Supporting Information

SPOTLIGHT: Structure-based Prediction and Optimization Tool for LIgand Generation on Hard-to-drug Targets - Combining Deep Reinforcement Learning with Physics-based de novo drug design
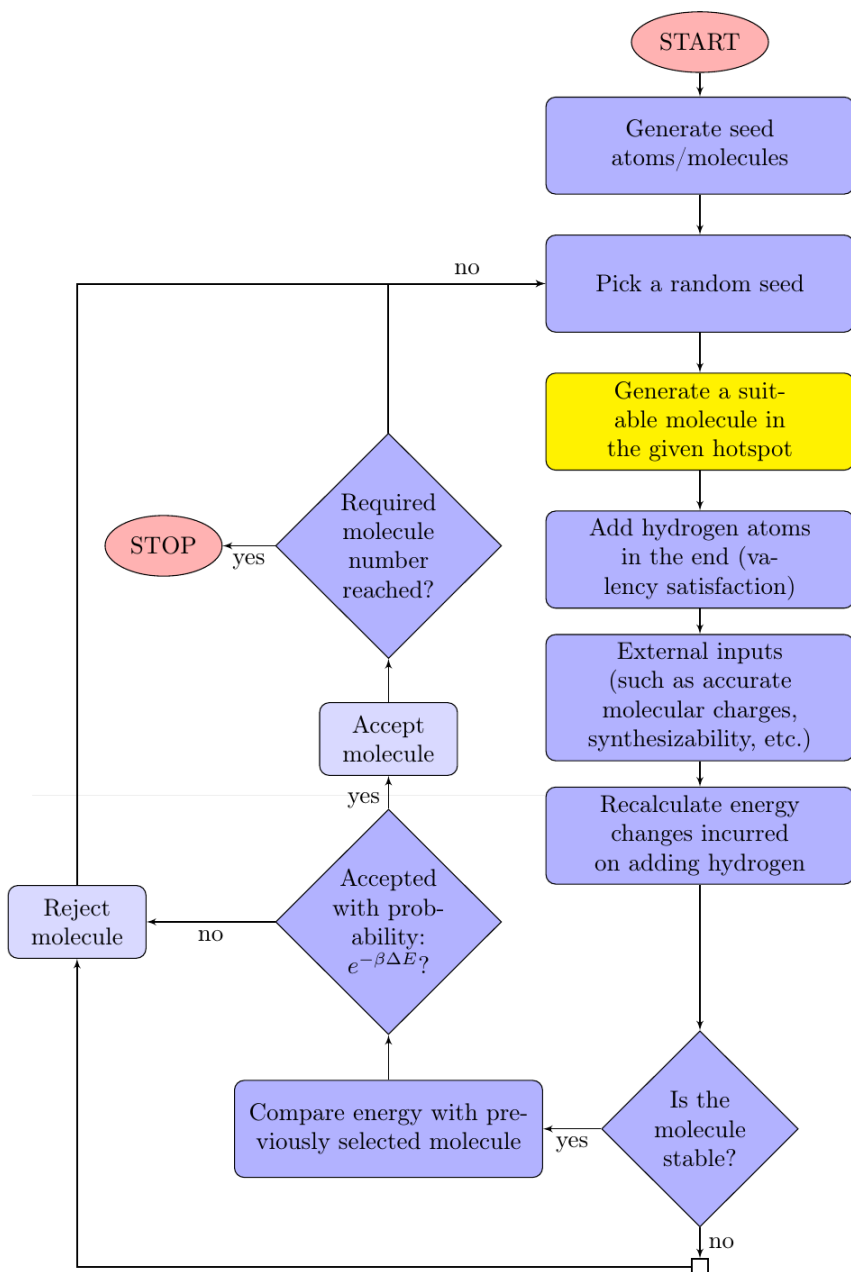
Venkata Sai Sreyas Adury and Arnab Mukherjee*

Figure S1 : Broad outline of the SPOTLIGHT algorithm's steps. Finer details of the atomistic construction are given in figure 2 below.
