

# **Predicting Binding from Screening Assays with Transformer Network Embeddings**

Paul Morris, Rachel St. Clair, William Edward Hahn, and Elan Barenholtz

Center for Complex Systems and Brain Sciences

Florida Atlantic University

Boca Raton, Florida 33431, United States

E-mail: [pmorris2012@fau.edu](mailto:pmorris2012@fau.edu)

# Supporting Information

---

## SI #1: Functional groups for deletion and swapping reactions in RDKit

'C=O', '[H+]', '[CX3]=[OX1]', '[CX3](=[OX1])C', '[OX1]=CN', '[CX3](=[OX1])O',  
'[CX3](=[OX1])[F,Cl,Br,I]', '[CX3H1](=O)[#6]', '[CX3](=[OX1])[OX2][CX3](=[OX1])',  
'[NX3][CX3](=[OX1])[#6]', '[NX3][CX3]=[NX3+]', '[NX3,NX4+][CX3](=[OX1])[OX2,OX1-]',  
'[CX3](=O)[OX2H1]', '[NX3][CX2][NX1]', '[#6][CX3](=O)[#6]', '[OD2](#[6])[#6]',  
'[NX3;H2,H1;!\$(NC=O)]', '[NX3,NX4+][CX4H]([\*])[CX3](=[OX1])[O,N]',  
"\$([\*-NX2-]-NX2+)[#NX1])\$([\*-NX2]=[NX2+]=[NX1-])", "[NX2]=N", '[NX3][NX3]',  
'[NX3][NX2]=[\*]', '[CX3;\$([C](#[6])[#6])\$([CH](#[6]))=[NX2][#6]', '[NX3+]=[CX3]',  
'[CX3](=[OX1])[NX3H][CX3](=[OX1])',  
"\$([NX3](=[OX1])(=[OX1])O)\$([NX3+](=[OX1-])(=[OX1])O)", '[NX1][CX2]',  
"\$([NX3](=O)=O)\$([NX3+](=O)[O-])[!#8]", '[NX2]=[OX1]',  
"\$([#7+][OX1-])\$([#7v5]=[OX1]);!\$([#7](~[O])~[O]);!\$([#7]=[#7])", '[OX2H]', '\$([OH]-\*=[!#6])',  
'[OX2,OX1-][OX2,OX1-]',  
"\$([P](=[OX1])([OX2H])\$([OX1-])\$([OX2P]))([OX2H])\$([OX1-])\$([OX2P]))\$([OX2H])\$([OX1-])\$([OX2P]))\$([P+][OX1-])([OX2H])\$([OX1-])\$([OX2P]))([OX2H])\$([OX1-])\$([OX2P]))\$([#16X2H])\$([#16X2H0])\$([#16X3](=[OX1])[OX2H0])\$([#16X3+](=[OX1-])[OX2H0])\$([#16X4](=[OX1])=[OX1])\$([#16X4+2](=[OX1-])[OX1-])\$([#16X3]=[OX1])\$([#16X3+][OX1-])\$([#16X4](=[OX1])(=[OX1])(OX2H,OX1H0-)[OX2][#6])\$([#16X4+2](=[OX1-])(OX1-)(OX2H,OX1H0-)[OX2][#6])\$([#16X4](NX3)(=[OX1])(OX2)[#6])\$([#16X4+2](NX3)(OX1-)(OX1-)[OX2][#6])\$([#16X2][OX2H,OX1H0-])\$([#16X2][OX2H0])\$([CX3](=[OX1])[F,Cl,Br,I])"