Supplementary Material

Target Specific De Novo Design of Drug Candidate Molecules with Graph Transformer-based Generative Adversarial Networks

Atabey Ünlü[1,2], Elif Çevrim[1,2], Ahmet Sarıgün[3,4], Hayriye Çelikbilek[1], Heval Ataş Güvenilir[1,5], Altay Koyaş[5], Deniz Cansen Kahraman[5], Abdurrahman Olğaç[6], Ahmet Rifaioğlu[7,8], Tunca Doğan[1,2,*]

[1]Biological Data Science Lab, Dept. of Computer Engineering, Hacettepe University,
[2]Dept. of Bioinformatics, Graduate School of Health Sciences, Hacettepe University,
[3]Dept. of Chemistry, Middle East Technical University
[4]Dept. of Physics, Middle East Technical University
[5]Cancer Systems Biology Lab, Graduate School of Informatics, Middle East Technical University
[6]Dept. of Pharmaceutical Chemistry, Faculty of Pharmacy, Gazi University
[7]Dept. of Electrical and Electronics Engineering, İskenderun Technical University, Hatay, Turkey
[8]Heidelberg University, Faculty of Medicine, and Heidelberg University Hospital, Institute for Computational Biomedicine, Bioquant, Heidelberg, Germany

*To whom the correspondence should be addressed: Tunca Doğan (tuncadogan@gmail.com)

Supplementary Figures

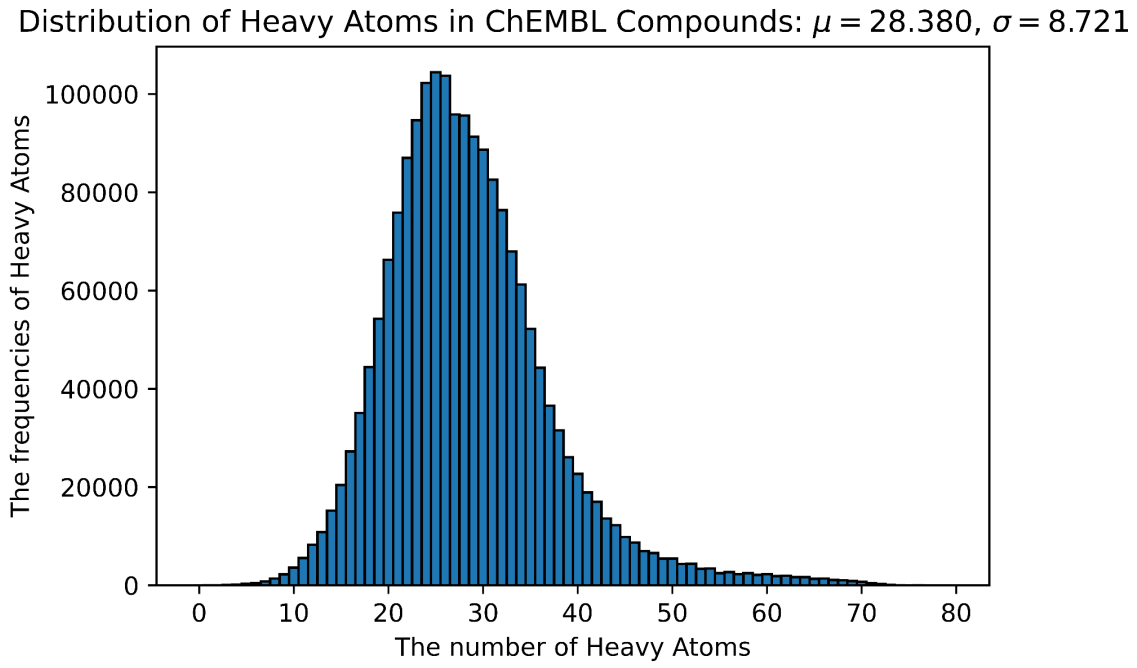


Fig. S1. The heavy atom distribution histogram of the small molecules in the ChEMBL (v29…) database.

