

*Figure 1: Description of the “group 1” and “group 2” environment and ring structures that have been used in the SMARTS\_RX definitions. In the structures, at least one “\*” of each labelled “het” ring should be a Heteroatom (defined as any other atom then carbon).  
Thus, “Phe” refers as purely carbon membered rings while “het” means heteroaromatic. All these rings can be substituted or not or being part of more complex ring entities.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Chemical functions hierarchical classification** | | | | **Structures sketch** | **SMARTS matched atom** |
| **Level 1** | **#*SMARTS-RX*** | **Level2** | **Level3 = (*“SMARTS-RX”*)** |
| Thioacid | 4 | Thioacid\_Aliphatic | Thioacid\_SaturatedAliphatic | R: Group 1 | O, S of OH, SH |
| Thioacid\_UnsaturatedAliphatic |
| Thioacid\_Aromatic | Thioacid\_Aromatic |
| Thioacid\_Heteroaromatic | Thioacid\_Heteroaromatic |
| 3 | Thioacid\_Carbamic | Thioacid\_Carbamic | Both A are any atoms (including H) | O, S of OH, SH |
| Thioacid\_Carbonic | Thioacid\_Carbonic | A is any atoms (including H) | O, S of OH, SH |
| Thioacid\_Thiocarbonic | Thioacid\_Thiocarbonic | Both R are any atoms (including H) | O, S of OH, SH |
| Acid | 4 | Acid\_Aliphatic | Acid\_SaturatedAliphatic | R: Group 1  NB: For Acid\_SaturatedAliphatic: R not starting N, O, S and  match also formic acid | O of OH |
| Acid\_UnsaturatedAliphatic |
| Acid\_Aromatic | Acid\_Aromatic |
| Acid\_Heteroaromatic | Acid\_Heteroaromatic |
| 3 | Acid\_Carbamic | Acid\_Carbamic | Both A are any atoms (including H) | O of OH |
| Acid\_Carbonic | Acid\_Carbonic | A is any atoms (including H) | O of OH |
| Acid\_Thiocarbonic | Acid\_Thiocarbonic | A is any atoms (including H) | O of OH |
| AcidX | 3 | Xformamide | Xformamide | A = H or carbon  X = [F, Cl, Br, I, “Triflate:” OC(F)(F)F] | X |
| XThioformate | XThioformate | A is any atoms (including H)  X = [F, Cl, Br, I, “Triflate:” OC(F)(F)F] | O sp3, S sp3 |
| XFormate | XFormate | A is any atoms (including H)  X = [F, Cl, Br, I, “Triflate:” OC(F)(F)F] | O sp3 |
| 20 | AcidX\_Aliphatic | AcidX-Fluoride\_SaturatedAliphatic | X = [F, Cl, Br, I, “Triflate:” OC(F)(F)F]  R: Group 1 | X |
| AcidX-Chloride\_SaturatedAliphatic |
| AcidX-Bromide\_SaturatedAliphatic |
| AcidX-Iodide\_SaturatedAliphatic |
| AcidX-Triflate\_SaturatedAliphatic |
| AcidX-Fluoride\_UnsaturatedAliphatic |
| AcidX-Chloride\_UnsaturatedAliphatic |
| AcidX-Bromide\_UnsaturatedAliphatic |
| AcidX-Iodide\_UnsaturatedAliphatic |
| AcidX-Triflate\_UnsaturatedAliphatic |
| AcidX\_Aromatic | AcidX-Fluoride\_Aromatic |
| AcidX-Chloride\_Aromatic |
| AcidX-Bromide\_Aromatic |
| AcidX-Iodide\_Aromatic |
| AcidX-Triflate\_Aromatic |
| AcidX\_Heteroaromatic | AcidX-Fluoride\_Heteroaromatic |
| AcidX-Chloride\_Heteroaromatic |
| AcidX-Bromide\_Heteroaromatic |
| AcidX-Iodide\_Heteroaromatic |
| AcidX-Triflate\_Heteroaromatic |
| Anhydride | 1 | Anhydride | Anhydride | A are any C, O is not in ring | O bonds to 2 carbonyls |
| XKetone | 10 | XKetoneAromaticFluoride | XKetoneAromaticFluoride | R is group starting by aliphatic carbon  Or  R is any aromatic | C from the carbonyl |
| XKetoneAromaticChloride | XKetoneAromaticChloride |
| XKetoneAromaticBromide | XKetoneAromaticBromide |
| XKetoneAromaticIodide | XKetoneAromaticIodide |
| XKetoneAromaticTriflate | XKetoneAromaticTriflate |
| XKetoneAliphaticFluoride | XKetoneAliphaticFluoride |
| XKetoneAliphaticChloride | XKetoneAliphaticChloride |
| XKetoneAliphaticBromide | XKetoneAliphaticBromide |
