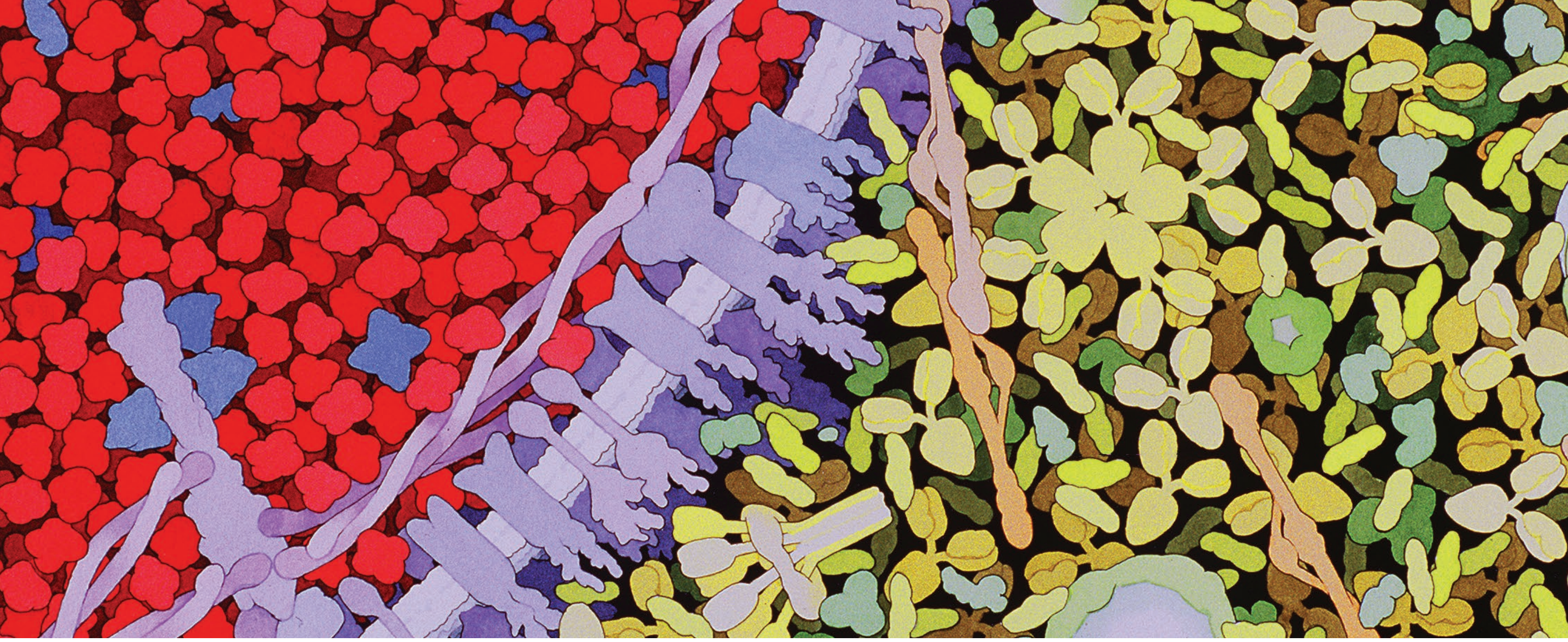

Virtual symposium celebrating the 50th anniversary of the Protein Data Bank

PROTEIN DATA BANK 50 Years

May 4–5, 2021

Day 1 Abstracts

ASBMB

WORLDWIDE PDB PROTEIN DATA BANK

Worldwide Protein Data Bank Foundation

Day 1: Tuesday May 4, 2021

POSTER SESSION 1
Tuesday May 4, 2021 1:00pm-1:40pm EDT

Identification of GntR/HutC Transcriptional Regulator MoyR from Mycobacterium tuberculosis that responds to monovalent cations in DNA binding and Oligomerization

Thanusha D. Abeywickrama[1], Inoka C. Perera[1]

1 Department of Zoology and Environment Sciences, Faculty of Science, University of Colombo

Mycobacterium tuberculosis is a well-known bacterium for its mounting of drug resistance and the modern strain survived over 70,000 years claiming millions of lives each year. Transcriptional regulators play a key role in the survival of bacteria under varying environmental conditions. The transcriptional regulator MoyR encoded by the gene moyR (Rv0792c) and the intergenic operator region moyO is shared by divergently transcribed Rv0789c-Rv0790-Rv0791c-moyR operon and Rv0793 where protein blast and the homology modeling revealed that these genes could be encoding a set of monooxygenases. A 10 bp imperfect direct repeat was predicted as the binding motif of MoyR within the 119 bp intergenic region and binding of MoyR to the intergenic region was identified using EMSA and in the presence of 100 to 300 mM concentration of KCl, it showed that three MoyR-DNA complexes were formed indicating binding of dimeric, tetrameric and hexameric MoyR with its cognate DNA. In the concentrations higher than 300 mM it showed that both dimeric and hexameric MoyR complexes were dissociated where only the tetrameric MoyR-DNA complex was stabilized in higher concentrations. Oligomeric states of MoyR were determined using size exclusion chromatograph in two different monovalent cation concentrations, 150 mM and 500 mM). Calculated molar weight of MoyR form calibration curve indicates a dimer at 100 mM concentration of NaCl and at 500 mM concentration the obtained values corresponds to a tetramer and a hexamer. This implies that MoyR is predominantly a dimer in solution whereas increasing monovalent cation concentration led to formation of higher oligomer states of MoyR. This was further analyzed by glutaraldehyde mediated crosslinking. In silico analysis of MoyR oligomer states were identified using GalaxyWeb server and higher similarity scores with templates reveals that MoyR protein can be present as homodimers and tetramers in the solution. GntR proteins bind to DNA as homo-dimers and tetrameric MoyR could be the functional state which can be crucial for its regulatory role and identifying its importance to *M. tuberculosis*. Identifying small molecule responsiveness of MoyR and stability of oligomers could be with an importance to find drug leads to combat TB.

The investigation of evolutionary events leading to substrate diversification in type I GPAs by the use of ancestral sequence reconstruction

Mathias H. Hansen[1], Max Cryle[1]

[1]Monash University, Clayton, Australia

Glycopeptide antibiotics (GPAs), a distinct class of antibiotics, are essential for the treatment of infections caused by Gram-positive bacteria. GPAs and other antibiotics today are natural products derived from soil microbes. Glycopeptide antibiotics are encoded by biosynthetic gene clusters that include genes supporting their regulation, synthesis, export, and resistance. Key to the peptide synthesis process is the actions of adenylation (A)-domains. These are responsible for the selection of amino acid substrates and their activation prior to incorporation of these residues into the growing peptide chain and are thus the main site for substrate selectivity. During the course of evolution GPAs have developed a wide variety of chemical and biosynthetic diversity. Vancomycin and pekiskomycin type GPAs evolved to contain aliphatic amino acids at position 1 and 3 of the heptapeptide instead of non-proteinogenic phenylglycine residues. Here, we explore the evolutionary history that gave rise to the selection and activation of alanine and leucine at position 1 of GPAs through phylogenetic analysis and ancestral sequence reconstruction (ASR). These predictions allowed us to express and purify four ancestral A-domains marking the evolution of type I GPAs and explore their biochemical activity, proving that diversification of the first module A-domain occurred through evolution of an intermediate ancestral pocket with low specificity for leucine, which subsequently diverged into the modern A-domains specialised into accepting either leucine or alanine. Additionally, we used the first A-domain from Tcp9 as a model to mutate key binding pocket residues to explore the evolutionary changes seen in the four ancestral A-domains and to structurally characterise them. These structures illustrate the structural change required to progress from the flat 4-Hydroxyphenylglycine (Hpg) binding pocket to a deepened cavity that can accommodate aliphatic amino acids, and then subsequently to be partitioned, amplified and refined through natural selection.

Structure-based drug repositioning exploiting PLIP interaction patterns uncovered known drugs as novel treatments for Chagas disease

Melissa F. Adasme[1], Sarah Naomi Bolz[1], Lauren Adelmann[1], Sebastian Salentin[1], V. Joachim Haupt[1], Adriana Moreno-Rodríguez[2], Benjamín Nogueda-Torres[3], Verónica Castillo-Campos[3], Lilián Yepes-Mulia[4], José A. De Fuentes-Vicente[5], Guildardo Rivera[6], Michael Schroeder[1]

[1]Biotechnology Center (BIOTEC), Technische Universität Dresden, 01307 Dresden, Germany, [2]Facultad de Ciencias Químicas, Universidad Autónoma Benito Juárez, Oaxaca 68120, Mexico, [3]Departamento de Parasitología, Instituto Politécnico Nacional, Ciudad de México 11340, Mexico, [4]Unidad de Inv. Méd. en Enf. Infecciosas y Parasitarias, Instituto Mexicano del Seguro Social, México, [5]Laboratorio de Investigación y Diagnóstico Molecular, Univ. de Ciencias y Artes de Chiapas, Mexico., [6]Laboratorio de Biotecnología Farmacéutica, Instituto Politécnico Nacional, Reynosa 88710, Mexico

Chagas disease, caused by the parasite *Trypanosoma cruzi (T. cruzi)*, affects millions of people in South America. The current treatments are limited, have severe side effects, and are only partially effective. In light of this, drug repositioning, defined as finding new indications for already approved drugs, has the potential to provide new therapeutic options for Chagas disease. In this study, a structure-based drug repositioning approach has been conducted with over 130,000 3D protein structures which were analyzed with the tool PLIP for the detection of protein-ligand non-covalent interactions. The repositioning approach exploits the PLIP interaction patterns defining the binding mode of drugs that bind to known therapeutic targets (Chagas targets) to detect others with similar binding mode and thus, identify potential new Chagas treatments. Three of the new drug candidates—ciprofloxacin, naproxen, and folic acid—showed an *ex vivo* growth inhibitory activity in the micromolar range when tested on *T. cruzi* trypomastigotes and an *in vivo* parasitemia inhibition when tested in mice infected with *T. cruzi* strain Ninoa. As over the counter drugs, they are perfect for further validation in models and in humans, and could eventually lead to more efficient and better-tolerated treatments of Chagas disease. Overall, the structure-based drug repositioning approach was able to pinpoint relevant drug candidates at a fraction of the time and cost of a conventional drug discovery screening. Additionally, the results of this study demonstrate the potential of the approach in the context of neglected tropical diseases where the pharmaceutical industry has little financial interest in the development of new drugs.

This abstract is based on the following published articles: Repositioned drugs for Chagas disease unveiled via structure-based drug repositioning.International journal of molecular sciences 21.22 (2020): 8809. https://doi.org/10.3390/ijms21228809 and PLIP: fully automated protein–ligand interaction profiler. Nucleic acids research 43.W1 (2015): W443-W447. https://doi.org/10.1093/nar/gkv315

Structure of Enterohemorrhagic Escherichia coli O157:H7 Intimin Virulence Factor Bound to Nanobodies

Angham M. Ahmed[1], Cory L. Brooks[1]

[1]California State University, Fresno

Enterohemorrhagic E. coli O157:H7 (EHEC) presents a significant risk to human health. EHEC resides in the intestinal tract of cattle but can be transmitted to humans when contaminated food is ingested. Once EHEC colonizes the human intestinal tract, it secretes Shiga-like toxins causing serious illnesses. Antibiotics are not recommended for treatment of EHEC as they can induce enterotoxin release, increasing the risk of HUS. There is currently no effective treatment or vaccine against EHEC. EHEC pathogenicity involves intimin, which mediates bacterial colonization of the GI tract. Blocking the interaction of intimin with its cogent receptor, Tir, could neutralize EHEC pathogenesis. One novel solution for treating EHEC infection comes from antibodies known as nanobodies or V_HHs, which are the antigen binding domain of heavy chain antibodies produced by the Camelid family. Intimin specific nanobodies were generated by immunizing a llama followed by phage display and selection for high specificity and affinity for intimin binding nanobodies. We hypothesize that these nanobodies (V_HH1-V_HH5) could specifically bind the Tir-binding domain of intimin neutralizing EHEC infection. The goal of this project is to determine the nanobodies-intimin complex crystal structures and examine the physical interaction occurring between key residues of the two proteins. Understanding such interaction could form the bases for identifying and developing novel therapeutics for treating EHEC pathogenesis. The nanobodies-intimin have been purified and co-purified using size exclusion chromatography. Crystals complexes were harvested, sent out for x-ray data collection, diffraction measurements of V_HH2-intimin and V_HH3-intimin complexes were collected and processed in space groups P1 and P212121 with the resolution of 2.1 and 1.8 Å, respectively. Initial analysis of the structures shows that CDR3 and a non-CDR loop interact with intimin via polar interaction. Next, we are planning to preform inhibition assays using mammalian cells to evaluate the potency of the nanobodies against the bacteria.

Understanding SARS-CoV-2 Protein Evolution in 3D During the COVID-19 Pandemic: ORF7a

Luz H. Alfaro[1], Thejasvi Venkatachalam[2], Helen Zheng[5], Elliott Dolan[2], Changpeng Lu[2], Vidur Sarma[2], Zhuofan Shen[2], Maria Szegedy[2], Lingjun Xie[2], Christine Zardecki[3], Sagar Khare[2], Stephen K. Burley[3]

[1]Grinnell College, [2]Rutgers IQB, [3]RCSB PDB, [4]Rutgers University, [5]Watchung Hills Regional

The year 2019 took an unprecedented turn as SARS-CoV-2 (COVID-19) ravaged the world, taking over half a million lives (as of July 2020), leaving vulnerable communities to take the hardest hits, exacerbating systemic injustices towards marginalized people, and damaging economies. By examining proteins and their respective mutations in SARS-CoV-2, a better understanding of its structural biology can aid in drug and vaccine discovery. Further applications of a deepened structural understanding of SARS-CoV-2 include assisting in broader research efforts and valuable public health efforts to mitigate and prevent harm caused by SARS-CoV-2 and other coronaviruses worldwide.As part of a virtual summer research experience with the RCSB PDB, we studied how SARS-CoV-2 proteins protein evolved during the first six months of the COVID-19 pandemic by exploring amino acid sequence and 3D atomic-level structure using various structural bioinformatics tools, including Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo/) for sequence alignments and phylogenetic trees; Mol* (molstar.org) for 3D molecular visualization; and Foldit (fold.it) for structural/energetic effects of sequence mutations.The focus of this poster is SARS-CoV-2 Orf7a, a viral protein that interferes with N-linked glycosylation of the cellular protein BST-2, an inhibitor of coronavirus release (Taylor et al. 2015). Orf7a has also been shown to arrest cells in the G0/G1 cell checkpoint during SARS-CoV infection, thereby preventing cell cycle progression. The G0/G1 cell checkpoint is a phase in the cell division process wherein DNA is checked for damage before allowing progress into the S phase. If DNA damage cannot be repaired cell death ensues. Prevention of cell cycle progression is thought to provide more time for coronaviruses to replicate at higher rates with increased nucleotide pools and allow for higher virus pathogenicity (Yuan et al. 2005). Studying Orf7a can provide insights into SARS-CoV-2 pathogenesis, including what proteins are involved in regulating viral propagation and what SARS-CoV-2 proteins affect host cell cycle regulation. RCSB PDB is funded by the National Science Foundation (DBI-1832184), the US Department of Energy (DE-SC0019749), and the National Cancer Institute, National Institute of Allergy and Infectious Diseases, and National Institute of General Medical Sciences of the National Institutes of Health under grant R01GM133198.

Crystal Structures of KPC-2 β-Lactamase in Complex with Triazolyl Boronic Acid Transition State Inhibitors

Tahani A. Alsenani[1], Samantha Viviani[1], Magdalena A. Taracila[2, 7], Emilia Caselli[8], Fabio Prati[8], Robert A. Bonomo[1, 2, 3, 4, 5, 6, 7], Focco Van den Akker[1]

[1]Department of Biochemistry, Case Western Reserve University School of Medicine, OH,USA, [2]Louis Stokes Cleveland Department of Veterans Affairs Medical Center, Cleveland, OH, USA, [3]the CWRU-Cleveland VAMC Center for Antimicrobial Resistance and Epidemiology (Case VA CARES), [4]Departments of Pharmacology, Biochemistry, Case Western Reserve University School of Medicine,OH,USA, [5]Department of Molecular Biology and Microbiology,CWRU,School of Medicine, Cleveland, OH, [6]Department of Proteomics and Bioinformatics, CWRU,School of Medicine, Cleveland, OH, US, [7]Department of Medicine, Case Western Reserve University School of Medicine, Cleveland, OH, [8]Department of Life Sciences, University of Modena and Reggio Emilia, Modena, Italy

Abstract

Antimicrobial resistance (AMR) is one of the most urgent public health problems in the world. One of the most prescribed and time-honored class of antibiotics used to fight bacterial infections are the β-lactams (i.e., penicillins, cephalosporins, and carbapenems). Presently, Klebsiella pneumoniae Carbapenemase-2 (KPC-2) presents a significant clinical threat as this b-lactamase confers resistance to carbapenems, b-lactamase inhibitors (e.g., clavulanate) and oxyimino-cephalosporins. Unexpectedly, KPC-2 variants have also arisen in the clinic that seem to confer resistance to ceftazidime/avibactam, a novel b-lactam b-lactamase inhibitor therapy that has been used successfully to overcome serious infections caused by KPC-2 harboring bacteria. Developing novel inhibitors of KPC-2 and these variants is a strategic goal. To this end, triazolyl-containing boronic acid transition state inhibitors (BATSI) of KPC-2 were designed with these targets in mind. In cell-based assays, the MIC for cefepime (FEP) decreased from 32 to 0.5 mg/L with the addition of 4 mg/L of MB_076, and to 2 mg/L using PCF-003. The IC50's values are in the nM range for all inhibitors (IC50 MB076 = 135 nM and IC50PCF-003 = 460 nM). To probe their mechanism of inhibition of KPC-2, we determined the crystal structures of KPC-2 with these triazolyl boronic acid compounds. The structures of KPC-2 with MB_076, S17079 and PCF-003 were determined at 1.38, 1.07 and 1.10 Å resolution, respectively. The three BATSIs are covalently bonded to the catalytic S70 in the active site of KPC-2. The three boronic acid hydroxyl moieties form stabilizing interactions and occupy the oxyanion hole and the deacylation water pocket in each of the complexes. The triazolyl moieties in S17079 and PCF-003 each hydrogen bond with N132. Each inhibitor also made van der Waals interactions with N170. These crystal structure of KPC-2 BATSI complexes provide insights into the molecular basis of the BATIS's high affinity for KPC-2 and could aid in the design of future BATSI inhibitors of KPC-2 and variants

A Structure-Based Drug Discovery Approach to Stop the Witchweed

Amir Arellano-Saab[1, 2], Duncan Holbrook-Smith[1], Alexei Savchenko[2, 3], Peter J. Stogios[2], Peter McCourt[1]

[1]Department of Cell and Systems Biology, University of Toronto, Toronto, ON. Canada, [2]Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON, Canada, [3]Department of Microbiology, Immunology and Infectious Diseases, University of Calgary, Calgary, AB.

Striga hermonthica is the most destructive food crop parasite in the developing world. It invades its hosts by perceiving a group of host-derived hormones called strigolactones. The receptors in charge of sensing strigolactones are encoded by a group of *α/β* hydrolases that not only recognize strigolactones but also hydrolyze them.
To understand this mechanism and stop the receptor from sensing its ligand, which results in parasite germination, we obtained the crystal structure of the most sensitive strigolactone receptor, in complex with a non-hydrolyzable antagonist molecule. Using this structural data, we have employed a drug-discovery approach to find, design, and optimize a series of antagonist molecules for the efficient inhibition of the protein receptor. Utilizing a series of molecular dynamics simulations, biochemical characterization, and a deep structural analysis, we discovered two small molecules with the ability to reduce the germination of *Striga* by 90%. Ultimately, we believe this approach will be crucial in the fight against crop-invading parasitic plants and the development of targeted strigolactone antagonists.

In silico structural analysis of L-asparagine from Bacillus sp CH11 isolated from Chilca salters in Peru

Annsy C. Arredondo[1], MSayuri Kina-Ysa[1], Carol N. Flores-Fernandez[1], Marvin J. Bayro[2], Amparo I. Zavaleta[1]

[1]Faculty of Pharmacy and Biochemistry, Universidad Nacional Mayor de San Marcos, [2]Faculty of Natural Sciences, University of Puerto Rico

L-asparaginase (EC 3.5.1.1) hydrolyzes L-asparagine into L-aspartate and ammonium. It is used in medicine for the treatment of various types of cancer, and in the food industry to prevent the formation of acrylamide in thermo-processed foods. The aim of this work was to characterize L-asparaginase from *Bacillus sp.* CH11 isolated from Chilca salterns in Peru, by in silico structural analysis. The gene encoding L-asparaginase (*ansA*) was amplified and cloned in *Escherichia coli* DH5α. Then, the nucleotide and amino acid sequences were analyzed using the programs available in the Protein Data Bank. The protein did not present a signal peptide, had a molecular weight of 36.49 kDa, a pI of 4.7, and an instability index of 39.5. The secondary structure showed a conformation with a probability of 78.4%; 49.5% and 12.8% of α helices, folded β sheets and β turns, respectively. In addition, the multiple sequence alignment showed 99% of identity with other *Bacillus subtilis* L-asparaginases. The differences found were Lys4Leu and Thr174Ala. The prediction of the tertiary structure analyzed by Verify3D showed 92.55% of residues <0.2 (3D/1D) and 91.2% of residues in energetically favorable regions according to the Ramachandran diagram. The protein corresponds to a homodimer that includes the residues Thr11, Ser54 and Asp56 at its catalytic site. The in silico analysis of L-asparagine from *Bacillus sp.* CH11 allowed us to know the conformation and structural characteristics of this recombinant protein obtained from an extreme environment.

Studies on the N-terminal RNA-binding domain of the SARS-CoV-2 Nucleocapsid Phosphoprotein

Janet Gonzalez[1], Manfred Philipp[2, 3, 4], Ezekiel Olumuyide[4], Blessing Babalola[5]

[1]Natural Sciences, LaGuardia Community College, 31-10 Thomson Avenue, Long Island City, NY 11101, [2]Chemistry Doctoral Program, Graduate Center of CUNY, [3]Biochemistry, Program, Graduate Center of CUNY, [4]Chemistry Department, Lehman College, Bronx, NY 10468, [5]Biological Sciences, Lehman College, Bronx, NY 10468

We describe an investigation of NMR-based structures of the N-terminal RNA-binding domain of the SARS-CoV-2 nucleocapsid phosphoprotein. These structures were deposited in the 6YI3 pdb file by Veverka, & Boura and were described in (2020) PLoS Pathog 16: e1009100-e1009100. The SARS-CoV-2 nucleocapsid phosphoprotein is of interest because it has been described as a possible drug target that binds both single-stranded and double-stranded viral RNA.

The 6YI3 structure features a beta-pleated sheet region and alpha helical components. 40 NMR structures are present in the 6YI3 structure file, providing considerable variations in backbone and amino acid side chain structures. These variations facilitate the application of computational and molecular visualization tools to observe and measure the dihedral angle backbone movement of the whole domain, as well as by the individual amino acid residues on a bond-by-bond basis. Animations were prepared for the peptide chain and for plots of Phi, Psi, and Omega dihedral angle that were calculated for each model. There was a considerable variability of dihedral angles with amino acid position, even for the omega dihedral angle.

This work was done by undergraduate students using open access software and programs within a Windows/Mac environment. One program used Linux command lines to create models from the PDB file. Students used this work to fulfill the laboratory component of their undergraduate independent study program at in the Biological Sciences and Chemistry Departments of Lehman College, a 4-year college within the CUNY system. This kind of computational work allows undergraduates to complete their required undergraduate research even when college research laboratories are closed due to the current pandemic.

Interface size drives cotranslational assembly of protein complexes

Mihaly Badonyi[1], Joseph A. Marsh[1]

[1]MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh

Assembly pathways of protein complexes must be precise and efficient to minimise misfolding and unwanted interactions with other proteins in the cell. One way to achieve this is by seeding complex assembly during translation via nascent chain engagement. Here, we considered the possibility that the propensity of subunits to cotranslationally assemble is ingrained within the interface hierarchy of protein complexes. Using a combination of proteome-specific structure data and assembly onset positions determined by ribosome profiling, we show that larger interfaces are prioritised in the course of cotranslational assembly. We observe that this effect is not exclusive to homomeric complexes, but appears to drive the assembly of heteromeric subunits to the extent where interface size differences are detectable between N and C-terminal locations, with the former being larger on average. We provide explanations to this phenomenon and discuss its importance in a biological context.

A comprehensive report exploring viroporins as targets for structure-based antiviral development with special emphasis on their electrophysiology and structure determination

Sarika Bano[1], Nishi Jain[1, 2], Saptamita P. Choudhury[1, 3], Arisha Arora[4], Sanjay K. Dey[1]

[1]Dr. B.R. Ambedkar Centre for Biomedical Research, University of Delhi, Delhi-110007, India, [2]Department of Biotechnology, Amity University-Noida, Uttar Pradesh-201313, India, [3]KIIT School of Biotechnology, Kalinga Institute of Industrial Technology, Bhubaneswar, Odisha-751024, [4]Dept. of Biotechnology, Motilal Nehru National Institute of Technology Allahabad, Prayagraj-211004

Viroporins are a diverse group of small, hydrophobic virus-encoded transmembrane proteins that functions as ion-channels. Some of them aid in viral replication and exert pathogenicity thereby holds scope for antiviral drug discovery. Thus, solution and solid state NMR, X-ray crystallography, and single particle cryo-EM have been utilized to resolve three-dimensional structures of viroporins. Considering the significance of this family of viroporin to develop antivirals and availability of biochemical evidence for thousands of viroporins across diverse viral families, the existence of just 12 structures deposited so far in the protein data bank reflects the immense potential for future structural research in this underexplored field. Furthermore, many of these viroporins are yet to be explored in terms of their ion-channel activities and ion-sensitivities to accomplish these goals. Thus, we curated hundreds of infectious viruses containing viroporins or ion-channels from existing literatures and highlighted their scientific proof to be potent therapeutic targets to develop antiviral agents. This will inspire and help scientists to determine the three-dimensional structures of specific viroporins with priority. We have also meta-analysed the importance of specific oligomerization preference by specific viroporins to exert their ion-transport activity and/or ion selectivity at an optimum level. We also reviewed the dual function of these viroporins: one way they help the viral particles thrive by transferring specific ions when present on the viral membranes, and on the other hand they will inhabit the host cell membrane and worsen or alter those cellular functions for the viruses' own survival. Current study will also provide streamlined protocols to express, handle those membrane proteins; and characterise them electro-physiologically, and finally to test blockers/inhibitors both biophysically and biochemically. Combination of all of these efforts can certainly save millions of human lives by curbing viral diseases such as COVID-19, HIV, and others, as well as future viral epidemics, if any. A prime example of the same is the recently solved structure of SARS-CoV-2 envelope protein (PDB-ID:7K3G) to develop potential anti-COVID-19 drugs by blocking its $Ca2+$ ion-channel activity. Our lab is also currently involved in the structural and electrophysiological profiling of some of these viroporins from the present study.

Structure-based functional analysis of BRCA1 RING domain variants: Concordance of computational mutagenesis, experimental assay, and clinical data

Majid Masso[1], Anirudh Bansal[2], Arnav Bansal[3], Andrea Henderson[1]

[1]George Mason University, [2]Thomas Jefferson High School for Science and Technology, [3]Harvard University

A significant impediment to the improvement of clinical outcomes in treating breast and ovarian cancers rests with the lack of available interpretations for BRCA1 variants of unknown significance. Two research groups recently implemented large-scale functional assays for quantifying effects of single missense mutations on homology-directed DNA repair activity of BRCA1 variants, which is critical for tumor suppression and strongly correlates with cancer risk, and their results are significantly concordant with each other as well as with known pathogenic and benign variant clinical data. In this work, we implemented an established computational mutagenesis procedure to characterize structural impacts of single residue replacements to the BRCA1 RING domain. The computational data showed similarly strong concordance with known clinical data as well as with experimental data from both functional assays. Predictions made by models trained on our computational data offer a complementary and orthogonal approach for classifying all remaining unexplored BRCA1 RING domain variants.

Structural and Dynamic Insights into the HNH Nuclease of Divergent Cas9 Species

Helen B. Belato[1], Alexandra D'Ordine[1], Gerwald Jogl[1], George P. Lisi[1]

[1]Department of Molecular Biology, Cell Biology & Biochemistry, Brown University

CRISPR-Cas9 is a widely utilized biochemical tool with applications in biotechnology and precision medicine. In this study, we report the structural and dynamic properties of the HNH domain of the extensively studied Cas9 from Streptococcus pyogenes (SpCas9), and a recently discovered thermophilic Cas9 homolog from Geobacillus stearothermophilus (GeoCas9). Novel constructs of SpHNH and GeoHNH were engineered from their respective full-length proteins, and despite low sequence similarity, the X-ray structures determined for these constructs reveal that the core structure of HNH is highly conserved. However, we detect significant differences in multi-timescale protein dynamics within the structurally conserved regions of SpHNH and GeoHNH, which were recently shown to be critical for biological signaling. Residues experiencing ms-ms dynamics (experimentally characterized with NMR spin relaxation experiments) are present throughout SpHNH, corresponding to a “pathway” that spans the entire domain. GeoHNH lacks nearly all ms-ms timescale motions, and instead displays evidence of ps-ns dynamics in NMR order parameters assessing bond vector fluctuations. Multiple sequence alignment of HNH domains from various Cas9 species highlight conserved residues that surround the catalytic histidine in the core of the structure, but our data indicate that in the case of GeoHNH and the canonical SpHNH, residues that are structurally and sequentially conserved have strikingly different dynamic profiles. We speculate it is possible that GeoHNH and SpHNH share basic similarities of HNH proteins that would be found across all organisms, however the way in which HNH is incorporated into Cas9 proteins can modulates its dynamic properties. Such a circumstance is supported by different dynamic profiles, and likely differences in overall architecture of full-length SpCas9 and GeoCas9 proteins. Thus, even in local environments like the HNH domain where there is structural conservation, differences related to their behavior in full-length Cas9 can still be observed. That is, the intrinsic protein motions that are suggested to regulate Cas9 may be significantly different in the mesophilic SpCas9 and thermophilic GeoCas9 systems, as indicated through studies of their respective HNH domains.

Structural insights into outer membrane protein biogenesis in Neisseria gonorrhoeae

Evan Billings[1], Nicholas Noinaj[1]

[1]Department of Biological Sciences, Purdue University, West Lafayette, IN 47906

Neisseria gonorrhoeae (Ngo), is an obligate human pathogen and the causative agent of the disease gonorrhea. If left untreated, gonorrhea can lead to serious health issues including ectopic pregnancy and infertility. In addition to unsuccessful vaccine development, Ngo has rapidly become resistant to almost all classes of antibiotics. Recent data published by the Center for Disease Control stated over 500,000 drug resistant Ngo infections occur each year in the United States. Due to this widespread antibiotic resistance, the current and only treatment option is limited to a combination of ceftriaxone and azithromycin. However, resistance to even these drugs has emerged and the disease will likely become untreatable in the near future. This has created a desperate need for the development of novel antibiotics and vaccines in order to fight this disease. Like many bacterial pathogens, Ngo is Gram-negative having both an inner and outer membrane. Outer membrane proteins (OMPs) have been recently identified as a novel class of antibiotic and vaccine targets. The biogenesis of these β-barrel OMPs is mediated by a multi-component protein complex, known as the β-barrel assembly machinery (BAM) complex. Conserved across all Gram-negative bacteria, BAM is required for viability and represents a powerful potential antibiotic or vaccine target. In *E. coli*, this 200 kDa complex is comprised of five proteins: BamA, an OMP itself, and four lipoproteins, BamB through E. A clear mechanism for how substrate OMPs are folded and inserted into the membrane by BAM is still poorly understood. My research focuses on understanding the mechanism of BAM from Ngo (*Ng*BAM). Previous studies have indicated that Neisseria do not possess a homolog of BamB in their genome and *Ng*BAM may function as a four component complex. We are currently working to characterize *Ng*BAM using a combination of cryo-electron microscopy and in vivo studies. We recently determined the cryo-EM structure of *Ng*BAM to 6.5 Å. Our structural studies have revealed interesting features distinct from *E. coli*, and we plan to move into in vivo functional studies soon. This work will lay the foundation for characterizing the complex for future vaccine and drug development against gonorrhea.

