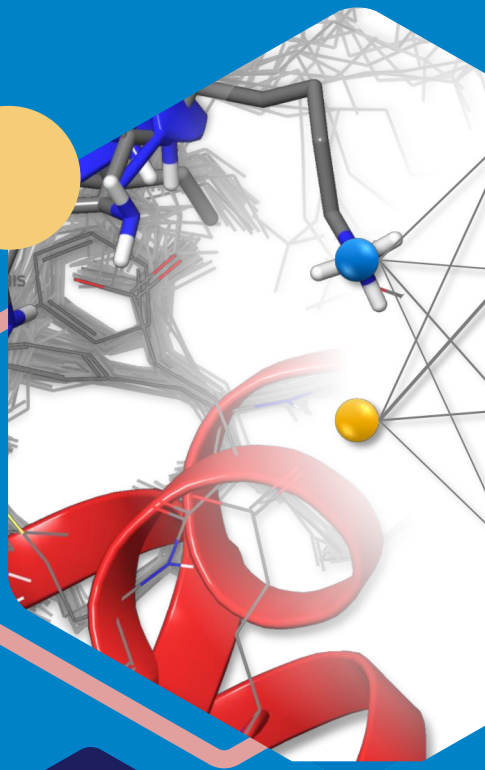
Apotekarsocieteten
Swedish Pharmaceutical Society



The 2nd Nordic Conference on Computational Chemistry

18-19 March 2025

Gothenburg, Sweden

Sponsors

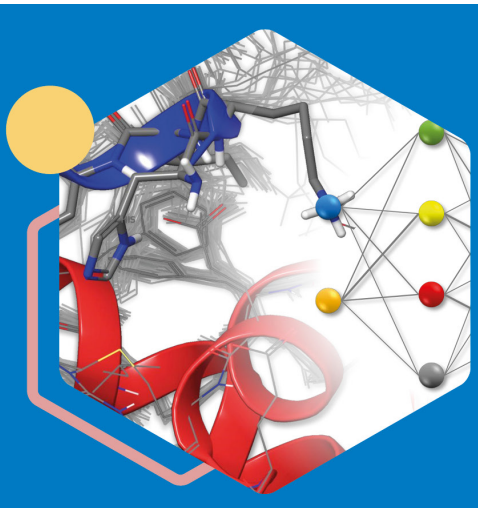
Table of contents

Sponsors 2
Program 3
List of poster presenters 5
Poster Abstracts 6
List of participants 44
Scientific Organizing Committee 46

Apotekarsocieteten

The 2nd Nordic Conference on Computational Chemistry

18-19 March 2025



Program

Tuesday 18th March	
09.15 – 10.00	Registration
10.00 – 10.10	Welcome
10.10 – 12.10 *10.10 – 10.50* *10.50 – 11.30* *11.30 – 12.10*	**Session 1:** Chair: Simone Fulle, Novo Nordisk • **The power and pitfalls of artificial intelligence in the design of new drugs,** Charlotte Deane, Oxford University • **Computational biophysics at the proteome scale,** Kresten Lindorff-Larsen, Copenhagen University • **An integrative approach to study intrinsically disordered proteins: From solution behavior to interaction with surfaces,** Marie Skepö, Lund University
12.10 – 12.40	**Poster Flash presentations**
12.40 – 13.40	**Lunch**
13.40 – 15.00 *13.40 – 14.20* *14.20 – 15.00*	**Session 2:** Chair: Jens Carlsson, Uppsala University • **Current Developments in Chemical Space Exploration and Molecular Design,** Mathias Rarey, University of Hamburg • **Trends in GPCR drug discovery,** David Gloriam, Copenhagen University
15.00 – 15.20	**Selected talk:** **Discovery of small molecule modulators of FZD_7 using in silico docking screens,** Magdalena M. Scharf, Karolinska Institutet
15.20 – 16.20	**Break & poster session**
16.20 – 17.20 *16.20 – 16.40* *16.40 – 17.00* *17.00-17.20*	**Selected talk:** **The ribosome lowers the entropic penalty of protein folding,** Julian O. Streit, University College London **Innovative computational drug design strategies at the intersection of automated molecule generation, machine learning, and free energy calculations,** Aniket Magarkar, Boehringer Ingelheim Pharma GmbH & Co **Finding better drugs: advice for optimizing multi-parameter optimization scores,** Jenny Viklund, AstraZeneca
19.00 –	**Dinner**

Wednesday 19th March	
08.30 – 09.00	Entrance to conference
09.00 – 10.20 *09.00 – 09.40* *09.40 – 10.20*	**Session 3:** Chair: Ruth Brenk, University of Bergen • **Engineering Molecules to Specification with Generative AI,** Rocio Mercado, Chalmers • **Novel Technologies Impacting Drug Discovery across Modalities at Lundbeck,** Henrik Keränen, Lundbeck
10.20 – 11.10	**Break & poster session**
11.10 – 11.50	• **The holy grail of generative AI – from hit to candidate drug,** Eva Nittinger, AstraZeneca
11.50 – 12.10	**Selected talk:** **Bimodal Data Fusion of Compound Structures and Cell Morphology Features for AI-driven Mechanism of Action Prediction in Cancer Drug Discovery,** Osheen Sharma, Karolinska Institutet
12.10 – 13.15	**Lunch**
13.15 – 14.35 *13.15 – 13.55* *13.55 – 14.35*	**Session 4:** Chair: Ola Engkvist, AstraZeneca • **Exploring Chemical Space with Assembly Theory & Chemputation,** Lee Cronin, University of Glasgow • **Computational studies of permeation enhancers for orally administered peptide drugs,** Per Larsson, Uppsala University
14.35 – 15.00	**Poster Award and closing**
15.10 –	**Tour of AstraZeneca**

List of poster presenters

P1	Mikkel Andreasen	Coarse-Grained System Builder (COBY) – An advanced tool for accurate creation of CG simulation systems
P2	Nils Anton Berglund	Accurate prediction of protein-ligand binding affinity and unbinding kinetics
P3	Israel Cabeza Vaca	Rapid Traversal of Vast Chemical Space using Machine Learning Guided Docking Screens
P4	Amanda Dyrholm	Exploring Insulin-Receptor Dynamics: Stability and Binding Mechanisms
P5	Catia Ferreira	The role of chirality and ring size in cyclic choline-analogues TMA lyase inhibitors
P6	Francesco Fontanive	P2Y12 Receptor Activation Mechanism
P7	Andrey Frolov	Peptide-Tools.com – Web server for calculating physicochemical properties of natural and modified peptides
P8	Gökçe Geylan	PepINVENT: generative peptide design beyond the natural amino acids
P9	Gian Marco Ghiandoni	On the Shoulders of Data: AstraZeneca's Predictive Insight Platform
P10	Enric Herrero	Ultra large virtual screening using 3D QM-derived hydrophobicity descriptors
P11	Lina Humbeck	Prediction of *in vivo* PK Profiles from Chemical Structures and in vitro ADME Experiments
P12	Iris Hättestrand	Evaluation of MSNovelist for structural assignment of chemicals and its applicability on environmental samples
P13	Yvonne Kreutzer	MS2Tox: Discovering Endocrine Disruptors via Molecular Networking
P14	Phong Lam	MolSanitizer: An Open-Source Python Pipeline for Preparing Large Databases of Small Molecules for Drug Discovery
P15	Renne Leini	Binding modes of the KRAS(G12C) inhibitors GDC-6036 and LY3537982 revealed by all atom molecular dynamics simulations
P16	Jacopo Manigrasso	Visualizing small molecules targeting pre-miR-377 by integrative structural biology
P17	Peter Maas	Rule-Based Retrosynthesis for Accessible Hits
P18	Sneha Menon	Characteriza*on of small-molecule mediated modula*on of an intrinsically disordered protein
P19	Lasse Messell Desdorf	Capturing Bicarbonate: The Transport Mechanism of Ncbe
P20	Laust Moesgaard	Rapid Virtual Screening of Ultra-Large Combinatorial Spaces with SpaceGA
P21	Jody Pacalon	Data-Driven Drug Design to Optimize Antipsychotics: Targeting Therapeutic Over Adverse Effects
P22	Tatu Pantsar	Observing RAS through the computational microscope – lessons learned from MD simulations
P23	Sonja Peter	Comparative Study of Allosteric GPCR Binding Sites and Their Ligandability Potential
P24	Dipti Potdar	Exploring a new alternative to polyethylene glycol (PEG) in drug delivery: Computational investigation of chiral poly(2-oxazolines).
P25	Jochen Sieg	Practical Molecular Property Prediction and Explainability with MolPipeline
P26	Christopher Southan	Using ontologies to make bioassay protocols machine readable
P27	Katarina Stanciakowa	Development-ability of Antibody Therapeutics
P28	Elise Tammekivi	Unravelling lignin structure with mass spectrometry and generative modelling
P29	Angeles Pulido	An Active Learning FEP workflow to identify the most promising molecules to make and test
P30	Alessandro Tibo	Does high quality data improve protein-conditioned ligand generation?
P31	Ovidiu Todoran	Structure-Based Development of FMN Riboswitch-Ligands: Paving the way for the Design of Novel, Selective, and Potent RNA-Based Antibiotics
P32	Susana Tomásio	Completing the medicinal chemists ´toolbox with cloud based deep learning models
P33	Ainoleena Turku	SpaceHASTEN: A structure-based virtual screening tool for nonenumerated virtual chemical libraries
P34	Jenny Viklund & Rajendra Kumar	Finding better drugs: optimizing multi-parameter optimization scores – and IT implementation of workflow
P35	Susanne Winiwarter	Augmentation of ACD/Labs pKa prediction algorithm with proprietary high-quality data
P36	Jiyeon Min	Thermodynamics of Arginine Interactions with Organic Phosphates
P37	Henrik Hupatz	Sustainable design of chemical reagents for the sensitive detection of pesticides using a machine learning workflow
P38	Farzaneh Jalalypour	Protein structure prediction, deep learning, alphafold, targeted protein degradation, generative AI

Poster Abstracts

P1 Coarse-Grained System Builder (COBY) – An advanced tool for accurate creation of CG simulation systems

Mikkel Dahl Andreasen*, Paulo Cesar Telles de Souza, Birgit Schiøtt, Lorena Zuzic

Department of Chemistry, Aarhus University, Denmark, mda@chem.au.dk

Aim:

To create a program capable of building simulation systems primarily in the Martini 3 force field. The program should ensure greater accuracy regarding lipid counts and should not contain arbitrary constraints surrounding the number of system components that can be included in a given system.

Methods:

The program (COBY) was written in Python 3 and has been compared to *insane*[1] which is the main program currently used for generating systems in the Martini 3 force field. The precision regarding lipid insertion and the speed of the program was compared to *insane*, while several systems have been created to showcase the various functionalities of the program.

Results and Discussion:

COBY was able to create systems without the same errors that are present when similar systems are created by *insane*, though COBY did prove to be slower than *insane* as a result of it using more complex algorithms. COBY is also able to create systems that are much more complex, (e.g., contains more system components) than *insane*. The heightened accuracy allows for more complex systems to be created without the need to worry about the accuracy of the individual components. The slower speed is deemed an acceptable trade-off for the higher accuracy and greater number of features.

Conclusions:

COBY has shown to be an improvement compared to *insane*, though with the caveat of being slower to run. COBY is available on GitHub at https://github.com/MikkelDA/COBY.

References:

1. Wassenaar, T. A., Ingólfsson, H. I., Böckmann, R. A., Tieleman, D. P., Marrink, S. J., 2015, Computational Lipidomics with insane: A Versatile Tool for Generating Custom Membranes for Molecular Simulations, J. Chem. Theory Comput. 11, 5, 2144–2155

P2

Accurate prediction of protein-ligand binding affinity and unbinding kinetics is a cornerstone of computational drug discovery, yet achieving high accuracy with physics-based methods often comes with significant computational expense.

Kvantify Koffee is a new approach that significantly accelerates the calculation of binding affinities and unbinding kinetics compared to molecular dynamics (MD) and alchemical methods. Koffee is purely physics-based, eliminating the need for machine learning models, pretraining, or reliance on datasets, providing reliable results that are generalizable across a diverse range of protein-ligand systems. Calculations can be performed in minutes and are easy to set up: atom typing and topology is fully automated and no pairwise comparison networks are required. This makes our methods accessible even to non-expert users.

Validation against a diverse set of experimentally characterized protein-ligand systems demonstrates that our methods provide binding affinity and unbinding kinetics predictions with high accuracy. Notably, our approach achieves these results with significantly reduced computational costs, enabling applications like hit expansion and large scale prioritization of molecules that would not be feasible using more computationally heavy methods.

In this presentation, we will outline the theoretical foundation of our method, and present case studies showcasing practical applications. Kvantify Koffee's fast physics-based approach offers a middle ground between between coarse, approximate methods such as Machine Learning and AI, and computationally demanding techniques like alchemical free energy calculations.