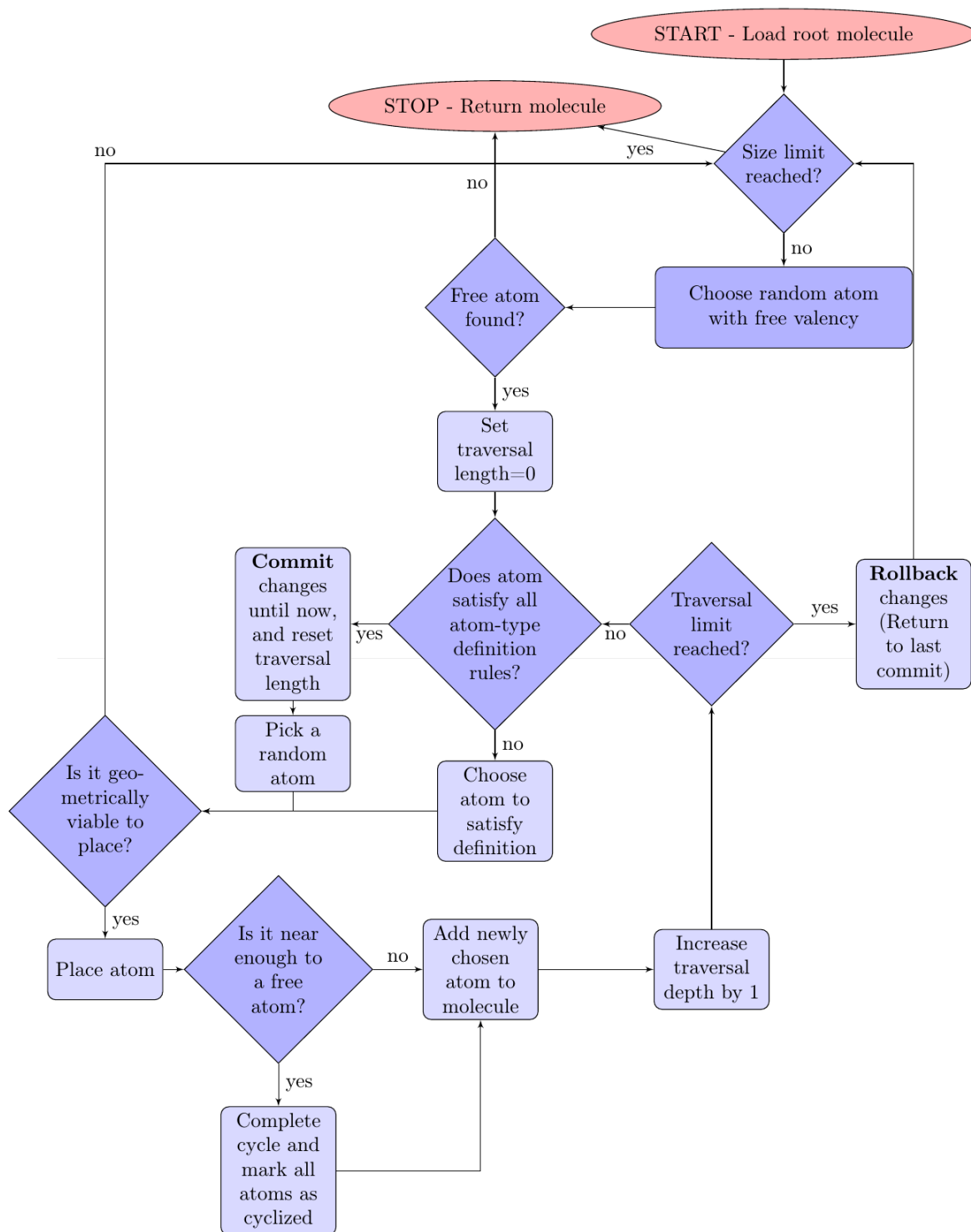
Figure S2 : A flowchart of the finer details of atom-level construction of molecules.
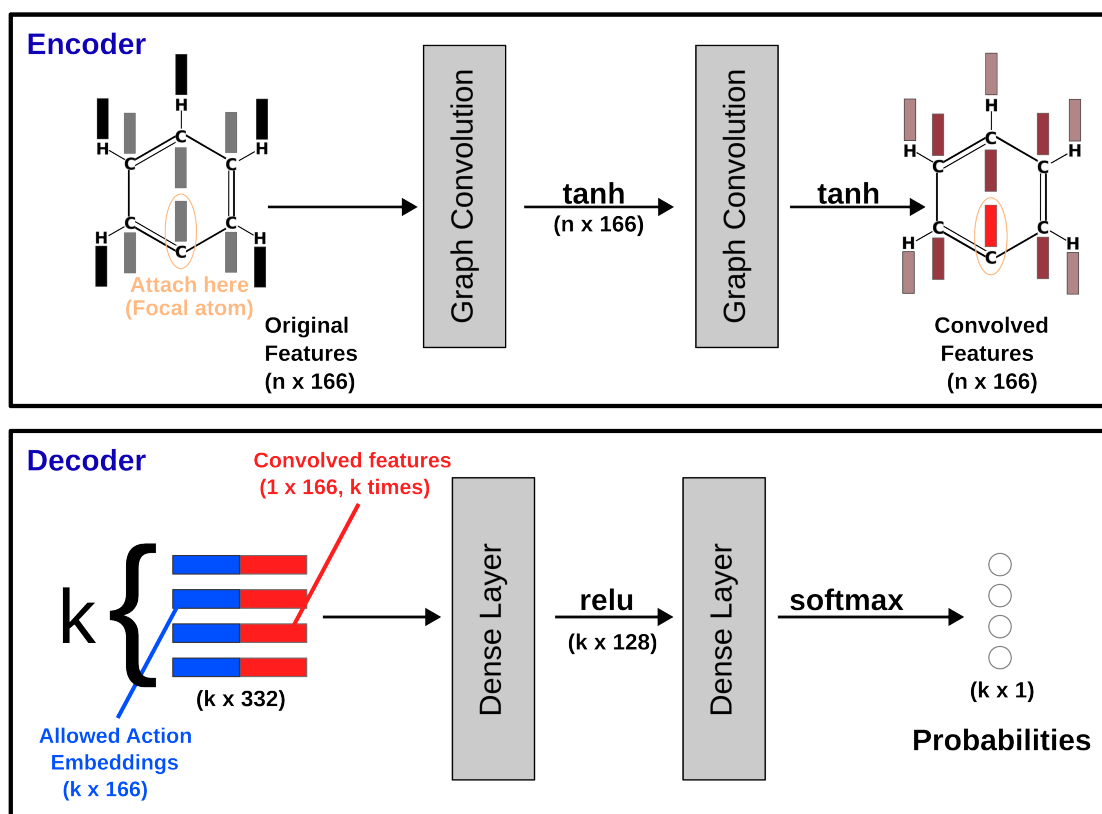
Figure S3 : Model Scheme for the encoder and decoder. The encoder uses the tanh activation function and the decoder uses the relu activation function. To generate the input for the decoder, the convolved features for the focal atom are *repeated k* times and attached to the action embeddings for each "allowed" action.
