This project involved the development of a computer protocol using the open-source chemical library RDKit and Python as the programming language. From the structure of a chemical backbone (Pyridinophane macrocycles), the protocol generates different derivative 2D structures by varying the attaching groups and their positions on the backbone structure, up to 18,000 final structures. The protocol then screens these structures for many desired chemical and pharmacological properties in order to narrow down the best few candidates for synthesis. This protocol can save time and resources for synthesis labs while providing a way to optimize chemical backbones in drug design.

Development of a Virtual Screening Protocol for Pyridinophane Macrocycle Derivatives as Therapies for Oxidative Stress

Minh Nhat Pham, Benjamin G. Janesko*

Department of Chemistry and Biochemistry, Texas Christian University, Fort Worth, TX 76129

Acknowledgements. MNP gratefully acknowledges Prof. Green and Prof. Janesko for their support. This work was supported by NIH 2R15GM123463-02.

INTRODUCTION

Oxidative stress refers to the imbalance between free radical activities and antioxidant activities in the body, and is known to be involved in the development of various neurodegenerative diseases (Alzheimer's and Parkinson's). To help the body target and rebalance this process, the Green group at TCU has designed pyridinophane macrocycle frameworks (PyN_3, Py_2N_2) for the development of a small multimodal molecule with direct targeting of oxidative stress. To enhance antioxidant activity, the group proposed a library of ligands as modifications to the pyridinophane frameworks.

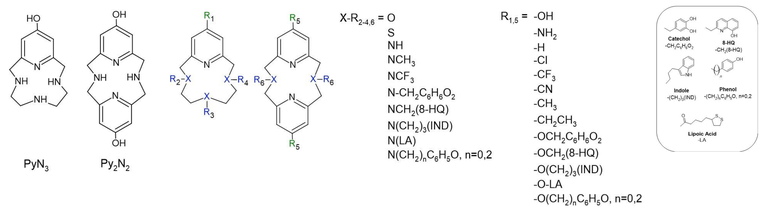


Figure 1. TCU Green group's successfully synthesized pyridinophane macrocycle frameworks (PyN_3 and Py_2N_2) and their proposed library of ligands as modifications to the backbone structures.

Due to the resulting large number of possible structures, computational pre-screening is essential to select the most promising candidates for synthesis as well as for biological and chemical testing. The ligand screening utilized the RDKit open-source cheminformatics toolkit,[1] with Python3 interface. The developed protocol is capable of generating the **2D structures** of all the possible modified structures of the pyridinophane frameworks and screen them for "drug-likeness" based on Lipinski's "Rule of Five." The properties computed and used as screening conditions in this project were **molecular weight** (MW), **ring count**, **octanol-water partition coefficient** (cLogP), number of **hydrogen bond donors** (HBD) and **acceptors** (HBA), and **polar surface area** (PSA).

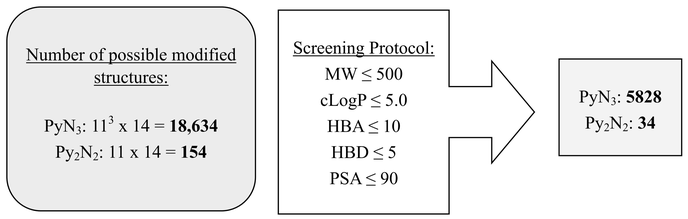


Figure 2. A general scheme of the screening protocol.

METHOD

The frameworks and ligands' structures were written as SMILES strings.

Ex: PyN_3: "**c1cc2nc(c1)CNCCNCCNC2**"

NCH_2(8-HQ): "**N(Cc5nc(c(O)ccc6)c6cc5)**"

For each attaching position, loop over the items of each ligand libraries and attach them to the frameworks as SMILES strings.

Ex: PyN3 :"**c1('+R1+')cc2nc(c1)C'+R2+'CC'+R3+'CC'+R4+'C2**"

Convert the SMILES strings to 2D molecular structures

Compute the descriptors for each resulting 2D structure (MW, ring count, cLogP, HBD, HBA, PSA)

MW ≤ 500 | **cLogP ≤ 5.0** | **HBD ≤ 5** | **HBA ≤ 10** | **PSA ≤ 90**

Count the total number of Lipinski' rule violations for each generated 2D structure (up to 5)

Count and report the number of derivative structures that has 0 violations of the Lipinski's rules

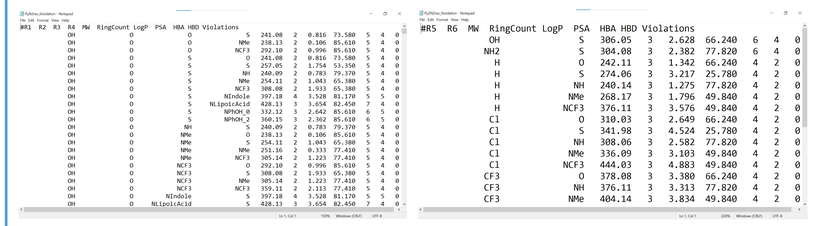
Lipinski's Rule of Five:

In 1977, after assessing various approved drugs and clinical candidates at that time, Lipinski and co-workers proposed the "**Rule of Five**" for orally active compounds.