P3 Rapid Traversal of Vast Chemical Space using Machine Learning Guided Docking Screens

Andreas Luttens, Israel Cabeza de Vaca*, Leonard Sparring, Jose Brea, Anton Leandro Martínez, Nour Aldin Kahlous, Dmytro S. Radchenko, Yurii. S. Moroz, Maria Isabel Loza, Ulf Norinder, Jens Carlsson
*Department of Cell and Molecular Biology, Uppsala University, Sweden
israel.cabezadevaca@icm.uu.se

Aims: This study addresses the challenge of screening multi-billion-scale chemical libraries for drug discovery, which is infeasible for even the fastest structure-based docking methods. By combining machine learning and molecular docking, the goal was to develop a workflow enabling rapid virtual screening of ultralarge databases to efficiently identify potent compounds.
Methods: The workflow involved training a machine learning classification algorithm on docking scores obtained from an initial set of one million compounds screened against a target protein. A conformal prediction framework was used to prioritize compounds from multi-billion-scale libraries, significantly reducing the number of compounds required for molecular docking. Various machine learning algorithms, including gradient boosting, deep neural networks, and Transformer, were evaluated, with a particular emphasis on optimizing the balance between accuracy and computational efficiency. The approach was benchmarked across eight diverse protein targets for validation.
Results: The CatBoost classifier achieved the best balance of speed and accuracy and was integrated into the workflow for ultralarge library screening. Using this optimized approach, >90% of the top-scoring molecules were identified from a library of 235 million compounds by docking only 3-5% of the set. Application of the method to a 3.5 billion compound library reduced the computational cost of structure-based virtual screening by over 1000-fold. Experimental validation confirmed the identification of prospective dual-target ligands for the A2A adenosine and D2 dopamine receptors.
Conclusions: This study introduces a scalable and efficient machine learning-based workflow for virtual screening of multi-billion-scale chemical libraries. The freely available protocol enables rapid exploration of the largest commercial libraries, significantly reducing computational resources while maintaining high accuracy, thereby facilitating the identification of starting points for drug development. The code is publicly available at https://github.com/Carlssonlab/conformalpredictor

References:

Luttens, A., Cabeza de Vaca, I., Sparring, L., Brea, J., Martínez, A. L., Kahlous, N. A., Radchenko, D. S., Moroz, Y. S., Loza, M. I., Norinder, U., Carlsson, J., 2025. Rapid Traversal of Vast Chemical Space using Machine Learning Guided Docking Screens. Nat. Comput. Sci. *Just accepted*
Shafer G., Vovk V., 2007. A tutorial on conformal prediction. arXiv preprint arXiv, https://doi.org/10.48550/arXiv.0706.3188
Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., Gulin, A. 2013. CatBoost: unbiased boosting with categorical features. arXiv preprint arXiv, https://doi.org/10.48550/arXiv.1706.09516

P4 Exploring Insulin-Receptor Dynamics: Stability and Binding Mechanisms

Amanda Dyrholm Stange*, Lorena Zuzic, Birgit Schiøtt, Nils A. Berglund
Department of Chemistry, Aarhus University, Aarhus, Denmark
E-mail: ads@chem.au.dk

Aim:
To uncover the dynamic processes involved in insulin binding and the activation of the insulin receptor (IR) through molecular dynamics (MD) simulations. This study focuses on the structural stability and conformational changes of the IR-insulin complex, offering critical insights into receptor activation and therapeutic advancements.

Methods:
We performed MD simulations of IR-insulin complexes starting from experimentally resolved structures under physiological insulin concentrations. The simulations focused on the behaviour of insulin at hybrid binding sites, conformational transitions in the αCT region, and inter-domain stabilization. Computational analysis included monitoring changes in secondary structure, binding interactions, and receptor domain dynamics.

Results and Discussion:
Our results revealed that insulin binding at the hybrid sites initiates the opening of Site 1, a critical step in receptor activation. Insulin's presence at Site 1 promotes the extension of the α-helix in the αCT region and enhances stabilization between receptor domains. These findings support a novel "ladder-climbing" model of activation, where insulin transitions sequentially from Site 2 to Site 1, orchestrating a stepwise conformational change. This dynamic perspective contrasts with static structural data and highlights the cooperative nature of insulin binding and receptor activation.

Conclusions:
This study provides a comprehensive understanding of the molecular events driving IR activation. The "ladder-climbing" model offers valuable insights into insulin's role in modulating receptor function. These findings are critical for the rational design of therapeutics targeting IR, offering potential advancements in diabetes treatment by enabling more precise modulation of receptor activity.

References:

McKern, N.M., Lawrence, M.C., Streltsov, V.A., Lou, M.Z., Adams, T.E., Lovrecz, G.O., et al., 2006. Structure of the insulin receptor ectodomain reveals a folded-over conformation. Structure 14, 705–716.

Li, J., Park, J., Mayer, J.P., Webb, K.J., Uchikawa, E., Wu, J., Liu, S., Zhang, X., Stowell, M.H.B., Choi, E., Bai, X.C., 2022. Synergistic activation of the insulin receptor via two distinct sites. Nat. Struct. Mol. Biol. 29, 357–368.

P5 The role of chirality and ring size in cyclic choline-analogues TMA lyase inhibitors

Cátia A. Bonito[1, *], Giuseppina La Sala[1], Anders Gunnarsson[2], Mariell Pettersson[1] Jens Petersen[2], Margareta Ek[2], Tove Sjögren[2] and John Liddle[1]

[1]Medicinal Chemistry, eCVRM, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden
[2]Protein Science, Structure and Biophysics, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden.
*catia.ferreira3@astrazeneca.com

Aim: TMA lyase is a gut bacteria enzyme that cleaves choline to TMA (trimethylamine) and a promising pharmaceutical target for cardiovascular diseases. A drug discovery campaign led us to the identification of cyclic choline-analogues inhibitors endowed of nanomolar binding affinity. Herein, we aim at rationalizing SAR by using computational approaches.

Methods: Docking and molecular dynamics (MD) simulations were performed using internal X-ray structures of TMA lyase in complex with cyclic choline-analogues. The simulations were analysed using several criteria to identify stable and populated binding modes over the trajectories. Differential Scanning Fluorimetry (DSF) experiments were performed for measuring ligands affinity. As thermal shift changes (ΔTm) are not a direct measure of affinity, it is critical to establish confidence in the relationship between ΔTm and binding affinity, determined by ITC. A linear relationship between ΔTm and affinity was observed which suggested that ΔTm can be used as a surrogate to estimate affinity.

Results and discussion: The X-ray structures show that our cyclic choline-like inhibitors mimic the interactions of choline, including two key H-bonds to E491 and D216. Firstly, we focused on the role of chirality in binding affinity, by comparing the S-quinuclidinol (Kd=15 nM) and R-quinuclidinol (Kd= 483 nM) enantiomers. The MD simulations show that the S-enantiomer adopts one energetically favoured binding mode in all our replicas, while the R-enantiomer shows multiple binding modes. This high flexibility might explain its lower affinity. Similar behaviour was observed for the C1-bridge piperidine compounds, where the S-enantiomer (Kd=90 nM) was more stable over our trajectories than the R-enantiomer (Kd=1.3 µM). The second approach was to explore the role of ring size by comparing the S-quinuclidinol, S-C1-bridge piperidine and S-piperidine (NA). Unlike the S-quinuclidinol, the S-C1-bridge can transiently lose the H-bond with E491 leading to a second energetically less favourable binding mode, explaining its lower affinity. The removal of the bridge in the S-piperidine resulted in a loss of both H-bonds, leading to multiple binding modes which may explain its inactivity.

Conclusions: This work provides a rationale for the role of chirality and ring size in the binding affinity of cyclic choline-analogues inhibitors. MD simulations can be employed to rationalize the SAR of our compounds and can drive the rational design of new inhibitors.

References:

Bollenbach M. et al., 2020. Discovery of a Cyclic Choline Analog That Inhibits Anaerobic Choline Metabolism by Human Gut Bacteria. ACS Med. Chem. Lett. 11(10):1980-1985.

P6

P2Y12 RECEPTOR ACTIVATION MECHANISM

Francesco Fontanive*, Alejandro Giorgetti

Strada Le Grazie 15, Biotechnology Dept., University of Verona
francesco.fontanive@univr.it

AIM

The research aim is to investigate the activation process from a dynamics perspective, shedding light on the transition from the agonist-bound inactive[1] to the fully active state[2] of P2Y12 receptor.

METHODS

We conducted a 1 μs Molecular Dynamics simulation of the receptor bound to the full agonist 2MeSADP using GROMACS software. PCA was performed with gmx covar and gmx anaieg, and the Free energy landscape was generated with gmx sham using PC1 and PC2. The Gromos algorithm with a cut-off of 0.1 nm was used for the clustering scheme.

RESULTS AND DISCUSSION

We examined the interaction between the full agonist and the residues in the orthosteric binding site. We searched the trajectory for molecular switches[3] and identified three types indicating changes within the receptor – such as histidine rotamerization. We applied Principal Component Analysis to pinpoint the receptor's essential dynamics[4], highlighting the movement of helix 5 from a bent to a vertical position, and the slight outward movement of helix 6. These movements are crucial for creating a crevice for G-protein accommodation. Using PCA, we built a scatter plot representing all the receptor conformations during the dynamics. We also generated the Free Energy Landscape. To compare conformations from the energy minima with available crystallographic structures, we built sub-trajectories and applied a clustering scheme. Hence, we superimposed the cluster representatives onto the crystallographic structures. We found conformations aligned with both the agonist-bound inactive and fully active states, highlighting significant positional differences between helices 5 and 6.

CONCLUSIONS

Our finding supports the conformational selection theory, as our simulation started from the agonist-bound inactive state and identified the fully active conformation without simulating the G protein.

REFERENCES

1. Zhang, J., Zhang, K., Gao, Z.G., *et al*., 2014. Agonist-bound structure of the human P2Y12 receptor. Nature. 509, 119–122.

2. Li, B., Han, S., Wang, M., *et al*., 2023. Structural insights into signal transduction of the purinergic receptors P2Y1R and P2Y12R. Protein Cell. 14(5), 382-386.

3. Trzaskowski, B., Latek, D., Yuan, S., *et al*., 2012. Action of molecular switches in GPCRs--theoretical and experimental studies. Curr. Med. Chem. 19(8), 1090-109.

4. Amadei, A., Linssen, A.B., Berendsen, H.J., 1993. Essential dynamics of proteins. Proteins. 17(4), 412-25.

P7 Peptide-Tools.com – Web server for calculating physicochemical properties of natural and modified peptides

*Andrey I. Frolov**

andrey.frolov@astrazeneca.com. Medicinal Chemistry, Research and Early Development, Cardiovascular, Renal and Metabolism (CVRM), BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden

Peptides are increasingly significant in medicine and material sciences. Their structures are often artificially modified with non-canonical amino acids or chemical alterations to enhance their properties. Designing, purifying, and analyzing peptides often involves calculating their physicochemical properties. We present "Peptide-Tools," a free web server with a user-friendly interface for calculating the isoelectric point [1] and common formulation liabilities of natural and modified peptides. The server accommodates various input formats and considers the ionization of non-canonical amino acids and other molecular fragments. Additionally, it calculates the extinction coefficients of peptides at 205, 214, and 280 nm, limited to canonical amino acids. We believe this server will aid in peptide optimization and workups in chemistry and analytics laboratories.

[1] Frolov, A. I., Chankeshwara, S. V., Abdulkarim, Z., Ghiandoni, G. M., 2023. pIChemiSt—Free Tool for the Calculation of Isoelectric Points of Modified Peptides. J. Chem. Inf. Model., 63, 1, 187–196

P8 PepINVENT: generative peptide design beyond the natural amino acids

Gökçe Geylan[ab*], Jon Paul Janet[a], Alessandro Tibo[a], Jiazhen He[a], Atanas Patronov[c], Mikhail Kabeshov[a], Florian David[b], Werngard Czechtizky[d], Ola Engkvist[ae] and Leonardo De Maria[d]

a. Molecular AI, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden.
b. Division of Systems and Synthetic Biology, Department of Life Sciences, Chalmers University of Technology, Gothenburg, Sweden.
c. Quantitative Biology, Discovery Sciences, BioPharmaceuticals R&D AstraZeneca, Gothenburg, Sweden
d. Medicinal Chemistry, Research and Early Development, Respiratory & Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden
e. Department of Computer Science and Engineering, Chalmers University of Technology and University of Gothenburg, Gothenburg, Sweden
* Email: gokce.geylan@astrazeneca.com

Aim: Incorporating non-natural amino acids (NNAAs) has enabled designing peptides beyond the natural amino acids, with improved properties such as cell permeability or half-life. However, the expansion of the amino acids (AAs) to a vast small molecule-like chemical space with NNAAs makes exploring attempts exhaustive. Our tool, PepINVENT[1], aims to enable novel peptide designs with enhanced properties through reinforcement learning (RL)-driven generative model.

Methods: A transformer-based generative model was pretrained with semi-synthetic peptide data consisting of natural AAs and previously enumerated NNAAs[2]. The data contained various topologies, lengths, and common AA modifications. CHUCKLES was used to encode the granularity and chemistry of the peptides[3]. The model learned to propose a set of AAs given a query peptide with masked AA positions. Next, this model was coupled with RL and was used in various *in silico* experiments to steer the generation process in multi-parameter optimization (MPO) setting to desirable peptide properties.

