Supporting Information

Automated Quantum Chemistry for Estimating

Nucleophilicity and Electrophilicity with Applications to

Retrosynthesis and Covalent Inhibitors

Nicolai Ree,[†] Andreas H. Göller,[*,‡] and Jan H. Jensen[*,†]

†*Department of Chemistry, University of Copenhagen, Universitetsparken 5,*

2100 Copenhagen Ø, Denmark

‡*Bayer AG, Pharmaceuticals, R&D, Computational Molecular Design,*

42096 Wuppertal, Germany

E-mail: andreas.goeller@bayer.com; jhjensen@chem.ku.dk

Overview of the QM-based workflow:

Input SMILES

Run a set of SMIRKS to generate possible MCA and MAA products

Generate 1 + 3 · (number of rotatable bonds) or a maximum of 20 conformers using RDKit

1. Generate 1 + 3 · (number of rotatable bonds) or a maximum of 20 conformers using RDKit
2. **Run GFNFF-xTB ALPB(DMSO) optimizations**
3. **Check input/output connectivity and flag if mismatch**
4. **Extract conformers with relative energies ≤ 10 kJ/mol**
5. **Select unique conformers using Butina clustering (RMSD > 0.5 Å)**
6. **Run GFN1-xTB ALPB(DMSO) optimizations**
7. **Check input/output connectivity and flag if mismatch**
8. **Identify the lowest energy conformer**
9. **Run single-point r^{2}SCAN-3c SMD(DMSO) calculations using ORCA**
10. **Check convergence and flag if error**

Calculate MCA and MAA for all nucleophilic and electrophilic sites, respectively

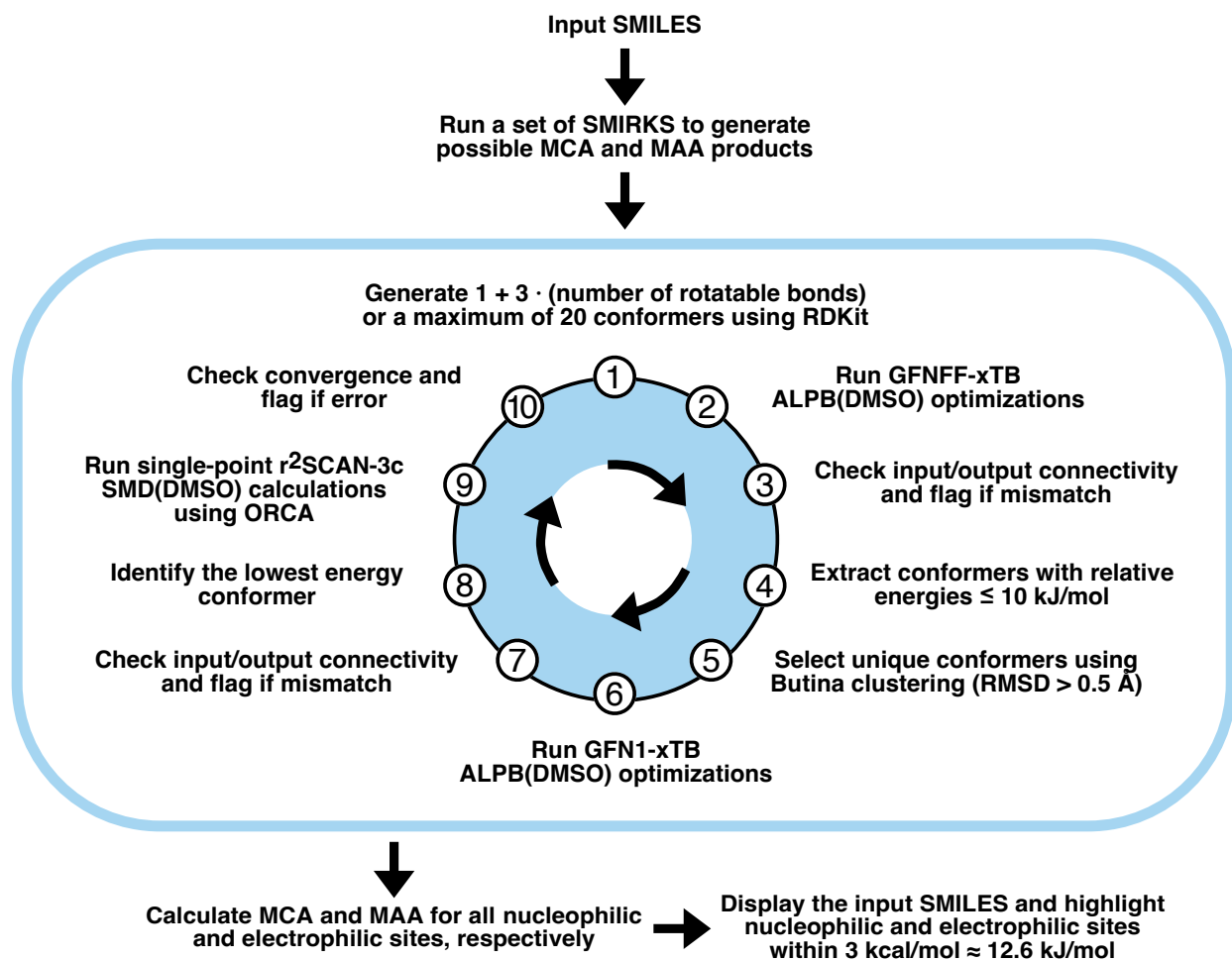**Display the input SMILES and highlight nucleophilic and electrophilic sites within 3 kcal/mol ≈ 12.6 kJ/mol**

Figure S1: Flowchart describing the automated QM-based workflow for identifying possible nucleophilic and electrophilic sites and computing methyl cation affinities (MCAs) and methyl anion affinities (MAAs) to quantify nucleophilicity and electrophilicity, respectively. An example of the visual output can be seen in Figure S3.

