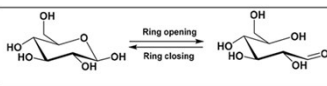
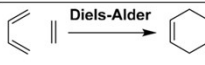
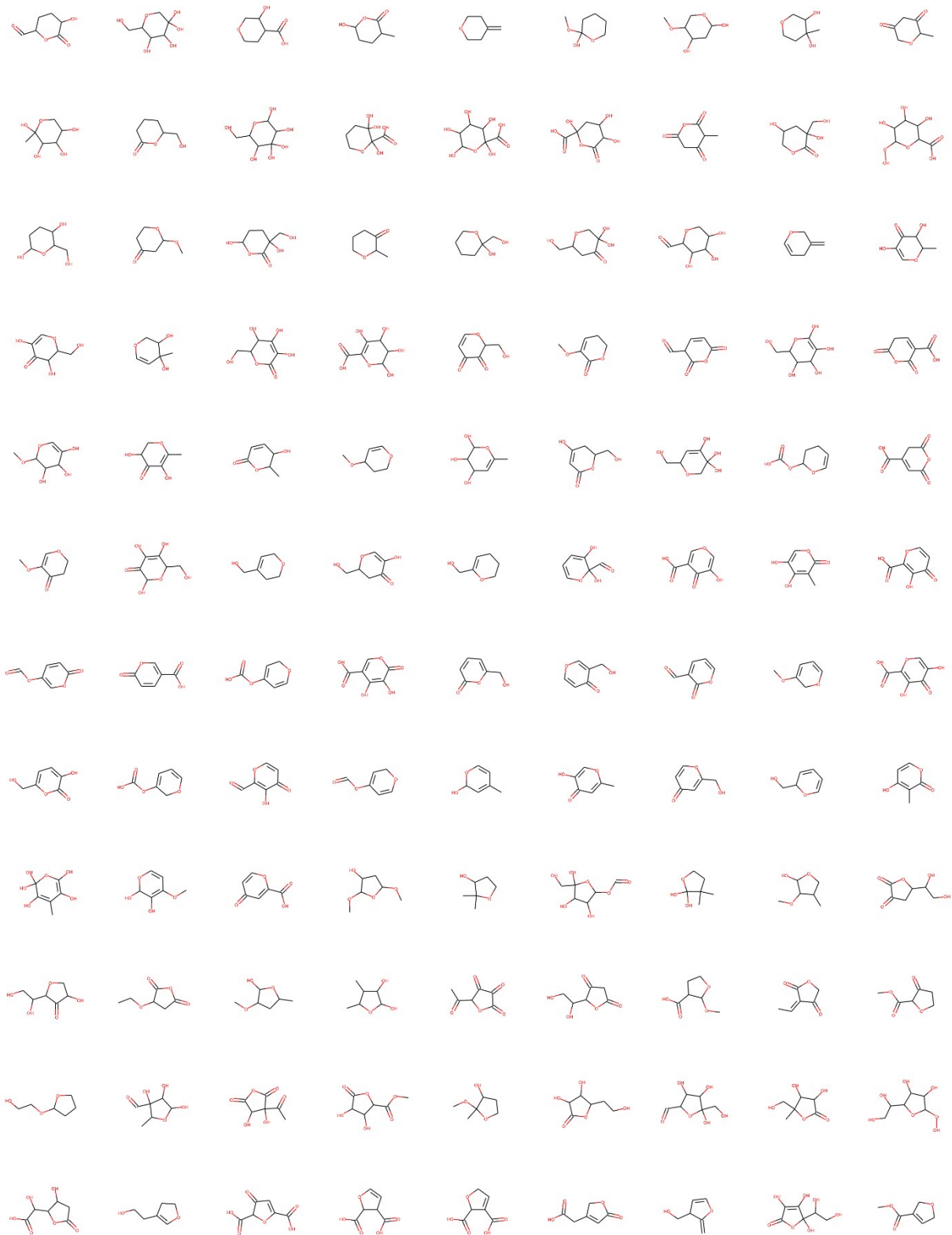
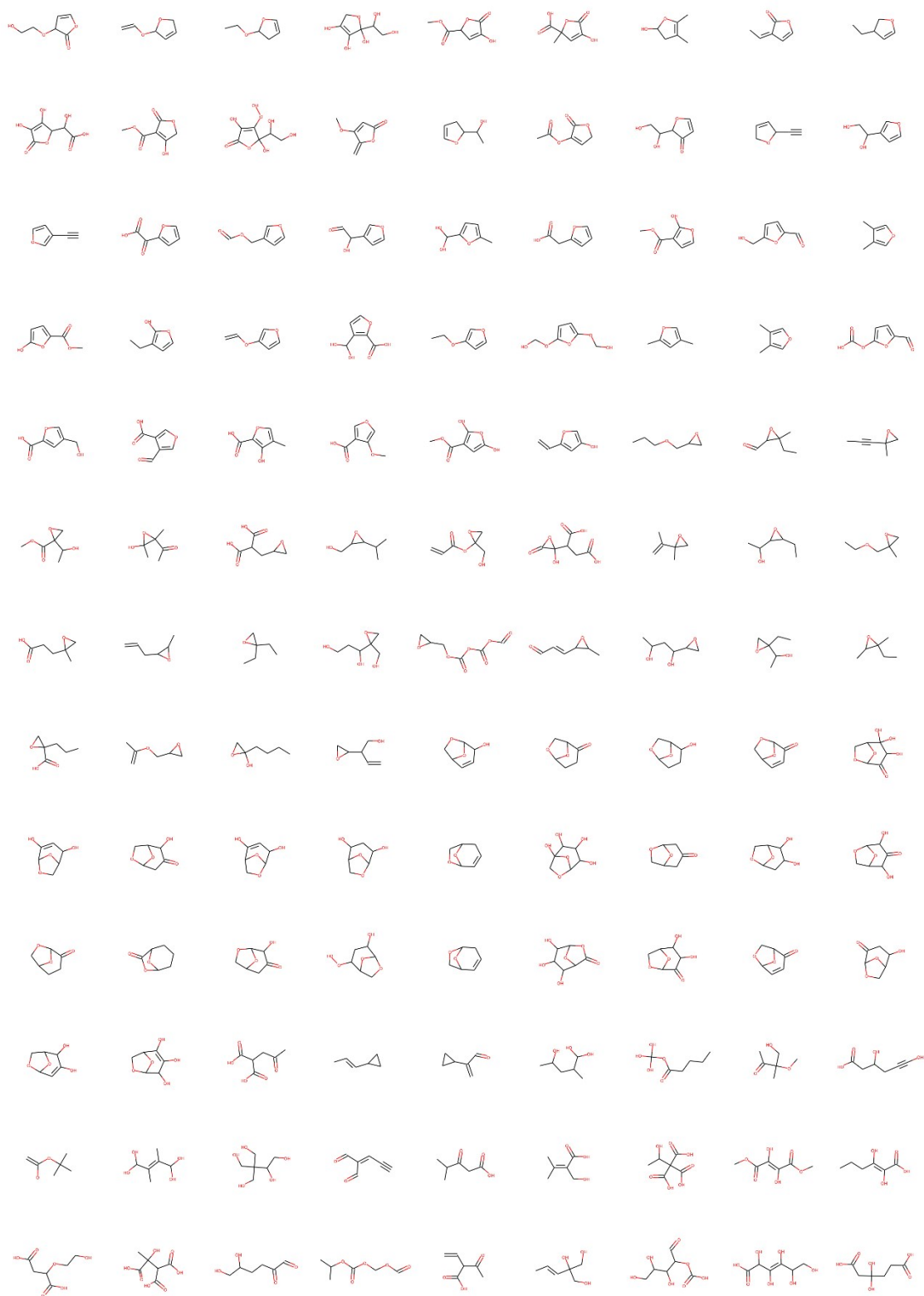


