List of publications

Part of this thesis

Liu X, Ye K, van Vlijmen HWT, IJzerman AP, van Westen GJP. An exploration strategy improves the diversity of de novo ligands using deep reinforcement learning: a case for the adenosine A_{2A} receptor. Journal of cheminformatics. 2019;11 (1):35.

Liu X, IJzerman AP, van Westen GJP. Computational approaches for *de novo* drug design: past, present, and future. Methods Mol Biol. 2021;2190:139-65.

Liu X, Ye K, van Vlijmen HWT, Emmerich MTM, IJzerman AP, van Westen GJP. DrugEx v2: *de novo* design of drug molecules by Pareto-based multi-objective reinforcement learning in Polypharmacology. Journal of cheminformatics 2021:13(1):85.

Sicho M, **Liu X**, Svozil D, van Westen GJP. GenUI: interactive and extensible open source software platform for de novo molecular generation and cheminformatics. Journal of cheminformatics. 2021;13(1):73.

Liu X, Ye K, van Vlijmen HWT, IJzerman AP, van Westen GJP. DrugEx v3: Scaffold-Constrained Drug Design with Graph Transformer-based Reinforcement Learning. (*preprint*)

Not part of this thesis

Yang Z, Wang C, Wang T, Bai J, Zhao Y, **Liu X**, et al. Analysis of the reptile CD1 genes: evolutionary implications. Immunogenetics. 2015;67(5-6):337-46.

Liu X, Yang S, Li C, Zhang Z, Song J. SPAR: a random forest-based predictor for self-interacting proteins with fine-grained domain information. Amino Acids. 2016;48(7):1655-65.

Chen Z, **Liu X**, Li F, Li C, Marquez-Lago T, Leier A, et al. Large-scale comparative assessment of computational predictors for lysine post-translational modification sites. Brief Bioinform. 2019;20(6):2267-90.

Chen Z, He N, Huang Y, Qin WT, **Liu X*,** Li L. Integration of A Deep Learning Classifier with A Random Forest Approach for Predicting Malonylation Sites. Genomics Proteomics Bioinformatics. 2018;16(6):451-9..

Chen Z, **Liu X**, Zhao P and Song J. iFeature^{Omega}: a comprehensive platform for generating, analyzing and visualizing various representations for biological sequences, structures and ligand. (*Submitted*)

Curriculum Vitae

Xuhan Liu was born on September 28th, 1989 in Jia County, Henan province, China. From 2007 to 2012, he studied in the College of Life Sciences in Henan Agricultural University. Besides following the majority of subjects in biological sciences, he also took a large amount of time in learning computer sciences and designed the first website for the biological lab where he took his internship. In the end, he obtained a Bachelor degree in biology with one year extension when he planned to further study the cross domain between computer sciences and biology.

In the following three years, he studied in the college of biology in China Agricultural University and was supervised by Prof. Dr. Ziding Zhang in the protein bioinformatics group. In this group, he obtained extensive knowledge of artificial intelligence. He also applied what he learned in the project he undertook about self-interacting protein prediction. In this project, he proposed a novel feature encoding scheme and successfully improved the performance of the predictive model. He also tried to apply deep learning with ambitious faith in spite of the failure in the end. In addition, he also designed a webserver to manage and release the information for this lab. In 2015, he obtained the Master degree in biology (bioinformatics).

From 2015 to 2017, he worked as a bioinformatician in Oriental YAMEI Gene Technology Research Institute Co., Ltd. His main work was the study of association between human genotype and phenotype under the lead of Prof. Dr. Li Shen. He developed the online office system for the company to generate and manage the gene detection reports of customers. Moreover, he also collaborated with Dr. Zhen Chen to study the prediction and functional analysis of lysine post-translational modification in proteins with deep learning.

