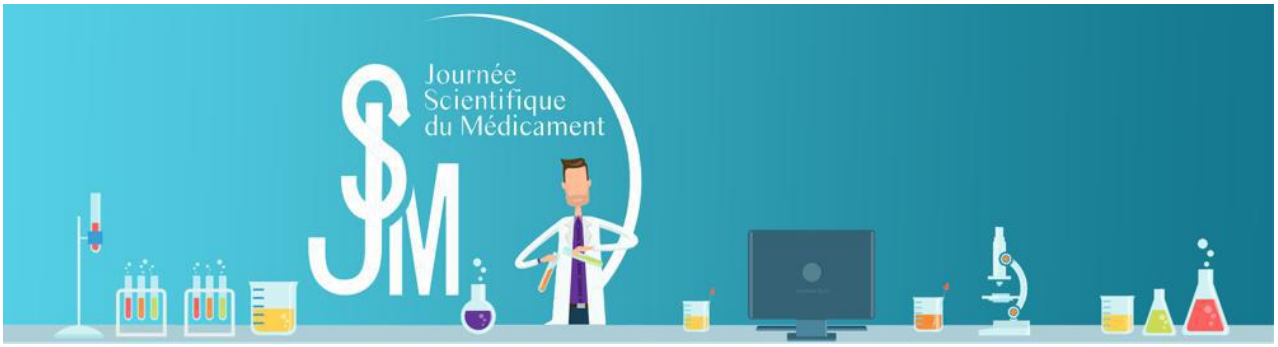


June 12, 2025, UGA

"Artificial Intelligence in Pharmaceutical Science"

Scientific committee:

Brigitte Bidegaray (LJK), Sergei Grudinin (LJK), Kliment Olechnovič (LJK), Basile Peres (DPM), Corinne Ravelet (DPM), Nelle Varoquaux (TIMC), Yung-Sing Wong (DPM)

Organisation committee :

Pascale Battistoni (DPM), Isabelle Baussanne (DPM), Sergei Grudinin (LJK), Meven Jobic (DPM), Matis Moretti (DPM), Khan-Chi Nguyen-Pham (DPM), Corinne Ravelet (DPM), Jean-Marc Sache (DPM), Yung-Sing Wong (DPM)

The **Department of Molecular Pharmacochemistry** (DPM), UMR 5063, a public laboratory depending on the **CNRS** and the **University of Grenoble Alpes** (UGA) in the frame of the **UFR of Pharmacy**, organizes multidisciplinary scientific days dedicated to the drug.

The aim of the "**Journée Scientifique du Médicament**" (JSM) is to bring together specialists in various fields: synthesis of drug candidates, structure-activity relationship studies, modelling, analytical chemistry, artificial intelligence (AI), pharmacochemistry, structural biology, imaging, vectorization, controlled release, toxicology… around a theme related to the drug.

This year, the 14th Scientific Day on Drugs will take place on **June 12**, **2025** in Grenoble (UGA) on the theme "**Artificial Intelligence in Pharmaceutical Science**".

This event is also co-organized by the **Jean Kunstmann Laboratory** (LJK)

JSM 2025 covers all aspects of Artificial Intelligence (Nobel Prize in Physics 2024) in conjunction with advances in Drug Design using Alpha fold (Nobel Prize in Chemistry 2024) to predict protein structures, their interaction with drugs and up to database processing.

JSM's History

The first Journées Scientifiques du Médicament were held at the UFR de Pharmacie et de Médecine, on the themes of "Vectorisation and genetic material transfer" (2011), "Biomolecules, aging and neurodegenerative diseases" (2012), "Natural substances of therapeutic interest" (2013), "Biotechnology for health" (2014), "Drug design, membrane proteins and membranes" (2015), "Therapeutic perspectives for rare and/or neglected diseases" (2016), "Epigenetic : toward new therapeutic targets" (2017), 'Ethnopharmacology: from traditional practices to today's medecines' (2018) and 'Département de Pharmacochimie Moléculaire 20th anniversary' (2019), 'Vectorization and Theranostics' (2021), 'Challenges in analysis and diagnostics' (2022), 'Click chemistry and light in Chemical Biology' (2023), 'Drugs by 3D printing' (2024) brought together a large number of participants around internationally renowned speakers.