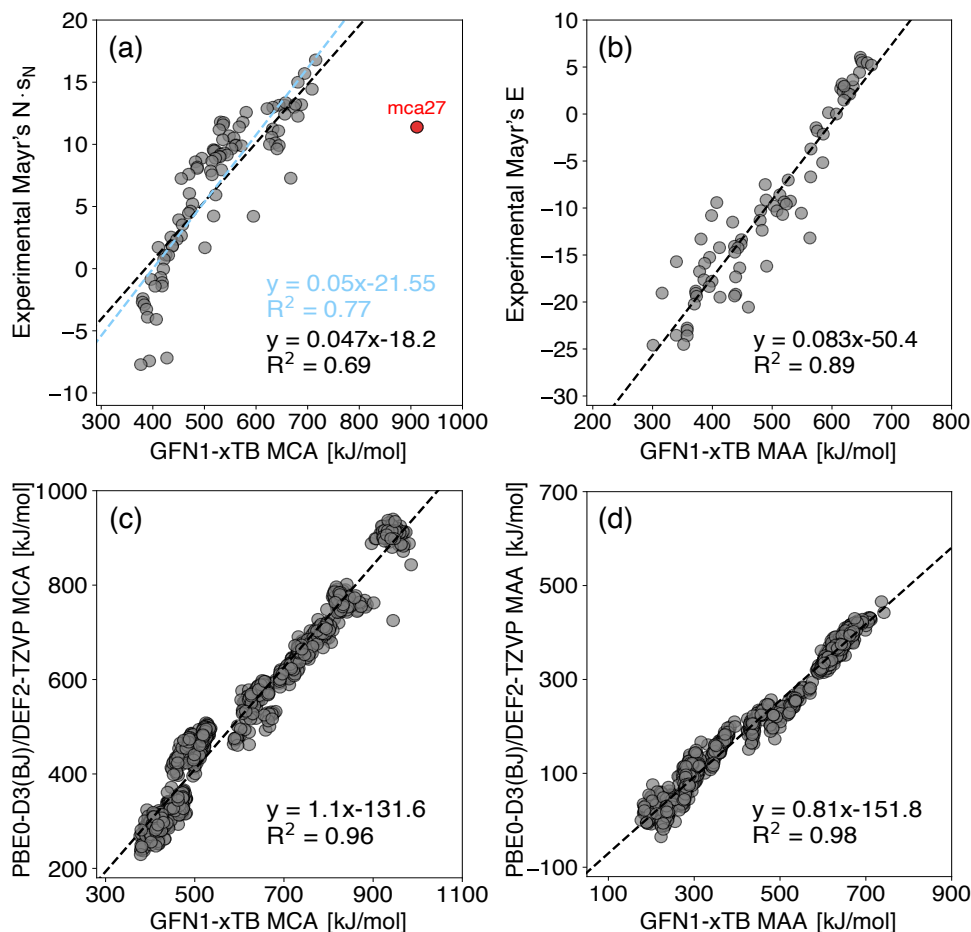# Results of QM-based workflow w/o DFT single-point calculations:

Figure S2: Correlation plots between (a) experimental Mayr's $N \cdot s_N$ and calculated MCAs, (b) experimental Mayr's E and calculated MAAs, (c) calculated MCAs at different levels of theory, and (d) calculated MAAs at different levels of theory. Calculations on the x- and y-axes are performed at the GFN1-xTB ALPB(DMSO) and PBE0-D3(BJ)/DEF2-TZVP COSMO(∞) level of theory, respectively. The black regression lines consider all points, whereas the blue regression line omits compound mca27 (cyanide: [C-]#N).

Chemical selectivity predictions for Raltegravir:

SMILES:
Cc1nnc(C(=O)Cl)o1.Cn1c(C(C)(C)N)nc(C(=O)NCc2ccc(F)cc2)c(O)c1=O

Electrophilicity (MAA)			
Atom ID	MAA Value [kJ/mol]	Error Log (Reactant, Product)	Type
1.5	395.43	None, Connectivity	Acyl halide
2.16	305.65	None, Connectivity	Double bond
1.1	288.53	None, None	Double bond
1.4	211.75	None, Connectivity	Double bond
1.3	201.21	None, None	Double bond
2.9	186.05	None, Connectivity	Amide
2.22	180.44	None, Connectivity	Double bond
2.20	156.85	None, None	Double bond