Results & Discussion: The generative model successfully proposed unique, diverse, and novel peptides with valid chemical designs. PepINVENT, with the learnt chemical language, went beyond the combinatorial sequence space bounded by traditional AA libraries, proposing novel NNAAs. PepINVENT offers structural flexibility to explore diverse topologies or topology-constrained generation. In MPO, it suggested soluble and permeable cyclic peptides. To connect the generative designs with novel NNAAs and synthesis of these peptides, we have also created a tool to assess the synthetic feasibility of the NNAAs.

Conclusion: We introduce RL-driven generative AI tool, PepINVENT, as an extension to the small molecule molecular design platform, REINVENT[4]. The open-source tool generates diverse and novel peptides with improved properties, exploring chemical spaces beyond enumerated AAs and integrating NNAAs to advance peptide therapeutics.

References:

Geylan, G., Janet, J.P., Tibo, A., He, J., Patronov, A., Kabeshov, M., David, F., Czechtizky, W., Engkvist, O., De Maria, L., 2024. PepINVENT: Generative peptide design beyond the natural amino acids. https://doi.org/10.48550/arXiv.2409.14040

Amarasinghe, K.N., De Maria, L., Tyrchan, C., Eriksson, L.A., Sadowski, J., An, D., Petrović*, P., 2022. Virtual Screening Expands the Non-Natural Amino Acid Palette for Peptide Optimization. J Chem Inf Model 62, 2999–3007. https://doi.org/10.1021/ACS.JCIM.2C00193

Siani, M.A., Weininger, D., Blaney, J.M., 1994. CHUCKLES: A Method for Representing and Searching Peptide and Peptoid Sequences on Both Monomer and Atomic Levels. J Chem Inf Comput Sci 34, 588–593. https://doi.org/10.1021/CI00019A017

Loeffler, H.H., He, J., Tibo, A., Janet, J.P., Voronov, A., Mervin, L.H., Engkvist, O., 2024. Reinvent 4: Modern AI–driven generative molecule design. J Cheminform 16, 1–16. https://doi.org/10.1186/S13321-024-00812-5

P9 On the Shoulders of Data: AstraZeneca's Predictive Insight Platform

Gian Marco Ghiandoni[1*], Emma Evertsson[2*], David Riley[1], Prakash Chandra Rathi[1], Christian Tyrchan[2]

1. Augmented Design Test Make Analyse Platform, R&D IT, AstraZeneca, Cambridge, United Kingdom
2. Medicinal Chemistry, Respiratory and Immunology, BioPharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden
gianmarco.ghiandoni@astrazeneca.com
emma.evertsson@astrazeneca.com

Drug discovery is an information-based process which relies on iterating through the stages of Design, Make, Test, and Analyse, also known as DMTA, until drug candidates with suitable properties to be tested in clinical settings are found.[1] The integration of predictive methods within DMTA aims to accelerate the discovery process by steering the prioritisation of compounds that fit desired criteria.[2] In particular, using predictions is key in virtual de novo design, library screening, and when testing chemical hypotheses is expensive or time-consuming.

The Predictive Insight Platform (PIP) is a modelling platform built at AstraZeneca that hosts a large variety of predictive models and algorithms.[3] PIP currently serves machine learning models, molecular descriptors, structural filters, and synthetic accessibility scores.

PIP allows computational chemists to easily build, deploy, and automatically retrain models. These models are then used by all discovery scientists to generate predictions at scale across the tools used in different stages of the DMTA cycle. Models can be either defined as "general", hence created using data sets with wide applicability domains, or "local", developed using project-specific data.

Since its release in 2021, PIP has produced almost 10 billion calculations, recently surpassing 500 million in a single month. Herein, we describe the use of PIP at AstraZeneca, its evolution and architecture, and we speculate on the role of predictive modelling in the future of drug discovery.

References:

1. Wesolowski, S. S. & Brown, D. G. The Strategies and Politics of Successful Design, Make, Test, and Analyze (DMTA) Cycles in Lead Generation. 487–512 (2016) doi:10.1002/9783527677047.CH17.
2. Catacutan, D. B., Alexander, J., Arnold, A. & Stokes, J. M. Machine learning in preclinical drug discovery. *Nature Chemical Biology 2024 20:8* 20, 960–973 (2024).
3. Ghiandoni, G. M., Evertsson, E., Riley, D. J., Tyrchan, C. & Rathi, P. C. Augmenting DMTA using predictive AI modelling at AstraZeneca. *Drug Discov Today* 29, 103945 (2024).

P10 Ultra large virtual screening using 3D QM-derived hydrophobicity descriptors

Javier Vázquez, Brian Medel, Fernando Martin, Albert Herrero, Enric Herrero*
Pharmacelera, c/ Baldiri Reixac, 8, 08028 Barcelona, ES (enric.herrero@pharmacelera.com)

Aim:
Commercial libraries of chemical compounds have experienced an exponential growth in the last decade, offering a vast chemical space that comprises billions of molecules. These libraries represent an opportunity to find novel and more diverse hits but also pose a significant challenge due to the immense computational resources required for an accurate exploration, the greater need for storage, and concerns about synthetic accessibility. Exploiting synthon-based search methods could overcome these limitations by enabling an efficient navigation through this unexplored domain, ensuring both precision and synthesizability.

Methods:
Here, we present a 3D virtual screening method employing commercial building blocks and combinatorial chemistry to effectively navigate huge chemical spaces. The method can be used in ligand-based screenings (exaScreen) and in structure-based screenings. The combinatorial enumeration protocol allows to perform a 3D search and enhance the level of theory used in the analysis of the proposed molecules at a reasonable cost. In particular, the similarity measurement between compounds uses the hydrophobic molecular profile derived from the fractional hydrophobic contributions of the quantum mechanical version of the Miertus – Scrocco – Tomasi (MST) solvation model [1,2]. The resulting molecules are synthesizable and can be purchased directly from the library provider (e.g., Enamine Ltd, WuXi AppTec…).

Results and discussion:
The proposed methodology has been evaluated in a retrospective search of known antagonists/inhibitors. Including GPCR, transferases, nuclear receptors, and kinases as validation targets. Comparison of the proposed compounds to known actives shows the ability of the method to find structurally unrelated molecule classes and effectively mine huge chemical spaces in 3D. The study covers the advantages and limitations of using synthons instead of a fully enumerated library.

Conclusions:
Novel methods are required to mine ultra large chemical spaces in 3D with a reasonable computational cost and without sacrificing the quality of the obtained results. In this work we show that the usage of synthon-based methods provide a scalable and efficient way to address this problem, obtaining similar performance than brute force and with 6000x less computational cost.

References:
[1] Forti, F. et al., 2008. Extension of the MST Continuum Solvation Model to the RM1 Semiempirical Hamiltonian. J. Comput. Chem. 29, 578–587.
[2] Vazquez, J. et al. 2018. Development and Validation of Molecular Overlays Derived from Three-Dimensional Hydrophobic Similarity with PharmScreen. J. Chem. Inf. Model. 58, 1596–1609.

P11 Prediction of *in vivo* PK Profiles from Chemical Structures and *in vitro* ADME Experiments

Moritz Walter[1], Bettina Gerner[2], Hermann Rapp[2], Hannes Wendelin[3], Christofer Tautermann[1], Miha Skalic[1], Jens Markus Borghardt[2], Lina Humbeck[1*]

[1, 2] Boehringer Ingelheim Pharma GmbH & Co. KG, Birkendorfer Straße 65, Biberach 88397, Germany
[1] Medicinal Chemistry, email: lina.humbeck@boehringer-ingelheim.com
[2] Drug Discovery Sciences
[3] Boehringer Ingelheim RCV GmbH & Co KG, Cancer Research DDS ADME, Dr. Boehringer-Gasse 5-11, 1121 Wien

Aim: A successful drug needs to combine several properties including high potency and a good pharmacokinetic (PK) profile to sustain efficacious plasma concentration over time. To estimate required doses for preclinical animal efficacy models or for the clinics, *in vivo* PK studies need to be conducted. However, for compound selection, *in silico* prediction might be sufficient. Here, we present machine learning (ML) models to predict *in vivo* PK profiles for two PK species after i.v. and p.o. administration based on chemical structures solely or in combination with *in vitro* data.

Methods: Based on complete plasma concentration-time profiles of more than 7k compounds a multi-task machine learning model was trained to predict six compartmental model parameters (Vc, Vp, CL, Q, ka, F) and auxiliary tasks (1). Moreover, a model predicting ADME (absorption, distribution, metabolism, excretion) parameters was trained to allow for PK predictions before synthesis (2). The models were evaluated in a realistic setting using multiple temporal training-test splits.

Results and discussion: Model predictions were of good accuracy with e.g. 71% of purely *in silico* predicted rat p.o. profiles having a geometric mean fold error (GMFE) of <3-fold to the fitted profiles. These models enable the prioritization of new compound designs before synthesis. Predictions can be further improved when compounds have been tested in *in vitro* ADME assays (79% with GMFE <3-fold).

Conclusion: The predicted profiles can be used in combination with potency information to prioritize compounds in drug discovery projects with the lowest projected human efficacious doses.

References:

(1) Walter, M., Aljayyoussi, G., Gerner, B., Rapp, H., Tautermann, C. S., Balazki, P., Skalic, M., Borghardt, J. M., Humbeck, L., 2024. In silico PK predictions in Drug Discovery: benchmarking of Strategies to integrate machine learning with empiric and mechanistic PK modelling, bioRxiv
(2) Walter, M., Borghardt, J. M., Humbeck, L., Skalic, M., 2024. Multi-Task ADME/PK prediction at industrial scale: leveraging large and diverse experimental datasets. Mol. Inform. 43, e202400079

P12 Evaluation of MSNovelist for structural assignment of chemicals and its applicability on environmental samples

Iris Hättestrand*, Henrik Hupatz, Anneli Kruve

Department of Chemistry, Stockholm University, Svante Arrhenius väg 16, 114 18 Stockholm, Sweden. iris.hattestrand@su.se

Aim: *De novo* structure elucidation from mass spectra offers an alternative to library matching for identification but remains challenging, with a recent benchmark showing 0% accuracy for top-10 structures [1]. MSNovelist [2], which uses predicted fingerprints as input to a recurrent neural network model to generate structures, needs extensive evaluation beyond its original study. This project aims to assess its performance, especially for environmentally relevant chemical space.

Method: Literature data was curated from MassSpecGym and combined with MassBank EU, extending the dataset to other ionization techniques and adduct types. A pre-processing workflow for both databases has been developed to address environmentally relevant chemicals and pinpoint the limitations of MSNovelist. This workflow included standardization of SMILES, deduplication, and removal of chemicals containing As, Si, B, Se, or having precursor mass over 850 Da. Spectra for chemicals with the same SMILES were merged into one spectrum and analysed with SIRIUS+CSI:FingerID [3] and MSNovelist.

Results and discussion: The combined datasets contained 353,616 spectra. After pre-processing the final dataset consisted of a total of 300,199 unique spectra with 29,717 unique SMILES. The generated structures from MSNovelist were compared to the correct structures to calculate the accuracy for top-*k* structures to assess the performance of the model. Additionally, the accuracy of MSNovelist is compared to SIRIUS+CSI:FingerID database matching.

Conclusions: The emerging generative models for mass spectrometry hold promise to elucidate the unidentified chemicals in environmental samples, which today still exceeds 90% of detected chemicals. We develop a workflow to evaluate these models on literature data as well as experimental wastewater samples.

References

[1] Bushuiev, R., Bushuiev, A., de Jonge, N.F., Young, A., Kretschmer, F., Samusevich, R., et al., 2024. MassSpecGym: A benchmark for the discovery and identification of molecules. arXiv, 2410.23326 [q-bio.QM].

[2] Stravs, M.A., Dührkop, K., Böcker, S., Zamboni, N., 2022. MSNovelist: de novo structure generation from mass spectra. Nat Methods 19, 865–870.

[3] Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A.A., Melnik, A. V., Meusel, M., Dorrestein, P.C., Rousu, J., Böcker, S., 2019. SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. Nat Methods 16, 299–302.

P13 MS2Tox: Discovering Endocrine Disruptors via Molecular Networking

Yvonne Kreutzer*, Ida Rahu, Ulf Norinder, Anneli Kruve

Department of Materials and Environmental Chemistry, Stockholm University, Svante Arrhenius väg 16, 106 91 Stockholm, Sweden; yvonne.kreutzer@mmk.su.se

Aim: This study compares three *in silico* approaches to enhance prioritization of liquid chromatography high-resolution mass spectrometry features potentially associated with endocrine-disrupting activity in non-targeted screening.

Methods: Seven nuclear receptor endpoints from the Tox21 Data Challenge and MS^2 spectra from MassBank, MoNA, GNPS, and NIST23 were utilized to construct molecular networks (MN) as well as apply a conformal prediction approach (Arvidsson McShane et al., 2024). MS2Tox models (Rahu et al., 2024), leveraging structural fingerprints, were retrained on the same dataset. The performance of the approaches was compared utilizing a false positive rate at 90% recall ($FPR_{TPR=0.9}$) as a key metric. Lastly, the methods were applied to prioritize features corresponding to potentially active compounds in wastewater samples.

