Abstracts

Computer-Aided Drug Design Symposium
University of Maryland, Baltimore
School of Pharmacy
May 25, 2023
Gallery and N103, School of Pharmacy
20 N. Pine Street.
University of Maryland, Baltimore, MD

https://www.pharmacy.umaryland.edu/about/depts/psc/cadd-symposium/

The symposium is primarily an in-person event. However, it will be possible to participate in the lecture portion of the symposium virtually. When registering, please indicate you participation as in person or virtual.

8:00 to 8:50: Breakfast and set up posters. All Posters will be available for viewing throughout the day.

8:50 to 9:00: Introduction from Prof. Paul Shapiro, Associate Dean for Research and Graduate Education

9:00 to 9:15: Drug discovery and research activity at the UMB School of Pharmacy
Prof. Alex MacKerell: UMB Computer-Aided Drug Design Center
Prof. Hongbing Wang: Program Chair, Department of Pharmaceutical Sciences

9:15 to 9:45: Rescuing destabilized protein mutants using small molecules

John Karanicolas
Fox Chase Cancer Center

Clinical genetics points to many human diseases for which the underlying pathology can be traced to mutations that map to the interior of a folded protein: we hypothesize that these mutations act by destabilizing an otherwise folded protein, such that the protein loses activity because an insufficient amount of the cellular population is correctly folded. However, extant drug discovery expertise is centered around inhibitors of enzyme activity and modulators of cell surface receptors, leaving our community ill-equipped to tackle the challenge of designing compounds that restore the function of proteins deactivated in this manner.

The Karanicolas group has been developing computational tools specifically catered to the shallow surface pockets typical of sites that are not naturally evolved for small-molecule binding. We have recently applied these tools to identify novel druggable sites on the surfaces of two different tumor suppressor proteins, and we identified small drug-like molecules that bind to these sites. By treating cancer cell lines that harbor these mutant tumor suppressors with the corresponding stabilizers, we find that these compounds refold these destabilized mutant proteins and restore WT activity.

We anticipate that the compounds described here may serve as a starting point for new classes of cancer therapeutics. More broadly, however, these represent first proof-of-concept for a new therapeutic modality: using small molecules to revert loss-of-function induced by mutations that act by disrupting protein stability.

9:45 to 10:15: Applications of Hydrogen-Deuterium Exchange Ensemble Reweighting to Aid Computational Drug Discovery.

Daniel Deredge
Department of Pharmaceutical Sciences
School of Pharmacy, Univ. of Maryland, Baltimore

Hydrogen-deuterium exchange mass spectrometry (HDX-MS) is a powerful approach to probe the structural dynamics of a protein in solution. It simultaneously provides time-resolved information on the structural environment of the amide hydrogen of every residue of a given protein except Prolines. As such, HDX-MS has been successfully applied to provide insight into the characterization of native state ensembles of proteins, macromolecular interactions, protein-small molecule interactions or protein folding and unfolding studies. However, the structural information reported by HDX-MS remains at peptide-level resolution. To overcome the limited structural resolution, efforts have aimed at developing quantitative and integrative approaches that leverage HDX-MS data to assess and refine structural ensembles generated by modeling and/or molecular dynamics simulations. HDX ensemble reweighting (HDXer) is one such approach which uses a *post hoc* maximum entropy approach to adjust the weights of individual frames of a structural ensemble to conform to HDX-MS data. Here, we demonstrate the applicability of HDXer within the perspective of small molecule drug discovery in two ways. First, we used HDXer to model the native ensemble of the putative therapeutic target PhuS, the cytoplasmic heme binding protein from *P.aeruginosa*, prior to structure-based drug design approaches to successfully target novel, cryptic drug binding sites that would have otherwise not been targetable using HDX-MS or MD simulations alone. Second, we demonstrate the specific applicability of HDXer in structure-based drug design in leveraging lower resolution structural information from HDX-MS to model protein-small molecule drug interactions in the absence of high-resolution drug bound structures.

10:15 to 10:30: Elucidating Excipient-Protein Interactions with Site-Identification by Ligand Competitive Saturation (SILCS)-Biologics Approach for Computationally Guided Biologics Formulation

Asuka A. Orr[1], Ahmad Kiani Karanji[1], Xun Li[1], Fang Wang[1], Daniel J. Deredge[1], Stephen W. Hoag[1], Aoxiang Tao[2], Olgun Guvench[2], Alexander D. MacKerell, Jr.[1]
1. Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland Baltimore, Baltimore, MD, USA. 2. SilcsBio LLC, 1100 Wicomico Street, Suite 323, Baltimore, MD, USA

Protein-based therapeutics, or biologics, typically require high concentrations of the active protein, which can lead to unfavorable solution behaviors such as phase separation and high viscosity due to protein-protein interaction (PPI). Such solution behaviors can be ameliorated through the inclusion of excipient molecules in the formulation of biologics to improve their stability, bioavailability, and manufacturability. Despite the importance of formulation in the development of biologics, excipients are often selected based on platform trial and error. Here, we present an approach based on site identification by ligand competitive saturation (SILCS) technology, termed SILCS-Biologics, that can

computationally guide excipient selection for biologics formulations and its application. Using precomputed 3D SILCS FragMaps, representing protein-functional group interaction patterns, the SILCS-Biologics approach maps potential excipient binding sites and PPI probability for the entire protein surface at atomistic level detail, providing a range of data to explore structure-based hypothesis for excipient impact on protein stability, aggregation, and viscosity. Computational data collected from SILCS-Biologics, including the distribution and affinity of excipients and the relation of the interactions to PPI, are used to predict excipient binding and impact on solution behavior. In addition, the SILCS-Biologics approach can account for the effect of the excipients, ions, and buffers on the effective charge of proteins and generate 3D structures of ion-bound proteins to facilitate the characterization of protein dipole moments and surface charge distribution in different salt environments. Overall, the SILCS-Biologics approach constitutes a promising structure-based method to elucidate the mechanisms by which excipients, ions, and buffers affect protein solution behavior and ultimately optimize biologics formulations. Conflict of Interest: ADM is Co-founder and CSO of SilcsBio LLC.

10:30 to 10:45: Coffee Break and Posters

10:45 to 11:15: Modeling immune recognition with deep learning

Brian Pierce
Institute for Bioscience and Biotechnology Research, University of Maryland

Accurate structural modeling of antigen recognition by antibodies and T cell receptors (TCRs) is a major challenge in computational biology, and effective approaches in this area can have a major impact on the design and development of immunotherapeutics and vaccines. We recently tested the use of the deep learning method AlphaFold to model antibody-antigen complex structures, and while it was highly successful in some cases, it failed to generate accurate models for the majority of complexes. Additional benchmarking using an expanded benchmark of over 400 recently released antibody-antigen complex structures has provided insights into determinants of success by AlphaFold. To predict the structures of TCRs in complex with peptide-MHC antigen targets, we recently developed the method TCRmodel2, which utilizes the AlphaFold pipeline with several adaptations to improve speed and accuracy. Based on benchmarking with a set of recently released TCR-peptide-MHC complex structures, TCRmodel2 outperformed AlphaFold as well as previously developed methods for the generation of near-native models for this class of structures, and it is available to the community through a web server.

11:15 to 11:45: Rational Design of Novel PIEZO1 Modulators

Yun Lyna Luo
Department of Pharmaceutical Sciences and Biotechnology
Western University of Health Sciences

Chemical modulators of mechanosensitive Piezo channels would help better understand physiology and are anticipated to yield many clinical benefits. A promising small molecule to expand the currently limited Piezo channel pharmacome is Yoda1, a Piezo1-selective activator identified from blind high-throughput screening. We previously uncovered an allosteric Yoda1 binding pocket in Piezo1, sandwiched between two hydrophobic transmembrane domains called Repeats A and B. Here, we show

that crosslinking these Repeats with disulfide bridges inhibits the effects of Yoda1, supporting a mechanism in which Yoda1 acts like a wedge by separating these two domains from each other. Using open and shut Piezo1 structural models and binding free energy calculations, we next show that Yoda1 interacts with this binding region with a higher affinity for the open state, as expected for an allosteric activator. Our free energy calculations further recapitulate structure-activity-relationships established for seven Yoda1 analogs. Among 100 purchased compounds virtually selected against a SILCS pharmacophore map of the Yoda1 binding site in the open state, several act as bona fide Piezo1 activators with chemical scaffolds distinct from Yoda1 or other known Piezo1 modulators. This work provides a structural and thermodynamic framework for the activation of Piezo1 by Yoda1 and demonstrates the possibility of computationally exploiting this pharmacological region to rationally design novel molecules with potential clinical value.

11:45 to 12:00: Unlocking Proteome-Wide Potential for Covalent Inhibition Using Machine Learning

Ruibin Liu
Department of Pharmaceutical Sciences
School of Pharmacy, Univ. of Maryland, Baltimore

Covalent inhibitors offer potentially enhanced selectivities and potencies compared to non-covalent inhibitors and may be used to target the traditionally undruggable pockets. The dynamic and complex nature of protein-ligand interactions makes the identification of covalently-ligandable sites challenging. In recent years, activity-based chemoproteomics has emerged as a powerful tool to probe proteome-wide ligandability; however, it is associated with high cost. Here we developed machine learning models based on the co-crystal structures of covalently labeled proteins in the entire Protein Data Bank. The prediction accuracy of the tree-based models reaches above 80%, while that of the convolutional neural network models is somewhat lower. Remarkably, external validation on the proteomic dataset shows similar accuracy. Our work suggests that our ML models may be used to accelerate covalent drug discovery and expand the druggable proteome space.