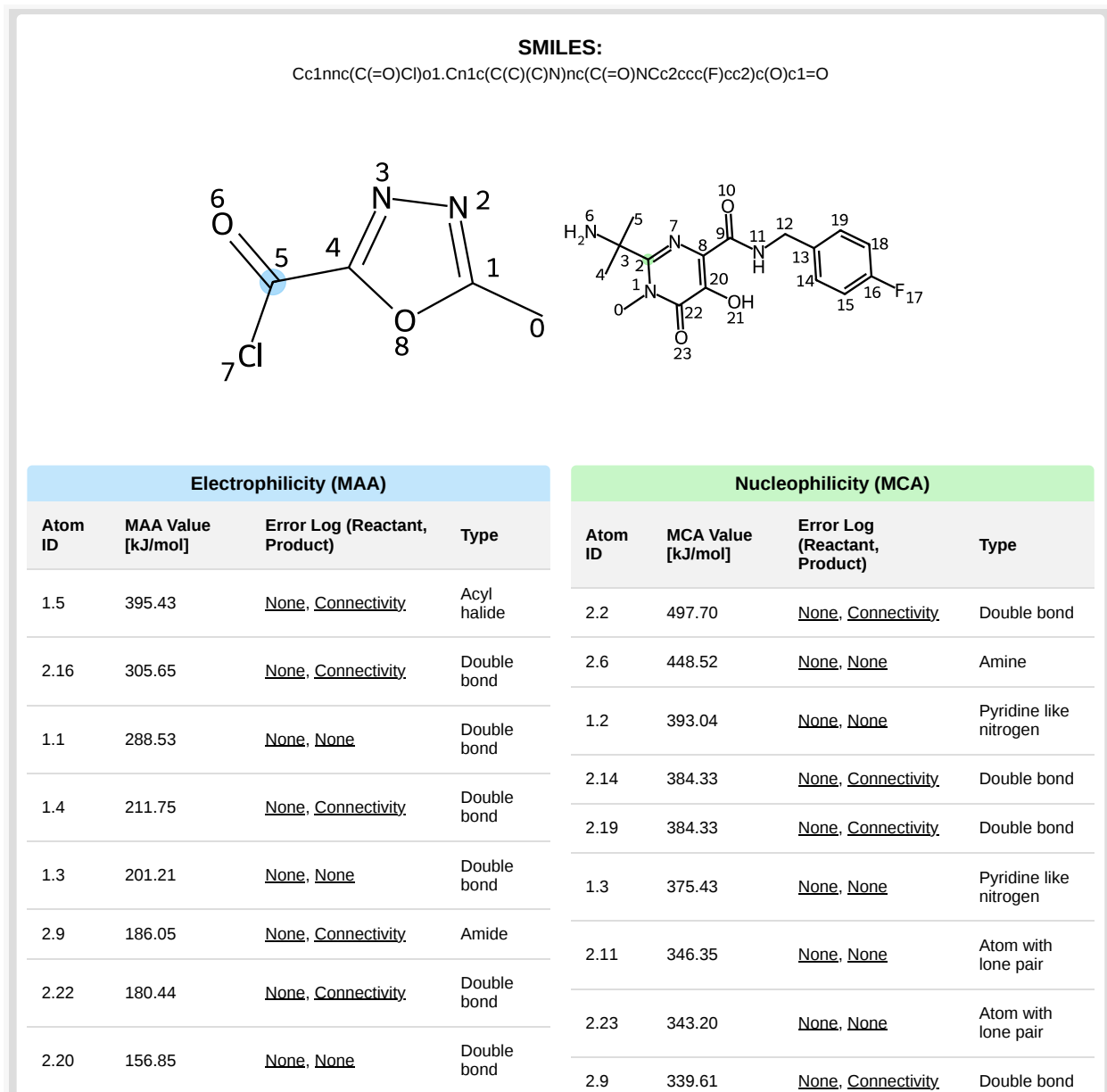| Nucleophilicity (MCA)			
Atom ID	MCA Value [kJ/mol]	Error Log (Reactant, Product)	Type
2.2	497.70	None, Connectivity	Double bond
2.6	448.52	None, None	Amine
1.2	393.04	None, None	Pyridine like nitrogen
2.14	384.33	None, Connectivity	Double bond
2.19	384.33	None, Connectivity	Double bond
1.3	375.43	None, None	Pyridine like nitrogen
2.11	346.35	None, None	Atom with lone pair
2.23	343.20	None, None	Atom with lone pair
2.9	339.61	None, Connectivity	Double bond

Figure S3: Chemical selectivity predictions in the synthesis of Raltegravir. Electrophilic and nucleophilic sites with MAAs and MCAs within 3 kcal/mol $\approx$ 12.6 kJ/mol are highlighted in blue and green, respectively. The MAAs and MCAs are obtained at the r^{2}SCAN-3c SMD(DMSO)//GFN1-xTB ALPB(DMSO) level of theory. Atom site 2.2 has a connectivity failure for the MCA product corresponding to the $-C^+(CH_3)(CH_3)NH_2$ group being detached from the ring, as the CH_3^+ takes its position. Hence, atom site 2.6 is the most nucleophilic site.

2.2	151.45	None, None	Double bond
1.6	129.58	None, None	Double bond
1.2	111.14	None, Connectivity	Double bond
2.8	106.81	None, None	Double bond
2.14	97.80	None, None	Double bond
2.19	97.80	None, None	Double bond
2.15	76.75	None, None	Double bond
2.18	76.75	None, None	Double bond
2.13	60.26	None, None	Double bond
2.23	0.64	None, None	Double bond
2.7	0.43	None, None	Double bond
2.10	-9.56	None, None	Double bond

2.7	333.88	None, None	Pyridine like nitrogen
2.21	329.96	None, Connectivity	Phenol
2.10	323.50	None, None	Amide
2.15	299.03	None, None	Double bond
2.18	299.03	None, None	Double bond
2.13	293.56	None, None	Double bond
2.22	289.71	None, Connectivity	Double bond
2.20	282.17	None, None	Double bond
2.8	278.52	None, None	Double bond
2.1	261.22	None, None	Atom with lone pair
2.16	256.31	None, None	Double bond
1.6	243.13	None, None	Atom with lone pair
1.4	233.88	None, None	Double bond
1.1	179.15	None, None	Double bond
1.7	176.05	None, None	Atom with lone pair
1.5	158.16	None, Connectivity	Double bond
1.8	137.65	None, None	Ether
2.17	130.80	None, None	Atom with lone pair

