check in (Hotel Lobby) Meeting registration materials will be provided at check-in **Posters** should be up for the entire meeting (Bay Room).
5:00 – 6:00 PM	Welcome Wine and Cheese (Granhall)
6:00 – 7:30 PM	Dinner (Granhall)
7:30 – 7:45 PM	Welcome Introduction (Mountain Lake) **Dorit Hanein**, Program Chair Professor, Departments of Biological Engineering, of Chemistry and Biochemsitry University of California, Santa Barbara, CA USA
7:45 – 8:45 PM	**Keynote: David Mastronarde** *Computational challenges in electron tomography, past and present* Professor, Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, CO, USA
8:45 – 11:00 PM	Social/Cash Bar (Cedar House)

THURSDAY, MARCH 14TH

7:00 – 8:00 AM Breakfast Buffet (Granhall)

Session 1: 3D Comformational Heterogeneity (Mountain Lake)

8:00 – 8:20AM	Introduction, **José-María Carazo Garcia**, Session Chair Professor, Biocomputing Unit Centro Nacional de Biotecnología, Madrid, SPAIN
8:15 – 8:40 AM	**Steven Ludtke** *Interrogating Systems with Complex Dynamics using Deep Learning GMMs* Professor, Dept. of Biochemistry and Molecular Pharmacology, Director, CryoEM/ET Core Baylor College of Medicine, Houston, TX, USA
8:40 – 9:05 AM	**Slavica Jonic** *Methods based on molecular dynamics simulations for extracting continuous conformational landscapes from single particle images* (MDSPACE) and subtomograms (MDTOMO) Research Director CNRS, IMPMC-UMR 7590 CNRS, Sorbonne University, MNHN, Paris, France
9:05 – 9:30 AM	**Pilar Cossio** *Simulation-based inference to identify conformations in cryo-EM Professor* Center for Computational Mathematics, Flatiron Institute, Simons Foundation, New York USA
9:30 – 9:55 AM	**Amit Singer** *RECOVAR: A Bayesian Framework for Heterogeneity Analysis using Regularized Covariance Estimation* *Professor,* Department of Mathematics, The Program in Applied and Computational Mathematics, Princeton University, Princeton USA
10:00 –10:20 AM	Coffee/Tea Break

Session 2: Protocols for Automation (Mountain Lake)

10:20 –10:35 AM	Introduction, **Peijun Zhang,** Session Chair Professor of structural biology and wellcome trust investigator University of Oxford, UK

10:35 –11:00 AM	**Mark Basham** *Traveling the Long Road to CryoET Automation* Science Director, Artificial Intelligence and Informatics The Rosalind Franklin Institute, Oxfordshire, UK
11:05 - 11:25AM	**Colin Palmer** *Automated data processing with the CCP-EM Pipeliner* Scientist, CCP-EM Project Rutherford Appleton Laboratory, Oxfordshire, UK
11:30 – 11:50AM	**Anchi Cheng** *Data pipeline at CZII* Team Leader Chan Zuckerberg Imaging Institute, San Francisco, USA
11:55AM 12:15PM	**Dimitry Tegunov** Principal Scientist & Group Leader Genentech, South San Francisco, USA
12:15 – 1:15 PM	Lunch Break (Granhall) Afternoon free until Poster Session

Poster Session

4:00 – 6:00 PM PM Poster Session/ Wine & Cheese Reception (Bay Room)

6:00 – 7:30 PM Dinner (Granhall)

Session 3: AI Deep Learning (Mountain Lake)

7:30 – 7:45 PM	Introduction, **Steven Ludtke**, Program & Session Chair Professor, Dept. of Biochemistry and Molecular Pharmacology, Director, CryoEM/ET Core Baylor College of Medicine, Houston, TX, USA
7:45 – 8:05 PM	**Joey Davis** *Visualizing ribosome assembly through cryo-EM, cryo-ET, and deep learning* Associate Professor, Biology Massachusetts Institute of Technology, Cambridge, USA
8:10 – 8:30 PM	**Carlos Oscar Sanchez Sorzano** *Advancements in Deep Learning for Single Particle Analysis through Cryo-electron Microscopy* Professor, Biocomputing Unit Centro Nacional Biotecnología (CSIC), Madrid, Spain
8:35 – 8:55 PM	**Min Xu** *Unsupervised deep learning for subtomogram clustering and alignment* Associate Professor, Computational Biology Carnegie Mellon University, Pittsburgh, USA
9:00 – 9:20 PM	**Hong Zhou** *Deep learning approaches to preferred orientation and missingwedge problems* Professor, Microbiology, Immunology and Molecular Genetics California NanoSystems Institute at University of California, Los Angeles, USA
9:30 – 11:00 PM	Social Hour/Cash bar (Cedar House)

FRIDAY, MARCH 15TH

7:00 – 8:00 AM Breakfast Buffet (Granhall)

Session 4: Data Validation, model building and model quality (Mountain Lake)

8:00 – 8:15 AM	Introduction, **Masahide Kikkawa**, Session Chair Professor, Department of Cell Biology & Anatomy Graduate School of Medicine, The University of Tokyo, Japan
8:20 – 8:40AM	**Koji Yonekura,** *Measurement of charges and chemical bonding in cryo-EM* SPA and 3D ED structures; Professor, RIKEN SPring-8 Center/IMRAM, Tohoku University, Japan
8:45 – 9:05AM	**Tristan Croll,** *Physics-based interactive model building in ISOLDE* Principle Scientist, Altos Lab, Cambridge, United Kingdom
9:10 -9:30 AM	**Daisuke Kihara** *Protein and DNA/RNA Structure Modeling and validation for Cryo-EM Using Deep Learning* Professor, Biological Sciences, Computer Science, Purdue University, USA
9:35-9:55 AM	**Tilman Franke** *An Acquisition and Reconstruction Pipeline with On-The-Fly Quality Assessment for Cryo-Electron Tomography* Product Manager for the Cryo-ET Software Workflow, Thermo Fisher Scientific Planegg, Germany
10:00 – 10:15 AM	Group Photo & Coffee/Tea Break

Session 5: In-situ Feature Annotation and Segmentation (Mountain Lake)

10:20 –10:35 AM	Introduction, **Niels Volkmann**, Session Chair Professor, Departments of Biological Engineering; Electrical and Computer Engineering; and Biomolecular Science and Engineering, University of California, Santa Barbara, CA USA

10:40 – 11:00AM	**Bronwyn Ayla Lucas**, *Methods to characterize macromolecular structure in situ using high-resolution 2D cryo-EM images*. Assistant Professor of Biochemistry, Biophysics and Structural Biology University of California, Berkeley, USA
11:05 –11:25 AM	**Antonio Martinez Sánchez**, *Efficient object detection for in situ quantitative analysis in cryo-electron tomography* Professor in Computer Sciences University of Murcia, Spain
11:20 –11:50 AM	**Nicholas Sauter** *Interactive Machine Learning for Tomogram Segmentation* Computer Senior Scientist, Molecular Biophysics and Integrated Bioimaging LBNL, Berkeley, California, USA
11:55 –12:15 PM	Vendor Talk: **Naoki Hosogi** *Introduction of efficient workflow with a new cryo-FIB/SEM system for cryo-TEM observation* CryoEM Application Expert, Cryo Application Support Group, JEOl Ltd, Akishima, Japan
12:15 – 1:25 PM	Lunch Buffet (Granhall)
	Free afternoon until the poster session.

Poster Session

4:00 – 5:15 PM	Poster Session / Wine & Cheese Reception **(**Bay Room)

Round Table Discussion : Data archival and public data repository policies

5:15 – 6:00 PM

Edward Egelman Moderator
Professor, Biochemistry and Molecular Genetics
Universty of Virginia, School of Medicine
Charlottesville, VA, USA

Roundtable Participants:

Kyle Morris
EMDB team leader
Hinxton, England, United Kingdom

Jason Kaelber
Professor and Director of the RCNF
Rutgers Institute for Quantitative Biomedicine , New Jersey , USA

Adam Marko
Director of Life Sciences, Panasas
San Jose, California USA

6:00 – 7:30 PM

Gala Dinner (Granhall)

Session 6: Short Talks / Poster Lectures (Mountain Lake)

Selected poster talks, 10 min + 5 min for Q&A

7:30 – 9:30 PM

Dorit Hanein, Program & Session Chair
Professor, Departments of Biological Engineering, of Chemistry and Biochemsitry
University of California, Santa Barbara, CA USA

End of Scientific Program

9:30 –11:00 PM

Social Hour / Cash Bar (Cedar House)

SATURDAY, MARCH 16TH

7:30 – 9:00 AM Breakfast Buffet (Granhal)

Check out / Departures

Poster Abstracts

Cryo-EM Structure of Human RECQL5 Helicase Associated with RNA polymerase II Elongation Complex

Alfredo Florez Ariza*, Nick Lue*, Patricia Grob, Ben Kaeser, Susanne Kassube, Eva Nogales‡

RecQ helicases constitute a highly conserved 3' → 5' helicases that are involved in the maintenance of genomic stability. The human genome contains five helicases of the RecQ family: RecQ1 to -Q4 and RECQL5. Mutations in genes encoding these helicases are associated with a variety of syndromes and diseases. RECQL5 has shown to play a crucial role at the interface of the cellular replication, transcription, and recombination. Moreover, *in vitro* experiments have shown that RECQL5 is the only RecQ helicase that can interact directly with Pol II and regulate the transcription throughput by inhibiting both the initiation and the elongation steps. Also, low-resolution structural studies showed that RECQL5 helicase and IRI domains are both important for a stable engagement with the polymerase. However, hitherto, the structural mechanism by which RECQL5 inhibits transcription elongation, remains unknown. In this work, we reveal a high-resolution cryo-EM structure of the RECQL5 helicase bound to an arrested Pol II elongation complex. We observed that RECQL5 IRI domain, engage with the polymerase Rpb1 lower-jaw, mainly through non-polar and H-bonding interactions. Meanwhile, the RECQL5 helicase domain, exhibits a highly dynamic positioning relative to the Pol II. After sorting out the local heterogeneity of this complex, we observed that RECQL5 helicase domain can stablish polar interactions with the Rpb1 upper-jaw region of the polymerase, as well as form extensive H-bonding contacts with phosphate backbone and sugar of several nucleotides in the 3'-extended ssDNA downstream region of the elongation complex. Altogether, we propose that RECQL5 acts as a roadblock, precluding the polymerase of transcribing the mRNA in the 5'→3' sense, and at the same time, occluding the binding site of other transcription factors like TF-IIS, that could resume the transcription elongation.

Dissecting the pathways and mechanisms of drug resistance evolution in HIV

Avik Biswas[1], Indrani Choudhuri[2,3], Allan Haldane[2,4], Zelin Shan[1], Ronald M. Levy[2,3,4], Dmitry Lyumkis[1]

Institution: [1]Laboratory of Genetics, The Salk Institute for Biological Studies, La Jolla, CA, [2]Center for Biophysics and Computational Biology, [3]Department of Chemistry, [4]Department of Physics, Temple University, Philadelphia, PA

Abstract:
In the absence of curative strategies, antiretroviral therapy (ART) presents the primary treatment option against HIV/AIDS. Although, Integrase (IN) Strand Transfer Inhibitors (INSTIs) form an integral part of current first-line ART, the emergence of drug resistance against INSTIs remains a pervasive problem. The mechanisms and pathways through which resistance to INSTIs emerge and eventually get evolutionarily trapped or “entrenched” in the patient population remain poorly understood and form a major stumbling block in the design of better combative strategies. Here, we use recently developed methodology using the Potts protein sequence-covariation model of the fitness landscape combined with kinetic Monte Carlo simulations to first define the temporal ordering of mutations along constrained mutational pathways leading to drug resistance starting from specific, wild-type laboratory molecular clones of the virus, making use of the model’s ability to effectively predict mutational likelihoods in the evolving sequence background. We then rationalize and provide the mechanistic bases for the preferred pathways to drug resistance by ordering the mutant structures derived by high-resolution cryo-electron microscopy along a logical trajectory. We show that the preferred pathways can vary widely between specific molecular clones to within and between HIV patient populations. With this work, we present the initial step towards a comprehensive understanding of the distinct pathways of emergence of patient-derived, clinically relevant complex resistant variants, contrasting them with pathways evolving from specific viral clones used in the laboratory. Probing these pathways by which resistance develops in varied, evolving viral sequence backgrounds will provide a framework to surveil the response of viral systems in general to external selection pressure from therapeutics, rationalize the mechanisms of resistance, and elucidate potential opportunities for therapeutic interventions along preferred pathways.