12:00 to 1:00 Lunch and Posters

1:00 to 2:00: Ellis S. Grollman Lecture in Pharmaceutical Sciences

Leveraging Molecular Dynamics Simulations and AI Tools to Inform the Development of Safer Opioid Medications

Marta Filizola
Department of Pharmacological Sciences
Icahn School of Medicine at Mount Sinai, New York, NY

The opioid epidemic has created a critical need for effective medications to treat opioid use disorder (OUD), a chronic and relapsing illness that affects millions of people worldwide. However, the development of such medications is hindered by the complex nature of opioid receptor activity and physiological functions. To address this challenge, cutting-edge technologies such as enhanced molecular dynamics (MD) simulations and AI/machine learning-based methodologies have emerged as

powerful tools to advance the discovery of safer and more effective opioids. By providing new insights into the molecular, dynamic, and kinetic basis of opioid receptor signaling, and extracting maximal information from the wealth of recent high-resolution experimental structures of opioid receptors, functional selectivity data, and ultra-large chemical libraries containing tens of billions of molecules, these tools can inform optimal guidelines for rational drug discovery. In this talk, I will discuss our recent efforts using MD and AI tools to generate accurate dynamic models of opioid receptor signaling and novel chemical probes with specific pharmacological profiles that have a higher likelihood of being developed into effective OUD medications. By combining these cutting-edge technologies with a deep understanding of opioid receptor signaling, we believe we are making good progress towards a brighter future for those struggling with addiction.

2:00 to 2:30: Are there extracellular conformational changes commonly associated with activation of aminergic G protein-coupled receptors?

Lei Shi
National Institute of Drug Addiction

Aminergic receptors are G protein-coupled receptors (GPCRs) that transduce signals from small endogenous biogenic amines to regulate intracellular signaling pathways. Agonist binding in the ligand binding pocket (LBP) on the extracellular side opens and prepares a cavity on the intracellular face of the receptors to interact with and activate either G proteins or β-arrestins. Here, by analyzing all the available aminergic receptor structures, we seek to identify the activation-related conformational changes independent of the specific scaffold of bound agonist, which we define as "activation conformational changes" (ACCs). While some common ACCs on the intracellular side have been well-documented, ACCs on the extracellular side, including those in the LBP are much less clear. They are complicated by local adjustments to different ligand scaffolds. However, the rotamer toggle switch ($Trp^{6.48}$) and the PIF motif at the bottom of the LBP have previously been proposed to mediate the conformational consequences of ligand binding to the intracellular side of the receptors. Our analysis shows that common ACCs in the LBP are primarily associated with the PIF motif and nearby residues, including $Trp^{6.48}$. We found no common ACCs among the extracellular subsegments, but we did identify small common rearrangements between the extracellular and middle subsegments. Thus, we propose a novel "activation switch" motif that includes the middle subsegments of TMs 3, 5, and 6, which integrates both the PIF motif and the $Trp^{6.48}$, for its common and critical role in aminergic receptor activation.

2:30 to 3:00: A single residue controls spike export and maturation in SARS-CoV-2 assembly

S. Saif Hasan[1,2,3]
[1]Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine, Baltimore MD 21201; [2]Center for Biomolecular Therapeutics, University of Maryland School of Medicine, Rockville MD 20850; [3]University of Maryland Marlene and Stewart Greenebaum Cancer Center, University of Maryland Medical Center, Baltimore, MD 21201

The spike (S) protein of SARS-CoV-2 is delivered to the virion assembly site in the ER-Golgi Intermediate Compartment (ERGIC) from both the ER and cis-Golgi in infected cells. However, the relevance and modulatory mechanism of this bidirectional trafficking are unclear. Here, using a combination of X-ray crystallography, single particle cryoEM, NMR, and functional analyses, we show that S incorporation into virions and viral fusogenicity are determined by coatomer-dependent S

delivery from the cis-Golgi and restricted by S-coatomer dissociation. Although S mimicry of the host coatomer-binding dibasic motif ensures retrograde trafficking to the ERGIC, avoidance of the host-like C-terminal acidic residue is critical for S-coatomer dissociation and therefore incorporation into virions or export for cell-cell fusion. The altered trafficking and re-localization of S modified in the coatomer binding residues leads to dramatic global changes in the maturation of immune-modulatory glycans. Because the C-terminal residue is the key determinant of SARS-CoV-2 assembly and fusogenicity, our work provides a framework for the export of S protein encoded in genetic vaccines for surface display and immune activation.

3:00 to 3:30: Coffee Break and Posters

3:30 to 4:00: Structural Studies of the hnRNP A18 Protein and Development of Small Molecule Inhibitors

Kristin Varney
Center for Biomolecular Therapeutics
School of Medicine, University of Maryland, Baltimore

The heterogeneous ribonucleoprotein A18 (hnRNP A18) is a new regulator of protein translation that is over-expressed in hypoxic regions of several solid tumors including prostate, breast, melanoma and colon cancers, as compared to normal tissues, and promotes tumor growth via the coordination of mRNA transcripts associated with pro-survival genes. Additionally, hnRNP A18 has been shown to act as a regulator of an immune checkpoint. HnRNP A18 recognizes a specific RNA signature motif in the 3'UTR of transcripts associated with cancer cell progression (Trx, VEGF, RPA) and an immune checkpoint (CTLA-4). Because hnRNP A18 regulates the translation of a variety of transcripts devoted to confer growth advantages to cancer cells, it provides a new handle on controlling tumor growth by counteracting backup mechanisms that are often used by cancer cells to bypass the inhibition of a particular pathway. This represents a significant advantage over targeting a single molecule performing a cellular function for which alternative mechanisms can be deployed. Targeting hnRNP A18 could thus provide a new mechanism to specifically inhibit the translation of an entire network of transcripts that are essential for tumor growth.

It is therefore our goal to develop small molecule inhibitors that target hnRNP A18 to control and possibly stop cancer progression as a therapeutic agent. Towards achieving this goal, we are studying the structure (X-ray) and dynamic properties (NMR) of hnRNP A18 and using these data together with computer-aided drug design (CADD) techniques to engineer inhibitors that are highly efficacious and specific for targeting A18 versus other RNPs.

4:00 to 4:30: Targeting *Pseudomonas aeruginosa* heme sensing and utilization as a viable therapeutic strategy

Aziza Frank, Garrick Centola, Fengtian Xue and Angela Wilks
Department of Pharmaceutical Science
School of Pharmacy, University of Maryland, Baltimore

Pseudomonas aeruginosa is a versatile opportunistic pathogen causing a wide variety of acute and chronic infections, especially in immunocompromised individuals. Iron is a central micronutrient

required for survival and adaptation within the host. The transition from acute to chronic infection coincides with a decrease in siderophore biosynthesis, while simultaneously increasing the ability to utilize heme. Our previous in vitro genetic and biochemical analysis of the heme uptake systems characterized the Pseudomonas heme uptake (Phu) system as the high capacity transport system, with the heme assimilation system (Has) as being essential for extracellular heme sensing and signaling. Transcriptome analysis of *P. aeruginosa* clinical isolates identified both the Has system and Phu systems as being critical for survival and adaptation within the host. We have further shown the heme metabolite BVIXβ, the product of heme degradation by HemO, is a signaling molecule linking the switch to heme utilization with cooperative behaviors associated with chronic infection and biofilm formation. Therefore, targeting heme signaling and utilization represents a multipronged therapeutic strategy to combat *P. aeruginosa* chronic infections that are intractable to current therapies. We have developed a series of highly soluble gallium salophen (GaSal) scaffolds as dual targeting inhibitors of HasAp and HemO. The compounds have inhibitory activity toward HasAp/HasR-dependent heme signaling, are taken up into the bacterial cell via xenosiderophore systems where they further inhibit HemO, and decrease the levels of BVIXβ. We propose the GaSal analogs that target multiple virulence traits will show increased efficacy while simultaneously decreasing the selective pressure to develop drug resistance in the treatment of chronic *P. aeruginosa* infections.

4:30 to 5:00: Targeted Protein Degradation: Computational Strategies and Challenges

Xiao “Sean” Zhu
Kymera Therapeutics

Heterobifunctional degraders are molecules designed to induce degradation of target proteins by harvesting the cell’s natural ability to dispose of and recycle proteins. This novel class of molecules holds promise to drugging the so-called “undruggable” class of therapeutically relevant proteins. Here, we will discuss some of the unique challenges in the discovery of therapeutically relevant heterobifunctional degraders and how existing and future computational methods can help accelerate the design of potent, safe, and orally bioavailable degrader drugs.

5:00 to 7:00: Poster session with Beer, Wine and Cheese. Note that the posters will be available throughout the day.

Poster Abstracts

Assignment of posters to specific boards will NOT be made.