Programme

(June 12, 2025, Amphi Ouest, UGA Campus)

08h30 - 09h00: Reception of participants / coffee

09h00 - 09h10: Introduction by **Philippe Roux**, Vice Président Recherche UGA and **Yung-Sing Wong**, DPM director

Chairmen: Yung-Sing Wong and Khan-Chi Nguyen-Pham

09h10 – 09h50: Conference 1: **Bruno Correia** (LPDA, EPFL, Lausanne)
"*Exploring the functional landscape of proteins by computational design*"

09h50 – 10h05: Oral Communication (OC) 1: **Anna Song** (Owkin, Paris)
"A glimpse of some pocket comparison approaches"

10h05 – 10h20: OC2: **The-Chuong Trinh** (LRB, UGA/Inserm, Grenoble)
"Ligand-based drug discovery leveraging state-of-the-art machine learning methodologies exemplified by CDR1 inhibitor prediction"

Chairmen: Elodie Laine and Meven Jobic

10h20 – 10h45: Coffee / Tea break, poster Sessions

10h45 – 11h25: Conference 2: **Philippe Nghe** (ESPCI, Paris)
"Comparing the power of generative models of RNA with an experimental system"

11h25 – 11h40: OC3: **John Rendu** (CHUGA, GIN, UGA, Grenoble)
"An example to resolve genetic dealocks: from RYR1 structural modeling to AI-based variant classification"

11h40 – 11h55: OC4: **Sandrine Py** (DCM, UGA/CNRS, Grenoble)
"C-Branched iminosugars as selective pharmacological chaperones of lysosomal alpha-glucosidase for the treatment of Pompe disease"

12h00 – 13h30: Lunch Buffet / Poster Sessions

Chairmen: Sergei Grudinin

13h30 – 14h10: Conference 3: **Elodie Laine** (IBPS, Paris Sorbonne Université)
"Computational Prioritisation of Genetic Variants and Therapeutic Targets: Linking Protein Sequences, Dynamics, and Functions"

14h10 – 14h50: Conference 4: **Maria Kadukova** (Astex Pharm, Cambridge)
"Can classical template-based protein-ligand docking challenge modern co-folding methods? An industrial perspective"

14h50 – 15h10: Coffee /Tea break / Poster Sessions

Chairmen: Kliment Olechnovic and Matis Moretti

15h10 – 15h25: OC5: **Marc Jamin** (IBS, UGA, Grenoble)
"AI-driven Design of human Respiratory Syncytial Virus Miniprotein Inhibitors"

15h25 – 15h40: OC6: **Christophe Battail** (IRIG, INSERM-CEA-UGA, Grenoble)
"Integrative Pharmacogenomic Models and Signatures for Patient Therapeutic Response Stratification in Metastatic Kidney Cancer"

15h40 – 16h20: Conference 5: **Julien Volle** (SynapCell, Saint-Ismier)
"Drug Discovery and Preclinical EEG: A promising playground for AI"

16h20 – 16h25: Closing ceremony

Thank you to our Sponsors

Conference abstracts

EXPLORING THE FUNCTIONAL LANDSCAPE OF PROTEINS BY COMPUTATIONAL DESIGN

Bruno CORREIA[1]

[1] *Laboratory of Protein Design and Immunoengeineering, EPFL, Lausanne*
bruno.correia@epfl.ch

The machine learning tool box has revolutionized our ability to design novel molecular entities (*e.g.* proteins) well beyond what the natural repertoire has explored. Despite the incredible advances, the *de novo* generation of functional molecules in biological concepts remains an incredible challenge.

In this talk I will present some of the efforts in our group in designing both proteins and small molecules. Particularly emphasizing different modalities of molecular representation and the interplay with generative ML to facilitate the exploration of unimaginably large spaces of possibilities. Importantly, many ML-based approaches for molecular design fall short in terms of generalization and sampling off the learned distribution, I will present some ideas on how representation can help to overcome some of these limitations.

Finally, I will present some of the approaches developed in our group to embed function into the designer proteins and how we are using these components in cellular systems to control the output of these complex biological devices.

DESIGNING FUNCITONAL RNAS WITH MACHINE LEARNING: THE CASE STDY OF GROUP I

Philippe NGHE[1]

[1] *UMR Chemistry Biology Innovation 8231 CNRS-ESPCI*
Philippe.nghe@espci.fr

Keywords: RNA; Machine Learning; Molecular design; high-throughput screening; Structure prediction

Generative models trained on evolutionary data offer a route to design functional RNA. We developed methods to integrate probabilistic modeling, high-throughput experimental validation, and iterative feedback to explore and exploit the functional sequence space of structured RNAs. Group I introns serve here as our workhorse, motivated by origin of life question of self-reproduction. However, these methods are expected to apply this specific RNA family.