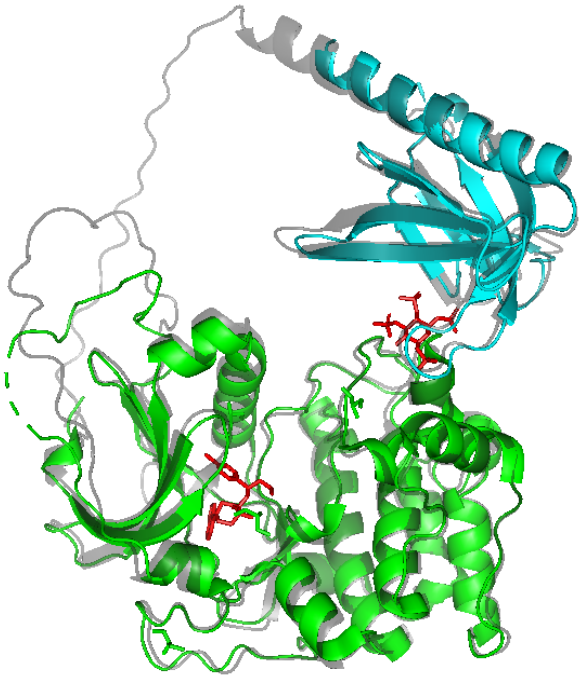


Fig. S2. General overview of AKT1 protein structure (AlphaFold structure based on Uniprot ID: P31749: gray, kinase domain: green (PDB: 4GV1), PH domain: cyan (PDB: 1H10), ligands: red).

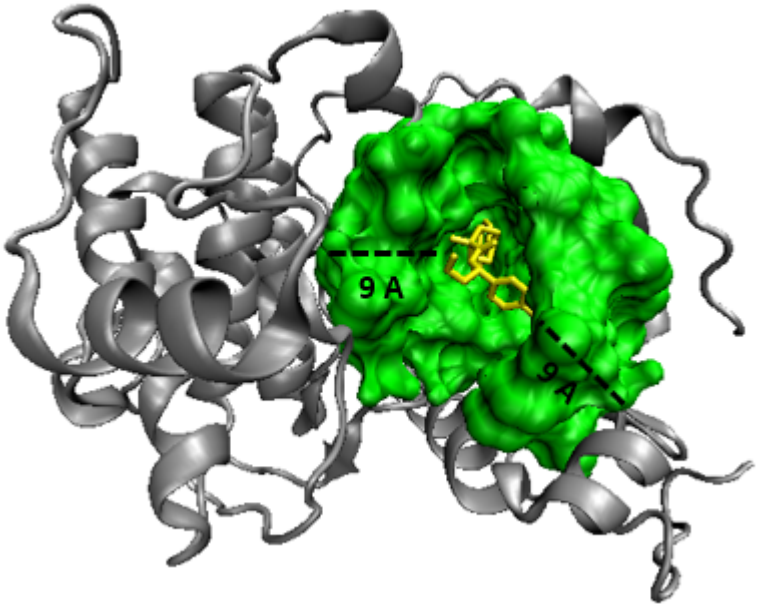


Fig. S3. Representation of the binding pocket of AKT1 including all atoms located at a maximum distance of 9 Angstroms from the ligand atoms; binding site: green, ligand: yellow, the rest of the kinase domain: gray (PDB: 4GV1).