Exploring Roles of PDB in Protein Structure Prediction: ProSPr and Multi-model Ensembles

Wendy M. Billings,[1,*] Connor J. Morris,[1] Bryce H. Hedelius,[1] Dennis Della Corte[1]

[1] Department of Physics and Astronomy, Brigham Young University, Provo UT 84602
[*] Presenting Author

Many protein structure prediction methods use a form of machine learning known as supervised learning. This approach requires having many examples of data for which the true outcomes are already known. In the case of protein structure prediction, these true outcomes are the experimental structures for given amino acid sequences. The large and ever-growing collection of structures in the PDB has played a significant role in enabling the success of current prediction techniques. Here, we examine the crucial role of PDB structures in our development of ProSPr, a supervised deep-learning method for protein distance prediction.[1] We also outline the augmentation strategies we used to increase the effective number of known examples and explore additional approaches. Although most structure prediction methods are trained using a similar set of protein structures available in the PDB, we demonstrate significant diversity among the contact predictions made by each. We further show that ensembling these diverse predictions made by models of different groups increases overall contact prediction accuracy and outperforms any individual method.[2] These results highlight the continued need for diverse approaches to the problem as well as inter-group collaboration in order to make the most of the protein structure data available in the PDB.
[1] Billings et al., BioRxiv, 2019 (https://doi.org/10.1101/830273).
[2] Billings et al., Nature Scientific Reports, Under Review (https://doi.org/10.22541/au.160317361.15075213/v1).

Biochemical Analysis of the Orf86 Protein in Staphylococcus aureus

Alison E. Bond[1], Faith D. Northern[1], Coy Eakes[2], Krystle J. McLaughlin[1], Matthew R. Redinbo[2], Willam Walton[2]

[1]Department of Chemistry, Vassar College, Poughkeepsie NY 12604, [2]Department of Chemistry, University of North Carolina at Chapel Hill, Chapel Hill NC 27599

Multidrug-resistant *Staphylococcus aureus* infections present an emergent threat to global health. Bacterial plasmids, known for transferring antibiotic resistance, play a major role in the survival of these strains. In the conjugative multiresistance plasmid pSK41, orf86 is a transcription factor believed to play a role in modulating pSK41 replication. We hypothesize that orf86 functions in an analogous manner to the controller protein in Restriction-Modification systems. This study focuses on the structural and functional characterization of orf86, to determine the molecular basis of its role in antibiotic multiresistance transfers.

Identification of a new hotspot of gelsolin instability: biochemical and biophysical characterization of three amyloidogenic variants

Irene Boniardi[1, 2], Matteo de Rosa[1, 2], Toni Giorgino[1, 2], Michela Bollati[1, 2]

[1]Università degli Studi di Milano, Milano, Italy 20133, [2]Istituto di Biofisica CNR, Milano, Italy, 20133

Gelsolin amyloidosis (AGel) is a rare hereditary disease caused by point mutations on the gene encoding gelsolin (GSN)1, an actin-binding protein composed of six homologous domains. Several amyloidogenic GSN mutations in the G2 domain have been discovered over the years, but only the most common mutants (D187N/Y) have been thoroughly characterized2. Recent clinical research has identified three pathological variants located at the interface between the G4 and G5 domains (A551P3, E553K4, and M517R5), suggesting the existence of a new amyloidogenic hotspot distant from the site of the already known cluster of mutations. With this work, we showed that new variants do not follow the canonical amyloidogenic mechanism of variants in G2, while analytical gel filtration and dynamic light scattering suggested an intrinsic tendency of the variants to aggregate and circular dichroism showed a reduction in the variants' melting temperature. Therefore, both X-ray crystallography experiments and 1 µS molecular dynamics simulations of the G4 and G5 domains were conducted to determine the structural implications underlying these biophysical variations. Hence, we suggest that these mutations may destabilize the full-length protein's structure through a shared mechanism based on loss of structural organization of the two β strands of the G4-G5 interface. This would appear to be based on variant-specific molecular implications, and may lead eventually to an aggregation hotspot's exposure, prone to establish aberrant protein-protein contacts.

1. Meretoja J. Familial systemic paramyloidosis with lattice dystrophy of the cornea, progressive cranial neuropathy, skin changes and various internal symptoms. A previously unrecognized heritable syndrome. Ann Clin Res. 1969;1(4):314-324.
2. de la Chapelle A. Gelsolin–derived familial amyloidosis caused by asparagine or tyrosine substitution for aspartic acid at residue 187. Nat Genet. 1992;2(2):157-160.
3. Sridharan M. A Patient With Hereditary ATTR and a Novel AGel p.Ala578Pro Amyloidosis. Mayo Clin Proc. 2018;93(11):1678-1682.
4. Potrc M. Clinical and Histopathological Features of Gelsolin Amyloidosis Associated with a Novel GSN Variant p.Glu580Lys. Int J Mol Sci. 2021;22(3):1084.
5. Cabral-Macias J. Clinical, histopathological, and in silico pathogenicity analyses in a pedigree with familial amyloidosis of the Finnish type (Meretoja syndrome) caused by a novel gelsolin mutation. Mol Vis. 2020;26:345-354.

Structural Characterization by Solution NMR Spectroscopy of a DnaX Mini-intein Derived from Spirulina platensis

Soumendu Boral[1], Woonghee Lee[2], Soumya De[1]

[1]School of Bioscience, Indian Institute of Technology Kharagpur, Kharagpur, West Bengal 721302, India, [2]Department of Chemistry, University of Colorado Denver, Denver, CO 80217, USA

Inteins or intervening proteins are enzymes that catalyze their excision from a precursor protein and facilitate ligation of the flanking external proteins (exteins). To date, >600 intein genes have been discovered, but only a few have been thoroughly characterized. Due to their wide range of applications, it is important to characterize more intein sequences that may find novel applications.
We have solved the solution NMR structure of a 136-residue intein enzyme derived from DNA polymerase III (DnaX) gene of a cyanobacterium *Spirulina platensis*.[1] An ensemble of 15 structures has a backbone RMSD of 0.27 Å and heavy atom RMSD of 0.52 Å (PDB code: **7CFV**) across all ordered residues.[2] Backbone Φ and Ψ torsion angles of 97.1% residues are in favored regions of the Ramachandran plot. None of the residues are in disallowed regions. Although the level of sequence identity is low (<33%), *Spl* DnaX mini-intein has very good structural homology with other inteins. The *Spl* DnaX mini-intein has a highly symmetric structure, comprising of 13 β-strands, one α-helix, and one 3_{10}-helix, which are arranged in a compact horseshoe-shaped fold commonly found in the HINT (hedgehog intein) domain superfamily. NMR-based hydrogen exchange experiments reveal the presence of a highly stable core in the *Spl* DnaX mini-intein. Backbone ^{15}N-dynamics experiments show the presence of conserved motions. Also, these dynamic motions were detected in symmetric positions within the intein structure, which is most likely a result of the symmetrical structure of the enzyme.
To assess *in vivo* activity, the intein enzyme is fused with two small soluble proteins at its N- and C-termini. Western blots and SDS-PAGE are used to detect the spliced products. The precursor protein is not detected at any timepoint of the assay, indicating rapid catalysis of the splicing reaction immediately upon synthesis in the cell. The small size, high stability, and high catalytic activity make the *Spl* DnaX intein enzyme an ideal candidate for protein engineering that may lead to novel applications.

References:

1. **Boral, S.** et al., *Biochemistry*, **59** (50), pp. 4711-4724. (https://doi.org/10.1021/acs.biochem.0c00828)
2. **Boral, S.** et al., *RCSB Protein Data Bank*, DOI: 10.2210/pdb7CFV/pdb (https://www.rcsb.org/structure/7CFV)

Autoinhibition of Von Willebrand Factor A1 domain: The story unfolds. A 2.1Å crystal structure of VWFA1 domain - Caplacizumab complex

Alexander K. Brown[1], Jonas Emsley[1], Renhao Li[2], Nicholas Arce[2]

[1]University of Nottingham, Nottingham, NG72RD, [2]Emory University, Atlanta, GA 30322

Background – The first ever approved nanobody therapy – Caplacizumab targets the discontinuous autoinhibitory module (AIM) of Von Willebrand Factor (VWF) which consists of flanking regions of the A1 domain that unfold under shear force and activate VWF. A mechanism for this has been suggested but no structural evidence exists.
Aims – To capture VWF-A1 domain in a novel conformation using the monomeric form of Caplacizumab - VHH81 and solve the associated structure.
Methods - Single crystal X-ray crystallography was used to determine the structure of a complex of VHH81 and a VWF-A1 fragment containing residues 1238-1481 to a 2.1 Å resolution. The structure was solved using homology models from both existing structures as well as a novel prediction from the PHYRE2 server.
Results – In the structure, many residues, including residues 1262-1267 and some in the A1 domain, are in direct contact with all three complementarity-determining regions (CDR) loops of VHH81. For instance, VWF residue R1274 forms a salt bridge to the side chain of E105 in CDR3, and residues 1262-1267 pack around Y32 in CDR1. Also pertinent to this study is the hydrogen bond between side chains of E1264 and R1341, which has not been observed in any previous structures of A1. Moreover, comparison of the AIM-A1/VHH81 complex structure to previously reported A1 structures, for example when bound to GPIbα (PDB:4C2B) reveals that the largest difference in VWF conformation lies in the α1β2 loop, as well as N and C-terminal AIM residues, such that these residues appear to move away from the α1β2 loop upon binding of GPIbα. In the AIM-A1/VHH81 complex, with residues 1463-1466 in position close to the α1β2 loop, the unresolved residues beyond 1466 (i.e. residues 1467-1481) would clearly interfere with GPIBα binding to A1.
Conclusions – We provide the first structural evidence of the tensile conformation changes that activate VWF-A1 as well as demonstrating how VHH81 exploits this to non-competitively inhibit the binding of VWFA1 to GPIbα.

Exploring the Mechanisms of Antiviral Therapeutics: A 3D Model of SARS-CoV-2 RdRp in Complex with Remdesivir

Michael Davis[1], Tra'Mya Lauderdale[1], Ryan Billings[1], LeShaundria Brown[1], Valeanna Adams[1], Serena Barnhill[1], Ntirenganyi Karamba[1], Melanie Van Stry[1], Candace Jones-Carter[1]

[1]Lane College, Jackson, TN, 38301

The recently discovered severe acute respiratory syndrome coronavirus 2 (SARS-Cov-2) causes coronavirus disease 2019 (COVID-19), a respiratory disease affecting the human population worldwide. SARS-CoV-2 replication requires the RNA dependent RNA polymerase (RdRp) complex, composed of non-structural proteins (nsp) 7, 8, and 12. Nsp 7 and 8 function as a primase, whereas nsp12 functions as RdRp for replication and transcription. This polymerase is the target of the antiviral drug remdesivir, an adenosine monophosphate analog. During RNA replication catalyzed by RdRp, remdesivir is covalently attached to the growing RNA strand, resulting in chain termination. Here, we designed a 3-dimensional (3D) model of the SARS-COV-2 nsp12-nsp7-nsp8 RdRp complex bound to template-primer double-stranded (ds)RNA and remdesivir using Jmol and PDB 7BV2 cryo-EM structure (Yin et al., 2020). The 3D model shows the nsp12 subunit bound to dsRNA template and growing RNA strand that forms a corkscrew-like structure within the center channel. The model highlights specific residues aspartic acid 760, valine557, and serine861 within the active site and the interactions of the template-primer RNA strands with remdesivir. An additional 3D model illustrates the structural similarity of remdesivir to adenosine monophosphate. These 3D models enable students to visualize complex biomolecules and understand mechanisms of therapeutics. Support or Funding Information NSF-DUE 1725940 and NSF-DUE 1833960

Educational Tools for Drug Discovery and Design Focusing on the SARS-CoV-2 Main Protease nsp5

Dalton M. Dencklau[1], Michel P. Evertsen[1], Isabelle G. Juhler[1], Kallista A. Larson[1], Alexa S. Pinos[1], Jesse D. Rothfus[1], Bonnie L. Hall[1]

[1]Department of Chemistry and Physics, Grand View University, Des Moines, IA 50316

A new coronavirus causing an acute respiratory syndrome (SARS-CoV-2) emerged in late 2019 and has been responsible for the outbreak of coronavirus disease 2019(COVID-19). Symptoms of COVID-19 range from mild to severe, including respiratory symptoms, systemic inflammatory responses, and even death. Currently,remdesivir is the only FDA-approved therapeutic agent to treat COVID-19, although it shows limited efficacy. We were interested in learning about the drug designprocess for a novel virus. We focused on drug research targeting the SARS-CoV-2 main protease, known as nsp5 or Mpro. The nsp5 protease is an enzyme critical forviral replication, with ebselen, cinanserin, and N3 being lead compounds identified as possible effective drugs (Jin, 2020). Starting with these chemical drug structures,we used an in silico drug design process to explore how to modify and refine these drugs. Our goal was to make the compounds more drug-like, according to Lipinski'sand Veber's rules. Using the crystal structure file 6YB7 from the Protein Data Bank (PDB), we focused on the protease active site, defined by amino acids His41,Met49, Asn142, Cys145, His164, Met165, Glu166, Pro168, and Gln189. Cys145 is especially important, as it is the target for irreversible covalent inhibitors. Based onthis information, we used molecular docking software to first re-dock these drugs into the nsp5 protease. We then made changes to the structures of the leadcompounds, docked those new compounds, and used the energetic information to continue the refinement process. Once we completed our own drug designexplorations, we created a three-dimensional printed model of the full nsp5 protease, with the key amino acids highlighted. We also created a three-dimensional modelzoomed in on the active site, with the different drugs fitting into that active site model. Here, we present what we learned about the drug design process for a novelvirus such as SARS-CoV-2. Our docking protocols and models are useful for teaching the fundamentals of drug design and about modern drug discovery and design processes.

Three-Dimensional Visualization of Remdesivir and the SARS-CoV-2 RdRp

Asal Eid[1], Karshana J. Kalyanaraman[1], Binh Nguyen[1], Sierra Dale[1], Laura A. Rusch-Salazar[1]

[1]Grand View University

The SARS-CoV-2 virus, which causes the COVID-19 disease, has been the focus of concern around the world, setting records as a modern-day pandemic. The World Health Organization (WHO) has confirmed over 111 million cases in the world with over 2.5 million confirmed deaths as of February 23, 2021. SARS-CoV-2 continues to be a serious threat to global public health, leaving a distinct footprint in human history. Our research visualizes and contextualizes the mechanism of remdesivir, an antiviral nucleotide analog prodrug. Remdesivir targets the viral RNA-dependent RNA polymerase (RdRp) to inhibit replication of the virus. This process is facilitated by a subunit replication-and-transcription complex of three major nonstructural proteins (nsps): nsp7, nsp8, and nsp12. With our three-dimensional model (PDB ID 7BV2), we depict the mechanism of remdesivir inhibition of the copying of the viral RNA template through the central channel of RdRp by terminating chain elongation. When deciding the project, remdesivir showed promise as an inhibition drug for SARS-CoV-2, as it was effective for MERS-CoV. The National Institutes of Health (NIH) and the WHO have conflicting results about the effectiveness of remdesivir for treating COVID-19. Further comparison between the RdRp structures of SARS-CoV-2 and MERS-CoV may elucidate why remdesivir may be ineffective and could lead to better drug design. In our RNA model, the hydrogen bonds between the RNA primer strand and remdesivir are highlighted to show how remdesivir acts as a nucleotide to inhibit further viral replication. In the full RdRp complex, the hydrogen bonds between remdesivir and key residues of the nsp proteins are highlighted. Our model illustrates the viral RdRp of SARS-CoV-2 and how remdesivir should impact viral RNA replication through inhibition.

Synthesis, Characterization and Biomedical Application of Ferrocene-Hormone Complexes

Mariola M. Flores-Rivera[1], José A. Carmona-Negrón[1], Dalice M. Dalice-Piñero[2], Enrique Meléndez[1]

[1]Department of Chemistry, University of Puerto Rico, Mayagüez Campus, [2]Department of Chemistry, University of Puerto Rico, Rio Piedras Campus

Approximately 52% of these cases are related to over-expression of estrogen receptor (ERα). Conventional metal-based therapeutics drugs, such as cisplatin and derivatives, are still used in order to inhibit this abnormal cellular proliferation rate. However, cisplatin-based drugs are highly cytotoxic, triggering a series of side effects that become detrimental to the body, due to their lack of selectivity between healthy and cancerous tissue. In 1984, Köpf-Maier, and his co-workers first reported the anticancer properties of ferrocene. This organometallic compound leads to the formation of radical oxygen species that cause oxidative damage to DNA, inducing apoptosis of the cell. Our research group has recently successfully incorporated ferrocene with estrone, and estradiol at estrogen's rings A and D. These ferrocene complexes have been found to show micromolar cytotoxic activity similar to conventional therapeutic agents such as cisplatin and tamoxifen.8 This work seeks to deliver a new approach to enhance the selectivity to effectively target hormone-dependent cancers, specifically, ER+ breast cancer. We present a series of novel ferrocene-hormone complexes with its characterization by X-Ray Diffraction (XRD). Computational studies of the interaction of the ferrocene-hormone complexes with ERα protein were performed which demonstrated the possibilities of docking interaction of these drugs in the ligand binding pocket of the ERα.

Characterization of a Putative Short Chain Dehydrogenase from M. smegmatis

Joseph Harrington[1], Krystle J. McLaughlin[1]

[1]Department of Chemistry, Vassar College, Poughkeepsie NY, 12604

Proteins from the bacterium responsible for Tuberculosis, Mycobacterium tuberculosis, have historically been difficult to crystalize and study. This has been a major hurdle in efforts to combat one of the leading causes of death worldwide from a single infectious agent. Mycobacterium smegmatis is a non-pathogenic mycobacterium with a similar proteome to M. tuberculosis, sharing almost 2000 homologous genes. Thus, structural and functional characterization of the M. smegmatis proteome is of interest as it can enable the discovery of new M. tuberculosis treatments. This study is focused on the structural and biochemical analysis of a putative short chain dehydrogenase from M. smegmatis.

The Flavoprotein Database Houses Structural and Thermodynamic Data and Reveals Insight into the Structure-Function Relationship in Flavoenzymes

Emily Kang[1], Sophia Heo[1], Hao Huang[3], Joshua Li[1], Andrew McGill[2], Owen O'Neil[1], Ethan Parr[1], Yi Lin Sim[1], Yixin Xao[1], Catherine Zhu[1], Bruce Palfey[1]

[1]University of Michigan, Ann Arbor, MI, [2]The University of Texas at Austin, Austin, TX, [3]WuHan University of Science and Technology, WuHan, China

The Flavoprotein Database is a graph database and web interface designed for flavoenzymologists to investigate structural, thermodynamic, and spectral properties of flavoproteins. At the active site of each flavoenzyme is an isoalloxazine moiety, a three-ring heterocycle that enables versatile chemistry including redox chemistry and covalent catalysis. Protein Data Bank (PDB) structures of flavoproteins were mined for information about the protein matrix surrounding this flavin prosthetic group. Each isoalloxazine atom and every neighboring atom within a 5-angstrom distance are unique nodes in the flavoprotein graph. Users can search for enzymes by isoalloxazine interactions, solvent-accessible surface area for each atom of isoalloxazine, torsion angles of the ribityl chain of bound flavins, reduction potentials of enzymes and substrates. Reduction potentials and absorbance spectra were mined from the literature, lists of reactions were mined from the KEGG and Uniprot databases, and binding affinity data were mined from the PDBbind database and BindingDB. The graph currently holds data on 3,698 PDB structures, 2,106 small molecules, and 655 flavoenzyme reduction potential values for 486 flavoproteins.

The Flavoprotein Database is a resource that biochemists and other researchers can use to gain a better understanding of flavin-chemistry and flavoenzyme mechanisms. We are using the database to explore questions such as: how does the protein matrix surrounding the flavin influence its reduction potential? What effects does the electrostatic environment have? How do reduction potentials of substrate/product couples compare to reduction potentials of their corresponding flavoenzymes? In preliminary analysis we have observed trends of protein structures clustering by reaction type. A greater understanding of the structure-function relationship of these enzymes could lead to advancements in biocatalysis, medicine, and environmental restoration.

Remdesivir and the RNA-dependent RNA polymerase of SARS-CoV-2

Muskan Kanungo[1], Amber Sabin[1], Evan Connelly[1], Natalie Evans[1], Ryan Nickel[1], Keagan Schmidt[1]

[1]Milwaukee School of Engineering, Milwaukee WI 53202

COVID-19 is a contagious disease caused by the SARS-CoV-2 strain of coronavirus, a variant of the 2002 SARS pathogen. Scientists began studying the effects of known antiviral medications on COVID-19 in hopes of quickly finding a suitable drug. One such medication was remdesivir, which was previously found to be somewhat successful in treating SARS and MERS coronaviruses in clinical trials. Although authorized for emergency treatment, it was deemed ineffective after studies found that remdesivir doesn't affect the mortality rate of patients with COVID-19 in a statistically significant way.

Remdesivir is a competitive inhibitor that resembles adenosine and works by binding to the RNA-dependent RNA polymerase (RdRp) that coronaviruses use to replicate their genomes. Remdesivir then causes an instability that terminates the RNA replication 3 base pairs upstream from the binding site, thus inhibiting the RdRp's function. The RdRp of SARS-CoV-2 is made up of three non-structural proteins: NSP7, NSP8, and NSP12. NSP7 and NSP8 act as coenzymes to NSP12, allowing the conformational changes needed in the protein so that it can properly bind to the RNA bases. The RdRp also contains two magnesium ions in the active site that help stabilize the negative charge from the RNA backbone. Remdesivir itself has a few key differences from adenosine, the most prominent being a cyano group (CN) attached to C1' of the ribose sugar. It is this group that is thought to cause the structural instability that terminates RNA replication.

Using the protein databank to visualize the effect of minor amino acid differences on the function of SARS- CoV monoclonal antibodies

Sanjana Likki, Feza Abbas, Lyla Abbas, Carolina Alzamora, Matthew Hunt, Pujita Julakanti, Carol Mannikkuttiyil, Christo Mannikkuttiyil, Emily Schmitt Lavin, Arthur Sikora

Nova Southeastern University

To visualize how minor differences in amino acids can affect function in SARS-CoV antibodies, undergraduate students cooperated in teams as part of the CREST (Connecting Researchers, Educators, and STudents) Program with the Center for Biological Modeling. Students from 8 universities collaboratively applied their knowledge to build 3D printed models to explain a particular protein-based molecular story using protein crystal structures deposited in the Protein Data Bank (PDB). Within the CREST program, the PDB is utilized to find the crystal structures of different proteins and/or antibodies related to a yearly theme, ideally to the highlighted scientist's laboratory. Teams survey the range of available structure files, conduct further research, and determine which structure to base their model on. Students utilized protein visualization programs (Pymol and Jmol) to configure models depicting the 80R (PDBID:2GHW) and 362 antibodies binding to their corresponding spike proteins. Using data on point mutation differences between these 2 antibodies, a potentially more universal antibody (NSU1) was designed. This molecule is expected to bind more effectively to future SARS spike protein mutants. Due to the trend of smaller amino acids in the MAb362 binding interface, it was hypothesized that more space in this area could allow antibodies to be more resistant to future SARS-CoV spike protein structure variations. NSU1 was based on MAb362 with 4 additional mutations: Asp103Gly, Trp104Leu, Gly170Ser, and Arg211Val. All except for Gly170 are mutations that decreased size and polarity of amino acid residues within the binding interface. Position 170 is Asp on the 80R structure; thus a Ser mutation is still expected to maintain the smaller antibody residues trend. This additional space created by amino acid substitutions in the binding region between the antibody and spike protein's RBD, NSU1 is predicted to be more resistant to spike protein mutations. The PDB structures allowed for a much deeper understanding of the impact that mutations have on binding interactions with viral proteins. The visualization and modeling tools available provide insight into the molecular structure. Leading to a potentially more adaptable antibody against variations in SARS-CoV.

Mammalian F0F1-ATP synthase can produce and hydrolyze inorganic polyphosphates – a view from molecular docking experiments

Albert R. Makhmudov[1, 2], Kamila S. Nebesnaya[1, 3], Artyom Y. Baev[1, 3]

[1]Laboratory of Experimental Biophysics, Centre for Advanced Technologies, Tashkent, Uzbekistan, [2]Faculty of Chemistry, National University of Uzbekistan, Tashkent, Uzbekistan, [3]Department of Biophysics, Faculty of Biology, National University of Uzbekistan, Tashkent,Uzbekistan

Inorganic polyphosphates (polyPs) are the ancient homopolymers that consist from orthophosphate residues linked together via high-energy phosphoanhydride bonds like in ATP molecule. This molecule plays a crucial role in all living organisms. Enzymes responsible for synthesis and degradation of polyP in microorganisms and unicellular eukaryotes have been isolated, purified, cloned and using nowadays as genetic tools for polyP level control in cells of interest. Homology search haven't shown any enzymes similar to bacterial in mammalian genome. Thus, taking into account a big role of polyP in physiological and pathophysiological processes in mammalian cells, one of the main question regarding metabolism of polyP in mammalian cells is an enzyme that synthesize this polymer.
Recently we showed that one of the possible enzymes that could synthesize polyP is mitochondrial F_0F_1-ATP synthase. This unique enzyme is responsible for the synthesis of majority of cellular ATP. In our study we showed that polyP, ATP alike, can be produced in the F_0F_1-ATP synthase and as well can be used as an energy source in the F_0F_1-ATPase. To understand whether polyP can bind to catalytic sites of F_0F_1-ATP synthase we performed docking experiments. For this purpose, we used bovine F_0F_1-ATP synthase in three conformational states, structure of which was published in protein data bank in 2020 (6ZPO, 6ZQM, 6ZQN). Receptor and ligand preparation was performed in UCSF Chimera, for docking we used Autodock Vina. Our calculations showed that SpolyP (14 orthophosphate residues in a chain) selectively binds to the catalytic site of F_1 with $\Delta G_{binding}$ varying from -9.2 to -10 for different receptors. Original structures of ATP synthase (6ZPO, 6ZQM, 6ZQN) contained ADP and ATP in active sites, which we deleted during the process of receptor preparation. To understand that our calculations are adequate we performed docking experiments with ADP molecule. When we merged our docking experiments with original structures, we saw that docked ADP superimposed ADP from the original structure, the position of docked ADP differed from the original one less than 1 Å. Thus, this study provides evidence of a polyP binding with mitochondrial F_0F_1-ATP synthase.

Molecular Dynamics Simulation of 6PEY.pdb a Novel Mutation in the Enzyme Methylenetetrahydrofolate Reductase.

Nicholas Mamisashvili[1], Daniel Williams[1]

[1]Shelter Island High School

Methylenetetrahydrofolate reductase (MTHFR) is a protein that has been implicated in many human blood clotting diseases. We crystallized a novel mutant of MTHFR and solved the x-ray diffraction structure by molecular refinement and the structure was deposited in the Protein Data Bank (6PEY). This structure, however, is based on a frozen protein, not in its natural aqueous environment. GROMACS (GROningen MAchine for Chemical Simulations), is a molecular dynamics software that enables molecular models from the Protein Data Bank to be simulated in their natural aqueous environment. Through GROMACS, solvent and ions were added to the PDB model, while energy was minimized. The published crystal structure was then compared to the GROMACS simulated structure examining changes in RMSD using xmgrace plotter and Pymol to examine the structural variations. We observed differences in the molecular dynamics simulation with the novel mutation's interaction with the ligand potentially affecting it's binding affinity. This supports published binding affinity studies on this mutation.

Mechanistic Investigation of HNO Formation from Hydroxyurea Catalyzed by Horseradish Peroxidase

Erika R. McCarthy[1], Yong Zhang[1]

[1]Department of Chemistry and Chemical Biology, Stevens Institute of Technology, Hoboken, New Jersey

Nitroxyl (HNO) plays an important role in many biological processes related to vasodilation, enzyme activity regulation, and neurological function regulation. However, the understanding of potential in vivo HNO releasing pathways from HNO precursor drugs is relatively scarce. The sickle cell anemia treatment hydroxyurea is reported to be an HNO donor; however, the mechanism of generation of HNO from the drug is thus far unclear. A quantum chemical investigation was performed to provide the first detailed mechanism of HNO generation from hydroxyurea catalyzed by hydrogen peroxide treated horseradish peroxidase (HRP). The results are consistent with experimental studies and provide insights into the electronic driving force, energetics, and structural features that govern this reaction. These results could promote future development of more effective HNO donor drugs.

Identification of Interface Residues of the Rab6/BicD2 Complex

David Moraga[1], Sozanne R. Solmaz[1]

[1]Binghamton University, Binghamton, NY 13905

Rab6 is a GTPase that is embedded in distinct secretory and Golgi-derived vesicles, where it recruits factors involved in vesicle transport and fusion. One of these proteins is the dynein adaptor BicD2, a homodimeric coiled-coil, which forms a complex with Rab6, thereby linking Rab6-decorated vesicles to the minus-end directed motor dynein for transport of the vesicles. How these two proteins interact remains unknown. In this work, a cluster of several Rab6 key interface residues was identified, which is important for the interaction with BicD2. We show that mutation of these residues abolishes binding to BicD2. These residues are also highly conserved and are located in a region that undergoes structural changes upon GTP-binding. Our preliminary data promises to reveal key mechanistic details about the poorly understood interaction between Rab6 and BicD2.

Protein Structure Refinement with Restrained Molecular Dynamics

Connor J. Morris[1], Wendy M. Billings[1], Dennis Della Corte[1]

[1]Brigham Young University, Provo, UT

Protein structure refinement is an important part of the computational prediction of protein structures from their amino acid sequence. The goal of refinement is to improve structures generated by various computational structure prediction methods. We recently created a structure refinement method and benchmarked it against other top groups in the CASP14 protein structure prediction competition. Our method takes a protein structure predicted by other groups as input, runs a set of molecular dynamics simulations with flat-bottom harmonic restraints on C-alpha atoms, then generates an averaged structure from the frames of the molecular dynamics simulation with the best energy scores. In the refinement category of CASP14, our group ranked 1st in refinement of structures with less than 100 amino acids, 1st in estimation of residue-wise error, and 3rd place overall.

Structural Analysis of Thioredoxin Reductase from Mycobacterium smegmatis

Eden O'Connell[1], Krystle McLaughlin[1]

[1]Vassar College, Poughkeepsie NY 12604

Mycobacterium smegmatis is a close relative of Mycobacterium tuberculosis, but is itself non-pathogenic. M.smegmatis is a useful model for behaviour of M.tuberculosis, and other members of the Mycobacterium family, and developing a better understanding M.smegmatis will allow for more effective modelling. Thioredoxin reductases are the enzymes responsible for the reduction of the redox protein thioredoxin, they consist of two binding domains, one for NAD+ and one for FAD. In this study the structure of M. smegmatis thioredoxin reductase is characterized.

X-ray and Antioxidant Determination of Butein and 2',4'-dihydroxy-3,4-dimethoxychalcone to Examine their Antimalarial Activity by Binding to Falcipain-2 (PDB Code: 3BPF).