We compared diverse generative methods based on machine learning (Direct Coupling Analysis, Variational AutoEncoders), compared their generative power with more naïve methods, or on the opposite, with hybrid methods including structure prediction. These methods successfully generate functional variants across a broad mutational landscape, but with varying degree of success in terms of accessible diversity, which we tested experimentally using a high-throughput screen for Group I intron activity.

We further show that reintegrating experimental feedback into the training of generative models demonstrates a substantial increase in the success rate of functional RNA design, highlighting our ability to produce artificial data at scale for machine learning algorithms.

References:

1. Calvanese, F., Lambert, C. N., Nghe, P., Zamponi, F., & Weigt, M. (2024). Towards parsimonious generative modeling of RNA families. Nucleic Acids Research, 52(10), 5465-5477.2.
2. Lambert, C. N., Opuu, V., Calvanese, F., Zamponi, F., Hayden, E., Weigt, M., Smerlak, M. & Nghe, P. (2024). Expanding the space of self-reproducing ribozymes using probabilistic generative models. bioRxiv, 2024-07.
3. Calvanese, F., Peinetti, G., Pavlinova, P., Nghe, P., & Weigt, M. (2025). Integrating experimental feedback improves generative models for biological sequences. arXiv preprint arXiv:2504.01593.

PRIORITISATION OF GENETIC VARIANTS AND THERAPEUTIC TARGETS: LINKING PROTEIN SEQUENCES, DYNAMICS, AND FUNCTIONS

Elodie, LAINE[1,2]

[1] *Sorbonne Université, CNRS, IBPS, Laboratory of Computational and Quantitative Biology (LCQB), UMR 7238, Paris, 75005, France*

[2] *Institut universitaire de France (IUF)*

elodie.laine@sorbonne-universite.fr

Keywords: protein mutation, protein motion, variant effect prediction, deep learning, protein evolution

I will present our recent work in which we leveraged protein language models to address questions such as how proteins move and deform to ensure their functions. Or what is the structural and functional impact of sequence variations? I will discuss the trade-off between model complexity and predictive performance, as well as the contribution of 3D structural information with respect to across-species sequence variations. I will show how our methods can help identify new therapeutic targets and guide CRISPR genome editing experiments.

References:

1. Abakarova M., M.I. Freiberger, A. Lierhmann, M. Rera, and E. Laine (2025). Proteome-wide Prediction of the Functional Impact of Missense Variants with ProteoCast. doi: 10.1101/2025.02.09.637326.
2. Lombard V., S. Grudinin, and E. Laine (2025). PETIMOT: A Novel Framework for Inferring Protein Motions from Sparse Data Using SE (3)-Equivariant Graph Neural Networks. *Learning Meaningful Representations of Life workshop at the thirteenth International Conference on Learning Representations (ICLR), Singapore, SG*.
3. Lombard V., D. Timsit, S. Grudinin, and E. Laine (2024). SeaMoon: Prediction of molecular motions based on language models. doi: 10.1101/2024.09.23.614585.
4. Marquet C., J. Schlensok, M. Abakarova, B. Rost, and E. Laine (2024). Expert-guided protein Language Models enable accurate and blazingly fast fitness prediction. *Bioinformatics* 40:btae621 doi: 10.1093/bioinformatics/btae621.
5. Lombard V., S. Grudinin, and E. Laine (2024). Explaining Conformational Diversity in Protein Families through Molecular Motions. *Sci. Data* 11:752 doi: 10.1038/s41597-024-03524-5.
6. Abakarova M., C. Marquet, M. Rera, B. Rost, and E. Laine (2023). Alignment-based protein mutational landscape prediction: doing more with less. *Gen. Biol. Evol.* 15:evad201 . doi: 10.1101/2022.12.13.520259.
7. Laine E., Y. Karami and A. Carbone. (2019) GEMME: a simple and fast global epistatic model predicting mutational effects. *Mol Biol Evol.* 36:2604–2619.

CAN CLASSICAL TEMPLATE-BASED PROTEIN-LIGAND DOCKING CHALLENGE MODERN CO-FOLDING METHODS? AN INDUSTRIAL PERSPECTIVE

Maria Kadukova, Lucian Chan, Chris Murray, Carl Poelking, Marcel Verdonk

Astex Pharmaceuticals, Cambridge, UK
maria.kadukova@astx.com

Keywords: fragment-based drug discovery; deep learning; co-folding; protein-ligand docking; AlphaFold

Methods that predict the structure of protein-ligand complexes are a cornerstone of computer-aided drug discovery. Going beyond traditional docking approaches in scope and complexity, deep-learning-driven co-folding techniques have made a significant advance in structure prediction. Here we outline how we use machine learning and docking at Astex and focus on some insights from benchmarking and validating co-folding techniques on our internal dataset. We assess the out-of-sample performance of co-folding models and explore the potential of co-folding for virtual fragment screening as a particularly challenging application area.