S1. Reaction Families Reproduced from Zhou et al.

Hydrogenation	$\begin{array}{c} \text{R} \text{---} \text{CH}=\text{CH} \text{---} \text{R}' + \text{H}_2 \xrightarrow{\text{Hydrogenation}} \text{R} \text{---} \text{CH}_2\text{---CH}_2\text{---} \text{R}' \\ \text{R}_1 \text{---} \text{C}(=\text{O})\text{---} \text{R}_2 + \text{H}_2 \xrightarrow{\text{Hydrogenation}} \text{R}_1 \text{---} \text{CH}_2\text{---} \text{CH}(\text{OH})\text{---} \text{R}_2 \end{array}$
Hydrogenolysis of C-O bond	$\text{R-OR}' + \text{H}_2 \xrightarrow{\text{Hydrogenolysis}} \text{R-H} + \text{R}'\text{OH}$
Hydrodeoxygenation	$\text{R-OH} + \text{H}_2 \xrightarrow{\text{Hydrodeoxygenation}} \text{R-H} + \text{H}_2\text{O}$
Hydrolysis	$\begin{array}{c} \text{R} \text{---} \text{C}(=\text{O})\text{---} \text{OR}' + \text{H}_2\text{O} \xrightarrow{\text{Hydrolysis}} \text{R} \text{---} \text{C}(=\text{O})\text{---} \text{OH} + \text{R}'\text{---} \text{OH} \\ \text{R-O-R}' + \text{H}_2\text{O} \xrightarrow{\text{Hydrolysis}} \text{R-OH} + \text{R}'\text{---} \text{OH} \end{array}$
Dehydration	$\text{R} \text{---} \text{CH}_2\text{---CH}(\text{OH})\text{---} \text{R}' \xrightarrow{\text{Dehydration}} \text{R} \text{---} \text{CH}=\text{CH} \text{---} \text{R}' + \text{H}_2\text{O}$
Hydration	$\text{R} \text{---} \text{CH}=\text{CH} \text{---} \text{R}' + \text{H}_2\text{O} \xrightarrow{\text{Hydration}} \text{R} \text{---} \text{CH}_2\text{---CH}(\text{OH})\text{---} \text{R}'$
Decarbonylation	$\text{R} \text{---} \text{C}(=\text{O})\text{---} \text{H} \xrightarrow{\text{Decarbonylation}} \text{R-H} + \text{CO}$
Decarboxylation	$\text{R} \text{---} \text{C}(=\text{O})\text{---} \text{OH} \xrightarrow{\text{Decarboxylation}} \text{R-H} + \text{CO}_2$
Ketonization	$\text{R} \text{---} \text{C}(=\text{O})\text{---} \text{OH} + \text{R}' \text{---} \text{C}(=\text{O})\text{---} \text{OH} \xrightarrow{\text{Ketonization}} \text{R} \text{---} \text{C}(=\text{O})\text{---} \text{C}(=\text{O})\text{---} \text{R}' + \text{CO}_2 + \text{H}_2\text{O}$
Esterification	$\text{R} \text{---} \text{C}(=\text{O})\text{---} \text{OH} + \text{R}'\text{---} \text{OH} \xrightarrow{\text{Esterification}} \text{R} \text{---} \text{C}(=\text{O})\text{---} \text{OR}' + \text{H}_2\text{O}$
Keto-enol tautomerization	$\text{R}' \text{---} \text{C}(=\text{O})\text{---} \text{CH}_2\text{---} \text{R}'' \rightleftharpoons \text{R}' \text{---} \text{C}(\text{OH})=\text{CH} \text{---} \text{R}''$
Ring-opening/closing	
Aldol condensation	$\text{R} \text{---} \text{C}(=\text{O})\text{---} \text{CH}_3 + \text{H} \text{---} \text{C}(=\text{O})\text{---} \text{R}'' \xrightarrow{\text{Aldol condensation}} \text{R} \text{---} \text{C}(=\text{O})\text{---} \text{CH}(\text{OH})\text{---} \text{CH}(\text{R}'')\text{---} \text{R} + \text{H}_2\text{O}$
Diels-Alder	
Ketone amination	$\text{R}_1 \text{---} \text{C}(=\text{O})\text{---} \text{R}_2 + \text{H} \text{---} \text{N}(\text{R}_3)\text{---} \text{R}_4 \xrightarrow[\text{Reductant}]{\text{Ketone amination}} \text{R}_1 \text{---} \text{CH}_2\text{---} \text{N}(\text{R}_3)\text{---} \text{R}_4 + \text{H}_2\text{O}$
Epoxidation	$\text{R}_1 \text{---} \text{C}(\text{R}_2)=\text{C}(\text{R}_3)\text{---} \text{R}_4 \xrightarrow{\text{Oxidant}} \text{R}_1 \text{---} \text{C}(\text{R}_2)\text{---} \text{C}(\text{R}_3)\text{---} \text{R}_4$
Selective oxidation	$\begin{array}{c} \text{R} \text{---} \text{CH}_2\text{---} \text{CH}(\text{OH})\text{---} \text{R}' \xrightarrow{\text{Oxidant}} \text{R} \text{---} \text{CH}_2\text{---} \text{CHO} \\ \text{R} \text{---} \text{CH}(\text{OH})\text{---} \text{CH}_2\text{---} \text{R}' \xrightarrow{\text{Oxidant}} \text{R} \text{---} \text{CHO} \\ \text{R} \text{---} \text{C}(=\text{O})\text{---} \text{H} \xrightarrow{\text{Oxidant}} \text{R} \text{---} \text{C}(=\text{O})\text{---} \text{OH} \end{array}$

