Reinforcement Learning based on XAI methods for de novo drug generation

PhD Proposal

PhD Supervisors: Thomas Papastergiou <papastergiou@lipn.univ-paris13.fr> and **Céline Rouveirol** <rouveirol@lipn.univ-paris13.fr>

1. Context

In this PhD we will develop new **deep reinforcement learning generative models** based on **reward functions derived from XAI methods**. In deep reinforcement learning methods, the policies and/or reward functions can be modeled by deep neural networks that in general act as black boxes lacking explanations on why a specific inference has been made. Our goal is **to exploit XAI methods to drive the generation process of the reinforcement learning generative models through the search space to generate output with the desired properties**. A possible important application of the aforementioned methods could be in the field of in silico de novo drug design (i.e. using AI generative methods to propose new molecules that can be effective against specific targets).

For a new drug to acquire marked authorization, about 10-14 years and 1 billion USD, are needed [1]. The first step in medicinal chemistry is the identification of new candidate compounds that will be synthesized and tested *in vitro*. To speed up this process *in silico* strategies, exploiting Machine Learning (ML) and Deep Learning (DL), have been extensively used [2] and AI Generative Models (GMs) have been proposed for *de novo* Drug Design [3]. Although their efficiency, the lack of their interpretability, makes the underlying decision-making mechanisms nontransparent, does not reveal inserted bias and correct predictions for the wrong reasons. Thus, domain experts (outside AI) find these models not trustable. In the last few years, a new field of research arose: **eXplainable Artificial Intelligence (XAI)** that aims to provide the end user of an AI model with transparency, justification, informativeness, and uncertainty estimation [4].

An important application of the methods developed in this PhD can be in the field of discovering new antibiotics. The excessive use of antibiotics led to the increase of bacterial resistance over the past 30 years, causing only in Europe about 33,000 deaths per year. World Health Organization recently (2019) declared that bacterial resistance is one of the **top 10 global public health threats that humanity faces**[5]. The most important treatment against bacterial infections nowadays, relies on antibiotics of the β-lactam class (the core of penicillin). **Bacteria resistance to these antibiotics consists in their ability to express specific enzymes (β-lactamase) that deactivate their antimicrobial ability** [6]. New-Delhi-Metallo-β-lactamase-1 (NDM-1), (first observed in India in 2008), and its variants, have spread worldwide, producing highly resistant pathogens (super-bugs), that represent a major threat on human health [7], since they are capable to deactivate the antimicrobial action of a broad spectrum of antibiotics (including last resort antibiotics). Thus it is crucial to address this menace. An approach to tackle this problem, is the use of effective inhibitors, for deactivating the NDM-1 enzymes, along with a β-lactam antibiotic agent. Currently, there are no NDM-1 inhibitors available in clinical practice, but numerous studies report on the effectiveness of different compounds. **It is an important task and a challenge to propose new compounds with desirable properties that can act as inhibitors for all or a broad number of NDM-1 variants**.

XAI methods used in drug discovery can be categorized in **Feature attribution** (FA), **Instance based** and **Graph-Convolution** based methods. Several feature attribution methods (mapping the features of the input to a contribution value) have been proposed: e.g. in [8] a gradient based FA approach that

revealed suspicious correlations in the explanations. Another study [9] exploited the SHAP methodology on molecular fingerprints (i.e. substructures) based ML models. Although this approach can deliver molecular sub-structures explanations (i.e. substructures potentially linked to the activity of a molecule), molecular fingerprints pose efficiency problems in the models' training because of their high dimensionality. Protein-ligand binding affinity predicted by 3D CNNs, based on 3D molecular structures, was interpreted by 3D visualization methods [10], an approach based on re-docking ligands, a very time consuming, and sometimes not accurate procedure. Studies based on character-wise representations (SMILES) of molecules, like in [11], can provide only atom-wise explanations. Recently, an explainability method exploiting the SHAP-DNN methodology was proposed [12]. Based on GNN models for SARs explanations this method cannot provide substructure explanations and lack of efficiency of the underlying predictive models, possibly due to the lack of training data. **Instance based approaches**, identify the presence of a set of features that must be present to retain, or be absent to change the prediction of a model. They can be employed for making DL models explainable. Examples of such methods are **anchor algorithms**[13], providing if-then rules for a set of features that are required for retaining the same prediction; **counterfactual instance search**[14], that aim to find data points as close to the original one, that make the model change its prediction; and c**ontrastive explanation methods**[15], that rely on the one hand on discovering 'pertinent positive' sets, the smallest set of features to be present in an instance for the model to predict a positive label. Furthermore, they can also identify and 'pertinent negative' sets, the smallest set of features to be absent in order for the model to distinguish an instance from other classes. **To the best of our knowledge, instance-based approaches are not yet applied to the drug discovery domain**. Finally, **Graph Convolution-based** approaches have been proposed, like a subgraph identification approach [16] or an Attention-based method [17] that also lack substructure explanations.