Our benchmark highlights that commonly used baselines such as free docking are overly pessimistic in a real-world structure-based setting; robust and efficient template-based docking remains underutilized and sets a higher bar for co-folding methods. Inspired by template-based techniques, we propose strategies to inject prior structural knowledge into co-folding models and introduce post-processing approaches that correct common structural artifacts affecting ligand poses.

DRUG DISCOVERY AND PRECLINICAL EEG: A PROMISING PLAYGROUND FOR AI

<<One blank>>

Julien, VOLLE,[1]

<<One blank>>

[1] *SynapCell SAS, 38330 Saint-Ismier, France*

jvolle@synapcell.fr

<<One blank>>

<<One blank>>

<<One blank>>

Keywords: Electrophysiology, Machine learning, Drug discovery, Neuroscience, Artificial intelligence

<<One blank>>

<<One blank>>

Drug discovery, particularly for central nervous system (CNS) disorders, faces high attrition rates and escalating costs [1]. Preclinical electroencephalography (EEG) offers a rich source of neurophysiological data crucial for evaluating drug candidates. Each class of chemicals elicit different EEG fingerprints, suggesting that EEG yields pharmaco-dynamic signatures specific to pharmacological action and can be used to enable classification based on the effects of the EEG [2]. However, the complexity and volume of EEG signals often limit traditional analytical approaches. Artificial Intelligence (AI), especially machine learning (ML), is emerging as a powerful catalyst in this domain. AI algorithms can help to discern subtle patterns, identify novel translational biomarkers, and predict drug efficacy (e.g., seizure occurrence) with enhanced precision [3]. This synergy holds the potential to accelerate drug candidate selection, refining dose-response understanding, and ultimately de-risking clinical translation. The integration of AI with preclinical EEG thus holds considerable potential to transform early-stage CNS drug development, making it faster and more predictive.

References:

1. Vamathevan, J., Clark, D., Czodrowski, P. et al. Applications of machine learning in drug discovery and development. Nat Rev Drug Discov 2019, 463–477.
2. Drinkenburg, Wilhelmus H I M et al. “Pharmaco-EEG Studies in Animals: A History-Based Introduction to Contemporary Translational Applications.” Neuropsychobiology vol. 72,3-4 2015, 139-50.
3. Edoho, M.; Mooney, C.; Wei, L. AI-Based Electroencephalogram Analysis in Rodent Models of Epilepsy: A Systematic Review. Appl. Sci. 2024, 14, 7398.

Oral Communication abstracts

A GLIMPSE OF SOME POCKET COMPARISON APPROACHES

Anna SONG,[1] Christian ESPOSITO,[1] Floriane MONTANARI[1]

[1] *Owkin, Inc*
{anna.song, christian.esposito, floriane.montanari}@owkin.com

Keywords: pockets, similarity, selectivity, off-target toxicity, representation

Finding the right representation of protein pockets is key to several tasks in drug discovery. In particular, off-target toxicity must be prevented early in the discovery pipeline by assessing the similarity between the binding site of interest and binding sites on off-targets that represent a toxicity risk. Here the choice of the representation is crucial to determine if differences can be exploited in order to achieve enough selectivity. In this short talk, we will give a glimpse of a few pocket comparison approaches, including a CNN method (DeeplyTough[1]) that showed very high performances on some benchmark datasets (TOUGH-M1[2], Vertex[3], ProSPECCTs[4]), a GNN approach (HCGNet[5]) that explicitly handles the chemical and geometric interactions in a multiscale fashion, and some baseline models that rely on handcrafted features or geometric alignment to compute similarity. We will finally suggest that similarity labels in benchmarking datasets are somehow ill-posed and can only partially address the question of off-target toxicity.