S2. Training Set for Random Forest Model





S3. Revised C6 Bioprivileged Candidates' CID and SMILES

10129968, C1C=CC(C(O1)CO)O
10154128, COC1CC(C(O1)CO)O
10464455, COC(=O)CCC(=O)CO
10866346, C1C2C(C=CC(O1)O2)O
10942437, C1C2C(C=CC(O1)O2)O
11062673, CC(CC(=O)C(=O)O)C(=O)O
11073541, CC(=O)C1(CCOC1=O)O
11083904, COC1=CC=C(O1)CO
11126217, C=CCC(C1CO1)O
11469087, CC1C(C(=O)O1)C(=O)C
11535562, C1=CC(=O)OC1CCO
11679816, C1C=C(OC1=O)CCO
12043215, COC1CC(C(O1)CO)O
12179941, C1C=CC(C(O1)CO)O
12305536, C1=COC(=CC1=O)C(=O)O
12794200, CC(CC=O)C(=O)OC
13009278, COC1CC(C(O1)CO)O
13327277, COC(=O)CC(CC(=O)O)O
13435928, C1COCC=C1C=O
14088632, CC(=O)C=CCC(=O)O
14321879, C1C=CC(=O)OC1CO
14337696, C1C(CC(=O)OC1CO)O
14607361, COCC1C(CC(O1)O)O
14693661, C1C(C2C(C(O1)CO2)O)O
14985560, CC1=C(COC1=O)C(=O)O
15565217, C1=COC=C1C(C(=O)O)O
17874494, C1C=CC(=CO1)C(=O)O
17905461, C1C(=CC=CO1)C(=O)O
18317855, C1=COC=C(C1=O)C(=O)O
18506391, C1C(O1)CCC(CO)O
18506424, CC1C(O1)CC(CO)O
18506451, CC(CC1C(O1)CO)O
18730707, CC(=O)C(CC(=O)O)OC
18959439, CCC(=O)C(C(=O)O)OC
19089753, C1C=CC(=CO1)C=O
193530, C(CC(C(=O)O)O)CC(=O)O
19778718, C1C=CCOC1C=O
19899299, CC1(OCC(O1)C=O)OC
20039212, C1C=C(C=CO1)C(=O)O
20061471, C1C(=CC=CO1)C=O
20591616, C=C1CC(CC(=O)O1)O
20623870, C1COC=C(C1=O)C=O
21299965, C=CC(=O)CCOC=O
21699374, C1C=C(C=CO1)C=O
21987554, C1=COC(=C1CO)CO
22015953, CC(=CC(=O)C)C(=O)O
22467004, CC(=O)C(=C)CC(=O)O
22768481, CCC(C(=O)C(=O)O)C(=O)O
22845691, C1CC(=O)OC1C(CO)O
229740, COC1C2C(O2)C(O1)CO
23198947, C1CC(OC1)C(C(=O)O)O
23628471, CC(CC(=O)O)C(=O)C(=O)O