| XKetoneAliphaticIodide | XKetoneAliphaticIodide |
| XKetoneAliphaticTriflate | XKetoneAliphaticTriflate |
| Bocanhydrate | 1 | Bocanhydrate | Bocanhydrate | Exact-match | O bonds to 2 carbonyls |
| Mesylate | 2 | Mesylate | Mesylate | A is any carbon  R is any heavy atom  Do not match triflate or dimesylate | O sp3 |
|  | DiMesylate | DiMesylate | A are any carbon | O sp3 |
| XF | 16 | X\_Aliphatic | X-Fluoride\_Aliph | X-R  X = [F]  R: Group 2 | F |
| X-Fluoride\_Vinyl |
| X-Fluoride\_Alkyne |
| X-Fluoride\_Nitrile |
| X\_Fluoride\_Aromatic | X-Fluoride\_Phe |
| X\_Fluoride \_Heteroaromatic | X-Fluoride\_Het6 |
| X\_Fluoride\_Heteroaromatic | X-Fluoride\_Het5 |
| X\_Fluoride\_Heteroaromatic | X-Fluoride\_Het6Het5 |
| X\_Fluoride\_Heteroaromatic | X-Fluoride\_PheHet5 |
| X\_Fluoride\_Heteroaromatic | X-Fluoride\_Het5Het6 |
| X\_Fluoride\_Heteroaromatic | X-Fluoride\_Het5Phe |
| X\_Fluoride\_Heteroaromatic | X-Fluoride\_Het5Het5 |
| X\_Fluoride\_Aromatic | X-Fluoride\_PhePhe |
| X\_Fluoride\_Heteroaromatic | X-Fluoride\_Het6Phe |
| X\_Fluoride\_Heteroaromatic | X-Fluoride\_PheHet6 |
| X\_Fluoride\_Heteroaromatic | X-Fluoride\_Het6Het6 |
| X | 64 | X\_Aliphatic | X-Chloride\_Aliph | X-R  X = [Cl, Br, I, “Triflate:” OC(F)(F)F]  R: Group 2 | X |
| X\_Aliphatic | X-Bromide\_Aliph |
| X\_Aliphatic | X-Iodide\_Aliph |
| X\_Aliphatic | X-Triflate\_Aliph |
| X\_Aliphatic | X-Chloride\_Vinyl |
| X\_Aliphatic | X-Bromide\_Vinyl |
| X\_Aliphatic | X-Iodide\_Vinyl |
| X\_Aliphatic | X-Triflate\_Vinyl |
| X\_Aliphatic | X-Chloride\_Alkyne |
| X\_Aliphatic | X-Bromide\_Alkyne |
| X\_Aliphatic | X-Iodide\_Alkyne |
| X\_Aliphatic | X-Triflate\_Alkyne |
| X\_Aliphatic | X-Chloride\_Nitrile |
| X\_Aliphatic | X-Bromide\_Nitrile |
| X\_Aliphatic | X-Iodide\_Nitrile |
| X\_Aliphatic | X-Triflate\_Nitrile |
| X\_Aromatic | X-Chloride\_Phe |
| X\_Aromatic | X-Bromide\_Phe |
| X\_Aromatic | X-Iodide\_Phe |
| X\_Aromatic | X-Triflate\_Phe |
| X\_Heteroaromatic | X-Chloride\_Het6 |
| X\_Heteroaromatic | X-Bromide\_Het6 |
| X\_Heteroaromatic | X-Iodide\_Het6 |
| X\_Heteroaromatic | X-Triflate\_Het6 |
| X\_Heteroaromatic | X-Chloride\_Het5 |
| X\_Heteroaromatic | X-Bromide\_Het5 |
| X\_Heteroaromatic | X-Iodide\_Het5 |
| X\_Heteroaromatic | X-Triflate\_Het5 |
| X\_Heteroaromatic | X-Chloride\_Het6Het5 |
| X\_Heteroaromatic | X-Bromide\_Het6Het5 |
| X\_Heteroaromatic | X-Iodide\_Het6Het5 |
| X\_Heteroaromatic | X-Triflate\_Het6Het5 |
| X\_Heteroaromatic | X-Chloride\_PheHet5 |
| X\_Heteroaromatic | X-Bromide\_PheHet5 |
| X\_Heteroaromatic | X-Iodide\_PheHet5 |
| X\_Heteroaromatic | X-Triflate\_PheHet5 |
| X\_Heteroaromatic | X-Chloride\_Het5Het6 |
| X\_Heteroaromatic | X-Bromide\_Het5Het6 |
| X\_Heteroaromatic | X-Iodide\_Het5Het6 |
| X\_Heteroaromatic | X-Triflate\_Het5Het6 |
| X\_Heteroaromatic | X-Chloride\_Het5Phe |
| X\_Heteroaromatic | X-Bromide\_Het5Phe |
| X\_Heteroaromatic | X-Iodide\_Het5Phe |
| X\_Heteroaromatic | X-Triflate\_Het5Phe |
| X\_Heteroaromatic | X-Chloride\_Het5Het5 |
| X\_Heteroaromatic | X-Bromide\_Het5Het5 |
| X\_Heteroaromatic | X-Iodide\_Het5Het5 |
| X\_Heteroaromatic | X-Triflate\_Het5Het5 |
| X\_Aromatic | X-Chloride\_PhePhe |
| X\_Aromatic | X-Bromide\_PhePhe |
| X\_Aromatic | X-Iodide\_PhePhe |
| X\_Aromatic | X-Triflate\_PhePhe |
| X\_Heteroaromatic | X-Chloride\_Het6Phe |
| X\_Heteroaromatic | X-Bromide\_Het6Phe |
| X\_Heteroaromatic | X-Iodide\_Het6Phe |
| X\_Heteroaromatic | X-Triflate\_Het6Phe |
| X\_Heteroaromatic | X-Chloride\_PheHet6 |
| X\_Heteroaromatic | X-Bromide\_PheHet6 |
