



Universiteit
Leiden
The Netherlands

To explore drug space smarter: artificial intelligence in drug design for G protein-coupled receptors

Liu, X.

Citation

Liu, X. (2022, February 15). *To explore drug space smarter: artificial intelligence in drug design for G protein-coupled receptors*. Retrieved from <https://hdl.handle.net/1887/3274010>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3274010>

Note: To cite this publication please use the final published version (if applicable).

To explore drug space smarter

Artificial intelligence in drug design for

G protein-coupled receptors

Proefschrift

ter verkrijging van

de graad van doctor aan de Universiteit Leiden,

op gezag van rector magnificus prof.dr.ir. H. Bijl,

volgens besluit van het college voor promoties

te verdedigen op 15 februari 2022

klokke 10:00 uur

door

Xuhan Liu

刘 许晗

geboren te Henan, China

in 1989

Promotores:

Prof. dr. Adriaan P. IJzerman
Prof. dr. Gerard J. P. van Westen
Prof. dr. Kai Ye



**Universiteit
Leiden**
The Netherlands

Promotiecommissie:

Prof. dr. Hubertus Irth
Prof. dr. Joke Bouwstra
Prof. dr. Mario van der Stelt
Prof. dr. Iwan J.P. de Esch (Vrije Universiteit Amsterdam)
Dr. Francesca Grisoni (Technische Universiteit Eindhoven)

LACDR



Copyright © 2022 Xuhan Liu

All rights reserved

Thesis lay-out & Cover design: Xuhan Liu

Printing: PrintSupport4U

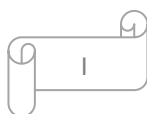
ISBN: 978-94-92597-94-6

This research is financially supported by the China Scholarship Council (CSC), CSC No. 201709110143

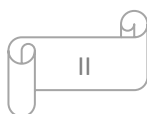
The research described in this thesis was performed at the Division of Drug Discovery and Safety of the Leiden Academic Centre for Drug Research (LACDR), Leiden University (Leiden, The Netherlands).

Content

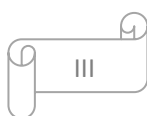
1.	General introduction	1
1.1.	AI in drug discovery	4
1.2.	G protein-coupled receptors	8
1.3.	Adenosine receptors	9
1.4.	Research questions	9
1.5.	Thesis outline	10
2.	Computational approaches for <i>de novo</i> drug design: past, present and future	17
2.1.	Introduction	19
2.2.	<i>De novo</i> drug design	21
2.2.1.	Molecular representations	22
2.2.2.	Multiple objectives	23
2.3.	Optimization methods	24
2.3.1.	Evolutionary algorithms	25
2.3.2.	Particle swarm optimization	28
2.3.3.	Simulated annealing	29
2.4.	Deep learning algorithms	30
2.4.1.	Recurrent neural networks	30
2.4.2.	Variational autoencoders	33
2.4.3.	Deep reinforcement learning	35
2.4.4.	Generative adversarial networks	37
2.5.	Competition or cooperation?	40
2.6.	Conclusion and perspective	42
3.	An exploration strategy improves the diversity of <i>de novo</i> ligands using deep reinforcement learning: a case for the adenosine A_{2A} receptor	49
3.1.	Introduction	51
3.2.	Dataset and methods	52
3.2.1.	Data source	52
3.2.2.	Prediction model (QSAR)	53
3.2.3.	Generative model	53
3.2.4.	Reinforcement learning	56



3.2.5.	Exploration strategy	56
3.2.6.	Molecular diversity	57
3.3.	Results and discussion.....	58
3.3.1.	Performance of predictors	58
3.3.2.	SMILES syntax learning	58
3.3.3.	Conditional SMILES generation.....	60
3.3.4.	Exploration strategy	62
3.3.5.	Comparison with other methods	64
3.4.	Conclusion and future prospect.....	71
4.	<i>DrugEx v2: de novo design of drug molecules by Pareto-based multi-objective reinforcement learning in polypharmacology</i>	81
4.1.	Introduction	83
4.2.	Materials and methods.....	85
4.2.1.	Data source.....	85
4.2.2.	Prediction model	85
4.2.3.	Generative model	87
4.2.4.	Reinforcement learning.....	88
4.2.5.	Algorithm extrapolation	91
4.2.6.	Exploration strategy	92
4.2.7.	Molecular diversity	93
4.3.	Results and discussion.....	94
4.3.1.	Performance of predictors	94
4.3.2.	Model optimization	95
4.3.3.	Performance comparisons	96
4.4.	Conclusion and future prospect.....	102
5.	<i>DrugEx v3: scaffold-constrained drug design with graph Transformer-based reinforcement learning</i>	113
5.1.	Introduction	115
5.2.	Materials and methods.....	116
5.2.1.	Data source.....	116
5.2.2.	Molecular representations	118
5.2.3.	End-to-end deep learning	119
5.2.4.	Multi-objective optimization.....	122



5.2.5.	Reinforcement learning.....	124
5.2.6.	Performance evaluation.....	125
5.3.	Results and discussion.....	126
5.3.1.	Fragmentation of molecule	126
5.3.2.	Pre-training & fine-tuning.....	127
5.3.3.	Policy gradient	131
5.3.4.	Generated molecules	134
5.4.	Conclusion and Future Perspective	135
6.	GenUI: interactive and extensible open source software platform for <i>de novo</i> molecular generation and cheminformatics.....	143
6.1.	Introduction	145
6.2.	Software architecture.....	148
6.3.	Frontend.....	150
6.3.1.	Graphical user interface (GUI).....	150
6.3.2.	Projects.....	151
6.3.3.	Compounds	151
6.3.4.	QSAR models	151
6.3.5.	Generators	153
6.3.6.	Maps.....	157
6.3.7.	JavaScript API.....	158
6.3.8.	Model components.....	159
6.3.9.	REST API components	160
6.4.	Backend	161
6.4.1.	Python API.....	161
6.4.2.	Extensions	162
6.4.3.	Automatic code discovery.....	163
6.4.4.	Generic views and viewsets	163
6.4.5.	Asynchronous tasks.....	164
6.4.6.	Integration of new features with the two APIs.....	165
6.4.7.	Compounds import.....	165
6.4.8.	QSAR models	165
6.4.9.	Molecular generators.....	166
6.4.10.	Chemical space maps	167



6.5.	Deployment	167
6.6.	Future directions.....	168
6.7.	Conclusions	169
7.	Conclusions and future perspectives	177
7.1.	Conclusions	179
7.2.	Further perspectives	182
7.2.1.	New AI technologies.....	182
7.2.2.	Different constraint conditions.....	183
7.2.3.	Designing various kind of molecules	184
7.3.	Final notes	185
Appendix	191
	Summary	193
	Samenvatting.....	195
	中文总结.....	199
	List of publications.....	201
	Curriculum Vitae.....	203
	Acknowledgements	205