Supporting Information

An Evaluation of 3-Dimensionality in Approved and Experimental Drug Space

Kathleen E. Prosser, Ryjul W. Stokes, Seth M. Cohen*

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093, United States

*scohen@ucsd.edu

Table of Contents

General Procedures

DrugBank Version 5.1.4, structures were downloaded November 20, 2019 as 3D SDF files and loaded in the Molecular Operating Environment (MOE) as a database. Structures were downloaded in their 3D format, generated by the DrugBank[1] using Molconvert by ChemAxon (https://www.chemaxon.com) and used without further modification. Only molecules having molecular weights >100 g/mol were maintained in the library, providing 8532 unique structures. The PMI calculation was carried out on the structures as downloaded without further filtering/optimization. For the DrugBank molecules in the protein databank (PDB), ligand codes associated with each drug were identified using the Drug and Drug Target Mapping tool (https://www.rcsb.org/pdb/ligand/drugMapping.do) accessed on November 20, 2019. For ligands (molecular weights >100 g/mol) all instances from each PDB entry were downloaded, structures were not filtered based on target, providing 7411 structures. The PMI analysis was carried out in MOE for all conformations. The data was reduced by averaging the PMI values of each compound replicate entry to give 502 chemically unique entries. For the analysis of the compound diclofenac, the energy of each conformation was carried out in MOE using AM1_E to give total energy in kcal/mol. These values were then normalized to evaluate the relative changes in energy across the different conformations. Structures for the HIV targeting therapies were accessed directly from the PDB through searching their ligand codes, and the NIH Basic Local Alignment Search Tool (BLAST) was used to determine sequence identity.

Additional PMI Analyses of the DrugBank and PDB Entries

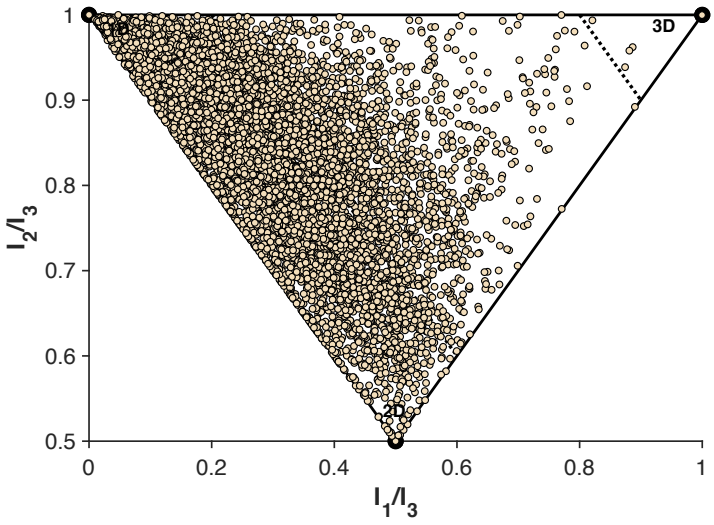

Figure S1. PMI plot for all 8532 structures in the DrugBank downloaded as 3D minimized structures.