Ijeoma P. Okoye[1], Miriam Rossi[1], Francesco Caruso[1]

[1]Department of Biochemistry, Vassar College, Poughkeepsie, NY 12604

Malaria is a huge global health burden as annually, 300-500 million people are infected with malaria and 1-3 million people worldwide die from malarial infections. Currently, there are no 100% effective vaccines against malaria and the antimalarial drugs in use are now becoming less effective as malaria parasites are gaining resistance to them. Interestingly, some of the most effective antimalarial drugs such as chloroquine and artemisinin combination therapies (ACTs) are derived from traditional medicinal plants that are typically local to malaria-endemic regions. This study focuses on Erythrina abyssinica, a leguminous tree that is native to Southern Africa, Eastern Africa, and Democratic Republic of Congo, and is used to treat malaria in these regions. Previous studies have shown that the stem extract of E. abyssinica contains two compounds, 5-prenylbutein and homobutein, that have moderate antiplasmodial activity. Thus, the objective of this project was to structurally characterize butein and 2',4'-dihydroxy-3,4-dimethoxychalcone (another chalcone structurally similar to butein) to elucidate possible molecular mechanisms by which these compounds clear malaria parasites and to also explore their potential as starting points for development of novel antimalarial drugs. Single Crystal X-ray Diffraction technique was used to determine the 3D crystal structure of both compounds and Rotating Ring-Disk Electrode Voltammetry was used to measure the antioxidant activity of both compounds. We hypothesized that both compounds would chemically interact with the cysteine residue in the catalytic site of Falcipain-2 (a cysteine protease that has been demonstrated to be involved in the degradation of hemoglobin in malaria-infected RBCs to allow for the survival and growth of malaria parasites). Results from the diffraction analysis and subsequent docking simulations showed that both butein and 2',4'-dihydroxy-3,4-dimethoxychalcone are in the right orientation and distance to allow for strong binding to the cysteine residue of Falcipain-2. The RRDE results also demonstrated that butein and 2',4'-dihydroxy-3,4-dimethoxychalcone have high antioxidant activity. Thus, both compounds could serve as potential competitive inhibitors of Falcipain-2 via their iron chelation ability, high antioxidant activity, and strong affinity to the active site of Falcipain-2.

Small Molecule Inhibition of CRISPR Cas9 (4un3.pdb)

Theodore Olinkiewicz[1], Daniel Williams[1]

[1]Shelter Island High School, Shelter Island, NY 11964

The Protein Data Bank is an incredible resource for computer based molecular simulations that can be executed in a high school classroom setting, including modeling of Cas9 inhibitors. Cas9 (4un3), an enzyme found in the CRISPR gene editing system has the ability to cut and edit DNA based on a programmable guide RNA sequences. Small molecule inhibitors of Cas9 are able to control the molecule turning off nuclease activity, making it more precise for gene editing. There are many known inhibitors of the Cas9 enzyme, however they are large and difficult to work with. Artificially designed small molecule inhibitors have been proposed and work under laboratory conditions, however, it is unknown where they bind and how they work. Autodock is a computer program that simulates the binding of molecules using Protein Data Bank files. Through Autodock simulations, these small molecule inhibitors were tested for binding affinity and location, providing potential insight to their mechanism of inhibition.

Enzymatic and Structural Analysis of MsFabG4

Prashit Parikh[1], Krystle J. McLaughlin[1]

[1]Department of Chemistry, Vassar College, Poughkeepsie NY, 12604

Fatty acid metabolism has been linked to streptomycin resistance in *Mycobacterium tuberculosis* and the consequent onset of tuberculosis. Tuberculosis is a leading cause of death worldwide, thus a better understanding of its basic biosynthetic pathways would provide valuable insight. FabG4 catalyzes the reduction of β-oxoacyl-ACP to β-hydroxyacyl-ACP in fatty acid synthesis. This study focuses on *Ms*FabG4, a homologous protein from *Mycobacterium smegmatis*. More specifically, the enzymatic activity of *Ms*FabG4 in the presence of various electron carrier molecules and its structure will be highlighted.

Do Economic Factors Affect the Numbers of Coronavirus PDB Protein Structure Determinations?

Janet Gonzalez[2], Matthew K. McDevitt[1], David Roman[1], Manfred Philipp[1,3,4]

[1]Chemistry Department, Lehman College, Bronx, NY 10468, [2]Natural Sciences, LaGuardia Community College, 31-10 Thomsom Avenue, Long Island City, NY 11101, [3]Chemistry Doctoral Program, Graduate Center of CUNY, [4]Biochemistry Program, Graduate Center CUNY

The current worldwide coronavirus pandemic follows on other, less widely distributed coronavirus epidemics that affected some, but not all countries. This study uses the deposition of pdb structures as an indicator of how countries react to the pandemic and compares the current response to the responses to the past coronavirus epidemics, in particular SARS and MERS.

This study uses population data as well as national and regional Gross Domestic Product metrics to analyze the factors that contribute to the generation of pdb files that describe coronavirus proteins and nucleic acids. Are there correlations between these economic and demographic metrics and the output of scientific data related to past and current epidemics? The output of pdb files provides a controlled and reproducible metric that can be related to these other factors. All are contained in publicly available formats.

This work was done in part as an effort to provide research experiences to chemistry undergraduate students who, due to the pandemic, cannot acquire necessary research experience in college laboratories. Such laboratories were closed by the current pandemic. It was important to use freely available data and analytical tools, tools that will run on commonly available desktop computers under Windows 10, as well as tools available on freely available web sites.

An Examination of the Torsion Angle Variability in a Viral RNA Genome Domain Bound to Retroviral Protein

Janet Gonzalez[3], Ezekiel Olumuyide[4], Manfred Philipp[1,2,4,] Jarlene Roman[4]

[1]Chemistry Doctoral Program, Graduate Center of CUNY, [2]Biochemistry, Program, Graduate Center of CUNY, [3]Natural Sciences, LaGuardia Community College, 31-10 Thomson Avenue, Long Island City, NY 11101, [4]Chemistry Department, Lehman College, Bronx, NY 10468

Solution NMR determinations can provide unique experimental insights into the internal mobility of nucleic acids. The 2MQV pdb file that has been published for the U5-primer binding site of murine leukemia virus RNA provides data on such a topic of study. This pdb file was deposited by D'Souza and Yildiz and published in Nature, 2014 Nov 27;515(7528):591-5.

In the current study, the ten models contained in the pdb file were each examined by calculating torsion angles using a server described in Nucleic Acids Research, 2019, 47, W1, W26–W34.

Graphs of torsion angle vs sequence were animated, and three-dimensional images of the RNA hairpin loop were used to prepare animations that illustrate protein and RNA motion. These investigations demonstrate how the interaction with protein affects the change in RNA structure as described by molecular motion and changes in the RNA torsion angles.

Biochemical Analysis of a Putative Chitin Deacetylase from Bacteroides ovatus

Lilith A. Schwartz[1], Krystle J. McLaughlin[1]

[1]Department of Chemistry, Vassar College, Poughkeepsie NY, 12604

Bacteroides ovatus is a microbe involved in complex carbohydrate catabolism within the human gut microbiome. *B. ovatus* is implicated in a variety of autoimmune disorders, but not much is known about this connection. The crystal structure of a novel putative carbohydrate esterase from *B. ovatus*, BoCDA, was recently solved. Based on sequence analysis, structural characterization, and enzymatic assays we demonstrate that BoCDA is a member of the carbohydrate esterase 4 (CE4) superfamily, and a putative chitin deacetylase. We also investigate a jelly-roll motif within the structure whose function is unknown. Based on enzymatic assays, the enzyme requires the presence of nickel or zinc to perform deacetylation of chitin. Studying BoCDA's function as a chitin deacetylase may advance our understanding of its role in the gut microbiome.

S309 Binding to SARS-CoV-2 Spike Protein

Behrgen P. Smith[1], Anna Debruine[1], Jude Ingham[1], Alexa De La Sancha[1], Delila Payton[1], Krishna Persaud[1], Josephine Selvik[1], Everen Wegner[1], Joshua Zieman[1]

[1]Milwaukee School of Engineering, Milwaukee, Wisconsin 53202

S309 is a monoclonal antibody (mAb) that potentially neutralizes the SARS-CoV-2 virus, a respiratory coronavirus, causative agent of the COVID-19 pandemic. mAbs differ from vaccines in that they are a treatment and do not activate antibody production or confer long-term immunity. Due to the conserved structures across SARS-CoV variants, an understanding of protein-mAb interactions and subsequent impacts on disease offer a pathway to developing future treatments for novel vaccine resistant variants. S309 was designed from an antibody that inhibits the spike protein of SARS-CoV, a transmembrane protein that mediates entry into a host cell, which makes this monoclonal antibody cross-reactive for SARS-CoV-2 as well. The S309 mAb binds noncompetitively to a conserved glycan (sugar) on the spike protein as opposed to the ACE-2 receptor binding motif. Along with the glycan, S309 recognizes several amino acid residues within the epitope, binding and subsequently neutralizing the virus, although the specific conformational change that leads to neutralization is unknown. Using Jmol, we were able to identify significant intramolecular interactions by isolating amino acid residues that were involved in the binding of the antibody to the spike protein. Significant interactions were defined as being within a proximity of five Å to other binding residues and its type was determined by the strongest attraction. Of the major interactions: two hydrophobic pockets, one ionic bond, two hydrogen bonds and two dipole-dipole interactions were considered integral components of the S309-spike protein binding, along with three LDF interactions that position the glycan.

Improving Molecular Docking Programs through Consensus Docking with Machine Learning

Brenden Stark[1], Connor Morris[1], Dennis Della Corte[1]

[1]Brigham Young University

Molecular docking programs are computational tools widely used in modeling protein-ligand interactions, especially during drug discovery. These programs predict the binding pose and energy for the interaction and rate ligand conformations using a scoring function, but they can at times produce inaccurate predictions. One method to improve the use of these programs is consensus docking, which evaluates protein-ligand binding using multiple docking programs and/or scoring functions to provide a more accurate model of the protein-ligand complex. We present a process using machine learning algorithms to improve the results and simplify the analysis of consensus docking.

Deep dive into the giant gutless beard worm

Olivia Krajewski[1], Kara Speakman[1], Pachoua Vang[1], Trisha Xiong[1], Kathleen Boyle[1], Colleen Conway[1], Maureen Leonard[1]

[1]Mount Mary University, Milwaukee, WI 53222

Hemoglobin is an essential, iron-containing protein found in the blood of a wide variety of species including vertebrates, invertebrates, fungi and plants. For most vertebrates, hemoglobin functions to carry oxygen in red blood cells. In humans, hemoglobin is composed of four subunits which work in concert to transport oxygen from the lungs to tissues. However, in invertebrates, the structure and function of hemoglobin may vary. Both the size (monomer to oligomer) and function (transportation of other small molecules such as carbon dioxide, nitric oxide, hydrogen sulfur, or sulfur) of hemoglobin are variable. Some of the most interesting hemoglobin molecules are found in worms that live at the bottom of the ocean. The marine gutless beard worm (*Oligobrachia mashikoi*) was discovered thriving in heated sulfuric acid vents 2,600 meters below in the Pacific Ocean near Indonesia. Gutless beard worms do not have a mouth, a gut, or an anus. To obtain nutrients, the gutless beard worms have established a symbiotic relationship with sulfur-oxidizing bacteria that live in the trophosome of the worm. These symbiotic bacteria obtain sulfur from the worm and in return, supply the worm carbon. The worms used their hemoglobin to transport oxygen and sulfide simultaneously. Thus, the hemoglobin of the gutless marine beard worm is quite unique. We have modeled the 24 subunit hemoglobin protein from the pogonophoran *O. mashikoi* using Jmol to facilitate visualization as a three-dimensional entity. This molecular modeling allowed us to highlight the oxygen and sulfur binding pockets to better understand the structure and function of this unusual protein. Our results demonstrate the beauty and diversity of proteins.

Running on empty; Greyhound hemoglobin

Ke'Layah Guyton[1], Hannah Lara[1], Jade Vue[1], Heather Young[1], Kathleen A. Boyle[1], Colleen Conway[1], Maureen Leonard[1]

[1]Mount Mary University, Milwaukee, WI 53222

Hemoglobin is an essential ubiquitous protein involved in transporting oxygen in the bloodstream of vertebrates via red blood cells. As vertebrates display different environmental living niches, where levels of oxygen availability and demand vary, hemoglobin proteins display a flexibility in their structure to adapt to the required function. Greyhound dogs are considered sight hounds, in that they hunt and pursue game through sight rather than sounds or smell. Greyhounds have been bred for greyhound racing in which they sprint around an oval track chasing an artificial rabbit. Performance of this strenuous task demands high levels of oxygen. This demand may be met by mutation of the greyhound's hemoglobin protein to reflect a higher oxygen affinity. Amino acid alignment of the hemoglobin sequences of generic dogs versus greyhounds revealed that they are 100% identical. To identify if changes in their tertiary structure impact their oxygen affinity, the hemoglobin proteins were crystallized in an oxygenated, deoxygenated and a transitional phase (R2) forms. It was in this transition phase (R2) form that a conformational change was revealed. A hemoglobin protein with higher oxygen affinity was demonstrated. This increased oxygen affinity can be attributed to five amino acids in the greyhound hemoglobin. We will present our data to demonstrate the reflective changes in structure that impact the change in function.

SARS-CoV-2 ORF8 Accessory Protein Dimerization Domains and Protein-Protein Host Interaction

Elizabeth Wiles[1, 2], Ryan Peterson[1, 2], Allison Robinson[1, 2], Ryan Vazquetelles[1, 2], Leeann Stearns[1, 2], Bailey Hart[1, 2], Karen Guzman[1, 2]

[1]Campbell University, Buies Creek, NC, [2]MSOE Center for Biomolecular Modeling

As the world continues to fight the coronavirus pandemic caused by SARS-CoV-2, we are gaining valuable insights by comparing this virus to other human coronaviruses such as SARS-CoV and MERS-CoV that have also caused outbreaks. Viral accessory proteins are generally implicated in increased infectivity, pathogenicity, and virulence. The accessory protein open reading frame 8 (ORF8), although not essential for viral replication, appears to play a critical role in disease severity by inhibiting multiple pathways of the immune response. Furthermore, the ORF8 gene is found within a highly variable portion of the genome, increasing the possibility of dangerous mutant forms arising. Structure-based design of therapeutics holds great promise for effective targeting of such rapidly evolving threats. Thus, we designed physical models and computer representations of ORF8, based on the published structure 7JTL, to better evaluate the structural features responsible for viral pathogenicity. ORF8 exists as a homodimer held together by amino acid residues that interact through hydrophobic interactions, hydrogen bonding, a salt bridge, and a disulfide bond. These interactions occur in a region referred to as the covalent interface. Another interface, referred to as the noncovalent interface, exists between separate homodimers and results in oligomerization. The homodimers in the oligomer are held together by an intermolecular beta sheet that is stabilized by an extensive hydrophobic surface. These two dimerization interfaces are unique to SARS-CoV-2 and have been proposed to be involved in the protein's ability to evade and suppress the host immune response. Computer and physical representations allow for better visualization of ORF8 structure and evaluation of the role of the multimeric structure. ORF8 has been found to interact with various host proteins, including Interleukin-17 Receptor A (IL17RA), a pro-inflammatory cytokine. To further corroborate the existence of this interaction and to better understand the role oligomerization might play in pathogenicity, we assessed the possible interactions of ORF8 (7JTL) and IL17RA (5N9B) through theoretical modeling using available tools such as HDOCK. Investigation of these inter-subunit interactions may improve our understanding of the unique mechanism the virus uses to evade the immune system and may be instrumental in the development of effective therapeutics.

POSTER SESSION 2
Tuesday May 4, 2021 1:40pm-2:20pm EDT

Structural comparison of cruzipain subtypes models revealed differences in protease subsites that might impact substrate specificity and drug discovery projects against Chagas disease

Augusto César B. Campos[1], Viviane C. Santos[1], Antônio Edson R. Oliveira[3], João Luíz R. Cunha[2], Daniella C. Bartholomeu[2], Santuza Maria R. Teixeira[1], Ana Paula C A. Lima[4], Rafaela S. Ferreira[1]

[1]Departamento de Bioquímica e Imunologia, Universidade Federal de Minas Gerais, [2]Departamento de Parasitologia, Universidade Federal de Minas Gerais, [3]Departamento de Análises Clínicas e Toxicológicas, Universidade de São Paulo, [4]Instituto de Biofísica Carlos Chagas Filho, Universidade Federal do Rio de Janeiro

Chagas disease, caused by the protozoan parasite *Trypanosoma cruzi*, is considered a neglected disease. The WHO estimates that 7 million people live with the infection worldwide, mainly in Latin America. The available drugs for treating the disease present low efficacy and severe side effects. Thus,new and more efficient drugs are urgently needed.
Cruzipain, a cysteine protease from *T. cruzi*, is involved in host cell invasion and modulation of the immune response. Due to its role, cruzipain is a validated target for drug development, and it has been studied for over 30 years.
Since 1994 the evidence of the existence of other cruzipain subtypes has been reported. But in most of the literature, the enzyme is still referred to as a single enzyme: cruzain. All the 26 complexes of cruzain-ligand currently available on PDB, with several classes of inhibitors, are structures of the same subtype.
In recent work, we reported the complete gene repertoire of cruzipain, and we characterize their expression pattern in the life stages of the parasite. Then, we proposed a new classification of cruzipains into four subfamilies.
To investigate structural differences among subfamilies, we modeled, by homology modeling approach in Modeller v9.4, their catalytic domain.All modeled sequences showed high identity (76% to 81%) to the template (PDB 1ME3 - cruzain) and total coverage. Assessment in QMEAN and ProsA-web server checked the high quality of the models. Further, the nAPOLI server analyzed the complexes of cruzain bound to their inhibitors to calculate the frequency of amino acid intermolecular interactions.
Differences among cruzipain subtypes impacted the shape, volume, and charge of S3, S2, and S1’ pockets of the active site. The S2 subsite is the best-defined site in cruzain, concentrating most residues that interact with ligands in over 50% of the known crystal structures. In Cruzipain 3, the S2 subsite is a smaller and solvent-exposed pocket, and in Cruzipain 4, it is a shallow and poorly defined one.
Our results suggest that cruzipain subtypes might have different substrates preferences and diverge affinities to cruzipain ligands.

Three dimensional structure determination of mutant translationally controlled tumor protein and binding site characterization

Maranda S. Cantrell[2, 3], Aaron D. Ajeti[2], Lisa R. Warner[2], Owen M. McDougal[2], Ken Fujise[1]

[1]Department of Cardiology, University of Washington, Seattle, WA 98195, [2]Department of Chemistry and Biochemistry, Boise State University, Boise, ID 83725, [3]Biomolecular Sciences Graduate Programs, Boise State University, Boise, ID 83725

Atherosclerosis is the leading cause of death worldwide. Many drugs to treat atherosclerosis target cholesterol formation, operating under the widely accepted lipid hypothesis, wherein high density lipoproteins (HDL) carrying cholesterol are found at high levels in atherosclerotic patient blood. However, recent studies suggest that approximately 50% of patients who have been diagnosed with atherosclerosis present with normal to low levels of cholesterol. Thus, a recent push in research efforts to design novel drugs that target different mechanisms of atherosclerotic lesion formation has been realized. Translationally controlled tumor protein (TCTP), also known as fortilin, is a pro-survival molecule and apoptotic inhibitor. It has been implicated in a novel mechanism for atherosclerosis, whereby its over-expression leads to foam cell aggregation by preventing macrophage apoptosis. Inhibition of fortilin in rodent models has demonstrated a potential new class of pharmaceutical intervention for atherosclerosis patients. Novel small molecule inhibitors of fortilin have been developed and tested in vitro and in vivo, however, the mechanistic details of inhibition are still unknown. Additionally, evidence suggests recombinantly expressed fortilin in E. coli with affinity tags for purification are un-cleavable without insertion of a GGS linker sequence. To free up the N- and C- termini, which are predicted binding sites for these small molecules, we constructed 3 mutant fortilin constructs where a streptavidin-affinity amino acid sequence was placed in the flexible domain of fortilin and tested to ensure that biological activity was retained and that these mutants would be viable for small molecule binding site determination via NMR spectroscopy.

Structural alphabets for conformational analysis ofnucleic acids available at dnatco.datmos.org

Jiri Cerny[1], Bohdan Schneider[1]

[1]Institute of Biotechnology of the Czech Academy of Sciences

A detailed description of the DNATCO (https://dnatco.datmos.org) web server implementing the universal structural alphabet of nucleic acids will be presented. It is capable of processing any mmCIF- or PDB-formatted files containing DNA or RNA molecules; these can either be uploaded by the user or supplied as the wwPDB or PDB-REDO structural database access code. The web server performs an assignment of the nucleic acid conformations and presents the results for the intuitive annotation, validation, modeling and refinement of nucleic acids.

Repurposing drugs against the main protease of SARS-CoV-2

Sohini Chakraborti[1], Sneha Bheemireddy[1], Narayanaswamy Srinivasan[1]

[1]Indian Institute of Science, Bengaluru, Karnataka 560012, India

'Drug Repurposing' deals with finding a new indication of an existing drug and thus, aiming to reduce the overall time and cost involved in discovering a new molecule. This study presents a computational analysis that exploits the structural similarities in protein 3-D and chemical 2-D space to suggest potential repurpose-able molecules against the main protease of SARS-CoV-2 (Mpro). It is known that proteins with similar structures may recognize similar ligands and vice-versa. Using the experimental structure of Mpro as the reference, we screened the PDB to obtain proteins with a similar structure known to bind to any approved/investigational drugs. Again, using the 2-D chemical structure of a small-molecule known to inhibit Mpro as the reference, we screened the DrugBank database to obtain compounds with similar chemical structures. Selected molecules were then subjected to docking studies to predict the feasibility of favorable interactions between the compounds and the binding pocket of Mpro. Subsequently, 29 drugs/drug-candidates have been prioritized for testing. Our analysis is based on an integrative framework that exploits available experimental data to provide possible mechanism-based insights. This can aid decision-making for further probing. Besides several approved/investigational anti-viral drugs, other hits from our study include anti-bacterial, anti-inflammatory, anti-cancer, and anti-coagulant drugs. Interestingly, many of these identified candidates are indicated to possess anti-viral properties in earlier published reports, and some of these are suggested to simultaneously modulate the functions of viral proteins and the host response system. However, the molecular mechanisms of action (MoA) of these drugs in viral infections are not clearly understood. Hence, this work opens a new avenue of research to probe into the MoA of drugs known to demonstrate anti-viral activity but are so far not known to target viral proteases. It is worth mentioning that the methodology we used in our study requires minimalistic resources and is a generic one; thus, it can be applied to any protein target if the necessary structural information is available.

Reference: Chakraborti, S., Bheemireddy, S. & Srinivasan, N. Repurposing drugs against the main protease of SARS-CoV-2: mechanism-based insights supported by available laboratory and clinical data. Mol.
Omi. (2020); 16(5):474-491.

Uncovering Structure-Function Relationship Between an Antifungal drug, Membranes and Sterols

Kevin J. Cheng[2], Ashley M. De Lio[1], Agnieszka Lewandowska[1], Corinne P. Soutar[1], Martin D. Burke[1], Chad M. Rienstra[3], Taras V. Pogorelov[1, 2]

[1]Department of Chemistry, University of Illinois at Urbana-Champaign, Champaign, IL, USA,
[2]Center for Biophysics, University of Illinois at Urbana-Champaign, Champaign, IL, USA,
[3]Department of Chemistry, University of Wisconsin-Madison, Madison, WI, USA

Amphotericin (AmB) is a potent and effective drug to combat life-threatening fungal infections. However, it is extremely toxic to humans and can lead to fatal complications including kidney damage or heart failure. It was experimentally shown that AmB's anti-fungal activity originates from its binding to ergosterol in fungal cells. Since ergosterol is similar to human cholesterol, the toxicity purported to come from AmB's non-specific binding to both sterols. To develop AmB derivatives that maintain potency while reducing toxicity, structural elucidation of sterol interactions is necessary. The goal of our study is to characterize unique binding modes between sterols and AmB and reveal an atomistic mechanism behind its fungicidal activity. To this end, we use all-atom molecular dynamics (MD) simulations to capture how an NMR-derived AmB lattice structure interacts with the membrane bilayer. We efficiently explore the structural landscape of AmB-Sterol complexes using extensive ~80 μs of replica exchange MD (REMD). Resulting AmB membrane-embedded structures are then subjected to equilibrium MD with the highly mobile-mimetic (HMMM) model that uses a novel sterol-compatible, in-silico solvent. We connect our results to solid-state NMR experiments by calculating theoretical chemical shifts. Our methodology produced an ensemble of atomistically resolved dynamic AmB-Sterol complexes that are within 8 ppm of experimental chemical shifts. We further support our computational model by showing structures of AmB complexes to satisfy within 3.5 Å the majority of distance restraints derived from NMR-REDOR experiments. This study offers a novel and generalizable workflow that combines distinct enhanced sampling methods, and bridges derived structures with NMR observables. Most importantly, the impact of this work will guide future experiments that focus on the design of AmB derivatives to selectively kill fungal cells and not to harm human cells. Such derivatives have the potential to save many lives.

X-ray crystallographic fragment screening of HIV-1 reverse transcriptase with a novel small halogenated library (Halo)

Ashima Chopra[1, 2], Joseph D. Bauman[1], Francesc X. Ruiz[1], Edward Arnold[1,2]

[1]Center for Advanced Biotechnology and Medicine, Rutgers University, [2]Department of Chemistry and Chemical Biology, Rutgers University

X-ray crystallographic fragment screening (XCFS) uses fragment-sized molecules (~200-300 Da) to access binding sites on proteins that may be inaccessible to larger drug-like molecules (~500-700 Da). Fragments containing halogen atoms (F/Cl/Br/I) are shown to have a high hit rate, compared to traditional fragment libraries. So, we designed the Halo Library containing 46 halogenated fragments (including the “universal fragment” 4-bromopyrazole). Fragments were selected based on a high binding occurrence in the PDB. The library was screened against crystals of HIV-1 reverse transcriptase with drug rilpivirine, yielding a high hit rate of 28%. Three new binding sites were discovered. Multiple fragments were found to be inhibitory towards RT polymerization, and will be pursued for fragment development. The library provides a convenient method of finding fragment binders of protein targets, discovering new sites, and paves the way to developing drug leads.

Current computational methods for enzyme design

Talmage L. Coates[1], Dennis Della Corte[1], Naomi Young[1], Austin J. Jarrett[1], Connor J. Morris[1], James D. Moody[1]

[1]Brigham Young University

Computational enzyme design has made massive progress in recent years, in large part thanks to the explosive growth in protein data availability. Computational enzyme design has emerged as a powerful tool which predicts how mutations modify the properties of enzymes. Computational methods are both faster and cheaper for evaluating large numbers of mutations. This can decrease the cost and length of the design process. This presentation draws on our recently published review “Current computational methods for enzyme design” in covering the wide variety of computational tools used over the last five years to design enzymes.

Cryo-EM Structure of Monomeric Cyanobacterial Photosystem I at 3.2 Å Resolution with A Cryo-ARM200

Orkun Çoruh[1], Anna Frank[2], Akihiro Kawamoto[1], Takayuki Kato[1], Keiichi Namba[3], Christoph Gerle[1], Marc M. Nowacyzk[2], Genji Kurisu[1]

[1]Institute for Protein Research, Osaka University, Suita, Osaka, Japan, [2]Plant Biochemistry, Faculty of Biology and Biotechnology, Ruhr-University Bochum, Bochum, Germany, [3]Graduate School of Frontier Biosciences, Osaka University, Suita, Osaka, Japan

Photosystem I (PSI), using the energy from light efficiently, propels the Calvin cycle by transferring the electrons from plastocyanin in lumenal side, across the thylakoid membrane to ferrodoxin in the stromal side. Understanding how this molecular machine orchestrates the efficient light harvesting and excitation energy transfer along with the electron transfer is of the utmost importance in the context of agriculture and bio-inspired applications such as photovoltaic energy production and photocatalysis. Following the illustrious structural analysis of trimeric cyanobacterial PSI from Thermosynechococcus elongatus as the first atomic resolution model of PSI in 2001, extensive work has been done on elucidating the structure of PSI from various organisms. However, despite all the supporting information from biochemical experiments, no conclusive theories have been established neither about the mysteries about the oligomeric presence nor the harmonious energy conversion ability of the protein under variable light conditions. Biochemical studies about cyanobacterial PSI have shown that monomeric PSI holds relevant information about oligomerization and the energy governance of the protein. Accordingly, more structural information about a functional cyanobacterial monomeric PSI transpires to be vital to comparatively investigate the intermingled details of the processes. Here, we describe the structure of monomeric PSI from Thermosynechococcus elongatus BP-1 [1]. Comparison with the trimer structure revealed monomerization-induced disorders in both the central trimerization domain and the peripheral regions of the complex. Encountered disorders are related to spectroscopic data, leading to insights about oligomerization mechanism and the identity of red chlorophylls, entailing hypotheses about the function of PsaX, effect of red chlorophylls on electron transfer, and a switch mechanism for activating the red chlorophylls.

[1] Çoruh et al., Commun Biol., 4:304, 2021

A game changer: First structure of a multiheme c-type cytochrome from a Gram-positive bacterium shows unanticipated relations between anaerobic respiratory metabolisms.

Nazua L. Costa[1], Bianca Hermann[2], Ricardo M. Soares[1], Oliver Einsle[2, 3], Catarina M. Paquete[1], Ricardo O. Louro[1]

[1]Instituto de Tecnologia Química e Biológica António Xavier, Universidade NOVA de Lisboa, Oeiras, Po, [2]Institut für Biochemie, Albert-Ludwigs-Universität Freiburg, Freiburg im Breisgau, Germany, [3]BIOSS Centre for Biological Signalling Studies, Albert-Ludwigs-Universität Freiburg, Freiburg im Bre

The discovery of microbial electroactivity happened more than a century ago, when Potter demonstrated that a polarized electrode inserted in media containing bacteria and fungi was able to generate electrical current. This landmark discovery, that boosted the bioelectrochemical technology as we know today, has been ever since linked to Gram-negative bacteria due to their optimal cell structural arrangement featuring three well differentiated subcellular localizations. Conversely, the thick peptidoglycan layer and narrow periplasmic space characteristic of Gram-positive bacteria argued for the inability of this class of microorganisms to be considered capable of extracellular electron transfer (i.e., exchange catabolic electrons with solid electron acceptors outside the cell). Recently, a thermophilic Gram-positive bacterium, Thermincola potens JR was shown to establish direct contact with the anode of an operating microbial fuel cell providing unique evidence for extracellular electron transfer performed by this class of microorganisms. Furthermore, protein identification studies pointed to the involvement of surface exposed multiheme c-type cytochromes (MHC) in conducting electrons from the cellular metabolism to the cell surface as previously described for Gram-negative bacteria. Here we report for the first time the structure of an outer cell wall associated MHC (OcwA) from this Gram-positive bacterium. The structure of OcwA from T. potens JR, obtained by X-ray with 2.2 Å resolution, showed that the overall fold and organization of the hemes in OcwA are not related to the “staggered cross” architecture found in the Gram-negative outer-membrane metal reductases. Instead, the OcwA structure is similar to those of MHC involved in the biogeochemical cycles of nitrogen and sulfur. In line with this observation OcwA was also shown reduce nitrite and reduce and oxidise hydroxylamine, in addition to efficiently transferring electrons to electrodes and iron oxides. These data reveal that OcwA can take the role of a respiratory “Swiss Army knife,” allowing T. potens JR to prosper in environments with rapidly changing availability of terminal electron acceptors without the need for transcriptional regulation and protein synthesis. Our data also reveal that terminal oxidoreductases of soluble and insoluble substrates are evolutionarily related, providing novel insights into the evolution of metabolic pathways relying on MHC.

Solubilization and Purification of the Membrane Associated Cyr1p-LRR from the Human Commensal Candia. albicans

Geneva Crump[1], Catherine Leimkuhler Grimes[1], Sharon Rozovsky[1]

[1]University of Delaware

The human commensal *Candida. albicans* is the most common fungal resident of the human microbiome. In healthy individuals this fungus occupies a budding morphology. During dysbiosis or immune incompetence, this fungus undergoes a morphological transition from a budding yeast to a hyphal format which results in long filamentous tails that invade human epithelial cells and if not successfully treated, can cause lethal systemic infections. One of the key molecular signals responsible for this morphological switch is the binding of bacterial peptidoglycan fragments to the leucine rich repeat domain of the adenyl cyclase Cyr1p. This binding event results in a cAMP signaling cascade event that perpetuates the upregulation of hyphal specific genes and thus fosters human cell invasion. To understand the principles that govern the binding and regulation of this morphological transition, it is imperative to obtain purified recombinant protein for biochemical analysis such as binding and inhibition studies. Previous studies that aimed to characterize binding of this protein to its peptidoglycan receptors although successful, were severely limited by the ability to obtain soluble, stable, and tag free purified recombinant protein. We recently discovered that CYR1P largely resides in the cellular membrane and can be solubilized with nondenaturing detergents, thus avoiding the laborious and unpredictable previous denaturing and refolding purification procedures utilized. After membrane fractionization and a series of detergent screens, the Cyr1p-LRR domain was solubilized and purified under native conditions in the nondenaturing detergent FOS-Choline with a small noninterfering 1.4Kda epitope tag. The ability to obtain the entire LRR domain of purified recombinant Cyr1p from *Candida. albicans* under native conditions, with an inert tag will enable the ability to establish binding parameters and set the stage for future inhibitor screens.

