Enumerator – A python package for the virtual generation of chemical space

Babak Mahjour

**Abstract**

The enumeration of chemical structures is a common task in cheminformatics.

**Introduction**

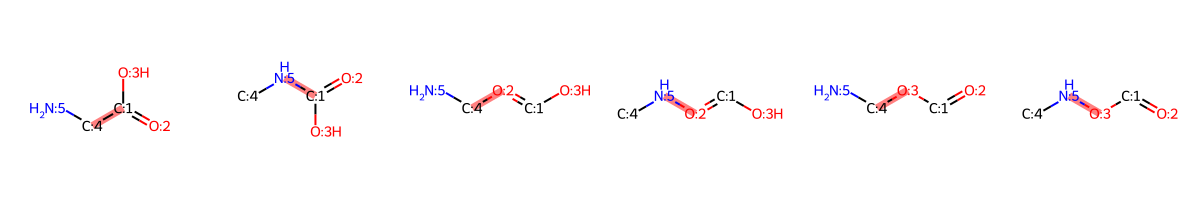
**Implementation**

* Original 320 Amine Acid
* Abstracted Cross Coupler
  + From templates

temp1 = Chem.MolFromSmarts("[C;H0;+0:1](=[O;H0;+0:2])-[O;H0;+0:3]")

temp2 = Chem.MolFromSmarts("[C;H2;+0:4]-[N;H2;+0:5]")

mol1, mol2, products, hit\_bonds = enumerate\_bonds\_between\_templates(temp1,temp2)



* + From Maximum Common Substructure

amines = ["NC(C(C)=O)=C", "CN"]

acids = ["CC(O)=O", "ClC(O)=O"]

mol1, mol2, products, hit\_bonds = combinate(amines,acids)

A group of molecules with numbers and letters

Description automatically generated with medium confidence

* Fusion Coupler



A group of red and blue molecules

Description automatically generated with medium confidence

* Bond order enumerator

mol = Chem.MolFromSmiles("O=C(C(C)C)C1=C(CCC2=CC=CC=C2)C(NC3=CC=CC=N3)=CC=C1")

mols, new\_bonds = bond\_order\_enumerator(mol, initial\_bond\_order="single", target\_bond\_order="double")

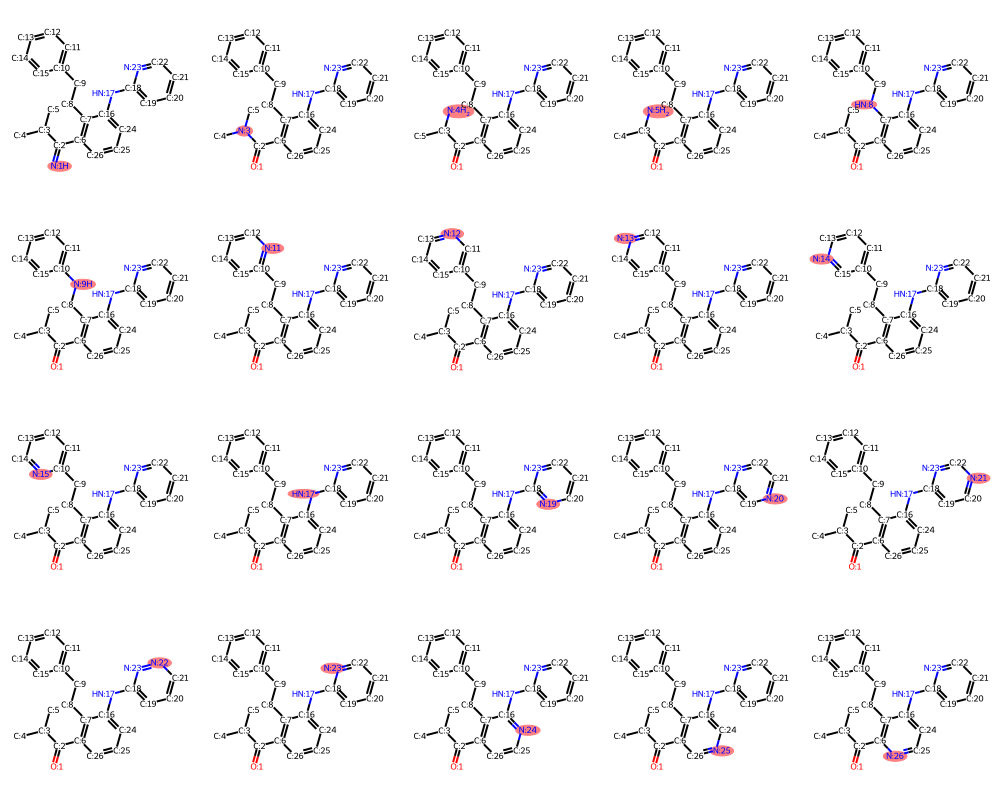
A molecule structure with numbers and letters

Description automatically generated with medium confidence

* Skeletal Editor
  + Atom swapper

mol = Chem.MolFromSmiles("O=C(C(C)C)C1=C(CCC2=CC=CC=C2)C(NC3=CC=CC=N3)=CC=C1")

mols, new\_bonds = atom\_swap(mol, "N")



* + Atom remover

mol = Chem.MolFromSmiles("O=C(C(C)C)C1=C(CCC2=CC=CC=C2)C(NC3=CC=CC=N3)=CC=C1")

new\_molecules, mapped\_mol, bond\_counter, removed = deleteAtom(mol)



A group of molecules on a white background

Description automatically generated

* + R-Walker

mol = Chem.MolFromSmiles("C=C[C@H]1C[N@@]2CC[C@H]1C[C@H]2[C@@H](C3=CC=NC4=CC=C(C=C34)OC)O")

mols, new\_bonds = atom\_walk(mol, "C")

A group of molecules in different directions

Description automatically generated with medium confidence

* Bag of Atoms
* Forward Predictors
  + VF
  + MCR enumeration of ASKCOS models
* Enzymatic templates??

**Results**

A diagram of chemical formulas

Description automatically generated

**Conclusion**