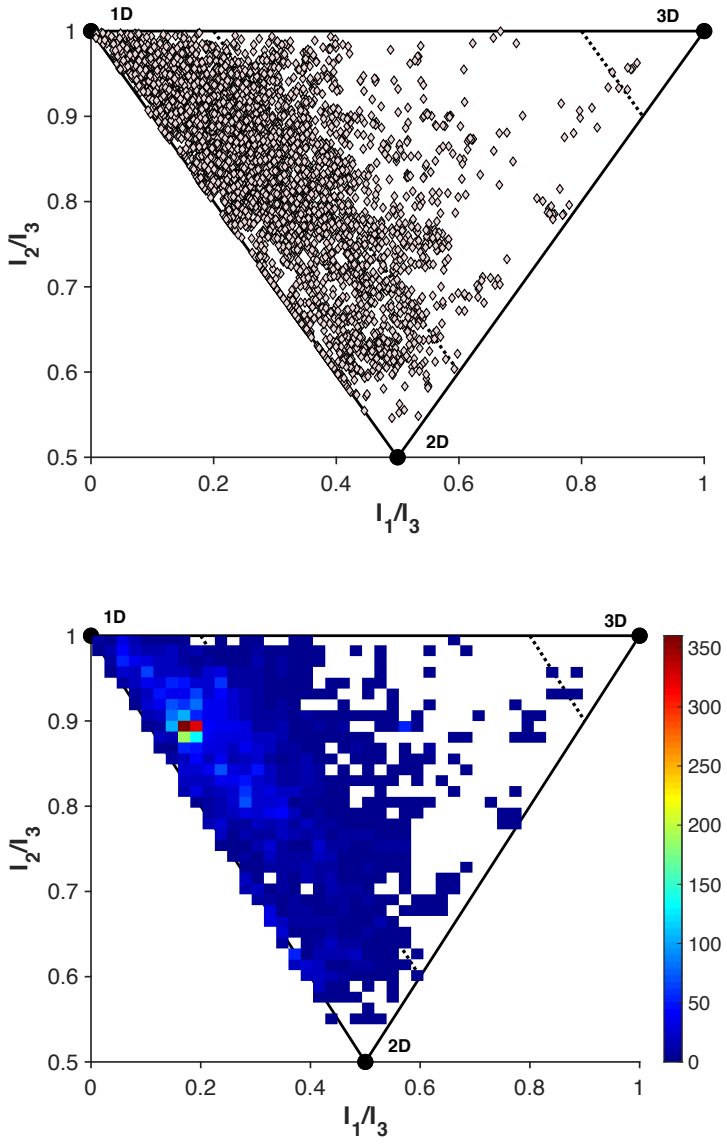

Figure S2. PMI values for PDB ligands that are DrugBank entries, having MW >100 g/mol. N = 7411

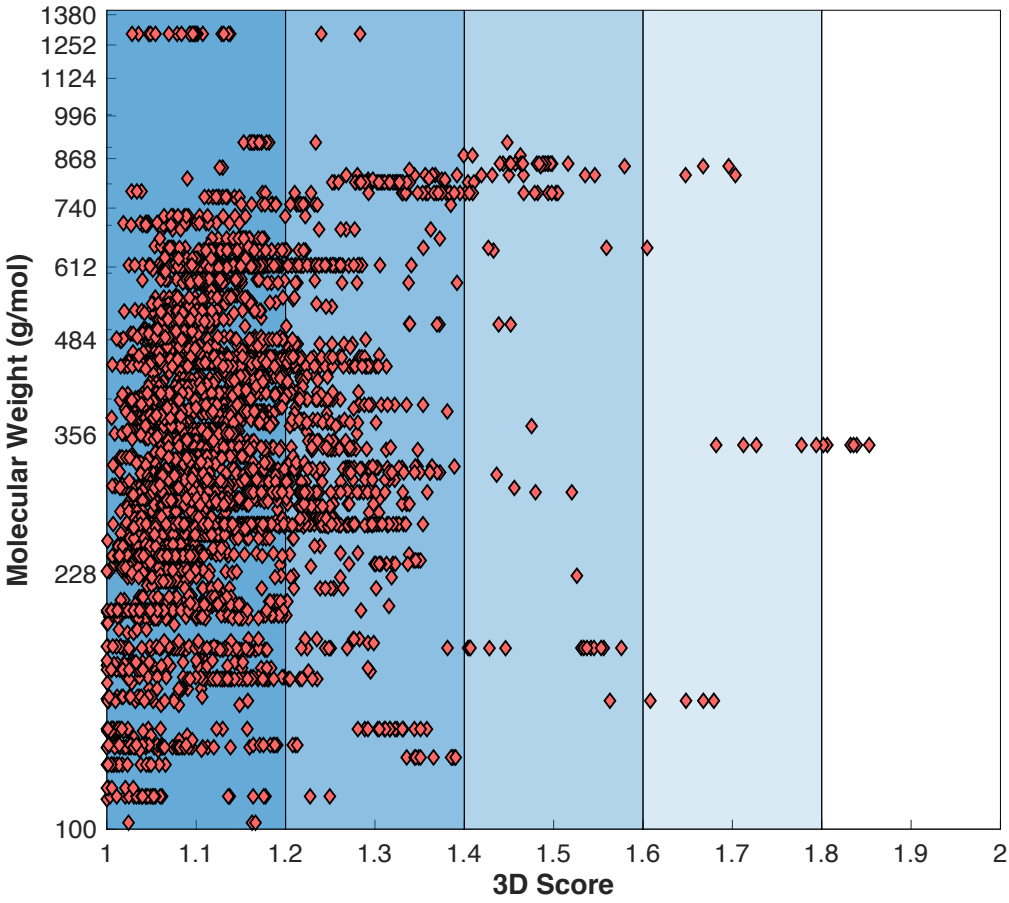

Figure S3. The 7411 entries for ligands in the PDB that have DrugBank identities evaluated by MW and their 3D score.