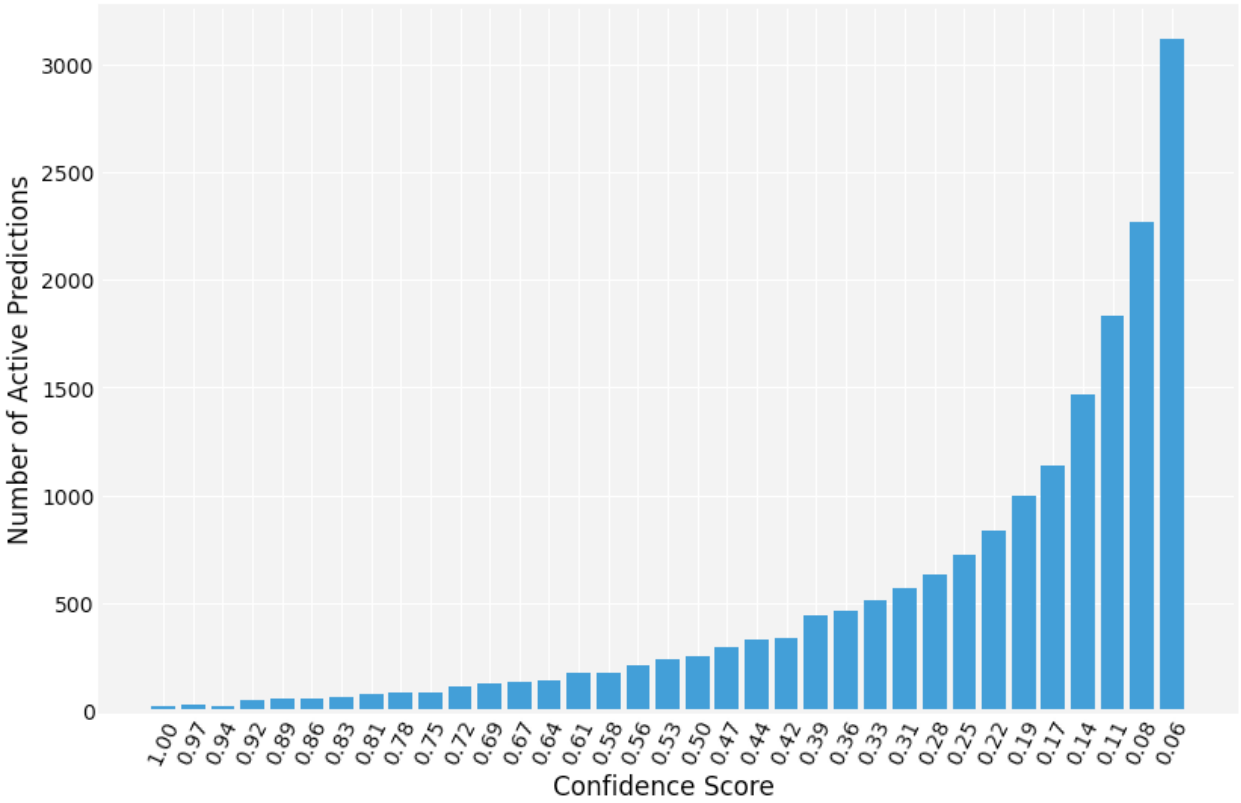
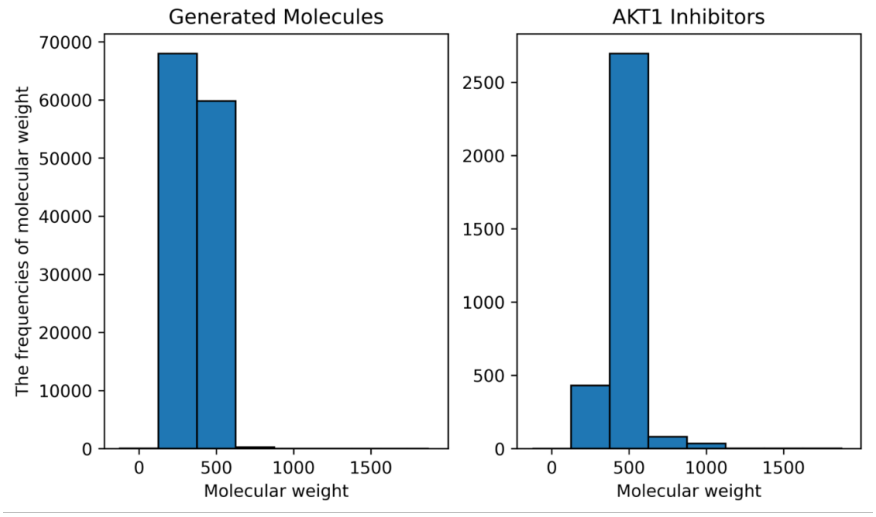
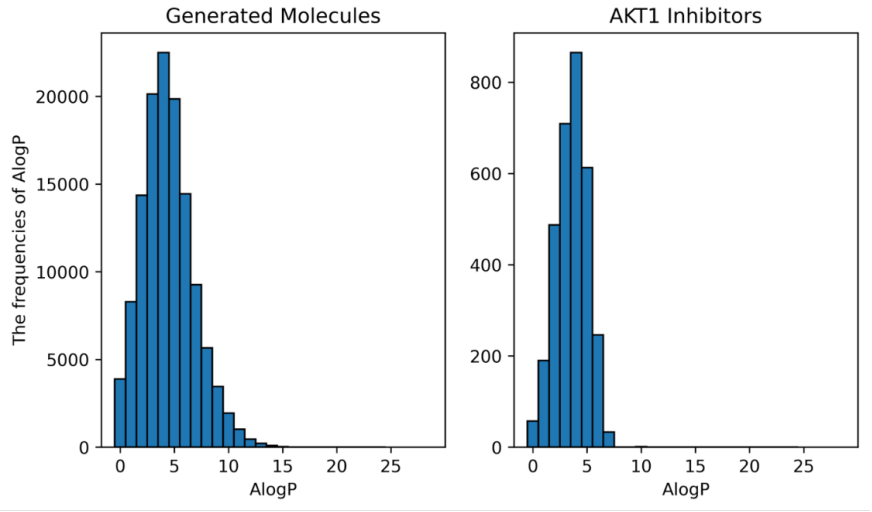
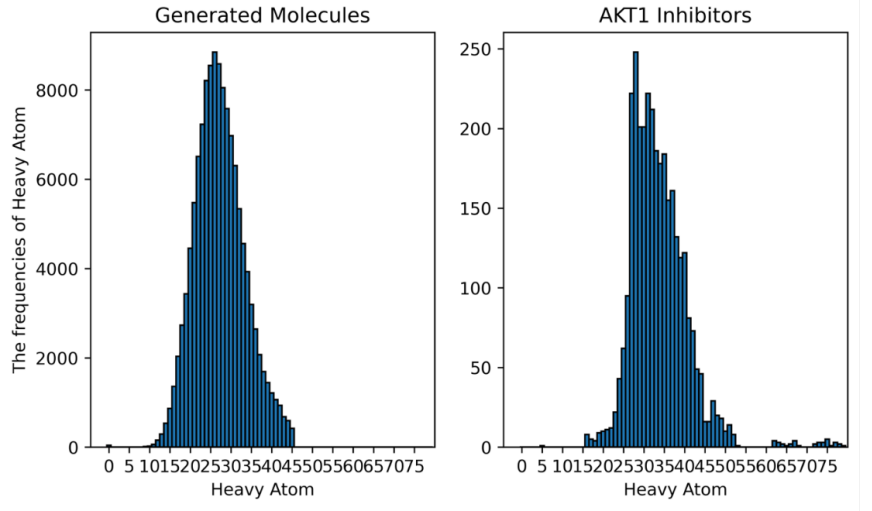
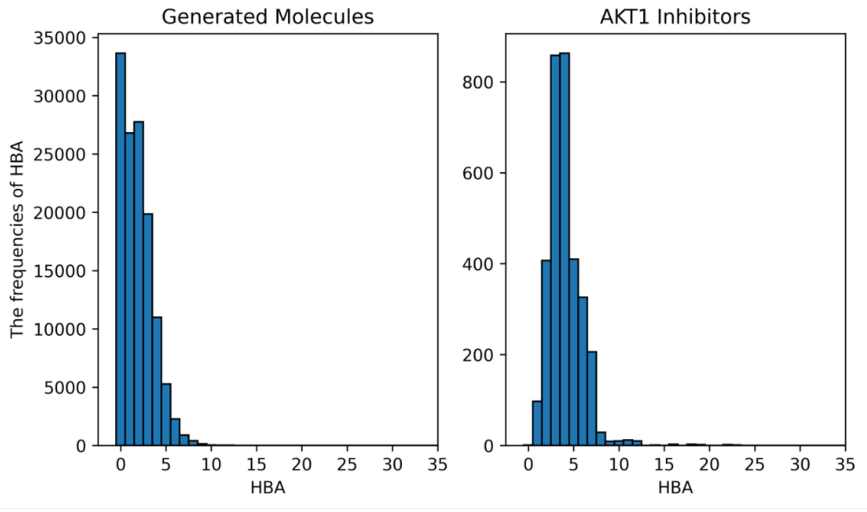