Results and discussion: MS2Tox models outperformed other approaches, achieving the lowest $FPR_{TPR=0.9}$ of 0.53 on aryl hydrocarbon receptor (AhR) activity. Results highlight that spectral similarity alone is insufficient for accurately assessing bioactivity. Prioritization of features based on AhR activity in wastewater samples helped to reduce the number of potentially relevant features. Prioritized features' structural analysis demonstrated that they contained scaffolds previously linked to the AhR activation. One candidate could be confirmed as zolpidem carboxylic acid.

Conclusions: Currently supervised models are more reliable compared to MNs for prioritizing features related to endocrine-disrupting activity in environmental samples.

References:

Arvidsson McShane, S., Norinder, U., Alvarsson, J., Ahlberg, E., Carlsson, L., Spjuth, O., 2024. CPSign: conformal prediction for cheminformatics modeling. Journal of Cheminformatics 16, 75. https://doi.org/10.1186/s13321-024-00870-9

Rahu, I., Kull, M., Kruve, A., 2024. Predicting the Activity of Unidentified Chemicals in Complementary Bioassays from the HRMS Data to Pinpoint Potential Endocrine Disruptors. J. Chem. Inf. Model. 64, 3093–3104. https://doi.org/10.1021/acs.jcim.3c02050

P14 MolSanitizer: An Open-Source Python Pipeline for Preparing Large Databases of Small Molecules for Drug Discovery

Phong Lam*, Szymon Pach, Israel Cabeza de Vaca, Ruth Brenk, Jens Carlsson
*Department of Cell and Molecular Biology, Uppsala University, Sweden.
phong.lam@icm.uu.se

Aims: Existing tools for accurate preparation of chemical databases often come with restrictive licenses or paywalls, while open-source alternatives usually underperform in critical tasks or require intensive manual intervention. Here, we introduce a rule-based, open-source Python pipeline for preparing small molecule databases. MolSanitizer offers versatile functionalities, including protonation, tautomerization, unwanted substructure filtering, and conformational sampling. It supports popular file formats (SMILES, SDF, Mol2, PDBQT, DB2), enabling its integration into diverse structure-based drug discovery workflows.
Methods: MolSanitizer is built on RDKit, comprising three core modules: The first module predicts tautomeric and protonation states of molecules across different pH values using rule-based SMARTS reactions. The second module performs filtering by using SMARTS matching to remove undesirable substructures. The third module focuses on conformational sampling, generating initial conformers via RDKit and exploring the conformational space stochastically using the Torsion Library v3.
Results: Preliminary tests on the set of DrugBank revealed a good agreement with the experimentally obtained data. Conformational sampling benchmarks using the Platinum Diverse Set achieve 98% reproduction of bioactive conformations (≤2Å RMSD). Enrichment analysis with DOCK3.8 on the DUDE-Z dataset showed superior performance of MolSanitizer-generated conformations, achieving a mean logAUC of 18.66 compared to 15.06 for the original ZINC-22 pipeline. Additionally, MolSanitizer is computationally efficient, running 15 times faster than RDKit ETKDGv3 and matching the speed of the original conformer generator pipeline provided for DOCK3.8.
Conclusions: MolSanitizer is an accurate, easy-to-use, and open-source Python pipeline for chemical database preparation, offering high performance and flexibility as a viable alternative to proprietary workflows.

References:

Friedrich, N.-O., Meyder, A., de Bruyn Kops, C., Sommer, K., Flachsenberg, F., Rarey, M., Kirchmair, J., 2017. High-Quality Dataset of Protein-Bound Ligand Conformations and Its Application to Benchmarking Conformer Ensemble Generators. J. Chem. Inf. Model. 57, 529–539.
Penner, P., Guba, W., Schmidt, R., Meyder, A., Stahl, M., Rarey, M., 2022. The Torsion Library: Semiautomated Improvement of Torsion Rules with SMARTScompare. J. Chem. Inf. Model. 62, 1644–1653.
Stein, R.M., Yang, Y., Balius, T.E., O'Meara, M.J., Lyu, J., Young, J., Tang, K., Shoichet, B.K., Irwin, J.J., 2021. Property-Unmatched Decoys in Docking Benchmarks. J. Chem. Inf. Model. 61, 699–714.
Wang, S., Witek, J., Landrum, G.A., Riniker, S., 2020. Improving Conformer Generation for Small Rings and Macrocycles Based on Distance Geometry and Experimental Torsional-Angle Preferences. J. Chem. Inf. Model. 60, 2044–2058.

P15 Binding modes of the KRAS(G12C) inhibitors GDC-6036 and LY3537982 revealed by all atom molecular dynamics simulations

Renne Leini[1]*, Jonas Kapp[2], Kari Kopra[3], Tatu Pantsar[1]
[1] School of Pharmacy, Faculty of Health Sciences, University of Eastern Finland, Kuopio, Finland. *Email: renne.leini@uef.fi
[2] Department of Biochemistry, University of Zurich, Zurich, Switzerland.
[3] Department of Chemistry, University of Turku, Turku, Finland.

Aim: Mutated RAS GTPases are emerging as key drug targets in oncology due to their frequent involvement in various cancers. For years, these proteins were considered undruggable until the discovery of the Switch-II pocket (SII-P) in 2013. This pocket is cryptic in the sense that it is only induced in the presence of a suitable ligand. Since then, multiple KRAS inhibitors have entered clinical trials with indications for different cancers. These inhibitors and their respective co-crystal structures have shed light on the function of the RAS SII-P. However, much remains unknown. Two orally bioavailable KRAS(G12C) inhibitors, divarasib (GDC-6036) and olomorasib (LY3537982), have shown promising results in the clinical trials, but their binding modes (co-crystal structures) remain undisclosed. Here, we employ microsecond-timescale molecular dynamics (MD) simulations to investigate their interactions and binding modes.

Methods: We conducted 200 μs of Desmond MD simulations (100 μs per inhibitor) with OPLS4 force field (NPT ensemble, 300K) to explore the binding modes of divarasib and olomorasib [**1**]. Additionally, biochemical assays, WaterMap simulations, and quantum mechanical (QM) calculations were performed to complement the MD simulations and assess binding affinity, characterize atropisomerism, and examine key binding interactions.

Results and discussion: Our simulations reveal putative binding modes of divarasib and olomorasib, highlighting key interactions with Glu62, Asp69, and Tyr96. We further analyze the Thr58-associated hydration site [**2**], demonstrating its high-energy in complex with either inhibitor. QM calculations provide insights into the atropisomer conversion barriers of both molecules. Biochemical assays confirm their strong binding affinity for KRAS(G12C) and reveal that co-mutations H95L (for divarasib) and Y96D (for both inhibitors) negatively impact binding, potentially affecting drug efficacy.

Conclusions: Our MD simulations reveal stable binding modes of divarasib and olomorasib, offering valuable insights for the further optimization of KRAS SII-P inhibitors.

References:
[**1**] Leini, R., Kapp, J.N., Kopra, K., Pantsar, T. Binding Modes of the KRAS(G12C) Inhibitors GDC-6036 and LY3537982 Revealed by All Atom Molecular Dynamics Simulations. *Manuscript in preparation.* (old version available at ChemRxiv) https://doi.org/10.26434/chemrxiv-2023-mq2p7
[**2**] Leini, R., Pantsar, T., 2023. In Silico Evaluation of the Thr58-Associated Conserved Water with KRAS Switch-II Pocket Binders. J. Chem. Inf. Model. 63 (5), 1490–1505. https://doi.org/10.1021/acs.jcim.2c01479

P16

Visualizing small molecules targeting pre-miR-377 by integrative structural biology

Jacopo Manigrasso[1,*], Emeline Mestdach[2], Celestine Chi[2], Giuseppina La Sala[1]

[1] Medicinal Chemistry, Research and Early Development, Cardiovascular, Renal and Metabolism (CVRM), BioPharmaceuticals R&D, AstraZeneca, 431 50 Mölndal, Sweden

[2] Biophysics, Discovery Sciences, BioPharmaceuticals R&D, AstraZeneca, 431 50 Mölndal, Sweden

[*]Presenting author, jacopo.manigrasso@astrazeneca.com

Targeting micro-RNA precursors (pre-miRNAs) with small molecules is a promising drug discovery strategy to address many therapeutic areas.[1] However, the conformational heterogeneity of pre-miRNA-ligand complexes has limited our structural understanding of these interactions, hampering the rational design of potent binders.[2]

Aim & Methods: Here, we have integrated NMR and enhanced sampling molecular simulations[3] to characterize the atomistic-resolution structural ensemble of the pre-miR-377 hairpin both in isolation and bound to the small molecule binder C1, previously reported to inhibit pre-miR377 maturation by Dicer.

Results and discussion: We show that the Dicer cleavage site of pre-miR377 exhibits high structural flexibility due to dynamic reorientation of the A12 bulge nucleotide, which can fold to form five distinct structural motifs. Additionally, our NMR-informed structural models suggest that pre-miR377's Dicer site and apical loop are dynamically linked, influencing major structural properties of the RNA, such as the hairpin bending. Most importantly, our simulations show that C1 stabilizes the A25•U46-A24 RNA triple at the Dicer site via conformational selection, locking pre-miR377 in a configuration that hampers Dicer's recognition due to steric clashes.

Conclusions: In summary, this work reports the in-solution structural ensemble of a pre-miRNA by integrative structural biology, both in isolation and in complex with a small molecule. These atomistic-resolution structural insights provide a foundation for structure-based drug design strategies targeting pre-miR377, an important therapeutic target for major cardiovascular disorders.

References:

1. Childs-Disney, J.L., et al., 2022, Targeting RNA structures with small molecules, Nat. Rev. Drug Discov., 21, 736-762
2. Manigrasso, J., Marcia, M., De Vivo, M., 2021, Computer-aided design of RNA-targeted small molecules: A growing need in drug discovery, Chem, 11, 2965-2988
3. Oxenfarth, A., et al., 2023, Integrated NMR/Molecular Dynamics Determination of the Ensemble Conformation of a Thermodynamically Stable CUUG RNA Tetraloop, J.A.C.S., 145, 16557-16572

P17 Rule-Based Retrosynthesis for Accessible Hits

An Open-Source Python Package

Oh-hyeon Choung[c], Esperanza Pear[c], Peter Maas[a*], Benedikt Wanner[b], Paula Nichols[c], Guillaume Coin[c], Samuele Bordi[b], Jeffrey Bode[b]

[a]eMolecules, 2712PB Zoetermeer, The Netherlands. e-mail:peter.maas@specs.net
[b]Synple Chem AG, 8310 Kemptthal, Switzerland.
[c]Laboratory of Organic Chemistry, Department of Chemistry and Applied Biosciences, ETH Zürich, 8093 Zürich, Switzerland

Aim:

Given the vast majority of new regions of virtual chemical space using multiple reaction step sequences and generative models, the challenge remains to determining the accessibility of these virtual compounds. To facilitate easy access to tractable hits, the development of an automated retrosynthesis tool is essential.

Methods:

An automated retrosynthesis tool has been developed using an open-source python package. The tool restricts retrosynthetic analysis to predefined, reliable reaction classes and commercially available building blocks that are predicted, with a high probability of success, to give the desired reaction outcome.

Results and discussion:

A case study has been conducted on a molecule library generated by REINVENT[3]. The tool provides rapidly accessible molecules with reaction schemes and commercially available building blocks.

Conclusion:

The developed tool provides rapid retro-synthesis of virtual library hits and allows for the identification of the necessary reaction protocols and building blocks for their preparation.

References:

1. Jiang, T.: Bordi, B., McMillan, A. E., Chen, K-Y., Saito, F., Nichols, P. L., Wanner, B. M. Bode, J. W., 2021. An integrated console for capsule-based, automated organic synthesis. Chem. Sci. 2, 6977-6982

2. McMillan, A. E., Wu. W. W. X; Nichols, P. L., Wanner, B. M., Bode, J. W., 2022. A vending machine for drug-like molecules – automated synthesis of virtual screening hits. Chem Sci. 13, 14292-14299.

3. Loeffler, H.H., He, J., Tibo, A. et al., 2024. Reinvent 4: Modern AI–driven generative molecule design. J. Cheminform. 16, 20.

P18 Characterization of small-molecule mediated modulation of an intrinsically disordered protein

Sneha Menon*†, Subinoy Adhikari‡, Jagannath Mondal‡
‡Tata Institute of Fundamental Research, Hyderabad, India
†University of Southern Denmark, Odense, Denmark
Email: sneha@sdu.dk

Aim: The misfolding and aggregation of Intrinsically Disordered Proteins (IDPs) such as α-synuclein is implicated in several neurodegenerative diseases(Wright and Dyson, 2015). Unlike folded proteins, IDPs exist as dynamic ensembles of multiple conformational states, making conventional structure-based drug design approaches ineffective(Metallo, 2010). This study aims to delineate the fuzzy ensemble of α-synuclein in the presence of a small-molecule drug fasudil and to provide insights into the ensemble modulatory effects mediated by the small molecule.