Generative models to design new kinase inhibitors

Grigorii Andrianov, John Karanicolas
Fox Chase Cancer Center, Philadelphia, PA

Protein kinases are an exciting class of targets for therapeutic intervention because of their central role in mediating cellular signaling. However, the fact that the kinase catalytic domain is highly conserved makes design of selective inhibitors is challenging. Although virtual screening approaches have been designed to facilitate drug discovery process, often the success of screening campaigns depend on the size of screening libraries. In this work, we present two new approaches for building large computational libraries of novel kinase-focused chemical matter. One method uses a traditional model (a graph-based genetic algorithm), and the other uses generative deep learning (a recurrent neural network). We show that both methods can produce libraries enriched in compounds with 3D properties that complement the kinase active site.

Optimization of Polarizable Drude Nucleic Acid Force Field using Parameter Reweighing via Machine Learning

Prabin Baral,[1] Mert Yigit Sengul,[2] Alexander D. MacKerell Jr.[1]
[1]Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, Maryland 21201, United States
[2]Roivant Discovery, Roivant Sciences, Inc., New York City, New York, 10036, United States

Molecular Dynamics (MD) simulations play an essential role in determining the structural and dynamical properties of biomolecular systems. The underlying force field in MD simulations helps to determine the accuracy of the simulations. The CHARMM classical Drude oscillator polarizable force field has been implemented and developed to more accurately model the electronic response in MD simulations. To improve the MD studies of nucleic acids additional optimization of the force field has been undertaken to more accurately treat the equilibrium between the folded and unfolded states of RNA and DNA hairpins along the canonical DNA and RNA structures and quantum mechanical (QM) data on small model compounds. Trajectories from MD simulations of the RNA and DNA hairpin structures are first generated and utilized as inputs along with user selected order parameters that are clustered and selected to generate a set of reaction coordinates (RC) using machine learning to sample folding/unfolding transitions. These RCs act as the basis of enhanced sampling simulations to characterize the ensembles of conformations associated with folding and unfolding dynamics of the RNA and DNA hairpins. The distribution of conformations are then used to optimize selected dihedral parameters to produce equal populations of folded and unfolded states at the respective melting temperatures along with structural properties of canonical RNA and DNA and potential energy scans from the QM calculations. The final force field is anticipated to be of utility for simulation studies of a range of nucleic acids in different environments, including in the presence of various ions.

Using molecular dynamics simulations to gain insight into the hinge flexibility of monoclonal antibodies

Violetta Burns[1], Christina Bergonzo[2], Harold W. Hatch[3]
[1] Biomolecular structure and function Group, BMD, UMD
[2] Biomolecular structure and function Group, BMD, NIST
[3] Chemical Informatics Group, CSD, NIST.

Monoclonal antibodies (mAbs) have become one of the key biomolecules in the pharmaceutical industry for targeted therapies, such as cancer and immunology treatments and vaccines. One of the goals of pharmaceutical companies is to produce and deliver high concentration mAb formulations; however, mAbs tend to aggregate and cluster at high concentrations, increasing viscosity and decreasing the effective shelf-life of the product. One of the key elements of mAbs that is still not well understood is the flexibility conferred by the hinge region, which is believed to be critical for effective close-packing and cluster formation of mAbs. Experimental techniques such as CryoEM or X-ray crystallography are not able to provide much information about the hinge itself due to being unstructured and highly flexible. In this work, we use molecular dynamics simulations to investigate the effects of chain length, amino acid sequence and number of disulfide bonds on the dynamics of the hinges of NISTmAb and 1IGY. Preliminary results show that 1IGY samples a smaller radius of gyration compared to NISTmAb, indicating that 1IGY is more conformationally restricted. Additionally, while the distribution of angles between Fab domains for NISTmAb and 1IGY overlap, our results show that NISTmAb has a broader distribution of angles between the Fab and Fc domain, while 1IGY's angle distribution is narrower. We correlate these differences to the presence of an extra disulfide bond in the hinge region, fewer prolines and shorter amino acid sequence of the hinge region of 1IGY. These results enhance our understanding of the factors that affect the hinge flexibility, aid in the development of multiscale simulation of high concentration solutions of mAbs, and expand our knowledge on the aggregation and clustering of mAbs that has so far inhibited pharmacological advances in the field.

A Polypharmacological Approach to Relapse/Refractory Multiple Myeloma: Dual Inhibition of the Proteasome and Aggresome Pathways

Alexandria M. Chan,[1] Andrea Cottingham,[1] Ashley Mitchell,[2] Asaminew Haile Aytenfisu,[1] Alexander D. MacKerell, Jr.,[1] and Steven Fletcher[1]
[1]Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, Baltimore, MD 21201; [2]University of Maryland Baltimore County, Baltimore, MD, 21250

Multiple myeloma is a cancer of the plasma cells with a 5-year survival rate of 58%. These malignant cells overproduce abnormal proteins that must be cleared to maintain survival. As a result, proteasome inhibitors such as bortezomib have emerged as a first-line treatment for the disease. These chemotherapies target the ubiquitin-proteasome system (UPS), the main method of protein degradation; the cancer cells become reliant on this system for cell survival, whereas when inhibited, the abnormal proteins accumulate resulting in apoptosis. Unfortunately, resistance to these treatments often occurs, pushing patients into the relapse/refractory population, having much less success with treatment in addition to increased comorbidities and inability to tolerate harsh therapeutics. One such resistance mechanism to proteasome inhibitors is the upregulation of alternative methods of protein degradation, such as the aggresome pathway. Along with other cellular functions, histone deacetylase-6 (HDAC6) plays a role in the accumulation of protein aggregates into aggresomes that are then degraded through autophagy. As expected, HDAC6 is upregulated in these resistant populations. Herein, we present our work on dual HDAC6/proteasome inhibitors which have been designed in collaboration with the MacKerell lab using site identification by ligand competitive saturation (SILCS). These dual inhibitors

include alternative zinc binding groups to the traditional hydroxamic acid HDAC pharmacophore, which have proven toxic in clinical settings. Implementation of a polypharmacologic method has several advantages over traditional combination treatments including the potential for increased therapeutic effect due to the simultaneous presence of both pharmacophores in target tissues, though this is mainly displayed in vivo. In vitro results shown may give key insights towards the use of a polypharmacologic method of treatment as well as additional mechanistic information on multiple myeloma.

Limitations for using open data for PK/PD modeling

Vladimir Chupakhin
MLMol Consulting

Pharmacokinetic/pharmacodynamic (PK/PD) modeling is an important piece of the computer-aided support for the compound optimization. The growing availability of open data sources has provided new opportunities for cheminformatics and computational chemistry to contribute to PK/PD modeling. However, the use of open data in this context presents unique challenges that must be addressed to ensure reliable and accurate model development. In this review, we explore the cheminformatics and computational chemistry limitations associated with using open data for PK/PD modeling. Key challenges include the heterogeneity and quality of chemical structure representations, the variability in bioactivity data and assay conditions, and the lack of standardized metadata for computational modeling. We conclude by emphasizing the potential of open data to advance PK/PD modeling and the critical role of cheminformatics and computational chemistry in unlocking the full value of these data resources.

Understanding dimer-potent BRAF V600E inhibitors: studying a modification to pontanib

Joseph Clayton, Aarion Romany, Jana Shen
University of Maryland Baltimore, School of Pharmacy

BRAF is a kinase within the MEK signaling pathway and is hyperactive in many cancers, with 8% of all human tumors carrying a BRAF mutant and 90% of these carrying the same mutation, V600E. Current FDA-approved drugs targeting this mutation, however, target monomeric BRAF V600E and suffer heavily from tumor resistance due to allosteric effects and enhanced dimerization. A new class of inhibitors, deemed type-II, has emerged that avoids this resistance by binding to both protomers in the BRAF V600E dimer. A recent work revealed that a modification to the type-II inhibitor ponatinib (dubbed ponatinib hybrid inhibitor 1, or PHI1) reduces inhibitory activity in cells expressing monomeric BRAF V600E while maintaining activity against dimeric BRAF V600E. Here, we utilize molecular dynamic (MD) simulations to answer why the modification causes a selectivity towards dimeric BRAF V600E. We find that the αC helix is flexible even in dimeric BRAF V600E, but is restrained by PHI1. In addition, through emperical calculations we find PHI1 binds stronger when a salt bridge is formed between two conserved residues, Lys483 and Glu501; our simulations show this salt bridge is less present in monomeric BRAF V600E compared to dimeric BRAF V600E. We thus conclude that the selectivity of PHI1 is driven by the presence of this salt bridge.