Figure S3: Continued.
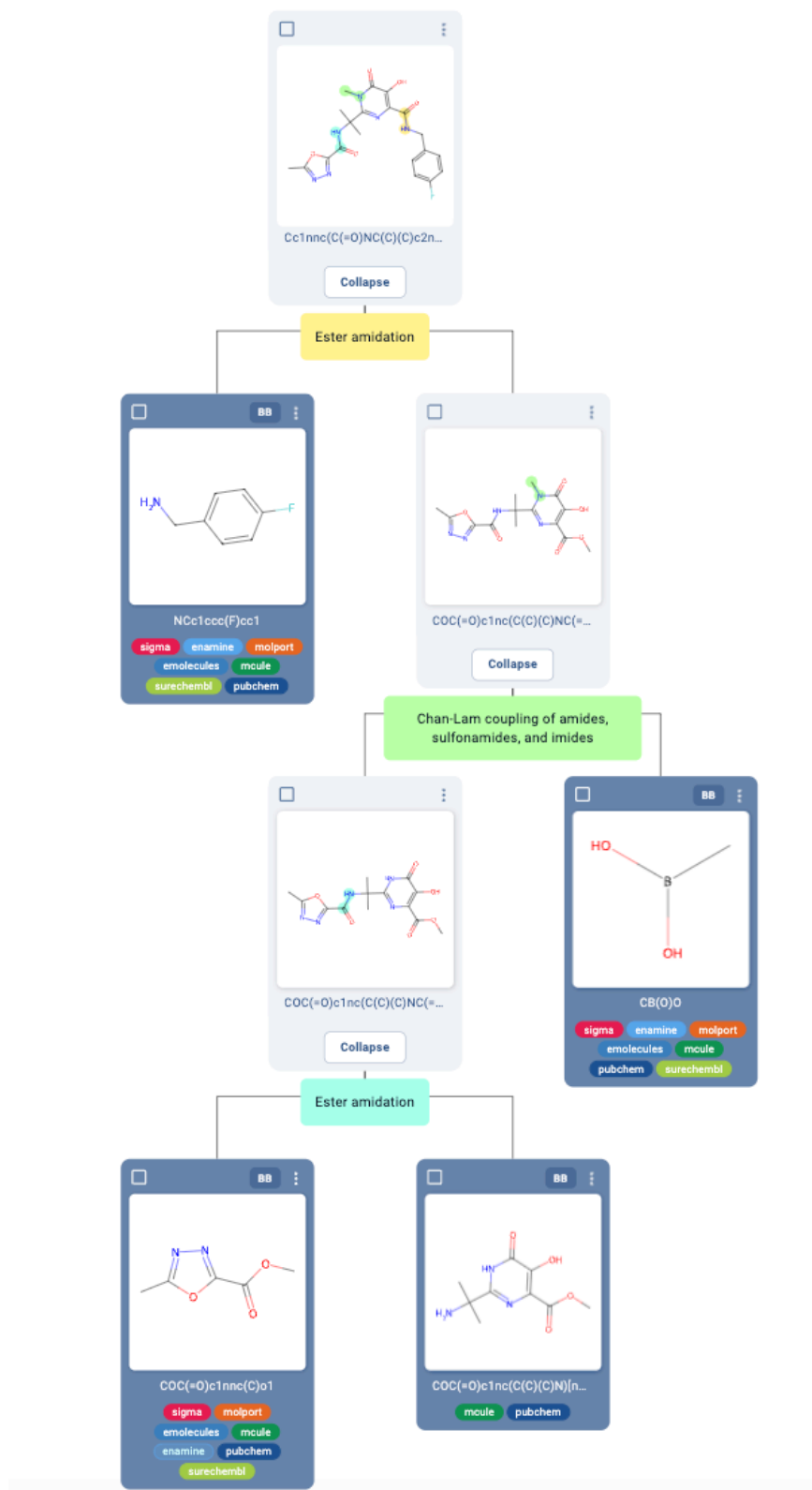
Figure S4: Retrosynthetic route for synthesizing Raltegravir predicted by Manifold from
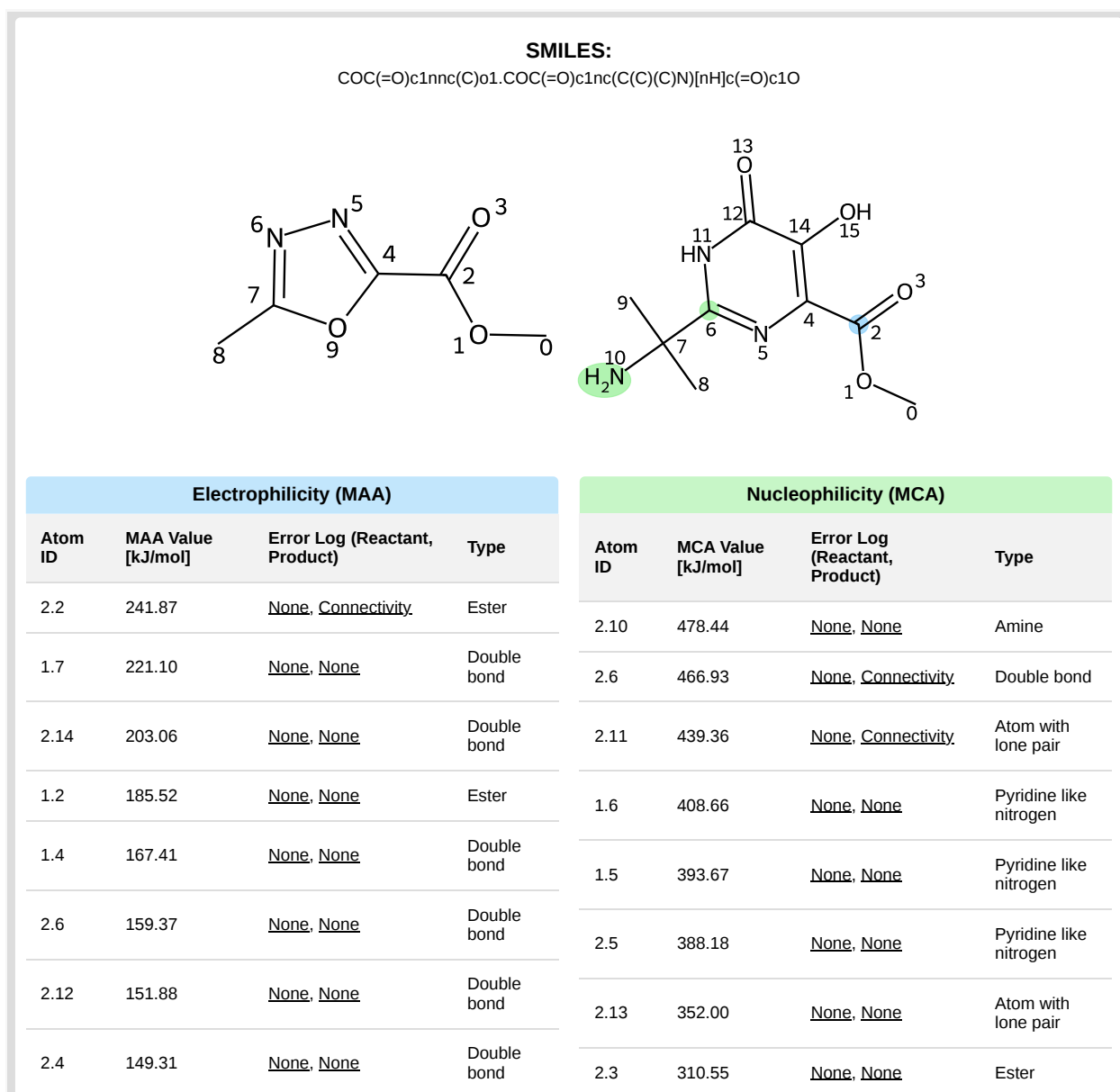`PostEra`.