Optimizing data quality with an event-based direct detector

Benjamin Bammes[1], Erik Anderson[2], Steve Ludtke[2]
[1] Direct Electron LP, USA
[2] Baylor College of Medicine, USA

Advancements in the technology and methods of transmission electron microscopy (TEM) have often required advancements in detector technology to fully realize their potential. The recent introduction of the Apollo camera—an ultra-fast, large-format event-based electron counting direct detector (Bammes et al., 2022)—represents another significant advancement in detector technology. Rather than using image processing to count sparsely-distributed detection events on an integrating direct detection sensor, the Apollo camera performs electron counting in hardware using a novel sparse-binary-readout direct detection sensor, resulting in lower noise and a higher detection rate than has been achieved with conventional direct detection sensors. This results in significantly lower coincidence loss and thus improved data quality and increased throughput for high-resolution single-particle cryo-EM (Peng et al., 2023).

Using a Thermo Fisher Glacios TEM, we have evaluated the data quality from the Apollo camera compared to a conventional direct detector (the Thermo Fisher Scientific Falcon 4 camera) for cryo-tomography. Results show a general trend of improved total signal-to-noise from Apollo, with the difference widening at as the exposure rate increases.

Additionally, we have begun demonstrating and testing the Apollo's new "centroided event streaming" output, which delivers coordinates of detected electrons with <400 μs time resolution. This may be useful to improve motion correction and to enable continuous-rotation tomography with very fine angular sampling (assuming that the microscope stage is sufficiently stable during rotation).

REFERENCES

Bammes B., Fu X., Spilman M., Peng R., & Stagg S. (2022). Apollo: A novel event-based direct detector for cryo-EM. Microscopy and Microanalysis, 28(S1): 1176-1177.

Peng R., Fu X., Mendez J.H., Randolph P.S., Bammes B.E., & Stagg S.M. (2023). Characterizing the resolution and throughput of the Apollo direct electron detector. Journal of Structural Biology: X, 7: 100080.

Advancements in Deep Learning for Single Particle Analysis through Cryo-electron Microscopy

C.O.S. Sorzano, D. Herreros, J.M. Carazo
Biocomputing Unit, Natl. Center of Biotechnology, CSIC

In Structural Biology, Cryo-electron microscopy (cryoEM) has emerged as a pivotal technique for understanding the intricate structures of biological macromolecules at near-atomic resolution. The advent of deep learning has revolutionized the processing of cryoEM images, offering unprecedented accuracy and efficiency. This presentation delves into our contributions to applying deep learning methodologies to enhance single particle analysis (SPA) in cryoEM, underscoring the transformative impact of these technologies.

Deep learning, a subset of machine learning characterized by neural networks with multiple layers, has shown exceptional aptitude in handling large datasets and extracting complex features. This capability is particularly beneficial in cryoEM image processing, where the high volume of data and the subtlety of vital features present considerable challenges. Traditional methods often grapple with issues like low signal-to-noise ratios and structural heterogeneity, which deep learning algorithms can easily accommodate. By learning from vast datasets, these algorithms can effectively identify and interpret the critical features of cryoEM images, leading to more accurate and detailed structural elucidations.

Our research introduces novel deep learning algorithms focused on two pivotal aspects of SPA in cryoEM: 3D image alignment and heterogeneity analysis. The development of these algorithms is driven by the need for more precise and reliable analysis tools capable of handling the nuances of cryoEM data.

Firstly, the 3D image alignment algorithms employ advanced deep-learning techniques to align and average thousands of particle images. This process is crucial for enhancing the signal-to-noise ratio and resolving detailed structures.

Secondly, our heterogeneity analysis algorithms are designed to tackle the challenge of structural variability among particles. This variability can obscure critical details and hinder accurate structure determination. The algorithm employs sophisticated deep-learning models to classify particles into distinct conformational states, facilitating the reconstruction of multiple structures from a heterogeneous sample. This approach enhances the accuracy of the structural models and provides insights into the dynamic range of macromolecular structures.

Papillomavirus cryoEM structures reveal unknown protein density

Caroline Langley[1,2], Dan Goetschius[3], Carol Bator[2], Neil Christensen[4,5,6], Cayce Dorrier[7], Joshua Wang[7], Susan Hafenstein[1,2]

1 Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, Falcon Heights, MN 55108, USA
2 The Hormel Institute, University of Minnesota Medical Research Center, Austin, MN 55912, USA
3 Department of Medicine, The Pennsylvania State University College of Medicine, Hershey, PA 17033, USA
4 Department of Pathology, The Pennsylvania State University College of Medicine, Hershey, PA 17033, USA
5 The Jake Gittlen Laboratories for Cancer Research, Hershey, PA 17033, USA
6 Department of Microbiology and Immunology, The Pennsylvania State University College of Medicine, Hershey, PA 17033, USA
7 VerImmune Inc, 7144 13th PL NW Suite 2200, Washington DC, 20012

Human papillomavirus (HPV) is a significant health burden and leading cause of virus-induced cancers. The HPV capsid is comprised of 360 copies of the major structural protein (L1) arranged in 72 capsomers and an unknown number of the minor structural protein (L2) in a T=7d icosahedral architecture. It has been biochemically determined that each capsid contains approximately 12-36 copies of L2, although the exact amount and distribution of L2 within the capsid is unknown. Determining the structure of the minor capsid protein, L2, has been difficult due to the asymmetric incorporation of L2 and inherent flexibility of the HPV capsid.

An HPV quasivirus capsid composed of L1 and L2 capsid proteins with a packaged cottontail rabbit papillomavirus genome was incubated with heparin. In the resulting 1.9 Å resolution structure, the asymmetric unit of HPV16, containing 6 L1 chains, was able to be built unambiguously. Two protein-like unfilled densities with discernible side chains were identified inside each capsomer, and these possible L2 densities are present asymmetrically.

In a collaboration with VerImmune, a smaller T=1 icosahedral MusPV capsid containing only 60 L1 copies (henceforth Virus-inspired Particles or “ViPs") was vitrified and data were collected for single particle analysis. Icosahedral refinement resulted in a 2.1 Å structure. Similar unknown densities observed in the HPV-heparin map were present in this ViP structure. As L2 is not present in ViPs, the unknown densities cannot be L2. Previously, unfilled densities mapping to the same location were seen in a 3.1 Å HPV quasivirus structure and a 3.3 Å authentic T=7d mouse papillomavirus (MusPV) structure and misidentified as L2. Notably, the viruses used to solve these structures originated from three different sources: E. coli (ViPs), 293TT cells (Quasiviruses), murine tail warts (MusPV).

The presence of two unknown densities in these papillomavirus structures suggests that the unknown proteins are necessary for papillomavirus capsid assembly, maturation, or stability. To identify these unknown proteins, further studies are ongoing, including cross-linking mass spectrometry.

Discovery of a New Binding Partner of the Synaptic Vesicle V-ATPase by Cryo-ET/Cryo-EM

Chuchu Wang

Neuronal synaptic vesicles (SVs) are fundamental components in vertebrate cognition and motor functions, and their effectiveness relies heavily on the membrane proteins and lipids. Although the roles of these components in neurotransmitter uptake and release are well-documented, the in-situ protein and lipid architectures still need to be explored, especially because the pre-synaptic proteins exist in a highly crowded environment. Here, by employing a hybrid of cryo-electron tomography (cryo-ET) and single-particle cryo-electron microscopy (SPA cryo-EM) techniques, we determined different conformational states of vacuolar (H+)-ATPases (V-ATPases) structures at ~4 Å resolution in isolated SVs from mouse brains. These maps, including both intact and partially disassembled V-ATPases, all revealed unknown densities at ~6 Å resolution. Protein mass spectrometry and AlphaFold2-predicted models fitting screening led us to attribute this density to the synaptophysin, the most abundant SV protein. Subsequent synaptophysin knockout experiments in mice confirmed its association with V-ATPase, as the binding density was absent in these altered SVs V-ATPase maps. This interaction potentially influences each SV's V-ATPase copy number, which opens new avenues for understanding synaptophysin function and SV dynamics. Our study highlights the power of combining cryo-ET and SPA cryo-EM for ex vivo high-resolution structural analysis, and also presenting an alternative method for resolving the high-resolution structures of small membrane proteins.

Building and Validating Protein and DNA/RNA Structure Models from medium to low resolution cryo-EM Maps

Daisuke Kihara

Cryo-electron microscopy (cryo-EM) has become one of the main experimental methods for determining biomolecular structures including proteins and nucleic acids. Structure modeling from cryo-EM is challenging when the resolution of maps is around 3 Å or lower, which is often not high enough to specify atom positions. For maps at medium resolution (up to around 5 Å), deep learning can detect characteristic local density features of amino acids and nucleotides, which can be used to guide structure modeling. For maps at medium resolution, we developed DeepMainmast and CryoREAD for protein modeling and DNA/RNA modeling, respectively. In these methods, following local structure detection with deep learning, full structure is built through intuitive optimization procedure that mimics manual modeling. Local density features detected by deep learning can also be used for validating existing protein structure models in PDB. The protein model quality assessment score, DAQ, works with that idea and detects potential errors in the model. In a large-scale analysis of protein models from cryo-EM, we found that a substantial small number of models may have some errors. We will also discuss our most recent method, DiffModeler, which is for structure fitting of Alphafold models in a lower resolution of up to ~15 Å. All the tools we developed are available at https://kiharalab.org/emsuites/ and https://em.kiharalab.org/

Automated processing pipelines at cryoEM facilities

Daniel Hatton

The UK national cryoEM facility, eBIC (Clare *et al*., 2017), offers users live processing for both SPA and cryoET sessions. For SPA this extends to batched 2D classification and attempts at 3D classification on a subset of particles selected based on 2D classification images using the RELION class ranker tool (Kimanius *et al*., 2021), and for tomography tomograms are reconstructed using AreTomo (Zheng *et al*., 2022) and denoised with Topaz (Bepler *et al*., 2020). Users do not need to provide any parameters with the majority of those required extracted from acquisition metadata allowing processing to be triggered within seconds of acquisition including after grid changes. The extracted metadata is stored as a record for the users and the gathering of usage statistics for analysis. Processed data is archived to tape following the completion of a session. Modern cloud technologies are leveraged to allow a scalable system which can be adapted to utilise a wide variety of cryoEM software while maintaining facility information management system integration (Delagenière *et al*., 2011). We deploy a service based system which receives processing requests at the earliest possible. This has greatly increased the responsiveness of the processing pipelines. The system currently supports the academic user program of eBIC across four Krios microscopes and two 200 keV systems.

Helical classification with deep learning

Daoyi Li

Many macromolecules in biological systems exist in the form of helical polymers. However, the inherent polymorphism and heterogeneity of samples complicate the reconstruction of helical polymers from cryo-EM images. Currently, available 2D classification methods are effective at separating particles of interest from contaminants, but they do not effectively differentiate between polymorphs, resulting in heterogeneity in the 2D classes. As such, it is crucial to develop a method that can computationally divide a dataset of polymorphic helical structures into homogenous subsets. In this work, we utilized deep-learning language models to embed the filaments as vectors in hyperspace and group them into clusters. Tests with both simulated and experimental datasets have demonstrated that our method – HLM (**H**elical classification with **L**anguage **M**odel) can effectively distinguish different types of filaments, in the presence of many contaminants and low signal-to-noise ratios. We also demonstrate that HLM can isolate homogeneous subsets of particles from a publicly available dataset, resulting in the discovery of a previously unreported filament variant with an extra density around the tau filaments.