References:
1. M. Simonovsky, J. Meyers, *J. Chem. Inf. Model.* **2020**, *60*, 2356.
2. R. G. Govindaraj, M. Brylinski, *BMC Bioinformatics* **2018**, *19*, 91.
3. Y.-C. Chen, R. Tolbert, A. M. Aronov, G. McGaughey, W. P. Walters, L. Meireles, *J. Chem. Inf. Model.* **2016**, *56*, 1734.
4. C. Ehrt, T. Brinkjost, O. Koch, *PLOS Computational Biology* **2018**, *14*, e1006483.
5. Y. Lin, L. Pan, Y. Li, Z. Liu, X. Li, *IEEE Journal of Biomedical and Health Informatics* **2024**, *28*, 1927.

LIGAND-BASED DRUG DISCOVERY LEVERAGING STATE-OF-THE-ART MACHINELEARNING METHODOLOGIES EXEMPLIFIED BY CDR1 INHIBITOR PREDICTION

The-Chuong, TRINH,[1] Pierre, FALSON,[2] Viet-Khoa, TRAN-NGUYEN,[3] and Ahcène, BOUMENDJEL[1]

[1] Univ. Grenoble Alpes, INSERM, LRB, Grenoble 38000, France

[2] Drug Resistance & Membrane Proteins Group, CNRS-Lyon 1 University Laboratory, UMR 5086, IBCP, 69367 CEDEX Lyon 07, France

[3] CNRS UMR8251, Université Paris Cité, INSERM U1133, F-75013 Paris, France

the-chuong.trinh@univ-grenoble-alpes.fr

Keywords: ligand-based drug discovery, machine learning, multi-instance learning, graph neural network, stacking ensemble.

Artificial intelligence (AI) is revolutionizing drug discovery with unprecedented speed and efficiency. In computer-aided drug design, structure-based and ligand-based methodologies are the main driving forces for innovation. In case no experimental structure or high-confidence homology/AlphaFold-predicted model of the target is available in 3D, ligand-based strategies are generally preferable. Here, we aim to develop and evaluate new predictive AI models for ligand-based drug discovery. To illustrate our workflow, we propose, as an example, an ensemble classification model for Cdr1 inhibitor prediction. We leverage target-specific experimental data from different sources, various molecular feature types, multiple state-of-the-art machine learning (ML) algorithms alongside a multi-instance 3D graph neural network (multiple conformations of a single molecule are considered). Bayesian hyperparameter tuning, stacked generalization and soft voting are involved in our workflow. The final target-specific ensemble model benefits from the classification and screening power of those constituting it. On an external test set structurally dissimilar to the training data, its average precision is 0.755, its F1-score is 0.714, the area under the receiver operating characteristic curve is 0.884, and the balanced accuracy is 0.799. It gives a low false positive rate of 0.1236 on another test set outside the training chemical space, indicating its ability to avoid false positives. The input of our tool was approved by recent synthesis and biological evaluation of some original and powerful Cdr1 inhibitors (results not shown, patent application pending). The present work highlights the potential of stacking ensemble ML and offers a rigorous general workflow to build ligand-based predictive AI models for other targets.

References:

1. Trinh, T.C.; Falson, P.; Tran-Nguyen, V.K.; Boumendjel, A., *J. Chem. Inf. Model.* **2025**, (in press), DOI: 10.1021/acs.jcim.5c00374.
2. Zankov, D. V.; Matveieva, M.; Nikonenko, A. V.; Nugmanov, R. I.; Baskin, I. I.; Varnek, A.; Polishchuk, P.; Madzhidov, T. I. QSAR Modeling Based on Conformation Ensembles Using a Multi-Instance Learning Approach. *J. Chem. Inf. Model.* **2021**, *61*, 4913−4923.
3. Tran-Nguyen, V. K.; Prasad, R.; Falson, P.; Boumendjel, A. Modulators of the Efflux Pump Cdr1p of Candida albicans: Mechanisms of Action and Chemical Features. *Curr. Med. Chem.* **2017**, *24*, 3242−3253.
4. Nim, S.; Baghel, P.; Tran-Nguyen, V.-K.; Peres, B.; Nguyen, K.A.; Pietro, A. D.; Falson, P.; Prasad, R.; Boumendjel, A. Make azoles active again: chalcones as potent reversal agents of transporters-mediated resistance in Candida albicans. *Future Medicinal Chemistry* **2018**, *10*, 2177−2186.