SMILES:

COC(=O)c1nnc(C)o1.COC(=O)c1nc(C(C)(C)N)[nH]c(=O)c1O

Electrophilicity (MAA)			
Atom ID	MAA Value [kJ/mol]	Error Log (Reactant, Product)	Type
2.2	241.87	None, Connectivity	Ester
1.7	221.10	None, None	Double bond
2.14	203.06	None, None	Double bond
1.2	185.52	None, None	Ester
1.4	167.41	None, None	Double bond
2.6	159.37	None, None	Double bond
2.12	151.88	None, None	Double bond
2.4	149.31	None, None	Double bond

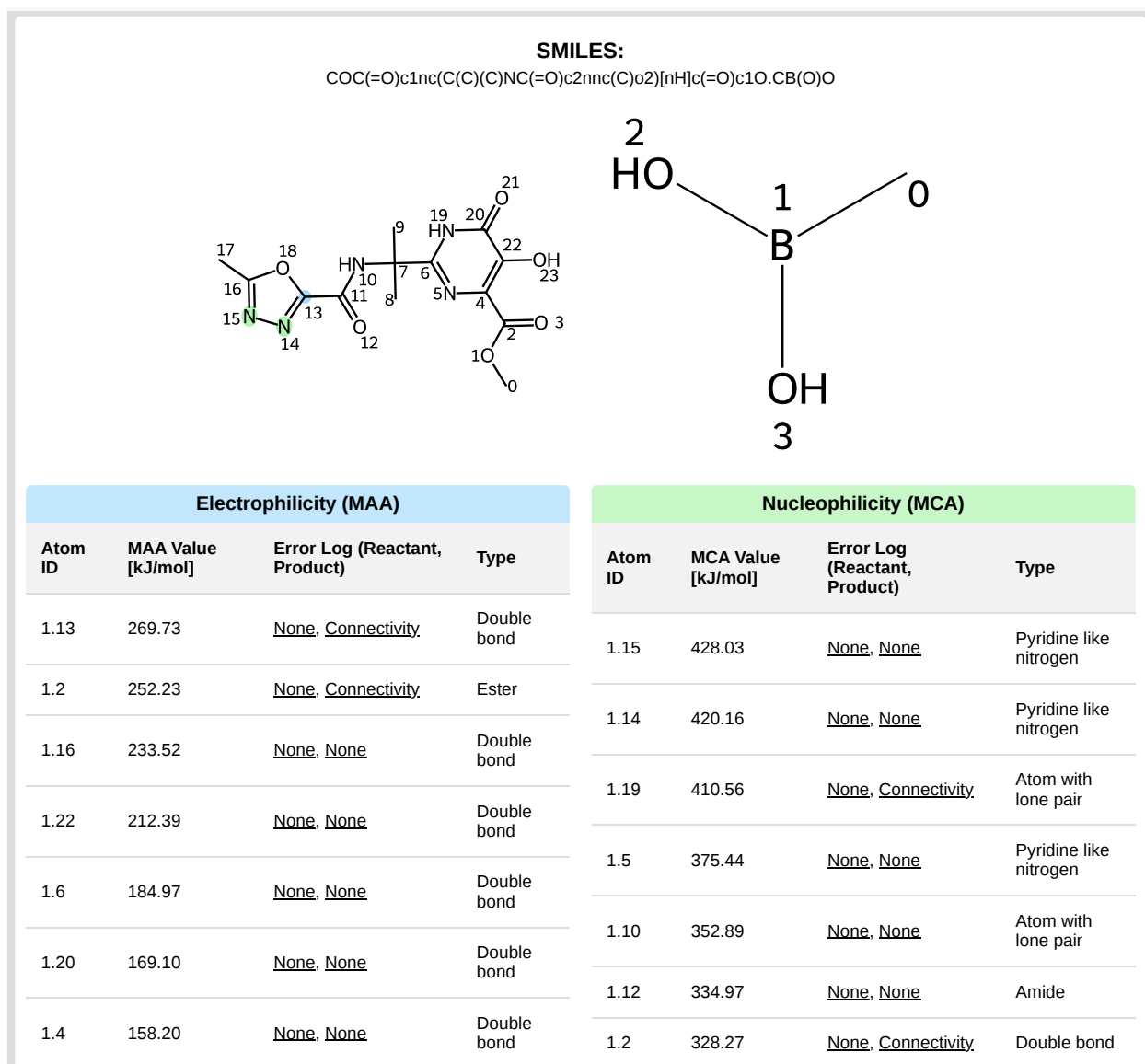| Nucleophilicity (MCA)			
Atom ID	MCA Value [kJ/mol]	Error Log (Reactant, Product)	Type
2.10	478.44	None, None	Amine
2.6	466.93	None, Connectivity	Double bond
2.11	439.36	None, Connectivity	Atom with lone pair
1.6	408.66	None, None	Pyridine like nitrogen
1.5	393.67	None, None	Pyridine like nitrogen
2.5	388.18	None, None	Pyridine like nitrogen
2.13	352.00	None, None	Atom with lone pair
2.3	310.55	None, None	Ester