Methods: An integrative approach combining machine learning and Markov State Model (MSM) was employed to characterize the conformational ensembles of α-synuclein derived from atomistic molecular dynamics simulations(Robustelli et al., 2022). A β-variational autoencoder (β-VAE) was built for dimension reduction of a feature set of inter-residue pairwise distances from simulation trajectories of α-synuclein in neat water and in aqueous fasudil solution. The reduced 2-dimensional latent space obtained from β-VAE was then used to build MSMs enabling discretization of the ensemble into kinetically distinct metastable states.

Results and discussion: A general framework describes the effect of small-molecule binding to IDPs by modulation of the disordered ensemble, either increasing or decreasing its conformational entropy(Heller et al., 2015). Comparing the metastable states of αS with and without fasudil revealed that small-molecule binding navigates the structural landscape of the protein. Notably, more number of metastable states are populated in the presence of fasudil, indicating that small-molecule mediated conformational modulation led to entropy expansion. Additionally, the ensembles exhibited strong conformation dependence in residue-wise interactions with the small molecule.

Conclusions: The study presented sheds light on the intricate interplay between small molecules and IDPs, offering insights into entropic modulation and ensemble expansion as key biophysical mechanisms that may drive potential therapeutic strategies.

References:

Heller, G.T., Sormanni, P., Vendruscolo, M., 2015. Targeting disordered proteins with small molecules using entropy. Trends Biochem Sci 40, 491–496.

Metallo, S.J., 2010. Intrinsically disordered proteins are potential drug targets. Curr Opin Chem Biol 14, 481–488.

Robustelli, P., Ibanez-de-Opakua, A., Campbell-Bezat, C., Giordanetto, F., Becker, S., Zweckstetter, M., Pan, A.C., Shaw, D.E., 2022. Molecular basis of small-molecule binding to α-synuclein. J Am Chem Soc 144, 2501–2510.

Wright, P.E., Dyson, H.J., 2015. Intrinsically disordered proteins in cellular signalling and regulation. Nat Rev Mol Cell Biol 16, 18–29.

P19 Capturing Bicarbonate: The Transport Mechanism of Ncbe

Lasse Messell Desdorf*, Anna Louise Duncan, and Jeppe Prætorius

Department of Chemistry, Aarhus University, Aarhus, Denmark. lade@chem.au.dk

Aim: Ncbe (*SLC4A10*) is a membrane protein that performs electroneutral sodium dependent bicarbonate transport. Ncbe appears to be an ideal target for correcting the dysregulated secretion of cerebrospinal fluid accompanied by increased intracranial pressure which can cause severe brain dysfunction and even death. Current treatments either require surgery to drain the excess cerebrospinal fluid or the administration of medicine with low efficiency and patient compliance. Understanding the transport mechanism is key to reveal the involved components and their interactions with the protein. Furthermore, inhibiting the transport of sodium and bicarbonate through Ncbe could have potential therapeutic effects.

Methods: By utilising available structures from other *SLC4A* transporters in various conformations, in combination with lowering the MSA depth, we were able to generate multiple AlphaFold2 models representing distinct states of Ncbe during its transport process. Enhanced sampling methods allowed us to drive between the generated conformations revealing the underlying energy landscape of these transitions.

Results and discussion: Enhanced sampling was able to find the energetically stable conformation of an inhibitor, which did not correspond to the docking pose. The free energy of binding for the inhibitor was obtained from the PMF curves. The energy landscape of transition between outwards-open and inwards open was revealed, which highlighted two distinct energy minima alongside the path of least energy.

Conclusions: Docking and subsequent molecular dynamics simulations allowed us to capture the functions of these inhibitors, paving the way for novel potential interventions to relieve intercranial pressure.

References:

Damkier, H.H., Brown, P.D., Praetorius, J., 2013. Cerebrospinal fluid secretion by the choroid plexus. Physiol. Rev. 93, 1847-1892.

Praetorius, J., Damkier, H.H., 2017. Transport across the choroid plexus epithelium. Am. J. Physiol. Cell Physiol. 312, 673-686.

Freeman, W.D, 2015. Management of intracranial pressure. ANN Continuum: Lifelong Leaning in Neurology. 21, 1299-1323.

Jumper, J., Evans, R., Pritzel, A., et al, 2021. Highly accurate protein structure prediction with AlphaFold. Nature. 596, 583-589.

P20 Rapid Virtual Screening of Ultra-Large Combinatorial Spaces with SpaceGA

Laust Moesgaard[1]*, Jacob Kongsted[1]

[1]Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Odense, Denmark

moesgaard@sdu.dk

Aim:

The continued expansion of the purchasable small molecule chemical space has greatly enhanced the usefulness of virtual screening for hit discovery. However, traditional virtual screening tools require substantial computational resources to leverage this growth. This study aims to develop an easy-to-use computational tool, SpaceGA, that enables fast identification of high-scoring molecules with minimal computational resources.

Methods:

SpaceGA[1] combines the strong optimization capabilities of a genetic algorithm[2] with the fast similarity search tool SpaceLight[3] to perform virtual screens in fragment spaces without the requirement for full database enumeration. This setup allows SpaceGA to explore trillion-sized spaces and be customized with user-specified scoring functions, filters, and machine learning for further acceleration. The code for SpaceGA is available on GitHub (https://github.com/lmoesgaard/SpaceGA).

Results and discussions:

SpaceGA was tested for its ability to identify lead-like molecules with excellent docking scores within the Enamine REAL space and benchmarked against active learning (AL) and the raw genetic algorithm (GA). SpaceGA identified top scorers comparable to those found using AL and superior to those identified using GA, all while requiring only a few GB of disk space through on-the-fly lead-like filtering. Additionally, SpaceGA efficiently screened the trillion-sized eXplore space as quickly as the 100-times smaller Enamine REAL space.

Conclusions:

SpaceGA's robust and efficient capabilities allow it to navigate trillion-sized combinatorial spaces, simplifying virtual screening. Although there is potential for further optimization to achieve more diverse hits, SpaceGA already facilitates parallel exploration of multiple binding site conformations and co-optimization of key properties.

References:

[1] Moesgaard, L., & Kongsted, J. 2024. Introducing SpaceGA: A Search Tool to Accelerate Large Virtual Screenings of Combinatorial Libraries. Journal of Chemical Information and Modeling, 64(21), 8123-8130.

[2] Jensen, J. H. 2019. A graph-based genetic algorithm and generative model/Monte Carlo tree search for the exploration of chemical space. Chemical science, 10(12), 3567-3572.

[3] Bellmann, L., Penner, P., & Rarey, M. 2020. Topological similarity search in large combinatorial fragment spaces. Journal of Chemical Information and Modeling, 61(1), 238-251.

P21 Data-Driven Drug Design to Optimize Antipsychotics: Targeting Therapeutic Over Adverse Effects

Dr. Jody Pacalon*[1], Pr. David Gloriam[1]
Department of Drug Design and Pharmacology, University of Copenhagen, Copenhagen, Denmark, Email: jody.pacalon@sund.ku.dk

Aim

The project aims to design optimized ligands tailored to (1) selectivity, (2) mechanism and (3) pathway bias to address the limitations of current antipsychotics. We aim to reduce on-target and off-target side effects while improving efficacy in the treatment of positive and negative symptoms of schizophrenia, particularly in patients with concomitant substance use disorders who have special needs (e.g. avoiding aversive D2 antagonism). The first objective is to identify amino acid interaction hotspots that govern selectivity, pathway bias and receptor mechanisms.

Methods

We selected candidate antipsychotic ligands targeting dopamine (D2/D3) and serotonin (5-HT2A) receptors, with various selectivity, mechanism and bias characteristics. Molecular dynamics simulations were used to extract ligand-receptor interaction frequencies. Principal component analysis (PCA) was used to classify ligands by mechanism of activity, and correlations with binding affinity and selectivity were used to identify key residues that determine receptor activity and ligand specificity.

Results and Discussion

PCA on interaction frequencies effectively separated ligands by pharmacological activity, distinguishing agonists, partial agonists, and antagonists. This classification revealed key amino acid interactions driving ligand-specific signaling. Correlations with affinity and selectivity identified critical residues at D2/D3 receptors, offering targets to enhance selectivity and reduce off-target side effects. These findings highlight key hotspots for designing safer, more effective antipsychotics.

Conclusions

The integration of PCA-based interaction frequency analysis and correlation with ligand properties has allowed us to identify mechanistic hotspots that influence receptor activity, selectivity, and pathway bias. These insights will guide the design of novel ligands with tailored target profiles and reduced side effects, addressing major limitations in current antipsychotic therapies. Our computational pipeline, validated through experimental data, presents a significant step toward developing safer and more effective drugs for schizophrenia and its comorbidity with substance use disorder.

P22 Observing RAS through the computational microscope – lessons learned from MD simulations

Tatu Pantsar[1]*, Renne Leini[1], Milla Kurki[1], Antti Salo[1], Randa Mahran[2], Jonas Kapp[3], Kari Kopra[2]

[1] School of Pharmacy, Faculty of Health Sciences, University of Eastern Finland, Kuopio, Finland. *Email: tatu.pantsar@uef.fi

[2] Department of Chemistry, University of Turku, Turku, Finland.

[3] Department of Biochemistry, University of Zurich, Zurich, Switzerland.

Aim: Mutated small GTPase RAS is a key driver of oncogenesis across human cancers, and nearly 20% of all tumors harbor a RAS alteration. Targeting RAS has proven difficult. With two FDA-approved drugs, the KRAS switch-II pocket (SII-P) has offered one of the most invaluable tools to target this protein. To date, dozens of SII-P binder co-crystal structures are available, which have considerably improved our understanding of this pocket [**1**]. However, the SII-P is very enigmatic in nature as it is enclosed by the highly dynamic switch-II loop region [**2**]. Since the "frozen" structural data provides only a limited view of this pocket, we used MD simulations to gain further insights into SII-P under different conditions.

Methods: We applied microsecond timescale all-atom Desmond MD simulations (OPLS4) to study SII-P with and without ligands, resulting in a total of ~180 μs (adagrasib, sotorasib; K/NRAS, WT, G12C, G12C/Y96D) [**3**]; > 1 ms (MRAS/RRAS/RRAS2, WT & mutants) [**4**]. Biochemical and cellular experiments were applied to support our computational findings.

Results and discussion: First, we show that adagrasib is strictly KRAS- but not KRAS(G12C)-specific due to its strong and unreplaceable interaction with H95. Unlike adagrasib, sotorasib is less dependent on H95, making it a RAS isoform-agnostic compound. Also, our results aid in understanding the molecular mechanism behind the clinically observed drug resistance, associated especially with secondary mutations on KRAS H95 and Y96. Our MRAS/RRAS/ RRAS2 simulations suggest deviant behavior of their SII-Ps compared to KRAS, and the outlook of transferring the SII-P targeting strategy to other RAS GTPases is discussed.

Conclusions: Microsecond timescale MD simulations provide various useful insights to the enigmatic RAS SII-P, which can aid future drug discovery and design efforts towards small RAS GTPases.

References:

[**1**] Leini, R., Pantsar, T., 2023. In Silico Evaluation of the Thr58-Associated Conserved Water with KRAS Switch-II Pocket Binders. J. Chem. Inf. Model. 63 (5), 1490–1505. https://doi.org/10.1021/acs.jcim.2c01479

[**2**] Pantsar, T., 2020. The Current Understanding of KRAS Protein Structure and Dynamics. Comput. Struct. Biotechnol. J. 18, 189–198. https://doi.org/10.1016/j.csbj.2019.12.004

[**3**] Mahran, R., Kapp, J.N., Valtonen, S., Champagne, A., Ning, J., Gillette, W., Stephen, A.G., Hao, F., Plückthun, A., Härmä, H., Pantsar, T.*, Kopra, K.*, 2024. Beyond KRAS(G12C): Biochemical and Computational Characterization of Sotorasib and Adagrasib Binding Specificity and the Critical Role of H95 and Y96. ACS Chem. Biol. 19, 2152-2164 https://doi.org/10.1021/acschembio.4c00315

[**4**] Kurki, M., Salo, A., Kopra, K., Pantsar, T. G-Domain Dynamics of Small GTPases MRAS, RRAS and RRAS2 Revealed by All-Atom Molecular Dynamics Simulations: Focus on Switch-II Pocket Druggability. *Manuscript in preparation.*

P23 Comparative Study of Allosteric GPCR Binding Sites and Their Ligandability Potential

Sonja Peter*, Lydia Siragusa, Morgan Thomas, Tommaso Palomba, Simon Cross, Noel M. O'Boyle, Dávid Bajusz, György G. Ferenczy, György M. Keserű, Giovanni Bottegoni, Brian Bender, Ijen Chen, Chris De Graaf

*University of Urbino Carlo Bo, Italy & Nxera Pharma U.K., U.K., s.peter@campus.uniurb.it

Aim:

The steadily growing number of experimental G-protein-coupled receptor (GPCR) structures has revealed diverse locations of allosteric modulation, and yet few drugs target them. This gap highlights the need for a deeper understanding of allosteric modulation in GPCR drug discovery.

Method:

The current work introduces a systematic annotation scheme to structurally classify GPCR binding sites based on receptor class, transmembrane helix contacts, and, for membrane-facing sites, membrane sublocation.