A robust, viable, and resource sparing HPLC-based logP method applied to common lipophilic drugs to help in expanding in silico training datasets

Ana Luisa Coutinho, James E. Polli

Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, Baltimore, Maryland, USA

Background: The partition coefficient P (logP) is an important lipophilicity index for medicinal chemistry, dissolution, pharmacokinetics, and pharmacodynamics modeling. Reliable, experimentally determined logP for most drugs are often unavailable in the literature but are essential for future logP computer model training sets.
Goal: This study aimed to measure the logP of twelve drugs commonly used in non-clinical and clinical studies using reverse-phase high-performance liquid chromatography (RP-HPLC).
Methods: Measurements were performed on a Waters AcquityH Class UPLC system. LogP was derived from the log of a compound's retention factor at 100% water (logkw). Our compounds were too lipophilic to use 100% water as mobile phase; therefore, the extrapolated value of logk was obtained from at least five runs using buffers with various percentages of methanol. The logkw and logP of well-established reference substances were used in the calibration curves at pH 6 and 9. HPLC-based logP values were compared against literature logP values. Half of the literature logP values were in silico based, the other half was experimentally determined.
Results: Most drugs had a measured logP that was similar to or larger than previously reported logP. A possible reason for the higher logP findings here compared to in silico literature values is that computational results reflect compounds in their training sets. In silico predictions could be biased low for highly lipophilic compounds since these compounds are missing from many training sets. LogP values here generally agreed with the other few HPLC-based literature logP values. The HPLC-based logP values found here agreed partially with literature logP values found using other methodologies (± 10%).
Conclusion: A robust, viable, and resource-sparing method to measure logP was developed using RP-HPLC. This method has excellent promise to provide reliable logP values of commonly used drugs available in the literature.

Structural and molecular basis of SARS-CoV-2 spike trafficking during assembly and infection

Debajit Dey[1], Enya Qing[2], Yanan He[3], Yihong Chen[3], Benjamin Jennings[4], Whitaker Cohn[5], Suruchi Singh[1], Lokesh Gakhar[6,7,8], Nicholas P. Schnicker[7], Brian G. Pierce[3,9], Julian P. Whitelegge[5,10,11], Balraj Doray[4], John P. Orban[3,12], Tom Gallagher[2], S. Saif Hasan[1,13,14]

[1]Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine, Baltimore MD; [2]Department of Microbiology and Immunology, Loyola University Chicago, Maywood IL; [3]W. M. Keck Laboratory for Structural Biology, University of Maryland Institute for Bioscience and Biotechnology Research, Rockville MD; [4]Department of Internal Medicine, Hematology Division, Washington University School of Medicine, St. Louis MO; [5]Pasarow Mass Spectrometry Laboratory, The Jane and Terry Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine, University of California, Los Angeles CA; [6]Department of Biochemistry, Carver College of Medicine, University of Iowa, Iowa City IA; [7]Protein and Crystallography Facility, Carver College of Medicine, University of Iowa, Iowa City IA; [8]PAQ Therapeutics, Cambridge MA; [9]Department of Cell Biology and Molecular Genetics, University of Maryland, College Park MD; [10]Molecular Biology Institute, University of California, Los Angeles, Los Angeles CA; [11]Jonsson Comprehensive Cancer Center, University of California, Los Angeles, Los Angeles CA; [12]Department of Chemistry and Biochemistry, University of Maryland, College Park MD; [13]University of Maryland Marlene and Stewart Greenebaum Cancer Center, University of Maryland Medical Center, Baltimore MD; [14]Center for Biomolecular Therapeutics, University of Maryland School of Medicine, Rockville MD

SARS-CoV-2 displays complex intra-cellular trafficking of the newly synthesized spike (S) protein in infected cells. This includes bidirectional S delivery from ER and cis-Golgi to the virion assembly site in ER-Golgi Intermediate Compartment (ERGIC). However, little is known about which organelle supplies S to generate infectious, fusion-competent SARS-CoV-2 particles, and how this S delivery is modulated at the atomic-level. Here we show that S incorporation into virions and associated viral fusogenicity is determined primarily by coatomer-dependent S delivery from cis-Golgi and is restricted by S-coatomer dissociation. Our structure-function analyses find that although mimicry of the coatomer-binding dibasic motif in the S tail ensures coatomer hijacking, avoiding host-like acidic residues at the S C-terminus is critical for S-coatomer dissociation. Hence, neutral-to-acidic C-terminal mutant, which enhances S mimicry of coatomer-interacting sequences, prevents S-coatomer dissociation, virion assembly, and S export for cell-cell fusion. In contrast, S proteins lacking the dibasic motif partially bypass virion incorporation sites and traffic to PM. Thus, we identify a single residue at the S C-terminus as the key and unconventional determinant of SARS-CoV-2 assembly and fusogenicity. This provides the atomic-level framework to improve release of genetic-vaccine encoded S from the coatomer for surface display and immune activation.

High resolution cryo-EM structures of solute carrier 22 (SLC22) organic cation transporters reveal substrate and inhibitor binding mechanism

Tongyi Dou, Tengfei Lian, Xiaofeng Tan, Yi He and Jiansen Jiang
Biochemistry and Biophysics Center, National Heart, Lung, and Blood Institute, Bethesda, MD 20892, USA.

Solute carrier 22 (SLC22) organic cation transporters (OCTs) are polyspecific, bidirectional, carriers that transport organic cations like nutrients, metabolites, and drugs. Due to such polyspecificity, OCTs also mediated a mess of clinically reported drug-drug interactions (DDIs). Here we show, by optimizing sample preparation as well as data processing procedures of cryogenic electron microscopy (cryo-EM) technology, we were able to solve high-resolution (~2.3A) OCTs cryo-EM structures. In detail, structures of endogenous substrate thiamine-bound, and inhibitor abacavir-bound OCT1 were obtained. All the twelve transmembrane helices (TMs) and the periplasmic domain are clearly solved, while the density of ligands, thiamine and abacavir, are well characterized. The new information helps us understand how OCTs recognize substrates and bind to inhibitors, as well as the mechanism of corresponding DDIs. In addition, our work demonstrated that it is practical to solve structure of ~50 kDa membrane proteins using cryo-EM technology without the help of nanobodies.

Dynamics and Conformational Heterogeneity of the PIM1 G-Quadruplexes From Polarizable Molecular Dynamics Simulations

Rebekah Fogarty[1], Haley Michel[1], and Justin A. Lemkul[1,2]
[1]Department of Biochemistry and [2]Center for Drug Discovery
Virginia Tech, Blacksburg, Virginia 24061

Reproductive cancers, such as triple-negative breast cancer (TNBC), are often challenging to treat due to a lack of specific chemotherapeutic targets. Overproduction of PIM1, an anti-apoptotic protein, is common in TNBC. Thus, counteracting this overproduction could serve as a viable therapeutic strategy against TNBC. The promoter region of the PIM1 gene can adopt two G-quadruplexes (GQs) with distinct topologies. GQs are higher-order, noncanonical nucleic acid structures that are rich in guanine and have become intensely explored as potential novel drug targets due to their proposed involvement in

transcriptional regulation and other critical cellular processes. Two GQ-duplex hybrid structures have been found to exist in equilibrium in the promoter region of PIM1. The dominant GQ structure, form 1, contains conventional guanine tetrads and a duplex loop region of three Watson-Crick base pairs, while form 2 adopts a fold containing a mixed guanine-cytosine (GCGC) tetrad via slippage of two canonical Watson-Crick base pairs, and also contains a duplex loop consisting of two additional Watson-Crick pairs. We performed unbiased simulations and Gaussian-accelerated molecular dynamics simulations of both GQs with the Drude-2017 polarizable force field to better understand the properties of these GQs and how they might interconvert. Given the role of ion binding and noncanonical base pairing in stabilizing GQ structures, we monitored ion coordination and electronic properties (base dipole moments and electric fields). We observed an increased dipole moment for cytosine bases in the GCGC mixed tetrad of form 2, suggesting electronic plasticity as a factor in structural rearrangement between forms 1 and 2. We also report on conformational sampling in each structure to provide atomistic details that will be essential in the future design of more specific chemotherapeutic agents that target PIM1 on the nucleic acid level.

Pharmacological re-activation of mutant pVHL in renal cell carcinoma

Mariam Fouad[1], Christopher Parry[2], John Karanicolas[1]
[1] Department of Cell Signaling and Tumor Microenvironment, Fox Chase Cancer Center, Philadelphia, PA [2] Biomedical Sciences, Lewis Katz School of Medicine at Temple University, Philadelphia, PA

Kidney cancer is the ninth most dominant cancer type world-wide with US on top of the list About 40% of cases harbor a missense mutation in tumor suppressor VHL. Upon mapping the locations of these mutations to the protein structure, we hypothesized that many of these mutations act by thermodynamically destabilizing pVHL, such that the mutant protein is not properly folded in cells. To test this, we developed small molecule CP4.29: a compound that binds to the natively folded conformation of VHL and provides back the stability lost from mutation. My studies show that CP4.29 binds to mutant pVHL in renal carcinoma cells, leading to its accumulation. CP4.29 also restores downstream activities of this mutated tumor suppressor: VHL is an E3 ligase that ubiquitinates HIF2a and ZHX2 to induce their degradation, and treating cells with CP4.29 restores these activities in cells harboring mutant VHL. It also restores lesser-known hypoxia-independent activities of VHL: these include ubiquitin-mediated degradation of AURKA, ubiquitin-mediated translocation of clusterin, and ubiquitin-independent interactions with fibronectin-1. My observations support the hypothesis that certain cancer-associated mutations act by destabilizing VHL, and position CP4.29 for further therapeutic development in VHL-mutated kidney cancer. These studies also serve as proof-of-concept that "refolder" drugs might prove useful for other cancers driven by destabilizing mutations to tumor suppressors.