Figure S5: Calculated MAAs and MCAs for the first reaction step in the retrosynthetic route by Manifold (see Figure S4). The MAAs and MCAs are obtained at the r^{2}SCAN-3c SMD(DMSO)//GFN1-xTB ALPB(DMSO) level of theory.

1.5	91.82	<u>None</u>, <u>None</u>	Double bond
2.13	28.99	<u>None</u>, <u>None</u>	Double bond
1.3	26.07	<u>None</u>, <u>None</u>	Double bond
2.3	12.72	<u>None</u>, <u>None</u>	Double bond
2.5	-10.35	<u>None</u>, <u>None</u>	Double bond
1.6	-27.10	<u>None</u>, <u>None</u>	Double bond

2.15	301.03	<u>None</u>, <u>Connectivity</u>	Phenol
2.12	300.48	<u>None</u>, <u>Connectivity</u>	Double bond
2.14	300.24	<u>None</u>, <u>Connectivity</u>	Double bond
2.2	296.71	<u>None</u>, <u>Connectivity</u>	Double bond
2.4	272.02	<u>None</u>, <u>None</u>	Double bond
1.3	271.56	<u>None</u>, <u>None</u>	Ester
1.2	268.66	<u>None</u>, <u>Connectivity</u>	Double bond
1.4	255.16	<u>None</u>, <u>None</u>	Double bond
2.1	246.00	<u>None</u>, <u>None</u>	Atom with lone pair
1.7	206.90	<u>None</u>, <u>None</u>	Double bond
1.1	197.01	<u>None</u>, <u>None</u>	Atom with lone pair
1.9	151.30	<u>None</u>, <u>None</u>	Ether

Figure S5: Continued.

SMILES:

COC(=O)c1nc(C(C)(C)NC(=O)c2nnc(C)o2)[nH]c(=O)c1O.CB(O)O

Electrophilicity (MAA)

Atom ID	MAA Value [kJ/mol]	Error Log (Reactant, Product)	Type
1.13	269.73	None, Connectivity	Double bond
1.2	252.23	None, Connectivity	Ester
1.16	233.52	None, None	Double bond
1.22	212.39	None, None	Double bond
1.6	184.97	None, None	Double bond
1.20	169.10	None, None	Double bond
1.4	158.20	None, None	Double bond

Nucleophilicity (MCA)

Atom ID	MCA Value [kJ/mol]	Error Log (Reactant, Product)	Type
1.15	428.03	None, None	Pyridine like nitrogen
1.14	420.16	None, None	Pyridine like nitrogen
1.19	410.56	None, Connectivity	Atom with lone pair
1.5	375.44	None, None	Pyridine like nitrogen
1.10	352.89	None, None	Atom with lone pair
1.12	334.97	None, None	Amide
1.2	328.27	None, Connectivity	Double bond

Figure S6: Calculated MAAs and MCAs for the second reaction step in the retrosynthetic route by Manifold (see Figure S4). The MAAs and MCAs are obtained at the r^{2}SCAN-3c SMD(DMSO)//GFN1-xTB ALPB(DMSO) level of theory.