Despite the HLM demonstrating efficacy in distinguishing different types of helical structures, its performance is dependent upon the quality of 2D classification. Thus, our current research efforts are directed towards the development of the image encoder designed to encode helical image. This approach has shown promise in addressing certain limitations inherent in the previous pipeline.

Advanced deep learning algorithms for enhanced cryoEM image analysis and molecular variability elucidation

D:Herreros, C.O.S. Sorzano, J.M. Carazo
Biocomputing Unit, Natl. Center of Biotechnology, CSIC

Cryo-electron microscopy (cryoEM) has emerged as an integral technique for deciphering the complexities of biological macromolecule structures at near-atomic resolution. The fusion of advanced image processing techniques with substantial data volumes facilitates the discernment of subtle, yet crucial, biomolecular features. However, traditional cryoEM methodologies frequently grapple with challenges such as low signal-tonoise ratios and the presence of structural heterogeneity in data. To surmount these inherent limitations of classical cryoEM methods, there has been a paradigm shift in cryoEM image analysis, prominently featuring the incorporation of deep learning into image processing frameworks. Deep learning, a sophisticated branch of machine learning known for its multilayered neural networks, is particularly adept at handling extensive datasets and extracting complex, nuanced features. Given the necessity to process large data volumes and the imperative to accurately identify subtle biomolecular features, the application of deep learning in cryoEM has become increasingly pivotal. In this context, our research presents a comprehensive suite of cutting-edge deep learning algorithms, meticulously designed to address two critical dimensions of single particle analysis (SPA) in cryoEM: three-dimensional image alignment and the analysis of structural heterogeneity. Our image alignment tools harness advanced deep learning techniques specifically for the determination of 3D image alignments, encompassing shifts and rotations. This approach is instrumental in uncovering detailed structural insights. Concurrently, we have engineered sophisticated deep learning algorithms for heterogeneity analysis. These algorithms are adept at segregating particles into distinct 3D conformational states. Such classification is crucial for approximating experimental conformational landscapes, thereby illuminating the dynamic range and diversity inherent in macromolecular structures.

Cryo ET analysis reveals location of Human Papillomavirus minor capsid protein L2

Ebere Precious Orji[1,2], Amanda Gramm[3], Caroline Langley[1,2], Kyle Messina[2], Matthew T. Swulius[4], Neil Christensen[4,5,6], Susan L. Hafenstein[1,2]

AFFILIATIONS:
[1]Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Twin Cities, Minneapolis, Mn, 55455, USA
[2]The Hormel Institute, University of Minnesota, Austin, Mn, 55912, USA
[3]Department of Biochemistry, Microbiology, and Molecular Biology, The Pennsylvania State University, State College, PA, 16801, USA
[4]Department of Pathology, The Pennsylvania State University College of Medicine, Hershey, PA, 17033, USA
[5]The Jake Gittlen Laboratories for Cancer Research, Hershey, PA, 17033, USA
[6]Department of Microbiology and Immunology, The Pennsylvania State University College of Medicine, Hershey, PA, 17033, USA

Human papillomaviruses are non-enveloped, double-stranded DNA viruses responsible for 5% of all cancers. HPV encodes two structural proteins, the major capsid protein L1 and the minor capsid protein L2, arranged in a T=7d icosahedral structure. Although L1 is sufficient for capsid formation, L1-only capsids are not infectious. Past studies revealed that L2 plays a significant role in virus entry, replication, and assembly. Yet very little is known about the structure, location, and copy number of L2, although it is known that the N-terminus is located outside the capsid. Determining the structure of L2 has been difficult due to its flexibility and asymmetric incorporation as well as the flexibility of the HPV capsid. In this study, we used gold-labeled anti-L2 antibodies to probe the location of L2 in HPV16 quasivirus. Negative stain TEM and cryo-electron tomography identified a specific region of L2 outside the capsid and confirmed that L2 is flexibly extruded to the outside of the capsid. Additionally, results from hydrogen-deuterium exchange mass spectrometry show that both the N and C terminals of L2 are extruded from the capsid shell. Ongoing limited proteolysis-mass spec experiments will complement the cryo-EM to elucidate the presence and location of L2 on the capsid. Next, we plan to use other anti-L2 antibodies that recognize different regions of L2 to determine where L2 exits the capsid. Furthermore, cryo-EM single particle analysis will aid the determination of L2 structure. This innovative mapping project will significantly advance our understanding of HPV and expedite the development of next-generation HPV antivirals and vaccines.

High-Throughput Single Particle Analysis using the JEOL CRYO ARM

Emmanuel Smith

Structural studies of protein complexes that orchestrate dynamic cellular processes are often limited due to their inherent flexibility and dynamic nature. Consequently, SPA workflows that allow efficient screening, processing, and data collection are becoming increasingly vital, especially in attempts to solve important, yet challenging structures. Here we present the JEOL CRYO ARMs, JEOL's dedicated 200kV and 300kV cryo-EM Transmission Electron Microscopes, and we demonstrate how the CRYO ARM can be used for high-throughput analysis of protein samples leading to high-resolution structural determination.

Structural and Computational Insights Drive Potent MTA-Synergic PRMT5 Inhibitor Design Using Cryo-EM

Gaya P. Yadav[1,3,4,5], Wei Zhou[2], Xiaozhi Yang[2], Feng Qin[3], Chenglong Li[2] & Qiu-Xing Jiang[1,3,4]

[1]Department of Microbiology and Cell Science, University of Florida, Gainesville, FL 32611, USA. [2]Department of Medicinal Chemistry, College of Pharmacy, University of Florida, Gainesville, FL 32610, USA. [3]Department of Physiology and Biophysics, the State University of New York at Buffalo, Buffalo, NY 14214, USA. [4]Laboratory of Molecular Physiology and Biophysics, Hauptman-Woodward Medical Research Institute, Buffalo, NY 14203, USA. [5]Present address: G.P.Y at the Department of Biochemistry and Biophysics, Texas A &M University, College Station, TX 77843, USA.

The "resolution revolution" in cryo-electron microscopy (cryo-EM) has transformed structural biology, yielding biomolecular structures at resolutions of 2.5–4.5 Å [1-3]. While cryo-EM often lags behind crystallography, recent breakthroughs in achieving atomic resolutions present opportunities for structure-based drug design (SBDD) [4]. Computational advancements in ligand configuration analysis further enhance cryo-EM SBDD [5].

PRMT5, pivotal in epigenetic processes and overexpressed in various cancers, is an attractive anti-cancer target [6-8]. In particular, cancer cells lacking 5′-methylthioadenosine phosphorylase (MTAP) exhibit sensitivity to partial PRMT5 inhibition, offering a targeted therapeutic approach. Current PRMT5 inhibitors fall into distinct classes, with challenges in achieving specificity and synergy with 5′-methylthioadenosine (MTA), a key metabolite [9-10].

Despite structural data limitations, including the elusive apo human complex, cryo-EM structures at 3.1 Å resolution illuminate PRMT5's catalytic mechanism. The MTA-synergic inhibitor 11-2 F induces positive cooperativity by shifting the cofactor-binding pocket, as revealed in cryo-EM structures. Computational analysis predicts subtype specificity among PRMTs, leading to the identification of three analogs with enhanced potency. One synthesized analog proves four times more efficient than 11-2 F in inhibiting PRMT5 catalysis, highlighting the potential of cryo-EM structures and computational analysis in drug design.

In conclusion, the combination of cryo-EM's resolution capabilities and computational tools unlocks avenues for precision drug design. The focus on PRMT5 in cancer therapy exemplifies the promise of this approach, offering hope for more effective and targeted treatments in the evolving landscape of precision medicine.

Structural studies of the interaction between Ctf18-RFC clamp loader and PCNA *via* cryo-EM.

Giuseppina Briola

In DNA replication, clamp loaders are vital for loading the PCNA (Proliferating Cell Nuclear Antigen) sliding clamp onto double-stranded DNA, utilizing ATP binding and hydrolysis. Once loaded, PCNA binds to DNA polymerases, significantly enhancing their processivity. The Ctf18-RFC (chromatin transmission fidelity 18-replication factor C) clamp loader, specialized for PCNA loading and DNA polymerase epsilon interaction, has a molecular weight of 324 kDa. It comprises two main sections linked by a flexible linker: the catalytic RFC module (including 5 subunits, RFC1-5) essential for PCNA loading, and the Ctf18-8-1 module, crucial for binding to DNA polymerase epsilon.

Using cryogenic electron microscopy (cryo-EM), we captured the first structure of the human Ctf18-RFC module in a complex with PCNA at a resolution of 2.9 Å. This complex is engaged with a closed PCNA ring, where RFC1, RFC2, and RFC3 directly interact with PCNA, whereas RFC4 and RFC5 are positioned further away. This structure reveals an autoinhibited conformation of the PCNA-catalytic RFC complex, indicating a sophisticated regulatory mechanism within DNA replication. The significance of this autoinhibited state and the specific interactions it entails are the focus of ongoing research, aiming to enhance our understanding of the complexities of human DNA replication.

Two distinct archaeal type IV pili structures formed by the same pilin

Gunnar Eastep

To survive and thrive in their environments bacteria and archaea depend on protein filaments, oligomeric chains of hundreds to thousands of protein molecules. Type IV pili are a ubiquitous class of protein filaments that are utilized by infectious bacteria for diverse functions. For instance, *pseudomonas aeruginosa* use type IV pili to move through host cells whereas uropathogenic *escheria coli* express filaments that allow them to adhere to uroepithelial cells and resist being swept away by the flow of urine. Type IV pili and the protein machinery they depend on for secretion are an ancient, evolutionarily conserved system amongst archaea and bacteria. Because of the many notable differences between pili in the two domains (perhaps the most striking being that archaeal flagella and type IV pili have very similar structures, while bacteria use a wholly separate type of flagella), the comparative study of archaeal and bacterial pili may reveal fundamental requirements for pili formation.
Here, we present the structures of filaments formed from a pilin protein of the archaeon *Sulfolobus islandicus*, determined using cryo-electron microscopy (cryo-EM). Pilin proteins contain an *N*-terminal a-helical domain that serves as the hydrophobic core of the filament by interacting with the helical domains of other pilin copies. They also contain a *C*-terminal domain which faces towards the outside of the filament and dictates how the filament interacts with its environment, such as by binding specific surfaces. The first filament identified in our study, the *S. sulfobolus* mono-pilus, was previously found to be produced by archaea under well-fed conditions. However, under conditions of nutrient-deprivation we observed that archaea make a pilus composed of three intertwined, left-handed helical filaments, referred to as the tri-pilus. Strikingly, we found that those two distinct helical polymers, mono and tri-pili, were formed by the same pilin protein of *S. islandicus*. In these filaments, the pilin's *N*-terminal and *C*-terminal domains took on markedly different orientations. Genetic knockout of *S. islandicus* secretion machinery (TadC and ATPase proteins) revealed that the same secretion system produces both filaments. Additionally, structural analysis of cryo-EM density maps demonstrated that the protein bears similar glycosylation in both the mono and tri-pilus, suggesting that alternative post-translational modifications are not the basis of tri-pilus expression.

To our knowledge, our study represents the first case of a pilin possessing the capacity to give rise to two structurally-distinct type IV pili. Conventional wisdom suggests that pilins must take on a specific conformation in the secretion machinery during assembly – our findings complicate this notion by showing the same secretion system produces the mono and tri-pilus. Further work is needed to understand the mechanisms by which the two filaments are assembled, including the possibility of unidentified secretion components or intrinsic promiscuity of the secretion system. The function of the pili is also of great interest given their expression by *S. islandicus* only under conditions of stress. Studies concurrent with ours suggest a potential function of twitching motility for the tri-pilus, perhaps allowing the archaeon to escape an inhospitable environment.