Results and discussion:

This GPCR specific annotation scheme was applied to 107 GPCR structures bound by small molecules contributing to 24 distinct allosteric binding sites for comparative evaluation of three binding site detection methods (BioGPS, SiteMap, and FTMap). BioGPS identified the most in 22 of 24 sites. In addition, our property analysis showed that extrahelical allosteric ligands and binding sites represent a distinct chemical space characterized by shallow pockets with low volume, and the corresponding allosteric ligands showed an enrichment of halogens. One challenge regarding site prediction is the ligand shaping effect on the observed binding site, especially for extrahelical sites. Furthermore, we demonstrated that combining receptor and ligand similarity can be a viable method for ligandability assessment.

Conclusion:

To our knowledge, this is the first study presenting a binding site annotation scheme standardized for GPCRs, and it allows a comparison of allosteric binding sites across different receptors in an objective way. This work provides a framework for future GPCR binding site studies and highlights the potential of targeting allosteric sites.

References:

Peter et al., 2024. Comparative Study of Allosteric GPCR Binding Sites and Their Ligandability Potential. J. Chem. Inf. Model. 64, 8176-8192

P24 Exploring a new alternative to polyethylene glycol (PEG) in drug delivery: Computational investigation of chiral poly(2-oxazolines).

Dipti Potdar* [a], Josef Kehrein [a,b], Robert Luxenhofer [b], and Alex Bunker [a].

[a] Faculty of Pharmacy, University of Helsinki, Finland

[b] Department of Chemistry, University of Helsinki, Finland

email: dipti.potdar@helsinki.fi

Aim:

Poly(2-oxazolines), which mimic pseudo-polypeptides, exhibit exceptional biocompatibility, tunable properties, and high functionalization potential. These properties make them a promising alternative to polyethylene glycol (PEG) for drug delivery applications[1- 2]. In this study, we employed molecular dynamics (MD) simulations to investigate the structural dynamics of chiral poly(2-oxazoline) polymers with varying side-chain compositions in solvents of different polarity. Inspired by the experimental work of Yang M. et.al. [3], we aim to understand how these properties contribute to their effectiveness in drug formulation and delivery systems.

Methods:

All-atom MD simulations were conducted to analyze enantiopure and racemic conformers of poly(2-propyl-4-methyl-2-oxazoline) (pPrMeOx), poly(2-ethyl-4-ethyl-2-oxazoline) (pEtEtOx), and poly(2-butyl-4-ethyl-2-oxazoline) (pBuEtOx). The simulations were performed in methanol, ethanol, and butanol using the GROMACS simulation package with the OPLS-AA force field.

Results and Discussion:

Our findings reveal that both the sidechain composition and solvent environment significantly influence the structural and dynamic properties of poly(2-oxazolines) in enantiopure and racemic forms. The enantiopure forms exhibited greater stability and a higher degree of structural order whereas racemic conformers demonstrated higher flexibility compared to their enantiopure counterparts, suggesting improved solvation properties. These results provide insights into how the stereochemistry affects polymer-solvent interactions and polymer flexibility.

Conclusions:

The study underscores the critical role of chain length and enantiomeric composition in determining the flexibility and solubility of poly(2-oxazolines) in various solvents during drug formulation.

References:

[1] Hoogenboom R., 2009. Poly(2-oxazoline)s: A polymer class with numerous potential applications. Angewandte Chemie, 48, 43, 7978-94

[2] Bloksma M.M., Rogers S., Schubert U.S., Hoogenboom R., 2011. Main-chain chiral poly(2-oxazoline)s: Influence of alkyl side-chain on secondary structure formation in solution. J. Polym. Sci. A Polym. Chem., 49, 2790-2801

[3] Yang M., Haider M.S., Forster S., Chen H., Luxenhofer R., 2022. Synthesis and investigation of chiral poly(2,4-disubstituted-2-oxazoline)-based triblock copolymers, their self-assembly, and formulation with chiral and achiral drugs. Macromolecules, 2022, 55, 14, 6176-90.

P25 Practical Molecular Property Prediction and Explainability with MolPipeline

Jochen Sieg*, Christian W. Feldmann, Jennifer Hemmerich, Conrad Stork, Frederik Sandfort, Philipp Eiden, Miriam Mathea

BASF SE, Ludwigshafen, Germany. Email: jochen.sieg@basf.com

Aim:
Molecular Property Prediction (MPP) is crucial for BASF in crop protection and environmental science. Accurate prediction of molecular properties is essential for prioritizing compounds for testing and decisions making. Explainability methods (XAI) can help interpret why a model made a prediction, providing an important basis for building trust in models and informing the decision-making process. This poster gives an overview of interpreting machine learning models with XAI methods using the open-source MolPipeline package.
Methods:
MolPipeline extends scikit-learn's machine learning capabilities for chemical compound tasks by leveraging RDKit. We integrated XAI methods from the SHAP library with MolPipeline to enable easy and effective interpretabilty of property prediction models. The integration includes automatic extraction of chemical information from a model pipeline and communicating explanations through visualizing of important contributions on the molecular structure and other explanatory information.
Results and discussion:
We discuss the insights gained from integrating XAI methods into MolPipeline and the significance of visualizing and communicating XAI results for decision-making. We also examine the variations in SHAP interpretations from different machine learning models like random forest and neural networks using real-world project data.
Conclusions:
MolPipeline offers a scalable, interpretable Python package for molecular property prediction, improving decision-making in crop protection and environmental science.

References:
Muratov, Eugene N., et al. "QSAR without borders." Chemical Society Reviews 49.11 (2020): 3525-3564.
Harren, Tobias, et al. "Interpretation of Structure–Activity Relationships in Real-World Drug Design Data Sets Using Explainable Artificial Intelligence." Journal of chemical information and modeling 62.3 (2022): 447-462.
Sieg, Jochen, et al. "MolPipeline: A python package for processing molecules with RDKit in scikit-learn." Journal of chemical information and modeling 64.24 (2024): 9027-9033.
Pedregosa, Fabian, et al. "Scikit-learn: Machine learning in Python." the Journal of machine Learning research 12 (2011): 2825-2830.
Landrum, Gregory, et a. "RDKit: Open-source cheminformatics." https://www.rdkit.org
Lundberg, Scott M., and Lee, Su-In. "A Unified Approach to Interpreting Model Predictions." In Advances in Neural Information Processing Systems 30 (2017): 4765-4774.

P26 Using ontologies to make bioassay protocols machine readable

Christopher Southan *[1], Alex Clark[2] and Barry Bunin[2]
1)Deanery of Biomedical Sciences, University of Edinburgh, UK
2)Collaborative Drug Discovery, San Mateo, CA
cdsouthan@hotmail.com

Aim: Collaborative Drug Discovery (CDD Vault) has engaged in an open-source project, BioAssay Express, as part of the Pistoia DataFAIRy project.
Methods: we have developed templates to organize ontologies for curating data and marked up assays from PubChem and using an automated model plus expert curation workflow [1].
Results and discussion: Assays can be searched for analysis and model building from small molecule , peptide and protein screening results including bioactivity assays from medicinal chemistry projects. These tools have been integrated into our CDD product.
Conclusion: This work enables machine-readable data to be processed as standard practice.

References: [1] Balakrishnan R, *et. al*. 2024. Bioassay protocol metadata annotation: Proposed standards adoption. SLAS Discovery, Volume 29, Issue 8, 100188, PMID39427714.

P27 Development-ability of Antibody Therapeutics

Katarina Stanciakova[1], Usman Ghani[2], Adison Smith[2], Jesper Sørensen[2]
[1] OpenEye, Cadence Molecular Sciences, Ebertplatz 1, 50 668, Cologne, Germany
[2] OpenEye, Cadence Molecular Sciences, 9 Bisbee Ct. Santa Fe, NM 87508, USA
jesper@eyesopen.com

Aim

Developing an antibody from a hit or lead into a clinical candidate involves significant design efforts. The issues encountered in development deal less with target binding yet focus on whether the antibody has optimal biophysical properties making it amenable to be easily manufactured and a long stable shelf-life.

Methods

We have developed an antibody suite of tools in our cloud-native modeling platform Orion®[1], to assist in biophysical characterization and antibody property optimization using a combination of physics and AI.

Results and discussion

The process of modeling antibody structure on Orion begins with taking sequences from a selection campaign analyzed using AbXtract[2], other methods or e.g. internal or public databases. 3D structures are then generated with AI-driven structure predictors. From these structures a wide variety of physico-chemical properties are calculated to enable selection of those antibodies/sequences most likely to be successfully developed into human therapies. Structural diversity in the complementarity determining region (CDR) can be explored using knowledge-based loop modeling, while molecular dynamics and enhanced sampling forms are used to explore global conformational diversity, and the effect on physical properties descriptors. Structural families within computed ensembles can be identified by clustering using similarity based on a physically rigorous shape and chemical feature distribution of the CDRs. Individual sequences can be optimized by single- or multi-point mutations and new hypotheses submitted to physico-chemical property profiling and conformational exploration. In addition, using large language models trained on experimental data, we can generatively improve antibody development properties.

Conclusions

This talk will go over the key elements of the suite highlighting how antibodies can be optimized to avoid development problems.

1) Sørensen, J., Bannan, C.C., Calabrò, G., Jain, V., Ovanesyan, G., Smith, A., Zhang, S., Bayly, C.I., Darden, T.A., Geballe, M.T., LeBard, D.N., McGann, M., Moon, J.B., Muddana, H.S., Shewmaker, A., LaFon, J., Tolbert, R.W., Skillman, A.G. and Nicholls, A. 2024. Orion ® A Cloud-Native Molecular Design Platform . In Computational Drug Discovery (eds V. Poongavanam and V. Ramaswamy).
2) Erasmus, M.F., Ferrara, F., D'Angelo, S., Spector, L., Leal-Lopes, C., Teixeira, A.A., Sørensen, J., Nagpal, S., Perea-Schmittle, K., Choudhary, A., Honnen, W., Calianese, D., Antonio Rodriguez Carnero, L., Cocklin, S., Greiff, V., Pinter, A., Bradbury, A.R.M., 2023. Insights into next generation sequencing guided antibody selection strategies. Sci Rep. 13, 18370

P28 Unravelling lignin structure with mass spectrometry and generative modelling

Eliise Tammekivi*, Avinash R. Pai, Joseph S. M. Samec, Anneli Kruve
Svante Arrhenius väg 16C, 114 18 Stockholm, Sweden; eliise.tammekivi@mmk.su.se

Aim

Lignin is part of lignocellulosic biomass, the most abundant natural material on Earth and the primary constituent of wood, stalks, and energy crops. After isolating cellulose and hemicellulose in the paper and ethanol industries, the residual lignin is used mainly as an energy source to produce heat by burning. However, lignin is an aromatic macromolecule that could be valorised as biofuel or used as a starting material to produce novel biomaterials, which would improve sustainability. Developing efficient valorisation approaches benefits from establishing structure-function-performance correlations, where the structural characterisation of the final or intermediate products is essential. Therefore, in this project, we aim to shed light on the structure of depolymerised lignin by using LC-MS together with a lignin structural library compiled with generative machine learning models.

Methods

An analytical method was developed combining LC-MS with MS fragmentation and cyclic Ion Mobility (IM). For MS data processing, fingerprint prediction, and tentative identifications, mzMine and SIRIUS+CSI:FingerID[1] software were used. The generative model was based on the model developed by Eswaran *et. al*.[2]

Results and discussion

A depolymerised lignin sample obtained from the reductive catalytic fractionation (RCF) of birch sawdust lignin was analysed using the developed LC-IM-MS^2 method. The state-of-the-art data analysis with SIRIUS demonstrated the lack of lignin compounds, especially dimers, trimers and higher oligomers, in spectral and structural databases. As lignin is formed from the radical polymerisation of mainly three phenolic monomers through five main linkages, the structures of potential lignin oligomers could be predicted *in silico* by developing a generative model. Based on the generated structures, fingerprints denoting the structural fragments of the compound could be predicted. As SIRIUS+CSI:FingerID enables prediction of structural fingerprints from measured MS^2 spectrum, the fingerprints of the detected compound could be used to find a match within the generated lignin library.

Conclusions

The LC-IM-MS^2 analysis of the depolymerised lignin sample with SIRIUS demonstrated the current limitations for identifying lignin oligomers using databases such as PubChem. Therefore, the developed approach using generative modelling to propose lignin structures and fingerprint prediction to find matches is needed to improve the structural characterisation of lignin compounds.

References

[1]Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A. A., Melnik, A. V., Meusel, M., Dorrestein, P. C., Rousu, J., Böcker, S., 2019. SIRIUS4: a rapid tool for turning tandem mass spectra into metabolite structure information. Nat. Methods 16, 299-302.
[2]Eswaran, S.c.d., Subramaniam, S., Sanyal, U., Rallo, R., Zhang, X., 2022. Molecular structural dataset of lignin macromolecule elucidating experimental structural compositions. Sci. Data 9, 647.