1.11	147.33	None, Connectivity	Amide
1.14	105.53	None, None	Double bond
1.12	62.28	None, None	Double bond
1.21	41.41	None, None	Double bond
1.3	26.83	None, None	Double bond
1.15	9.55	None, None	Double bond
1.5	-18.58	None, None	Double bond

1.11	325.66	None, Connectivity	Double bond
1.21	323.21	None, None	Atom with lone pair
1.3	311.41	None, None	Ester
1.23	293.93	None, Connectivity	Phenol
1.13	274.10	None, None	Double bond
1.4	269.92	None, None	Double bond
2.2	269.87	None, None	Atom with lone pair
2.3	269.87	None, None	Atom with lone pair
1.20	264.93	None, Connectivity	Double bond
1.22	262.63	None, None	Double bond
1.6	254.13	None, Connectivity	Double bond
1.1	247.26	None, None	Atom with lone pair
1.16	241.09	None, None	Double bond
1.18	181.68	None, None	Ether

Figure S6: Continued.

SMILES:

NCc1ccc(F)cc1.COC(=O)c1nc(C(C)(C)NC(=O)c2nnc(C)o2)n(C)c(=O)c1O
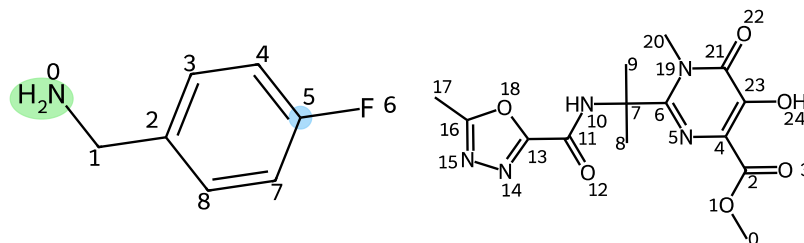
Electrophilicity (MAA)			
Atom ID	MAA Value [kJ/mol]	Error Log (Reactant, Product)	Type
1.5	310.44	None, Connectivity	Double bond
2.2	222.56	None, Connectivity	Ester
2.21	196.20	None, Connectivity	Double bond
2.23	195.82	None, None	Double bond
2.16	184.28	None, None	Double bond
2.6	173.14	None, None	Double bond
2.13	161.22	None, None	Double bond
2.4	138.50	None, None	Double bond

Nucleophilicity (MCA)			
Atom ID	MCA Value [kJ/mol]	Error Log (Reactant, Product)	Type
1.0	488.97	None, None	Amine
2.15	411.26	None, None	Pyridine like nitrogen
2.14	403.50	None, None	Pyridine like nitrogen
1.2	400.10	None, Connectivity	Double bond
2.22	329.57	None, None	Atom with lone pair
2.5	326.60	None, None	Pyridine like nitrogen
2.12	319.65	None, None	Amide
1.5	317.49	None, Connectivity	Double bond
1.7	314.44	None, None	Double bond

Figure S7: Calculated MAAs and MCAs for the third reaction step in the retrosynthetic route by Manifold (see Figure S4). The MAAs and MCAs are obtained at the r^{2}SCAN-3c SMD(DMSO)//GFN1-xTB ALPB(DMSO) level of theory.

2.11	110.08	None, None	Amide
1.8	82.22	None, None	Double bond
1.3	82.22	None, None	Double bond
1.4	71.43	None, None	Double bond
1.7	71.43	None, None	Double bond
1.2	55.35	None, None	Double bond
2.14	42.93	None, None	Double bond
2.22	24.46	None, None	Double bond
2.5	10.16	None, None	Double bond
2.12	4.25	None, None	Double bond
2.15	-3.30	None, None	Double bond
2.3	-4.51	None, None	Double bond

|			
1.4	314.44	None, None	Double bond
2.10	304.59	None, None	Atom with lone pair
1.3	303.20	None, None	Double bond
1.8	303.20	None, None	Double bond
2.3	298.20	None, None	Ester
2.21	289.68	None, Connectivity	Double bond
2.2	285.58	None, Connectivity	Double bond
2.23	272.59	None, Connectivity	Double bond
2.13	270.05	None, None	Double bond
2.24	268.37	None, Connectivity	Phenol
2.4	256.98	None, None	Double bond
2.11	242.67	None, Connectivity	Double bond
2.6	240.50	None, Connectivity	Double bond
2.1	239.26	None, None	Atom with lone pair
2.19	229.57	None, None	Atom with lone pair
2.16	225.10	None, None	Double bond
2.18	150.97	None, None	Ether
1.6	143.47	None, None	Atom with lone pair