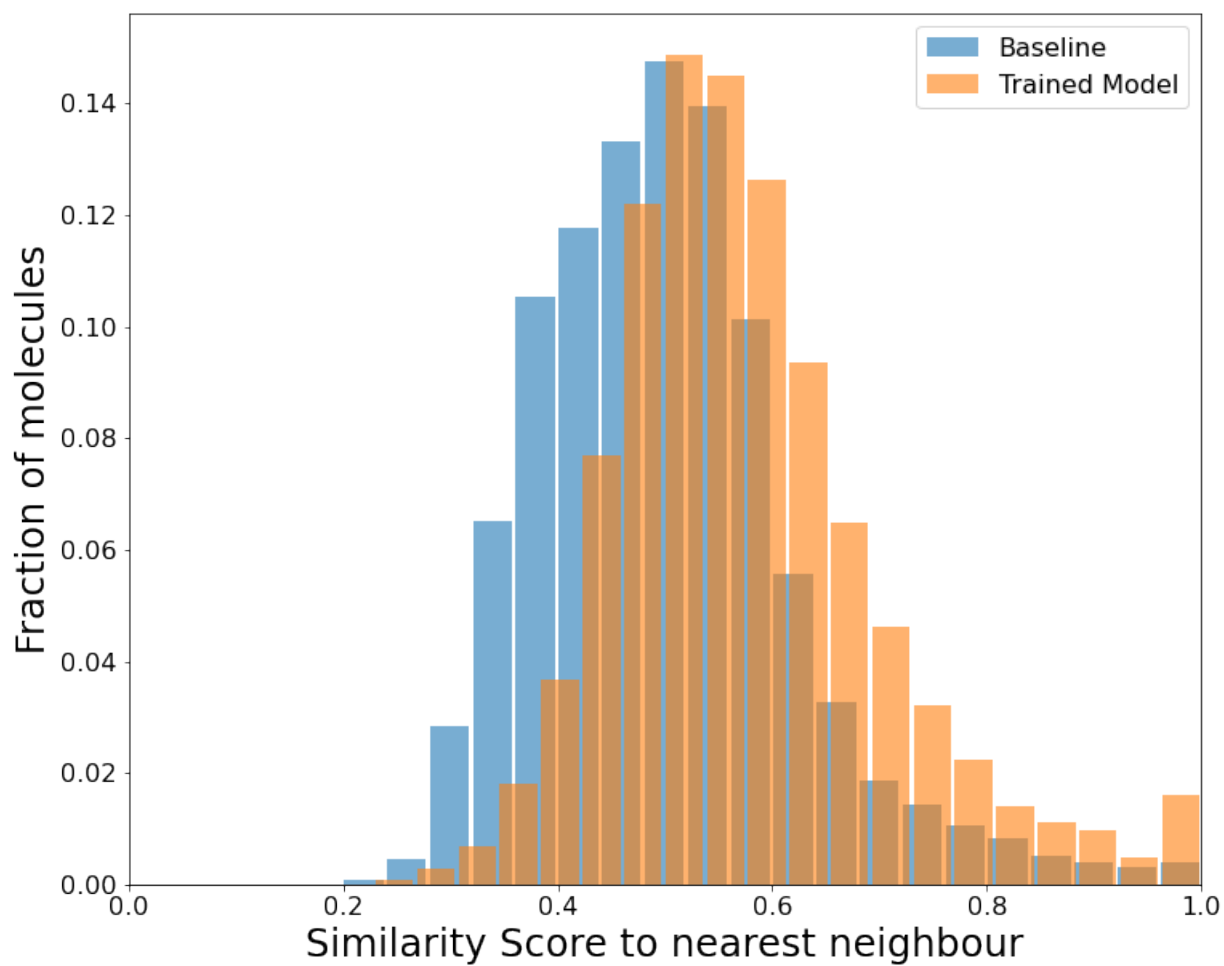
Figure S4 : Self-similarity distributions with 12000 molecules. This is expected to bring out long-range similarities. After training the mean similarity shifts from 0.5 to 0.58
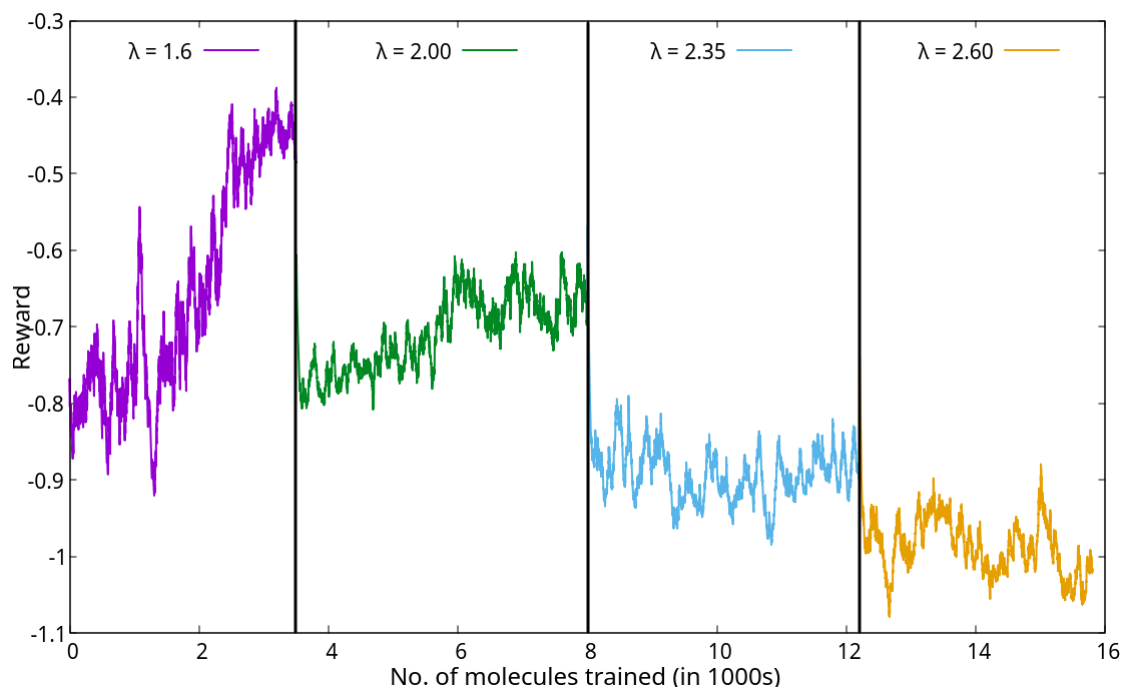
Figure S5 : Evolution of the reward function while training. $\lambda = 1.6$ was used for the first 3500 molecules and slowly increased after every 4500 molecules thereafter. The vertical lines indicate the separated instances of training when λ was adjusted. Towards the end, the reward function stabilizes.