P29 An Active Learning FEP workflow to identify the most promising molecules to make and test

Giovanna Tedesco*, Matthew Habgood, Venkata K. Ramaswamy, and Mark D. Mackey
Cresset, New Cambridge House, Bassingbourn Road, Litlington, Cambridgeshire, SG8 0SS, UK
Giovanna.tedesco@cresset-group.com

Aim

Bioisostere replacement is a powerful and popular tool used to optimize the potency and selectivity of candidate molecules in drug discovery. In this poster we present an active learning workflow to identify the strongest-binding bioisosteric replacements from a pool of several hundred candidates.

Methods

Using a dataset of human aldose reductase (ALR2) inhibitors, we assembled an active learning workflow to prioritize molecules from a pool of hundreds of bioisostere replacements generated by Cresset's Spark[1]. The workflow used two rigorous computational methods: 3D-Quantitative Structure Activity Relationships (3D-QSAR) using electrostatic and shape descriptors[2] and Free Energy Perturbation in Flare[1].

Results and Discussion

The ROC-AUC for selection of known actives in 80 top-ranked candidates improved to 0.88 from 0.64, and the top picks were enriched with highly potent ALR2 inhibitors, including the well-known clinical candidate Zopolrestat[3,4].

Conclusions

We will show that this workflow can rapidly identify the strongest-binding bioisosteric replacements with a relatively modest computational cost.

References

1. Flare™, Cresset®, Litlington, Cambridgeshire, UK, https://www.cresset-group.com/software/flare/; Spark™, https://www.cresset-group.com/software/spark/
2. Cheeseright, T., et al., Molecular field extrema as descriptors of biological activity: definition and validation. J Chem Inf Model. **2006**, 46, 2, 665-676. DOI: 10.1021/ci050357s.
3. Mylari, B.L., et al., Novel, potent aldose reductase inhibitors: 3,4-dihydro-4-oxo-3-[[5-(trifluoromethyl)-2-benzothiazolyl] methyl]-1-phthalazineacetic acid (zopolrestat) and congeners. J Med Chem., **1991**, 34, 1, 108-122. DOI: 10.1021/jm00105a018.
4. Mylari, B.L., et al., Potent, orally active aldose reductase inhibitors related to zopolrestat: surrogates for benzothiazole side chain. J Med Chem., **1992**, 35, 3, 457-465. DOI: 10.1021/jm00081a006.

P30 Does high quality data improve protein-conditioned ligand generation?

Alessandro Tibo*, Helen Lai, Ross Irwin, Jon Paul Janet, Ola Engkvist
Affiliations: (all) Molecular AI, Discovery Sciences, R&D, AstraZeneca, Gothenburg, Sweden – (Ross Irwin): Chalmers University of Technology Gothenburg, Sweden.

Address: Pepparedsleden 1, 431 83 Mölndal, Sweden – alessandro.tibo@astrazeneca.com

Aims: The field of structured-based drug discovery (SBDD) has been impacted by deep generative models like diffusion or flow matching models which have shown the potential to generate de novo ideas into protein pockets. As deep learning models typically require a large amount of data to be properly trained, one of the bottlenecks in SBDD is the quantity and quality of available datasets. Currently, PDBBind is one of the most widely used datasets for protein-conditioned ligand generation. Still, it suffers from inconsistencies such as not all the ligands being drug-like molecules and typically consists of unique ligand-protein pairs, with few congeneric series that would support learning structure-based drug design. This study aims to apply diffusion and flow matching models to SBDD tasks while leveraging a curated proprietary crystallography data from AstraZeneca.

Methods: The study focuses on diffusion and flow-matching models for pocket-conditioned molecule generation. Diffusion models gradually add noise to input data and learn to reverse this process. In contrast, flow-matching models rely on the principles of probability transport and optimal transport theory to transform simple distributions into complex ones. We compared the performances of both models on PDBBind and our internal dataset.

Results and conclusions: We assess the models' performance using several metrics including docking scores and number of interactions to the protein. The generated ligands are also evaluated in terms of chemical validity, novelty and quantitative estimate of drug likeness (QED) scores. We show the importance of good quality data.

References:

[1] Irwin, R., Tibo, A., Janet, J. P., Olsson, S. (2024). Efficient 3D Molecular Generation with Flow Matching and Scale Optimal Transport. arXiv preprint arXiv:2406.07266.
[2] Hoogeboom, E., Satorras, V. G., Vignac, C., Welling, M. (2022). Equivariant diffusion for molecule generation in 3d. In International conference on machine learning (pp. 8867-8887).

P31 Structure-Based Development of FMN Riboswitch-Ligands: Paving the way for the Design of Novel, Selective, and Potent RNA-Based Antibiotics

Ovidiu Todoran*[1], Ruth Brenk[1]
*Ovidiu.Todoran@uib.no
[1]Department of Biomedicine, Faculty of Medicine, University of Bergen, Bergen, Norway

Aim:
Antibiotic resistance is projected to become one of the biggest health crises in the near future, presenting an urgent need for antibiotics with novel mechanisms of action.[1] Riboswitches are promising antibacterial drug targets, as they regulate key processes in bacteria and present a novel mechanism of action that can be targeted by small molecules.[2] This project aims to establish a structure-based drug discovery workflow by integrating computational and experimental methodologies. We will apply this to help develop a new generation of selective and potent RNA-based antibiotics targeting the FMN Riboswitch.
Methods:
The first phase of the project focused on developing and validating computational models of the FMN Riboswitch. Using DOCK3.7, molecular docking was done to dock crystallized FMN ligands. RMSDs between the docked and experimental poses were calculated to evaluate the highest scoring ligand conformations. The ability of the models to distinguish active ligands from decoys was quantified by plotting ROC curves and calculating the AUCs.
Results:
Initial results show that multiple of our RNA models can redock FMN ligands with poses that are within 2.00 Å of the experimental structures. The ROC plots show that all models successfully distinguish actives from decoys, with AUC values ranging from 80% to 88%. A final FMN Riboswitch model was selected based on having a low RMSD and high AUC value.
Conclusions:
The initial results regarding the validation of computational RNA models were encouraging. In the next steps of the project, Molecular Dynamics simulations will be run to test the stability of RNA-ligand complexes and more accurately predict binding free energies. This will be followed by a prospective fragment-based virtual screening campaign to identify FMN Riboswitch ligands with novel scaffolds, which will be experimentally tested, optimized, and developed into selective and potent RNA-ligands.

References:
1. Church, N.A., McKillip, J.L, 2021. Antibiotic resistance crisis: challenges and imperatives. Biologia 76, 1535–1550.
2. Panchal, V., Brenk, R., 2021. Riboswitches as Drug Targets for Antibiotics. Antibiotics, 10(1), 45.

P32 Completing the medicinal chemists´toolbox with cloud based deep learning models

Susana Tomásio, PhD

Senior Application Scientist

stomasio@collaborativedrug.com

The field of drug discovery has witnessed remarkable advancements with the integration of artificial intelligence (AI) and deep learning (DL) approaches. Herein, we highlight an innovative cloud-based DL module (available within Collaborative Drug Discovery's CDD Vault®) based on technology developed by the Research Informatics Group at Collaborative Drug Discovery. We introduce the concept of **Chemically Rich Vectors** (CRVs), which serve as concise numerical summaries of chemical properties. CRVs allow for the reversible reconstruction of original molecules, enabling the encoding of comprehensive structural-activity information in a multi-dimensional numerical space. The introduction of CRVs greatly improved the autonomous robustness and interoperability of our DL model and can be additionally used to predict chemical properties, design new compounds, assess their retrosynthetic feasibility and query whether a commercial source is available. Based on these innovations, CDD's DL module empowers drug discovery research teams with a cutting-edge tool to perform similarity searches in the ChEMBL database, featuring ~2.5 million synthesizable compounds within the secure CDD platform.

P33 SpaceHASTEN: A structure-based virtual screening tool for nonenumerated virtual chemical libraries

Tuomo Kalliokoski, Ainoleena Turku* and Heikki Käsnänen
Orion Pharma, Orionintie 1A, 02101 Espoo, Finland
ainoleena.turku@orionpharma.com

Aim: In drug discovery, the relevant chemical space is vast rendering fully enumerated compound libraries unfeasible to work with. This has been addressed by developing nonenumerated virtual chemical spaces; in these, compounds are described as building blocks connected by rules. Prominent examples of such nonenumerated chemical spaces are those available from BioSolveIT (.space file format). Searching these space files has been possible with tools based on several two-dimensional (2D) similarity metrics, however, there was no software available that allows utilizing three-dimensional (3D) protein structures as the input. To address this, we developed SpaceHASTEN (Kalliokoski *et al.*, 2025), a hybrid ligand/structure-based virtual screening tool for nonenumerated chemical spaces.

Methods: We built SpaceHASTEN based on free software (chemprop) and commercial tools from BioSolveIT (SpaceLight, FTrees) and Schrödinger (LigPrep, Glide). The method requires a chemical space file, seed molecules, Glide docking grid, and Glide docking settings as the input. Prior to the screening, a set of random molecules (seeds) is acquired from the screened chemical space and docked to the target. A screening round starts by training a machine learning model based on the docking scores provided by the previous screening round (on the first round this means the docking scores of the seeds), followed by three subsequent similarity searching cycles. The screening round ends by docking the top million compounds found by the similarity searching cycles and ranked by their predicted docking scores. For a complete virtual screening campaign, we recommend two SpaceHASTEN screening rounds.

Results and discussion: Validation work was conducted with three public targets selected from the DUD-E data set. SpaceHASTEN was able to retrieve a substantial number of high-scoring compounds (virtual hits) from nonenumerated chemical spaces of billions of molecules, after docking just a few million compounds. The retrieved virtual hits were both diverse and novel, and the screening process was robust between different seed molecule sets.

Conclusions: SpaceHASTEN is a ready-to-use tool for both commercial and academic users. It shows promise for future virtual screening campaigns with the compound libraries, which are currently beyond the reach of reasonable enumeration efforts.

References:

Kalliokoski T., Turku A., Käsnänen H., 2025. SpaceHASTEN: A structure-based virtual screening tool for nonenumerated virtual chemical libraries. J. Chem. Inf. Model. 65 (1), 125–132.

P34 Finding better drugs: optimizing multi-parameter optimization scores - and IT implementation of workflow

Jenny Viklund, Cardiovascular, Renal and Metabolism, Biopharmaceuticals R&D

Kumar, Rajendra, R&D IT

The current advances in AI generative design and machine learning (ML) methods have set high expectations on accelerated delivery of drug candidates. When comparing the Design-Make-Test-Analysis workflows for the medicinal chemists within two large pharma companies (AstraZeneca and Sanofi) we see that the workflows and challenges in this process are similar: Predictive models, combined into multi-parameter optimization scores (MPOS) are crucial tools in this process, for filtering and prioritizing design sets for synthesis. The predictive power of the computational models is obviously important, for selecting the most promising compounds. However, we have both experienced that the art of optimizing the MPOS is often neglected.

We propose general and transparent guidelines for building optimized MPOS that can be easily implemented in any project. Advice is given for how to combine and set filtering thresholds and weights on each optimization parameter on series basis, based on the power of the corresponding predictive models, also considering the continuous learning of ML models updated with new data. The examples we show are mainly aiming at prioritizing virtual compounds for synthesis, but the methods can equally well be applied when prioritizing already synthesized compounds to *in vitro* or *in vivo* studies.

We will also show how fine-tuning the MPOS on series-basis matters significantly for the outcome when filtering or prioritizing both small and large sets of compounds.

Finally, we implemented this advice system as the user-friendly interactive dashboard in a Jupyter-Notebook and deployed in-house using Voilà server making it accessible to anyone within the AstraZeneca. The dashboard presents an excellent example where data-science and IT implementations are independent, which allows maintaining and updating the application more accessible.

P35 Augmentation of ACD/Labs pKa prediction algorithm with proprietary high-quality data

Susanne Winiwarter[1], Alexander Proskura[2], Eduard Kolovanov[2]

1 ... Drug Metabolism and Pharmacokinetics, Research and Early Development Cardiovascular, Renal and Metabolism (CVRM), BioPharmaceuticals R&D, AstraZeneca, Gothenburg, 431 83, Sweden; 2 ... ACD/Labs, 8 King Street East, Toronto, ON, M5C 1B5, Canada

Aim:

To increase the predictivity for in-house proprietary compounds by augmenting the existing pKa prediction algorithm with additional data.

Methods:

AstraZeneca internal pKa data was gathered and curated and shared with ACD/Labs (Advanced Chemistry Development, Inc.) pKa experts to augment the existing structure of Hammet equations for pKa predictions. The updated ACD/Labs pKa software was used to check the predictivity change on the dataset and on additional, newer compounds within the AstraZeneca data store.

Results and discussion:

The curated dataset will be characterised, and structural fragments not previously captured in the software exemplified. It was found that the new model version did enhance predictivity for compounds with specific structural features and maintained good predictivity for compounds that were well predicted with the previous version.

Conclusions:

Augmenting the ACD/Labs pKa software with in-house compounds improved the predictivity of the software on compounds not included in the original shared dataset.