Pharmacophore-based Virtual Screening of Ultra-large Virtual Libraries with vScreenML

Thomas J. A. Graham, [1] Grigorii V. Andrianov, [2] Glenn Doyle,[2] John Karanicolas[1,2]
[1] Program in Molecular Therapeutics, Fox Chase Cancer Center, Philadelphia PA 19111
[2] Moulder Center for Drug Discovery Research, Temple University School of Pharmacy, Philadelphia, PA 19140

Chemical space accessible through purchasable make-on-demand chemical libraries has undergone explosive growth in recent years. In parallel, the cost associated with utilizing computational resources has become more affordable. As a result, virtual screening with ultra-large virtual libraries has become

increasingly feasible. Pharmacophore-based virtual screening (PBVS) relies on key structural features (shape, H-Bond donors/acceptors, etc) within the active binding pose of a query and can often be performed with GPU-acceleration. These properties make PBVS ideal for identifying potential hits that recapitulate key ligand-protein interactions required for biological activity. To take advantage of this, we developed a pipeline for conducting pharmacophore-based virtual screening that enables the use of >1B-compound virtual libraries. Our approach relies on the use of FASTROCS for pharmacophore-screening, followed by ligand-protein energy minimization with Rosetta, and scoring with vScreenML. As a preliminary test-case we have applied this approach to find inhibitors of select cancer-driving RNA-binding proteins.

In silico approach to identify the potential multi-targeting ligands for the treatment of Alzheimer's Disease

Kailash Jangid, Vijay Kumar, Ashish Ranjan Dwivedi, Naveen Kumar, Vinay Kumar, Vinod Kumar
Laboratory of Medicinal and Organic synthesis, Department of Chemistry, Central University of Punjab, Ghudda, Bathinda, Punjab

Abstract: Alzheimer's disease (AD) is one of the multifactorial and fatal disorders. Memory loss and cognitive decline occur due to the death of brain cells. The main pathophysiological processes in AD are protein aggregation, neurodegeneration, and brain inflammation. The main pathological characteristics are β-amyloid plaques, β-amyloid oligomers formation and neurofibrillary tangles formation due to hyperactive phosphorylated tau protein, oxidative stress in cells, and deficiency of acetylcholine1. The single-target molecule strategy has failed in the treatment of multifactorial diseases including neurodegenerative disorders, cancer, and infectious diseases, so current research is focused on developing the potential multitarget directed ligands to show activity in more than one molecular target with advantage of synergistic pharmacological activity of the single molecule2. Our research has been focused to develop novel components that favourably inhibits acetylcholinesterase (AChE) and simultaneously inhibit the monoamine oxidase B (MAO-B), which are associated with the treatment of AD. The computational approaches are robust and unfailing tools for identifying novel therapeutic ligands. In this research, the ASINEX database was used to identify the novel molecular using computational approaches including molecular docking, free energy calculation, ADMET, molecular dynamics etc. We found 3 lead compounds as dual inhibitors are AOP 19078710, BAS 00314308 and BDD 26909696, which are having good binding interaction with both targeted enzymes.

Integrating Computational Docking, Molecular Dynamics, and HDX-MS through Maximum Entropy Reweighting to Yield Atomistic Resolution Models of Protein-Small Molecule Binding

Kyle Kihn, Olivia Purdy, Paul Shapiro, Daniel Deredge
University of Maryland Baltimore School of Pharmacy

In the field of structure-based drug design (SBDD), obtaining a high-resolution structure of a protein in complex with a lead compound is often a critical milestone. A high-resolution structure provides detailed, three-dimensional information about the structural contacts between the small molecule and the target protein, revealing crucial information on the structural determinant of affinity and specificity and hypothesis generating information for lead optimization. However, obtaining high-resolution structures is not always achievable. In the absence of structures, computational modeling or predictive methods have been developed as an alternative with varying degrees of accuracy and reliability. Some of these

methods aim to leverage low resolution experimental data with computational modeling in an attempt to garner atomistic resolution. Hydrogen Deuterium Exchange coupled Mass Spectrometry (HDX-MS) is a particularly attractive experimental method to apply this integrative approach as intrinsic H/D exchange rates are well defined and phenomenological equation which allow for the prediction of H/D exchange from computational predictions has been developed. In this work, we developed a workflow to model high resolution drug bound complexes which includes HDXer, an integrative method of HDX-MS and molecular dynamics experiments, in combination with various small molecule docking and structure-based drug design approaches. To that end, we used the model system of Mitogen-activated protein kinase 1 (ERK2) and four of its known small molecule inhibitors resulting in atomistic resolution poses. We then validated these selected poses to the known high-resolution structures of each drug-protein complex, finding that the selected poses for 3 out of the 4 drugs come withing 3Å of corresponding crystal structure.

SILCS-RNA: Structure-based drug design approach for targeting RNAs with small molecules

Abhishek A. Kognole, Alexander D. MacKerell Jr.
UMB and SilcsBio LLC

RNA molecules can act as potential drug targets in different diseases, as their dysregulated expression or misfolding can alter various cellular processes. Noncoding RNAs account for ~70% of the human genome, and these molecules can have complex tertiary structures that present a great opportunity for targeting by small molecules. The site identification by ligand competitive saturation (SILCS) computational approach is extended to target RNA, termed SILCS-RNA. Extensions to the method include an enhanced oscillating excess chemical potential protocol for the grand canonical Monte Carlo calculations and individual simulations of the neutral and charged solutes from which the SILCS functional group affinity maps (FragMaps) are calculated for subsequent binding site identification and docking calculations. The method is developed and evaluated against seven RNA targets and their reported small molecule ligands. SILCS-RNA provides a detailed characterization of the functional group affinity pattern in the small molecule binding sites, recapitulating the types of functional groups present in the ligands. The developed method is also shown to be useful for identification of new potential binding sites and identifying ligand moieties that contribute to binding, granular information that can facilitate ligand design. While limitations are present, the SILCS-RNA approach uniquely enhances the drug discovery efforts for targeting RNAs with small molecules.

Development of SILCS kinetics methodology for the determination of ligand dissociation pathways and free energy barriers Development of SILCS kinetics methodology for the determination of ligand dissociation pathways and free energy barriers

Shashi Kumar, Mingtain Zhao, Wenbo Yu, ALexander D. MacKerell Jr.,
School of Pharmacy, Department of Pharmaceutical Sciences, University of Maryland, Baltimore.

The fast and accurate assessment of unbinding kinetics of ligands from proteins remains challenging due to high computational requirements and the lack of the information about the molecular transition states due to limited conformational sampling. Therefore, in the present study we investigate the extension of the site-identification by ligand competitive saturation (SILCS) methodology towards estimation of ligand unbinding kinetics. The SILCS methodology involves the precomputation of the distribution of a collection of solutes and water in and around a protein from which a map of the free

energy landscape of functional around proteins is obtained. This landscape, termed FragMaps, was tested for its ability to identify and quantify of transition states for ligand dissociation from which estimations of koff may be made. The proposed SILCS-kinetics (SILCS-KIN) method is implemented to sample the free-energy landscape of drug dissociation pathways. Target data for method validation was a manually curated compilation of koff values of ligands reported in the BindingDB database and in the published research literature. Proteins for model development were chosen based on multiple criteria such as ligand binding site geometry, knowledge about ligand dissociation pathways, and biochemical context, including membrane proteins. The proteins were subjected to combined SILCS Grand Canonical Monte Carlo/Molecular Dynamics (GCMC/MD) simulations using the additive CHARMM36m force field. From these simulations, grid free energy (GFE) FragMaps for functional groups were generated and then ligands were subjected to the SILCS Monte Carlo (SILCS-MC) optimization in the field of the pre-computed GFE Fragmaps along comprehensive dissociation pathways to estimate the absolute free-energy profiles. Methodological details and results from the training set of proteins and ligands will be presented to determine the potential utility of the SILCS-KIN methodology as a potential tool for the discovery and design of drug-like compounds with optimized ligand dissociation properties.

Generalizing Drude Polarizable Force-Field to Drug-like Molecules

Anmol Kumar and Alexander D. MacKerell Jr.
School of Pharmacy, University of Maryland Baltimore

CHARMM Drude force-field (FF) is a well-established atomistic FF for biomolecules such as proteins, nucleic acids, lipids, and carbohydrates. Its ability to capture electronic polarization effects via auxiliary particles (Drude oscillators) attached to non-hydrogen atoms sets it apart from commonly applied additive FFs, which rely on fixed charges. Extension of Drude FF to novel drug-like molecules is challenging as it requires a mechanism to assign electrostatic (atomic charges as well as polarizabilities), bonded and van der Waals parameters to non-biomolecular systems. We have already developed a deep neural network (DNN) model that can determine RESP charges on atoms as well as lone-pairs (an integral feature of Drude FF) and atomic polarizabilities (Alpha) and Thole scaling factors on non-hydrogen atoms. In addition, we have also put significant effort towards the determination of Lennard-Jones (LJ) parameters using a set of DNN models for various atom types belonging to conjugated alkenes, alkynes, nitriles, amines, nitro-benzyl species, bipyrroles and so on. The current set of bonded parameters are assigned based on analogy of bonded atoms in the given molecule with existing Drude FF parameters of biomolecules. However, wider diversity of chemical space in drug-like molecules requires significant expansion in the existing parameters. Because bonded parameter optimization requires quantum mechanically (QM) derived potential energy scans for a range of values, this data-intensive problem is handled by normal mode sampling (NMS) method for generating molecular conformations. NMS method significantly reduces the number of conformations to be dealt with, without compromising the sampling of energetically important conformations. To improve the performance of bonded parameters, Monte Carlo Simulated Annealing (MCSA) method is applied on batches of molecules sharing them. This iterative approach towards improvement of bonded parameters is validated by examining the RMSD between QM and MM optimized structures using renewed parameters. Our multifaceted approach towards optimizing all components of Drude FF using huge amount of chemical data will harbinger the extension of Drude force-field to novel drug-like molecules. Details of parameter optimization especially for drug-like molecules will be presented in the context of Drude FF.