| X\_Heteroaromatic | X-Iodide\_PheHet6 |
| X\_Heteroaromatic | X-Triflate\_PheHet6 |
| X\_Heteroaromatic | X-Chloride\_Het6Het6 |
| X\_Heteroaromatic | X-Bromide\_Het6Het6 |
| X\_Heteroaromatic | X-Iodide\_Het6Het6 |
| X\_Heteroaromatic | X-Triflate\_Het6Het6 |
| Aldehyde | 4 | Aldehyde\_Aliphatic | Aldehyde\_SaturatedAliphatic | R: Group 1  NB: Aldehyde\_SaturatedAliphatic match also formaldehyde | C of C=O |
| Aldehyde\_UnsaturatedAliphatic |
| Aldehyde\_Aromatic | Aldehyde\_Aromatic |
| Aldehyde\_Heteroaromatic | Aldehyde\_Heteroaromatic |
| Ketone | 7 | KetoneAromaticCyclic | KetoneAromaticCyclic | Carbonyl C is in ring or not  A is aliphatic or aromatic carbon  “Mix” one A is aliphatic carbon, one is aromatic carbon  Both A’s have same carbon type in other cases (aliphatic or aromatic) | C of C=O |
| KetoneAromaticAcyclic | KetoneAromaticAcyclic |
| KetoneAliphaticCyclic | KetoneAliphaticCyclic |
| KetoneAliphaticAcyclic | KetoneAliphaticAcyclic |
| KetoneMixCyclic | KetoneMixCyclic |
| KetoneMixAcyclic | KetoneMixAcyclic |
| Thioketone | Thioketone | Both A are any carbon but not attached to any double bond (excluded its own function C=S) | C of C=S |
| Nitrogen\_X | 10 | NitrogenFluorideCyclic | NitrogenFluorideCyclic | N-X  X = [F, Cl, Br, I, “Triflate:” OC(F)(F)F]  N is within a ring or not | X |
| NitrogenFluorideAcyclic | NitrogenFluorideAcyclic |
| NitrogenChlorideCyclic | NitrogenChlorideCyclic |
| NitrogenChlorideAcyclic | NitrogenChlorideAcyclic |
| NitrogenBromideCyclic | NitrogenBromideCyclic |
| NitrogenBromideAcyclic | NitrogenBromideAcyclic |
| NitrogenIodideCyclic | NitrogenIodideCyclic |
| NitrogenIodideAcyclic | NitrogenIodideAcyclic |
| NitrogenTriflateCyclic | NitrogenTriflateCyclic |
| NitrogenTriflateAcyclic | NitrogenTriflateAcyclic |
| Imine | 2 | AromaticImine | AromaticImine | A-N=C  A is aliphatic or aromatic carbon. For aliphatic, A can be H | N |
| AliphaticImine | AliphaticImine |
| Thioester | 2 | ThioEsterAcyclic | ThioEsterAcyclic | O, S sp3 and C sp2 in ring of not  A is Carbon | S or O sp3 |
| ThioEsterCyclic | ThioEsterCyclic |
| Ester | 2 | EsterAcyclic | EsterAcyclic | O sp3 and C sp2 in ring of not  A is Carbon | O sp3 |
| EsterCyclic | EsterCyclic |
| EsterAromatic | EsterAromatic | O and C in aromatic ring (e.g. ...oc(=O)...) | O aromatic |
| Nitrile | 4 | Nitrile\_Aliphatic | Nitrile\_SaturatedAliphatic | R: Group 1 | C |
| Nitrile\_Aliphatic | Nitrile\_UnsaturatedAliphatic |
| Nitrile\_Aromatic | Nitrile\_Aromatic |
| Nitrile\_Heteroaromatic | Nitrile\_Heteroaromatic |
| Sulfur | 1 | Sulfur | Sulfur | S (not attached to any Heavy atoms) | S |
| Sulfonyl | 10 | Sulfonyl\_Aliphatic | Sulfonyl-Fluoride\_Aliphatic | X = [F, Cl, Br, I, “Triflate:” OC(F)(F)F]  R is heavy atoms, aliphatic or aromatic groups | X |
| Sulfonyl\_Aliphatic | Sulfonyl-Chloride\_Aliphatic |
| Sulfonyl\_Aliphatic | Sulfonyl-Bromide\_Aliphatic |
| Sulfonyl\_Aliphatic | Sulfonyl-Iodide\_Aliphatic |
| Sulfonyl\_Aliphatic | Sulfonyl-Triflate\_Aliphatic |
| Sulfonyl\_Aromatic | Sulfonyl-Fluoride\_Aromatic |
| Sulfonyl\_Aromatic | Sulfonyl-Chloride\_Aromatic |
| Sulfonyl\_Aromatic | Sulfonyl-Bromide\_Aromatic |
| Sulfonyl\_Aromatic | Sulfonyl-Iodide\_Aromatic |
| Sulfonyl\_Aromatic | Sulfonyl-Triflate\_Aromatic |
| 2 | Sulfonic\_Aliphatic | Sulfonic\_Aliphatic | R is heavy atoms, aliphatic or aromatic groups | O from OH |
| Sulfonic\_Aromatic | Sulfonic\_Aromatic |
| Sulfinyl | 10 | Sulfinyl\_Aliphatic | Sulfinyl-Fluoride\_Aliphatic | X = [F, Cl, Br, I, “Triflate:” OC(F)(F)F]  R is heavy atoms, aliphatic or aromatic groups | X |