A new integrative NMR-SAXS approach reveals the structure of a trans VS ribozyme in solution

Pierre Dagenais[1], Geneviève Desjardins[1], Pascale Legault[1]

[1]Université de Montréal

Divide-and-conquer strategies are commonly used in several scientific fields to solve problems by breaking them down into simpler parts. Although RNA molecules are modular by nature, the application of this type of strategy for high-resolution structure determination of RNAs has surprisingly been limited in the past. Here, we have developed a new integrative approach based on the divide-and-conquer strategy and applied it to determine high-resolution solution structures of a *trans* Neurospora VS ribozyme. As a first step, NMR and SAXS studies were conducted on several isolated subdomains which define the overall architecture of the ribozyme. Structural models of a *trans* ribozyme were then built by combining the subdomain structures using fragment assembly and filtering these structures based on their fit to the experimental SAXS data. The resulting hybrid NMR-SAXS structural ensemble shares several important structural properties with the reported crystal structures of the VS ribozyme. However, a local structural difference was observed in stem III, which affects the relative orientations of the two three-way junctions of the ribozyme and the positioning of key residues for its catalytic activity. This study uncovers a global conformational change in the VS ribozyme structure that is likely associated with substrate binding and highlights how integrative approaches can be used to study the structure and dynamics of large RNAs in solution.

Structural insights of human N-acetyltransferase (NAT10) and identification of Remodelin binding site

Mahmood H. Dalhat, Hisham N. Altayb, Mohammad Imran Khan, Hani Choudhry

Biochemistry Department, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia, Centre for Artificial Intelligence in Precision Medicine, KAU, SaudiArabia, Cancer and Mutagenesis Research Unit, King Fahd Medical Research Center, KAU, Saudi Arabia

N-acetyltransferase 10 (NAT10) is a nucleolar protein known to add acetyl group(s) to RNAs and proteins leading to acetylation mediated stability of RNAs and proteins. Deregulation in NAT10 function reported in diseases such as Hutchinson-Gilford progeria syndrome (HGPS), human immunodeficiency virus (HIV) and cancer showed its promising role as a potential therapeutic target. Remodelin is a small molecule inhibitor of NAT10 known to reverse the defect in HGPS, reduce HIV viral copies and reverse effect of cell proliferation, cell migration & attenuating doxorubicin resistance in cancer. Although several studies have used Remodelin as inhibitor of NAT10, but still the site of Remodelin-NAT10 interaction is unclear. Here we predicted the Remodelin-NAT10 interaction using human modeled NAT10. Molecular docking of Remodelin revealed it binds to the acetyl-CoA binding pocket of NAT10. To compare the substrate of acetyl-CoA binding pocket of NAT10 and Remodelin; we also performed docking for Acetyl-CoA. Binding scores revealed similarity in Acetyl-CoA (−5.691) and Remodelin (−5.3) as well as molecular interaction sites. Further, both Acetyl-CoA- and Remodelin- docked NAT10 showed stability in molecular interaction during molecular dynamic (MD) simulation. Since several studies have reported the functional role of NAT10 in cell proliferation, cell migration, rRNA biogenesis and DNA damage response through NAT10 mediated RNA or protein acetylation, we therefore presumed; that Remodelin inhibits NAT10 acetylation activity by occupying the acetyl-CoA binding site.

The Role of Conformational Change and Scaffold Stability in NF-κB Essential Modulator Function

Anthony M. DeMaria[1], Christopher J. Dirusso[2], Robert Shaffer[1], Adrian Whitty[1], Thomas D. Gilmore[2], Karen N. Allen[1]

[1]Department of Chemistry, Boston University, Boston, MA 02215, [2]Department of Biology, Boston University, Boston, MA 02215

Scaffold proteins help mediate interactions between protein partners, enhancing intracellular signaling, but are often assumed to play passive roles. We are investigating the structural basis behind how the extended coiled-coil scaffold protein NF-κB Essential Modulator NEMO can rapidly stimulate the NF-κB inflammatory pathway, focusing on the central Intervening Domain (IVD) of NEMO. We have employed Small-Angle X-ray Scattering (SAXS) to investigate a ligand-induced conformational change, utilized Circular Dichroism (CD) to look at domain interplay on dimer stability, and analyzed the sequence of the IVD, using the results to conduct *in-vivo* mutational studies. SAXS showed a ligand-induced conformational change indicated by IVD compaction, and factors causing stabilization of the IVD as measured by CD lead to destabilization of adjacent domains. Sequence analysis is consistent with the globular character of the IVD, with *in-vivo* mutants retaining IVD hydrophobicity preserving the ability for NF-κB pathway activation. This indicates a conformational change dependent on IVD conversion to a more globular form, at the expense of destabilizing surrounding domains, overcome via ligand binding. Together, these findings indicate nuanced participation of a scaffold protein in activation of a vital pathway and present alternative targeting areas for therapeutic modulation.

Structural and Biochemical Insights into Adenine Excision by N146S Gs MutY

Merve Demir[1], Peyton Russelburg[2], Jonathan Lin[1], Carlos H. Trasvina-Arenas[1], Martin P. Horvath[2], Sheila S. David[1]

[1]Department of Chemistry, University of California Davis, Davis, CA, [2]School of Biological Sciences, University of Utah, Salt Lake City, UT

The DNA glycosylase MutY prevents G:C to T:A transversion mutations by removing misincorporated adenine across 8-oxo-guanine (OG), an oxidation product of guanine. In this study the role of an active site residue N146 in the adenine excision mechanism has been explored in *Geobacillus stearothermophilus* (*Gs*) MutY using site-directed mutagenesis, kinetic studies and X-Ray crystallography. The crystal structure of *Gs* MutY bound to the transition state analog (1N) was instrumental in identifying hydrogen bond contacts N146 forms with the DNA phosphate backbone and catalytic residue D144. Interestingly, mutation of this residue to a serine corresponds to a MUTYH-Associated Polyposis (MAP) variant that predisposes individuals to colorectal cancer. The catalytic activity of N146S was hindered significantly in comparison to wild-type, and the pH-dependent kinetics showed an increase in the pK_a of D144. This result suggests that N146S impacts the efficiency of the reaction by altering certain electrostatic interactions between catalytic residues and the reaction intermediates. In order to understand the role of N146 in the adenine excision mechanism at a molecular level, we used an acid labile substrate analog, purine (P), for biochemical analysis and crystallography. Surprisingly, the crystal structure of N146S with purine shows coordination of a Ca^{2+} ion in the additional space formed by N146S substitution. Furthermore, the purine base is still intact in the active site despite the ability of N146S *Gs* MutY to excise P. We believe Ca^{2+} coordination in the active site prevents the cleavage of P. Hence, crystallization of N146S with P in the absence of Ca^{2+} resulted in abasic site product. Our aim is to remove Ca^{2+} from the active site in crystalline state to initiate and trap different steps of the reaction including intermediates that have been elusive thus far. Finally, the preliminary results of N146S-OG:P crystals soaked in EDTA containing solution showed lowered Ca^{2+} occupancy in the active site. In conclusion, by monitoring the base excision mechanism of N146S *Gs* MutY, we hope to gain insights into why this MAP variant has altered activity and suggest potential treatments.

Role of protein HelD in bacterial transcription – cryo-EM complex of HelD and RNA polymerase

Jan Dohnálek[1], Tomáš Kouba[2], Tomáš Koval[1], Petra Sudzinová[3], Jirí Pospíšil[3], Barbora Brezovská[3], Jarmila Hnilicová[3], Hana Šanderová[3], Martina Janoušková[3], Michaela Šiková[3], Petr Halada[3], Michal Sýkora[4], Ivan Barvík[5], Jirí Novácek[6], Mária Trundová[1], Jarmila Dušková[1], Tereza Skálová[1], URee Chon[7], Katsuhiko S. Murakami[7], Libor Krásný[3]

[1]Institute of Biotechnology of the Czech Academy of Sciences, 252 50 Vestec, Czech Republic, [2]EMBL Grenoble, 71 Avenue des Martyrs, Grenoble, France, [3]Institute of Microbiology of the Czech Academy of Sciences, Prague, Czech Republic, [4]Institute of Molecular Genetics of the Czech Academy of Sciences, Prague, Czech Republic, [5]Faculty of Mathematics and Physics, Institute of Physics, Charles University, Prague, Czech Republic, [6]CEITEC, Masaryk University, Brno, Czech Republic, [7]The Center for RNA Molecular Biology, Pennsylvania State University, University Park, PA 16802, USA

RNA synthesis is central to life, and RNA polymerase (RNAP) depends on accessory factors for recovery from stalled states and adaptation to environmental changes. Here, we report a cryo-EM structure of a complex between the *Mycobacterium smegmatis* RNAP and HelD (Kouba et al., Nat Commun 2020, 11(1):6419). The crescent-shaped HelD simultaneously penetrates deep into two RNAP channels that are responsible for nucleic acids binding and substrate delivery to the active site, thereby locking RNAP in an inactive state. Three structural classes suggest three states of the process of HelD binding to RNAP and of its interference with transcription. HelD is expected to help release RNAP from stalled elongation complexes. While the mechanism of its binding and nucleic acids release is largely explained by the current results, the mechanism of HelD unbinding from RNAP is yet to be explained.
This work was supported by MEYS (LM2015043 and CZ.1.05/1.1.00/02.0109), CSF (20-12109S and 20-07473S), NIH (grant R35 GM131860), AS CR (86652036), ERDF (CZ.02.1.01/0.0/0.0/16_013/0001776 and CZ.02.1.01/0.0/0.0/15_003/0000447), EMBL (EI3POD) and Marie Sklodowska-Curie grant (664726).

Arsenite oxidase bound to new intermediates help to elucidade enzymatic reaction mechanism

Filipa S. Engrola[1], Márcia Correia[1], Maria João Romão[1], Teresa Santos-Silva[1]

[1]UCIBIO, FCT-NOVA

Arsenic (As) and antimony (Sb) are two metalloids that, due to anthropogenic and natural causes, pose an environmental threat, considered as priority pollutants by the World Health Organisation and the United States Environmental Protection Agency. Although the safety guards recommend a maximum of 10 µg/L of As and Sb in drinking water, these values are exceeded in many regions worldwide, with no remediation approach that is simultaneously effective, clean and economically sustainable [1,2]. The ancient bioenergetic enzyme arsenite oxidase (Aio), from microorganisms Rhizobium sp. NT-26 (NT-26 Aio) and Alcaligenes faecalis (A.f. Aio), is currently being studied for its use as a biosensor and in bioremediation processes. Both Aio enzymes contain a large subunit (AioA) that harbours a molybdenum centre and a [3Fe-4S] cluster, and a small subunit (AioB) that possess a Rieske [2Fe-2S] cluster and have demonstrated to oxidise AsIII, as well as SbIII, into the easier to remove and less toxic forms of AsV and SbV, respectively [3,4]. Aiming to elucidate the catalysis mechanism of the enzymes, a combination of expression and purification of the proteins, crystallisation, structural analysis, enzyme kinetics and affinity tests were conducted. X-ray structures of the ligand-free form of the enzyme had been previously determined (PDB: 4AAY, 5NQD and 1G8K [3,5,6]). In our work, Aio crystals in complex with two different forms of the substrate analogue - Sb oxyanions, with a reaction kinetic 6500 times slower than AsIII [6] - diffracted up to ca 1.8 Å resolution. The structures show the reaction intermediates bound at the active site, with a µ-oxo bridge binding Sb to the Mo atom. Analysis of bond lengths and geometry of the ligands at the Mo active site allowed us to revisit the catalytic mechanism of As oxidation [7], contributing to the understanding and future biotechnological application of this family of enzymes in water treatment.

Exploring Beta Structures of the TOC159 Membrane Anchor and their Role in Chloroplast Targeting Specificity

Michael Fish[1, 2], Simon Chuong[3], Masoud Jelokhani-Niaraki[2], Matthew Smith[1]

[1]Department of Biology, Wilfrid Laurier University, Waterloo, ON, Canada, [2]Department of Chemistry and Biochemistry, Wilfrid Laurier University, Waterloo, ON, Canada, [3]Department of Biology, University of Waterloo, Waterloo, ON, Canada

The dynamic function of plastids relies on the fidelity of protein targeting pathways which shuttle plastid precursor proteins to the organellar surface where they are recognized and imported. The targeting of TOC159 receptors and other key components of the translocation machinery is essential to processes of plastid biogenesis and photomorphogenesis and therefore plant growth and plant development. Due to the ambiguous nature of targeting signals between plastid and mitochondrial precursor proteins, establishing key players in targeting specificity remains an important undertaking in the field. Previous fluorescent targeting experiments have established that the TOC159 membrane domain carries an unconventional membrane anchor, composed of β-strands. Our most recent fluorescent targeting data implicates the anchor in plastid targeting with the proposed signal composed of a β-hairpin at the anchor's C-terminus. Bioinformatic analyses have identified some homology to the lipid binding β-helix domain of UDP-3-O-acyl-glucosamine N-acetyltransferase (LpxD), but our most recent structural predictions using iTASSER and Phyre 2 suggest the anchor could take on a β-propeller or β-prism fold. β-propellers and β-prisms are known for binding sugar groups and irreversibly associating with biological membranes. Interestingly, plastid membranes are the only membranes in plant cells that contain galactolipids which may provide binding sites for such β-folds and could be involved in plastid targeting specificity. Future studies will attempt to resolve the structure of the membrane anchor and its interactions with galactolipids using CD spectroscopy and x-ray crystallography.

Identification of protein-ligand interaction diversity by alternative splicing and populational genome variants

Ziyang Gao, Mo Sun, Dmitry Korkin

Department of Bioinformatics and Computational Biology, Worcester Polytechnic Institute

Protein-ligand binding is often an initial step behind many biological processes. Ligands, such as ions or small organic compounds form molecular complexes with the receptors to facilitate key physiological functions. Over several decades, thousands of protein-ligand complexes have been solved and deposited to Protein Data Bank, allowing the scientists to determine key regions of receptors, ligand binding sites, which include atoms that are in direct physical contact with the ligand molecule. Here, we view the protein-ligand interaction, not as a mere biophysical process, but a complex regulatory event that can be altered by genetic changes such as mutations in the ligand binding sites of a receptor due to genetic variation in healthy populations or disease phenotypes. In this project, we aim to study two regulatory mechanisms affecting the protein-ligand interactions, genetic variation across ethnic group and post-transcriptional variation due to tissue-specific alternative splicing.

The human genetic variants influence the protein-ligand binding by producing the non-synonymous mutations of receptor proteins. The frequency of the variants is differentially distributed across the populations; thus, each ethnic group may have unique common mutations affecting protein-ligand binding. We took a recently released and most complete dataset on human genome variation to analyze the mutations in the ligand-binding regions and classify the ligands into their functional groups to create the population-wise profile of genetic variants in ligand binding sites. Furthermore, we studied how alternative splicing (AS)—a process of making multiple protein products from a single gene—can affect ligand binding due to differentially expressed isoforms across different tissues. The ligand-binding sites of a protein-ligand interaction can be affected by AS because the binding region might be spliced out. We built a novel theoretical model to quantify the impact of AS on protein- ligand binding, by integrating the isoform ligand-binding function and the tissue-wise isoform expression levels.

Cryo-EM based structural study of influenza stem immunogens displayed on MsDps2 protein nanoparticles for vaccine design

Priyanka Garg[1], Uddipan Kar[1], Somnath Dutta[1], Raghavan Varadarajan[1]

[1]Molecular Biophysics Unit, Indian Institute of Science, Bangalore-560012, India

The influenza virus is an important human pathogen. Hemagglutinin (HA), the primary surface glycoprotein on the Influenza virus, is an attractive target for vaccine design. Due to the high mutational rate of viral surface glycoproteins, we recently designed a bacterially expressed'headless' HA mini-stem immunogen pH1HA10-Foldon, where Foldon is a trimerization domain from phage T4 fibritin. The HA stem was derived from the H1N1 A/California/04/2009isolate (PDB-3LZG). We engineered genetic fusion of pH1HA10-Foldon onto DNA binding protein from stationary-phase cells of Mycobacterium smegmatis (MsDps2, PDB-2Z90) platform to improve immunogenicity through a multivalent display.

In our current study, our target is to determine the structural details of MsDps2-pH1HA10 fusion immunogen using cryo-EM. Initially, the complex is visualized by negative staining EM, and monodispersed complexes are observed. However, the low-resolution TEM images and 2D class averages of MsDps2-pH1HA10 failed to detect any immunogens displayed because of the flexibility of the pH1HA10. Therefore, we probed the structure of MsDps2-pH1HA10 using single-particle cryo-EM. Reference-free 2D class averages and 3D reconstruction, show the extra density of pH1HA10 at the periphery of the nanoparticle (MsDps2). The 3D classification study shows that the immunogens are firmly attached to the three-fold symmetry axes on the MsDps2 nanoparticle. This result supports our engineering strategy of genetic fusion of pH1HA10-Foldon onto MsDps2. The immunogens are connected to the nanoparticle core by 9 amino acid flexible linkers, which hinder high-resolution structural characterization of nanoparticles. However, the structure of MsDps2- pH1HA10 fusion is solved at a resolution of 7Å, whereas the core MsDps2 is resolved at 2.6-3.42 Å. This is the first cryo-EM based structural characterization of immunogens displayed protein nanoparticles at near atomic resolution, where the core is resolved at below 3Å .The structure is validated by fitting the crystal structure into the cryo-EM density map. We are currently probing the immunogencity in mouse immunization and challenge studies

Novel Targets for Iron-Starving Resistant Staphylococcus aureus: in Silico Characterization of the IsdB-Human Hemoglobin Complex

Eleonora Gianquinto[1], Omar De Bei[2], Marialaura Marchetti[2], Loretta Lazzarato[1], Mariacristina Failla[1], Stefano Bettati[2, 3], Luca Ronda[3], Barbara Campanini[4], Francesca Spyrakis[1]

[1]Dept. of Drug Science and Technology, University of Turin, Via Giuria 9, 10125, Turin, Italy, [2]Interdept. Center Biopharmanet-TEC, University of Parma, Parco Area delle Scienze, Bldg. 33, 43124,, [3]Dept. of Medicine and Surgery, University of Parma, Via Gramsci 14, 43126, Parma, Italy, [4]Dept. of Food and Drug, University of Parma, Parco Area delle Scienze, 27/A, 43124, Parma, Italy

The acquisition of iron from the human host environment is an essential step for *Staphylococcus aureus* to survive. In infections, *S. aureus* obtains iron primarily from methemoglobin (metHb), by expressing the Iron-regulated surface determinants (Isd), among which IsdB plays a key role in hemoglobin recognition and binding, and heme extraction. In particular, recent works demonstrated that without IsdB, *S. aureus* loses its virulence [1]. The aim of this project is to investigate the interaction between IsdB (composed by two neat domains $IsdB^{N1}$ and $IsdB^{N2}$ connected by a linker) and Hb by means of *in silico* methodologies, to get insights on the dynamics of the system, the mechanism of IsdB:metHb interaction and heme extraction, and to provide essential information for new potential antimicrobials. Extensive Molecular Dynamics (MD) simulations (about 6 μs) were run to explore the IsdB:Hb interaction and identify the critical residues involved in the complex formation and in the first steps of heme extraction. The composite pattern of electrostatic interactions that we observed in IsdB:metHb complex might stabilize the $IsdB^{N1}$:metHb interface, allowing the $IsdB^{N2}$ domain approaching the heme binding site and starting the extraction process. In particular, these interactions can have a key role in stabilizing the transition state, accomplished through a disruption of the F helix of Hb, and a consequent weakening of the coordination bond between the heme iron and the proximal histidine. Moreover, specific key residues mediating the interaction with the heme have been identified and suggested as mediators of heme extraction.

[1] Grigg J. C. et al., *J. Inorg. Biochem*. **2010**, 104 (3):341-8.
REFERENCE PUBLICATION: Gianquinto, E., Moscetti, I., De Bei, O. et al., *Sci Rep*. **2019**, 9, 18629.

Exploration of a novel mechanism of cardiac troponin release: in silico evidence of non-classical secretion

Jose M. Gonzalez-Rayas[1], Jose A. Hernandez-Hernandez[1], Rosa D. Lopez-Sanchez[1], Ana L. Rayas-Gomez[2], Jose M. Gonzalez-Yanez[2], Carlos Jerjes-Sanchez[1]

[1]Tecnologico de Monterrey, School of Medicine and Health Science, [2]Hospital San Jose de Queretaro

Cardiac troponin is one of the most important biomarkers in cardiology. Traditionally, it has represented a biochemical sign of necrosis, a characteristic of acute myocardial infarction (AMI). However, with the arrival of high-sensitivity detection assays, cardiac troponin is now detected in healthy patients and in a number of uncommon scenarios unrelated to AMI (such as strenuous exercise, rapid atrial pacing and COVID-19). This suggests that troponin could be released as a consequence of alternative processes other than the well-stablished necrotic cell death. To try to explain this phenomenon, multiple mechanisms of troponin release have been postulated, although there is still no consensus on this matter. Furthermore, the term *myocardial injury* (a troponin determination above the upper reference limit, not necessarily associated with a specific diagnosis) was included in the Fourth Universal Definition of Myocardial Infarction. Nevertheless, the histological or pathobiological process responsible for *myocardial injury* has not been determined. Here, we studied the possibility of troponin release through non-classical secretion, a pathway specially triggered by cellular stress and inflammation. The protein sequences of the eight different isoforms of troponin (I, T and C of fast-twitch skeletal, slow-twitch skeletal and cardiac muscle) were retrieved from UniProtKB and the Protein Data Bank. *SecretomeP 2.0*, a neural networks-based program, was used to predict if any of these proteins undergoes non-classical release. *SignalP 5.0* and *TMHMM 2.0* were also applied to rule out the possibility of classical (endoplasmic reticulum/Golgi dependent) and Type IV (transmembrane) non-classical secretion. After analyzing the sequences, the cardiac T isoform was predicted to be non-classically secreted, as well as the I, T and C subunits of the fast-twitch skeletal muscle and the I subunit of the slow-twitch skeletal muscle. None of the troponin isoforms was found to be subject to classical or Type IV secretion. We believe that troponin T release through a non-classical pathway may contribute to the pathobiological definition of *myocardial injury*. There are already some experimental studies showing the release of troponin through extracellular microvesicles (a type of non-classical secretion) in various cell lines. Nonetheless, more evidence is needed to fully understand this intriguing phenomenon in cardiomyocytes.

Clinical and Immunophenotyping of young ethnic south Indian population with diabetes mellitus

Sarayu Gopal S S , Sadishkumar Kamalanathan , Jayaprakash Sahoo , Negi V S , Dukhabandhu Naik

JIPMER

ABSTRACT

Background: Diagnosis of Type 1 diabetes(T1D) is largely clinical and the distinction between diabetic types is uncertain, especially in adolescence and youth. Report on antibody status along with immunophenotyping will strengthen the diagnosis and treatment decisions for young diabetes mellitus. **Objective:** To characterize the profile of autoantibodies, T follicular helper cells (TFH) and c-peptide levels in young ethnic south Indian population with diabetes mellitus. **Methodology:** The study participants were T1D(82), T2D(48) and healthy controls, HC(30). Islet cell antibodies such as GAD65A, IA2A and znT8A were measured by ELISA. Random C- peptide levels were measured by Chemiluminescence Immunoassay after confirmation of euglycemic status. Flowcytometry analysis were done to find the frequency of TFH population using surface protein markers such as CD4, CD45RA, CXCR5, ICOS, PD1 **Results:** Islet cell autoantibodies were positive in 61(74%) of clinical type 1 DM patients. GAD antibody was the most prevalent seen in 54(66%)of them. Autoantibody negative T1D was documented in 21(26%). IA2A and znT8A positivity were positive in 7(8%) of the non GADA positive T1D. Autoantibody positivity in T2D was seen in 4(8%)of clinical type 2 DM patients. Increased frequency of circulating TFH (CD4+ CD45RA- CXCR5+ Tcells) were found in T1D compared to HC. PD1+CXCR5+ T cells and ICOS+ CXCR5+ T cells were correlating with diabetic ketoacidosis. Random C- peptide was undetectable in most type 1 DM and above cutoff of 0.6 ng/ml in type-2 DM patients.

Conclusion: One fourth of young south Indian diabetic patients were negative for islet cell autoantibodies. IA2A positivity was found to be low compared to other ethnicities. Random C-peptide also contributes to distinction between T1DM and T2DM. Antibody and C-peptide status should be always considered with clinical features for definite diagnosis for young diabetes patients. Increased circulating TFH cells are likely to be an etiological factor for T1D and could be considered as an immunotherapeutic target.

Chemical biology and structural studies on the bidirectional allosteric mechanism of regulation of phosphoinositide-dependent protein kinase 1 (PDK1)

Lissy Z. Gross[1], Mariana Sacerdoti[1], Alejandro E. Leroux[1], Abhijeet Ghode[4], Ganesh S. Anand[4], Jorg O. Schulze[2], Sebastian Klinke[3], Ricardo M. Biondi[1, 2]

[1]IBioBA – CONICET – Partner Institute of the Max Planck Society, Argentina, [2]Department of Internal Medicine I, Universitätsklinikum Frankfurt, Germany, [3]Fundación Instituto Leloir, IIBBA-CONICET, Argentina, [4]Department of Biological Sciences, National University of Singapore, Singapore

Phosphoinositide-dependent protein kinase 1 (PDK1) is a master AGC kinase of the PI3K signalling pathway that phosphorylates at least other 23 AGC kinases, being PKB/Akt the most relevant substrate for growth and cell survival, and therefore a potential drug target for cancer treatment. Over the years, our laboratory used a chemical and structural biology approach to study and characterize in detail the allosteric regulation of the catalytic domain of PDK1. We developed small compounds that bind to a regulatory site, termed the *PIF-pocket*, and activate PDK1, mimicking the mechanism of activation of AGC kinases by phosphorylation. Using an integrative approach between biochemistry, crystallography and molecular dynamics, we first demonstrated an allosteric regulation from a regulatory site to the active site, and the existence of the reverse process. This bidirectional allosteric mechanism of regulation between both pockets can therefore be modulated by small molecules that bind to their specific orthosteric site and either enhance or inhibit interactions at the allosteric site. Taking this into consideration, it is not surprising that while the pharmaceutical industry has been developing compounds that bind at the ATP-binding site of kinases, they unwillingly developed drugs that affect protein kinase–protein interactions. Therefore, two compounds that equally inhibit the activity of a given kinase could have drastic differences in cells and patients if they differentially affect the formation of protein kinase complexes. We now provide further evidence of the bidirectional system using hydrogen/deuterium exchange (HDX) experiments and present a rather complete structural model for a kinase that can be modulated bidirectionally with small compounds. On the other hand, could metabolites bind at the active site of protein kinases and physiologically regulate the formation of protein kinase complexes? We found that adenosine binds at the ATP-binding site and allosterically enhances interactions in the *PIF-Pocket* regulatory site of PDK1. The findings open the possibility that the physiological regulation of the kinase complexes may be modulated by metabolites and implies that the metabolic state of cells could directly feedback to the regulation of cell signalling.

In Silico Gene Expression Analysis of P450 87A3 Candidate Gene for Withanolides Production in Datura metal plant.

Madhavi M. Hewadikaram[1, 2], Sanjaya Bathige[1], Veranja Karunaratne[3]

[1]Sri Lanka Institute of Nano Technology, Sri Lanka, [2]University of Sri Jayewardenepura, Sri Lanka, [3]University of Peradeniya, Sri Lanka

Datura metal (family *Solanaceae*) involved in the synthesis of wide varieties of triterpenoids via isoprenoid biosynthesis pathway. The pharmacological values of these triterpenoids are immense, hence these are potent bioactive molecules that needed to be investigated widely. Transcriptomic analysis of *Datura metal* revealed that P450 87A3 like pivotal enzyme might be an ideal candidate involved in the biosynthetic pathway of withanolides in this plant. The present study entails the *in silico* characterization of P450 87A3 from the *Datura metal* plant. The full length of the open reading frame of CYP 87A3 is 1470bp, encoding 490 amino acid residues. The Secondary structure of CYP87A3 is present predominantly in the α-helical form with 47.44% of random coils in 36.34%. Moreover, extended strands (12.88%) and least dominant β-turns (5.32%) were also observed through the Self-Optimized Prediction Method with Alignment (SOPMA) online tool. Three –dimensional structure of CYP87A3-like was determined by the Phyre2 server and 3D structure prediction was obtained using the crystal structure of cytochrome P450 90b1 (PDB: c6a18A) as a template with 100% coverage. Ligand binding sites were predicted using the GALAXY web server and predicted HEME ligands have the highest potential to be bind in 105Y 118L 119M 126H 130R 281F 285A 289T 293T 355I 361R 420A 421F 422G 426R 428C 429V 430G 434S binding residuals. Predicted molecular weight (55.92kDa) and theoretical isoelectric point (8.88) were calculated from the ProtParam tool of ExPASY software. SMART and Conserved Domain Database in NCBI tools revealed characteristic P450 superfamily domain in Pro 33- Gly 470 protein sequences. Four strong possible transmembrane domains were predicted in 1-21, 2-174, 3-205, 4-272 proteins using the TMpred program. Phylogenetic tree constructed by the neighbour-joining (NJ) method (bootstrap 1000) revealed, a discrete evolutionary pathway for *Nicotiana* subfamily and higher similarity to *Solanaceae* subfamily in *Solanaceous* genera. Moreover, Homology analysis showed nearly 88.518% similarity with *Capsicum* that already produced Withanolides as *Datura* plants. These empirical data suggest that the *in silico* characterization of CYP87A3 like might be an ideal candidate for the biosynthesis of withanolides in *Datura metal* plants.

Understanding condensation domain selectivity in non-ribosomal peptide biosynthesis: structural characterization of the acceptor bound state

Candace Ho[1, 2, 3], Thierry Izoré[1, 2], David L. Steer[1, 4], Robert J. Goode[1, 4], Ralf B. Schittenhelm[1, 4], Julien Tailhades[1, 2, 3], Max J. Cryle[1, 2, 3]

[1]Department of Biochemistry and Molecular Biology, Monash University, Clayton, Victoria 3800, Austral, [2]EMBL Australia, Monash University, Clayton, Victoria 3800, Australia, [3]ARC Centre of Excellence for Innovations in Peptide and Protein Science, [4]Monash Proteomics and Metabolomics Facility, Monash University, Clayton, Victoria 3800, Australia

Many medically relevant complex peptides are naturally produced by enzymes known as non-ribosomal peptides synthases (NRPSs). Within these multi-modular assembly lines, condensation (C) domains perform the central function of chain assembly, typically by forming a peptide bond between two peptidyl carrier protein (PCP)-bound substrates. As many antibiotics are produced by NRPS assembly lines in vivo, producing improved versions of these antibiotics requires engineering of the NRPS assembly line. Therefore, it is crucial to understand the mechanism and substrate selectivity exhibited by C-domains given their central role in NRPS biosynthesis.
A universal feature in NRPS biosynthesis is the pantetheine linker between the carrier protein shuttle domains and their substrates. Here, we developed and utilised novel chemical tools focused on this linker to enable structural characterization of key condensation domain complexes. A series of coenzyme A (CoA) probes, precursors that are able to be loaded onto the carrier proteins enzymatically to afford modified pantetheine linkers, were synthesized and loaded on carrier protein domains from the NRPS that produces the siderophore fuscachelin. Structural snapshots of a C-domain in complex with an aminoacyl-carrier protein acceptor substrate were obtained (1.9Å). These structures allow the identification of a mechanism that controls access of acceptor substrates to the active site in condensation domains. The structures of this previously uncharacterized complex also allow us to demonstrate that condensation domain active sites do not contain a distinct pocket to select the side chain of the acceptor substrate during peptide assembly but that residues within the active site motif can instead serve to tune the selectivity of these central biosynthetic domains.