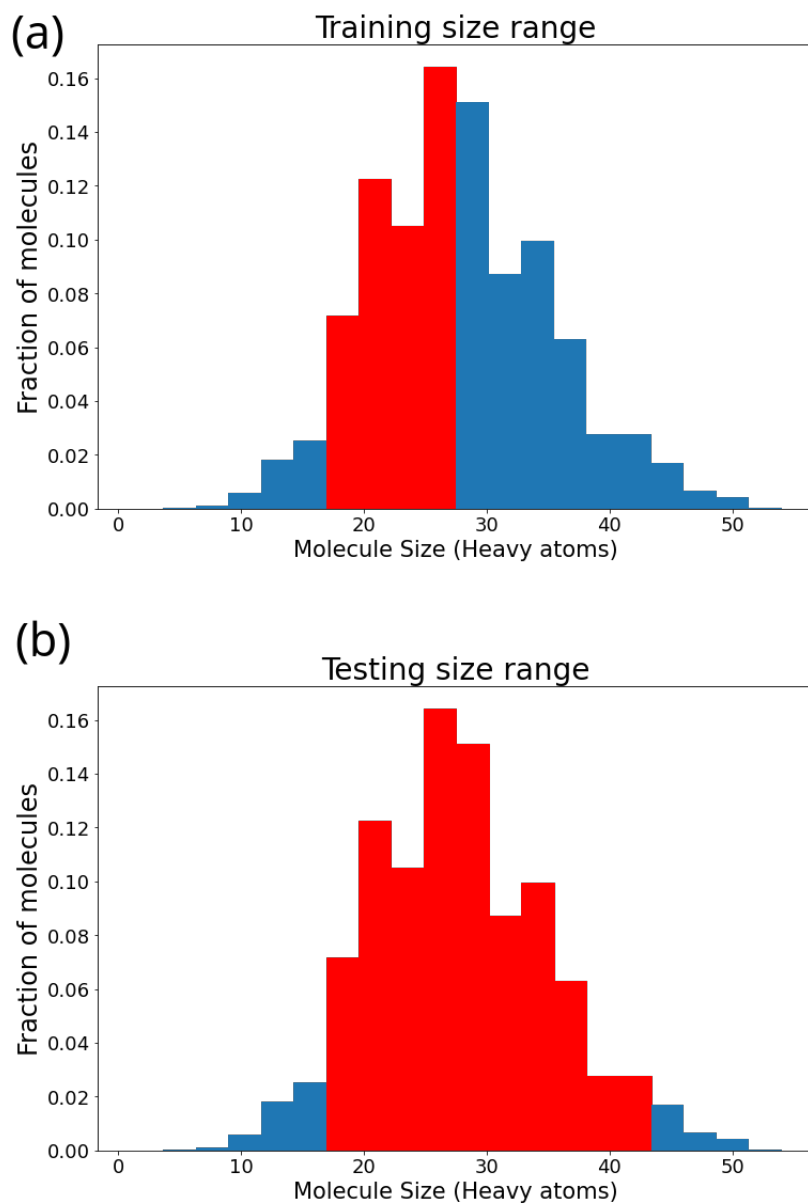
Figure S6 : The plots are the distribution of the sizes of molecules in the CheMBL database. The red-colored regions indicate the size ranges targeted for training and testing respectively. We trained on smaller size ranges, but ensured that we included the peak of the CheMBL size distribution. The model was validated on larger sizes as well (upto a target size of 41).