| Sulfinyl\_Aliphatic | Sulfinyl-Chloride\_Aliphatic |
| Sulfinyl\_Aliphatic | Sulfinyl-Bromide\_Aliphatic |
| Sulfinyl\_Aliphatic | Sulfinyl-Iodide\_Aliphatic |
| Sulfinyl\_Aliphatic | Sulfinyl-Triflate\_Aliphatic |
| Sulfinyl\_Aromatic | Sulfinyl-Fluoride\_Aromatic |
| Sulfinyl\_Aromatic | Sulfinyl-Chloride\_Aromatic |
| Sulfinyl\_Aromatic | Sulfinyl-Bromide\_Aromatic |
| Sulfinyl\_Aromatic | Sulfinyl-Iodide\_Aromatic |
| Sulfinyl\_Aromatic | Sulfinyl-Triflate\_Aromatic |
| SulfonylImine | 10 | SulfonylImine\_Aliphatic | SulfonylImine-Fluoride\_Aliphatic | X = [F, Cl, Br, I, “Triflate:” OC(F)(F)F]  A is any atom including H  R is heavy atoms, aliphatic or aromatic groups | X |
| SulfonylImine\_Aliphatic | SulfonylImine-Chloride\_Aliphatic |
| SulfonylImine\_Aliphatic | SulfonylImine-Bromide\_Aliphatic |
| SulfonylImine\_Aliphatic | SulfonylImine-Iodide\_Aliphatic |
| SulfonylImine\_Aliphatic | SulfonylImine-Triflate\_Aliphatic |
| SulfonylImine\_Aromatic | SulfonylImine-Fluoride\_Aromatic |
| SulfonylImine\_Aromatic | SulfonylImine-Chloride\_Aromatic |
| SulfonylImine\_Aromatic | SulfonylImine-Bromide\_Aromatic |
| SulfonylImine\_Aromatic | SulfonylImine-Iodide\_Aromatic |
| SulfonylImine\_Aromatic | SulfonylImine-Triflate\_Aromatic |
| Nitro | 4 | Nitro\_Aliphatic | Nitro\_SaturatedAliphatic | R: Group 1 | N |
| Nitro\_Aliphatic | Nitro\_UnsaturatedAliphatic |
| Nitro\_Aromatic | Nitro\_Aromatic |
| Nitro\_Heteroaromatic | Nitro\_Heteroaromatic |
| PFAS | 4 | PFAS\_Aliphatic | CF2\_SaturatedAliphatic | R-CF2 (not PFAS)  R: Group 1 | C |
| PFAS\_Aliphatic | CF2\_UnsaturatedAliphatic |
| PFAS\_Aromatic | CF2\_Aromatic |
| PFAS\_Heteroaromatic | CF2\_Heteroaromatic |
| 4 | PFAS\_Aliphatic | CF3\_SaturatedAliphatic | R-CF3 (not PFAS)  R: Group 1 | C |
| PFAS\_Aliphatic | CF3\_UnsaturatedAliphatic |
| PFAS\_Aromatic | CF3\_Aromatic |
| PFAS\_Heteroaromatic | CF3\_Heteroaromatic |
| 4 | PFAS\_Aliphatic | PFAS\_SaturatedAliphatic | R-C(C(F)(F))nC(F)(F)F n=1-11 | First atom of R attached on CF2 |
| PFAS\_Aliphatic | PFAS\_UnsaturatedAliphatic |
| PFAS\_Aromatic | PFAS\_Aromatic |
| PFAS\_Heteroaromatic | PFAS\_Heteroaromatic |
| Nitroxy | 1 | Nitroxy | Nitroxy | A is any atoms including H (does not match nitro) | N |
| Diazo | 1 | Diazo | Diazo | A-C=[N+]=[N-]  A is any atoms including H | C |
| ThioCyanate | 2 | ThioCyanate\_Aliphatic | ThioCyanate\_Aliphatic | R is aliphatic or aromatic  R can be H | S |
| ThioCyanate\_Aromatic | ThioCyanate\_Aromatic |
| ThioisoCyanate | 4 | ThioisoCyanate\_Aliphatic | ThioisoCyanate\_SaturatedAliphatic | R: Group1 | N |
| ThioisoCyanate\_Aliphatic | ThioisoCyanate\_UnsaturatedAliphatic |
| ThioisoCyanate\_Aromatic | ThioisoCyanate\_Aromatic |
| ThioisoCyanate\_Heteroaromatic | ThioisoCyanate\_Heteroaromatic |
| Cyanate | 2 | Cyanate\_Aliphatic | Cyanate\_Aliphatic | R is aliphatic or aromatic  R can be H | O |
| Cyanate\_Aromatic | Cyanate\_Aromatic |
| IsoCyanate | 4 | IsoCyanate\_Aliphatic | IsoCyanate\_SaturatedAliphatic | R: Group 1 | N |
| IsoCyanate\_Aliphatic | IsoCyanate\_UnsaturatedAliphatic |
| IsoCyanate\_Aromatic | IsoCyanate\_Aromatic |
| IsoCyanate\_Heteroaromatic | IsoCyanate\_Heteroaromatic |
| Boroxine | 3 | AromBoroxine | AromBoroxine | All A aromatic | One B |
| AliphaticBoroxine | AliphaticBoroxine | All A aliphatic heavy atoms | One B |
| MixBoroxine | MixBoroxine | At least one A aromatic, heavy atoms if not | One B |
| Boronate | 4 | Boronate\_Aliphatic | Boronate\_SaturatedAliphatic | A = Any heavy atoms  R: Group1 | B |
