
GlobalChem
Suliman Sharif
Organic Synthesis
Aziza Frank
Named Reactions
Aziza Frank, Bettina Lier
Amino Acids
Protecting Groups
Organic
Solvents
Common Monomers
Polymers
Materials
Asuka Orr
Clay
Montmorillonite adsorption
Dyes
Medical Dyes
Medicinal Chemistry
Rings
Rings In Drugs
Alexander MacKerell Jr.
IUPAC Blue Book
Cannabis
Ian Jones, Bettina Lier
Phase 2 Heterocyclics
Shaoqi Zhan
Warheads
Common
Electrophilic
Scaffolds
Common R Groups
IUPAC Blue Book
Privileged
Sunhwan Jo
Environment
Emerging Perfluoroalkyls
Miscellaneous
Space
Interstellar Space
Vitamins
OpenSmiles
Amino Acids
Regex Patterns
Chris Burke
Proteins
Kinases
Scaffolds
Privileged
Ruibin Liu
BRAF
Common By Pocket
Aarion Romany
Lube
Daniel Khavrutskii
Sex
Narcotics
Schedule One
Schedule Two
Schedule Three
Schedule Four
Schedule Five
Asuka Orr
Formulation
Monoclonal Antibodies
Biopharmaceutical Class 3
Asuka Orr

Data & Curation

Gathering Data

```
def get_smiles():

    smiles = {
        'water': 'O',
        'methanol': 'OC',
        'ammonia': 'N',
        'hydrogen chloride': 'Cl',
        'hydrogen fluoride': 'F',
        'ethane': 'CC',
        'hydrogen cyanide': 'C#N',
        'formic acid': 'C(=O)O',
        'hydrogen bromide': 'Br',
        'fluroform': 'C(F)(F)F',
        'propane': 'C(C)C',
        'toulene': 'Cc1ccccc1',
        'pyridine': 'c1ccccn1',
        'dimethylamine': 'N(C)C',
        'ethanol': 'OCC',
        'formamide': 'C(N)=O',
        'morpholine': 'N1CCOCC1',
        'nitro': '[N+](=O)[O-]',
        'isobutane': 'C(C)(C)C',
        'anisole': 'c1ccc(OC)cc1',
        'flurobenzene': 'c1ccc(F)cc1',
        'cyclohexane': 'C1CCCCC1',
        'acetic acid': 'CC(=O)O',
        'methyl formate': 'C(=O)OC',
        'butane': 'CCCC',
        'acetamide': 'NC(C)=O',
        'methanamine': 'NC',
        'acetaldehyde': 'C(C)=O',
        'chlorobenzene': 'c1ccc(Cl)cc1',
        '1-methylpiperazine': 'N1CCN(C)CC1',
        'phenylmethanol': 'OCc1ccccc1',
        'ethylbenzene': 'CCc1ccccc1',
        'piperidine': 'N1CCCCC1',
        'thiophene': 'c1cccs1',
        'cyclopropane': 'C1CC1',
        'phenol': 'Oc1ccccc1',
        'aniline': 'Nc1ccccc1',
```

IUPAC Natural Language

```
'cyclopropenylidene': 'C1=C[C]1',
'propadienylidene': 'C=C=C',
'protonated cyanoacetylene': 'C#CC#[NH+]',
'2-propynal': 'C#CC=O',
'acrylonitrile': 'C=CC#N',
```

```
'amadori_reaction_rearrangement': r'OC(/C=N/C)C.CC(CNC)
'meerwein-ponndorf-verley_reduction': 'CC(C)=O.OC(C)C',
'stephen_aldehyde_synthesis': 'CC#N.CC=O',
'diels-alder_cycloaddition': 'C=CC=C.C=C.C1=CCCCC1',
'nebes_rearrangement': 'CC(CC)=O.CC((C(N)C)=O)',
'balz-schiemann_reaction': 'NC1=CC=CC=C1.FC2=CC=CC=C2',
'dakin-west_reaction': 'OC(C(C)N)=O.CC(C(C)NC(C)=O)=O',
'stevens_rearrangement': 'CC[N+](C)(C)C.CC(N(C)C)C',
'nenitzescu_indole_synthesis': r'O=C1C=CC(C=C1)=O.O=C(C
'criegee_oxidation': 'OCCO.C=O',
```

```
'oxyterephthaloylhydrazine-terephthaloyl': 'OC(C1=CC=C(C(C(NNC(C2=CC=C(C(=O)[O-])C=C2)=O)=O)C=C1)=O',
'nitrilo-1,4-phenylenenitriloprop-2-en-3-yl-1-ylidene-1,4-phenyleneprop-2-en-1-yl-3-ylidene':
'oxycarbonylnitrilopropane-idenenitrilocarbonyl': 'OC(N=CCC=NC=O)=O',
'oxyethyleneiminomethylenesulfanediylethyleneiminocyclohexane': 'OCCCNCSCCNC1CCCCC1',
```