AN EXAMPLE TO RESOLVE GENETIC DEADLOCKS: FROM RYR1 STRUCTURAL MODELING TO AI-BASED VARIANT CLASSIFICATION

John RENDU[1,2,3]

[1] *Univ. Grenoble Alpes*
[2] *CHU Grenoble Alpes*
[3] *Grenoble Institute Neurosciences*

john.rendu@univ-grenoble-alpes.fr

Keywords: RYR1, VUS, machine learning, Structural modeling, myopathy

To date, more than half of patients affected by rare diseases remain without a definitive diagnosis. Despite advances in sequencing technologies, the diagnostic yield for identifying the genetic causes of rare diseases remains limited to approximately 40–50%. This means that nearly half of these patients remain in a state of diagnostic deadlock. The major contributor to this limitation is the inability to interpret genetic variants detected through sequencing, which prevents their classification as either benign or pathogenic.
In the end 2024, the ClinVar database contained over 1,355,555 variants, of which 908,752 (67%) were classified as variants of uncertain significance (VUS), underscoring the critical need for improved interpretative tools.
Beyond functional validations that are gene-specific and time-consuming, we propose to significantly reduce the number of VUS—particularly missense variants—by optimizing bioinformatics tools in line with international guidelines, such as those from the American College of Medical Genetics and Genomics (ACMG). However, even with these guidelines, over 90% of missense variants remain unclassified due to the inability of current algorithms to fully capture the biological complexity and uniqueness of protein function.
This project aims to develop a variant classification algorithm leveraging protein structural knowledge and advanced machine learning techniques to enable accurate and biologically relevant predictions. We focused our efforts on the *RYR1* gene, for which our diagnostic and research laboratory has over 25 years of expertise.
The initial phase involved homology modeling of the human RYR1 protein structure. Each variant was then characterized using novel, protein-specific structural descriptors. We constructed a curated dataset comprising both pathogenic and benign variants derived from literature and clinical diagnostics.
Using this dataset, we trained a machine learning model capable of predicting pathogenicity scores across all possible amino acid substitutions in *RYR1*.

Our model not only demonstrates strong predictive performance but also offers explainability—crucial for clinical interpretation and decision-making.
This approach represents a step forward in the use of AI for variant classification and may serve as a framework applicable to other clinically relevant genes.

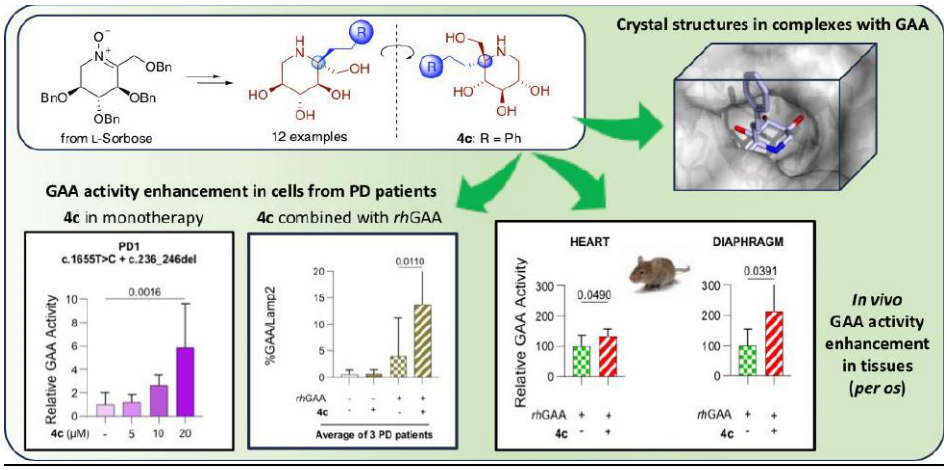
C-Branched Iminosugars as Selective Pharmacological Chaperones of Lysosomal Alpha-Glucosidase for the Treatment of Pompe Disease

Sandrine PY,[1] Jean-Bernard BEHR,[2] Marta ARTOLA,[3] Herman S. OVERKLEEFT,[3] Marco MORACCI,[4] Gerlind SULZENBACHER,[5] Giancarlo PARENTI[6]

[1] *Univ. Grenoble Alpes, CNRS, DCM, 38000 Grenoble, France*
[2] *Université de Reims Champagne-Ardenne, CNRS, ICMR, Reims, France*
[3] *Department of Medical Biochemistry and Department of Bio-organic synthesis, Leiden Institute of Chemistry (LIC), Leiden University, Leiden 2333 CC, The Netherlands*
[4] *Department of Biology, University of Naples "Federico II", Complesso Universitario di Monte S. Angelo, Naples, Italy and National Biodiversity Future Center (NBFC), 90133 Palermo, Italy*
5 *Architecture et Fonction des Macromolécules Biologiques (AFMB), CNRS, Aix-Marseille University, Marseille, France*
[6] *Telethon Institute of Genetics & Medicine, Pozzuoli, Italy and Department of Translational Medical Sciences, Federico II University, Naples, Italy*
E-mail: sandrine.py@univ-grenoble-alpes.fr