| Boronate\_Aliphatic | Boronate\_UnsaturatedAliphatic |
| Boronate\_Aromatic | Boronate\_Aromatic |
| Boronate\_Heteroaromatic | Boronate\_Heteroaromatic |
| 1 | PincerBoronate | PincerBoronate |  | B |
| 1 | AromDiboronate | AromDiboronate | A is any heavy atom  A : aliphatic or aromatic | B |
| 1 | AliphaticDiboronate | AliphaticDiboronate | B |
| 1 | TrialkylBoronate | TrialkylBoronate | A is any heavy atom  A is aliphatic | B |
| 1 | TrifluoroBoronateAromatic | TrifluoroBoronateAromatic | R is heavy atoms, aliphatic or aromatic | B |
| 1 | TrifluoroBoronateAliphatic | TrifluoroBoronateAliphatic |
| Boronic | 4 | Boronic\_Aliphatic | Boronic\_SaturatedAliphatic | B not in ring  R: Group 1 | B |
| Boronic\_Aliphatic | Boronic\_UnsaturatedAliphatic |
| Boronic\_Aromatic | Boronic\_Aromatic |
| Boronic\_Heteroaromatic | Boronic\_Heteroaromatic |
| 1 | BoronicCyclic | BoronicCyclic | A any heavy atoms. B in ring | B |
| Borane | 2 | AliphaticBorane | AliphaticBorane | A is sp3 carbon  R is any group not H | B |
| AliphaticDiborane | AliphaticDiborane | A are all sp3 carbon | One B |
| Protecting | 3 | Fmoc | Fmoc |  | N |
| Boc | Boc |  | N |
| Pivalic | Pivalic |  | O |
| Alkene/MichaelAcceptor | 2 | AlkeneTerminal | AlkeneTerminal | At least one A is any heavy atom | C with A |
| AlkeneNonTerminal | AlkeneNonTerminal | At least one A and R not H  R is Carbon or H  A is Any atom (not N, O, F, Cl, Br, I) | C with R |
| Thioether | 5 | ThioetherAliphaticCyclic | ThioetherAliphaticCyclic | R-S-R  R is aliphatic or aromatic carbon  “Mix” one R is Aliphatic, one R is Aromatic  Both R’s have same type in other cases  The S or s can be in ring or not | S |
| ThioetherMixCyclic | ThioetherMixCyclic |
| ThioetherAromaticAcyclic | ThioetherAromaticAcyclic |
| ThioetherAliphaticAcyclic | ThioetherAliphaticAcyclic |
| ThioetherMixAcyclic | ThioetherMixAcyclic |
| Ether | 5 | EtherAromaticAcyclic | EtherAromaticAcyclic | R-O-R  R is aliphatic or aromatic carbon  “Mix” one R is Aliphatic, one R is Aromatic  Both R’s have same type in other cases  The O can be in ring or not | O |
| EtherAliphaticCyclic | EtherAliphaticCyclic |
| EtherAliphaticAcyclic | EtherAliphaticAcyclic |
| EtherMixCyclic | EtherMixCyclic |
| EtherMixAcyclic | EtherMixAcyclic |
| Ketals | 4 | Ketals | Ketals | A, R1, R2 are any heavy atoms | One O |
| ThioKetals | A, R1, R2 are any heavy atoms | One S |
| Acetal | Acetal | A is Hydrogen  R1 is any atoms including hydrogen  R2 is any heavy atoms | One O |
| ThioAcetal | A is Hydrogen  R1 is any atoms including hydrogen  R2 is any heavy atoms | One S |
| Sulfone | 1 | Sulfone | Sulfone | A is Carbon | S |
| Alkyne | 5 | AlkyneTerminalAromatic | AlkyneTerminalAromatic | H on C#C is required  A is any heavy atom (aliphatic or aromatic) | C attached to R |
| AlkyneTerminalAliphatic | AlkyneTerminalAliphatic |
| AlkyneNonTerminalAromatic | AlkyneNonTerminalAromatic | A is any heavy atom (aliphatic or aromatic)  A is not halogen | C |
| AlkyneNonTerminalAliphatic | AlkyneNonTerminalAliphatic |
| AlkyneNonTerminalMix | AlkyneNonTerminalMix |
| Sulphonamide | 3 | Sulphonamide | Sulphonamide0H | R is any atom including H  A is any atom, not H | N |
| Sulphonamide1H | R is any atom including H  A is any atom, not H | N |
| Sulphonamide2H | R is any atom including H | N |
| Thiocarbamate | 1 | Thiocarbamate | Thiocarbamate | A is any atom including H  R is any heavy atoms | N |
| Carbamate | 1 | Carbamate | Carbamate | A is any atom  R is any heavy atoms | N |