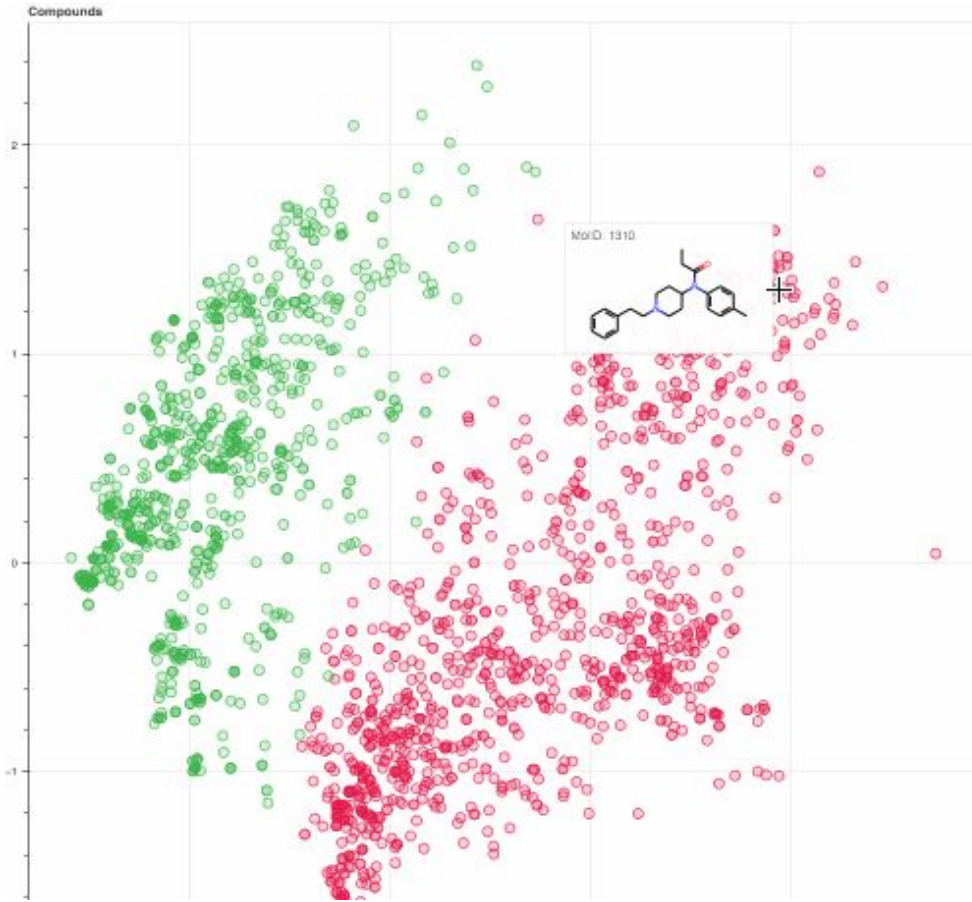
Rule Book & Language Standard

1. Remove numbers
2. Remove Site Point Language
3. Remove StereoChemistry
4. Must be “Natural” Sounding.
5. For Reactions Record in this format: R1.R2.[*:1].P1.P2
6. Common Name must have a history showing evidence of valuability