A long-range interaction impacts brightness and pH stability of a dark fluorescent protein

Mian Huang[1], Shu Zhang[2], Ye Zou[1], Mengying Deng[3], Michael Z. Lin[4], Ho L. Ng[1], Jun Chu[3, 5]

[1]Department of Biochemistry and Molecular Biophysics, Kansas State University, Manhattan KS 66502, [2]Department of Minimally Invasive Intervention, Peking University Shenzhen Hospital, Shenzhen 518036, [3]Laboratory for Biomedical Optics and Molecular Imaging, Shenzhen Institutes of Advanced Technology, [4]Departments of Neurobiology and Bioengineering, Stanford University, Stanford, CA 94305, [5]CAS Key Laboratory of Health Informatics, Shenzhen Institutes of Advanced Technology

As biosensors, fluorescent proteins (FPs) have revolutionized biological studies through visualizing *in vivo* and *in vitro* activities of proteins, ions, and other biomolecules. Besides the commonly known visible-color classification, FPs can also be categorized into bright (bFPs) and dark ones (dFPs) by fluorescence intensity. Recently, dFPs start drawing attention due to their low spectrum contamination in FRET experiments. Here we developed a novel monomeric orange dFP, darkmRuby, which is excited at 558 nm and emits fluorescence at 592 nm. The protein features a quantum yield (QY) of 0.05 and a bell-shaped pH profile peaked at pH 6.5. Derived from the bFP mRuby3 (QY = 0.45) containing an unusual *trans* chromophore, darkmRuby could be in a unique conformation leading to the dim fluorescence. To atomically decipher the features, we analyzed the X-ray crystal models of darkmRuby obtained at pH 5.0, 8.0, 9.0, respectively. Noticeable changes of His197 conformation and its interaction with the chromophore are observed in the models, implying its key role in affecting darkmRuby performance. After inspecting the models by B factor analysis, we hypothesized that Met94 and Phe96 influence His197 conformation at a location spatially 15 Å away from His197. Meanwhile, we find a channel-shaped cavity connecting the residues in the barrel. Site-directed mutagenesis shows that substitutions with Thr94 and Tyr96 (from mRuby3's recipe) can dramatically increase the brightness and pH stability of the protein. The crystal model of this mutant displays an expected conformational change of His197, which supports our long-range interaction hypothesis. Results of molecular dynamics simulations suggest that His197 protonation and water molecules hydrogen bonding to Thr94 and Tyr96 in the channel play vital roles in the His197-chromophore interaction, resulting in enhanced brightness and pH stability. To our best knowledge, this is the first long-range interaction reported in FPs. We propose that the interaction is accomplished under the mediation of water molecules in an interior channel of darkmRuby. It gives a new insight into the structure-function relationship of FPs. Moreover, the discovery will provide a new approach to manipulate fluorescence intensity when designing FP biosensors.

Intramolecular regulation of Fis1 activity in fission and mitophagy

ugochukwu k. ihenacho[1], John Egner[1], Megan Harwig[1], Blake R. Hill[1]

[1]Department of Biochemistry, Medical College of Wisconsin, Milwaukee WI 53226

Since its initial identification in yeast, Fis1, a 17kDa mitochondria-outer-membrane protein has been implicated in mitochondrial dynamics, mitophagy, and apoptosis. Despite this importance, the mechanisms by which Fis1 activity is governed in these processes remain unclear. Structural considerations suggests that a highly flexible N-terminal “arm” may mediate interactions with cognate binding partners. We recently demonstrated that the human Fis1 arm, contrary to previous reports, is capable of assuming conformations akin to its distant relative, yeast Fis1p, suggesting a conserved mode of activity regulation across kingdoms. With this premise, we searched for amino acid sequences using structure and homology-based sequence alignments, which identified a uniquely conserved region that we hypothesize to be important for regulating Fis1 function. Using MD simulations, IF-microscopy, and NMR, we show that this region is important for Fis1 fission and mitophagic function.

Assessing The Evolutionary Conservation Of Protein Conformational Diversity Using A Novel Metric Based On Difference Distance Maps

Mallika Iyer[1], Lukasz Jaroszewski[2], Zhanwen Li[2], Mayya Sedova[2], Adam Godzik[2]

[1]Graduate School of Biomedical Sciences, Sanford Burnham Prebys Medical Discovery Institute, CA 92037, [2]Biosciences Division, University of California Riverside School of Medicine, CA 92521

We know that homologous proteins share similar sequence, structure and function, and several measures have been developed to evaluate these similarities. However, we also know that in their native state, proteins are highly flexible and exist in not just one structure, but in a multitude of conformations forming a set called an ensemble. Although knowledge of the ensemble is necessary to fully understand protein function, the relationship between evolutionary distance and similarity of conformational ensembles of homologous proteins has not been systematically evaluated. To compare conformational ensembles we first need to develop a metric to compare the **differences** between pairs or groups of conformations. In this study, we developed such a metric by calculating a difference distance map (DDM) between two conformations and then comparing pairs of DDMs using the correlations between them. This allows us not only to recognize, but also to quantify the similarity of the conformational differences between two pairs of conformations. We used this approach to compare the differences between pairs of alternative conformations of homologous proteins in the Protein Data Bank (PDB). We first identified clusters of similar conformations from the multiple coordinate sets of individual proteins in the PDB and then compared DDMs between pairs of representative conformations of different, but homologous proteins. We show that, as expected, homologous proteins display similar conformational changes but also that this similarity correlates with their sequence identity. This result can be used to improve the relevance of homology modeling in analyzing protein function, where we can model not only individual structures, but also functionally relevant structural changes of the targets.

Preliminary Structural Elucidation of β-(1,3)-glucan Synthase from Candida glabrata Using Cryo-Electron Tomography

Jennifer Jiang[1, 2], Cristina Jiménez-Ortigosa[3], Muyuan Chen[4], Xuyuan Kuang[1, 2, 5], Kelley R. Healey[6], Paul Castellano[1, 2], Nikpreet Boparai[1, 2], Steven J. Ludtke[4], David S. Perlin[3], Wei Dai[1, 2]

[1]Institute for Quantitative Biomedicine, Rutgers University, Piscataway NJ 08854, [2]Department of Cell Biology and Neuroscience, Rutgers University, Piscataway NJ 08854, [3]Hackensack Meridian Health-Center for Discovery and Innovation, Nutley NJ 07110, [4]Department of Biochemistry and Molecular Biology, Baylor College of Medicine, Houston TX 77030, [5]Department of Hyperbaric Oxygen, Central South University, Changsha 410008, China, [6]Department of Biology, William Paterson University, Wayne NJ 07470

Echinocandin drugs have become a front-line therapy against *Candida* spp. infections due to the increased incidence of infections by species with elevated azole resistance, such as *Candida glabrata*. Echinocandins target the fungal-specific enzyme ß-(1,3)-glucan synthase (GS), which is located in the plasma membrane and catalyzes the biosynthesis of ß-(1,3)-glucan, the major component of the fungal cell wall. However, resistance to echinocandin drugs, which results from hotspot mutations in the catalytic subunits of GS, is an emerging problem. Little structural information on GS is currently available because, thus far, the GS enzyme complex has resisted homogenous purification, limiting our understanding of GS as a major biosynthetic apparatus for cell wall assembly and an important therapeutic drug target. Here, by applying cryo-electron tomography (cryo-ET) and subtomogram analysis, we provide a preliminary structure of the putative *C. glabrata* GS complex as clusters of hexamers, each subunit with two notable cytosolic domains, the N-terminal and central catalytic domains. This study lays the foundation for structural and functional studies of this elusive protein complex, which will provide insight into fungal cell wall synthesis and the development of more efficacious antifungal therapeutics.

Conservation of the structurally uncharacterized CtBP C-terminal domain across Metazoa

Dhruva Kadiyala[1], Madeline NIblock[2], Yahui Yang[1], Aanchal Jain[3], Kalynn Bird[2], Akshay Seenivasan[2], Kayla Bertholf[4], Ana-Maria Raicu[5], David N. Arnosti[1]

[1]Department of Biochemistry and Molecular Biology, Michigan State University, East Lansing, MI 48824, [2]Lyman Briggs College, Michigan State University, East Lansing, MI 48825, [3]Okemos High School, Okemos, MI 48864, [4]College of Wooster, Wooster, OH 44691, [5]Cell and Molecular Biology Program, Michigan State University, East Lansing, MI 48824

The C-terminal Binding Protein (CtBP) is a conserved eukaryotic transcriptional corepressor involved in development and disease. CtBP resembles D-2-hydroxyacid dehydrogenases and binds to the NAD(H) cofactor, a feature that links it to the metabolic state of the cell. The conserved catalytic core of the protein has been structurally characterized; however, the structure and function of the C-terminal domain (CTD) of the protein remains elusive. The human CtBP CTD has been shown to assist with, but not be necessary for, tetramerization of the protein. In Drosophila, CtBP isoforms are produced with both a short and long CTD, suggesting that this domain may be necessary only in certain contexts. To uncover the significance of the CTD, we used a comparative evolutionary approach and determined the conservation of this domain across Metazoa. CtBP-like proteins only exist within in Bilateria; vertebrates encode two or more paralogous genes, while most other animals encode a single gene. In many lineages, alternative splicing appears to generate unique isoforms, affecting both the N- and C-terminus of the protein. Metazoan isoforms of CtBP produce CTDs that range from 20 to 350 residues in length, suggesting both long and short versions are produced. Most invertebrates encode long CTDs, but some lineages including Diptera and Hymenoptera independently evolved short isoforms which are distinct from ancestral versions both in composition and in length. An ancestral sequence found across arthropods contains three very highly conserved motifs, one of which is also conserved in vertebrates. In comparison, among vertebrate genes, the C-terminus of CtBP1 is highly conserved, while CtBP2 is much more diverged, especially in Amphibia and fish. Secondary structure analysis of the CtBP CTDs reveals that all divergent forms share a property of being disordered even within the highly conserved motifs, with short α-helical segments predicted in some species. Predicted protein binding domains within conserved disordered regions suggest that the CTD may interact with partners of CtBP. Our comparative evolutionary approach coupled with structural analysis reveals possible functional and regulatory significance of CtBP's divergent C-terminal domain, which may play critical roles in the protein's function in health and disease.

Study of the structural features of cyclosporin D and G in chloroform

Polina P. Kobchikova[1], Sergey V. Efimov[1], Vladimir V. Klochkov[1]

[1]Kazan Federal University

It is well known that the structure of a molecule can influence its properties, both chemical and biological. Having established the Structure Activity Relationships, it can be further used to predict the biological activity of a molecule by its structure, which can be widely used in drug design. By using various cyclosporins, which differ slightly in chemical structure, one can try to better understand how the structure and properties of the molecule are related. Cyclosporins D and G were studied in chloroform solution. To study cyclosporins, high-resolution NMR methods were used, as well as the method of molecular dynamics. Each cyclosporin molecule consists of 11 amino acid residues, of which the second amino acid is the most variable. From the one-dimensional NMR spectra, it was clearly seen that minor conformers are present in CsG at room temperature. In the two-dimensional NOESY spectra one can see that some cross-peaks from certain pairs of atoms appear in different cyclosporins, which tells us about the similarity of their structures. For example, the contact between atoms of the first and sixth amino acid residues was obserevd in both peptides. Thus, using the NOESY spectra, the distances between the atoms were found for both cyclosoporins. Then, using these data, as well as the values of the spin-spin splitting for the amide groups, the structures of the molecules were established. Then, using the molecular dynamics method, the mobilities of the main and side chains were studied. Among the structures obtained by XPLOR, the largest discrepancy has been shown to occur in the region of the 7th residue. In turn, molecular dynamics simulation showed that nanosecond dynamics of the backbone occurs in both cyclosporins.

Structural studies of a GH30 xylanase provide insights into its mechanism of action as a potential biocatalyst

Efstratios Nikolaivits[1], Christos Kosinas[2], Christina Pentari[1], Christian Feiler[3], Manfred S. Weiss[3], Evangelos Topakas[1], Maria Dimarogona[2]

[1]School of Chemical Engineering, National Technical University of Athens, Athens, Greece,
[2]Department of Chemical Engineering, University of Patras, Patras, Greece,
[3]Helmholtz-Zentrum Berlin, Macromolecular Crystallography (HZB-MX), Berlin, Germany

Plant cell-wall polysaccharides are abundant in nature and can be utilized as a source to produce sustainable fine chemicals and biofuels [1]. Xylan, as a hemicellulosic component of plant biomass, is considered a major source of renewable carbon and can be degraded with the assistance of xylanolytic enzymes. The most important among them are endo-β-1,4-xylanases, which are glycoside hydrolases (GHs) that depolymerize the polysaccharide into smaller fragments. Knowledge of structural characteristics of such enzymes, provided by tools of structural biology, boost considerably the effort of designing industrial biocatalysts that can be implemented in recycling processes, as assets of groundbreaking green technologies.
This work is focused on a novel xylanase from Thermothelomyces thermophila, named TtXyn30A. It is a member of GH family 30 (CAZy database), which consists of enzymes specialized for the breakdown of xylans containing 4-O-methyl-D-glucuronic acid (MeGlcA) or D-glucuronic acid (GlcA) side residues. However, unlike most other GH30s, TtXyn30A is a bifunctional enzyme [2], exhibiting an exo-xylobiohydrolase activity on the non-reducing end of the glucuronoxylan chain, in complementarity to the endo-activity which leads to the generation of typical bacterial GH30 UXOS. We have determined the X-ray crystallographic structure of TtXyn30A inactive mutant, alone and in complex with its reaction product (4-O-methyl-α-D-glucuronyl-xylobiose), at 1.36 and 1.85 Å, respectively. The detailed analysis of the two structures, combined with mutagenesis studies, allowed the identification of the structural elements that determine the unique catalytic properties of TtXyn30A. The latter are discussed in the present work, providing a deeper understanding of the function of these biocatalysts.

[1] Kim JY, Lee HW, Lee SM, Jae J, Park YK. Overview of the recent advances in lignocellulose liquefaction for producing biofuels, bio-based materials and chemicals. Bioresour. Technol. Elsevier Ltd; 2019. p. 373–84.
[2] Katsimpouras C, Dedes G, Thomaidis NS, Topakas E. A novel fungal GH30 xylanase with xylobiohydrolase auxiliary activity. Biotechnol Biofuels. BioMed Central; 2019;12:120.

Enhancing protein dynamics studies through coevolution and smFRET

Aishwarya Krishnamohan[1], George Hamilton[2], Narendar Kolimi[2], Hugo Sanabria[2], Faruck Morcos[1]

[1]The University of Texas at Dallas, [2]Clemson University

Abstract:
Aishwarya Krishnamohan[1], George Hamilton[2], Narendar Kolimi[2], Hugo Sanabria[2], Faruck Morcos
[1]Department of Biological Sciences, University of Texas at Dallas, Richardson, TX, USA
[2]Physics and Astronomy, Clemson University, Clemson, SC, USA
Proteins are flexible molecules that adopt different structural conformations in varied time scales. Protein flexibility plays a major role in performing their functions, hence probing into the unique states of a protein and their transition pathways are of great significance to study protein's functions. There are number of experimental and computational methods to investigate the time varying dynamics of protein systems, however, there is still a need to uncover comprehensive information about the alternative functional conformations of a protein and the functionally important intermediates that a protein takes in its conformational dynamics. To confront these challenges, we utilize available sequence information collected in genomic databases to identify important amino acid interactions that help model the functionally relevant conformational space of proteins maintained by evolution. In the present study, we utilize a coarse-grained Structure Based Model (SBM) in synergy with Direct Coupling Analysis (DCA) (SBM+DCA), a coevolution-based hybrid model, to probe into the distinct functional conformations of proteins. In this model, the coevolutionary information of a protein family obtained by DCA is combined with the structural information of a protein from X-ray crystallography or NMR experiments which drives the model to identify an ensemble of distinct conformations of a protein. Using this methodology, we were able to explore the conformational diversity of Calmodulin, PDZ1-2 of PSD-95 protein without the presence of ligands or ligand constraints and discover novel states of Yeast Rad51 protein, PDZ1-2 that are not present in the protein structure repository. Further, we synergized the information from SBM+DCA to identify optimal sites for flurophore-coupling in single-molecule Forster Resonance Energy Transfer (smFRET) experiments and help design smFRET networks. Uncovering the conformational landscape of proteins and their pathways using this coevolutionary model is useful to characterize their functional relevance and may have potential applications in drug design.

Human DND1-RRM2 forms a non-canonical domain-swapped dimer

Pooja Kumari[1], Neel S. Bhavesh[1]

[1]International Centre for Genetic Engineering and Biotechnology (ICGEB), New Delhi, India

RNA Recognition Motifs (RRMs) being the most abundant RNA binding domain in eukaryotes, is a major player in cellular regulation. Several variations in the canonical βαββαβ topology have been observed. We have determined the 2.3 Å crystal structure of the human DND1 RRM2 domain. The structure revealed an interesting non-canonical RRM fold, which is maintained by the formation of a 3D domain swapped dimer between β1 and β4 strands across protomers. We have delineated the structural basis of the stable domain swapped dimer formation using the residue level dynamics of protein explored by NMR spectroscopy and MD simulations. Our structural and dynamics studies substantiate major determinants and molecular basis for domain swapped dimerization observed in the RRM domain.

Amyloid fibril structure of islet amyloid polypeptide by cryo-electron microscopy

Tatsiana Kupreichyk[1, 2, 3], Christine Röder[1, 2, 3], Lothar Gremer[1, 2, 3], Luisa U. Schäfer[1, 2], Karunakar R. Pothula[1, 2], Raimond B. G. Ravelli[4], Dieter Willbold[1, 2, 3], Wolfgang Hoyer[1, 2, 3], Gunnar F. Schröder[1, 2, 5]

[1]Institute of Biological Information Processing (IBI-7), Forschungszentrum Jülich, Juelich, Germany, [2]JuStruct, Centre for Structural Biology, Forschungszentrum Jülich, Juelich, Germany, [3]Institut für Physicalische Biologie, Heinrich Heine University Düsseldorf, Dusseldorf, Germany, [4]The Multimodal Molecular Imaging Institute, Maastricht University, Maastricht, The Netherlands, [5]Physics Department, Heinrich Heine University Düsseldorf, Dusseldorf, Germany

A critical role of the hormone islet amyloid polypeptide (IAPP) is vividly discussed for Type 2 Diabetes (T2D), where amyloid deposits in pancreatic islets consisting of fibrillar IAPP have been associated with beta cell loss. Here, we applied cryo-electron microscopy to elucidate the structure of IAPP fibrils prepared at physiological pH and reconstructed a density map of the main polymorph (PM1) of 4.2 Å resolution. Its atomic model comprising residues 13-37 reveals two S-shaped, intertwined protofilaments (PDB: 6Y1A). The segment 21-NNFGAIL-27, which is essential for IAPP amyloidogenicity, forms the protofilament interface together with tyrosine 37 and the amidated C-terminus. The results rationalize previous data on IAPP fibrils, help to elucidate mechanisms of amyloid formation and toxicity, and support the design of fibril growth inhibitors. Moreover, we revealed similarities between S-fold of IAPP and polymorphs of Alzheimer's disease (AD)-associated amyloid-β (Aβ) fibrils, which is of particular interest in context of epidemiological link between T2D and AD.

Celebrating the evolution of protein design through the lenses of the Protein Data Bank

Salvatore La Gatta[1], Linda Leone[1], Marco Chino[1], Vincenzo Pavone[1], Angela Lombardi[1], Luigi F. Di Costanzo[2]

[1]University of Naples Federico II, Department of Chemical Sciences, [2]University of Naples Federico II, Department of Agricultural Sciences

It is not a coincidence that protein design began in the late 1970s, few years after the institution of the Protein Data Bank (PDB), the archive serving as the global repository for open access of biological macromolecules. Back then, the opportunity of exploring available 3D structures of macromolecules inspired protein designers to draw small sequences that could fold in simple secondary structures, such as an alpha helix. Since the first structure of a short alpha helix named ALPHA-1 protein (1986), protein designers learned more complex folding rules.The design of sequences that could fold in entire new proteins, comprised of more than one secondary structure, and assembled together represents a great achievement.
These days protein designers use the PDB beyond gathering of new macromolecular structures. They explore the archive to get new structural patterns, or any information to compile libraries, or as needed for their designs. Then, they assemble novel configurations of secondary structures not present in natural proteins.
We explored the PDB to find unique structures of designed sequences not longer than 50 residues.We aimed to understand how protein designers are using strategies from de novo design and miniaturization, or by computational methods. This set of structures led us to understand the evolution of protein design architecture over the years and identify among them trends of function by analyzing the binding details of prosthetic groups. Our analysis shows that protein designers are still learning new rules to make computational design more reliable. On the other hand, they are making available examples that represent the application of protein design across several disciplines. These outstanding designed proteins include use in medicine (designed antibodies), nanotechnology (proteins able to interact with nanoparticles), advanced materials (minienzymes), bio- and environmental-technologies (designed photosynthetic proteins). Similar to a roundtrip journey that begins from exploring the PDB, protein designers are reaching continuously new goals, including the allosteric cooperation in the newest de novo designed protein.

Crystal Structure Of Androgen Dna Binding Domain Interacting With Its Response Element (C3(1) ARE)

Xiao Yin Lee[1], Christine Helsens[1], Arnout Voet[2], Frank Claessens[1]

[1]Department of Cellular and Molecular Medicine, KU Leuven, Belgium, [2]Department of Chemistry, KU Leuven, Belgium

Prostate cancer (PCa) is the second leading cancer in the world in men. The major therapeutic target in advanced PCa is the ligand-binding pocket (LBP) of the androgen receptor (AR). The anti-androgens used to treat cancer are binding to this domain and compete with the natural agonist. Over time the cancer grows resistant towards this therapy due to mutations in the AR. Nevertheless, the cancer still typically depends on the AR, indicating other sites of the AR may still serve as drug targets. The DNA binding domain (DBD) is a likely good therapeutic target as it is involved in essential processes such AR dimerization, translocation to the nucleus and binding to the androgen response elements (AREs). However the structure and the DNA binding interactions with AREs are not fully explored yet. Only one crystal structure (pdb: 1R4I) is available. In this structure the DBD is bound to an artificial direct repeat DNA sequences, resulting in an unexpected dimeric head-to-head arrangement. A better understanding on how the AR recognizes and binds to different native DNA sequences which are mainly organized as semi inverted repeats is crucial for rational drug design. We have recently discovered that the natural ARE described in the first intron of the androgen-regulated C3(1) gene also results in a head-to-head binding arrangement of two DBDs. The small variations between the two crystal structures gives us a new understanding of the way AR interacts with DNA, which is a prerequisite for its biological activities as a transcription factor.

Ligand bound structures of an alkanesulfonate monooxygenase provide insight into catalysis

Jeremy Liew[1], Daniel P. Dowling[1], Israa El Saudi[1]

[1]University of Massachusetts Boston, MA 02125

The necessity of sulfur for the production of cysteine and its derivative sulfur-containing biomolecules leads to the biological sourcing of sulfur from a number of pathways. Under the sulfate starvation response, certain bacteria assimilate and convert sulfur from organosulfur compounds via two-component flavin-dependent monooxygenases capable of cleaving the stable C–S bond to produce sulfite. One such enzyme is the monooxygenase MsuD, which converts methanesulfonate (MS–) to sulfite, but neither its structure nor its mechanism are entirely understood. To gain a better understanding, several crystal structure of MsuD from Pseudomonas fluorescens were solved in different liganded states with MS– and/or FMN. These structures represent the first alkanesulfonate monooxygenase structures with bound flavin to be solved, revealing a complex network of interactions connecting ligand binding to conformational changes within the monooxygenase that are likely important for catalysis. Additionally, the ternary MsuD structure with FMN and MS– bound has aided in a better understanding of the C–S cleavage mechanism and a heretofore unknown function of the extended protein C-terminus.

Interaction of rhodium compounds with proteins

Domenico Loreto[1], Antonello Merlino[1]

[1]University of Naples Federico II

Rhodium compounds do not have known physiological functions and for this reason their applications in biological contexts are quite rare. However, Rh complexes possess interesting chemical and physical properties, such as good solubility and stability in aqueous solution and good functional group tolerance. Paddlewheel dirhodium tetracarboxylate complexes have been considered as potential anticancer agents and have been employed for the development of artificial metalloenzymes because of their widespread catalytic activity.With the aim to unveil insight into the interaction of rhodium compounds with proteins, the X-ray structures of the adducts formed upon reaction of dirhodium tetraacetate with the model proteins hen egg white lysozyme (HEWL) and bovine pancreatic ribonuclease (RNase A) have been solved. Under the explored experimental conditions dirhodium tetracetate in part breaks down when reacts with HEWL. This contrasts with what happens when the dimetallic complex reacts with RNase A. In fact, in the dirhodium complex/RNase A adduct, the metal complex binds the protein via coordination of the imidazole ring of a His side chain to one of the dirhodium center axial site. Notably, additional binding of an imidazole to the metal compound can occur in the adduct at solid state upon replacing of an acetate ligand. These structures have been then compared with those of the Rh/protein adducts and of Rh compound/protein complexes reported in the Protein Data Bank. The results indicate that Rh complexes can be introduced in a protein scaffold through linking strategy or by direct coordination of metal center to a protein residue side chain. In latter case, the formation of the coordinative bond can occur both retaining or displacing the Rh ligands. The formation of a Rh/protein adduct does not significantly alter the overall structure of the macromolecule. Rh compounds usually interact with proteins on their surface, or in region of protein that are highly flexible. Rh mainly binds the side chain of His residues. Other residues involved in interaction with Rh complexes are Arg, Lys, Asp, Asn and the C-terminal tail. These results can be useful for rational design of new artificial Rh-containing enzymes.

Targeting HIV-1 from inside and out: HIV-1 gp120 antagonists also inhibit HIV-1 reverse transcriptase by bridging the NNRTI and NRTI sites

Natalie Losada[1, 2], Francesc X. Ruiz[2], Kalyan Das[4], Asim K. Debnath[3], Eddy Arnold[1, 2], Francesca Curreli[3], Alyssa Pilch[2], Kevin Gruber[2]

[1]Department of Chemistry and Chemical Biology, Rutgers University - New Brunswick, NJ, [2]Center for Advanced Biotechnology and Medicine, Rutgers University, New Brunswick, NJ, [3]Lab of Modeling & Drug Design, Lindsley F. Kimball Research Institute, New York Blood Center, NY, [4]Rega Institute for Medical Research, Department of Microbiology and Immunology, KU Leuven, Belgium

Human immunodeficiency virus (HIV) is a retrovirus that targets CD4 cells, which our immune system uses to fight infections. If left untreated, HIV progresses to acquired immune deficiency syndrome (AIDS). The most common type, HIV-1, is typically treated with a cocktail of two or more FDA-approved drugs that include at least one of the two classes of reverse transcriptase (RT) inhibitors – nucleos(t)ide RT inhibitors (NRTIs) and/or non-nucleoside RT inhibitors (NNRTIs). NRTIs, such as AZT (PDB 3V6D), bind at the polymerase active site as triphosphates, incorporate, and inhibit DNA elongation. On the other hand, NNRTIs, like rilpivirine (PDB 4G1Q), allosterically prevent nucleotide incorporation through binding at a pocket ~10 Å away from the active site. Current efforts prioritize compounds that inhibit RT through alternative mechanisms to avoid cross-resistance to existing drugs, e.g., bifunctional inhibitors that connect NNRTIs and NRTIs. HIV-1 research has also aimed to treat with entry antagonists and a new class has been under development by the Debnath lab since 2005. The entry of HIV-1 into host cells is mediated by the envelope glycoprotein gp120 binding to the host cell CD4 receptor. A class of CD4-mimetic HIV-1 entry Inhibitors, called NBD compounds, bind to the Phe43 cavity of gp120 and prevent Env-mediated cell-to-cell fusion of the virus to the host (Curreli et al., Antimicrob. Agents. Chemother., 58:5478-5491, 2014). Interestingly, some of these compounds also inhibit HIV-1 RT. Here, we have analyzed several NBD compounds using RT enzyme inhibition and antiviral activity assays, X-ray crystallography, and molecular docking, to establish structure-activity relationships for inhibiting HIV-1 RT. Intriguingly, two of the compounds interact with both the dNTP- and NNRTI-binding sites, inhibit RT (IC50 <5 μM), and show potent antiviral activity (EC50 <200 nM). Our results show potential for developing a new class of multi-target HIV drug candidates from the lead NBD compounds.

A zinc recruitment mechanism for the last-line antibiotic resistance enzyme MCR-1

Emily Lythell[1, 3], Reynier Suardíaz[2, 3, 4], Philip Hinchliffe[1], Adrian Mulholland[3], James Spencer[1]

[1]School of Cellular and Molecular Medicine, University of Bristol, UK, [2]Complutense University of Madrid, Spain, [3]Centre for Computational Chemistry, University of Bristol, UK, [4]School of Biochemistry, University of Bristol, UK

A combination of structural biology, biochemistry and computational enzymology work come together to offer a novel mechanism for Mobilised Colistin Resistance (MCR) enzymes. MCR enzymes, first identified in 2016, are the first plasmid-mediated resistance mechanism identified against polymyxin antibiotics in Gram-negative bacteria and carry a significant threat to global healthcare.
Conflicting structural evidence regarding zinc stoichiometry of the catalytic domains of MCR-1 and -2 is here resolved with a working mechanism for a mixed-metalation state mechanism. Structures have been deposited in the PDB, by our group and others, showing MCR-1 and -2 can support either a monozinc or a dizinc metalation state. Which state is relevant for the catalytic mechanism, and what are the roles of the zinc sites?
Mutagenesis of the catalytic residue Thr285 to both prevent (valine) and mimic (glutamate) phosphorylation has yielded novel structures to explore the link between phosphorylation, as a proxy for substrate binding, and metal stoichiometry. These structures are presented here for the first time. Biochemical data continues to support a mixed mode of zinc binding and phosphorylation, with both mono- and dizinc enzymes, present in protein preparations.
Complementary computational work, including molecular dynamics simulations and DFT cluster modelling, suggests one zinc ion is sufficient for the first step of the reaction, while two zincs are required for the second. We believe this enzyme operates via a zinc recruitment mechanism, where a second zinc ion, required only for the second step, binds the enzyme along with the second substrate. This mechanism explains the disparate biochemical and structural results, where MCR enzymes are seen in solution and in crystallo in both mono- and dizinc forms.
This novel mechanism along with key structural insights pave the way for the design of MCR inhibitor molecules. Inhibition of MCR enzymes will render colistin useful against multi- or pan-resistant Gram-negative bacterial infections, essential in the fight against antibiotic resistance.
Refs: Lythell *et al.*, 2020, Chem Comm; Suardíaz *et al.*, 2021, OBC; Hinchliffe *et al.*, 2017, Sci. Rep.

Disease-Associated Mutations of Pre-mRNA Splicing Factor U2AF2 Alter RNA Interactions and Splicing

Debanjana Maji[1], Dong Li[2], Eliezra Glasser[1], Mary J. Pulvino[1], Jermaine L. Jenkins[1], Clara L. Kielkopf[1]

[1]Center for RNA Biology, University of Rochester School of Medicine and Dentistry, Rochester NY, USA, [2]Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia PA, USA

Mutations of an essential pre-mRNA splicing factor, U2AF2, are associated with certain human diseases. Normally, U2AF2 recognizes the polypyrimidine (Py) tract preceding the 3 ´ splice site and initiates spliceosome assembly. We first investigated the structural and functional consequences of representative N196K or G301D mutants associated with leukemia or solid tumors. We determined crystal structures of the wild-type and mutant U2AF2 proteins each bound to a prototypical Py tract at 1.5, 1.4 and 1.7 Å resolutions. The N196K residue appears to stabilize the open conformation of U2AF2, consistent with an increased apparent RNA binding affinity of the N196K-substituted protein. The unfavorable promity of the G301D residue to the RNA backbone explains a decrease in the RNA binding affinity of the G301D-substituted protein. We found that expression of either N196K- or G301D-substituted U2AF2 alters splicing of representative transcripts. We next investigated de novo U2AF2 mutations associated with developmental delay and intellectual disabilities, R149W, R150H, R150C and K329del. We observe that these U2AF2 mutants reduce Py tract binding affinity. We determined crystal structures of R149W and R150H bound to the Py tract at 1.4 Å resolution. Preliminary analyses indicate that the bulky R149W residue repositions the N-terminal a-helix and thereby disrupts hydrogen bonds with the terminal Py tract nucleotide. The R150H residue is unable hydrogen bond with the Py tract due to the shortened length of the side chain. Altogether, our results demonstrate that disease-associated mutations of U2AF2 mark functional interfaces and are capable of dysregulating pre-mRNA splicing for disease progression.