Cryo-EM structure of human DICER in a cleavage-ready state

Hansol Lee[1,2*], Young-Yoon Lee[2,3], Haedong Kim[2,3], Vic Narry Kim[2,3], and Soung-Hun Roh[1,2]

[1] Institute of Molecular Biology and Genetics, Seoul National University, Seoul 08826, Republic of Korea;
[2] School of Biological Sciences, Seoul National University, Seoul 08826, Republic of Korea;
[3] Center for RNA Research, Institute for Basic Science (IBS), Seoul 08826, Republic of Korea

RNA interference (RNAi) is a vital gene regulation pathway that relies on small regulatory RNAs targeting messenger RNAs (mRNAs). DICER, an essential ribonuclease III protein with multiple domains, plays a central role in RNAi by processing precursor molecules into microRNAs (miRNAs) and small interfering RNAs (siRNAs). However, the lack of a high-resolution structure of vertebrate DICER in an RNA-cleavage competent state has limited our understanding of small RNA biogenesis mechanisms. In this study, we utilized cryo-electron microscopy (cryo-EM) to determine the structures of human DICER in its apo state and in complex with pre-miRNA, revealing a conformational state primed for cleavage and providing insights into microRNA processing.

During the dicing state, we observed conformational flexibility in the N-terminal helicase domain, which likely facilitates the binding of pre-miRNA near the catalytic sites. Subsequent rearrangement allows the C-terminal double-stranded RNA binding domain (dsRBD) to recognize the substrate, exhibiting both sequence-independent and sequence-specific interactions. Notably, we identified arginine side chains within the 5'-pocket that establish critical interactions with the 5' end base, thereby determining the cleavage site. Furthermore, our investigations revealed cancer-associated mutations in the 5'-pocket residues, which impair the activity of human DICER in vitro and disrupt miRNA biogenesis in cells. These findings deepen our understanding of how DICER achieves high specificity in recognizing pre-miRNAs and shed light on the mechanistic underpinnings of DICER-related diseases.

By elucidating the structural basis and functional implications of DICER-mediated miRNA processing, this study offers valuable insights into the molecular mechanisms underlying RNAi regulation and provides a foundation for future therapeutic approaches targeting DICER-related disorders.

Substrate Inhibition of a Polyketide Synthase

Heidi L. Schubert[1], Feng Li[2], Christopher P. Hill[1], Eric W. Schmidt[2]

Departments of Biochemistry[1] and Medicinal Chemistry[2], University of Utah

Animal fatty acid like polyketide synthase (AFPK) bridges the evolution gap between fatty acid synthase (FAS) and polyketide synthase (PKS), important to animal lipid metabolism. *Elysia chlorotica* PKS1 and PKS2, share a high sequence identity with EcFAS, yet synthesize polypropionates by condensing unusual animal primary substates while using methylmalonyl-CoA as the extender unit. Here we present the full structures of apo-EcPKS1 and EcPKS2 bound to malonyl-CoA and NADPH at the AT and KR domain, respectively. Initial steps of PKS assembly are further revealed as the ACP-pPANT is docked at the KS domain in the EcPKS2 structure. Remarkably, one entire linker between the KR and ACP domain is visible revealing interactions with the MT° domain. Biochemical experiments designed to disrupt the ACP linker interactions demonstrate how the ACP linker collaborates with the catalytic domains to regulate assembly products. Nominal rigidity of the linker enables flexibility analysis of the two domain mega-enzyme while the domains within the related apo-structure occupy a much larger conformational space. The discovery and characterization of the ACP linker are significant to understanding how FAS and type I iterative PKSs select substrates and regulate product formation.

Enabling Automation in SPA Data Acquisition with Smart EPU

Fanis Grollios[1], **Holger Kohr**[1], Alexandru Pustianu[1]
1Thermo Fisher Scientific, Materials & Structural Analysis, Eindhoven, the Netherlands

Single Particle Analysis (SPA) is an increasingly popular technique among structural biologists for visualizing the structures of proteins and protein complexes. Recent improvements in instrumentation and automation improve the efficiency and of high-quality data collection, which will broaden the applicability of this workflow to users that have limited microscopy expertise [1]. However, acquiring high-quality SPA data still largely depends on the expertise of the microscope operator. This limits the speed of adoption as researchers must invest considerable time into understanding microscopy and the technicalities of the workflow. Thermo Scientific introduced Smart EPU [2] as a software platform for workflow automation and customization, thereby lowering the entry barrier for novice or non-expert users while providing customizability to advanced users. Smart EPU includes a set of built-in smart plugins integrated into the EPU workflow that are designed to free operators from tedious and repetitive tasks like acquisition area selection, removal of unsuitable areas, etc. Automating these steps helps to reduce the chance of user error and improve the productivity of microscopes.

In this poster, we present the current state of Smart EPU, including the latest additions and improvements. Smart EPU now provides additional automation for screening, the workflow where manual interactions in session setup and execution traditionally present the most severe bottlenecks. As a prominent example, Smart EPU features a new AI-based automatic grid square selection plugin that recognizes large-scale defects at atlas magnification and removes unsuitable grid squares from the selection. Furthermore, the algorithm groups grid squares into appropriate categories for screening and automatically selects the best squares based on a user-provided number. This new plugin is intended to remove the need for the user to intervene in the selection of grid squares. Further AI automation features like foil hole detection and selection for fast screening will be discussed as well.

Apart from built-in plugins, Smart EPU provides an Open API that enables community members to connect custom algorithms to an ongoing acquisition and optimize it in terms of efficiency and quality by leveraging, in real time, the vast amounts of data and metadata produced by the microscope. This automation platform ensures that decisions made by algorithms are applied in a timely, conflict-free, and safe manner.

[1] Single-particle cryo-EM at atomic resolution, Nakane, T., Kotecha, A., Sente, A. et al. Nature 587, 152–156 (2020)
[2] Smart EPU: SPA Getting Intelligent, Yuchen D, Fanis G, Holger K et al. Microsc. Microanal. 27 (Suppl 1), 2021

TomoNet: A streamlined cryoET software pipeline with automatic particle picking on flexible lattices

Hui Wang1,2,3,*, Shiqing Liao2,3, Xinye Yu3, Jiayan Zhang2,3, and Z. Hong Zhou1,2,3,*
1Department of Bioengineering, University of California, Los Angeles (UCLA), Los Angeles, CA 90095, USA
2California NanoSystems Institute, UCLA, Los Angeles, CA 90095, USA
3Department of Microbiology, Immunology, and Molecular Genetics, UCLA, Los Angeles, CA 90095, USA
*Corresponding author: Z. Hong Zhou (Hong.Zhou@UCLA.edu, 1-310-694-7527)

Cryogenic electron tomography (cryoET) is capable of determining *in situ* biological structures of molecular complexes at near atomic resolution by averaging half a million subtomograms. While abundant complexes/particles often clustered in arrays, precisely locating and seamlessly averaging such particles across many tomograms present major challenges. Here, we developed TomoNet, a software package with a modern graphical user interface to carry out the entire pipeline of cryoET and subtomogram averaging to achieve high resolution. TomoNet features built-in automatic particle picking and 3D classification functions and integrates commonly used packages to streamline high-resolution subtomogram averaging for structures in one-, two- or three-dimensional arrays. Automatic particle picking is accomplished in two complementary ways: one based on template matching and the other employing deep learning. TomoNet's hierarchical file organization and visual display facilitate efficient data management as required for large cryoET datasets. Applications of TomoNet to three types of cryoET data demonstrate its capability of efficient and accurate particle picking on flexible and imperfect lattices to obtain high-resolution 3D biological structures: virus-like particles, bacterial surface layers within cellular lamellae, and membranes decorated with nuclear egress protein complexes. These results demonstrate TomoNet's potential for broad applications to various cryoET projects targeting high-resolution in situ structures.
Cryogenic electron tomography (cryoET) is capable of determining *in situ* biological structures of molecular complexes at near atomic resolution by averaging half a million subtomograms. While abundant complexes/particles often clustered in arrays, precisely locating and seamlessly averaging such particles across many tomograms present major challenges. Here, we developed TomoNet, a software package with a modern graphical user interface to carry out the entire pipeline of cryoET and subtomogram averaging to achieve high resolution. TomoNet features built-in automatic particle picking and 3D classification functions and integrates commonly used packages to streamline high-resolution subtomogram averaging for structures in one-, two- or three-dimensional arrays. Automatic particle picking is accomplished in two complementary ways: one based on template matching and the other employing deep learning. TomoNet's hierarchical file organization and visual display facilitate efficient data management as required for large cryoET datasets. Applications of TomoNet to three types of cryoET data demonstrate its capability of efficient and accurate particle picking on flexible and imperfect lattices to obtain high-resolution 3D biological structures: virus-like particles, bacterial surface layers within cellular lamellae, and membranes decorated with nuclear egress protein complexes. These results demonstrate TomoNet's potential for broad applications to various cryoET projects targeting high-resolution in situ structures.

Authentic B19 virions purified from human blood have active host proteins bound

Hyunwook Lee[1,2*], Ruben Assaraf[3*], Suriyasri Subramanian[4], Dan Goetschius[4], Jan Bieri[3], Nadia M. DiNunno[4], Remo Leisi[3], Carol M. Bator[1,2], Susan L. Hafenstein[1,2,4,5#], Carlos Ros[3#]

1The Pennsylvania State University, University Park, PA 16802, USA
2The Hormel Institute, University of Minnesota, Austin, MN 55912, USA
3Department of Chemistry, Biochemistry and Pharmaceutical Sciences, University of Bern, Switzerland
4The Pennsylvania State College of Medicine, Hershey, PA 17033 USA
5 University of Minnesota, Department of Biochemistry, Biophysics and Molecular Biology, Minneapolis MN 55455 USA
* These authors contributed equally
# Shared corresponding authors

Abstract (192 words)

Parvovirus B19 is a human pathogen causing a broad range of syndromes in adults varying in severity. The lack of a permissive cell culture system has limited high-resolution structural information of B19 to noninfectious virus-like particles (VLP). We have solved the first near-atomic, 2.6 Å resolution structure of authentic B19 purified from de-identified patient blood samples. There are distinct differences compared to noninfectious VLPs. However, more strikingly, associated with the capsids in all patient samples tested, we identified two distinct non-virus densities consistent with capsid-bound proteins. Local reconstruction techniques and mass spectrometry identified the bound human proteins as inter-alpha-trypsin inhibitor heavy chain 4 (ITIH4) and serpinA3. These two protease inhibitors are acute phase reactants that contribute to the host defense and other adaptive capabilities. ITIH4 binds at the icosahedral five-fold axis of the capsid and serpinA3 occupies the two-fold axis, saturating the surface and modifying the antigenic and physical properties of the virus. Protein-decorated virions remain infectious in cell culture and the capsid-associated serpins maintain protease inhibition activity. These findings suggest that virion binding of the acute phase reactants is beneficial for the virus in the context of a systemic infection.