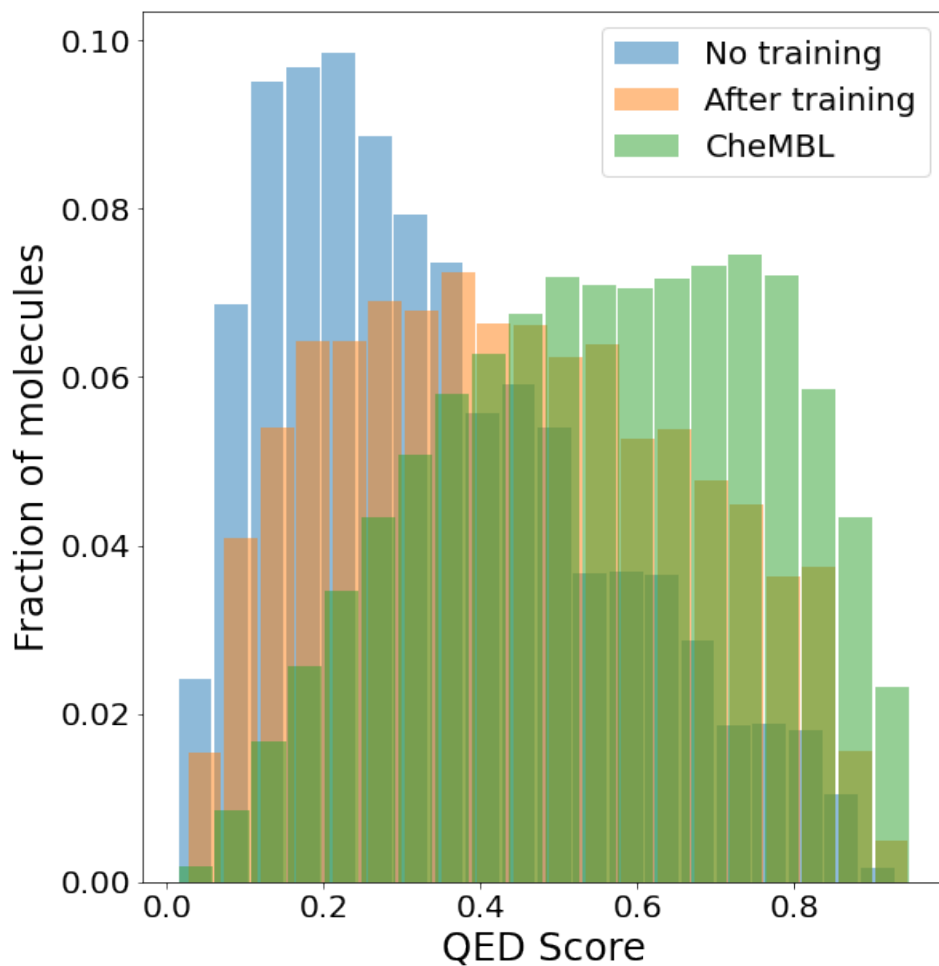
Figure S7 : Distribution of QED scores after training (2-layer model) as compared to no training. The mean shifts from 0.34 ± 0.2 to 0.44 ± 0.22. To 39% of the molecules have a score above 0.5 after training, compared to only 22% before. For reference, the green distribution shows the QED scores for the 1.8 million molecules from CheMBL. The median QED score is 0.56, which is surpassed by 31% molecules after training, nearly doubling from the original 16.6% before.