237332, C1=C(OC(=C1)C=O)CO
276679, C=C1C(COC1=O)C(=O)O
284624, C=C(C(=C)C(=O)O)C(=O)O
310, C(=CC(=O)O)C=CC(=O)O
392323, CC(C(=O)O)C(C)(C(=O)O)O
439619, CCC(C(C(=O)O)O)C(=O)O
439700, CC(C(=C)C(=O)O)C(=O)O
45039803, COC1CC(C(O1)CO)O
5280615, CC(C(=O)O)C(C)(C(=O)O)O
5325916, C=C1CC(OC1=O)CO
53641043, CC(=O)C1CC(=O)OC1=O
53799488, C1CC(OCC1O)C(=O)O
53953812, CC(=C(C(=O)O)OC)C(=O)O
53971420, C1C(C(OC1=O)CC(=O)O)O
54001154, C1C(C1=O)CCC(=O)O
54247055, CC(=CC=O)C(=O)OC
542, C1=CC(=O)OC1CC(=O)O
54405733, CC(CC(=O)O)(C(=O)OC)O
54456019, CC(C=CC(=O)O)C(=O)O
54527780, CC(=CC(=O)C)C(=O)O
54675757, CC1=CC(=CC(=O)O1)O
54686175, C1=C(C=C(OC1=O)CO)O
54728446, CC1=C(C(=CC(=O)O1)O)O
550907, C1COC(CC1=O)C(=O)O
55299096, C1C(COCC1O)C(=O)O
55301806, C1CC(OC1)(C=O)C(=O)O
55302884, C1COC(CC1O)C(=O)O
557816, C1COC(C=C1)C=O
558881, C1CC(=O)OC1C(CO)O
57121014, CCC(=CCO)C(=O)O
57127683, C1C(O1)C(C2C(O2)CO)O
57186783, CCC(CO)(C(=O)O)OC
57763310, CCC(=O)C=CCO
58069330, CC=C(COC=O)C(=O)O
58171640, CCC(COC=O)(C(=O)O)O
61870903, CC(CCOC)(C(=O)O)O
643506, COC1CC(C(O1)CO)O
67000802, CC1=C(OC(=O)C1(C)O)O
67454773, C1COCC(C1=O)C(=O)O
68141, C1=CC(=O)OC=C1C(=O)O
68179366, C1=COC(=O)C=C1C(=O)O
86860, C1=COC(=C1)C(C(=O)O)O
87528182, CC=CC(C)(C(=O)O)O
87617779, CCC(CC(=O)O)(O)OC
87885427, C1COC(C1C(=O)O)C=O
88465184, C=CCC(C=O)C=O
89987442, C1C(C(C(O1)C2CO2)O)O
9824904, CC1C(=O)C(=O)C=CO1
99884, CC(=CC(=O)O)CC(=O)O