Uncovering Biofilm Components in *Burkholderia cepacia* Complex Species Associated with Cystic Fibrosis

Jessie Lynda Fields[1,2,4], Caitlyn Sebastian[3,4], Zhangli Su[2,5], Megan Kiedrowski[3,4], Fengbin Wang[1,2,4]

[1] Department of Biochemistry and Molecular Genetics, Heersink School of Medicine, University of Alabama at Birmingham, Birmingham, Alabama, USA.
[2] O'Neal Comprehensive Cancer Center, Heersink School of Medicine, University of Alabama at Birmingham, Birmingham, Alabama, USA.
[3] Division of Pulmonary, Allergy, and Critical Care, Department of Medicine, Heersink School of Medicine, The University of Alabama at Birmingham, Birmingham, Alabama, USA
[4] Gregory Fleming James Cystic Fibrosis Center, The University of Alabama at Birmingham, Birmingham, Alabama, USA
[5] Genetics Research Division, Department of Genetics, Heersink School of Medicine, University of Alabama, Birmingham, Alabama, USA

Cystic Fibrosis (CF) is a common life-shortening genetic disorder brought on by a malfunction in the cystic fibrosis transmembrane conductance regulator (CFTR) protein, often leading to chronic lung infections. The *Burkholderia* genus, a complex array of Gram-negative bacteria, is of interest due to its inherent drug resistance and its pathogenic role in immunocompromised individuals, particularly in people with CF (pwCF). Within this genus is the *Burkholderia cepacia* complex (BCC), which contains species of *Burkholderia* frequently implicated in severe CF infections, including cepacia syndrome which is fatal for 75% of pwCF who contract it. BCC strains rely on biofilm formation for their virulence; this involves the production of large macromolecular assemblies such as extracellular filaments. This is done under varied conditions, as synthesizing these large assemblies is metabolically taxing. Nearly no pili have been studied in CF pathogens, presenting a knowledge gap in both the structure of these filaments and their role in host-pathogen interactions. Cryo-electron microscopy (EM) can be used to identify unknown, unpurified, protein structures within uncharacterized samples. This is significant as extracellular filaments are often composed of protein monomers and difficult to purify, which is needed in typical protein identification and structural determination. From a pure culture, we were able to identify the chaperon-usher pathway (CUP) pili of *B. contaminans*, a species of BCC seen in infections of pwCF, by leveraging recent advancements in cryo-EM and its semi-sequencing capabilities. A 3.2Å resolution structure of the *B. contaminans* CUP pili was achieved. CUP pili function in attachment to host surfaces and withstanding host generated shear mechanical forces, making it essential in evading the host pathogen clearing mechanisms. Further studies will be done in human cell culture to determine how deletions of genes encoding CUP pili in *B. contaminans* affect infectivity; preliminary data indicates that the formation of pili is a dynamic process in BCC species. Additionally, the impact BCC species CUP pili on biofilm formation will be studied to provide the context of chronic BCC infections. Further, cryo-tomography will be used to study large scale interactions between CUP pili and host cells to understand the role of CUP pili in BCC pathogenicity in pwCF. Understanding the structure and interactions of CUP pili produced by BCC species in the context of chronic CF infections can lead to the development of novel therapeutic targets.

THE PALISADE LAYER OF THE POXVIRUS CORE IS COMPOSED OF FLEXIBLE A10-TRIMERS

Jiasui Liu, Simon Corroyer-Dulmont, Vojtěch Pražák, Iskander Khusainov, Karola Bahrami, Sonja Welsch, Daven Vasishtan, Agnieszka Obarska-Kosińska, Sigurdur R. Thorkelsson, Kay Grunewald, Emmanuelle R. J. Quemin, Beata Turoňová, Jacomina Krijnse Locker

Although vaccinia virus (VACV) is the best studied poxvirus, the structure of the mature virion (MV) remains poorly understood. The asymmetric shape, size and compactness of the VACV mature virions cause a huge challenge for electron microscopy (EM) analysis, including cryo electron tomography. We thus analyzed the cores obtained by detergent-stripping of MVs to facilitate the analysis. We focused on the most prominent part of the core wall called palisade layer. We devised a new particle identification workflow which incorporates manual core segmentation and AlphaFold2 prediction to reliable obtain positions of the stakes that constitutes the palisade layer. We identified the stakes as trimers of the major core protein A10. The trimers do not form any regular arrangement and their classification indicate structural flexibility. To confirm physiological relevance of our findings, we compared the *in vitro* cores to the *in situ* ones, obtained from cellular cytoplasm at the early stages of infection. We show that the A10 trimers are organized in a similar manner. Altogether, our study identifies for the first time the structure and molecular composition of the palisade units and reveal more details on the arrangement of the whole core wall

EMhub: a web platform for data management and on-the-fly processing in scientific facilities

José Miguel De la Rosa Trevín (1) , Grigory Sharov (2) , Stefan Fleischman (3), Dustin Morado (3), John Bollinger (1) , Darcie Miller (1) , Daniel Terry (1) , Scott Blanchard (1) , Israel Fernández (4) , Marta Carroni (3)
(1) St. Jude Children's Research Hospital, Memphis, Tennessee, USA.
(2) Laboratory of Molecular Biology, Cambridge, United Kingdom
(3) Science for Life Laboratory, Stockholm, Sweden
(4) Instituto Biofisika, Basque Country, Spain

Scientific facilities generate vast amounts of data, not only large but also diverse and at a fast speed. Additional challenges are posed by the management of users, instruments, report generation, and invoicing. Here we present EMhub, a web platform designed to address these challenges by supporting daily operations of a scientific facility. It allows manipulation of users, instruments, bookings, and project tracking in a simple way. The application was originally inspired by the needs of a CryoEM facility, but it has been general enough that its implementation is extended to other domains. EMhub is implemented in a modular way, allowing the extension of the core functionality. Moreover, different external processes can be connected to the application via a REST API to automate various tasks (e.g. folder creation, user and password generation, etc) and for monitoring on-the-fly data processing independently of the software used. EMhub has been in use in the last few years at the Swedish National CryoEM Facility and more recently also installed in the CryoEM center at the Structural Biology Department at St. Jude Children's Research Hospital. There, a fully automated single particle pipeline has been implemented for on-the-fly data processing and analysis. At St. Jude, the Xray Crystallography Center and the Single-Molecule Imaging Center have already extended the platform to support their data management workflows.

A Better Metric for Template Matching in Cryo-EM Cellular Images

Kexin Zhang

Accurately placing macromolecular assemblies in the cellular context is an important step in understanding their mechanistic role inside the cell. Previously, our lab developed a 2D template matching (2DTM) approach to detect targets in cellular cryo-EM images with high positional and orientational accuracy. In 2DTM, we calculate a 2DTM z-score for every location in the cryo-EM image that depends on the agreement between the template and the image. A target is detected when the 2DTM z-score exceeds a statistically defined threshold. 2DTM offers a way to study macromolecular assemblies in a broader context of a cell, taking advantage of the increasing number of available high-resolution structures as templates.

Building on the success in locating and discriminating larger molecular species in cells, we seek to improve 2DTM to detect more challenging targets, such as aspherical and relatively small complexes. In our current work, we have developed a novel 2DTM metric, the 2DTM p-value, by considering cross-correlations between the template and the image across all frequency ranges and building a probabilistic target detection model via 2DTM in cellular cryo-EM images. We chose an aspherical and relatively small protein (clathrin) and simulated images with various solvent thicknesses and different types of structural noise. The 2DTM p-value detects more clathrin particles than the 2DTM z-score. We also show that the 2DTM p-value could robustly recover mature 60S particles in the yeast lamellae dataset when strong Gaussian noise is introduced to the image.

Structural Studies of Picornavirus Entry Complexes

Santosh Panthi[1,2], **Kyle J Messina[3*]**, Carol M. Bator[3], Susan L. Hafenstein[3,4]
[1]Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA
[2]The Huck Institute of the Life Sciences, The Pennsylvania State University, University Park, PA
[3]The Hormel Institute, University of Minnesota, Medical Research Center, Austin, MN 55912
[4]Department of Biochemistry, Molecular Biology and Biophysics, University of Minnesota, Minneapolis, MN 55455
*Presenting author

It has long been established that virus host entry is mediated by coxsackie-adenovirus receptor (CAR) recognition for the picornavirus, coxsackievirus B3 (CVB3)[1-3]. It is known that CAR binds into a specific cleft in the CVB3 capsid and that at physiological temperatures this binding event triggers both asymmetric and global conformational changes leading ultimately to release of the genome[3-4]. We have identified a specific strain of CVB3, called CVB3-RD (rhabdomyosarcoma), that is not triggered under typical conditions with CAR bound. Comparison between CVB3-RD and CVB-3RD-CAR complexes have revealed minor structural changes that give additional insights into the conformational triggers that set off the entry cascade.

References:

1. C. J. Cohen et al., *Proc. Natl. Acad. Sci.*, 98(26), 15191-15196 (2001).
2. Y. He et al., *Nature structural biology*, 8(10), 874-878 (2001).
3. L. J. Organtini et al., *Journal of virology*, 88(10), 5755-5765 (2014).
4. H. Lee et al., *Science Advances*, 2(8), e1501929 (2016).

Enhancement of Signal-to-Noise Ratio (SNR) via Distortion Modeling of Movie Frames with the Update Unblur Program

Lingli Kong,[1] Johannes Elferich,[1,2] Nikolaus Grigorieff[1,2]
[1]University of Massachusetts Chan Medical School, [2]Howard Hughes Medical Institute

Due to the nature of the specimen and the imaging method, radiation damage and beaminduced motion will lead to different distortions in the sample, which is one of the main obstacles to obtaining high-resolution tomograms. To solve this problem, a model that describes the distortion can be implemented to represent the motion pattern and correct the local misalignment. Since the distorted specimen can present a variety of local movements, high flexibility should be one of the most important features of the model to deal with the complexity of the distortion behavior. However, the low SNR of the movie frames will end up feeding the model with noisy patterns and cause overfitting problems at the same time. Thus, a balance between the model flexibility and overfitting should be achieved. In this work, we propose to implement a cubic B-spline model to describe the distortion behaviors and incorporate the Akaike Information Criterion to regulate the model's flexibility to avoid the overfitting problem. The method was embedded in the new version of the Unblur program in our cisTEM software.
The performance of the model was evaluated by examining the SNR of the detected target particles using template matching. The result shows that the new version of Unblur outperformed other models by higher SNR values, which indicates a more accurate local distortion correction and higher quality of the produced image.

Capturing peptide transporters in action

Maxime Killer

The uptake of peptides plays a crucial role in metabolism and inflammation. In mammals, peptide absorption and reabsorption are mediated by the proton coupled oligopeptide transporters (POTs) PepT1(SLC15A1) and PepT2 (SLC15A2), of the solute carrier family 15. POTs are one the most promiscuous transporters among solute carriers and constitute the main route of entry for orally administrated peptidomimetic drugs. SLC15 transporters are involved in various inflammatory diseases, and the paralogue PHT1 (SLC15A4) was recently identified as therapeutic target in Systemic Lupus Erythematosus (SLE). The three dimensional structures of several bacterial homologues have been determined in the past 10 years, but how these shuttle systems adapt to such an array of substrate remains poorly understood on the molecular level. In addition, these past snapshots represented exclusively ‘inward facing’ conformations, therefore limiting our molecular understanding of the transitions required to complete an entire transport cycle. In a first step, we determined high resolution three dimensional structures of the prototypical POT DtpB from *E. coli*, bound to 14 different di- and tripeptides, using macromolecular crystallography (MX). This work provides a profound basis for understanding promiscuity and ligand recognition in POTs at the molecular level. Second, I used single particle analysis cryogenic electron microscopy (SPA cryo-EM), to determine the first structures of the human peptide transporters 1 (HsPepT1) and 2 (HsPepT2). Human PepT1 and PepT2 were captured in four different states throughout the transport cycle, providing a dynamic molecular understanding of substrate uptake within the SLC15 family. Third, we continued using SPA cryo-EM to determine the first structure of PHT1, in an outward facing conformation. This work provides a framework to determine the structure of this newly identified target of SLE, which could be used to obtain high resolution data with various therapeutics. Last, I determined the first structure of the atypical POT DtpC, from *E. coli*. In this work, we explored various fiducial strategies, to improve the resolution of the reconstruction of MFS transporters in general, and provided a molecular explanation for the selectivity of DtpC towards positively charged dipeptides. In summary, this work delivers new insights into the working principles of proton coupled oligopeptide transporters, and will serve as a reference for future structure based drug design (SBDD) studies targeting members of this family.