3D Fragment library chemical structures

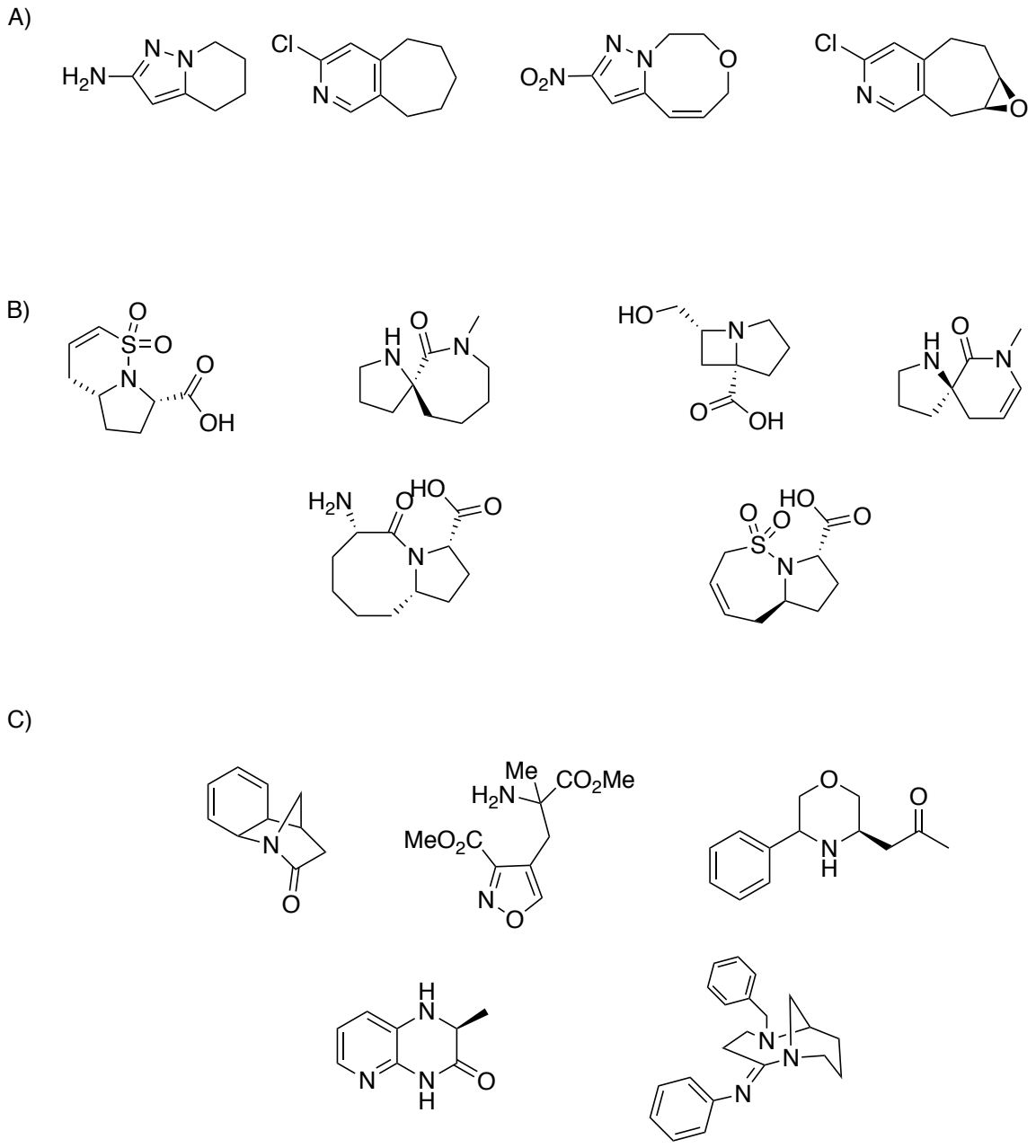

Figure S4. Set of 15 organic fragments from three literature reports (A^2 B^3 C^4) aimed at preparing 3D organic fragments.