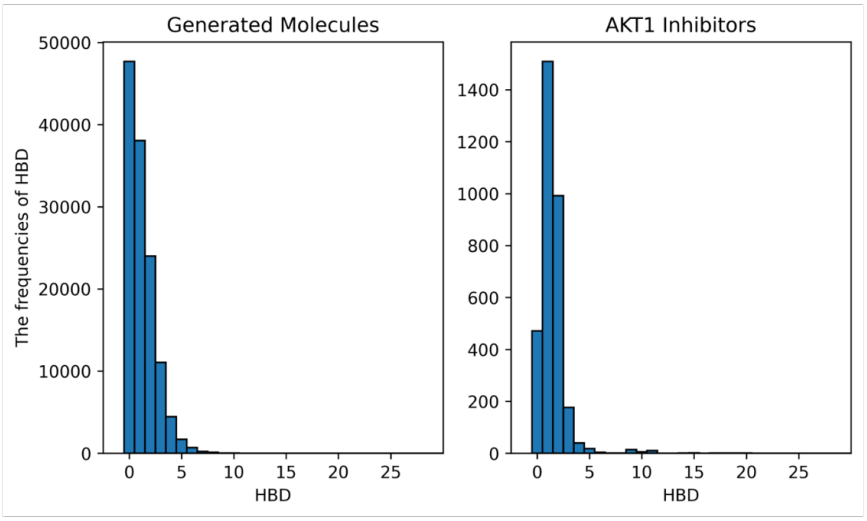
Fig. S4. Drug-target interaction prediction confidence level histograms of 18,114 compounds that are predicted to be bioactive against the AKT1 protein (out of 40,000 DrugGEN generated de novo molecules). The results are produced by the DEEPScreen system (see section 2.8 in the main text).