| Imide | 1 | Imide | Imide | A is any (not attached with double bond)  R is any C attached groups | N |
| TetraX | 1 | TetraX | TetraX | X = [F, Cl, Br, I] | One halogen |
| Phosphate | 1 | PhosphoricAcid | PhosphoricAcid | A is any atom including H  R is any atoms | O of OH |
| 4 | PhosphoricEster\_Aliphatic | PhosphoricEster\_SaturatedAliphatic | A is any atom including H  R group 1 | O with R |
| PhosphoricEster\_Aliphatic | PhosphoricEster\_UnsaturatedAliphatic |
| PhosphoricEster\_Aromatic | PhosphoricEster\_Aromatic |
| PhosphoricEster\_Heteroaromatic | PhosphoricEster\_Heteroaromatic |
| 3 | PhosphoricAmide\_Aliphatic | PhosphoricAmide\_Aliphatic | A is any including H  R1 is heavy atom  “Mix” one R is Aliphatic, one R is Aromatic  Both R’s have same type in other cases | N |
| PhosphoricAmide\_Aromatic | PhosphoricAmide\_Aromatic |
| PhosphoricAmide\_Mix | PhosphoricAmide\_Mix |
| Sulfinate | 1 | Sulfinate | Sulfinate | A is any atom including H | O bonds to A |
| Amide | 2 | Amide | Amide | R is any including H  A is any group attached by a carbon | N |
| Pyridone\_likeNH | in any aromatic ring (e.g. ...[nH]c(=O)...  Hydrogen on N imposed | N |
| Oxime | 1 | Oxime | Oxime | A is Carbon | N |
| Azide | 1 | Azide | Azide | A-[N+,-]=[N+,-]=[N+,-] or  A-[N]=[N+,-]=[N]  A is any atoms | N attached to A |
| Phenol | 3 | PhenolHeteroaromatic6membered | PhenolHeteroaromatic6membered | At least 1 “\*” is N | O |
| PhenolHeteroaromatic5membered | PhenolHeteroaromatic5membered | At least 1 “\*” is N | O |
| Phenol | Phenol |  | O |
| Thiophenol | 3 | ThiophenolHeteroaromatic6membered | ThiophenolHeteroaromatic6membered | At least 1 “\*” is N | S |
| ThiophenolHeteroaromatic5membered | ThiophenolHeteroaromatic5membered | At least 1 “\*” is N | S |
| Thiophenol | Thiophenol |  | S |
| Thiol | 10 | PrimaryThiolAromatic | PrimaryThiolAromatic | R is aliphatic or aromatic | S |
| PrimaryThiolAliphatic | PrimaryThiolAliphatic |
| SecondaryThiolAromaticCyclic | SecondaryThiolAromaticCyclic | R is aliphatic or aromatic  “Mix” one R is Aliphatic, one R is Aromatic  Both R’s have same type in other cases  The C attached to S can be in ring or not | S |
| SecondaryThiolAliphaticCyclic | SecondaryThiolAliphaticCyclic |
| SecondaryThiolMixCyclic | SecondaryThiolMixCyclic |
| SecondaryThiolAromaticAcyclic | SecondaryThiolAromaticAcyclic |
| SecondaryThiolAliphaticAcyclic | SecondaryThiolAliphaticAcyclic |
| SecondaryThiolMixAcyclic | SecondaryThiolMixAcyclic |
| TertiaryThiolCyclic | TertiaryThiolCyclic | At least 1 R is Carbon  The C attached to S can be in ring or not |  |
| TertiaryThiolAcyclic | TertiaryThiolAcyclic |  |
| Alcohol | 10 | PrimaryAlcoholAromatic | PrimaryAlcoholAromatic | R is aliphatic or aromatic  Match methanol | O |
| PrimaryAlcoholAliphatic | PrimaryAlcoholAliphatic |
| SecondaryAlcoholAromaticCyclic | SecondaryAlcoholAromaticCyclic | R is aliphatic or aromatic  “Mix” one R is Aliphatic, one R is Aromatic  Both R’s have same type in other cases  The C attached to O can be in ring or not | O |
| SecondaryAlcoholAliphaticCyclic | SecondaryAlcoholAliphaticCyclic |
| SecondaryAlcoholMixCyclic | SecondaryAlcoholMixCyclic |
| SecondaryAlcoholAromaticAcyclic | SecondaryAlcoholAromaticAcyclic |
| SecondaryAlcoholAliphaticAcyclic | SecondaryAlcoholAliphaticAcyclic |
| SecondaryAlcoholMixAcyclic | SecondaryAlcoholMixAcyclic |
| TertiaryAlcoholCyclic | TertiaryAlcoholCyclic | R is aliphatic or aromatic  The C attached to O can be in ring or not | O |