S4. C5 Bioprivileged Candidates' CID and SMILES

10219405, C1C=CC(=O)C(O1)O
10351658, CCCC(C=O)O
10464328, C1C=C(C(=O)O1)CO
10678218, COCOC(=C)CO
10909501, COCC(=C)C=O
10942356, C=CCC(CO)O
10948203, CC(C=O)C(C)O
10997012, CC(=O)C(C)(CO)O
10997016, CC(CCO)(CO)O
11051586, CC(C=O)C(=O)C
11126191, C=CC(CCO)O
11182580, C(CO)C=CCO
11446296, C=C(CCO)CO
11469098, CC1(C(COC1=O)O)O
11579, CC(=O)CCC(=O)O
12006600, CC(C(=O)C=C)O
12200912, C1=COC(=C1)C=O
12236733, COCCC(=O)CO
12514401, C1COCC(C1=O)O
12556265, CC(=O)C(=C)CO
126055, C1=COC(C1O)CO
12962854, COCC(C=O)OC
13199216, CCC(C)(C=O)O
13311642, CCC(CC=O)O
13536808, CC=CC(C(=O)O)O
13550245, C=CC(C1CO1)O
14141023, C1C(OC(=O)C1O)CO
14512532, CCC(CO)C=O
14660056, CC(C(C=C)O)O
150996, CC(C(=O)C)C(=O)O
157706, C1C(=CC(=O)O1)CO
15918561, C=CC(=O)CCO
15927954, CC(CO)C(=O)C(=O)O
169324, CC(C=C)(C(=O)O)O
171001, C1C2C(C(C(O1)O2)O)O
18000384, CC(COC)(C(=O)O)O
181583, CC(=C)C(CO)O
18172379, C1=COC(=CO1)C(=O)O
18506440, C1C(O1)(CCO)CO
185554, CC(CO)(C=C)O
18783249, C1C(=O)OCC1(CO)O
19382908, CC(C=C)(C=O)O
19384762, C1COC(=O)C1(CO)O
19707998, COC(=O)C(CCO)O
19884375, CC(CCO)C=O
20069322, C1=C(OC1=O)CCO
20618051, CC(=C)C(=O)CO
21328181, C1C(O1)CC(CO)O
21441687, COCCC(=O)C(=O)O
21474940, COCC(=C)C(=O)O
21537108, CCC(C(=O)CO)O
21911137, C1C(C(=O)CO1)C(=O)O

22481615, CC(C(COC)O)O
22600970, CC(C(C)(CO)O)O
22857038, C1C(OCC1=O)CO
290511, C1C(=CC(=O)O1)C(=O)O
3025731, C1C(=O)C=C(O1)CO
439684, CCC(=O)CC(=O)O
4961356, CC(C(=O)C(=O)O)C(=O)O
51418889, CC(CCC(=O)O)O
5312944, C=CC(=O)CC(=O)O
5314562, COCC(=O)C=C
5318286, C=C1C(COC1=O)O
53437016, CC=C(CO)C(=O)O
5368088, COC(=O)C=CCO
53806232, CC(=CCO)CO
53859644, C1=COC(=O)C1CO
53993215, C(C=CCO)C(=O)O
54104277, CC(=CCO)C(=O)O
54161810, CC(C=CCO)O
54173442, CCC(=O)C(=O)CO
54196276, COCC=CC=O
54232530, COC(C=O)C(=O)OC
54261087, CC=CC(CO)O
544596, CC(CC=O)CO
55255160, C1=COC(C1=O)CO
55269592, C1C(COC1C(=O)O)O
57007257, CC(C(=O)C(=O)OC)O
57023262, CC(=O)C=CCO
57028153, COC(=O)CC(=O)CO
57035808, C(CC(=O)O)COCO
57127623, CC(C(CO)OC)O
57214022, C1=CC(OC1CO)O
57290522, C(C1C2C(O2)C(O1)O)O
57444232, C1=COC(=CO1)C=O
57827549, COCC(C=C)O
57840988, C(C(=O)COC=O)C(=O)O
58104777, CC(CO)(C(=O)O)OC
58649439, C1COC(C1O)CO
58801266, CC(C(C)O)C(=O)O
59771871, CC1(C(=O)COC1=O)O
59771874, CC(C=O)(C(=O)OC)O
59947295, CC(CC(=O)O)C=O
638129, CC(=CC(=O)O)C(=O)O
72739553, C1C(O1)C=CC(=O)O
811, C=C(CC(=O)O)C(=O)O
91260, C1C(O1)C=CC=O
91419129, C1COCC(=O)C1O
91426886, CC1(COC(=O)C1O)O
976, C=CCC(=O)C(=O)O
9866019, C1C(=O)C(=O)C=CO1
9877343, CC(=O)C(=O)C=C