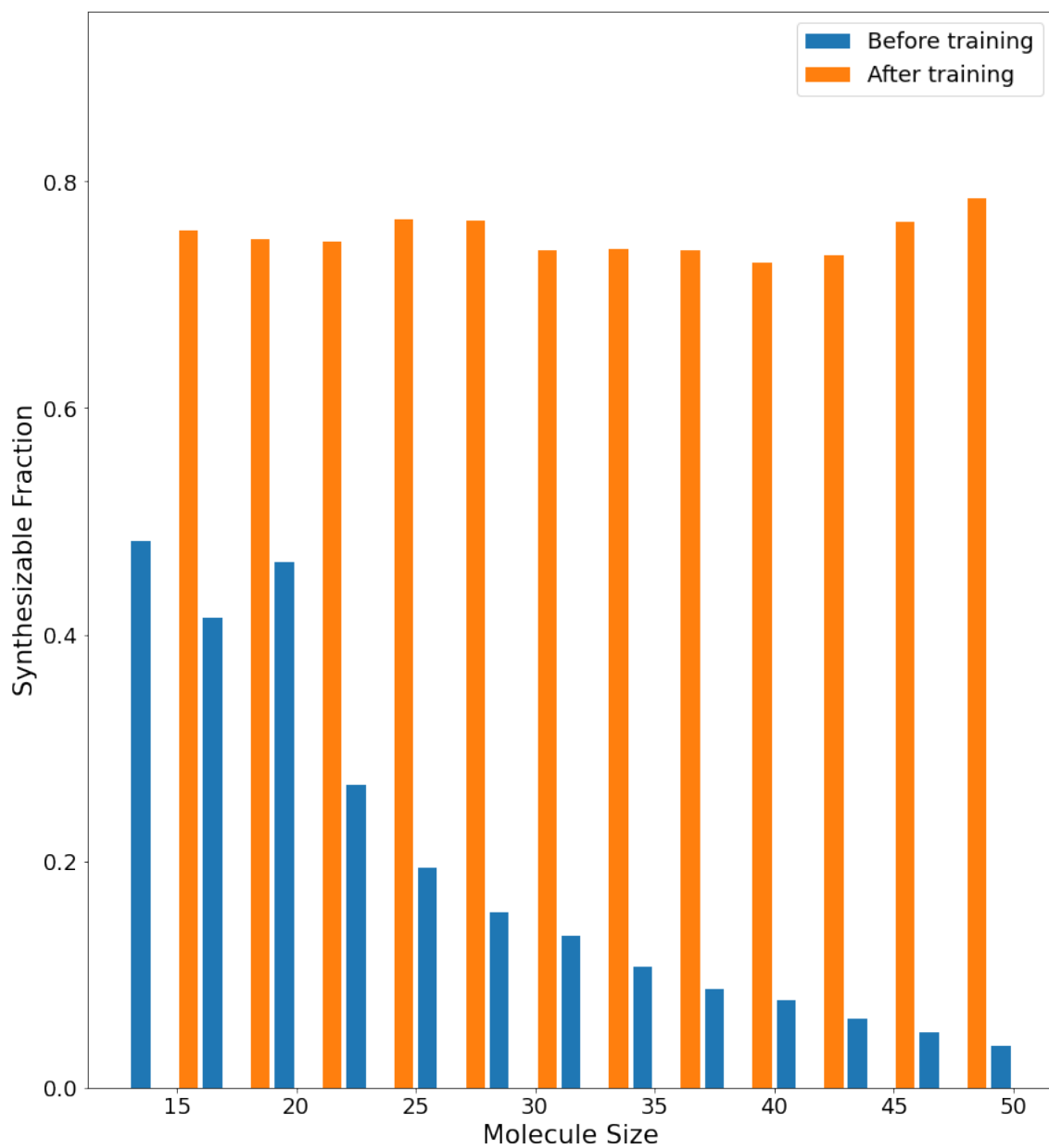
Figure S8 : Prediction of synthesizability by size of the molecule. Originally, larger molecules were found to be harder to synthesize, owing to the high odds of picking hard-to-synthesize atoms/linkages. After training, the odds are strictly higher and roughly steady for all molecule sizes.