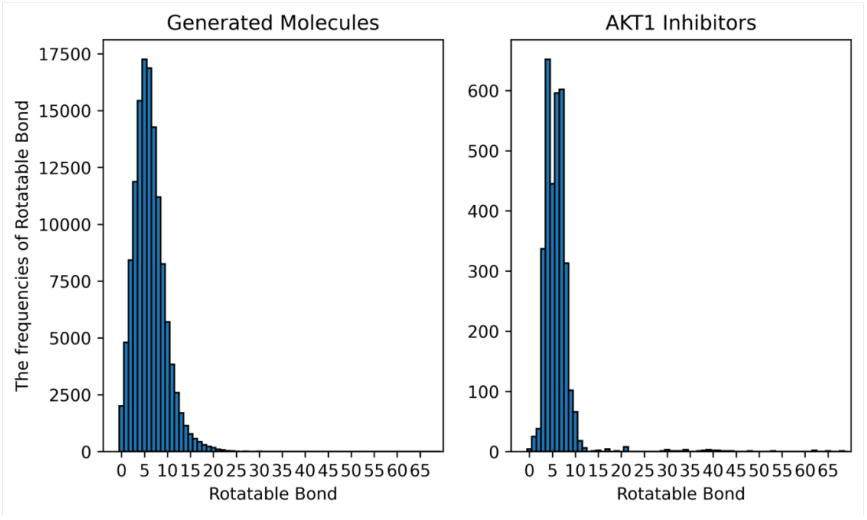

Generated Molecules
AKT1 Inhibitors
The frequencies of molecular weight
Molecular weight
Molecular weight


Generated Molecules
AKT1 Inhibitors
The frequencies of AlogP
AlogP
AlogP


Generated Molecules
AKT1 Inhibitors
The frequencies of Heavy Atom
Heavy Atom
Heavy Atom


Generated Molecules
AKT1 Inhibitors
The frequencies of HBA
HBA
HBA


Generated Molecules
AKT1 Inhibitors
The frequencies of HBD
HBD
HBD


Generated Molecules
AKT1 Inhibitors
The frequencies of Rotatable Bond
Rotatable Bond
Rotatable Bond