Pushing the resolution of mix-and-inject serial crystallography to the limit

Tek Narsingh Malla[1], Marius Schmidt[1]

[1]University of Wisconsin-Milwaukee, Milwaukee, WI 53211, [2]Rice University, [3]North Eastern Illinois University, [4]Cornell University, [5]LCLS

Structural enzymology is now possible at X-ray free electron lasers (XFELs) using mix-and-inject serial crystallography (MISC) technology. In time resolved (TR) MISC, reaction in enzymatic crystals is triggered by mixing with a substrate, and the resulting structural changes are probed by X-ray pulses. We have been studying Mycobacterium tuberculosis β-lactamase (BlaC) as a model system. Previous experiments using ceftriaxone (CEF) as substrate solution demonstrated the robust case for routine (TR) MISC at XFELs. Time resolution, however, has been diffusion limited due to large CEF molecule that has to diffuse into enzyme crystals. With recent experience of binding study with sulbactam (SUB), a 3 times smaller molecule than CEF, we planned to push the time resolution of the MISC to the limit. Besides, SUB, in addition, is an inhibitor that irreversibly binds to BlaC. Here we present the results of lowest time resolution ever achieved with MISC, and the enzymatic reaction of BlaC with SUB.

Putting the pieces together: Structural exploration of plant cellulose synthase oligomerization and function

Lynnicia Massenburg[1], Venu Vandavasi[2], Hugh O'Neill[2,3,] Manish Kumar[4], Tracy Nixon[1]

1 Pennsylvania State University, State College, PA 16802, 2 Oak Ridge National Laboratory, Oak Ridge, TN 37831, 3University of Tennessee, Knoxville, TN 37996, 4University of Texas at Austin, Austin, TX 78712

Arabidopsis thaliana CESA1 catalytic domain (AtCESA1catD) has been shown to form stable oligomers. Negative stain transmission electron microscopy of AtCESA1catD revealed a fit to CESA isoforms, including the *Populus tremula x tremula* CESA8 trimer. While the trimer interactions are stable, interest has shifted into stabilizing domains within the trimer that are poorly resolved, including the intrinsically disordered N-terminal region and the Class Specific Region (CSR), which functions have yet to be fully elucidated. This work seeks to determine structural insights on chemically-stabilized and scaffold-stabilized trimers. Current progress has shown stable trimers in the presence of chemical fixation. A screen on Arctica and Krios has revealed low resolution trimer structures of AtCESA1catD trimers. Future studies will pursue the chemical fixation of full- length CESAs. Scaffolds designed from stable soluble trimers are currently being tested for CESA trimer stabilization by utilizing the Spy-Dock system.

Structural studies of human Fis1 reveals a dynamic region important for Drp1 recruitment and mitochondrial fission

Kelsey A. Meacham[1], John M. Egner[1], Megan C. Harwig[1], Maxx H. Tessmer[2], Elizabeth L. Noey[3], Ugochukwu K. Ihenacho[1], Francis C. Peterson[1], Michael E. Widlansky[4], Blake Hill[1]

[1]Department of Biochemistry, Medical College of Wisconsin, Milwaukee, WI 53226, [2]Department of Microbiology & Immunology, Medical College of Wisconsin, Milwaukee, WI 53226, [3]Department of Biophysics, Medical College of Wisconsin, Milwaukee, WI 53226, [4]Department of Medicine, Medical College of Wisconsin, Milwaukee, WI 53226

Mitochondrial fission protein 1 (Fis1) and dynamin-related protein 1 (Drp1) were initially described as being evolutionarily conserved for mitochondrial fission. Yet in humans, the role of Fis1 in this process is unclear and disputed by many. In budding yeast where Fis1 helps to recruit Drp1 from the cytoplasm to mitochondria for fission, an N-terminal “arm” of Fis1 is required for function. The yeast Fis1 arm interacts intramolecularly with a conserved surface that governs in vitro interactions with yeast Drp1. In human Fis1, NMR and x-ray structures show different arm conformations and its importance for human Drp1 recruitment is unknown. We use molecular dynamics simulations and comparisons to experimental NMR chemical shifts to show the human Fis1 arm can adopt an intramolecular conformation akin to that observed with the yeast protein. This finding is further supported through intrinsic tryptophan fluorescence and NMR experiments on human Fis1 with and without the arm. Using NMR, the human Fis1 arm is also observed to be sensitive to environmental changes. The importance of these findings are revealed in cellular studies where the removal of the Fis1 arm reduces Drp1 recruitment and mitochondrial fission, suggesting an important role for Fis1 in human mitochondrial fission.

The structures of zinc solute binding proteins and accessory factors.

Ady B. Meléndez[1], Daniel Valencia[1], Erik Yukl[1]

[1]New Mexico State University

Bacteria use extracellular solute binding proteins (SBP) to specifically bind various essential nutrients for import into the cell via ATP binding cassette (ABC) transporters. These have been divided into groups based on structure, which also correlates with the specific substrate type. We have focused on the cluster A-I SBPs, which specifically bind either zinc, manganese or iron. These are essential for survival in metal limited environments such as the human host, making them attractive antibiotic drug targets. We have determined zinc-free and zinc-bound structures for the zinc-specific SBP AztC. This protein is unusual in that it works in conjunction with another zinc binding protein AztD, which acts as a metallochaperone delivering zinc to AztC through a direct, associative mechanism. Structures of AztD have allowed us to construct a docking model illustrating the mechanism of this transfer that is consistent with solution studies. Recently, we have begun the characterization of an entirely new family of SBPs that function in zinc import through ABC transporters. These share virtually no sequence homology to known SBPs and likely fold into entirely different structures. They bind zinc specifically and with high affinity in vitro, and we are currently optimizing crystallization conditions with the hope of determining a high resolution structure. Here we present an overview of our work in the field focused on how crystal structures of zinc SBPs have informed our knowledge of binding and transport mechanisms essential to bacterial biology.

POSTER SESSION 3
Tuesday May 4, 2021 2:20pm-3:00pm EDT

Computational analysis of the NEK family of proteins uncovers unique structural features and interactions with other cancer-related proteins

Andriele Dos Santos Silva

The NIMA-related kinases (NEK) are a family of serine/threonine kinases that have been functionally associated in, the assembly of the mitotic spindle, the regulation of the disjunction of the centrosome and the function of the primary cilium. Protein kinases that regulate the centrosome cycle are frequently abnormally controlled in tumor cells. NEK2A, for instance, is often upregulated in numerous types of human cancers. To fully exploit NEKs as therapeutic targets, we must comprehend their structure-function relationship and ascertain the foundation of their mechanism of action. In this study, we show a thorough computational analysis of all eleven NEK proteins. Three-dimensional modeling of the catalytic domains, using several approaches including homology modeling, threading, and ab initio techniques was performed, followed by in silico protein-protein predictions and docking analysis. Our analyses demonstrate unique features present in different NEK family members, such as the armadillo (ARM)-like fold in NEK 10 and the RCC1 repeats found in NEK 8 and NEK 9. We predict unique protein-protein interactions that have not been previously reported such as NEK7 and TNF. These findings are corroborated by our docking analysis that shows predicted interaction via hydrogen bonds and salt bridges in the E-F LOOP area of TNF and specific residues of NEK 7. Overall, this study offers novel and intriguing information about the NEK protein family, unique protein-protein interaction prediction, and lays the groundwork for elucidating their molecular mechanism to examine their potential as drug targets.

Understanding the stability of a chaperone- Gankyrin as a model.

Mukund Sudharsan MG[1], Somavally Dalvi[1], Prasanna Venkatraman[1]

[1]Tata Memorial Centre, Advanced Centre for Treatment, Research and Education in Cancer (ACTREC)

Gankyrin is a non-ATPase proteasomal chaperone, which assembles the 19S regulatory component of the proteasome. Gankyrin was classified as an oncogene after identifying it in a screen for genes overexpressed in hepatocellular carcinoma and its ability to transform NIH3T3 cells1. Overexpression of Gankyrin in cancers leads to poor prognosis in many cancers. It is an early-stage biomarker for oral cancer. Gankyrin is known to interact with Rb, CDK4, MDM2, RhoGDI and control different hallmarks of cancer to promote oncogenesis. In this current study, we have tried to understand the effect of heat, chaotropic agents and pH on the stability of Gankyrin using various approaches. We also report the crystal structure of Gankyrin at 1.8 Å, which is an update to the existing 2 Å structure.

Structure-function characterization of Streptococcus intermedius' surface antigen Pas

Joshua Mieher[1], Norbert Schormann[1], Ren Wu[1], Manisha Patel[1], Sangeetha Purushotham[1], Hui Wu[2], Jessica Scoffield[3], Champion Deivanayagam[1]

[1]Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham, [2]Oregon Health and Science University School of Dentistry, [3]Department of Microbiology, University of Alabama at Birmingham

Pas is an AntigenI/II-family (AgI/II) surface antigen on *Streptococcus intermedius*, a known contributor to subgingival dental plaque, and is associated with purulent infections in multiple organs and cystic fibrosis lungs. To establish potential roles in pathogenesis this study aims to (a) characterize the structure of Pas, (b) investigate its interactions with Gp340, (c) with fibrinogen and (d) inquire in dual-species biofilm studies the interaction of *S. intermedius* with *Candida albicans* and *Pseudomonas aeruginosa*.
The crystal structure of V^{Pas} with a 1.7Å resolution shows a lectin-like fold and a metal binding site, this is homologous to AgI/II V-regions found in other streptococci. The crystal structure of C_{123}^{Pas} was resolved to 2.7Å. Pas was found to adhere to Gp340 domains with nanomolar affinity, as well as to the blood protein fibrinogen with micromolar affinity. In dual-biofilm studies, the Pas deficient mutant with (i) *C. albicans* showed reduction in biofilm, and with (ii) *P. aeruginosa* demonstrated that Pas mediates the interaction and biofilm formation with an acute isolate (PAO1), but not the chronic isolate (FRD1).
Our findings highlight the multifunctionality of *S. intermedius'* surface antigen Pas where the adherence to fibrinogen may facilitate the movement from the oral cavity through the blood stream. Reaching the lungs, the high affinity interaction with Gp340/SRCR would localize *S. intermedius*, where subsequent biofilm formation with *C. albicans* and *P. aeruginosa*, opportunistic CF pathogens, could result in microbially induced disease conditions.

Conformational attuning of flexible loops in PD-1 determines the interplay among structural stability, plasticity and energetics: A route for identification of anti-PD-1/PD-L1 checkpoint inhibitors

Lovika Mittal[1, 2], Narayan Sugandha[1], Amit Awasthi[1], Rajiv K. Tonk[2], Shailendra Asthana[1]

[1]Translational Health Science and Technology Institute, Faridabad, Haryana-121001, India,
[2]Delhi Pharmaceutical Sciences and Research University, New Delhi, Delhi-110017, India

The dynamics and plasticity of the PD-1/PD-L1 axis are the stumbling blocks to finding small-molecule antagonists that can significantly disrupt this interaction interface. The exploration of this protein-protein interaction (PPI) interface at an atomic level is of fundamental biological importance using molecular structure-guided approaches. The mechanistic understanding of the PD-1 pathway and its modulators has greatly benefited from advances in structural biology, biophysical, biochemical, and *in-silico* approaches. Food and Drug Administration (FDA)-approved anti-PD-1 monoclonal antibodies (mAbs), i.e., Pembrolizumab and Nivolumab are the *first-in-class* with distinct binding modes to access this PD-1/PD-L1 axis clinically; however, their dynamical and energetic basis of interactions remained elusive. We have investigated the native plasticity of PD-1 at global (structural and dynamical) and local (residue side-chain orientations) levels with its endogenous ligand PD-L1 and the reported potent mAbs. We found that the structural stability and coordinated Cα movements are increased in the presence of PD-1's binding partners. The comprehensive computational biophysical approaches enabled a detailed study of PD-1 PPI interfaces that revealed the intrinsic plasticity of PD-1, its concerted loops' movement (BC, FG, and CC′), distal side-chain motions, and the thermodynamic landscape are significantly perturbed from its unbound to bound states. Based on intra-/inter-residues' contact networks and energetics, we have identified two *hot-spots* zones-'P-zone' and 'N-zone', that are found to be essential to arrest the dynamical motions of PD-1 significantly for the rational design of therapeutics by mimicking the mAbs binding signature. We have also identified and experimentally validated low molecular weight molecules with a considerable binding affinity and inhibition activity towards the PD-1/PD-L1 interface. Overall, we provide a novel insight into the rational design of therapeutic agents by mimicking the mechanisms of mAbs and suggests that the C'D, BC, and FG loops of PD-1 are critical regions to target for drug discovery. We present the essential atomic-level data for understanding PPI interfaces and their underlying mechanisms, which will aid in ongoing research on the PD-1/PD-L1 axis.

Protein design strategies for co-crystallization of pre-T-cell receptor and peptide bound major histocompatibility complex molecules

Réka Mizsei[1, 2], Wan-Na Chen[2], Xiaolong Li[1, 2], Jia-huai Wang[1, 2], Gerhard Wagner[2], Ellis L. Reinherz[1, 2], Robert J. Mallis[1, 2]

[1]Dana-Farber Cancer Institute, Boston MA 02215, [2]Harvard Medical School, Boston MA 02115

T lymphocytes discriminate between healthy versus infected or cancerous cells via T-cell receptor (TCR) mediated recognition of peptides bound and presented by cell-surface-expressed major histocompatibility complex molecules (MHCs). The thymic education creates functional T cell repertoire; during one early checkpoint in this process, the β chain first pairs with a surrogate pTα chain to form the preTCR. Although early studies suggested that preTCRs signal autonomously; it has been shown recently that peptide bound MHC molecules (pMHCs) are preTCR ligands that promote thymocyte development. However, the lack of appropriate tools has limited prior efforts to describe the atomic details of the preTCR-pMHC complex structure. To facilitate preTCR-pMHC complex formation we covalently linked preTCRβ chain, the sole ligand-binding subunit of the preTCR, to its pHMC ligands. In the presentation we demonstrate a small-scale linkage screening protocol using bismaleimide linkers for determining residue-specific distance constraints within the preTCR-pMHC complex. We also describe the concept of linkage specificity as a protein residue level affinity and distance measure. Using the linkage methodology an X-ray crystallographic structure of the preTCRβ-pMHC complex has been solved by our collective team. preTCRβ-pMHC orientation is in contrast to the established binding modality of mature TCRs, whereby both α and β chains bind in a vertical orientation to pHMC. The docking presented here points to the C-terminal part of the MHC-bound peptide as playing a key role in specific peptide recognition, which was further clarified by alanine screening of the peptide as monitored by changes in linkage specificities. The solution interaction geometry of preTCRβ-pMHC complex is also verified with paramagnetic pseudocontact chemical shift (PCS) NMR of the unlinked protein mixtures as an orthogonal approach. The docking geometry of the preTC-pMHC complexes help to solve the mystery of how functionally rearranged β chains can properly engage with pMHC; therefore, highlighting the physiological importance of an early thymocyte developmental stage that contributes to the diversification the T-lineage repertoire that eventually move into periphery to monitor and patrol the body.

A basement membrane protein Nephronectin plays important roles in tooth development

Kanji Mizuta[1], Keigo Yoshizaki[1], Tomomi Yuta[1], Kanako Miyazaki[1], Keita Funada[1], Tian Tian[1], Yao Fu[1], Yuta Chiba[2], Ichiro Takahashi[1], Satoshi Fukumoto[2]

1Section of Orthodontics, Faculty of Dental Science, Kyushu University, Fukuoka, Japan, 2Section of Pediatric Dentistry, Faculty of Dental Science, Kyushu University, Fukuoka, Japan

Tooth development is initiated by epithelial-mesenchymal interactions. The basement membrane between the epithelium and mesenchyme mediates as an essential scaffold that regulates various signals. We focused on Nephronectin (NPNT), which is specifically localized in the basement membrane during epithelial-mesenchymal interaction. Npnt has an EGF-like repeat domain on the N-terminus and an RGD domain, known as an integrin binding domain, on the C-terminus, and can be assumed to have a unique protein function.Previously we reported that the EGF-like repeat domain of NPNT promotes the induction of odontogenic epithelial stem cells into ameloblasts by repressing the expression of SOX2. In this study, we revealed that addition of BMP2 decreased the mRNA level of Egfr, so that the signaling of EGF-like repeat domain is downregulated. Pre-ameloblasts with reduced EGFR expression switch the function of NPNT to the RGD domain and accelerate their differentiation into ameloblasts. In addition, we analyzed the function of RGD domain of NPNT using Npnt-ΔEGF and Npnt-ΔRGD expression vectors, lacking the EGF-like repeats and the RGD domain of NPNT. Npnt-FL and Npnt-ΔEGF overexpression in dental epithelial cells (M3H1) resulted in increased expression of ameloblastin (Ambn), which is a differentiation marker for ameloblasts. However, overexpression of Npnt-ΔRGD didnot alter the expression of Ambn in M3H1 cells. These data suggest that the RGD domain of NPNT may play an important role for ameloblast differentiation.In summary, NPNT regulates the development of dental epithelial cells by using the function of EGF-repeat domain during morphogenesis stage, followed by switching to the function of RGD domain in differentiation stage. NPNT is a unique protein that has two different functions by using the two functional domains properly

Mining the Protein Data Bank for Carbohydrate Data using GlyFinder

David Montgomery[1], Robert J. Woods[1]

[1]Complex Carbohydrate Research Center, University of Georgia, 315 Riverbend Rd, Athens, GA 30602

Since its inception 50 years ago, nearly 180,000 structures have been deposited in the Protein Data Bank (PDB). Nearly 30% of these structures contain carbohydrates covalently attached to the proteins or interacting with them as ligands, and the ability to mine this wealth of data is of significant interest to the scientific community. Unfortunately, a lack of carbohydrate features in the crystallographic software and inconsistent community standards for naming carbohydrates has caused a significant number of structural errors and incorrect or inconsistent naming for the carbohydrates in the PDB. To help address these issues, our group assisted the PDB Biocuration Team with their carbohydrate remediation project, an effort that was completed last summer. Our software, which names carbohydrates based on their structure and reports any missing atoms or mismatches between the residue name and the detected sugar, was used to remediate the PDB archive and has been integrated into the PDB OneDep system to check incoming structures. In addition, we collected structural information, including atomic coordinates, glycosidic torsion angles, ring shape, and chemical derivatives, and created a database of the nearly 200,000 carbohydrates we detected in the PDB. To facilitate data mining of this carbohydrate structural data, we have developed GlyFinder (www.dev.glycam.org/gf), a web interface to search our database. Here we present an overview of the carbohydrate data available and highlight important statistics and analyses that are nearly impossible to perform without using GlyFinder.

Investigating the Structure and Function of the Arf-GEFs Gea1 and Gea2

Arnold Muccini[1, 2], J. Christopher Fromme[1, 2]

[1]Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, [2]Weill Institute for Cell and Molecular Biology

In eukaryotes, vesicle formation at the Golgi complex is initiated by the conserved small GTPase Arf1. Arf1 is responsible for multiple trafficking pathways at the Golgi including retrograde recycling to the ER, intra-Golgi tracking, and exit from the Golgi at the TGN. These pathways each occurring at different stages of the Golgi maturation process, require precise regulation of Arf1. In budding yeast, Arf1 localization and activity is controlled by three separate guanine nucleotide exchange factors (GEFs): Gea1, Gea2, and Sec7. The paralogs Gea1 and Gea2 are responsible for activating Arf1 at the early Golgi to commence COPI vesicle formation which is essential for retention of resident proteins at the endoplasmic reticulum and early Golgi. Yet, the mechanism by which Gea1 and Gea2 are recruited to the Golgi at the right place and time to activate COPI vesicle formation is unclear. Having a structure of the full-length protein would facilitate our understanding of how Arf1 activation is regulated, however past attempts at structure determination using X-ray crystallography have failed. In this poster I will present my work in determining the full-length structure of Gea2 bound to Arf1 using cryo-EM to 3.5Å resolution. This data has enabled me to build an atomic model of the complex, which I will be able to use to test hypotheses regarding how these proteins are regulated and controlled.

CopC from Pseudomonas fluorescens as a model system to explore the copper histidine brace

Sebastian J. Muderspach[1], Johan Ø. Ipsen[2], Cristina H. Rollán[3], Søren Brander[4], Andreas B. Bertelsen[3], Poul Erik Jensen[5], Morten H. Nørholm[3], Katja S. Johansen[4], Leila Lo Leggio[1]

[1]Department of Chemistry, University of Copenhagen, DK-2100 Copenhagen Ø, Denmark., [2]Department of Plant and Environmental Sciences, Copenhagen University, Frederiksberg, DK-1871, [3]The Novo Nordisk Foundation Centre for Biosustainability, Technical University of Denmark, DK-2800, [4]Department of Geosciences and Natural Resource Management, Copenhagen University, Frederiksberg, DK, [5]Department of Food Science, University of Copenhagen, Rolighedsvej 26, Frederiksberg, DK-1958

The copper binding histidine brace shows high diversity in terms of functionality between proteins. The histidine brace consists of an N-terminal histidine and an internal histidine that donates three nitrogen atoms for copper chelation. In some proteins, such as the Lytic Polysaccharide monooxygenases, the copper in the histidine brace is utilized for oxidative reactivity. This makes the enzymes able to degrade recalcitrant polysaccharides. In other proteins such as CopC the histidine brace is used to transport copper and reactivity is therefore highly undesirable. However, the differences leading to this functional diversification in the histidine brace is not very well understood. This study attempts to elucidate the structural differences that accounts for the functional diversity using CopC from Pseudomonas fluorescens as a model system. In addition to two histidine residues, an aspartic acid residue is also involved in Cu coordination in CopC. Several variants were produced with alterations to residues directly coordinating the copper (primary coordination sphere) and/or to residues that coordinate to the directly coordinating residues (secondary coordination sphere). These variants were characterized using isothermal titration calorimetry, reactivity assays, differential scanning fluorimetry and X-ray crystallography. The study identifies the residues important for copper binding both from the primary and the secondary coordination sphere. Variants that retained the histidine brace, but altered the aspartic acid of the primary coordination sphere showed a complete inability to bind copper(II), while a mutation to a strictly conserved secondary sphere residue shows a weakened copper binding. None of the 15 PfCopC variants in this study were redox active at ambient conditions with ascorbate as reductant.

Crystal structures of Scone, pseudosymmetric folding of a designer protein

Bram Mylemans[1], Laurens Vandebroek[1], Arnout Voet[1]

[1]Departement of Chemistry, KU Leuven, Leuven, Belgium

Recent years have seen a rise in the development of computationally designed proteins including symmetric ones. We recently developed a nine-fold symmetric β-propeller protein named Cake.This protein sequence possesses structural plasticity. When eight repeats are expressed it forms a stable eight-bladed β-propeller, when nine are expressed it will adopt a nine-bladed structure.
Here, we wanted to further engineer this protein to a three-fold symmetric nine-bladed propeller using computational design, with the idea that it would prevent the formation of the eight-fold variant. Two nine-bladed propeller proteins were designed, named Scone-E and Scone-R. Both proteins could be purified to a high yielding mono disperse fraction. Circular dichroism confirmed that the protein adopted a similar fold to the Cake proteins. However, Crystallography revealed that both designs folded as an eight-bladed propeller with distorted termini, leading to a pseudo-symmetric protein. In addition to this unexpected behavior, Scone-E could only be crystallized upon addition of the polyoxometalate, STA. This molecule interacts with multiple protein chains thus creating more crystal contacts, resulting in higher symmetric space groups. This research shows that the design of β-propeller with a specific symmetry remains difficult and that addition of STA can facilitate crystals formation and the crystal properties.

Integrative modeling of SARS-COV-2 envelope structure

Oleksandr Narykov[1], Weria Pezeshkian[2], Tsjerk Wassenaar[2], Senbao Lu[1], Fabian Grünewald[2], Siewert-Jan Marrink[2], Dmitry Korkin[1]

[1]Worcester Polytechnic Institute, [2]University of Groningen

The year 2020 brought one of the largest pandemic of the past century, caused by the SARS-CoV-2 virus. Scientific community responded with unprecedented collaborative efforts, unraveling genomic, structural and interactome information about this virus with the goal of finding a cure. Currently, even though we have a range of highly efficient vaccines that successfully combat the spread of the virus, it is unlikely that COVID would follow footsteps of smallpox - the only eradicated viral infection. Furthermore, we already see a wide range of variants arising in different locations across the globe. Thus, novel therapies might be needed, such as nanoparticles that mimic the virus delivering the drug. Besides a well-studied (S)pike protein, (M)embrane protein, being the main building block of the viral envelope, is another lucrative drug target and requires structural characterization. Even though a large part of SARS-CoV-2 structural repertoire has been experimentally resolved, we have a limited understanding on the functional complexes the proteins form; to date, there is no experimentally solved structure of M dimer. Even less is known about the overall morphology of the viral envelope, which consists of lipids, several dozens spike trimers, a few (2-3) (E)nvelope pentamers, and a large amount (~1,000) of M dimers constantly interacting with each other. We present a physically tractable mesoscale system of viral envelope that integrates the most recent information about the viral envelope (protein structures, stoichiometry, morphology, and geometry) into series of molecular dynamics simulations to study the behavior of envelope structure in solvent. The model was ultimately simulated for 4s by leveraging TACC Frontera, one of the most powerful supercomputers in the world, and remained stable. The simulation showed that M dimers on the surface exhibit self-ordering behavior, highly resembling previous experimental microscopy findings for SARS virus. It allows one to explain how virus can maintain sufficient flexibility for the membrane fusion while remaining stable. These results would establish structural basis to new antiviral drug targets at the whole virion scale, providing foundations for studying virion structures of other coronavirus.

Glucose-based molecular rotors “shine” at the catalytic site of glycogen phosphorylase fostering the design of new antidiabetic agents

Dionysios D. Neofytos[1], Michail-Panagiotis Minadakis[2], Kostantinos F. Mavreas[2], Maria Paschou[3], Michael Mamais[2], Panagiota Papazafiri[3], Thanasis Gimisis[2], Evangelia D. Chrysina[1]

[1]Institute of Chemical Biology, National Hellenic Research Foundation, [2]Department of Chemistry, National and Kapodistrian University of Athens, [3]Department of Biology, National and Kapodistrian, University of Athens

The increasing prevalence of diabetes has stimulated the market potential for new, cost-effective therapeutic agents. The structure-based drug design (SBDD) approach has provided the major impetus needed for shortening the drug discovery timeline. Glycogen phosphorylase (GP) is one of the key protagonists in glycogen degradation that controls sugar levels in the blood circulation in type 2 diabetes patients. GP binding sites have been portrayed with SBDD; specifically, its catalytic site has been probed with more than 200 ligands, glucose-based analogues and their complex structures have been deposited with the PDB. The knowledge that stems from the 3D structure of the complexes coupled with the geometry of the binding site has fostered the development of potent GP inhibitors in the low nM range. With the aim to articulate the structural features of the active site that prescribe ligand binding and develop new probes for monitoring ligands at cellular level, a new family of synthetic fluorescent compounds has been explored, named molecular rotors. Molecular rotors are a family of synthetic fluorescent compounds which undergo intramolecular twist, when subjected to photoexcitation. A new glucose-based molecular rotor (RotA) has been synthesized and characterized by spectroscopic and biochemical studies and its 3D structure has been determined in complex with GP at 1.8 Å resolution by X-ray crystallography at PETRAIII DESY/EMBL-Hamburg. The results showed that RotA is a potent inhibitor of GP, binds tightly at the catalytic site of the enzyme and exhibits molecular rotor properties as it becomes evident by fluorescent and preliminary cell-based assays performed.
The work has been supported by “INSPIRED-The National Research Infrastructures on Integrated Structural Biology, Drug Screening Efforts and Drug target functional characterization” (MIS 5002550); Action “Reinforcement of the Research and Innovation Infrastructure”, Operational Programme "Competitiveness, Entrepreneurship and Innovation" (NSRF 2014-2020) and co-financed by Greece and the European Union. This research is also co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning 2014- 2020» in the context of the project “Synthesis and study of fluorescent molecular Rotors as cellular probes for enzymatic function” (MIS 5048135).

Oligomerization of the Human Adenosine A2A Receptor Is Driven by the Intrinsically Disordered C-Terminus

Khanh D. Nguyen[1], Michael Vigers[2], Eric Sefah[3], Susanna Seppälä[2], Jennifer P. Hoover[1], Nicole S. Schonenbach[2], Blake Mertz[3], Michelle A. O'Malley[2], Songi Han[1, 2]

[1]Department of Chemistry and Biochemistry, University of California - Santa Barbara, CA 93106, [2]Department of Chemical Engineering, University of California - Santa Barbara, CA 93106, [3]C. Eugene Bennett Department of Chemistry, West Virginia University, Morgantown, WV 26506

G protein-coupled receptors (GPCRs) have long been shown to exist as oligomers with functional properties distinct from those of the monomeric counterparts, but the driving factors of GPCR oligomerization remain relatively unexplored. In this study, we focus on the human adenosine A2A receptor (A2AR), a model GPCR that forms oligomers both in vitro and in vivo. Combining experimental and computational approaches, we discover that the intrinsically disordered C-terminus of A2AR drives the homo-oligomerization of the receptor. The formation of A2AR oligomers declines progressively and systematically with the shortening of the C-terminus. Multiple interaction sites and types are responsible for A2AR oligomerization, including disulfide linkages, hydrogen bonds, electrostatic interactions, and hydrophobic interactions. These interactions are enhanced by depletion interactions along the C-terminus, forming a tunable network of bonds that allow A2AR oligomers to adopt multiple interfaces. This study uncovers the disordered C-terminus as a prominent driving factor for the oligomerization of a GPCR, offering important guidance for structure-function studies of A2AR and other GPCRs.

Bazedoxifene is a Micromolar SARS-CoV-2 Receptor-Binding domain/ACE2 complex Inhibitor

Olaposi I. Omotuyi[1, 2]

[1]Department of Biochemistry, Adekunle Ajasin University, Akungba-Akoko, [2]Bio-Simulation & Cheminformatics Unit, Mols&Sims, Ado Ekiti

Aims: Coronavirus spike glycoprotein (gp) and its interaction with host receptor angiotensin converting enzyme 2 (ACE2) is one of the most structurally understood but therapeutically (therapeutics) under-tapped pathogenesis of COVID-19; the current study aims at screening drug building-blocks for RBD/ACE2 interface binding; for a low-energy binding candidate, FDA-approved drug library was screened to identify drugs with similar scaffold; which is then re-evaluated *in vitro*.
Methods: Virtual screening was performed using consensus (Vina, and Glide Xp/Sp) scheme, molecular dynamics simulation was performed using GROMACS software, whilst in vitro evaluation of RBD/ACE2 inhibition was a competitive binding assay.
Findings: In this study, virtual screening of 12000 chemical building blocks into the interface of GP/ACE2 showed the benzimidazole (BAZ) scaffold as potent scaffold for possible GP/ACE2 inhibitor. Tanimoto-coefficient based comparison of filtered FDA-approved drug library identified 12 drugs with the benzimidazole-like substructure. When these compounds were re—docked into GP/ACE2 interface, the consensus docking identified bazedoxifene; in vitro GP-receptor-binding domain/ACE2 inhibition kinetics showed micromolar IC50 value (1.131 mM). Consequently, molecular dynamics simulation of RBD/ACE2 in the presence BAZ resulted in loss of contact and specific hydrogen-bond interaction required for RBD/ACE2 stability.
Significance: Taken together, These findings identified benzimidazole scaffold as a building block for developing novel RBD/ACE2 complex inhibitor and provided mechanistic basis for the use of bazedoxifene as a repurposeable drug for the treatment of COVID-19 acting at RBD/ACE2 interface.