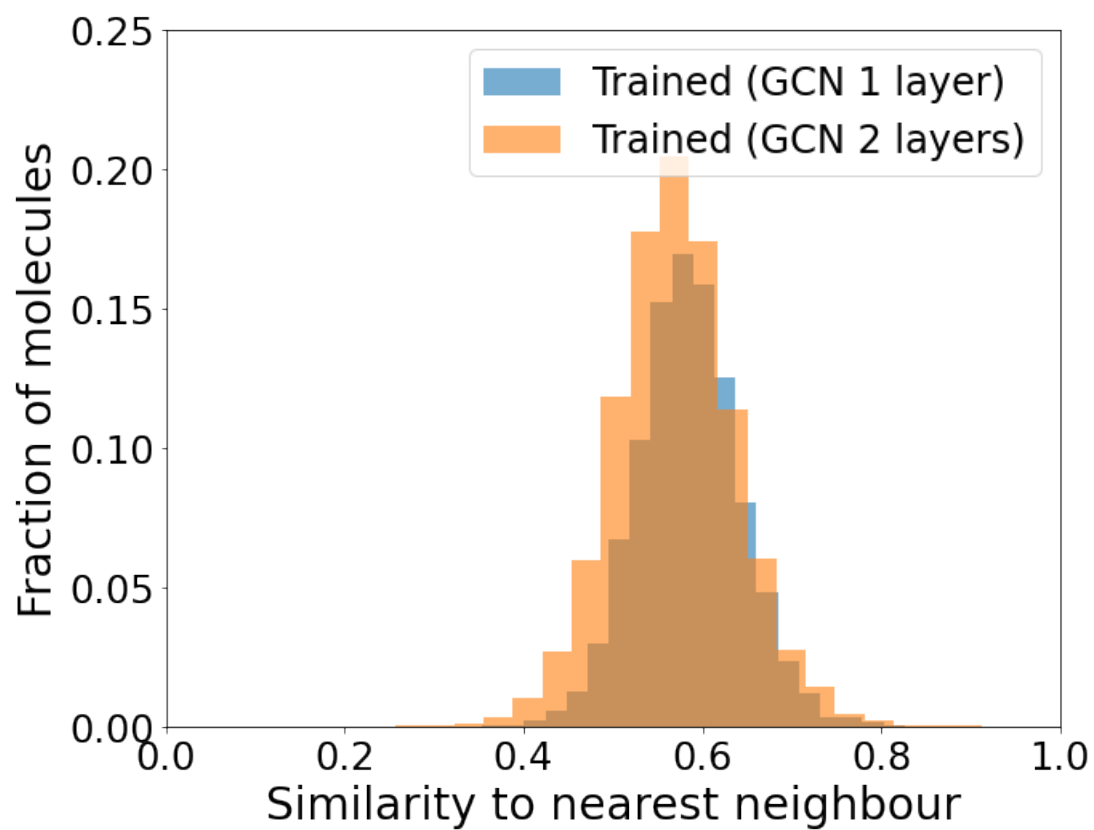
Figure S9 : Self-similarity for 12000 molecules using both the optimized model architectures - 1 layer and 2 layer. With only one layer, the SSD is 0.585 ± 0.06, which is comparable to the 0.569 ± 0.07
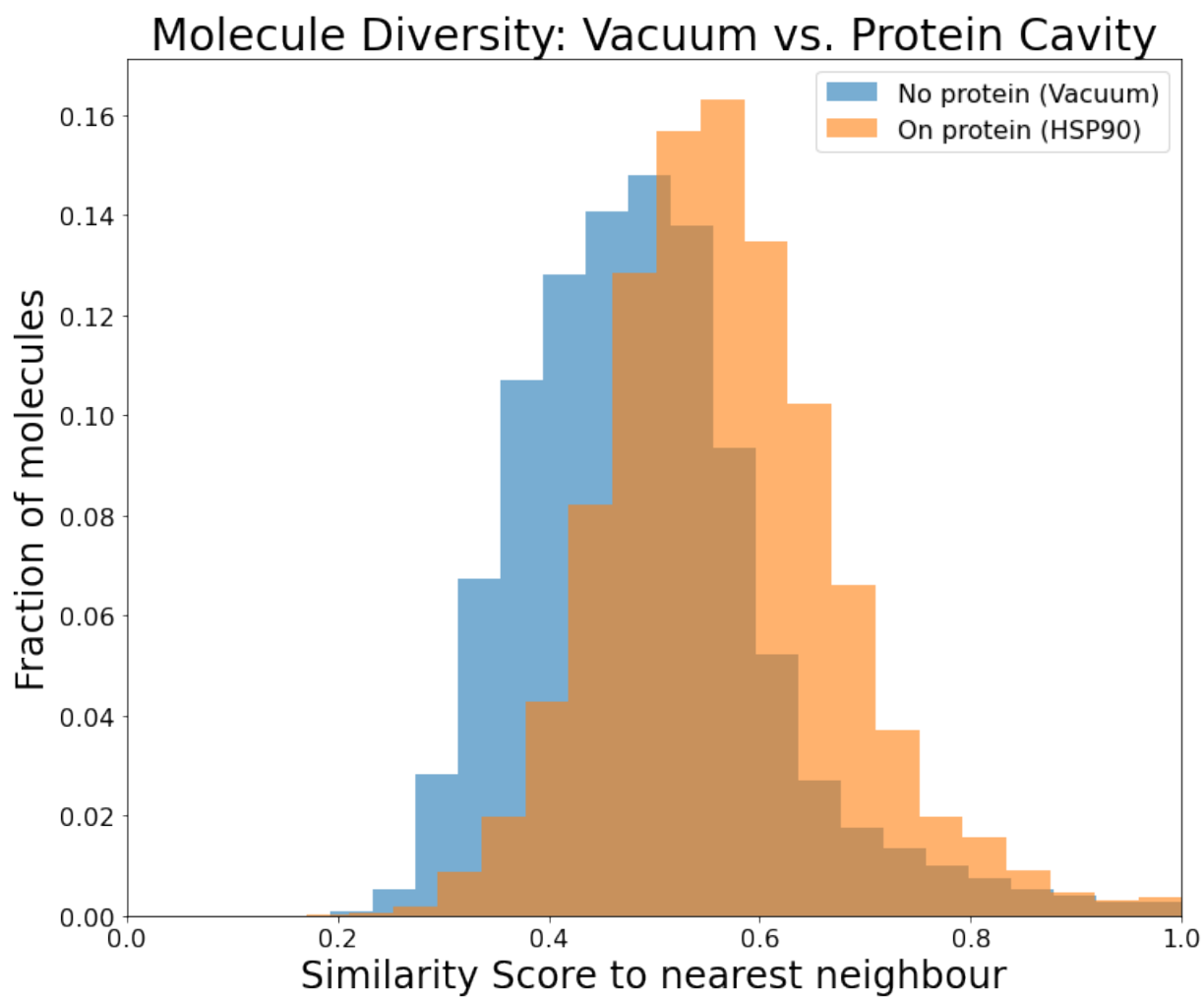
Figure S10 : Self-similarity distributions with 8500 molecules comparing vacuum generation with