Biophysical and structural characterization of a multifunctional viral genome packaging motor

Nikolai S. Prokhorov[1,2], Christal R. Davis[3], Kashyap Maruthi[2], Qin Yang[3], Michael B. Sherman[2], Michael Woodson[2], Mark A. White[2], Lohra M. Miller[4], Martin F. Jarrold[4], Carlos E. Catalano[3] and Marc C. Morais[1,2]

[1] Department of Molecular and Cellular Biochemistry, Indiana University, Bloomington, IN
[2] Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch at Galveston, Galveston, TX
[3] Department of Pharmaceutical Chemistry, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado Anschutz Campus, Aurora CO
[4] Department of Chemistry, Indiana University, Bloomington, IN

The large dsDNA viruses replicate their DNA as concatemers consisting of multiple covalently linked genomes. Genome packaging is catalyzed by a **terminase enzyme** that excises individual genomes from concatemers and packages them into preassembled procapsids. These disparate tasks are catalyzed by terminase alternating between two distinct states—a stable nuclease that excises individual genomes and a dynamic motor that translocates DNA into the procapsid. It was proposed that bacteriophage λ terminase assembles as an anti-parallel dimer-of-dimers nuclease complex at the packaging initiation site. In contrast, all characterized packaging motors are composed of five terminase subunits bound to the procapsid in a parallel orientation. Here, we present biophysical and structural characterization of the λ holoenzyme complex assembled in solution. Analytical ultracentrifugation, small angle X-ray scattering, and native mass spectrometry indicate that 5 subunits assemble a cone-shaped terminase complex. Classification of cryoEM images reveals starfish-like rings with skewed pentameric symmetry and one special subunit. We propose a model wherein nuclease domains of two subunits alternate between a dimeric head-to-head arrangement for genome maturation and a fully parallel arrangement during genome packaging. Given that genome packaging is strongly conserved in both prokaryotic and eukaryotic viruses, the results have broad biological implications.

Cryo- EM structure of the mycobacterial 70S ribosome in complex with ribosome hibernation promotion factor RafH

Niraj Kumar

Ribosome hibernation is a key survival strategy bacteria adopt under environmental stress, where a protein, hibernation promotion factor (HPF), transitorily inactivates the ribosome. *Mycobacterium tuberculosis* encounters hypoxia (low oxygen) as a major stress in the host macrophages and upregulates the expression of RafH protein, which is crucial for its survival. The RafH, a dual domain HPF, an orthologue of bacterial long HPF (HPFlong), hibernates ribosome in 70S monosome form. Whereas in other bacteria, the HPFlong induces 70S ribosome dimerization and hibernates its ribosome in 100S disome form. Here we report the cryo- EM structure of *M. smegmatis*, a close homolog of *M. tuberculosis*, 70S ribosome in complex with the RafH factor at an overall 2.8 Å resolution. The RafH N-terminus domain (NTD) binds to the decoding center, a similar binding for HPF long NTD, contrary to the HPFlong C- terminus domain (CTD), the RafH CTD, which is larger, binds to a unique site at the platform binding center of the ribosomal small subunit. The two domain connecting linker regions, which remain mostly disordered in earlier reported HPFlong structures, interact mainly with the anti-Shine Dalgarno sequence of the 16S rRNA.

Characterizing the Synuclein and Tau Structures in Post-Mortem Brains of Lewy Body Dementia Patients

N.Shafiei[1] , W.D.J.Van de Berg[2], H.Stahlberg[1], A.J.Lewis[1]

[1] Laboratory of Biological Electron Microscopy, Institute of Physics, SB, EPFL & Department of Fundamental Microbiology, Faculty of Biology and Medicine, UNIL, Switzerland.
[2]Department Anatomy and Neurosciences, section Clinical Neuroanatomy and Biobanking, Amsterdam UMC, Vrije University, Amsterdam Neuroscience, Amsterdam, The Netherlands.

Millions of people around the world suffer from neurodegenerative diseases. While specific treatments may be able to relieve some of the symptoms of neurodegenerative diseases, in most cases, these complex neurological disorders do not have a cure as we do not completely understand the cause of the disease. The current theory for the mechanism of neurodegenerative diseases is that the causative protein misfolds into a beta-sheet conformation triggering the formation of amyloid fibrils that aggregate into distinctive cellular inclusions. One such protein is alpha-synuclein (aSyn) which accumulates into fibrils in Parkinson's disease (PD), dementia with Lewy bodies (DLB), and multiple system atrophy (MSA). When aSyn amyloid fibrils clump together in diseases such as PD and DLB, they form Lewy bodies. As amyloid fibrils are thought to be the causative agent of the disease, many therapeutic strategies are being developed to inhibit the formation of the fibrils. To facilitate structure-based drug design, several methods have emerged in recent years to generate disease fibrils in vitro and obtain the atomic resolution 3D structure using cryo-electron microscopy. Of the many structures which are emerging from these studies, one major question is if these in vitro-generated structures are the same as what is seen in human disease. My aim is to obtain the high-resolution structure directly within post-mortem human brain samples by cryo-correlative light-electron microscopy (Cryo-CLEM). This will provide a direct representation of the pathological fibril structure, which can be used for drug discovery and mouse models. I can identify the specific fibril strain which makes up the pathology seen in DLB and PD. The results of this study may lead to the discovery of new therapeutic approaches for synucleinopathies and, in particular, for the prevention or slowing of neurodegeneration associated with Lewy Body diseases.

Computational detection and suppression of fiducial nanoparticles post reconstruction to reduce artifacts and improve visual clarity in cryogenic electron tomography datasets

Peter Van Blerkom[1], Armel Bezault[2], Cecile Sauvanet[3], Rowan Borsari[1], Dorit Hanein[1,4], Niels Volkmann[4]

[1]UC Santa Barbara, Department of Chemistry and Biochemistry
[2]Structural Image Analysis Unit, CNRS, Institut Pasteur
[3] Institut de Biologie Intégrative de la Cellule, CNRS, CEA, Université Paris Saclay
[4]UC Santa Barbara, Department of Biological Engineering

Gold nanoparticles with sizes in the range of 10-20 nm are a standard method of providing fiducial markers to assist with alignment during reconstruction in cryogenic electron tomography. However, due to their high electron density and resulting contrast when compared to standard cellular or biological samples, they introduce artifacts such as streaking in the reconstructed tomograms. Here we demonstrate a tool that automatically detects these nanoparticles and suppresses them by replacing them with a local background as a post-processing step, providing a cleaner tomogram without removing any sample relevant information or introducing new artifacts or edge effects from uniform density replacements. This tool can also be applied to single particle analysis when nanoparticles are present, mitigating attention issues in 2D and 3D classification. This is because nanoparticles typically generate a stronger signal than conventional biological samples like proteins.

Accurate and Rapid 2-D Class Averages in Single Particle Analysis using Bispectral Invariants

P. R. Baldwin and S Ludtke

In CryoEM single particle analysis, 2-D class-averages are used for at least four purposes. First, they may be used as a first stage assessment of particle quality, and to identify “bad particles”. Secondly, they may be used for initial assessment of structural variability/flexibility, by looking for class-averages in the same orientation with variability in specific domains. Third, they may be used as references for additional rounds of particle selection from raw micrographs. Finally, they may be used to generate initial 3-D models for subsequent refinement.

While some of these tasks, such as initial model generation, do not require use of the full data set, other tasks, such as bad particle identification, clearly require classification of the entire set of input particles. This task may therefore make use of hundreds of thousands or even millions of particles. Classification algorithms in most cryoEM softwares rely on rotational/translational registration of images, as a precursor to deciding which images are similar. We have implemented the classification of 2D images using a set of bispectral invariants, which are independent of image rotation/translation, but also retain much of the information content in the images. This greatly speeds the classification process, by omitting the computationally expensive registration step.

We discuss the mathematical underpinnings of the method, and test several single particle data sets from EMPIAR to demonstrate the efficacy of the particle quality assessment. In comparing to RELION, RELION typically produces a few very good classes, and is more effective at the very lowest signal to noise ratios, but EMAN produces a broader array of class averages in as much as 50 times faster. For that reason, it is nearly always advantageous to run the EMAN classification to acquire some feel for the views and conformations of a protein of interest.

The use of Feret signatures for the characterization of heterogeneous single-particle samples.

Pierre Nottelet [1], Xiao-Ping Xu [2], Peter Van Blerkom [1], Dorit Hanein [1,3], Niels Volkmann[3]

1UC Santa Barbara, Department of Chemistry and Biochemistry
2The Scintillon Institute, La Jolla
3UC Santa Barbara, Department of Biological Engineering

Feret diameters of an object are defined as the distance between two parallel tangents of the object's boundaries. Feret diameters are often used for the characterization of particle sizes and their distribution in powder samples or polycrystalline solids. We recently used distributions of Feret diameters, the Feret signature, to characterize and detect heterogeneity in nanodisc samples. Here, we explore the potential of Feret signatures to inform on the heterogeneity and classification performance within the framework of single-particle reconstruction of protein samples exhibiting a high degree of conformational flexibility. In particular, we investigate the Feret signatures of membrane-embedded human αIIbβ3 integrin complexes.

Cryo-electron tomography reveals the structural diversity of cardiac proteins in their cellular context

Rahel A. Woldeyes [1], Masataka Nishiga [2,3], Alison S. Vander Roest [4,5], Leeya Engel [6,7], Prerna Giri [2,4], Gabrielle C. Montenegro [1], Andrew C. Wu [1], Alexander R. Dunn [7], James A. Spudich [8], Daniel Bernstein [2,4], Michael F. Schmid [1,9], Joseph C. Wu [2,3] and Wah Chiu [1,9]*

1 Department of Bioengineering, Stanford University, Stanford, CA, USA
2 Stanford Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA, USA
3 Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, USA
4 Department of Pediatrics, Stanford University School of Medicine, Stanford, CA, USA
5 Department of Biomedical Engineering, University of Michigan, MI, USA
6 Faculty of Mechanical Engineering, Technion - Israel Institute of Technology, Haifa, Israel
7 Department of Chemical Engineering, Stanford University, Stanford, CA, USA
8 Department of Biochemistry, Stanford University School of Medicine, Stanford, CA, USA
9 Division of Cryo-EM and Bioimaging, SSRL, SLAC National Accelerator Laboratory, Stanford University, Menlo Park, CA, USA

Cardiovascular diseases are a leading cause of death worldwide, but our understanding of the underlying mechanisms is limited, in part because of the complexity of the cellular machinery that controls the heart muscle contraction cycle. Cryogenic electron tomography (cryo-ET) provides a way to visualize diverse cellular machinery while preserving contextual information like subcellular localization and transient complex formation, but this approach has not been widely applied to the study of heart muscle cells (cardiomyocytes). Here, we deploy a platform for studying cardiovascular disease by combining cryo-ET with human induced pluripotent stem cell–derived cardiomyocytes (hiPSC-CMs). After developing a cryo-ET workflow for visualizing macromolecules in hiPSC-CMs, we reconstructed sub-nanometer resolution structures of the human thin filament, a central component of the contractile machinery. We also visualized a previously unobserved organization of a regulatory complex that connects muscle contraction to calcium signaling (the troponin complex), highlighting the value of our approach for interrogating the structures of cardiac proteins in their cellular context.

Structural insights into a novel flavivirus using sub-volume extraction and refinement.

Sayan Das

Since the resolution revolution, the field of cryo-electron microscopy has made major strides towards developing new hardware and software for gaining deeper insights into the structures of biological macromolecules. However, we still rely heavily on global symmetry averaging when it comes to solving virus structures. This poses a major challenge in limiting the resolution for structural analysis of enveloped and flexible viruses like flavivirus. Flaviviruses are a family of positive sense single strand RNA enveloped viruses which are responsible for a significant global health burden. These arthropod-borne viruses infect humans resulting in serious diseases; some family members infect with high mortality rates. Although there are vaccines available for a few of these viruses there have been issues with vaccine safety and efficacy. Recent software advancements have introduced localized and block-based reconstructions that can allow extracting sub-volumes for refinement, while enforcing local symmetry. This advancement is particularly important as it allows us to study smaller virus sub-volumes at higher resolutions, compensating for potential capsid-wide flexibility and allowing the reconstruction to be performed at full pixel size. Our lab has previously developed a localized reconstruction software called Icosahedral Subvolume Extraction and Correlative Classification (ISECC) that can address the reconstruction problems resulting from the large flexible capsids and refine individual asymmetric units at higher resolution.