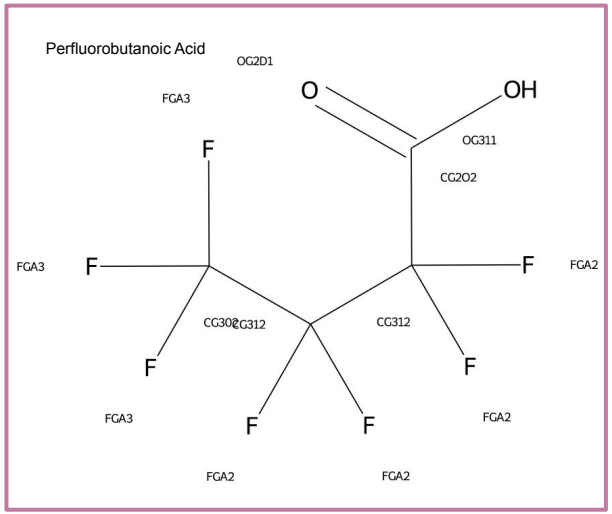
Charmm General ForceField Scoring

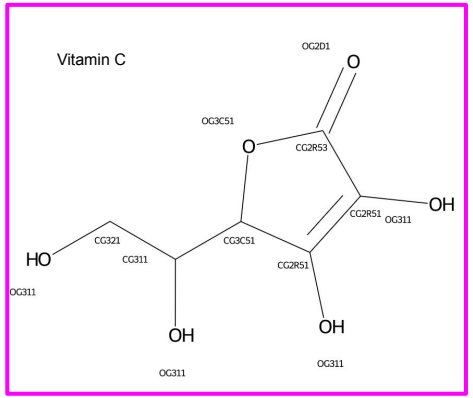
- CGenFF was built for things that are “drugs” or “drug-like”.
- Performance decreases on molecules it has not seen before
- Avenue for Selection

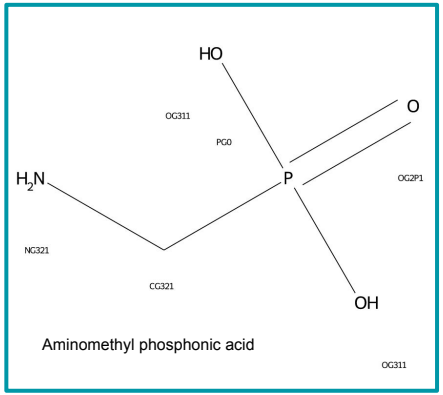
Global-Chem: A Chemical Knowledge Graph of common small molecules and their IUPAC/SMILES/SMARTS for selection of compounds relevant to diverse chemical communities

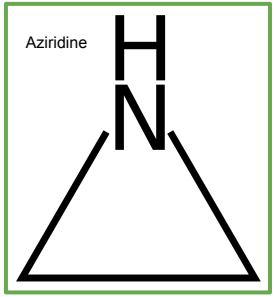


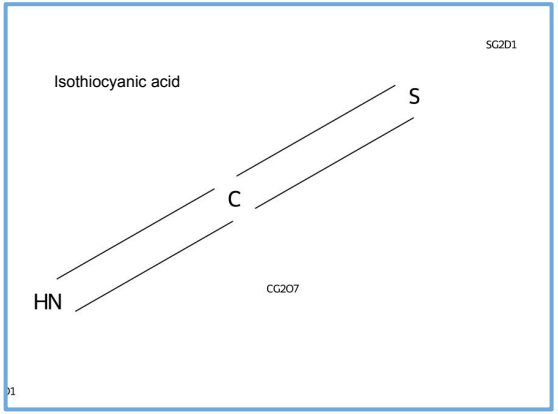
Conclusion Selection





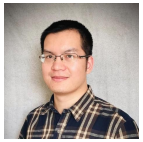






Acknowledgments


Aziza Frank


Ruibin Liu


Suliman Sharif

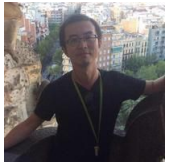

Chris Burke


Bettina Lier


Tyree Wilson


Anastasia Croitoru


Daniel Khavrutskii


Shaoqi Zhan


Alexander D. MacKerell Jr.


Aarion Romany


Sunhwan Jo


Jacob Weiner


Asuka Orr

Ian Jones

Elena Yi Chow

MacKerell Group

UNIVERSITY of MARYLAND
SCHOOL OF PHARMACY