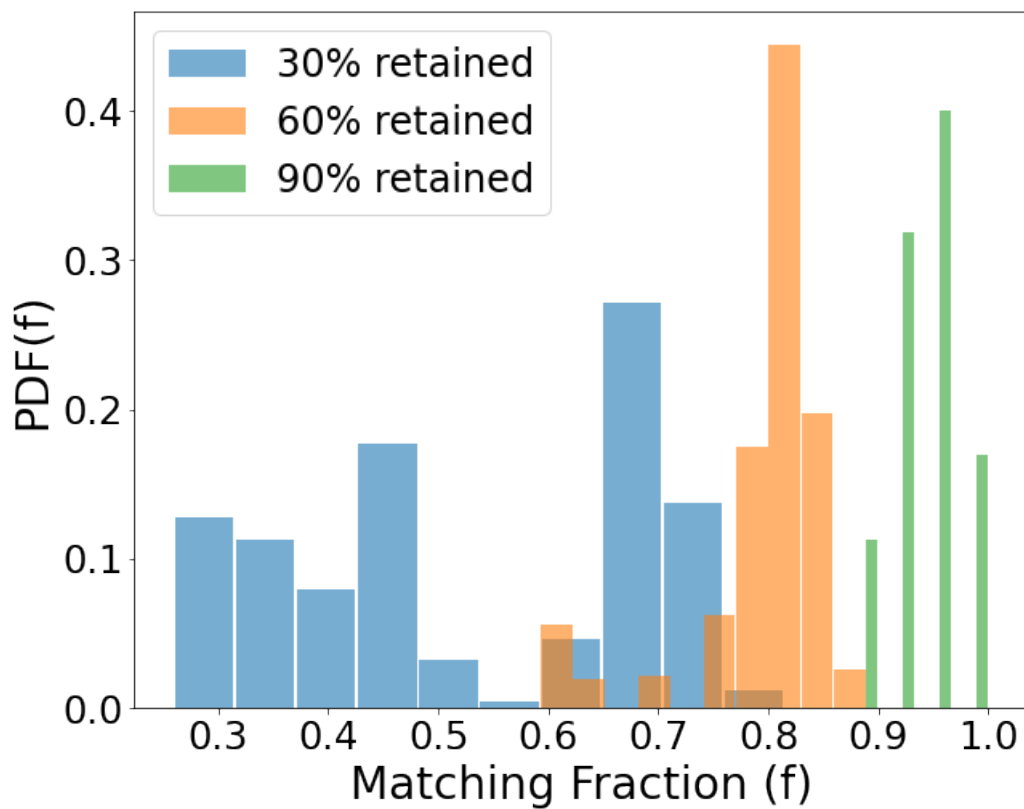the trained model to the same on HSP90. After training the mean similarity shifts from 0.5 to 0.57

Figure S11 : Distribution of the regenerated fraction for the ligands at different retention fractions. Higher retention fractions clearly reproduce higher retention, with lower variance in the final retention fraction as shown here.