| TertiaryAlcoholAcyclic | TertiaryAlcoholAcyclic |
| Amidine | 1 | Amidine | Amidine | R2 is carbon  R1, R3 is any heavy  Does not match guanidine | N in  (N=C) |
| Urea | 3 | AromaticUrea | AromaticUrea | At least 1 R and 1 A are aromatic | One N |
| MixUrea | MixUrea | 1 R or 1 A is aromatic, the other are aliphatic or H | One N |
| AliphaticUrea | AliphaticUrea | All R and A are aliphatic or H | One N |
| 3 | AromaticThiourea | AromaticThiourea | At least 1 R and 1 A are aromatic | One N |
| MixThiourea | MixThiourea | 1 R or 1 A is aromatic, the others are aliphatic | One N |
| AliphaticThiourea | AliphaticThiourea | All R and A are aliphatic or H | One N |
| Phosphine | 3 | AromaticPhosphine | AromaticPhosphine | R is aromatic or aliphatic carbon  “Mix” one R is aliphatic carbon, one R is Aromatic  Both R’s have same type in other cases | P |
| AliphaticPhosphine | AliphaticPhosphine |
| MixPhosphine | MixPhosphine |
| 3 | AromaticPhosphineOxyde | AromaticPhosphineOxyde | R is aromatic or aliphatic carbon  “Mix” one R is aliphatic carbon, one R is Aromatic  Both R’s have same type in other cases | P |
| AliphaticPhosphineOxyde | AliphaticPhosphineOxyde |
| MixPhosphineOxyde | MixPhosphineOxyde |
| Sulfinylamine | 2 | SulfinylamineAromatic | SulfinylamineAromatic | A is aromatic or aliphatic carbon  R is any atom including H | N |
| SulfinylamineAliphatic | SulfinylamineAliphatic |
| NitrogenHeterocycle | 4 | Heterocycle5\_1N | Heterocycle5\_1NH | pyrrole | N |
| Heterocycle5\_1N | Heterocycle5\_1NC | A is any not H | N |
| Heterocycle5\_polyN | Heterocycle5\_polyHetNH | At least one “\*” is [n,o,s] | N |
| Heterocycle5\_polyN | Heterocycle5\_polyHet | At least one “\*” is [n,o,s] | N |
| 4 | Heterocycle6\_1N | Heterocycle6\_1NC |  | N |
| Heterocycle6\_1N | Heterocycle6\_1N |  | N |
| Heterocycle6\_polyN | Heterocycle6\_polyHetNC | At least one “\*” is [n,o,s] | N |
| Heterocycle6\_polyN | Heterocycle6\_polyHet | At least one “\*” is [n,o,s] | N |
| OxygenHeterocycle | 1 | OxygenHeterocycle | Heterocycle5\_O | “\*” can be [c,n]  Match furane | o |
| SulfurHeterocycle | 1 | SulfurHeterocycle | Heterocycle5\_S | “\*” can be [c,n]  Match thiophene | s |
| Guanidine | 1 | Guanidine | Guanidine | Any substitution on any N | N single bond |
| Amonia | 1 | Amonia | Amonia | NH3  Exact match | N |
| Amine | 34 | Benzyl\_amine\_primary | Benzylamine\_primary | R is any including H | N |
| Benzyl\_amine\_primary | HeteroBenzylamine\_primary | As above but phenyl replaced by any 5 or 6 membered heterocycles | N |
| Amine\_Aliphatic | Amine\_Primary\_SaturatedAliphatic | R: Group 2  Excluding any Benzyl\_amine\_primary | N |
| Amine\_Aliphatic | Amine\_Primary\_UnsaturatedAliphatic |
| Amine\_Aromatic | Amine\_Primary\_Phe |
| Amine\_Heteroaromatic | Amine\_Primary\_Het6 |
| Amine\_Heteroaromatic | Amine\_Primary\_Het5 |
| Amine\_Heteroaromatic | Amine\_Primary\_Het6Het5 |
| Amine\_Heteroaromatic | Amine\_Primary\_PheHet5 |
| Amine\_Heteroaromatic | Amine\_Primary\_Het5Het6 |
| Amine\_Heteroaromatic | Amine\_Primary\_Het5Phe |
| Amine\_Heteroaromatic | Amine\_Primary\_Het5Het5 |
| Amine\_Aromatic | Amine\_Primary\_PhePhe |
| Amine\_Heteroaromatic | Amine\_Primary\_Het6Phe |
| Amine\_Heteroaromatic | Amine\_Primary\_PheHet6 |
| Amine\_Heteroaromatic | Amine\_Primary\_Het6Het6 |
| Amine\_Aliphatic | Amine\_AcyclicSecondary\_SaturatedAliphatic-SaturatedAliphatic | R: Group 2  N can be in a ring or not | N |
| Amine\_Aliphatic | Amine\_AcyclicSecondary\_SaturatedAliphatic-UnsaturatedAliphatic |