References:

https://www.acdlabs.com/

P36 Thermodynamics of Arginine Interactions with Organic Phosphates

Jiyeon Min*, Madolyn Britt, Bernard R. Brooks, Sergei I. Sukharev, Jeffery B. Klauda

Presenter Information: [1]Biophysics Program, University of Maryland, College Park, MD, USA, [2]The Laboratory of Computational Biology, National Institutes of Health, Bethesda, MD, USA, Email: Jiyeonmin7@gmail.com

Aim: To investigate the thermodynamics of arginine-phosphate binding in biomolecular interactions, and to evaluate the ability of the CHARMM force field to accurately reproduce these interactions. The study aims to refine the force field parameters for enhanced modeling accuracy in nucleic acids, lipids, and membrane proteins.

Methods: Molecular dynamics (MD) simulations were conducted to analyze arginine-phosphate interactions, employing small molecules and peptides as models. Simplified methyl guanidine (MGUA) was used to isolate guanidinium interactions, while glycerol phosphate (Gly3P), glucose phosphate (Glu6P), dimethyl phosphate (DMP), and inositol triphosphate (IP3) were used as phosphate-containing partners. Parameters for these molecules were derived from the CHARMM36 and CGenFF force fields. Simulations were carried out using OpenMM and NAMD, employing enhanced sampling methods (Replica Exchange Umbrella Sampling and Adaptive Biasing Force) and clustering algorithms (DBSCAN) to extract binding conformations. Free energy calculations were validated against experimental data from isothermal titration calorimetry (ITC). Adjustments to the Lennard-Jones R_{min} parameter were made to refine the force field for MGUA-DMP interactions.

Results and Discussion:

- ITC and MD simulations showed strong agreement for MGUA–Gly3P (-3.30 vs. -4.08 kcal/mol), MGUA–Glu6P (-3.89 vs. -4.20 kcal/mol), and RGR–IP3 (-8.96 vs. -9.17 kcal/mol).
- A significant discrepancy was found for MGUA–DMP binding (-2.24 vs. -0.51 kcal/mol).
- Refinement of the CHARMM force field by reducing the Lennard-Jones R_{min} parameter from 3.55 to 3.405 Å improved the accuracy for the MGUA–DMP system, addressing the underestimation of the binding energy.

Conclusions: The CHARMM force field accurately modeled monoester interactions but required refinement for MGUA–DMP diester interactions involving nitrogen and oxygen atoms. Adjusting Lennard-Jones parameters improved accuracy, highlighting the value of combining experiments with computations.

P37 *Sustainable design of chemical reagents for the sensitive detection of pesticides using a machine learning workflow*

Henrik Hupatz*, Miguel Rivero Crespo, Berit Olofsson, Anneli Kruve

Stockholm University Center for Circular and Sustainable Systems (SUCCeSS), Stockholm University, 106 91 Stockholm, henrik.hupatz@su.se

Aim: We develop and evaluate a machine learning (ML) workflow for the sustainable and tailored design of derivatizing reagents for the fast and sensitive detection of polar pesticides, such as glyphosate. We demonstrate how renewable resources and reagent toxicity, in addition to desired analytical properties can be incorporated into a reinforcement learning algorithm.

Methods: Sensitivity in electrospray ionization mass spectrometry depends on how well the analyte is ionized, which can be empirically described with the ionization efficiency (log*IE*). We train ML models on a log*IE* dataset of 419 chemicals [1], using an 80/20 random split and evaluating different algorithms (XGBoost, random forest, support vector regressor) and molecular representations (Mordred and RDKit descriptors, MACCS keys, and ECFP6 fingerprints), as well as the Chemprop model. As a generative model, we employ REINVENT4 [2] with a prior trained on PubChem. The scoring function integrates polarity (log*P*), predicted log*IE* values, and synthetic accessibility score (SAS). Structural constraints are used to direct the model to generate reagents synthesizable from vanillin, a lignin depolymerization product, and containing one carboxylic acid group.

Results and discussion: Log*IE* prediction models were evaluated based on the RMSE on the test set. Many algorithm–representation combinations performed similarly, with RMSEs ranging from 0.77 to 0.82. Across different models, log*IE* values were more accurately predicted in the range of 1.5 to 4.5. During the reinforcement learning, 600,000 structures were visited, of which 575,663 were unique. 365,464 contained all desired structural elements, and 27,561 exhibited a high score (> 0.9), corresponding to log*IE* > 3.8 and SAS < 2.3. These high-scoring structures contained functional groups that contribute to a high log*IE*, such as tertiary amines, amides, or pyridines and are promising candidates for experimental validation. These candidates were filtered based on their predicted toxicity [3]. The final three reagents were selected based on structural fingerprint diversity from the remaining candidates.

Conclusions: Green chemistry principles can be incorporated into a generative modeling workflow and promising candidates for new derivatizing reagents for glyphosate were obtained.

References:

[1] Sepman, H., Malm, L., Peets, P., MacLeod, M., Martin, J., Breitholtz, M., Kruve, A., 2023. Bypassing the Identification: MS2Quant for Concentration Estimations of Chemicals Detected with Nontarget LC-HRMS from MS^2 Data. Anal. Chem. 95, 12329.

[2] Löffler, H. H., He, J., Tibo, A., Janet, J. P., Voronov, A., Mervin, L. H., Engkvist, O. 2024. Reinvent 4: Modern AI-driven generative molecule design. J. Cheminform. 16, 20.

[3] Söderberg, E., von Borries, K., Norinder, U., Petchey, M., Ranjani, G., Chavan, S., Holmquist, H., Johansson, M., Cotgreave I., Hayes, M. A., Fantke, P., Syrén, P.-O., 2024 Toward safer and more sustainable by design biocatalytic amide-bond coupling. Green Chem. 26, 11147

P38

Keywords: protein structure prediction, deep learning, alphafold, targeted protein degradation, generative AI

Farzaneh Jalalypour

Accurate prediction of protein-ligand and protein-protein interactions is essential for computational drug discovery, yet remains a significant challenge, particularly for complexes involving large, flexible ligands. In this study, we assess the capabilities of AlphaFold 3 (AF3) for modeling ligand-mediated ternary complexes, focusing on PROteolysis-TArgeting Chimeras (PROTACs). PROTACs facilitate targeted protein degradation by recruiting E3 ubiquitin ligases to a protein of interest, offering a promising strategy for previously undruggable intracellular targets. However, their size, flexibility, and cooperative binding requirements pose significant challenges for accurate computational modeling. To address these challenges, we leverage AF3's inference code, which enables direct ligand incorporation, to predict 48 PROTAC-related complexes from the Protein Data Bank. We systematically evaluate AF3's predictive accuracy using RMSD, pTM, and DockQ scores, demonstrating that when ligand information is provided, AF3 achieves high structural accuracy, even for post-2021 structures that were not included in its training set. Additionally, we explore alternative input strategies—comparing molecular string representations (SMILES) versus explicit ligand atom positions—to refine ligand placement and improve interaction predictions. By analyzing the relationships between ligand positioning, protein-ligand interactions, and structural accuracy metrics, we provide insights into key factors influencing AF3's performance in modeling PROTAC-mediated ternary complexes.

List of Participants

First name	Last name	Organisation	City
Mikkel	Andreasen	Aarhus University	Aarhus
Renate	Baker	Collaborative Drug Discovery	Burlingame
Flavio	Ballante	Karolinska Institutet	Stockholm
Nils Anton	Berglund	Kvantify	Copenhagen
Vladimir	Berishvili	AstraZeneca	Mölndal
Ruth	Brenk	University of Bergen	Bergen
Michael	Brunsteiner	RCPE GmbH	Graz
Israel	Cabeza de Vaca	Uppsala University	Uppsala
Lili	Cao	AstraZeneca	Mölndal
Jens	Carlsson	Uppsala University	Uppsala
Bruna	Clemens	Karolinska Institutet	Stockholm
Lee	Cronin	University of Glasgow	Glasgow
Charlotte	Deane	Oxford University	Oxford
Lasse Messell	Desdorf	Aarhus University	Århus
Elisabeth	Dietze	AstraZeneca	Mölndal
Amanda	Dyrholm Stange	Aarhus University	Aarhus
Ola	Engkvist	AstraZeneca, Chalmers	Mölndal
Emma	Evertsson	AstraZeneca	Mölndal
Cátia	Ferreira	AstraZeneca	Mölndal
Francesco	Fontanive	University of Verona	Verona
Andrey	Frolov	AstraZeneca	Mölndal
Simone	Fulle	NovoNordisk	Bagsvaerd
Gökce	Geylan	AstraZeneca	Mölndal
Gian Marco	Ghiandoni	AstraZeneca	Cambridge
David	Gloriam	Copenhagen University	Copenhagen
Anna-Lena	Gustavsson	Karolinska Institutet, SciLifeLab	Stockholm
Phil	Hastings	Cresset	Litlington
Jiazhen	He	AstraZeneca	Mölndal
Anders	Hogner	AstraZeneca	Mölndal
Lina	Humbeck	Boehringer-Ingelheim	Biberach
Henrik	Hupatz	Stockholm University	Stockholm
Iris	Hättestrand	Stockholms Universitet	Stockholm
Farzaneh	Jalalypour	Chalmers	Göteborg
Jon Paul	Janet	AstraZeneca	Mölndal
Anna	Karawajczyk	Selvita S.A.	Krakow
Henrik	Keränen	Lundbeck	Valby
Keishi	Kohara	AstraZeneca	Mölndal
Yvonne	Kreutzer	Stockholms Universitetet	Stockholm
Yashraj	Kulkarni	University of Copenhagen	Copenhagen
Giuseppina	La Sala	AstraZeneca	Mölndal
Phong	Lam	Uppsala University	Uppsala University
Per	Larsson	Uppsala University	Uppsala
Renne	Leini	University of Eastern Finland	Kuopio
Kresten	Lindorff-Larsen	Copenhagen University	Copenhagen
Peter	Maas	Specs, eMolecules	The Netherlands
Aniket	Magarkar	Boehringer Ingelheim Pharma	Biberach
Erik	Malmerberg	AstraZeneca	Mölndal
Marcus	Malo	Integrative Research Laboratories Sweden	Göteborg
Jacopo	Manigrasso	Astrazeneca	Mölndal
Minttu-Maria	Martikainen	Selvita S.A.	Krakow
Josip	Matic	RCPE GmbH	Graz
Sneha	Menon	University of Southern Denmark	Odense
Rocio	Mercado	Chalmers	Göteborg
Jiyeon	Min	University of Maryland	Maryland
Laust	Moesgaard	University of Southern Denmark	Odense

Eva	Nittinger	AstraZeneca	Mölndal
Jonas	Nyman	The Cambridge Crystallographic Data Centre	Karlskoga
Jody	Pacalon	University of Copenhagen	Copenhagen
Tatu	Pantsar	University of eastern Finland	Kuopio
Sonja	Peter	University of Urbino / Nxera Pharma UK	Urbino
Elżbieta	Plebanek	Selvita S.A.	Krakow
Vasanthanathan	Poongavanam	SciLifeLab/DDDP	Uppsala
Dipti	Potdar	University of Helsinki	Helsinki
Angeles	Pulido	Cresset	Litlington
Mathias	Rarey	Univeristy of Hamburg	Hamburg
Ekaterina	Ratkova	AstraZeneca	Mölndal
Susanne	Sauer	AstraZeneca	Mölndal
Magdalena	Scharf	Karolinska Institutet	Stockholm
Osheen	Sharma	Karolinska Instiute	Stockholm
Jochen	Sieg	BASF SE	Ludwigshafen
Lukas	Sigmund	AstraZeneca	Mölndal
Arman	Simonyan	University of Copenhagen	Copenhagen
Marie	Skepö	Lund University	Lund
Christian	Sköld	Uppsala University	Uppsala
Christopher	Southan	Univeristy of Edinburg	Edingburg
Katarina	Stanciakova	OpenEye, Cadence Molecular Sciences	Santa Fe
Julian	Streit	University of Copenhagen	Copenhagen
Peder	Svensson	Integrative Research Laboratories Sweden	Göteborg
Istvan	Szabo	InterAx Biotech	Schweiz
Eliise	Tammekivi	Stockholm University	Stockholm
Alessandro	Tibo	AstraZeneca	Mölndal
Ovidiu	Todoran	University of Bergen	Bergen
Susana	Tomásio	Collaborative Drug Discovery	Burlingame
Anna	Tomberg	AstraZeneca	Mölndal
Ainoleena	Turku	Orion Corporation Orion Pharma	Espoo
Jenny	Viklund	AstraZeneca	Mölndal
Guillem	Vila	Uppsala University	Uppsala
Vincent	Vivien	OpenEye, Cadence Molecular Sciences	Santa Fe
Fredrik	Wallner	Integrative Research Laboratories Sweden	Göteborg
Özge	Yoluk	AstraZeneca	Mölndal

Scientific organizing committee

Jens Carlsson, Uppsala University, SE
Ruth Brenk, University of Bergen, NO
Ola Engkvist, AstraZeneca, Chalmers, SE
Emma Evertsson, AstraZeneca, SE
Simone Fulle, NovoNordisk, DK
Anna-Lena Gustavsson, Karolinska Institutet, SciLifeLab, SE
The Section for Pharmacokinetics and Drug Metabolism at Apotekarsocieteten