S5. C7 Bioprivileged Candidates' CID and SMILES

10034839, COC1=CC(=O)OC(=C1)CO
10531300, COCC1=CC(=O)C(=CO1)O
10942420, CCC(=O)C=CC(=O)C
1125227, COCC1=CC=C(O1)C(=O)O
11320989, CCC(=O)C=C(C)C(=O)O
11435181, COC(=O)C1=CC=C(O1)C(=O)O
117857437, CCC(C=CCO)C(=O)O
11804854, COC(=O)C=CCCC=O
11992679, COC1=COC(=CC1=O)C(=O)O
12524268, COC1=COC(=CC1=O)C=O
12957800, C(CCC(=O)O)CC(C(=O)O)O
13129932, CC(CO)C(=O)C(=C)C
13265879, COCC1=COC=C1C(=O)O
13504598, CCC(CO)(COC)C=O
13699526, COCC(=O)C1=CC=CO1
13793774, COCC1=CC=C(O1)CO
14196805, CC(=O)C(CCC(=O)O)C(=O)O
14383705, CC(C(=O)CC(=O)O)C(=O)OC
14814246, CCCC(C=O)C(=O)C
14869905, CC(=CCO)CCC(=O)O
15818585, COC(=O)C1=C(OC=C1)C=O
17801465, COCC1=C(OC=C1)C=O
17940644, COC(=O)CC(CCC(=O)O)O
18678805, CC(C1=C(C(=O)C=CO1)O)O
18767333, COC(=O)CCC(CC(=O)O)O
18983630, COC(=O)C1=COC(=C1)C=O
19424577, C1CCOC(C1)C(CO)O
19784052, CC(CC(=C)C(=O)O)C(=O)O
19870930, CCCC(=O)C(=O)C=C
19989345, COC(=O)CC(=O)C(=O)C(=O)OC
20431593, C1C(O1)CCC(=O)CCO
2049778, C1=COC(=C1)CC(=O)C(=O)O
20515422, C(CC(C(=O)O)O)CC(C(=O)O)O
20977963, C=C(CCC(=O)O)CC(=O)O
21422142, COCC1=C(OC=C1)C(=O)O
21520674, C1=C(OC(=C1)C(=O)O)CC(=O)O
21554958, CCCC(=O)C(=C)C(=O)O
21792544, COC1=C(OC=CC1=O)C(=O)O
21805186, C=CC(CCC(=O)O)C(=O)O
21877115, CCC(C)C(=O)C(=O)C(=O)O
22182133, CC(=C)C(=O)OC(C)(CO)O
22313166, CCCC(C(=O)C(=O)O)C(=O)O
22364041, CCC(C(=O)C=C)C(=O)O
22630163, CC(C=CCC(=O)O)C(=O)O
22965133, CC1C(C(C(O1)COC)O)O
23094424, COCC1=C(C(=O)C=CO1)O
234556, COC1=COC(=CC1=O)CO
23629603, CCC(CC(=O)O)C(=O)C(=O)O
292847, CC(=C(C)C(=O)O)CC(=O)O
4337439, C1=COC(=C1C(=O)O)CC(=O)O
4632710, CCCC(C(C(=O)O)O)C(=O)O
4737732, CC1=C(C(=CO1)C(=O)O)C(=O)O