After obtaining funding from China Scholarship Council in 2017, he began his study at the division of drug discovery and safety at LACDR/Leiden University under the supervision of Prof. Dr. A.P. IJzerman and Prof. Dr. G.J.P van Westen. In addition, he also spent time

in the academic activities organized by Prof. Dr. Aske Plaat and Dr. Mike Preuss in the reinforcement learning group to learn cutting-edge technologies of AI in LIACS/Leiden University. During these four years study, he thoroughly investigated the application of AI in drug *de novo* design. Based on these studies, he proposed a method named ***DrugEx*** containing relevant Python-based algorithms and released it on GitHub to facilitate its accessibility in the scientific community. During these four years, he pitched his work during different public academic occasions. In the last three months, he also got an opportunity to take an internship at Janssen Pharmaceuticals, also as a possible preparation for his career in the future.

Acknowledgements

My four years of research at Leiden University were not a smooth journey but filled with failures and setbacks. Especially, the pandemic outbreak of COVID-19 led to being forced to work from home, which made me doubt myself whether I can complete my PhD studies. Thanks to my friends who made endless efforts to help me overcome the hardships, I persisted in completing these projects and finishing the writing of this thesis. I will therefore give my full gratitude to the following people.

First of all I would like to thank my promoter Prof. Ad IJzerman. I am indebted to your suggestions about scientific directions which guided me to move forward and I also learned from your academic integrity embodied in your attitude towards every details of scientific research. Besides supervision of my study you also gave me advice about how to adapt myself to the culture of the Netherlands. I am also grateful for the help of my daily supervisor Prof. Gerard van Westen, who organized the academic activities and managed the computational equipment to provide an excellent scientific environment in our group. You also listened to me and helped me in all aspects of the life as a PhD student. In addition, I also thank my co-promoter Prof. Kai Ye and Prof. Herman W. T. van Vlijmen, who gave me useful comments about my works regularly. In addition, Prof. van Vlijmen provided me with an opportunity to take an internship at Janssen Pharmaceuticals.

In the division of drug discovery and safety, there are many kind-hearted colleagues, who establish a joyful atmosphere. I thank Martin that we cooperated to integrate ***DrugEx*** into web-based ***GenUI*** and successfully guided bachelor students into computational drug design. I also thank my master student Milan who successfully proved the effectiveness of ***DrugEx*** for other GPCRs. For Olivier, Bart, Willem, Brandon, Lindsey, Marina and Martin, thank you for discussing with me and giving me help when I met problems. Interestingly, I still remember the discussion about the name of ***DrugEx*** with Marina as a nice officemate. Thank you, Roelof, to fix the bugs on the computational servers. Thanks to Lia and Rongfang for arranging the daily routine and inform me on time. For Huub, Bert and

Sebastian, it was an unforgettable moment to take part in the Leiden Science run with you at my thirtieth birthday.

I was lucky to have a chance to learn many new AI methodologies from the reinforcement learning group in LIACS. Thanks to Prof. Aske Plaat and Dr. Mike Preuss for organizing regular meetings to discuss the fast developments in AI technologies. Thanks also to Dr. Michael Emmerich who provided me with the solutions of multi-objective optimization. I also thank Hui Wang for discussing the detailed application of AI approaches. Moreover, I also thank Dr. Erik van Geest from LIC and Zhenzheng Zhang from LACDR for collaborating with each other to undertake the projects about graphene and metabolism, respectively. Specially, I thank Prof. Zhen Chen in Henan Agricultural University who offered the opportunity to apply deep learning methods in biomedicine for the first time.

In Leiden I have made many truly good friends. Special thanks go to Xue Yang with whom I discussed many psychological theories and applied them to solve the problems we met in real life. I thank Wei Chen, Wensen Li, Xiangyu Liu, Xiaoyue Song and Li Jia for jogging 5km almost every afternoon together with me, which was great to keep my health state. I also thank my friends Di Wang, Yixin Luo, Wei Wang, Zhao Yang, Jing Zhang, Zhuoyi Luo and Ruochen Oyang for preparing delicious food and organizing parties frequently. Without you, my life in this foreign country would not have been so wonderful.

Last but not least, I thank the selfless support from my parents and my wife Dr. Lily Wang despite the remote distance between us. Because of your unrequited love, I feel strong to make up my mind and to accomplish my scientific research alone but not lonely. No matter what kind of achievements are ahead of me, I am more than willing to share these with you forever.