Machine learning for identification of protein-protein interactions

Victoria Mischley[1, 2], John Karanicolas[2, 3]
[1]. Drexel University, Molecular Cell Biology and Genetics, Philadelphia, PA 19103, USA [2]. Fox Chase Cancer Center, Molecular Therapeutics, Philadelphia PA, 19111 [3]. Temple University School of Pharmacy, Moulder Center for Drug Discovery Research, Philadelphia, PA 19140, USA

Identifying and modeling protein-protein interactions is crucial for understanding cellular pathways and has implications for our understanding of diseases and drug discovery. Traditionally, protein-protein interactions have been discovered and validated using a variety of experimental laboratory techniques. However, these classic laboratory techniques are often low throughput and have varying degrees of accuracy. As the human proteome consists of at least 20,000 proteins, building comprehensive interaction maps will require the use of methods that allow vastly higher throughputs. A natural choice for this task is to screen for protein-protein interactions in-silico. While many methods have been published that have provided in-silico methods for identifying protein-protein interactions, they typically lack the accuracy needed for confidently screening for new interactions de novo. AlphaFold2 is a deep learning method that far surpassed other protein modeling tools in its ability to accurately predict not only monomeric, but also multimeric protein structures. Further, AlphaFold2 also provides confidence metrics with each prediction, to access model quality. While AlphaFold2's confidence metrics are useful to access model quality, though, they do not directly allow one to distinguish true vs false ("decoy") pairs of interacting proteins. In this work, we show that by using machine learning, we can leverage confidence metrics from AF2 in combination with energetic features and interface composition features to accurately discriminate between true protein-protein interactions and non-interacting "decoy" pairs.

Balancing Monoatomic Ion-Biomolecular Interactions in the Polarizable Drude Force Field

Yiling Nan and Alexander D. MacKerell Jr
University of Maryland Computer-Aided Drug Design Center, Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD 21201 MD

Molecular dynamics (MD) simulation is a widely used tool to study molecular behavior with atomic precision. In MD simulations an accurate force field is required for convincing and rigorous results. Traditional additive force fields, which average the effect of atomic polarizability, have been challenged in capturing the electronic response to heterogeneous environments. As a result, additive fore fields usually yield bad reproducibility of experimentally observed thermodynamic properties in systems containing high ionic concentrations with heterogeneous environments. Such heterogeneity is ubiquitous in biological systems (i.e., ion channels, ion bridging) and other industrial applications. Better describing the heterogeneous environmental response of polarizable atoms (i.e., ions around different protein functional groups) can improve the atomistic understanding of molecules obtained from MD simulations. In the present work, we refine the interactions between monoatomic ions (Li+, Na+, K+, Rb+, and Cs+) and typical functional groups present in biomolecules using a polarizable force field based on the classical Dude oscillator model. Previously, ion parameters were optimized to reproduce the hydration-free energies and coordination geometries with water in the Drude-2013 force field. In the present work, instead of reparameterizing existing ion, biomolecule, and water parameters, we use the pair-specific LJ term (called NBFIX in CHARMM) and through-space Thole dipole screening term (NBTHOLE) to fit a combination of quantum mechanical (QM) and experimental target data. NBFIX overcomes the limitations due to LJ interactions being calculated using predefined combination rules. We target the gas-phase QM interaction energies as well as condensed-phase osmotic pressure under varying ion concentrations. In addition to single ion-model compound interaction energies, multiple ions, and model compound interactions are targeted for parameter optimization. The latter is shown to be more

representative of interactions occurring in an aqueous environment. This approach is anticipated to lead to an important refinement of the Drude polarizable FF parameters for modeling ion-biomolecular interactions.

SLICE: A No-Code Platform For Encoding Chemistry & Generating Virtual Libraries

Stefi Nouleho[1], Victorien Delannée[1], Philip Judson[2], Olga Grushin[1], Nadya Tarasova[3], Marc Nicklaus[1]
[1]Computer-Aided Drug Design Group, Chemical Biology Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD, 21702, USA.
[2]Heather Lea, Bland Hill, Norwood, Harrogate, HG3 1TE, England.
[3]Synthetic Biologics and Drug Discovery Group, Cancer Innovation Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD, 21702, USA

We present SLICE (Smarts and Logic In ChEmistry), a language for encoding chemistry with the goal of fast generation of on-demand virtual libraries of compounds. SLICE is part of the SAVI (Synthetically Accessible Virtual Inventory) project. To make the methodology easily accessible for both cheminformaticians and chemists, we have created SLICE Designer, which is an open-source No-Code Development Platform (NCDP) combining SMARTS with a logic language. The current version allows the SLICE transform writer to generate SMARTS by drawing patterns and configuring each atom and bond in the pattern, to add information about the reaction conditions, and to define chemical constraints/logic to create reactions with a high likelihood of synthetic success. At the end of using SLICE Designer, an XML SLICE file is generated, which can be used for à la carte molecule generation. Given a set of building blocks, the program checks their compatibility with the reactant patterns and then enumerates compatible reactions for easily synthesizable compounds.

Modeling Binding Pathways and Affinity Maturation of of de novo Designed Miniprotein Binders of H1N1 Hemagglutinin HA2 Using Markov State Models and Massively Parallel Free Energy Calculations

Dylan Novack,[1] Si Zhang,[2] Vincent A. Voelz[1]
[1] Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122, USA [2] Redesign Science, New York, NY 10014, USA

Hyperstable miniproteins can be de novo designed to tightly bind protein targets , a promising new avenue for treating infectious disease. Pioneering work from the Baker lab at University of Washington has resulted in de novo designed miniproteins that bind H1N1 hemagglutinin stem protein (HA2), inhibiting the conformation change of the fusion peptide region necessary for cellular infection. Unlike the variable HA1 head region typically targeted in influenza therapeutics, the HA2 region is highly conserved across influenza variants. After an initial high-throughput screening of computational designs to identify binders, site-saturated mutagenesis (SSM) and affinity maturation were used to improve the dissociation constants by two orders of magnitude. Interestingly, affinity-matured variants have mutations at non-interfacial residues; moreover, SSM predictions from ROSETTA differ substantially from the experimental measurements. To determine if all-atom molecular simulation methods may offer better prediction and/or explanation of these effects, we performed massively parallel GPU-accelerated simulations of ab initio binding for three pairs(wild-type and affinity- matured) of miniprotein HA2 binders on the Folding@home distributed computing platform. Analysis of these simulations using Markov state models (MSMs) can determine if conformational dynamics helps to explain how non-

interfacial mutations enhance binding affinity. To determine if all-atom simulation methods can more accurately predict affinity-matured variants, we are using massively parallel alchemical free energy simulations to perform in silico SSM for each of the wild-type miniproteins. These strategies present a new framework for computational miniprotein design that provides mechanistic understanding and predictive design of high-affinity binders.

A Structure-Based Drug Design Approach for Targeting the Dengue Virus Nonstructural 5 (NS5) Protein

Juliet O. Obi, Linfah McQueen, Daniel J. Deredge
Department of Pharmaceutical Sciences, University of Maryland School of Pharmacy, Baltimore, MD 21201

Dengue virus is the most prevalent arthropod-borne virus and is a part of the flavivirus family. There are four known dengue serotypes, and humans can be infected with more than one serotype which can lead to severe dengue. There are no clinically approved antiviral drugs for the treatment of dengue infection to date, and vaccines available are limited to seropositive individuals. The non-structural 5 (NS5) protein is the largest protein encoded by flaviviruses and is a major component of the viral replication complex. Dengue NS5 has an N-terminal methyltransferase (MTase) domain responsible for 5' RNA capping, and a C-terminal RNA-dependent-RNA-polymerase (RdRp) domain responsible for de novo RNA synthesis. Dengue NS5 is also known to interact with RNA elements in the 5'-untranslated regions to promote viral RNA synthesis, in particular the 5' stem loop A (SLA). However, there is limited structural characterization of the dengue NS5/SLA complex. Here, we characterized the interaction of the full length and individual domains of dengue serotype 2 (DENV2) NS5 with SLA at the 5'-UTR using surface plasmon resonance (SPR) studies, electromobility shift assays (EMSAs) and hydrogen-deuterium exchange coupled to mass spectrometry (HDX-MS). HDX-MS studies were used to probe the structural dynamics of DENV2 NS5 and its binding interface with SLA. Results from these studies suggest that coordination is required between the MTase and RdRp domains, which stabilize the interaction between NS5 and SLA. We leveraged the regions of protection on NS5 from our HDX data for structure-based drug design using the site identification by ligand competitive saturation (SILCS) approach for pharmacophore screening. We propose lead compounds that can be tested for activity against dengue NS5.