Nucleosome Invasion Mechanism of the Pioneer Transcription Factor Sox

Ramachandran Boopathi[1, 2, 3], Burcu Ozden[4, 5], Imtiaz N. Lone[4], Ayse Berçin Barlas[4, 5], Jan Bednar[1], Carlo Petosa[2], Seyit Kale[4], Ali Hamiche[6], Dimitar Angelov[3], Ezgi Karaca[4, 5], Stefan Dimitrov[1]

[1]Université Grenoble Alpes, CNRS UMR 5309, INSERM U1209, Institute for Advanced Biosciences (IAB), Si, [2]Université Grenoble Alpes, CNRS, CEA, Institut de Biologie Structurale (IBS), 38000 Grenoble, France, [3]Université de Lyon, Laboratoire de Biologie et de Mode´lisation de la Cellule LBMC, [4]Izmir Biomedicine and Genome Center, Dokuz Eylul University Health Campus, [5]Izmir International Biomedicine and Genome Institute, Dokuz Eylül University, [6]Université de Strasbourg,Institut de Génétique et Biologie Moléculaire et Cellulaire (IGBMC)

Pioneer transcription factors (PTFs) have the remarkable ability to bind to chromatin directly to stimulate vital cellular processes. Up to date, several studies have been carried out to dissect the interaction landscape between PTFs and chromatin. However, the precise invasion mechanism of chromatin by PTFs is yet to be resolved. Expanding on this, here, we aim to disclose the mechanistic principles of Sox PTF binding to the nucleosomes. For this, we have combined high-end experimental and in silico approaches, while concentrating on three super helical locations (SHLs) as binding sites, namely SHL0, SHL+2, SHL+4. Our results have shown that Sox can bind to any SHL site, while it can specifically recognize its binding motif only at off-dyad (off-SHL0) positions. We have found that histone-DNA interactions dictate this selective binding by facilitating different DNA thermal fluctuations and DNA bendability at on- and off-dyad SHLs.

Crystallography of Human Carbonic Anhydrases with Inhibitors for Drug Design

Vaida Paketuryte[1], Alexey Smirnov[1], Elena Manakova[2], Saulius Gražulis[2], Daumantas Matulis[1]

[1]Department of Biothermodynamics and Drug Design, Life Sciences Center, Vilnius University,
[2]Department of Protein – DNA Interactions, Life Sciences Center, Vilnius University

Carbonic anhydrases (CAs) are a family of zinc metalloenzymes that are important targets for the treatment of numerous diseases and conditions such as glaucoma, edema, and epilepsy. Many drugs have also been used as diuretics, and currently there are efforts to target CAIX selectively in cancer. In humans, there are 12 catalytically active isoforms, namely, CAI, CAII, CAIII, CAVII, CAXIII – cytosolic, CAIV, CAIX, CAXII, CAXIV - membrane-bound, CAVA and CAVB – mitochondrial, and CAVI - secreted. Their structural folds and active sites are very similar among isoforms making it difficult to design isoform-specific inhibitors. The first CA crystal structure of CAII was determined in 1972 by Liljas et. al. and now 1081 (reviewed on 08/03/2021) structures of catalytically active human isoforms have been deposited to the PDB. We are still missing structures of CAVA and CAVB.
In our group, we have determined and deposited to the PDB 101 X-ray crystallographic structures of more than 60 different compounds bound to five CA isoforms, CAI, CAII, CAIV, CAXII, and CAXIII. Most of the structures of CAXII (20 of 26) and CAXIII (13 of 16) in the PDB are from our laboratory. This structure determination is part of the dataset of our-designed 900 primary sulfonamide compound binding thermodynamic parameters for all human CA isoforms [1]. This effort has led to the design of compounds that interact with tumor-associated CAIX in the picomolar affinity range and exhibit thousand-fold selectivities over other CA isoforms. The 9 out 21 structures of CAIX in the PDB contain our inhibitors. The combined structural and thermodynamic investigation [2] may yield general principles of drug design for other target proteins.
References:
[1] Linkuviene, V. et. al. Thermodynamic, Kinetic, and Structural Parameterization of Human Carbonic Anhydrase Interactions toward Enhanced Inhibitor Design. Quarterly Reviews of Biophysics 2018, 51.
[2] Smirnov, A. et. al. Correlations Between Inhibitor Binding Thermodynamics and Co-Crystal Structures with Carbonic Anhydrases. In Carbonic Anhydrase as Drug Target: Thermodynamics and Structure of Inhibitor Binding; Matulis, D., Ed.; Springer International Publishing: Cham, 2019; Chapter 17, pp 249–261.

In silico structural & in vitro functional analysis of GPR120 receptor to screen potential anticancer candidates.

Ajay Pal[1, 2], James F. Curtin[1], Gemma K. Kinsella[1]

[1]School of Food Science and Environmental Health, College of Sciences and Health, TU Dublin,
[2]Environmental Sustainability and Health Institute (ESHI), Grangegorman, TU Dublin

The G-protein coupled receptor - GPR120 has recently been implicated as a novel target for colorectal cancer (CRC) and other cancer managements. In this study, a homology model of GPR120S (short isoform) was generated to identify potential anti-cancer compounds targeting the GPR120 receptor using a combined *in silico* docking-based virtual screening (DBVS), structure-activity relationship (SAR) study, molecular dynamics (MD) simulations and *in vitro* screening approach. ZINC and SPECS databases of synthetic chemical compounds were screened using the developed GPR120S model to identify molecules binding to the orthosteric pocket followed by an AutoDock SMINA rigid-flexible docking protocol. DBVS hit molecules were then tested *in vitro* to evaluate their cytotoxic activity against SW480 – human CRC cell line expressing GPR120. The test compound **1** (3-(4-methylphenyl)-2-[(2-oxo-2-phenylethyl)sulfanyl]-5,6-dihydrospiro(benzo[h]quinazoline-5,1'-cyclopentane)-4(3H)-one) showed ~90% inhibitory effects on cell growth with micromolar affinities (IC_{50} = 23.21-26.69μM). Finally, SAR analysis of compound **1** led to the identification of two more active compounds from the SPECS database showing better efficacy during cell-based cytotoxicity assay – **5** (IC_{50} = 5.89 - 6.715 μM) and **7** (IC_{50} = 6.789 - 7.502 μM). Simultaneously, MD simulations of DBVS hits interacting with a key residue Asn313 led to a predicted antagonism of GPR120S. The GPR120S homology model generated, and SAR analysis conducted by this work discovered a potential chemical scaffold, dihydrospiro(benzo[h]quinazoline-5,1'-cyclopentane)-4(3H)-one, which will aid future research on anti-cancer drug development for CRC management.

Studying the 50S Ribosomal Subunit Assembly Process for New Antibiotic Target Discovery

Armando Palacios[1, 2], Amal Seffouh[1, 2], Dushyant Jahagirdar[1, 2], Kaustuv Basu[1, 2], Joaquin Ortega[1, 2]

[1]Department of Anatomy and Cell Biology, McGill University, Montreal, Quebec H3A 0C7, Canada, [2]Centre for Structural Biology, McGill University, Montreal, Quebec H3G 0B1, Canada

The rapid emergence of multi- and pan-drug resistant bacteria across the globe is alarming. It's estimated that by 2050, ten million people will die each year from infections caused by drug-resistant bacteria. Widespread antibiotic resistance and the lack of new antibiotics has caused researchers to look for new targets to combat bacterial infections. Bacterial ribosome biogenesis is a promising target for new antibiotic development. Despite the ribosome being a common target for antibiotics, its assembly pathway has remained relatively unexplored. YphC is an under characterized and widely conserved enzyme that plays multiple roles in the late-stage assembly of the bacterial 50S ribosomal subunit. Previous studies have revealed biochemical and structural information pertaining to the GTPase function of YphC. We are using microscale thermophoresis and cryogenic electron microscopy to study how and when YphC intervenes in this assembly process. We have found out that it plays a major role in the formation of several important ribosomal landmarks, like the central protuberance and the E, P and A sites. Further, YphC depletion stalls the 50S subunit in an immature, "locked" conformation that is unable to progress through the canonical assembly pathway.

Redox-Fyn-c-Cbl (RFC) pathway Regulates O-2A/OPC Cell Cycle Exit and Differentiation

Yunpeng Pang[1], Christopher J. Folts[2], Ibro Ambeskovic[2], Mark D. Noble[2]

[1]Neuroscience graduate program, University of Rochester, Rochester, NY, [2]Department of biomedical genetics, University of Rochester, [3]Vertex Pharmaceuticals

Thyroid hormone (TH) plays an important role in the development of the central nervous system (CNS). In particular, TH signal is critical for the differentiation of oligodendrocyte precursor cells (O-2A/OPCs) into oligodendrocytes, which myelinate the CNS. Although gene regulations by TH and TH receptors has been the focus of the field, it is still not clear how TH initiates cell cycle exit in dividing OPCs. Previous work from the Noble laboratory revealed that intracellular redox status is a central regulator of OPC fate decisions between cell division and differentiation. Here, we report the critical role of c-Cbl, a tumor suppressor protein, in regulating the cell cycle exit of dividing OPCs in response to TH signal via the redox-Fyn-cCbl (RFC) pathway. We hypothesized that both Fyn and c-Cbl are required for the cell cycle exit of O-2A/OPCs in response to TH signaling. We performed shRNA knockdown of c-Cbl, Fyn, or scrambled control in primary OPC isolated from neonatal rats before exposure to TH and allow them to differentiate in vitro. We found that compared to scrambled control, both c-Cbl KD and Fyn KD OPCs failed to exit the cell cycle and did not express oligodendrocyte marker GALC. Additionally, we found that other glial differentiation signals including TGF-beta 1 and BMP4 converge on the activation of the RFC pathway and that they require generation of reactive oxygen species (ROS) to drive OPC cell cycle exit. Taken together, our study provides a novel molecular pathway by which ROS signaling from multiple distinct signaling molecules regulate OPC cell cycle in additional gene regulation and antigen expression.

Do LCP proteins represent the key to biofilm inhibition?

João Paquete-Ferreira[1], Jayaraman Muthukumaran[1], João Ramos[1], Filipe Freire[1], Marino F. A. Santos[1], Márcia A. S. Correia[1], Maria João Romão[1], Alexandra R. Fernandes[2], Teresa Santos-Silva[1]

[1]UCIBIO-NOVA, Dep. Química, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, [2]UCIBIO-NOVA, Dep. Ciências da Vida, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa

Bacterial biofilms represent a serious health threat worldwide, as they are many times associated with the failure of indwelling medical devices, and persistent and chronic infections. So far there are no drugs available to address this problem, making the development of new therapeutic strategies of pivotal importance.
Wall teichoic acids (WTAs) are important components of the bacterial cell wall crucial for bacterial survival and adherence to surfaces. Depleting bacteria from WTAs impairs their growth and division, increasing its susceptibility to antibiotic drugs and hence, contributes to fight infection. The biosynthesis pathway of WTAs is quite complex, yet the last step in this pathway occurs outside the cell and is accomplished by the LytR-CpsA-Psr - LCP - family of proteins, making this step a very interesting one to target.
The first structure of LytR from *Streptococcus dysgalactiae subsp. dysgalactiae* (SDSD) was solved at 2.80 Å resolution. Analysis of the structure revealed a high structural homology with other LCP proteins deposited at the PDB, as well as a long narrow hydrophobic pocket also present in the other homologous structure. Also, it allowed to understand that the conserved molecular determinants important for activity are also present in the determined structure. Characterization of the protein in solution was also performed using SAXS technique. Preliminary results suggest that the protein might adopt a different conformation in solution when compared to the crystal form. Further characterization of the protein is necessary to understand if the protein really has two distinct conformations, and if the presence of a ligand can induce this conformational change in solution.
Further characterization might help improve what is known about the LCP family of proteins, namely how the WTAs-lipid linked precursors enter the hydrophobic pocket and the mechanism of transfer the WTAs to the peptidoglycan. Hopefully this increase in knowledge will aid in the development of inhibitors for biofilm formation.

Study of Lipid Droplets Accumulation in Hepatocarcinoma Cell using Raman Micro-spectroscopy

Pradjna N. Paramitha[1], Anisa Maryani[1], Riki Zakaria[1], Yukako Kusaka[1], Kosuke Hashimoto[1], Bibin B. Andriana[1], Hidetoshi Sato[1]

[1]Kwansei Gakuin University, Sanda, Hyogo 669-1337, Japan

Abstract
Non-alcoholic fatty liver disease (NAFLD) is a pathological condition where an accumulation of excess fat occurs in the liver of people who drink little or no alcohol. About 80-90% of people who have NAFLD do not develop progressive liver disease and therefore have a relatively good prognosis. However, the others develop progressive liver disease; non-alcoholic steatohepatitis (NASH), fibrosis, cirrhosis, and eventually hepatocellular carcinoma (HCC). The factors that affect the progression of NAFLD to HCC are still unclear. This research aims to investigate the role of lipid accumulation that is the major symptom in NAFLD as well as in the carcinogenesis of hepatocellular carcinoma. Raman measurements are carried out in live-cultured cells by using confocal Raman micro-spectroscopy with an excitation wavelength of 785 nm. We employed HepG2 cells, a human liver cancer cell line. A steatosis model of HepG2 cells was developed by exposing cells to a medium containing a high concentration of glucose and different types of fatty acid: palmitic acid, stearic acid, oleic acid, and linoleic acid. High glucose and high fatty acid concentration in the culture medium induced accumulation of lipid droplets in HepG2 cells with varying chemical composition, demonstrated by Raman spectrum of the major lipid droplets present in each treatment group. The major variance was observed at 1260 and 1650 cm-1 which attribute alkyl -C—H cis stretches and C=C bond, respectively. This result shows that the abundance of glucose and fatty acid in the culture medium might induce alteration in the chemical composition of lipid droplets in HepG2 cells. The fatty acid is up-taken by the cells and gives effects to its chemical composition in the lipid droplet, while glucose induces fat synthesis inside the cells.

Identifying Variations in Drug Target Binding Sites that Potentially Contribute to Anti-Fungal Resistance in Candida auris

Jocelyn Petitto[1, 2], Asmaa Elkabti[3], Alicia Howell-Munson[1], Dmitry Korkin[1,4,5]

[1]Bioinformatics and Computational Biology, Worcester Polytechnic Institute, Worcester, MA 01609, [2]RNA Therapeutics Institute, University of Massachusetts Medical School, Worcester, MA 01605, [3]Biology and Biotechnology, Worcester Polytechnic Institute, Worcester, MA 01609, [4]Data Science Program, Worcester Polytechnic Institute, Worcester, MA 01609, [5]Computer Science Department, Worcester Polytechnic Institute, Worcester, MA 01609

Candida auris has been a growing concern in the hospital setting, particularly in long-term health care facilities, which has been exacerbated by the COVID-19 pandemic, with outbreaks occurring in COVID-19 dedicated units. *C.auris* is a multi-tiered threat: it is resistant to anti-fungal drugs and is easily misidentified using standard techniques for clinical isolates as other yeast strains such as *Candida haemulonii*.
We hypothesized that common anti-fungal drugs are ineffective against *C. auris* as a result of variation in the drug-target protein structures as compared to related drug-sensitive fungi. To test this hypothesis, we first considered sequence conservation and then utilized homology modeling to predict the structures of *C. auris* proteins targeted by anti-fungal drugs. Using the predicted model, we used computational methods to determine the binding ability of drugs to their target site. Finally, we compared the resulting model to those of related species as well as the *Homo sapiens* homolog.
Based on the availability of protein templates, we primarily focused our research on cytochrome p450, the target of azole class of anti-fungal drugs. We found that there was a high level of sequence conservation at the docking site between species. However, when comparing the *H. sapiens* and fungal structures, a shift in the heme group – a co-factor for the enzymatic activity of cytochrome p450 – was observed. Additionally, when the pocket forming the active site was visualized and inspected, differences in the size of the entrance of the pocket were observed. This suggests access to the heme group and/or steric hindrance of the target binding site may play a part in *C. auris*’ resistance to azoles.

Developing a dataset of tertiary structural motifs for metalloprotein design

Fabio Pirro[1], Thomas Lemmin[2, 3], Marco Chino[1]

[1]Department of Chemical Sciences, University of Napoli Federico II, 80126 Napoli, Italy, [2]Department of Computer Sciences, ETH Zurich, 8092 Zürich, Switzerland, [3]Institute of Medical Virology, University of Zurich, 8057 Zürich, Switzerland

The intricate relationship between metal ions and protein scaffolds plays a major role in most processes that are fundamental to Life. Metal cofactors are often essential for the proper folding of polypeptide chains or function of folded proteins. The birth and exponential growth of the PDB over the past 50 years has provided the scientific community with an invaluable resource for the study of such relationships, allowing for example to unveil the coordination number and binding propensity of coordinating residues. It has also been exploited by Machine Learning studies aimed at ab initio identifying and differentiating structural and functional metal sites. However, the full potential of protein ligands to bind any given metal cofactors and tune their chemistry has not been exhaustively explored by evolution. To this aim, the de novo design of metalloproteins by leveraging the wealth of knowledge contained in the PDB represents a promising strategy for the discovery of new catalysts.It has been shown that the protein structural space is highly degenerate and can be recapitulated with a finite set of structural units, such as elements of tertiary structure called TERMs (**TER**tiary **M**otifs).[1]

Here, we build on this concept and report the development of a dataset of highly specialized metal-binding elements: MetalTERMs. Over 100’000 MetalTERMs were identified from sites in which the metal was bound only by protein residues and water molecules. Subsequently, MetalTERMs were clustered according to their root-mean-square-deviation, the total number of residues and the number of non-contiguous segments. We find that the number of clusters rapidly drops with the increase in complexity of the tertiary arrangement, and that MetalTERMs composed of at most three segments can recapitulate about 90% of the whole dataset. This would indicate that medium to long-range mutations have most likely only a marginal effect on the metal coordination sphere, and would therefore corroborate the well-established adoption of the miniaturization approach for designing metalloproteins from scratch. Our analyses have also enabled us to identify new-to-nature combinations of most recurring MetalTERMs, which could finally lead to the design of unprecedented catalysts.

[1]Mackenzie et al, PNAS, 2016, 113 (47), E7438-E7447

Why we need high-resolution structures: hidden features revealed.

Marina Plaza-Garrido[1], Mª Carmen Salinas-Garcia[1], Ana Camara-Artigas[1]

[1]University of Almeria

To date, thanks to advances in synchrotron facilities, a growing number of protein structures deposited in the PDB are determined at high-resolution (HR). Before 2020 only the 5% of the structures were solved at a resolution higher than 1.5 Å. Last year, the 16% of the structures were obtained at atomic-resolution. Up to the next year, most of these structures where solved by X-ray diffraction (XRD). During the last year, cryo-EM was able to begin to compete with XRD in this field. HR data are essential to monitor characteristics in proteins that, in some cases, could have physiological significance. Our experience with atomic resolution structures showed us the importance of obtaining good quality data. It allowed us to identify modified residues, such as a succinimide in the PDZ3-PSD95 or isoaspartate in lysozyme. Up to date these modifications have been poorly identified and characterized. Besides, better data allows improved modelling of ligands showing features not displayed by lower resolution structures. For example, the dye bromophenol blue bound to lysozyme showed the dye's resonant form. Another improvement in the protein structures, reached with HR data, is the modelling of alternated conformations. The complexes of proline-rich motifs (PRMs) with SH3 domains from the Src family of kinases show different conformation of the highly conserved residue Leu, which seems to work as a switch between conformations that favours the binding of PRMs in different orientations. Finally, high-resolution structures allow better modelling of the solvent. Some water molecules have been proposed to play a key role in PRMs binding to modular domains. HR structures have allowed us to model sodium ions in the SH3 domain structures from the c-Src and Fyn kinases. Some of these ions are located at the same position that some water molecules proposed to play a critical role in the binding's energetics. Better modelling and characterization of the solvent might be critical in the development of new drugs in silico. HR structures, besides to facilitate the modelling of the structures, reveal hidden features in the proteins with unexpected roles.

Hypervariability of accessible and inaccessible conformational space of proteins

Ashraya Ravikumar[1], Narayanaswamy Srinivasan[1]

[1]Molecular Biophysics Unit, Indian Institute of Science, Bengaluru, India

Proteins undergo motions in a range of amplitudes, from large-scale domain motions to backbone rotational movements, leading to changes in (?,ψ) torsion angles and small-scale bond vibrations and bond angle bending. In this work, we study the extent of variations in (?,ψ) values in a given protein and the effects of minor variations in bond geometry due to vibrational motions in a protein on the accessible, steric clash-free (?,ψ) space. We perform 1-femtosecond timestep unconstrained molecular dynamics (MD) simulations on a super-high-resolution Lysozyme structure. The extent of variations observed in bond geometry during the simulation is within acceptable ranges of bond lengths and angles. However, using bond geometry-specific (?,ψ) steric maps we show that the steric clash-free (?,ψ) space is continuously changing, even for a given residue over the simulation time. (?,ψ) regions that are disallowed (or) involved in steric clash at one time point can become steric clash-free at a future time point through minor adjustments to the backbone bond lengths and angles. We also highlight an example where a residue conformation undergoes (?,ψ) transition from the left to the right half of the (?,ψ) map in consecutive snapshots of the trajectory. Although the two quadrants are separated by a steric clash-prone region which corresponds to a high energy barrier, the height of this barrier is brought down by adjusting the bond geometry such that a bridging region of steric clash-free, low energy (?,ψ) values is formed. This provides a likely route for the (?,ψ) transition. Through this work, we demonstrate the idea of dynamically varying nature of acceptable and accessible (?,ψ) steric space in a protein, which has implications for protein folding; proteins could sample (?,ψ) space which is originally considered to be inaccessible, during folding, through minor adjustments to their backbone bond geometry.

Understanding the structural basis for the N- demethylation of caffeine by NdmA

Pedro Rivera-Pomales[1], Amie K. Boal[1,2]

1 Department of Chemistry, The Pennsylvania State University, University Park, PA 16802, 2 Department of Biohemistry and Molecular Biology, The Pennsylvania State University, University Park

Rieske non-heme iron-dependent oxygenases (ROs) catalyze a variety of biologically relevant reactions, including the generation of cis-dihydroxylated aromatic compounds or O-/N-demethylated versions of aromatic precursors. NdmA is an N-demethylase involved in caffeine degradation. The reaction involves hydroxylation of the methyl group attached to N1 of caffeine to produce theobromine. The mechanism of demethylation is not well-understood. To investigate the reaction pathway of NdmA, we have solved structures of six different states in the catalytic cycle, including reactant, product, and intermediate states. Additionally, we have applied a vanadyl probe to mimic transient states in NdmA. This approach has been used previously in one other iron enzyme class, but here we show its utility in RO enzymes. Comparative analysis of NdmA and other aliphatic RO hydroxylases reveals a conserved location for the substrate binding pocket relative to aromatic dioxygenases in this superfamily. These differences may rationalize distinctions in reaction mechanism among ROs with divergent reaction outcomes.

Beyond acid catalysis: The importance of Glutamate 43 for adenine removal by MutY

Laura P. Russelburg[1], Merve Demir[2], Karina Cedeno[1], Sheila S. David[2], Martin P. Horvath[1]

[1]School of Biological Sciences, University of Utah, Salt Lake City, UT, [2]Department of Chemistry, University of California Davis, Davis, CA

In our efforts to outline the molecular steps encountered during base excision catalyzed by the DNA repair enzyme MutY, we obtained several crystal structures of *Geobacillus stearothermophilus* MutY E43 variants in complex with DNA. MutY removes adenine from 8-oxoguanine:adenine mispairs, thereby initiating repair and preventing G:C to T:A transversion mutations. We altered an active site residue critical for acid-base catalysis (E43); E43 acts as a general acid to protonate the adenine base and promote N-glycosidic bond cleavage. A previous structure of MutY in complex with a substrate analog shows E43 close to N7 of the adenine base (PDB ID 3G0Q) consistent with a role in acid / base catalysis. Our goal here was to illuminate the role of E43 in the mechanism of MutY through a combination of amino acid replacement and structure determination. We changed the glutamate residue to a conservatively shaped glutamine residue and a smaller serine residue. Substitutions at E43 do not entirely abolish catalytic power, as evidenced by lack of substrate base in some structures and corroborated by biochemical assays. Crystals diffracted anisotropically with strongest data along the b* axis ranging from 1.8 - 2.5Å. We obtained three categories of structures. First, structures obtained for E43S complexed with the cognate adenine base display the anti-conformation of the substrate adenine within the active site; previously reported substrate complexes feature the syn conformation. The anti-conformation appears unproductive as it disengages the base from close contact with active site residues. Replacement of E43 thus reveals that glutamate positions the substrate in the correct orientation for cleavage. Second, multiple structures obtained with purine in the active site display a reduced occupancy of the acid labile analog with both S and Q substitutions of E43, indicating a mixture of substrates and intermediates. Last, in some E43S structures obtained with the purine analog, we captured the MutY catalyzed product, an abasic site in the ring closed form. In conclusion, we have obtained structures across multiple molecular steps that suggest a key role for E43 in MutY's commitment to catalysis.

Atomic-Resolution Structures of Protein Assemblies by Integrating Magic-Angle-Spinning NMR Distance Restraints and Low-to-Medium Resolution cryo-EM Density Maps

Ryan W. Russell[1, 3], Chunting Zhang[1], Changmaio Guo[1], Manman Lu[2, 3], Juan R. Perilla[1, 3], Angela M. Gronenborn[2, 3], Tatyana Polenova[1, 3]

[1]Department of Chemistry and Biochemistry, University of Delaware, [2]Department of Structural Biology, University of Pittsburgh School of Medicine, [3]Pittsburgh Center for HIV Protein Interactions, University of Pittsburgh School of Medicine

Atomic-resolution protein structures determined by nuclear magnetic resonance (NMR) spectroscopy in solution and in the solid state rely on experimental distance restraints obtained through dipolar correlation experiments and dihedral restraints derived from NMR chemical shifts. While in smaller proteins NMR experimental restraints alone are sufficient for deriving the structures, for systems of increasing complexity/size as well as those possessing multiple domains, integration of several techniques is necessary. One example of such integration is combining experimental restraints from magic angle spinning (MAS) NMR experiments with medium- and low-resolution electron density maps obtained through cryoelectron microscopy (cryo-EM). We have undertaken this approach to devise a general protocol for atomic-resolution structure determination of large protein assemblies. We discuss the technical requirements of our integrated approach as well as its application to derive atomic-resolution structures of HIV-1 capsid assemblies and kinesin assemblies with polymerized microtubules.

Would Alpha Fold Predict Protein Misfolding?

Mari Carmen Salinas-Garcia[1], Marina Plaza-Garrido[1], Ana Camara-Artigas[1]

[1]Department of Chemistry and Physics, University of Almería, ceiA3.

Last year, we witnessed one of the most revolutionary achievements in solving the problem of protein folding. More than fifty years ago, Anfinsen experimentally demonstrated that the folding of the protein was in some way encoded in its sequence. Since then, a different approach has tried to predict in silico the final fold of the protein. Last year, Alpha Fold 2 won the first position in the last CASP14 challenge, reaching a score higher than 90 GDT in predicting the 3D fold of proteins. Finally, artificial intelligence (AI) seems to have reached the same efficiency as experimental techniques as NMR and X-ray crystallography. However, after solving the protein folding problem, would AI predict protein misfolding or other alternative foldings?
3D domain-swapping (3D-DS) is a process that occurs when two or more molecules of a protein exchange their secondary structure elements or protein domains, forming homodimers or higher-order oligomers. The increasing number of 3D-DS structures deposited at the PDB shows that this alternative folding is not just as crystallization artefact. 3D-DS phenomena can occur in various proteins of diverse function, sequence, size, and fold. Even though its biological role is not clear, this process has been related to another misfolding process, amyloid formation, which is the leading cause of neurodegenerative diseases. In this way, studying 3D-DS can help to characterize the initial steps of protein misfolding structurally. Some proteins have been isolated only in their domain-swapped form and have been suggested to perform some functional role. The PDZ2 of the protein Zonula Occludens (ZO) forms homodimers in ZO1 and ZO2 and heterodimers in ZO1-ZO2 and ZO1-ZO3. Here we show our results in 3D-DS from proteins that might have functional relevance and model proteins to characterize this phenomenon. Protein crystallography is the dedicated tool to characterize this kind of structures. Structural data of the monomeric an oligomeric form of these proteins can help to unveil the mystery of why some proteins, instead of folding as expected, are conducted to form oligomers through the exchange of their secondary structure elements.

Nucleolin-miRNA interactions are driven by specific elements on RNA binding domains of nucleolin: An sillico modeling and RNA-protein docking study

Avdar San[1, 2], Anjana Saxena[1, 2], Shaneen Singh[1, 2]

[1]City University of New York, The Graduate Center, [2]City University of New York, Brooklyn College

Nucleolin (NCL) is a stress responsive multifunctional RNA binding protein (RBP) that plays important roles in gene transcription and RNA metabolism. NCL levels and similarly many of NCL RNA targets are dysregulated in multiple human cancers. NCL- RNA interactions are driven by its four RNA binding domains (RBDs). Despite the myriad NCL functions that involve its RNA-binding properties, the mechanisms driving these interactions are poorly understood. Partial 3D structural information available on NCL RBDs in the RCSB PDB database, is not sufficient to provide a comprehensive understanding of NCL- RNA interactions. To investigate RNA binding specificity for its diverse targets, we have built upon the existing NCL-RBDs structures to generate multiple combinations of RBDs in tandem, using in silico modeling approaches. In this study, we analyze NCL-miRNA interactions in depth with a focus on a subset of miRNAs that are implicated in breast cancer. Using both template-based and ab initio approaches, we have generated complete and robust structural models of all 4 NCL RBDs in tandem and in combinations. All models were then tested using a variety of structural quality evaluation programs. Models with top validation scores were used in RNA-protein docking algorithms and assessed for interaction with specific miRNA. Our docking analyses have generated a comprehensive map of the miRNA-NCL protein interface sites for each individual miRNA and consistently predict specific RBDs in NCL-miRNA binding. We have also identified critical residues on the NCL-RBDs that may drive the miRNA binding in an RBD-type specific manner. Our results corroborate previous studies on RBDs from other RBPs. Structural information derived from this study provides a valuable perspective for future experiments. Identification of RNA binding motifs on NCL RBDs is critical in elucidating RNA target specificity by NCL. Computational analyses provide time and cost-effective benefits over experimental techniques; the in silico predictions are important in designing rational experiments for the future investigations confirming the NCL-miRNA interactions. Our study provides the foundational steps for establishing consensus motifs on NCL-RBDs that potentially direct RNA target specificity. Targeting these specific binding sites will provide new approaches to regulate NCL functions in gene expression during tumorigenesis.

ConPlot: Web-based application for the visualisation of protein contact maps integrated with other data

Filomeno Sanchez Rodriguez[1, 2], Shahram Mesdaghi[1], Adam J. Simpkin[1], Javier Burgos-Mármol[1], David L. Murphy[1], Ville Uski[3], Ronan M. Keegan[3], Daniel J. Rigden[1]

[1]Institute of Structural, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK, [2]Life Science, Diamond Light Source, Harwell Science and Innovation Campus, Oxfordshire, UK, [3]UKRI-STFC, Rutherford Appleton Laboratory, Research Complex at Harwell, Didcot OX11 0FA, UK

Recent developments in the field of evolutionary covariance and machine learning have enabled the precise prediction of residue-residue contacts and increasingly accurate inter-residue distance predictions. As a result, these data types have gained popularity in the field of protein bioinformatics, and are usually represented as two-dimensional binary matrices called contact maps and distograms. These typically omit contacts between sequential near neighbours resulting in a blank space on and near the diagonal axis of the matrix. A multitude of properties can be predicted by other sequence-based methods and researchers often need to consider diverse sources of information in order to form a complete and integrated picture for the inference of structural features. Here we present ConPlot, a web-based application which uses the typically empty space near the contact map or distogram diagonal to display multiple coloured tracks representing other sequence-based predictions. These predictions can be uploaded in various popular file formats. This integration enables researchers to easily analyse a variety of data simultaneously and facilitates discovery of structural features. This web application is currently available online at www.conplot.org, along with documentation and examples.

Backbone-based descriptors allow to identify metal-binding sites in proteins.