Keywords: Iminosugars • Selectivity • Acid alpha-glucosidase • Pharmacological Chaperones • Pompe Disease

An efficient synthesis of new iminosugars and their biological evaluation as acid α-glucosidase (GAA) pharmacological chaperones will be presented. The latter represent a novel class of orally available small-molecule stabilizers improving this enzyme stability. In particular, 5-*C*-phenethyl-DNJ (**4c**) was studied in human cells and *in vivo* (mice) as an outstandingly selective GAA stabilizer, paving the way for further development towards clinical investigations, for improving the care of patients suffering from Pompe disease, and to establish a cost-effective and scalable therapeutic approach.



Further work will involve molecular modeling of the interaction of chaperone **4c** with GAA and its impacts on enzyme 3D-structure, to generate useful data for computational analysis and AI treatment in view to hopefully correlate enzyme mutations to patient responsiveness to chaperones.

References:
Vieira Da Cruz, A., Perraudin, V., Minopoli, N., Iacono, R., Roig-Zamboni, V., Bossio, A., Tangara, S., Fayolle, M., Kanazawa, A., Philouze, C., Tarallo, A., Heming, J.J.A., Artola, M., Behr, J.-B., Overkleeft, H.S., Moracci, M., Sulzenbacher, G., Parenti, G., Py, S., *C*-Branched Iminosugars as Selective Pharmacological Chaperones of Lysosomal Alpha-Glucosidase for the Treatment of Pompe Disease. Submitted, manuscript under review.

AI-driven Design of human Respiratory Syncytial Virus Miniprotein Inhibitors

Maxime Bierre[1], Elodie Tacussel[1], Marie Galloux[2], Cédric Leyrat[3], Jean-Marie Bourhis[1], Marc Jamin[1]

[1]Institut de Biologie Structurale, Université Grenoble-Alpes.
[2]Unité de Virologie et Immunologie Moléculaires, INRAE Jouy-en-Josas
[3]Institut de Génomique Fonctionnelle, CNRS-INSERM-UM, Montpellier
marc.jamin@ibs.fr

Keywords: miniprotein design, viral replication, antiviral, respiratory syncytial virus, bronchiolitis

Human respiratory syncytial virus (hRSV) infections primarily affect infants, the elderly, and immunocompromised patients. They are a major cause of morbidity and mortality worldwide and represent a significant social and economic burden. However, although a new generation of monoclonal antibodies preventing severe hRSV infections in infants is now available, there are still no vaccine or treatment options for the vast majority of the population including vulnerable patients at risk of severe infections. The development of novel antivirals to treat hRSV infections, whether prophylactic and/or curative, remains a strong societal and medical need.

For many years we studied the encapsidation of viral RNA by the nucleoprotein (N) to form new nucleocapsids (NC), an essential step in the virus replication cycle. In this process, the production of unassembled RNA-free N (N^0) requires interaction with the phosphoprotein (P), followed by its transfer and polymerization onto a newly synthesized RNA. In previous works, we discovered that interfering with the formation of the N^0-P complex inhibits viral replication (1, 2).

In order to explore the possibility of developing a therapeutic compound in this line, we established a pipeline of available AI-based computational tools and used it to designed *de novo* miniproteins (<70 aa) targeting the human respiratory syncytial virus (hRSV) nucleoprotein. We selected and experimentally tested 12 miniproteins with different “hallucinated” folds. Of these, 10 miniproteins were successfully expressed in bacteria, 4 have so far crystallized and yielded high-resolution structures, and 6 showed nanomolar affinity for their targets, effectively inhibiting viral RNA synthesis in *in vitro* minireplicon assays. These results show that only a small fraction of computer-generated models require experimental validation to identify high-affinity ligands, making this approach both accessible and effective, within the funding capabilities of our laboratories.

References:

1. Galloux M, Gabiane G, Sourimant J, Richard CA, England P, Moudjou M, Aumont-Nicaise M, Fix J, Rameix-Welti MA, Eléouët JF. Identification and characterization of the binding site of the respiratory syncytial virus phosphoprotein to RNA-free nucleoprotein. *J Virol.* **2015**, 89, 3484-3496. doi:10.1128/JVI.03666-14.