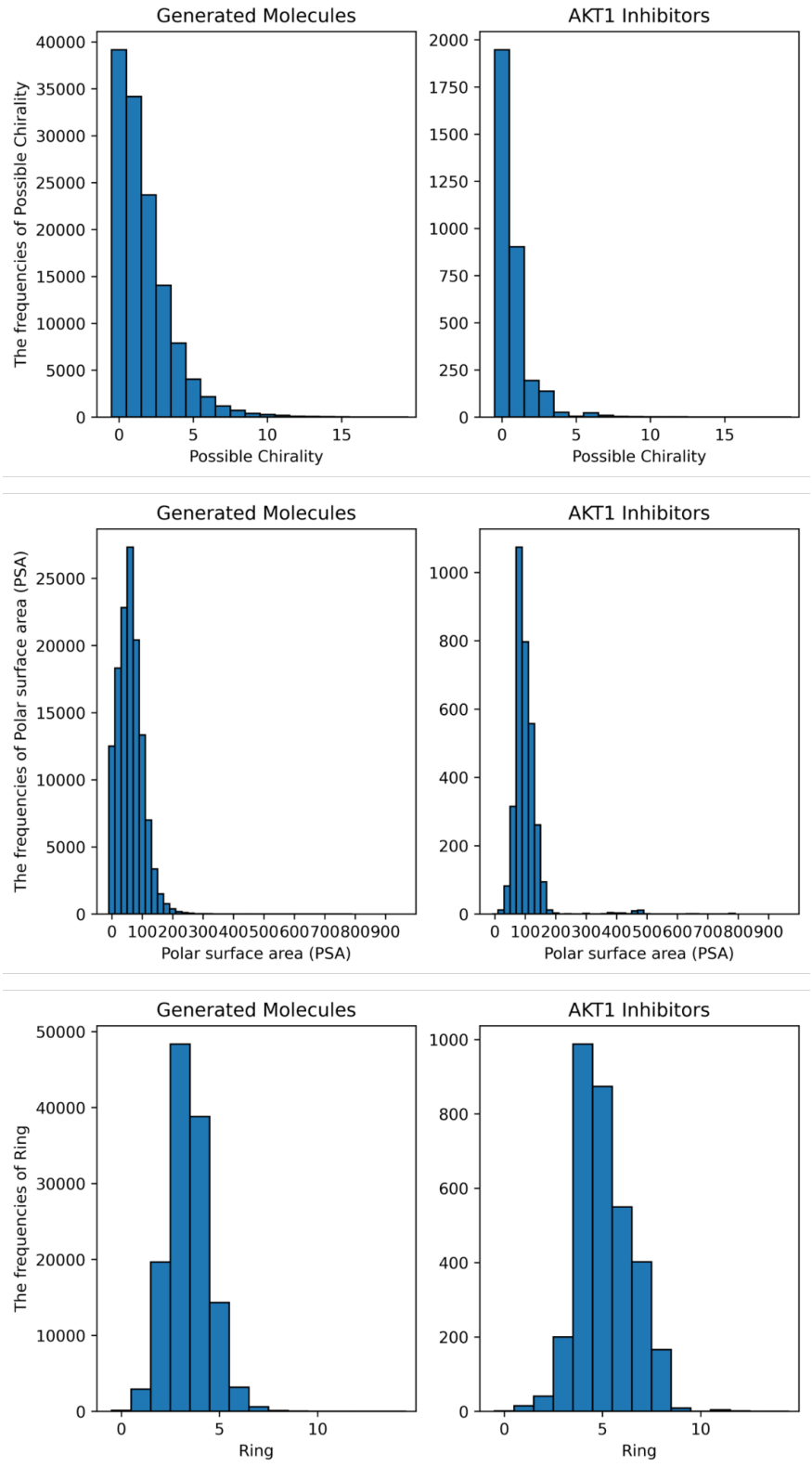


Fig. S5. Physicochemical properties of AKT1-specific de novo molecules generated by DrugGEN models and real AKT1 inhibitors.