The majority of the works on XAI for Structure Activity Relationships (SAR) prediction, are providing explanations on atoms and bonds, or/and features related to atoms/bonds. Thus, more research on natural explanations (e.g. substructure explanations), closer to the chemical intuition, is needed, since the activity of an inhibitor is not relying on atoms or bonds, but on chemical substructures. Compound optimization (changing or adding small molecular structures to a compound for optimizing their properties), can benefit from explanation strategies that can indicate the active/inactive substructures of a compound. Finally, there is a need to investigate further novel ways for guiding the generation procedure of GMs through the vast molecular space, towards new feasible compounds with the desired properties.

2. Research Directions

In this dissertation, we will rely on a database of **NDM-1 activities [18, 19]** for designing and evaluating the proposed models. We will adapt, XAI methods **not yet used in SAR prediction of NDM-1 inhibitors** (e.g. XAI MIL models and instance based approaches), to explain molecular substructure contributions on NDM-1 bioactivity and we will use SOTA (State-of-the-art) Data Mining (DM) methods **to analyze and potentially discover new SARs for NDM-1 inhibitors**. Furthermore, we will design a novel Generative Method (GM) using SOTA Deep Reinforcement Learning (DRL) methods (e.g. REINFORCE or the Cross Entropy Method), that will incorporate the knowledge acquired from the elaborate analysis of SARs through XAI. To the best of our knowledge, **none of the works done so far does not incorporate in the reward design elaborate information on substructures activity**. **A novelty of the approach is that the generative model will be designed and tailored for generating NDM-1 inhibitors, unlike other approaches where the design of the GM is general or in some cases can be adapted to specific tasks.**

The output compounds will be evaluated for their synthetic feasibility and docking ability to NDM-1 variants in collaboration with a multidisciplinary team of IBMM in Montpellier, specialized in medicinal chemistry.

3. Research Methodology and approach

Phase 1: Explanations' extraction and analysis for the activity classifier (12-15 months): In the first place, we will design, implement and evaluate a Deep Learning (DL) activity classifier that will take as input molecular substructures. We will investigate two different representations of molecules: as bags of (unordered) Mol2Vec[20] vectors (corresponding to substructures) that will lead to the design of a MIL Deep Learning Classifier (e.g. like in [21]), or as an ordered sequence of Mol2Vec vectors that will lead to an Recurrent Neural Network (RNN)-like classifier (Long-Short Term Memory (LSTM), Bi-directional LSTM, Gated Recurrent Unit Networks, Transformers etc.). The classifiers will be evaluated and the best performing will be used in the subsequent phases of the PhD.

Afterwards, we will adapt XAI approaches to the designed classifier. In particular, we will investigate **counterfactual** and **contrastive explanation instance** based methods, **adapted to the MIL or the RNN setting and to the molecular representation**. The latter is crucial, because we aim for activity explanations corresponding to molecular substructures, unlike the majority of the works that output atom- and bond-wise explanations.

Finally, we will analyze and evaluate the explanations extracted from the XAI framework with the objective to identify molecular substructures that contribute to the activity prediction. This will be done using classical statistical methods and data mining approaches e.g. pattern mining approaches like association rule mining, sequential pattern mining or sequence rule mining. We will use open-source DM libraries (e.g. SPMP [22]) or we will implement some approaches. For evaluating the mined rules, we will collaborate and consult the IBMM team, for comparing the extracted rules to the existing binding NDM-1 inhibitors information from the PDB-Protein DB [23].

Phase 2: Reinforcement Learning model and reward function using the knowledge from XAI (12-15 months) : In this phase**,** we will design, implement and evaluate a DRL model for generation of NDM-1 inhibitors. To this goal we will design and implement Cross Entropy based methods, as we will investigate other approaches as well: e.g. Q-Learning, REINFORCE etc. In the first place we will use as a reward function the classification activity model constructed in **Phase 1**. This will act as a general GM for NDM-1 inhibitors without the incorporation of the knowledge from the XAI methods.

Afterwards, we will design a reward function based on the knowledge acquired by exploring the explanations delivered by the XAI methods (**Phase 1**). Different multi-objective methodologies, like in [24] or multi-objective reward shaping [25] will be explored, will be adapted to our problem and evaluated. Furthermore, we will explore and exploit explicable reward functions (as in [26] or [27]), with the aim to design a combined multi-objective explicable reward function that will explain the drug generation process. By exploiting the explicable nature of the designed function we will be able to readapt the reward function, if needed, for better guiding the generation processes through the generation space.