For a drug molecule to be orally bioavailable, the molecule's lipophilicity and solubility are important factors in intestinal absorption. Molecules that are small (MW ≤ 500), less soluble in water (cLogP ≤ 5.0), has small total polar surface area (PSA ≤ 90) and hydrogen bond counts (HBD ≤ 5 and HBA ≤ 10) are **more likely to cross the lipid bilayers of the cell membrane**.[2]

Molecules failing to comply these rules may show poor absorption, faster rate of metabolism and excretion, unfavorable distribution, and toxicity.[3]

RESULTS

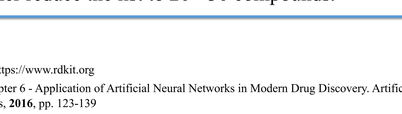


#R1	R2	R3	R4	MW	RingCount	LogP	PSA	HBA	HBD	Violations
OH			O		O		S	241.08	2	0.816	73.580	5	4	0
OH			O		O		NMe	238.13	2	0.106	85.610	5	4	0
OH			O		O		NCF3	292.10	2	0.996	85.610	5	4	0
OH			O		S		O	241.08	2	0.816	73.580	5	4	0
OH			O		S		S	257.05	2	1.754	53.350	5	4	0
OH			O		S		NH	240.09	2	0.783	79.370	5	4	0
OH			O		S		NMe	254.11	2	1.043	65.380	5	4	0
OH			O		S		NCF3	308.08	2	1.933	65.380	5	4	0
OH			O		S		NIndole	397.18	4	3.528	81.170	5	5	0
OH			O		S		NLipoicAcid	428.13	3	3.654	82.450	7	4	0
OH			O		S		NPhOH_0	332.12	3	2.642	85.610	6	5	0
OH			O		S		NPhOH_2	360.15	3	2.362	85.610	6	5	0
OH			O		NH		S	240.09	2	0.783	79.370	5	4	0
OH			O		NMe		O	238.13	2	0.106	85.610	5	4	0
OH			O		NMe		S	254.11	2	1.043	65.380	5	4	0
OH			O		NMe		NMe	251.16	2	0.333	77.410	5	4	0
OH			O		NMe		NCF3	305.14	2	1.223	77.410	5	4	0
OH			O		NCF3		O	292.10	2	0.996	85.610	5	4	0
OH			O		NCF3		S	308.08	2	1.933	65.380	5	4	0
OH			O		NCF3		NMe	305.14	2	1.223	77.410	5	4	0
OH			O		NCF3		NCF3	359.11	2	2.113	77.410	5	4	0
OH			O		NIndole		S	397.18	4	3.528	81.170	5	5	0
OH			O		NLipoicAcid		S	428.13	3	3.654	82.450	7	4	0

#R5	R6	MW	RingCount	LogP	PSA	HBA	HBD	Violations
OH			S	306.05	3	2.628	66.240	6	4	0
NH2			S	304.08	3	2.382	77.820	6	4	0
H			O	242.11	3	1.342	66.240	4	2	0
H			S	274.06	3	3.217	25.780	4	2	0
H			NH	240.14	3	1.275	77.820	4	2	0
H			NMe	268.17	3	1.796	49.840	4	2	0
H			NCF3	376.11	3	3.576	49.840	4	2	0
Cl			O	310.03	3	2.649	66.240	4	2	0
Cl			S	341.98	3	4.524	25.780	4	2	0
Cl			NH	308.06	3	2.582	77.820	4	2	0
Cl			NMe	336.09	3	3.103	49.840	4	2	0
Cl			NCF3	444.03	3	4.883	49.840	4	2	0
CF3			O	378.08	3	3.380	66.240	4	2	0
CF3			NH	376.11	3	3.313	77.820	4	2	0
CF3			NMe	404.14	3	3.834	49.840	4	2	0

(a) **(b)**

Figure 3. Results for PyN_3 (a) and Py_2N_2 (b) frameworks. PyN_3: **18634** to **5828** structures. Py_2N_2: **154** to **34** structures.



(a) **(b)**

Figure 4. 2D representations of the Green group's synthesized compounds PyN_3 (a) and Py_2N_2 (b) based on the RDKit 2D internal representation. These two-dimensional structures are automatically generated by the cheminformatics protocol, and do not perfectly represent the real 3D molecular strucures.

Discussion. Based on the results obtained from the protocol, the two compounds already synthesized by the Green group were among the candidates for synthesis. In recent years, there have been concerns whether Lipinski's rules still hold as a valid screening method for drug candidates. It was found that in 2016 and 2017, the average molecular weight of FDA -approved drugs was larger than 500 g/mol.[4] Moreover, there existed many expansions of Lipinski's rules that could give more accurate screening conditions such as Veber's or Ghose's rules.

CONCLUSION

From the starting 18,634 possible structures for the derivative of the pyridinophane framework PyN_3 , the developed protocol were able to narrow the list of candidates to 5828 structures with 0 violations of the Lipinski's rules. Similarly, the Py_2N_2 framework also yielded 34 best candidates from the original 154 options. In the next stage, 3D structures of these compounds can be generated and subjected to DFT (density functional theory) calculations for more advanced properties such as metal binding ability, pKa and pI, HAT, and redox potentials to further reduce the list to 20 - 30 compounds.

REFERENCES

1. RDKit: Open-source Cheminformatics. https://www.rdkit.org
2. Mandlik V.; Bejugam P. R.; Singh S. Chapter 6 - Application of Artificial Neural Networks in Modern Drug Discovery. Artificial Neural Network for Drug Design, Delivery and Disposition, Academic Press, **2016**, pp. 123-139
3. Araújo de Brito M. Pharmacokinetic study with computational tools in the medicinal chemistry course. *Brazilian Journal of Pharmaceutical Sciences* **2011**, vol. 47
4. Mullard, A. Re-assessing the rule of 5, two decades on. *Nat Rev Drug Discov* 17, 777 (2018). https://doi.org/10.1038/nrd.2018.197