In our study, we generated a chimeric flavivirus having the genetic backbone of Yellow Fever Virus (YFV) and the structural protein genes of Deer Tick Virus (DTV) to reduce the biosafety level from BSL-3 to BSL-2 for structural investigations. We used high-resolution cryo-electron microscopy and subvolume refinement to resolve the structure of the chimeric virus at a nominal resolution of 2.8 Å. Analysis of the density map has revealed the presence of two known lipid densities that were successfully built with confidence. These lipids likely mediate certain hydrophobic interactions within the virus structure. Furthermore, we observe another finger-like density that arises from the unstructured viral core. This unfilled density is located between the hydrophobic transmembrane helices of structural proteins. Previous studies on zika virus (ZIKV) in our lab revealed the presence of a novel pocket factor that we showed was essential for the ZIKV maturation pathway. This finding opens a new avenue of drug discovery since the lipid moiety may be used as a novel target for lipid-based antivirals. Hence, gaining insights into structural proteins of the yDTV chimera has future implications in structure-guided drug design studies against flaviviruses.

Hierarchical Assembly of Intrinsically Disordered Short Peptides

Shane Rich-New[1], Jiaqi Guo[2], Chen Liu[3], Yimeng Huang[2], Weiyi Tan[2], Hongjian He[1], Meihui Yi[2], Xixiang Zhang[3], Edward H. Egelman[4], Bing Xu[2], Fengbin Wang[1]

[1]Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, Birmingham, Alabama 35203
[2]Department of Chemistry, Brandeis University, Waltham, Massachusetts 02453, United States
[3]Physical Science and Engineering Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia
[4]Department of Biochemistry and Molecular Genetics, University of Virginia, Charlottesville, VA 22908, USA

Intrinsically disordered peptides (IDPs) are peptide sequences derived from the intrinsically disordered regions (IDR) of proteins. These small fragments are capable of self-assembling into higher-order nanostructures, but compared to the exploration of canonical peptide assemblies that adopt highly characterized secondary structures such as α-helices and β-sheets, the mechanisms governing the self-assembly of IDPs are poorly understood. Thus far, peptide nanomaterials have shown promise for use in material science, biotechnology, and therapeutic design. In the past, we have demonstrated that molecules akin to those used in this research acted as targeted, anti-cancer therapeutics. Therefore, it is critical that we understand these rules to inform rational design and adaption of these molecules for various use-cases. During our research, we demonstrated that our designed IDPs formed nanofibers. We next aimed to understand the main driving forces for the self-assembly of these molecules, such as regulatory effects enacted by post-translational modifications (PTMs), metal ion-enabled self-assembly, and the canonical forces governing protein/peptide folding, including hydrophobic collapse, and hydrogen bond networking by examination of each IDP's near-atomic resolution structure. To accomplish this, The Xu lab (Brandeis) conjugated an aromatic motif and introduced a phosphate group to a heptapeptide, YSPTSPS, that was originally derived from the disordered region of the C-terminal domain of human RNA polymerase II subunit A (POLR2A). Throughout this research, the effects of altered phosphorylation site, primary sequence, addition of metal ions, and conjugation of different hydrophobic motifs were examined to ascertain their relative importance for self-assembly. We determined the structures of one such peptide, with and without the presence of Ca^{2+} using cryo-EM, which reached 3.1 and 3.0 Å, respectively, and elucidated the roles that these modifications play on self-assembly of IDPs.

Cryo-ET and Subtomogram Averaging of the Nipah Virus Matrix Protein Lattice

Viraj Upadhye[1], Jean Lee[1], Nihan Ercarli[1], Hector C. Aguilar[1*] & Robert Dick[2*]
[1]Department of Microbiology and Immunology Cornell University College of Veterinary Medicine T8016A Veterinary Research Tower Ithaca, NY
[2] Center for ViroScience and Cure, Department of Pediatrics, Emory University School of Medicine HSRB E492, 1760 Haygood Drive NE Atlanta, GA 30322
*Co-Corresponding authors

Paramyxoviruses are pervasive human pathogens that include measles virus, mumps virus, human parainfluenza viruses, and the deadly zoonotic henipaviruses Nipah (NiV) and Hendra (HeV) viruses. All paramyxoviruses have a matrix protein (M), which orchestrates viral budding, a fusion protein (F), a receptor-binding protein (G/RBP/HN), and an RNA packaging nucleocapsid protein (N). For NiV, M or F can spontaneously induce the budding of enveloped particles that are commonly referred to as virus-like-particles (VLPs). The NiV M protein dimer has been solved via X-ray crystallography to 2.05Å, and M protein arrays for New Castle Disease and Measles viruses have been observed at ~30Å resolution. Currently, there are no publicly available sub-nanometer structures of any paramyxovirus M lattice. To address this, we developed a method to generate a high concentration of NiV M VLPs. Previously M VLPs were reported to be spherical, which we also observed by negative stain TEM. However, here we found by cryo-EM that M VLPs are tubular, with a diameter range of 50-110nm. Further, the co expression of NiV M with a very low amount of G resulted in a sample of sufficient quality and abundance for cryo-ET data collection. Subtomogram averaging of M-VLPs resulted in a 7.5 Å map that describes the quaternary arrangement of M in the viral lattice. The lattice is formed by 2-fold symmetric M-dimers that rotate by 90˚ with respect to each other resulting in two types of non-equivalent tetramers of dimers. Additionally, through extensive co-expression experiments of M with other structural proteins, we investigated the influence of NiV N, G and F on the M lattice and on particle morphology. High concentrations of G resulted in the formation particles that were more spherical and contained little to no M lattice. The presence of N resulted in an increase in particle polymorphism and we demonstrated that N oligomer packaging is efficient in the presence of M and independent of the viral genome. Interestingly, in F-only particles F appeared randomly distributed, F became highly organized in the presence of M, occupying one of the tetramers. In summary, we present the first sub-nanometer structure of any paramyxoviral M lattice and provide a structural bases that can explain paramyxoviral pleomorphism.

The Biomolecular cryo-EM Facility at UC Santa Cruz: What Should You Expect?

Vitor Hugo Balasco Serrão[1,2], Melissa Jurica[3,4], Sarah Loerch[1], Rebecca DuBois[5], Carrie Partch[1], Harry Noller[3,4], Jevgenij Raskatov[1], Alexander Ayzner[1], Joshua Arribere[3,4] and Seth Rubin[1]

1Department of Chemistry and Biochemistry - UCSC;
2Biomolecular cryo-Electron Microscopy Facility - UCSC;
3Department of Molecular Cell and Developmental Biology – UCSC;
4Center for Molecular Biology of RNA – UCSC;
5Department of Biomolecular Engineering - UCSC

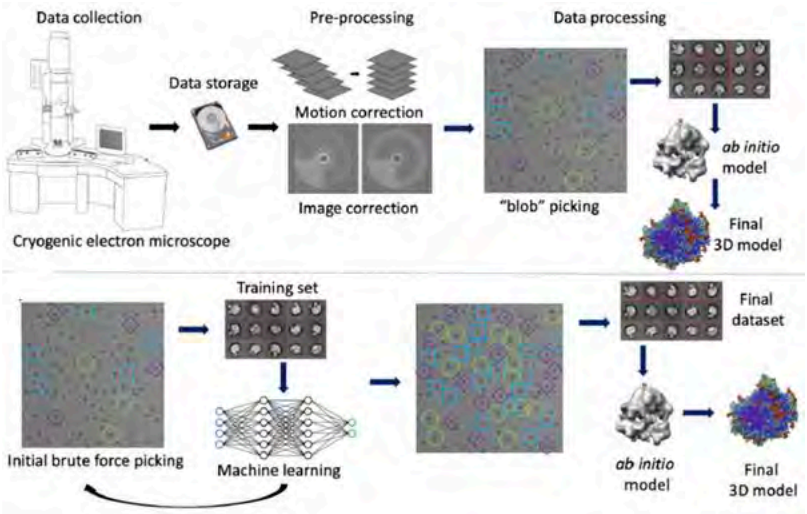

Figure 1. The biomolecular cryo-EM facility at UCSC workflow.

Cryo-EM has become a highly effective tool in the field of structural biology. In 2013, the 'resolution revolution', initiated by a faster, more sensitive detector, drastically changed the field and pushed the visible details of molecules to atomic resolution. This cutting-edge approach offers the unique possibility of investigating biological molecules in their native state and revealing how these cellular machines function.

The Biomolecular cryo-EM facility at the University of California, Santa Cruz, is a leading provider of cutting-edge structural biology solutions using transmission electron microscopy. We provide access and guidance to a complete workflow from sample analysis and optimization to grid preparation, automated high-resolution imaging, data analysis, and tridimensional reconstruction, which allows structure determination via model construction and validation.

Our Glacios 200 kV is coupled with a Gatan K2 Summit direct detector, which allows sample preparation quality to be checked using our screening time blocks (4 hours, up to 8 grids). Low-magnification maps are initially collected using SerialEM, allowing fast selection of the top grids based on the overall ice quality. Pre-screened grids can be imaged and auto-processed to obtain high-resolution 3D reconstructions. Overnight data collection typically provides ~2,000-5,000 movies and thousands of particles that are automatically processed to optimize the efficiency and quality of the datasets. Using SerialEM coupled to cryoSPARC-Live, our team is capable of obtaining near-atomic resolution within ~16 hours of unsupervised data collection and processing. Our workflow allows preferential orientation and image and particle quality analysis, increasing the data quality and resulting in an initial 3D reconstruction in-live.

For data processing, our facility manages and grants access to one of our six GPU-based (NVIDIA RTX-3090) workstations and storage. We have all of the most commonly used software packages for single-particle and subtomogram averaging analysis (Relion 4.0, cryoSPARC 4.1, EMAN 2.91, cisTEM, Phenix, Coot, CCPEM, Dynamo, PEET, etc). Multiple levels of service and access are available and correspond with the type of analysis desired. On-site individual training is provided for users from different levels, with no geographical restrictions. We currently have 30+ ongoing projects, including collaborators from four different countries and companies.
www.ucsccryoem.org – cryoem@ucsc.edu

nextPYP: a user-friendly and comprehensive solution for single-particle cryo-electron tomography

Ye Zhou[1], Hsuan-Fu Liu[2], Jeff Martin[1], Wendy Huang[1], Alberto Bartesaghi[1,2,3]
[1]Department of Computer Science,
[2]Department of Biochemistry,
[3]Department of Electrical and Computer Engineering at Duke University

Single-particle cryo-electron tomography (SP-CET) has emerged as the technique of choice to determine the structure of proteins imaged in situ at high-resolution. While high-throughput techniques for sample thinning and beam-image shift acquisition can produce tilt-series at unprecedented speeds, the demanding storage and computational requirements of the SP-CET data analysis pipeline poses major barriers for the development of this technology. Here, we present nextPYP, an end-to-end software platform designed to streamline the conversion of raw tilt-series into high-resolution structures using SP-CET. Featuring an easy-to-use web-based interface, nextPYP makes the tomography data analysis pipeline accessible to non-experts allowing them to analyze more data and faster. The ability to routinely analyze large datasets is critical for the success of structure determination efforts of proteins imaged under physiological concentrations. The software offers a wide range of tools and options covering all steps of the SP-CET processing pipeline, including tilt-series alignment, particle picking, 3D refinement and classification. In addition to reference-based refinement, nextPYP implements a robust algorithm for ab initio 3D reconstruction and methods to analyze conformational heterogeneity based on 2D projections. Advanced refinement algorithms based on constrained single particle tomography –including per-particle CTF refinement, region-based and per-particle movie-frame refinement– allow the efficient extraction of high-resolution information resulting in better quality maps. nextPYP was designed with computational and storage efficiency in mind, it features a small storage footprint and leverages multi-core CPU architectures for faster processing. With its unique set of features, nextPYP will contribute to expand the adoption of in situ structure determination using SP-CET.