53754763, CCCC=CC(=O)C=O
53764534, CC(=C)C(=O)OC=CC(=O)O
53798196, CC1=CC=C(O1)C(C(=O)O)O
54000959, CC(C(C(=O)C)C(=O)O)C(=O)O
54069022, CC1C(C(C(CO1)OC)O)O
54406529, CCC=C(CC(=O)O)C(=O)O
54410845, CC(=C)C=C(C)CO
54411426, CC=C(CCC(=O)O)C(=O)O
54446973, CCCC=C(CO)C(=O)O
57072290, CCC(CC(=O)O)(CO)OC
57082174, CC(=CCCC(=O)O)C(=O)O
57095184, COC1COC(C1OC)C=O
57137898, CCC(C(=O)CC(=O)O)C(=O)O
57314944, CC(C(=CCO)C)C(=O)O
574153, COC(CO)C1=CC=CO1
57517040, COC(=O)C1=COC=C1C=O
57599440, CC(C(C(=O)C)C(=O)O)OC
57723949, COCC1C(C2C(O2)CO1)O
5908441, CCCC(=CC(=O)O)C(=O)O
67962874, CC1C(C(C(O1)CO)OC)O
68421701, CC1=CC(=C(O1)C(=O)O)C(=O)O
70064563, COCC1=C(OC=C1)CO
71330544, COCCCCC(C(=O)O)O
74711, COCC1=CC=C(O1)C=O
84649887, COC(=O)C1=CC(=CO1)C=O
85853067, CC=CC(CC(=O)O)C(=O)O
87060121, CC(=C(CC=C)O)C(=O)O
87060123, CC(=CCC=CO)C(=O)O
87089770, C=C(CCC(=C)O)C(=O)O
87089771, CC(=CCC(=C)O)C(=O)O
87145573, CCC(=C(CO)C(=O)O)C
87224660, CC(=C(C)C(=O)O)C(=O)C(=O)O
87224751, CC(CCC(=O)O)C(=O)C(=O)O
87224819, CC(CC(=O)O)C(C(=O)OC)O
87224866, CC(C(C)C(=O)O)C(=O)C(=O)O
87270928, CC(=C(CC=C)C(=O)O)O
87270931, CC=C(CC=CO)C(=O)O
87329265, C=C(CCC=CO)C(=O)O
87466770, CCC(CC(C(=O)O)O)OC
87467117, CCC(CC(=O)C(=O)O)OC
87847309, CC(=C(C(=O)C)C(=O)O)OC
88001916, CC(C(=O)CCOC)C(=O)O
88399780, C(CC=CC(=O)O)COC=O
89056492, CCC(=O)C=CCOC
89419379, CC1C(C(C(O1)CO)O)OC
89686662, CC1(CCC(CO1)O)C(=O)O
91570464, CCC(C)C(C(=O)CO)O
9167334, CC1=C(C=C(O1)CO)C(=O)O

S5. C4 Bioprivileged Candidates' CIDs and SMILES

10129908, C=CC(C(=O)O)O
10154000, CCC(C=O)O
10338, C=CC(CO)O
10413, C(CC(=O)O)CO
10920591, C(C1C(O1)C(=O)O)O
10931385, COCOCC=O
11029762, CC(CO)C=O
11116042, COCOCC(=O)O
1112, C(CC(=O)O)C=O
111509, CC(=O)CCO
11217234, CC(CO)C(=O)O
12130083, C1C(=O)C(=O)C(=O)O1
12256008, COCC(C(=O)O)O
12369703, C=CC(=O)CO
13061653, CC(CC=O)O
13133780, C1C(OC1)C=O
137720, COCCC=O
145744, CCC(=O)C=O
14926651, COCC(=O)CO
15084311, C=C(CO)C=O
156179, COCC(CO)O
162804, C=CC(=O)C(=O)O
163325, C(C=CC=O)O
167605, C=CC(=O)C=O
17894303, CC(C(=O)O)(O)OC
18364059, C1C(C1O)C=O
18386435, C1C(=O)C(=O)OC1
18680744, C1OC=C(O1)C=O
18728307, COCC(=O)C=O
199534, C=C(CO)C(=O)O
19973299, C(CO)C(=O)OCO
20511974, COCC(=O)C(=O)O
20753240, C(C(=O)O)C(=O)OCO
20766, COC=CC=O
21184345, C1OC=C(O1)C(=O)O
21267695, C1C(O1)C(=O)C(=O)O
21301831, COC(C=O)C=O
21924930, CC(C(O)OC)O
222656, C(C(C(=O)O)O)C(=O)O
22621441, CC(O)OCC(=O)O
22997012, C1C(C(OC1)O)O
22999750, COC(CCO)O
23035738, CCOC(C(=O)O)O
23064804, CC(COC(=O)O)O
23159180, CC(=O)C(=O)CO
25935, C(COCC(=O)O)O
296, CC(C=O)C(=O)O
441, CC(CC(=O)O)O
54517555, COC(CO)C(=O)O
54675764, C(=C(C(=O)O)O)C(=O)O
57024947, CC(COCO)O
58038867, C(C(C(=O)O)O)OC=O

58523990, C(C(O)OC=O)C(=O)O
59432048, C1C(=O)OC(O1)CO
59770612, CC(=O)CC(=O)O
81958, C(C=O)OCC=O
91196040, C=C(C=O)C(=O)OO
93093, C(CC=O)CO
970, C(C(=O)C(=O)O)C(=O)O