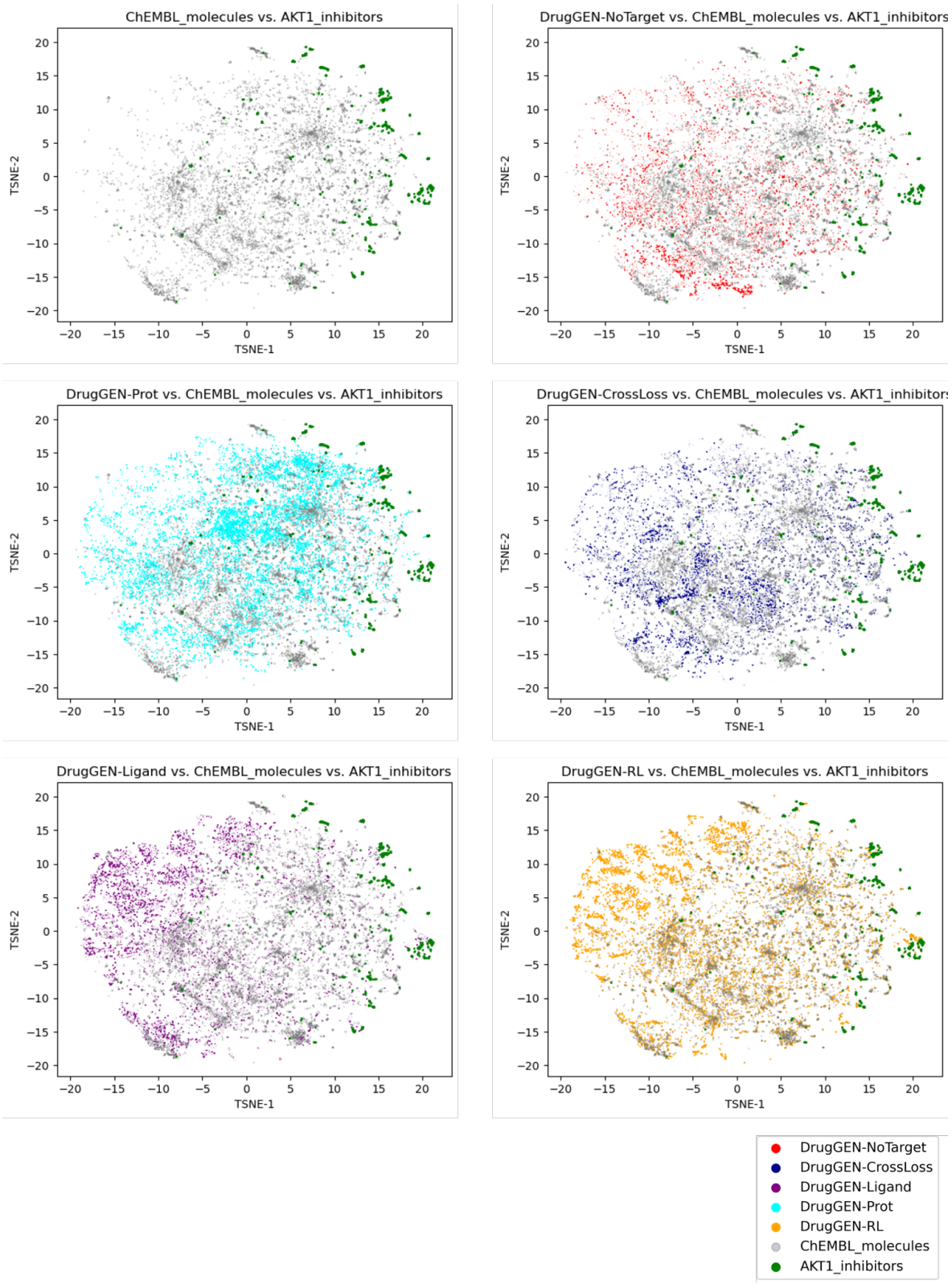


Fig. S6. Individual 2-D visualizations of each DrugGEN molecules compared with ChEMBL molecules and known/real AKT1 inhibitors, in the same t-SNE embedding as given in Figure 2B.