Analysis of the ERK Pathway Cysteinome for Targeted Covalent Inhibition of RAF and MEK kinases.

Aarion Romany,[1] Jana Shen[1,2]
[1]Department of Pharmaceutical Sciences, University of Maryland Baltimore, Baltimore, MD, United States [2] Institute for Bioscience and Biotechnology Research, University System of Maryland, Rockville, MD, United States.

The ERK pathway is one of the most important signaling cascades involved in tumorigenesis. So far, eight noncovalent inhibitors of RAF and MEK kinases in the ERK pathway have been approved by the FDA for the treatment of cancers; however, their efficacies are limited due to various resistance mechanisms. There is an urgent need to develop novel targeted covalent inhibitors. Here we report a systematic study of the covalent ligandabilities of the ERK pathway kinases (ARAF, BRAF, CRAF, KSR1, KSR2, MEK1, MEK2, ERK1, and ERK2) using constant pH molecular dynamics titration and pocket analysis. Our data revealed that the hinge GK (gate keeper)+3 cysteine in RAF family kinases

(ARAF, BRAF, CRAF, KSR1, and KSR2) and the back loop cysteine in MEK1 and MEK2 are reactive and ligandable. Structure analysis suggests that the type II inhibitors belvarafenib and GW5074 may be used as scaffolds for designing pan-RAF or CRAF-selective covalent inhibitors directed at the GK+3 cysteine, while the type III inhibitor cobimetinib may be modified to label the back loop cysteine in MEK1/2. The reactivities and ligandabilities of the remote cysteine in MEK1/2 and the DFG-1 cysteine in MEK1/2 and ERK1/2 are also discussed. Our work provides a starting point for medicinal chemists to design novel covalent inhibitors of the ERK pathway kinases.

Structure determination of human organic cation transporter 3 to characterize the mechanism of transport and basis of polyspecificity

Jonathan Scalabrini, Shi Shu, Priyanka Abeyrathne, Tongyi Dou, Jiansen Jiang
Laboratory of Membrane Proteins and Structural Biology, Biochemistry and Biophysics Center, National Heart Lung and Blood Institute, National Institutes of Health, Bethesda, MD

Organic cation transporters (OCTs) function in the translocation of endogenous and exogenous cationic, zwitterionic, and neutral substrates across the cell membrane in multiple tissues. Members of this protein family have a broad range of physiologically relevant, overlapping substrates; however, the exact mechanism of this substrate polyspecificity is not fully understood. Furthermore, the details of substrate transport are not fully described. To understand these functionalities, this study aimed to determine the apo and substrate-bound structures of human OCT3 (SLC22A3) by cryo-EM. Here, we report the successful purification of human OCT3 in detergent micelles and describe efforts to solve the apo structure of the transporter by cryo-EM. Particularly, negative stain electron microscopy indicates the presence of homogenous, mono-disperse particles, and 2D class averages from a preliminary cryo-EM dataset reveal secondary structural features. These findings offer early insight into the structure of human OCT3 and serve as the backbone for future studies into the mechanism of human OCT3 mediated substrate transport and polyspecificity. Additionally, given the relatively small size of this protein (62kDa), these results provide important methodological insight regarding cryo-EM structure determination of small membrane proteins.

Structural analysis of SARS-CoV-2 spike maturation by single particle cryoEM

Suruchi Singh[1], Yi Liu[2], Meghan Burke[2], Benjamin Jennings[3], Debajit Dey[1], Brian G. Pierce[4,5,6], Balraj Doray[3], Stephen Stein[2], and S. Saif Hasan[1,6,7]
[1]Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine, Baltimore MD; [2]Mass Spectrometry Data Center, Biomolecular Measurement Division, National Institute of Standards and Technology, Gaithersburg MD; [3]Department of Internal Medicine, Hematology Division, Washington University School of Medicine, St. Louis MO; [4]W. M. Keck Laboratory for Structural Biology, University of Maryland Institute for Bioscience and Biotechnology Research, Rockville MD; [5]Department of Cell Biology and Molecular Genetics, University of Maryland, College Park MD; [6]University of Maryland Marlene and Stewart Greenebaum Cancer Center, University of Maryland Medical Center, Baltimore MD; [7]Center for Biomolecular Therapeutics, University of Maryland School of Medicine, Rockville MD

The SARS-CoV-2 spike encoded in COVID-19 genetic vaccines from Pfizer-BioNTech, Moderna, and Astra Zeneca, is synthesized in the endoplasmic reticulum (ER), glycosylated in ER and Golgi

compartments, and then exported to the plasma membrane (PM). This ER-to-PM export is accompanied by maturation of spike glycans, which control the conformational accessibility of the receptor binding domain (RBD) and RBD-localized immunogenic epitopes. However, despite extensive structural analyses of mainly the truncated spike ecto-domain fragment, and the publication of over 1000 truncated spike structures, no atomic level insights are available into spike glycan maturation. This is a fundamental knowledge-gap in our understanding of how RBD conformations and epitope accessibility undergo glycan maturation-associated changes during spike biogenesis. Here we present the first single particle cryoEM analysis of a full-length spike maturation intermediate at a resolution of 3.3Å. We demonstrate that the three RBDs in this spike maturation intermediate display distinct extents of conformational changes. These RBD conformational changes are restricted to a limited volume around the consensus RBD positions. It is thus inferred that this spike maturation intermediate demonstrates a homogeneous and stable conformational ensemble of RBDs. Mass spectrometry analysis of this spike maturation intermediate shows higher homogeneity of glycans relative to secreted ecto-domain fragments. Collectively, these complementary investigations provide fundamental structural insights into spike biogenesis and maturation and open new avenues to elucidate conformational modulation in the spike by tissue and organ-specific glycan modifications.

Developing an HDX-MS-guided Adaptive Sampling Workflow for Molecular Dynamics Simulations

Ally K. Smith, Daniel J. Deredge
Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, MD 21201

Conventional molecular dynamics (cMD) simulations depict the structure and dynamics of biomolecules with atomistic resolution. Due to crystallographic biases, forcefield inaccuracies, and timescale limitations, however, cMD may inaccurately reflect a protein's behavior in solution. This discrepancy is compounded by the conformational heterogeneity of the molecule. For example, calcium-free calmodulin remains compact throughout long cMD simulations, despite the C-terminal domain's flexibility in solution. Enhanced sampling methods (eMD) bias the system to facilitate greater exploration of the protein conformational landscape. Some eMD techniques incorporate experimental data as a positional restraint to mitigate spurious conformations, though this typically requires a well-defined and highly resolved experimental observable. Hydrogen-deuterium exchange mass spectrometry (HDX-MS) is a solution-based technique that offers unique insights into native protein structure and dynamics. Yet, the low spatial resolution of HDX-MS, coupled with a poor understanding of how various structural features influence the rate of H/D exchange, limits its compatibility with these experiment-directed modes of eMD. Hence, an HDX-MS-guided adaptive sampling workflow was developed to better capture solution-based conformations. Adaptive sampling involves performing multiple short, independent simulations simultaneously and iteratively propagating a subset of these according to the selection criteria. Here, five unbiased simulations of apo-calmodulin were run in tandem for 200 ns each using the CHARMM36m forcefield. The average HDX rates of protein backbone amide hydrogens were predicted for each trajectory and the simulations were ranked by RMSE between the predicted and experimental data. Three 100 ns simulations were seeded from the two highest-ranked simulations. This algorithm was repeated until achieving a cumulative 2 μs trajectory. Whereas cMD simulations of apo-calmodulin diverged from the in-solution ensemble, adaptive sampling simulations showed improved agreement throughout the simulation. Further, HDX-MS-guided adaptive sampling simulated the C-terminal domain motion of apo-calmodulin more accurately than conventional MD.

Building Robust QSAR Models of Dopamine D2 and D3 Receptors' Ligands Using Machine and Deep Learning Approaches

Sung Joon Won, Kuo Hao Lee, and Lei Shi
Computational Chemistry and Molecular Biophysics Section, Molecular Targets and Medications Discovery Branch, National Institute on Drug Abuse, Intramural Research Program, National Institutes of Health, Baltimore, Maryland 21224, United States

Dopamine receptors are aminergic G protein-coupled receptors that play a significant role in regulating various physiological functions. Dysfunctions at dopamine D2 receptor (D2R) are associated with psychiatric and neurological conditions like Parkinson's disease, schizophrenia, and substance abuse disorder. Current primary treatments for these conditions are D2R antagonists, but these often cause motor side effects. As a result, recent research efforts have been devoted to dopamine D3 receptor (D3R) as a therapeutic target since it elicits similar effects without the motor side effects. However, the close structural homology of the D2R and D3R make it a challenging task to search for D3R-selective antagonists. In this study, we aimed to establish machine learning-based quantitative structure-activity relationship (QSAR) models to predict compounds' binding affinity for either D2R or D3R. We queried and collected the D2R and D3R binding affinity training datasets from ChEMBL, specifically developed a suite of filters to extract high-quality and specific datasets, and manually curated to exclude misannotated entries. We applied both deep neural network (DNN) and extreme gradient boosting (XGBoost) algorithms to build the QSAR models. For DNN models, we developed a protocol to systematically tune the hyperparameters by exploring the feasible ranges and considering their potential dependency, and this resulted in significantly improved benchmarks that are comparable to those of the XGBoost models. In addition, we evaluated the impact of compound protonation state on the model performance and found some improvements. Taken together, we established D2R and D3R QSAR models with robust power in predicting the binding affinity for each individual receptor, which can be used to identify the D3R- or D2R-selective compounds.