Utilizing random starting points in 3D classification to extract structural states from molecules with continuous motion

Yirui Guo [1,2], Dominika Borek [2,3], Raquel Bromberg [1,2], Zbyszek Otwinowski [2,3]
[1] Ligo Analytics, Dallas, TX.
[2] Department of Biophysics, University of Texas Southwestern Medical Center, Dallas, TX;
[3] Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX;

Capturing the continuous motion exhibited by macromolecules has posed a significant challenge in the field of cryo-EM single-particle analysis. To address this issue computationally, various methods have been developed. In practice, most of these methods perform optimally when the flexible region constitutes only a small portion of the entire macromolecule or when the overall scale of the motion is relatively small. However, when a substantial part of the macromolecule undergoes such motion, these refinement methods often fail to capture the states of motion at high resolution.
In this poster, we present an example of extracting distinct motion states of an ATPase by employing multiple rounds of 3D classification, starting from random points on different subsets of particles. Initially, the structure had a resolution worse than 4 Å and lacked secondary structure features. Through this approach, we successfully extracted four distinct states of motion for this protein, achieving a resolution ranging from 3.4 Å to 3.8 Å, which provided sufficient secondary structure information for atomic model building.

Misleading Features of Half-Maps-Based Validation

Zbyszek Otwinowski [1,2], Dominika Borek [1,2], Raquel Bromberg[1,3], Yirui Guo [1,3]

[1] Department of Biophysics, University of Texas Southwestern Medical Center, Dallas, TX;
[2] Department of Biochemistry, University of Texas Southwestern Medical Center, Dallas, TX;
[3] Ligo Analytics, Dallas, TX.

Half-maps, which are derived from two halves of particle image data, are mandatory components of SPR deposits to the Electron Microscopy Data Bank (EMDB) because during structure determination they are used to assess the resolution-dependent signal-to-noise-ratio (RSNR). While the RSNR per particle is a well-defined statistic, it has a problem when propagating to full reconstruction if preferred orientation is encountered. Currently, uneven coverage information is absent from deposits and ignored in downstream analysis such as molecular refinement, even though refinement programs could easily incorporate this information.

The RSNR for a map, derived from FSC curves between half-maps, is considered the central quality statistic, even if how to calculate it is unclearly defined. Firstly, Wiener-weighting has a large impact in the case of preferred orientation where weighted maps overestimate resolution while unweighted maps underestimate it. If coverage of Fourier space would be deposited, then all relevant statistics could be calculated.

Secondly, RSNR calculations are affected by noise which is proportional to the volume in which it accumulates. Thus, masked calculations are preferable, as unmasked calculations have unspecified box volume which can be larger or smaller, impacting its interpretation. Alternatively, volume correction based on molecular weight consideration can be used. Historical arguments against using masking are no longer applicable when data are of resolution sufficient for model building.

We present the impact of such factors on half-maps-based validation and discuss features of alternative and complementary approaches: (1) using FSC between the reconstruction and the deposited molecular model as a primary validation statistic; (2) modifying the requirement for half-maps deposition to include coverage information; and (3) using only FSC from masked reconstructions, as unmasked FSC curves lack statistical rigor.

Dissecting structural heterogeneity of co-chaperone, and chaperonin TRiC cooperation during the folding cycle

Junsun Park[1], Hyunmin Kim[1], Seyeon Lim, and Soung-Hun Roh[1*]

[1]Department of Biological Science, Seoul National University, Seoul, Republic of Korea.

Proper cellular proteostasis, essential for viability, requires a network of chaperones and cochaperones. ATP-dependent chaperonin TRiC/CCT partners with phosducin-like proteins (PhLPs) to facilitate the folding of essential eukaryotic proteins. Using cryoEM and biochemical analyses, we determine the ATP-driven cycle of TRiC-PFD-PhLP2A interaction. In the open TRiC state, PhLP2A binds to the chamber's equator while its N-terminal H3-domain binds to the apical domains of CCT3/4. ATP-induced TRiC closure rearranges the contacts of PhLP2A domains within the closed chamber. In the presence of substrate, actin and PhLP2A segregate into opposing chambers, each binding to the positively charged inner surfaces formed by CCT1/3/6/8. Notably, actin induces a conformational change in PhLP2A, causing its N-terminal helices to extend across the inter-ring interface to directly contact a hydrophobic groove in actin. Our findings reveal an ATP-driven PhLP2A structural rearrangement cycle within the TRiC chamber to facilitate folding.

Reference

Park, Kim, Gestaut et al., A structural vista of phosducin-like PhLP2A-chaperonin TRiC cooperation during the ATP-driven folding cycle,
bioRxiv 2023.03.25.534239; doi: https://doi.org/10.1101/2023.03.25.534239

Structural Studies of AMOT-NEDD4L-MERLIN Complexes in Hippo Signaling

Adam Campos, John McCullough, Akram Alian, and Wesley I. Sundquist
Department of Biochemistry, University of Utah School of Medicine, Salt Lake City UT, 84112

The conserved Hippo tumor suppressor pathway dictates organ size and regeneration, participates in stem cell self-renewal and expansion, is dysregulated in many human cancers, and is a leading target for anticancer therapeutic development. Here, we propose to define the biochemistry and structural biology of a key class of kinase/ubiquitin-responsive complexes that function as central nodes in Hippo signaling (and elsewhere). These complexes comprise motin family proteins (AMOT p130, AMOT p80, AMOTL1, or AMOTL2), the Merlin (Neurofibromin 2) tumor suppressor, the HECT ubiquitin E3 ligase NEDD4L, and YAP/TAZ transcriptional co-activators. The complexes transduce Hippo kinase signaling inputs into biological outputs that either favor homeostasis state (“Hippo ON”) or promote actin assembly and proliferative gene transcription (“Hippo OFF”).

In preliminary studies, we have shown that AMOT proteins form large, homopolymeric assemblies that appear to function as signaling platforms that bind NEDD4L, Merlin, and YAP/TAZ, and recruit LATS kinases. When Hippo signaling is on, these complexes activate the LATS and NEDD4L enzymes, promote phosphorylation of LATS, AMOT, and YAP/TAZ, and facilitate Merlin ubiquitination, thereby creating phospho-YAP/TAZ complexes that are retained in the cytoplasm and inducing conformational changes that allow Merlin to function as a tumor suppressor. When Hippo kinase signaling is off, unmodified YAP/TAZ subcomplexes move into the nucleus to stimulate transcription of proliferative genes, and AMOT-Merlin complexes nucleate actin polymerization to promote cell motility. The fundamental goal of our research is to gain a molecular understanding of this central switch in Hippo-mediated signaling.

Our initial focus targeted the structure of AMOT p80, the simplest of the AMOT assemblies. In our preliminary studies, we have: 1) overexpressed AMOT p80 in human HEK 293T cells, 2) optimized the purification of milligram quantities of native AMOT p80 assemblies, 3) demonstrated that AMOT p80 forms very large homomeric assemblies (~5 MDa) and 4) performed preliminary single particle cryo-EM reconstructions showing that AMOT p80 forms a striking “three-tiered” platform. We now propose to determine the structure of the AMOT p80 assembly to near-atomic resolution by collecting much more cryo-EM data of improved samples.

Transformation of Inorganic Nanoparticle Electron Microscopy Analysis

MaKayla Walker[1*], Christopher Ackerson[2]
[1]Colorado State University, Department of Chemistry, 1301 Centre Ave, 80523, Fort Collins, Colorado, US
[2]Colorado State University, Department of Chemistry, 1301 Centre Ave, 80523, Fort Collins, Colorado, US
[*]MaKayla.Marie.Walker@colostate.edu

Routine characterization of inorganic nanoparticles (iNPs) via transmission electron microscopy (TEM) imaging is primitive (analyzing at most a few hundred particle images by hand), compared to the cryo-EM methods developed for biomacromolecule analysis over the past decade (which allow 10^3-10^6 particle images to be collected and analyzed). We are working to adapt biological cryo-EM methods for routine characterization of iNPs as well as the total determination of iNP structures (including their ligand shells/interfacial layers). Both goals are facilitated by automated data acquisition software (e.g., SerialEM). Expected outcomes include: (1) recommendations to the iNP community for improved protocols for iNP data collection and analysis; (2) total structural determination of iNPs, including their organic ligand shells. Total structures will be determined by adapting biologically developed single particle analysis (SPA) software (e.g., cryoSPARC) for 3D structural determination. As developmental cryo-EM targets, we propose to investigate iNPs with high symmetry axes (e.g., gold nanorods) or alternatively anisotropic iNPs with confined spaces in which ligand movement is restricted (e.g., gold dendrimers). We will present initial results from our attempts to adapt biological cryo-EM methods to improve iNP characterization and determine iNP structure.

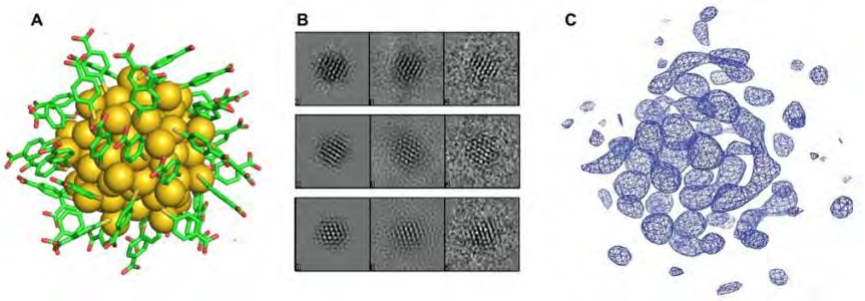

Fig. 1: A) X-ray crystal structure of $Au_{102}(pMBA)_{44}$. B) Class averages and C) electron density of Au_{68}, the first SPA of an iNP.[1,2]

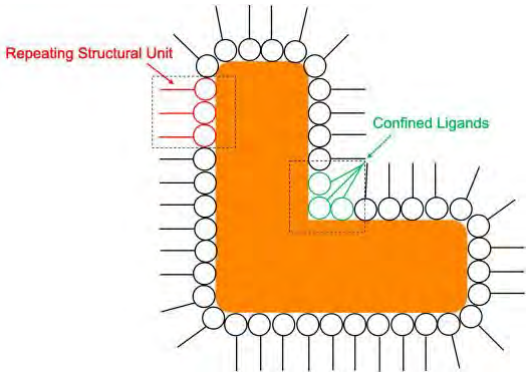

Fig. 2: A cartoon of a hypothetical anisotropic iNP where ligands are either in confined spaces (green) or a part of a repeating structural unit (red).

References:
[1] P.D. Jadzinsky et al., Science **318**, 430-433 (2007).
[2] M. Azubel et al., Science **345**, 909-912 (2014).

The Mechanism of SNARE recycling by a AAA+ supramolecular machine

Yousuf Khan

SNAREs, a diverse set of evolutionary conserved proteins with many domains that drive membrane fusion, have radical length, structural, and biochemical variations outside of their conserved 60-70 amino acid SNARE motif. Despite this, all SNAREs are hypothesized to be recycled by the same molecular machinery. This AAA+ complex, Sec18, is conserved in all eukaryotes and its dysfunction can range from lethality to neurological deficits. At present, there exists no mechanism to explain Sec18's all-encompassing substrate processing. Here we combine *in-vivo* mass-spectrometry, kinetically-informed Cryo-EM, and single-molecule FRET to structurally characterize the different states of Sec18 throughout SNARE recycling. apoSec18 exists as both a heptamer and a split D1/D2 hexamer, which is ready to accept substrate through an unexpected, ATP-hydrolysis independent side-loading mechanism. Substrate is then primed and disassembled in an ATP-hydrolysis dependent manner. Sec18 then breaks its D1/D2 ring once again and releases disassembled SNARE. This detailed characterization of SNARE recycling provides a fundamental basis as to how all SNAREs in the cell are recycled.

Unsupervised deep learning for cryo-ET structural pattern mining

Min Xu
Computational Biology Department
Carnegie Mellon University, USA

Recent developments in cryo-electron tomography (cryo-ET) allow direct visualization of native sub-cellular structures with unprecedented details. We have been developing methods to find structural patterns that may potentially represent macromolecular complexes captured by cryo-ET. However, cryo-ET has high structural complexity and imaging limits, which makes structural pattern mining difficult. Deep learning has been increasingly useful for facilitating pattern mining thanks to its capability to learn complex rules from large amounts of data. However, supervised deep learning requires a large amount of annotations which is often not available. We describe several recent works on developing unsupervised approaches to support structural pattern mining in cryo-ET data, including denoising, simulation, subtomogram clustering, and alignment.