José-Emilio Sánchez-Aparicio[1], Laura Tiessler-Sala[1], Lorea Velasco-Carneros[2], Lorena Roldán-Martín[1], Giuseppe Sciortino[1, 3], Jean-Didier Maréchal[1]

[1]Universitat Autònoma de Barcelona, [2]University of the Basque Country, [3]Institute of Chemical Research of Catalonia

Metal ions play a crucial role in many biological processes, such as oxygen transport or stabilization of protein structures. This is reflected in the Protein Data Bank, where around 30% of the deposited structures contain a metal ion. Understanding protein-metal interactions have therefore direct biomedical (metal related diseases, metallodrugs) and biotechnological (*de novo* metalloenzymes, biocatalysis, biocompatible material) impacts. In addition to experimental procedures, *in silico* methods are interesting to address the specific challenge of predicting where a metal ion binds to a protein.
To date, the vast majority of computational predictors are either sequence-based or focus on the first coordination sphere of the metal. However, other structural aspects are relevant in predicting where a metal can bind to a protein. One of those is related to the protein pre-organization for accommodating the metal. Here we present BioMetAll, a novel family of metal-binding site predictors where pre-organization is defined by the backbone geometry.
The software was developed first by analyzing the entire set of structures of the MetalPDB, focusing on geometric descriptors between the metal and the backbone atoms of the coordinating amino acids and second, by parameterizing these descriptors for each natural amino acid. The result is an open-source piece of software called BioMetAll, (https://github.com/insilichem/biometall). BioMetAll takes as inputs the protein structure in pdb format and a set of user-defined parameters. The output is a list of the areas in the protein that are considered feasible for metal binding. We report a benchmark of BioMetAll on a set of more than 90 metal-binding X-ray structures and its application on three challenging cases: i) the modulation of metal-binding sites during conformational transition in human serum albumin, ii) the identification of possible routes for metal migration in hemocyanins, and iii) the prediction of mutations to generate convenient metal-binding sites for *de novo* biocatalysts. This study shows that BioMetAll offers a versatile platform for numerous fields of research at the interface between inorganic chemistry and biology and allows to highlight the role of the pre-organization of the protein backbone as a marker for metal binding.

Investigating the effect of TGF-β1 and aberrant microRNAs in airway epithelial cell lines

Maria J. Santiago[1], Srinivasan Chinnapaiyan[1], Rajib K. Dutta[2], Hoshang Unwalla[1]

[1]Florida International University, [2]Miller School of Medicine, University of Miami

HIV-1 Tat and cigarette's smoke increase the expression of TGF- β1 in airway epithelial cells. Increased TGF-β1 signaling has a multifactorial effect on airway health leading to chronic inflammation observed in Chronic Bronchitis and COPD. Specifically, we are investigating the effects of TGF-β1 signaling on the airway microRNAome that leads to dysregulation of genes regulating mitophagy. This inhibits mitophagy leading to accumulation of damaged mitochondria which in turn leads to senescence. The senescence associated secretory phenotype leads to secretion of proinflammatory cytokines leading to neutrophil infiltration and chronic inflammation. To probe the role of TGF-β1 and identify aberrant microRNAs, we are treating airway epithelial cells with 10nM of TGF-β1 for a period of 48 hours followed by a Western Blot to check the expression of *SIRT1* that plays an important role in macroautophagy and mitophagy. We will also determine the levels of miR-145-5p and miR-449b which are upregulated by TGF-β1 and known to suppress *BNIP3* and *SIRT1*. We anticipate that microRNA mimics will reproduce our observations with TGF-β1 validating their role in impaired mitophagy and increased senescence. This project will allow us to better understand the mechanism by which TGF-β1 and aberrant microRNAs impair mitophagy and promote senescence with consequent inflammation with a long-term goal towards designing therapeutics for treatment of COPD.

Approaches to Structure Determination in Heterogeneous Amyloid Samples - Applying Computational Methods to Improve Structure Modelling in Medium Resolution Cryo-Electron Microscopy Maps

Luisa U. Schäfer[1, 2], Christine Röder[1, 2, 3], Tatsiana Kupreichyk[1, 2, 3], Lothar Gremer[1, 2, 3], Karunakar R. Pothula[1, 2], Raimond B. G. Ravelli[4], Dieter Willbold[1, 2, 3], Wolfgang Hoyer[1, 2, 3], Gunnar F. Schröder[1, 2, 5]

[1]IBI-7: Institute of Biological Information Processing, Research Center Jülich, 52425 Jülich, Germany, [2]JuStruct, Centre for Structural Biology, Research Center Jülich, 52425 Jülich, Germany, [3]Institute of Physical Biology, Heinrich-Heine Universität Düsseldorf, 40225 Düsseldorf, Germany, [4]Institute of Nanoscopy, Maastricht University, 6211 LK Maastricht, Netherlands, [5]Physics Department, Heinrich-Heine Universität Düsseldorf, 40225 Düsseldorf, Germany

Heterogeneity of cryo-electron microscopy (cryo-EM) data sets often poses challenges regarding reconstruction and model building. Here, we applied cryo-EM to elucidate the structure of human IAPP (islet amyloid polypeptide) fibrils at physiological pH and reconstructed densities of three dominant polymorphs. Furthermore, we used different computational methods to verify and optimize our results.

The atomic model of the most abundant polymorph, which reveals two S-shaped, intertwined protofilaments (PDB: 6Y1A), was built in a density map of 4.2 Å. Its stability was verified by MD-Simulations.

In contrast, the density map details of a second polymorph at 4.2 Å did not allow for atomic model building. Here we used our software DireX to automatically model putative sequences into the density and thus facilitate the interpretation of the density map.

RNA Improves its Posture: riboswitch stacking spines are nucleobase networks that communicate metabolite-sensing to distal domains that support gene regulation

Griffin M. Schroeder[1, 2], Joseph E. Wedekind[1, 2]

[1]Center for RNA Biology, University of Rochester School of Medicine & Dentistry, Rochester NY 14622, [2]Dept. of Biochemistry and Biophysics, University of Rochester School of Medicine & Dentistry

Riboswitches are naturally occurring RNA molecules that directly “sense” cellular levels of specific metabolites to control downstream genes. Although numerous riboswitch-effector complexes are known, a central challenge is to understand how molecular recognition of a specific ligand leads to conformational pathways that control transcription or translation. Recently, we determined crystal structures of the *Thermoanaerobacter tengcongensis* (*Tte*) riboswitch in the presence and absence of its cognate ligand $preQ_1$ [Schroeder et al. (2020) *NAR* **48**, 8146]. A structural comparison revealed major changes in the effector-binding pocket, where different conformations are observed for the L2 loop. The finding that L2 is well-ordered in both unbound- and bound-states allowed us to analyze low-energy pathways describing the interconversion of these conformations using an innovative Nudged Elastic Band (NEB) computational approach. NEB pathways reveal how A14 moves out of the binding pocket as C15 enters the pocket to recognize $preQ_1$. This change is accompanied by formation of a nucleobase “stacking spine” that spans the length of the riboswitch fold, bridging the ligand-binding domain and the distant expression platform, which contains the essential gene-regulatory Shine-Dalgarno sequence. We corroborated stacking-spine formation by substituting 2-aminopurine at various positions in the stacking-spine network, which quenched fluorescence with added $preQ_1$. Additional support for specific stacking-spine pathways was derived from chemical modification analysis coupled with next-generation sequencing. We then sought to identify comparable stacking spines in other riboswitches with an emphasis on those with high-resolution unbound- and bound-state structures. We identified analogous spines in the tetrahydrofolate (THF) and *add* adenine riboswitches, which form only in the presence of ligand. These spines are compared to a stacking spine located in the ribosome-sensing network of the viral tRNA-like-structure (TLS) [Hartwick et al. (2018) *Nat. Comm*. **9**, 5074]. Overall, this work shows how functional RNAs can sense effectors that induce long-range chemical networks to control gene regulation.

Catechol binding to the active site Fe(III) of the inactive chimera of catechol and protocatechuate dioxygenases suggests a role for the N-terminal domain in substrate specificity and selectivity.

Patrick Semana[1], Justin Powlowski[1, 2]

[1]Department of Chemistry and Biochemistry, Concordia University, Montreal, QC, Canada,
[2]Centre for Structural and Functional Genomics, Concordia University, Montreal, QC, Canada

Intradiol ring-cleavage dioxygenases are enzymes able to cleave a range of vicinal dihydroxylated aromatic substrates albeit with differing catalytic efficiencies. Although all members of the intradiol ring-cleavage family belong to the same evolutionary lineage, their substrate specificity appears to be reflected in their phylogenetic relationships. However, studies focused on identifying the molecular determinants of substrate specificity remain scarce to date. We have recently heterologously expressed, purified and characterized a catechol-specific (Anig_179) specific and a protocatechuate-specific (Anig_151) intradiol ring-cleavage dioxygenase from *Aspegillus niger*. Although, the amino acid sequences of these two enzymes are 54 % identical (over 88% of the sequence), they have no catalytic activity on similar substrates. Since the catalytic residues are conserved and homology-based models of these two enzymes suggest that they have similar structural folds, we hypothesized that the substrate specificity might be determined by residues from the dimerization domain interfacing the active site.

To examine this hypothesis, we have generated a chimeric variant (Anig_151Nt_179Ct) consisting of the C-terminal (catalytic) domain from Anig_179 fused with an N-terminal (dimerization) domain from Anig_151. The purified chimeric protein: has a UV-Vis absorbance spectrum reminiscent of intradiol ring-cleavage dioxygenases, but with the absorbance maximum in the visible-region shifted to the red by about 50 nm; is similarly homodimeric in solution; and also has secondary structure content similar to Anig_151 and Anig_179. In addition, we found that although the Anig_151Nt_179Ct variant has no catalytic activity towards either catechol or protocatechuate, it is able to bind the native substrate of Anig_179, catechol. Furthermore, titration experiments with catechol show saturation behavior with the derived-data best-described by a two-site binding model, with apparent K_D values 31.7 and 1720 µM. We propose that the lower K_D reflects catechol binding to the active site of the chimeric protein, while the higher K_D-value likely represents the affinity of Anig_151Nt_179Ct for the Fe(III)-catecholate complex. This is supported by the observation that catechol acts as a chelator for the active site-Fe(III) at millimolar concentrations. Furthermore, homology modeling is being used to identify residues likely to contribute to substrate selectivity and specificity, which will be verified by site-directed mutagenesis.

Molecular Dissection of Sulfoglycolytic degradation pathway of gut microbiota metabolism

Mahima Sharma[1], Yi Jin[2], Ethan D. Goddard-Borger[3], Spencer J. Williams[4], Gideon J. Davies[1]

[1]York Structural Biology laboratory, Department of Chemistry, University of York, [2]School of Chemistry, Main Building, Cardiff University, Park Place, Cardiff CF10 3AT, [3]CRF Chemical Biology Division, The Walter and Eliza Hall Institute of Medical Research, Parkville, V, [4]School of Chemistry and Bio21 Molecular Science and Biotechnology Institute, University of Melbourne

Sulfoquinovose (SQ), a sulfonated glucose analog liberated from plant sulfolipids (SQDG and its metabolites), is a major component of global sulfur cycle with its estimated production amounting to 10 billion tonnes annually [1]. Different bacterial sulfoglycolytic pathways analogous to fundamental glycolytic Embden-Meyerhof-Parnas (EMP) and Entner-Doudroff (ED) pathways have been reported [2,3], however, structural and biochemical characterization of core enzymes involved in the degradation of sulfoquinovose is lacking. Here we present, crystal structures of these enzymes, in complex with their proposed intermediates, as well as the kinetic studies to shed light on their mechanisms, the determinants of sulfo-sugar specificity and their selectivity over glycolysis intermediates. The sulfonate recognition sequences thus identified would further inform our search for sulfoglycolytic enzymes in different organisms in varied environmental niches, and possibly even heretofore unreported sulfoglycolysis pathways.

Discerning Ferredoxin Modules using Statistical Coupling Analysis (SCA)

Jan A. Siess[1], Vikas Nanda[2]

[1]Institute for Quantitative Biomedicine, Rutgers University-New Brunswick, [2]Center for Advanced Biotechnology and Medicine, Rutgers University-New Brunswick

Modern, extant proteins are typically composed of multiple, independently folding sections of polypeptides or domains. This agglomeration of domains imparts many structural and functional aspects and constitutes the main components of globular proteins. Domains themselves have gone through a multitude of recombination and duplication events which not only affect how proteins fold, but also the way in which they function and interact with other structural elements globally. While domains are considered as units of evolution, a definition for what comprises one is often variable. Take ferredoxin as an example. Ferredoxins are iron-sulfur proteins that are ubiquitous amongst a wide range of organisms. They are responsible for mediating electron-transfer in a host of metabolic reactions, and we have seen them present within larger complexes. However, isolating the ferredoxin by itself looks like its own domain, representing an assembly of two asymmetric units consisting of two halves. Thus, using ferredoxin as a test case, we seek to identify the fundamental folds that comprise ferredoxins. One approach includes the use of statistical coupling analysis, which identifies coevolving amino acids amongst an ensemble of aligned, homologous sequences. By distilling these proteins into their core functional components, we can identify key residue interactions contributing to these fundamental folds' formation. With ferredoxin as the base, we ultimately seek to extend this to larger, more complex ferredoxin-containing biomolecules.

Structural genomics and interactomics of SARS-CoV-2

Suhas Srinivasan[1], Oleksandr Narykov[2], Hongzhu Cui[3], Ziyang Gao[3], Ming Liu[3], Senbao Lu[3], Winnie Mkandawire[3], Mo Sun[3], Dmitry Korkin[1, 2, 3]

[1]Data Science Program, Worcester Polytechnic Institute, Worcester, MA 01609, [2]Computer Science Department, Worcester Polytechnic Institute, Worcester, MA 01609, [3]Bioinformatics and Computational Biology Prog, Worcester Polytechnic Institute, Worcester, MA 01609

In the 15 months since a pandemic novel virus was discovered, nearly 123 million people were infected and 2.7 million lives lost. The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a novel beta coronavirus suspected to originate from a bat coronavirus, is the cause of the coronavirus disease 2019 (COVID-19) pandemic. The pathogenicity of SARS-CoV-2 arises from the function of its proteins, carried out by forming intraviral and virus-host protein complexes. The publication of the first SAR-CoV-2 genome ignited a race to structurally characterize the protein armory of the virus. With novel viral epidemics becoming more frequent, it is crucial to examine the capabilities of computational modeling for the structural characterization of viral proteins and their fundamental complexes.
Computational structural genomics, typically characterizes the protein architecture of a new virus by performing comparative or template-based modeling, through which a protein's 3D structure is derived from experimentally resolved structures of its evolutionary homologs. In our recent work, we developed one of the first available and accurate comparative models of the viral proteins and their complexes, complementing it with integrative models of the key structural proteins, spike (S) and membrane (M), as well as homo-oligomers they form. In addition, we analyzed all comparative and *de novo* models of SARS-COV-2 proteins and protein complexes that have been experimentally solved to determine: the protein sequence coverage of the models, the timeline of the models and experiments, and the models' accuracies were compared to the experimental structures.
Structural modeling of the proteins and fundamental complexes of a novel virus provides strategic information many crucial months ahead of what experimental techniques can achieve. Moreover, recent efforts to develop AI-based *de novo* modeling tools is expected to provide further advancement in the prediction accuracy of individual protein structures and, eventually, larger protein complexes.

Two riboswitches with identical binding pockets and disparate RNA folds show different tolerances to equivalent mutations

Yoshita Srivastava[1], Jermaine L. Jenkins[1], Joseph E. Wedekind[1]

[1]Department of Biochemistry & Biophysics and Center for RNA Biology, URMC, Rochester, NY 14642, USA

Riboswitches are naturally occurring structured RNAs that directly "sense" the cellular levels of specific metabolites to regulate downstream genes. Although some riboswitches are being targeted for antibacterial development, others are found in bacteria critical for human health. For example, the Ruminococaceae family of commensal gut bacteria produce the nucleobase queuine, which is crucial for genetic decoding by tRNAs containing GUN anticodons (i.e., those that specify Asp, Asn, Tyr, and His). Faecalibacterium prausnitzii (Fpr) is a member of this family that appears protective against Crohn's disease. The Fpr preQ1-III (class III) riboswitch senses the queuine precursor, preQ1, to control the translation of a downstream transporter required for preQ1 salvaging. We showed previously that the class III effector binding pocket shares ten identical nucleotides with phylogenetically distinct preQ1-II (class II) riboswitches [Liberman et al. (2015) PNAS 112, E3485], even though each riboswitch class adopts a distinctive overall fold. Moreover, class II riboswitches are found in many human pathogens. Given the goal of targeting RNA motifs with small molecule therapeutics, it is important to understand how commensal class III and pathogenic class II riboswitches respond to mutations in or around their binding pockets that could lead to drug resistance or loss of health benefits for the host. Accordingly, we prepared A84G, A52G, Δ84, U8C/A8G mutations in the class III binding pocket based on our previous analysis of equivalent class II mutations, which were detrimental to function [Dutta & Wedekind (2020) JBC 295, 2555]. Interestingly, ligand binding analysis by isothermal titration calorimetry revealed that each mutation is tolerated better by class III riboswitches. X-ray crystallographic analysis of each class III mutant accounts for its ability to maintain preQ1 binding and is corroborated by chemical probing of riboswitch flexibility in solution. A take-home message is that the context of the binding pocket within the overall fold impacts the ability to tolerate mutations.

In silico structural survey of newly identified late embryogenesis abundant proteins (LEAPs) from R. serbica and their structure - function relationship

Strahinja Stevanovic[1], Marija Vidovic[1]

[1]Institute of Molecular Genetics and Genetic Engineering, Laboratory for Plant Molecular Biology, UB

Desiccation or extreme water loss leads to protein denaturation, aggregation, and degradation and impairs membrane lipid fluidity, resulting in loss of membrane integrity at the cellular level. The induction of late embryogenesis abundant proteins (LEAPs) is considered an essential component of desiccation tolerance strategy in so-called resurrection plants. This heterogeneous group of hydrophilic, non-globular proteins is characterized by a high structural plasticity that allows them to adopt a random conformation in aqueous solutions that transforms into α-helices during dehydration [1]. Therefore, LEAPs can interact with various ligands and partners, including ion sequestration and stabilization of membranes and enzymes during freezing or drying [2].

Our new transcriptome database of an endemic resurrection species Ramonda serbica allowed us to identify 153 members of the LEA gene family. LEAPs of this sample data have an average primary sequence similarity and identity of 10% and 6%, respectively, but with a high variance (141 and 108), which means that the sample proteins can be classified based on domain homology. The averaging is based on multiple sequence alignment and the variance is estimated using pairwise sequence alignment scores. Accordingly, all identified LEAPs were clustered into six groups based on protein families (PFAM). Among these groups, LEAPs differ significantly in their secondary structure, disorder propensity and aggregation potential. Furthermore, we built homology models using Protein Data Bank structure information as templates. For each group, an ensemble of superimposed 3D homology models was analyzed.

The information obtained from the representative structural models is key to understanding the function of LEAPs and the regulation of their intrinsic structural disorder-to-order transition during desiccation. This will pave the way for the identification of LEAPs endogenous partners and their targets in the cell and provide further insights into the protective mechanisms of desiccation tolerance.

References:
1. Hundertmark M, Hincha DK (2008). BMC Genom 9:1.
2. Chakrabortee S, Tripathi R, Watson M, Schierle GSK, Kurniawan DP, Kaminski CF, ..., Tunnacliffe A (2012). Mol. BioSyst. 8:210-19.
Acknowledgement:
This_work_was_supported_by_the_Science_Fund_of_the_Republic_of_Serbia_(PROMIS_project LEAPSyn-SCI, grant_number_6039663)_and_by_the_Ministry_of_Education,_Science_and_Technological Development, Republic_of_Serbia_(contract_no._451-03-9/2021-14/ 200042).

Dead but not defunct: Pseudokinases maneuver flexibility signatures associated with protein kinase fold to reinforce non-catalytic functions.

Anindita Paul[1], Seemadri Subhadarshini[1], Narayanaswamy Srinivasan[1]

[1]Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, Karnataka, India

Pseudokinases lack one or more catalytic residues, conserved in canonical protein kinases that are required for nucleotide binding, divalent cation binding or phosphoryl group transfer (Manning *et al.*, 2002). Hence, they are predicted to be enzymatically dead. However, pseudokinases have been implicated in regulation of diverse signaling pathways, where they function as signal integrators, allosteric modulators, scaffolds, competitors or decoy. Despite their sequence dissimilarities with protein kinases, these pseudoenzymes retain the bilobal protein kinase-like fold. In the active conformation of protein kinases, the activation segment is extended, rendering the enzyme catalytically competent. However, tertiary structures of pseudokinases reveal that their topologically equivalent activation segments adopt a well-folded configuration, packing against the pseudoenzyme core. Comprehensive analyses of kinase-like sequences have revealed that pseudokinases have undergone notable expansions in plants, animals, fungi and bacteria (Kwon *et al.*, 2019, Paul & Srinivasan, 2020). In plants, pseudokinases related to receptor-like kinase (RLK) family have experienced massive expansion, accompanied by domain shuffling. Phylogenetically, RLK family shares a common ancestral origin with animal tyrosine kinase-like (TKL) and tyrosine kinase (TK) families, with respect to their kinase domains (Shiu & Bleecker, 2001). Herein, Anisotropic Network Model (ANM) based Normal Mode Analysis (NMA) was conducted on 51 active conformation structures of protein kinases and 26 structures of pseudokinases, related to plant RLK and human TKL and TK families. Our observations indicate that although the backbone fluctuation profiles are similar for individual kinase-pseudokinase families, low intensity mean square fluctuations in pseudo-activation segment and other sub-structures impart rigidity to pseudokinases. Analyses of collective motions from functional modes reveal that pseudokinases in comparison to active kinases, undergo distinct conformational transitions using the same structural fold. All-atom NMA of protein kinase-pseudokinase pairs, sharing high amino acid sequence identities, yielded distinct clusters of communities, partitioned by residues exhibiting highly correlated fluctuations. It appears that atomic fluctuations from equivalent activation segments guide the community network topologies for respective kinase and pseudokinase. Our findings indicate that such adaptations in the backbone and all-atom fluctuations render the pseudokinase competent for catalysis-independent roles.

Mining for functional ribosomal variants in Saccharomyces cerevisiae.

Daniel Sultanov[1], Andreas Hochwagen[1]

[1]New York University, New York, NY 10003-6688

ds5639@nyu.edu
The ribosome is a large macromolecular RNA-protein complex that performs translation. The ribosomal RNA (rRNA) is the key structural component and catalytic effector of this ensemble. Recent research in translational heterogeneity has sparked interest in characterizing different ribosome variants and their role in translation. However, such studies have primarily focused on the protein constituents of the ribosome. By contrast, data on functional rRNA variants in eukaryotes are sparse because rRNA genes are presumed to be highly homogeneous.
We searched for naturally occurring rRNA variants by analyzing available deep-sequencing data from 918 *S. cerevisiae* isolates that belong to different ecological niches [**Bioproject ID: PRJEB13017**]. We then mapped the variants onto high-resolution crystal structures of ribosomes in two different conformations [**PDB IDs: 4v88** and **6woo**]. Two factors were considered to estimate how rRNA variants affect cell fitness. These factors include the frequency of variants among all rRNA genes within each genome and their distribution in highly conserved eukaryotic rRNA regions. Therefore, we found 226 mutable positions in the rRNA that are most likely non-deleterious. This number accounts for 4% of the total rRNA in the ribosome. Around 40% of those positions also mediate interactions between rRNA and ribosomal proteins.
Structural analysis revealed the overall spatial distribution of the mutable positions, with depletion in the conservative catalytic center and enrichment in the outermost part of the ribosome. This result is in line with evolutionary predictions where nucleotide diversity in the ribosome increases from the core outwards.
We also detected variants in the vicinity of structurally and functionally important regions. Nucleotide changes were located near a site that is responsible for the translational fidelity. Variants were also localized to positions responsible for ribosome integrity and coordination. Moreover, further superposition of the ribosome structures revealed a dynamic rearrangement in some of those positions. Such variability may contribute to ribosome performance in vivo.
In summary, the findings pinpoint naturally occurring structural variants in a eukaryotic ribosome that might reflect their functional importance in vivo. The data provide information on mutable regions in rRNA and make them candidates for computational modeling and experimental validation.

Crystallization of C-Phycocyanin from *Spirulina platensis* using the counter-diffusion technique

AM Taranu[1, 2], J Pozo-Dengra[3], JA Gavira[2], S Martínez-Rodríguez[1, 2]

[1]Dept. of Biochemistry and Molecular Biology 3 and Immunology. University of Granada. Granada. Spain., [2]Laboratorio de Estudios Cristalográficos, IACT-CSIC. Granada., [3]Biorizon Biotech. Almería.

C-Phycocyanin (PC) is a blue, light-harvesting pigment in cyanobacteria, Rhodophyta and Cryptophyta [1]. The typical bluish color of many cyanobacteria is due to this water-soluble phycobiliprotein, from which these organisms are also known as blue-green algae [2]. Spirulina platensis PC (SplPC) has been widely used in cosmetics, food, and the biomedical field because of its unique advantages including high nutritional value, excellent biosafety, good water solubility, unique color, and strong luminescence [3]. Very recently, PC has been found to have specific affinity to tumor-associated macrophages via the scavenger receptor-A, and thus, it has been proposed as a tumor-targeted vehicle [4].
We have selected Spirulina platensis PC as a target system to design new materials/supports for protein crystallization into microfluidic devices. In this work we sumarize the purification and crystallization of the wild-type PC. We have standardized the crystallization of this phycobiliprotein using the counter-diffusion technique, being able to improve the resolution of the structure deposited in the PDB [5,6] with the first datasets obtained (below 2Å). Further improvement and selection of other potential polymorphs is ongoing in order to gain insights in differences observed in the phycobilin environment.

[1] Kuddus et al., 2013. Biomed Res Int. 2013:742859.
[2] Eriksen, 2008. Appl Microbiol Biotechnol. 80(1):1-14.
[3] Wan et al., 2020. Bioconjug Chem. 31(5):1438-1448.
[4] Wan et al., 2017. Chem Commun. 53: 4112– 4115.
[5] Wang et al., 2001. Acta Crystallogr D Biol Crystallogr. 57(Pt 6):784-792.
[6] Padyana et al., 2001. Biochem Biophys Res Commun. 282(4):893-898.

Study of heme binding mechanisms at molecular level

Laura Tiessler-Sala[1], Giuseppe Sciortino[1], José-Emilio Sánchez-Aparicio[1], Lur Alonso-Cotchico[1], Laura Masgrau[1], Agustí Lledós[1], Jean-Didier Maréchal[1]

[1]Departament de Química, Universitat Autònoma de Barcelona, Cerdanyola del Vallès, E-08193, Spain

A large number of biological systems obtain their activity by the inclusion of protoporphyrin IX into one or several binding pockets, prototypically heme binding proteins. Heme bound structures, and to less extend apo forms, can be found in the Protein Data Bank (PDB). Comparisons of apo/holo forms of different systems suggest that different patterns can exist in nature for heme binding recognition (from rigid receptor to major induced fit). To date, little has been done to decode the molecular mechanism of most of heme binding processes. Molecular modeling tools could be particularly interesting for such although the simulation of binding processes including metal and macrocycle are at the edge of today's methodologies.
In the present work, we settle a molecular modeling approach to investigate the dynamical behavior of the unbound form of several proteins and how this correlates with the binding of the heme. The computational framework involves classical molecular dynamics (MD), Gaussian accelerated MD (GaMD), updated molecular dockings tools able to account for metals and metal binding predictors we recently developed. The work focuses on two representatives of heme binding systems: heme oxygenase HO and hemophore HasA.
Our results reveal how heme oxygenase presents a conformational selection heme binding mechanism with a very similar conformational landscape between bounded and unbounded structure. On the contrary, hemophore HasA presents a mix mechanism with apo structures that swings between different states and the X-ray holo structure is only researchable after heme binding.
In conclusion, this study provides both an original computational framework to simulate substrate binding processes using the available apo and holo structures in PDB, but furthermore demonstrates the variability of molecular mechanisms that could be needed for the heme to bind to its receptors.

Structural basis for the interaction between plant-type HO1 and Fd revealed by Integrative/Hybrid method

Rei Tohda[1, 2], Hideaki Tanaka[1, 2], Xuhong Zhang[3], Tsuyoshi Konuma[4], Takahisa Ikegami[4], Catharina T. Migita[5], Genji Kurisu[1, 2]

[1]Institute for Protein Research, Osaka University, Suita, Osaka 565-0871, [2]Department of Macromolecular Science, Osaka University, Toyonaka, Osaka 560-0043, [3]Graduate School of Medical Science, Yamagata University, Yamagata, Yamagata 990-9585, [4]Graduate School of Medical Life Science, Yokohama City University, Tsurumi-ku, Yokohama 230-0045, [5]Department of Biological Chemistry, Yamaguchi University, Yoshida Yamaguchi 753-8515

Heme oxygenase (HO) converts heme to carbon monoxide, biliverdin, and free iron, products that are essential in cellular redox signaling and iron recycling. In higher plants, HO is also involved in the biosynthesis of photoreceptor pigment precursors. Although the enzymatic reactions are common in all HOs, the amino acid sequence identity of the plant-type HO is exceptionally low (~19.5 %) to all the other types of HOs, and amino acids that are catalytically important in mammalian HO are not conserved in plant-type HOs. Structural characterization of plant-type HO was limited, and it remains unclear how the structure of plant-type HO differs from that of other HOs. Then, we focused on the structure of *Glycine max* (soybean) HO-1 (GmHO-1) and the interaction with heme and Fd by the combination of X-ray crystallography, ITC and Nuclear Magnetic Resonance (NMR) spectroscopy [1]. Here, we have solved the crystal structure of GmHO-1 at 1.06 Å resolution, and carried out the NMR spectroscopic studies of its interaction with ferredoxin (Fd), the plant-specific electron donor. The high-resolution X-ray structure of GmHO-1 reveals several novel structural components; an additional irregularly structured region, the new water tunnel from the active site to the surface, and the new hydrogen-bonding network, unique to the plant-type HOs. We also revealed that the interaction sites of Fd toward GmHO-1 is close to the [2Fe-2S] cluster by NMR experiment. And the experiment-based Fd:GmHO-1 complex structure can be proposed by NMR data and HADDOCK simulation. We propose that while the catalytically essential elements, such as hydrogen-bonding network and His ligation, are structurally conserved, the plant-type specific structural feature; new water channel near the heme pocket, unique electrostatic potential, and surface structure may be evolutionally optimized for the plant-type specific physiological demands.

[1] Tohda, R., Tanaka, H., Mutoh, R., Zhang, X., Lee, Y.-H., Konuma, T., Ikegami, T.,Migita, C. T., and Kurisu, G. (2020) Crystal structure of higher plant hemeoxygenase-1 and its mechanism of interaction with ferredoxin. J. Biol. Chem. 296,100217

Structure and function of phage-encoded SAM lyase enzymes

Silvia Trigüis[1]

[1]Uppsala University

The molecule S-adenosyl methionine (SAM) is used by methyltransferases (MTases) for the methylation of DNA, RNA, proteins, lipids and small molecules. MTases of bacterial restriction-modification (RM) systems protect bacteria against bacteriophages by methylating cognate DNA.

To protect themselves from restriction, phages have developed anti-restriction systems such as elimination or methylation of the restriction sites on their genomes, and degradation of SAM. The first SAM-cleaving enzyme (SAMase), encoded by the T3 bacteriophage, was described in the 1960s and annotated as a SAM hydrolase.

No other SAMase from phage had been characterized, until the Selmer group solved the first structure of a SAMase, called Svi3-3. This enzyme is a trimer of alpha–beta sandwiches with active sites formed at the trimer interfaces. Svi3-3 and another phage-encoded enzyme called Orf1, have the same function as the T3 SAMase but their sequences share only 10 % identity. The structures of Svi3-3 with MTA and inhibitor S-adenosyl homocysteine (SAH) showed a glutamic acid (E69) and a tyrosine (Y58) as possible catalytic bases, but no water was appropriately positioned in the structures. Kinetic studies on Svi3-3 mutants showed that E69 is essential for activity and its main role is to orient the substrate in the active site.

Further analysis of Svi3-3, Orf1 and T3 SAMase reaction products by thin layer chromatography and NMR revealed that the products of the reaction are actually homoserine lactone and MTA, which categorizes these proteins as SAM lyases and not SAM hydrolases.