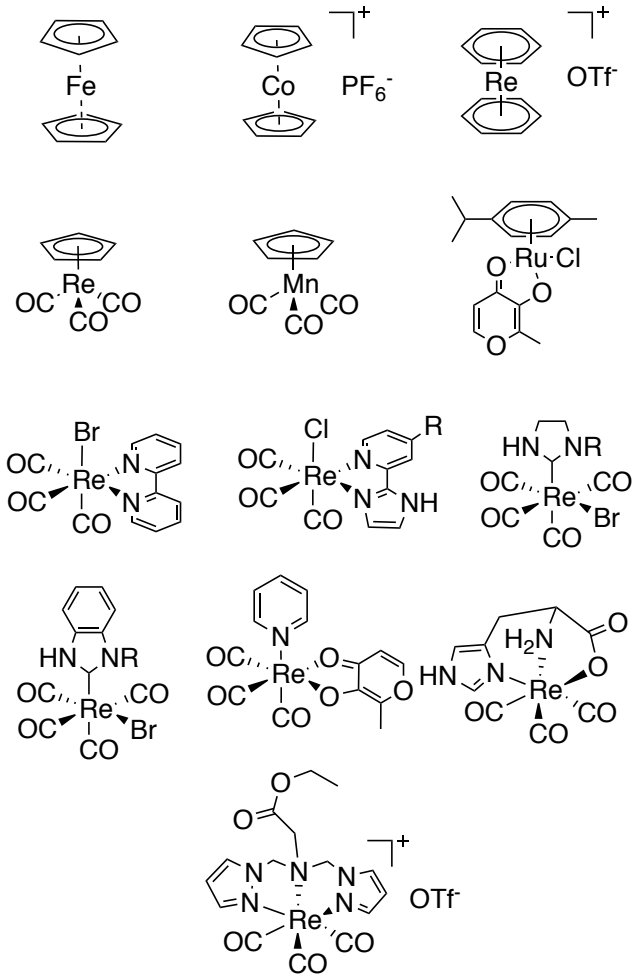

Figure S5. Inorganic metallofragments discussed in this work as having greater 3D topological diversity.[5]

PDB ID Values for HIV Drug Conformations

Table S1. The HIV therapies and the associated PDB ID values used to evaluate therapeutic conformations in the active site. Percent sequence identity for nevirapine and efavirenz-bound structures was calculated against 3LP1 for HIV reverse transcriptases. Percent sequence identity for nelfinavir, lopinavir, and saquinavir-bound structures was calculated against 3EKX for HIV proteases.

Compound	PDB IDs
Nevirapine	1FKP, 1JLB, 1JLF, 1LW0, 1LWC, 1LWE, 1LWF, 1S1U, 1S1X, 1VRT, 2HND, 2HNY, 3HVT, 3LP0, 3LP1, 3QIP, 5HBM.
Efavirenz	1FK9, 1FKO, 1IKV, 1IKW, 1JKH.
Nelfinavir	1OHR, 2PYM, 2PYN, 2Q63, 2Q64, 2QAK, 2R5Q, 3EKX, 3EL0, 3EL5.
Lopinavir	1MUI, 1RV7, 2O4S, 2Q5K, 2QHC, 2RKF, 2RKG, 2Z54, 4L1A, 6DJ1, 6DJ2.
Saquinavir	1C6Z, 1HXB, 2NMY, 2NMZ, 2NNK, 2NNP, 3CYX, 3D1X, 3D1Y, 3EKQ, 3EL4, 3K4V, 3NDT, 3NDU, 3OXC, 3PWR, 3S56, 3TKG, 3TL9, 3UFN, 4QGI, 5KQX, 5KR2.

References:

1. Wishart, D. S.; Feunang, Y. D.; Guo, A. C.; Lo, E. J.; Marcu, A.; Grant, J. R.; Sajed, T.; Johnson, D.; Li, C.; Sayeeda, Z.; Assempour, N.; Iynkkaran, I.; Liu, Y.; Maciejewski, A.; Gale, N.; Wilson, A.; Chin, L.; Cummings, R.; Le, D.; Pon, A.; Knox, C.; Wilson, M., DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* **2018,** *46* (D1), D1074-D1082.

2. Twigg, D. G.; Kondo, N.; Mitchell, S. L.; Galloway, W. R. J. D.; Sore, H. F.; Madin, A.; Spring, D. R., Partially Saturated Bicyclic Heteroaromatics as an sp3-Enriched Fragment Collection. *Angew. Chem., Int. Ed.* **2016,** *55* (40), 12479-12483.

3. Hung, A. W.; Ramek, A.; Wang, Y.; Kaya, T.; Wilson, J. A.; Clemons, P. A.; Young, D. W., Route to three-dimensional fragments using diversity-oriented synthesis. *PNAS* **2011,** *108* (17), 6799.

4. Kidd, S. L.; Osberger, T. J.; Mateu, N.; Sore, H. F.; Spring, D. R., Recent Applications of Diversity-Oriented Synthesis Toward Novel, 3-Dimensional Fragment Collections. *Front. Chem.* **2018,** *6*, 460.

5. Morrison, C. N.; Prosser, K. E.; Stokes, R. W.; Cordes, A. L.; Metzler-Nolte, N.; Cohen, S. M., Expanding Medicinal Chemistry into 3D Space: Metallofragments as 3D Scaffolds for Fragment-Based Drug Discovery. *Chem. Sci.* **2020,** *11*, 1216-1225.