Covalent Drug Design using the Site-Identification by Ligand Competitive Saturation (SILCS) Method

Wenbo Yu, Alexander D. MacKerell, Jr
Department of Pharmaceutical Sciences, School of Pharmacy University of Maryland, 20 Penn St, Baltimore, MD, USA

Covalent drug design becomes an important component in drug discovery nowadays. Traditional drug interacts with a target in an equilibrated state while covalent drug holds the promise of increasing the duration of interactions with a target by forming covalent bond with target residues, and thus may offer more effective therapeutics. For this aspect, computational methods can be used to help identify reactive cysteines on a target protein for covalent binding, to test various warhead groups on ligands for their effectiveness, and to predict covalent binding strengths for designed inhibitors. During the study of protein systems, we developed a functional group mapping approach based on explicit solvent all-atom molecular dynamics simulations (SILCS: Site Identification by Ligand Competitive Saturation) which intrinsically considers molecular flexibility and desolvation effects along with solute-solvent interactions. This method was extensively verified to be useful in many aspects of computer-aided drug discovery. In the current study, we applied this method to study covalent binding. By docking a set of molecules with representative warhead groups using SILCS-Monte Carlo (SILCS-MC), reactive

cysteines of tested protein systems can be correctly identified. Furthermore, a machine learning model was trained to quantify the effectiveness of various warhead groups for a protein target using metrics from SILCS as well as experimental warhead reactivity data. The ability of ranking small covalent molecular binders with the same warhead group using SILCS ligand grid free energy (LGFE) was also tested for multiple protein systems. The promising result indicates that the SILCS method can both qualitatively and quantitatively inform the covalent drug discovery.

Cyclic Peptide Linker Design and Optimization by Molecular Dynamics Simulations

Lei Yu, Stephanie A. Barros, Chengzao Sun, Sandeep Somani
Janssen Research & Development, LLC, Spring House, PA 19477, USA

Macrocyclic peptide is an emerging therapeutic modality that can target protein-protein interactions with high affinity and selectivity. Designing optimal linkers for macrocyclization of peptide hits is an important problem in rational design of peptide therapeutics. We present a method for predicting the impact on the conformation of a peptide by varying linker length and chemistry. The method is based on enhanced sampling molecular dynamics simulations and compares the conformational stability of series of cyclic peptides with different linker lengths, substitution patterns, and linker types. We applied it to three macrocyclic peptide series that bind to PCSK9, trypsin, and MDM2 with secondary structures spanning helix, β-sheet, and loop. In general, we observe high correlation between the structural rigidity and binding affinity of the peptides. This method may be used to design peptides with increased potency and chemical stability.

Acceleration of the SILCS Methodology through Porting of the Grand-Canonical Monte Carlo and Ligand Monte Carlo Sampling Algorithms to Graphical Processing Units

Mingtian Zhao[1], Wenbo Yu[1], Abhishek A. Kognole[2], Oliver Tao[2], Sunhwan Jo[2], Alexander D. MacKerell[1, 2]
[1]Department of Pharmaceutical Sciences, School of Pharmacy, University of Maryland, Baltimore, Maryland 21201, United States [2]SilcsBio LLC, Baltimore, Maryland 21230, United States

Accurate and rapid calculation of ligand-macromolecule binding free energies has many applications in computer-aided drug design (CADD) from finding binding hotspots, screening large drug-like molecule databases, and for optimization of ligands to approve their affinity to targets such as proteins and RNA. The site-identification by ligand competitive saturation (SILCS) ligand grid free energy (LGFE) evaluation is a both computationally efficient and accurate method for estimating ligand-macromolecule binding affinities. Central to the SILCS approach is the use of Grand Canonical Monte Carlo (GCMC) simulations of water and multiple small solutes such at benzene, propane, formamide, methylammonium and acetate in the condensed phase to sample their distribution around macromolecules from which, in conjunction with MD simulations, allow for their probability distributions to be calculated, which may then be normalized and converted to grid free energies (GFE) that are used in many facets of CADD. To facilitate the GCMC calculations we ported the method to graphical processing units (GPUs) including facilitating the acceptance rates though the use of cavity bias, configurational bias and system partitioning, showing significant speed up versus the CPU-based GCMC implementation, especially in large systems. In addition, the SILCS-MC docking algorithm was extended to GPUs. Using the multi-threading advantage of GPUs, multiple ligands can be docked simultaneously on a single or multiple binding sites on the target macromolecule. Test were also undertaken on the use of genetic algorithms for SILCS docking. These approaches will facilitate the

application of the SILCS methodology thereby potentially improving accuracy and the applicability of the method to large heterogeneous macromolecular systems.

Molecular mechanism of positive allosteric modulation at the dopamine D1 receptor

Alexander Goldberg, Lei Shi
Computational Chemistry and Molecular Biophysics Section, Molecular Targets and Medications Discovery Branch, NIDA IRP

G protein-coupled receptors (GPCRs) are involved in a variety of cellular processes and are targeted by approximately 40% of prescribed drugs. The dopamine D1 receptor (D1R) is a Gs-coupled class A rhodopsin-like GPCR. The D1R has been implicated as a potential target for treating many psychiatric disorders including Parkinson's disease, Alzheimer's disease, and schizophrenia. Upregulation of the D1R may help ameliorate motor and cognitive symptoms, however, orthosteric agonists have significant drawbacks that limit their potential in clinical applications. In the past few years, allosteric compounds have shown growing potential as therapeutics and several allosteric compounds that target the D1R have been discovered, including LY3154207, a highly potent and D1R-selective positive allosteric modulator. Recently, cryo-EM structures of the D1R in the active state have been reported in the presence and absence of LY3154207, which has provided great starting points to reveal the binding and allosteric mechanism of this compound. By integrating the information from all the relevant D1R structures, we built a relatively complete D1R-Gs complex model and performed atomistic molecular dynamics (MD) simulations. Our simulation results show that LY315207 prefers the horizontal binding orientation but not the vertical orientation above intracellular loop 2 (IL2). The analysis of our comparative simulations of the D1R-Gs models in the presence and absence of LY3154207 elucidated subtle but significant conformational differences. In particular, we found that LY3154207 stabilized the helical conformation of IL2 and compressed some regions in the transmembrane domain. The allosteric pathway we identified is likely associated with the positive allosteric effect of LY3154207 to enhance Gs protein coupling. Taken together, we propose an allosteric mechanism of LY3154207 at the D1R that may facilitate rational design of more effective allosteric modulators.

Are there extracellular conformational changes commonly associated with activation of aminergic G protein-coupled receptors?

Kuo Hao Lee[1], Jamie J. Manning[2,3,4], Jonathan Javitch[2,3,4], Lei Shi[1,]
[1]Computational Chemistry and Molecular Biophysics Section, Molecular Targets and Medications Discovery Branch, National Institute on Drug Abuse – Intramural Research Program, National Institutes of Health, Baltimore, MD 21224, USA
[2]Division of Molecular Therapeutics, New York State Psychiatric Institute, New York, NY 10032, USA
[3]Department of Molecular Pharmacology and Therapeutics, Vagelos College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA
[4]Department of Psychiatry, Vagelos College of Physicians and Surgeons, Columbia University, New York, NY 10032, USA

Aminergic receptors are G protein-coupled receptors (GPCRs) that transduce signals from small endogenous biogenic amines to regulate intracellular signaling pathways. Agonist binding in the ligand binding pocket on the extracellular side opens and prepares a cavity on the intracellular face of the

receptors to interact with and activate G proteins and β-arrestins. Here, by reviewing and analyzing all available aminergic receptor structures, we seek to identify activation-related conformational changes that are independent of the specific scaffold of bound agonist, which we define as "activation conformational changes" (ACCs). While some common intracellular ACCs have been well-documented, identifying common extracellular ACCs, including those in the ligand binding pocket, is complicated by local adjustments to different ligand scaffolds. Our analysis shows no common ACCs at the extracellular ends of the transmembrane helices. Furthermore, the restricted access to the ligand binding pocket identified previously in some receptors is not universal. Notably, the $Trp^{6.48}$ toggle switch and the $Pro^{5.50}$-$Ile^{3.40}$-$Phe^{6.44}$ (PIF) motif at the bottom of the ligand binding pocket have previously been proposed to mediate the conformational consequences of ligand binding to the intracellular side of the receptors. Our analysis shows that common ACCs in the ligand binding pocket are associated with the PIF motif and nearby residues, including $Trp^{6.48}$, but fails to support a shared rotamer toggle associated with activation. However, we identify two common rearrangements between the extracellular and middle subsegments, and propose a novel "activation switch" motif common to all aminergic receptors. This motif includes the middle subsegments of transmembrane helices 3, 5, and 6, and integrates both the PIF motif and $Trp^{6.48}$.