(About 3 to 6 months have been assigned to literature review as well as to the dissertation reduction).

4. Collaborations:

In the frame of this PhD we will collaborate with the interdisciplinary team of Institut des *Biomolécules Max Mousseron (IBMM)* of Montpellier (UMR 5247). More specifically Dr. Laurent Gavara who is specialized in medicinal chemistry (on discovering NDM inhibitors (i.e. compounds against resistant bacteria)) will provide access to specialized compounds activities databases, as well as he will provide input and evaluate the explanations and the generated molecules by the proposed models.

References:

[1] J. A. DiMasi, et al.(2016), «Innovation in the pharmaceutical industry: New estimates of R&D costs», J. Health Econ., vol. 47, p. 20-33
[2] S. Dara et al., (2022), Machine Learning in Drug Discovery: A Review, Artif. Intell. Rev., vol. 55, no 3, p. 1947-1999
[3] Mouchlis VD,et al. (2021) Advances in De Novo Drug Design: From Conventional to Machine Learning Methods. Int. J. of Mol. Sc.
[4] Jiménez-Luna, et al. (2020). Drug discovery with explainable artificial intelligence. Nature Machine Intelligence, 2(10), 573-584
[5] World Health Organization. (2021). Global antimicrobial resistance and use surveillance system (GLASS) report: 2021
[6] Shi, C., et al. (2019). Approaches for the discovery of metallo-β-lactamase inhibitors: A review. Chem. Biol. & D. D., 94(2), 1427-1440.
[7] Khan, et al., (2017) Genetics and Worldwide Spread of New Delhi Metallo-β-lactamase (NDM): a threat to public health. BMC Microbiol 17, 101.
[8] McCloskey, et al. (2019). Using attribution to decode binding mechanism in neural network models for chemistry. Proceedings of the National Academy of Sciences, 116(24), 11624-11629.
[9] Rodríguez-Pérez et al. (2019). Interpretation of compound activity predictions from complex machine learning models using local approximations and shapley values. Journal of medicinal chemistry, 63(16), 8761-8777
[10] Hochuli, J., et al. (2018). Visualizing convolutional neural network protein-ligand scoring. J. of Mol. Graph. and Modelling, 84, 96-108.
[11] Karpov, et al. (2020), Transformer-CNN: Swiss knife for QSAR modeling and interpretation. J Cheminform 12, 17
[12] Harren, T., et al. (2022). Interpretation of structure–activity relationships in real-world drug design data sets using explainable artificial intelligence. Journal of Chemical Information and Modeling, 62(3), 447-462
[13] Ribeiro, et al.(2018). Anchors: High-precision model-agnostic explanations. In Proc. of the AAAI conference on AI (Vol. 32, No. 1)
[14] Wachter, S., et al. (2017). Counterfactual explanations without opening the black box: Automated decisions and the GDPR. Harv. JL & Tech., 31, 841
[15] Dhurandhar, et al. (2018). Explanations based on the missing: Towards contrastive explanations with pertinent negatives. Advances in neural information processing systems, 31
[16] Ying, et al. (2019). Gnnexplainer: Generating explanations for graph neural networks. Adv. in Neural Inf. Processing Systems, 32.
[17] Veličković, et al. (2017). Graph attention networks. arXiv preprint arXiv:1710.10903.
[18] Papastergiou et al. (2022) Multiple Instance Learning Based on Mol2vec Molecular Substructure Embeddings for Discovery of NDM-1 Inhibitors. 16th PACBB.
[19] Papastergiou, T., et al. (2023). Discovering NDM-1 inhibitors using molecular substructure embeddings representations. Journal of Integrative Bioinformatics, 20(2)
[20] Jaeger,et al. (2018). Mol2vec: unsupervised machine learning approach with chemical intuition. J. of Chem. Inf. and Mod. 58(1), 27-35
[21] Ilse, M., et al. (2018). Attention-based deep multiple instance learning. In Int. Conf. on ML (pp. 2127-2136). PMLR.
[22] https://www.philippe-fournier-viger.com/spmf/
[23] https://www.rcsb.org/
[24] Nguyen, T. T., et al. (2020). A multi-objective deep reinforcement learning framework. Eng. Appl. of AI, 96, 103915
[25] Mannion, P. et al. (2017). Policy invariance under reward transformations for multi-objective reinforcement learning. Neurocomputing, 263, 60-73
[26] Devidze R., Radanovic G., Kamalaruban P., Singla A., Explicable Reward Design for Reinforcement Learning Agents. NeurIPS 2021: 20118-20131
[27] Saulières L., Cooper M. C., Bannay F., Reinforcement Learning Explained via Reinforcement Learning, Towards Explainable Policies through Predictive Explanation. ICAART (2) 2023: 35-44