| Amine\_Aromatic | Amine\_AcyclicSecondary\_SaturatedAliphatic-Aromatic |
| Amine\_Heteroaromatic | Amine\_AcyclicSecondary\_SaturatedAliphatic-Heteroaromatic |
| Amine\_Aliphatic | Amine\_AcyclicSecondary\_UnsaturatedAliphatic-UnsaturatedAliphatic |
| Amine\_Aromatic | Amine\_AcyclicSecondary\_UnsaturatedAliphatic-Aromatic |
| Amine\_Heteroaromatic | Amine\_AcyclicSecondary\_UnsaturatedAliphatic-Heteroaromatic |
| Amine\_Aromatic | Amine\_AcyclicSecondary\_Aromatic-Aromatic |
| Amine\_Heteroaromatic | Amine\_AcyclicSecondary\_Aromatic-Heteroaromatic |
| Amine\_Heteroaromatic | Amine\_AcyclicSecondary\_Heteroaromatic-Heteroaromatic |
| Amine\_Aliphatic | Amine\_CyclicSecondary\_SaturatedAliphatic-SaturatedAliphatic |
| Amine\_Aliphatic | Amine\_CyclicSecondary\_SaturatedAliphatic-UnsaturatedAliphatic |
| Amine\_Aromatic | Amine\_CyclicSecondary\_SaturatedAliphatic-Aromatic |
| Amine\_Heteroaromatic | Amine\_CyclicSecondary\_SaturatedAliphatic-Heteroaromatic |
| Amine\_Aliphatic | Amine\_CyclicSecondary\_UnsaturatedAliphatic-UnsaturatedAliphatic |
| Amine\_Aromatic | Amine\_CyclicSecondary\_UnsaturatedAliphatic-Aromatic |
| Amine\_Heteroaromatic | Amine\_CyclicSecondary\_UnsaturatedAliphatic-Heteroaromatic |
| Amine\_Aromatic | Amine\_CyclicSecondary\_Aromatic-Aromatic |
| Amine\_Heteroaromatic | Amine\_CyclicSecondary\_Aromatic-Heteroaromatic |
| Amine\_Heteroaromatic | Amine\_CyclicSecondary\_Heteroaromatic-Heteroaromatic |
| AmineT | 3 | AromaticTertiaryAmine | AromaticTertiaryAmine | R is aliphatic or aromatic carbon  “Mix” one R is Aliphatic, one R is Aromatic  Both R’s have same type in other cases | N |
| AliphaticTertiaryAmine | AliphaticTertiaryAmine |
| MixTertiaryAmine | MixTertiaryAmine |
| Hydrazine | 3 | HydrazineTerminalAromatic | HydrazineTerminalAromatic | NH and terminal NH2 is required  R is aliphatic or aromatic  Match hydrazine | N of NH2 |
| HydrazineTerminalAliphatic | HydrazineTerminalAliphatic |
| HydrazineTerminalCarbamoyl | HydrazineTerminalCarbamoyl | A is any atoms including H | N of NH2 |
| 3 | HydrazineNonTerminalAromatic | HydrazineNonTerminalAromatic | R is aliphatic or aromatic  “Mix” one R is Aliphatic, one R is Aromatic  Both R’s have same type in other cases  A is any atoms including H | One N |
| HydrazineNonTerminalAliphatic | HydrazineNonTerminalAliphatic |
| HydrazineNonTerminalMix | HydrazineNonTerminalMix |
| HydroxyAmidine | 1 | HydroxyAmidine | HydroxyAmidine | A is any not H | N |
| Hydroxylamine | 2 | HydroxylamineAromatic | HydroxylamineAromatic | R is aromatic | N |
| HydroxylamineAliphatic | HydroxylamineAliphatic | R is aliphatic  Does not match hydroxylamine | N |
| 2 | SubstitutedHydroxylamineAromatic | SubstitutedHydroxylamineAromatic | A is Carbon  R is aromatic | N |
| SubstitutedHydroxylamineAliphatic | SubstitutedHydroxylamineAliphatic | A is Carbon  R is aliphatic (can be H) | N |
| Grignard | 4 | Grignard\_Aliphatic | Grignard\_SaturatedAliphatic | R: Group 1 | Mg |
| Grignard\_Aliphatic | Grignard\_UnsaturatedAliphatic |
| Grignard\_Aromatic | Grignard\_Aromatic |
| Grignard\_Heteroaromatic | Grignard\_Heteroaromatic |
| Metals | 1 | Metals | Metals | [Zn,Mg,Li,Sn,Hg] | Metal |
| Silicium | 2 | Silicium | Silicium1 | Si- | Si |
| Silicium | Silicium2 | -Si- | Si |
| Hydroxylamine | 1 | Hydroxylamine | Hydroxylamine | NH2OH | N |
| Halogen | 1 | Halogen | Halogen | X2, (e.g. Cl2, Br2, I2) | X |
| Peroxyde | 1 | Peroxyde | Peroxyde | O-O (not in ring) | O |
| XPhosphorus | 1 | XPhosphorus | XPhosphorus | R is [F,Cl,Br,I]  A is any | P |