2. Yabukarski F, Lawrence P, Tarbouriech N, Bourhis JM, Delaforge E, Jensen MR, Ruigrok RW, Blackledge M, Volchkov V, Jamin M. Structure of Nipah virus unassembled nucleoprotein in complex with its viral chaperone. *Nat Struct Mol Biol.* **2014**, 21, 754-759. doi: 10.1038/nsmb.2868.

INTEGRATIVE PHARMACOGENOMIC MODELS AND SIGNATURES FOR PATIENT THERAPEUTIC RESPONSE STRATIFICATION IN METASTATIC KIDNEY CANCER

Florian JEANNERET, [1] Liangwei YIN,[1] KATY Consortium,[2] CANVAS Consortium,[3] Christophe BATTAIL [1]

[1] Univ. Grenoble Alpes, IRIG, Laboratoire Biosciences et Bioingénierie pour la Santé, UA 13 INSERM-CEA-UGA, 38000 Grenoble, France.
[2] European Union's Horizon 2020 RIA project (grant agreement No 101017453)
[3] European Union's Horizon Europe Twinning project (grant agreement No 10107951)
christophe.battail@cea.fr

Keywords: Pharmacogenomics, kidney cancer, RNA-seq, multi-omics, precision oncology.

The clinical response to targeted therapies and immune checkpoint inhibitors (ICIs) in patients with metastatic clear cell renal cell carcinoma (ccRCC) is highly heterogeneous and remains largely unpredictable (1). The increasing availability of bulk RNA sequencing data from clinical drug trials offers new opportunities for developing predictive pharmacogenomic models, stratifying patients, and identifying clinically translatable molecular signatures (2). We conducted three complementary studies demonstrating that transcriptomic data from clinical drug trials, when augmented with publicly available multi-omics datasets and curated biological knowledge, enabled the development of precision oncology approaches that were compatible with clinical constraints. (I) Tumor Microenvironment (TME): Transcriptome-based deconvolution of metastatic ccRCC samples allowed us to reveal TME cell fraction signatures associated with ICI response (3). We developed a Tumor-Immunity Differential (TID) score that integrates tumor-intrinsic and microenvironmental features, achieving high predictive performance (AUC-ROC = 0.88) for ICI response. YWHAE, a component of the TID score, emerged as a robust predictive biomarker and was validated in independent melanoma and lung cancer cohorts. (II) Gene Coexpression Networks: By applying a patient-specific gene coexpression network approach, we demonstrated that patients with similar therapeutic outcomes exhibit shared network topologies (4). Increased negative gene-gene correlations were linked to poor prognosis. Moreover, incorporating network-derived features into machine learning models outperformed traditional gene expression-based approaches and highlighted dysregulated pathways associated with therapy response . (III) DNA Methylation Subtypes: Through meta-analysis of 700 ccRCC tumors profiled by DNA methylation, we benchmarked multiple published probe-based signatures and identified two distinct subtypes (5). Hypermethylated tumors were associated with worse prognosis, increased tumor cell proliferation, and reduced activity of homeobox transcription factors. To enable clinical translation, we developed a simple yet accurate classifier based on two gene expression ratios (IGF2BP3/PCCA and TNNT1/TMEM88), achieving an AUC-ROC of 0.91. These methylation subtypes were significantly associated with differential responses to ICIs and targeted therapies. Together, our integrative pharmacogenomic framework, leveraging bulk transcriptomics from clinical trials enriched with single-cell RNA-seq, interactomic, or epigenomic data, uncovers robust, interpretable signatures to improve patient stratification and guide precision therapeutic strategies in metastatic ccRCC.

References:

1. Morad G, et al. Hallmarks of response, resistance, and toxicity to immune checkpoint blockade. Cell. 2021 Oct 14;184(21):5309–37.
2. Braun DA, et al. Interplay of somatic alterations and immune infiltration modulates response to PD-1 blockade in advanced clear cell renal cell carcinoma. Nat Med. 2020 Jun;26(6):909–18.
3. Jeanneret F, et al. Microenvironment of metastasis reveals key predictors of PD-1 blockade response in renal cell carcinoma. bioRxiv; 2024. p. 2023.07.17.548676.
4. Yin L, et al. Sample-specific network analysis identifies gene coexpression patterns of immunotherapy response in advanced kidney cancer. bioRxiv; 2024. p. 2024.11.02.621068.
5. Jeanneret F, et al. DNA methylation is a key determinant of response to targeted and immune checkpoint therapies in metastatic renal cell carcinoma. bioRxiv; 2024. p. 2024.11.05.622095.