Representing Molecules (SMILES) with substructures (SMARTS)

Written by Ghislaine Deutsch

Import

```
In [17]: from rdkit import Chem
from rdkit.Chem import Draw
from rdkit.Chem.Draw import rdMolDraw2D
```

Preparing the molecule

Inserting molecule SMILES to highlight one substructure

```
In [18]: smi = 'CCOC(=0)Nc1nc(C(N)=0)c(NC(N)=0)s1'
mol = Chem.MolFromSmiles(smi)
```

Preparing the substructure

Setting up the substructure that we want highlighted

```
In [19]: patt = Chem.MolFromSmarts('[#16R]') # MACCS keys https://github.com/rdkit/rdkit/blob/n
hit_ats = list(mol.GetSubstructMatch(patt))
```

Matching substructure on the molecule

Each bond is analyzed and the ones matching with the given substructure gets highlighted

```
In [20]: hit_bonds = []
for bond in patt.GetBonds():
    aid1 = hit_ats[bond.GetBeginAtomIdx()]
    aid2 = hit_ats[bond.GetEndAtomIdx()]
    hit_bonds.append(mol.GetBondBetweenAtoms(aid1,aid2).GetIdx())
```

Choosing colours

Simple step to insert which color to highlight with in RGB code

```
In [21]: colours = [(1.0,0.0,0.0)] #Colors https://pymolwiki.org/index.php/Color_Values
    atom_cols = {}
    for i, at in enumerate(hit_ats):
        atom_cols[at] = colours[0]
    bond_cols = {}
    for i, bd in enumerate(hit_bonds):
        bond_cols[bd] = colours[0]
```

Representing the molecule

```
In [22]: d = rdMolDraw2D.MolDraw2DCairo(500, 500) # or MolDraw2DSVG to get SVGs
```

Representing the subtructure on the molecule

In [23]: rdMolDraw2D.PrepareAndDrawMolecule(d, mol, highlightAtoms=hit_ats,highlightAtomColors=

Saving the structure as an image

```
In [24]: with open('V8.png', 'wb') as f: f.write(d.GetDrawingText()) # Saves the image where you
In []:
```