Figure S7: Continued.
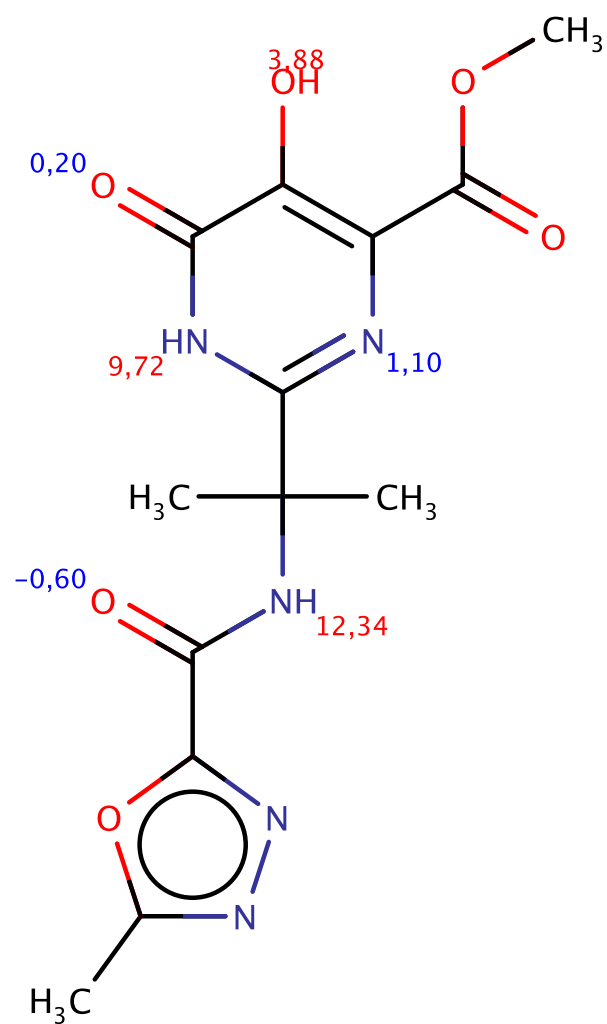
Figure S8: Predicted pKa values for the first reactant in the second reaction step of the retrosynthetic route by Manifold (see Figure S4). The predicted pKa values are obtained using MarvinSketch.
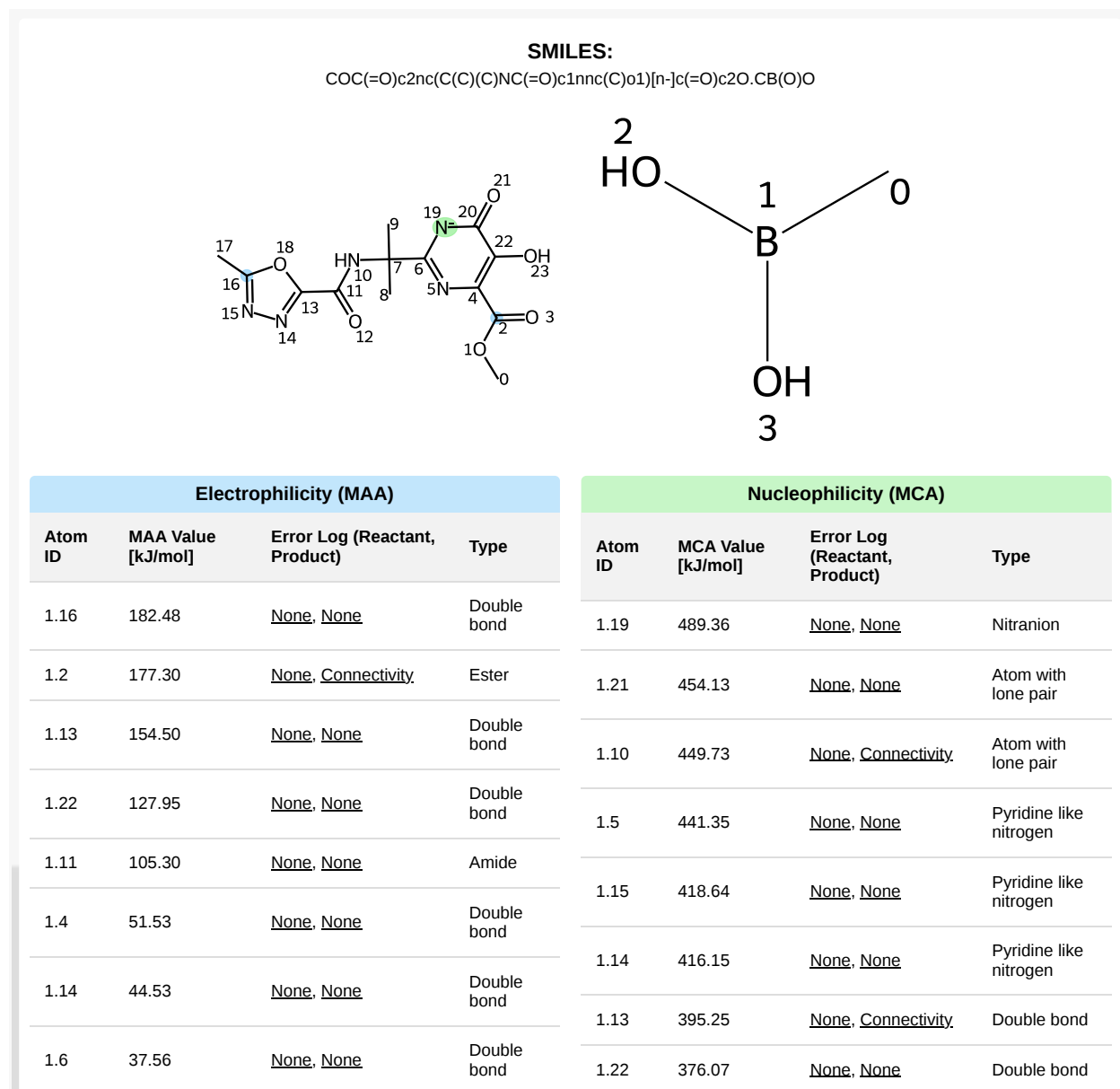
SMILES:

COC(=O)c2nc(C(C)(C)NC(=O)c1nnc(C)o1)[n-]c(=O)c2O.CB(O)O

Electrophilicity (MAA)			
Atom ID	MAA Value [kJ/mol]	Error Log (Reactant, Product)	Type
1.16	182.48	None, None	Double bond
1.2	177.30	None, Connectivity	Ester
1.13	154.50	None, None	Double bond
1.22	127.95	None, None	Double bond
1.11	105.30	None, None	Amide
1.4	51.53	None, None	Double bond
1.14	44.53	None, None	Double bond
1.6	37.56	None, None	Double bond

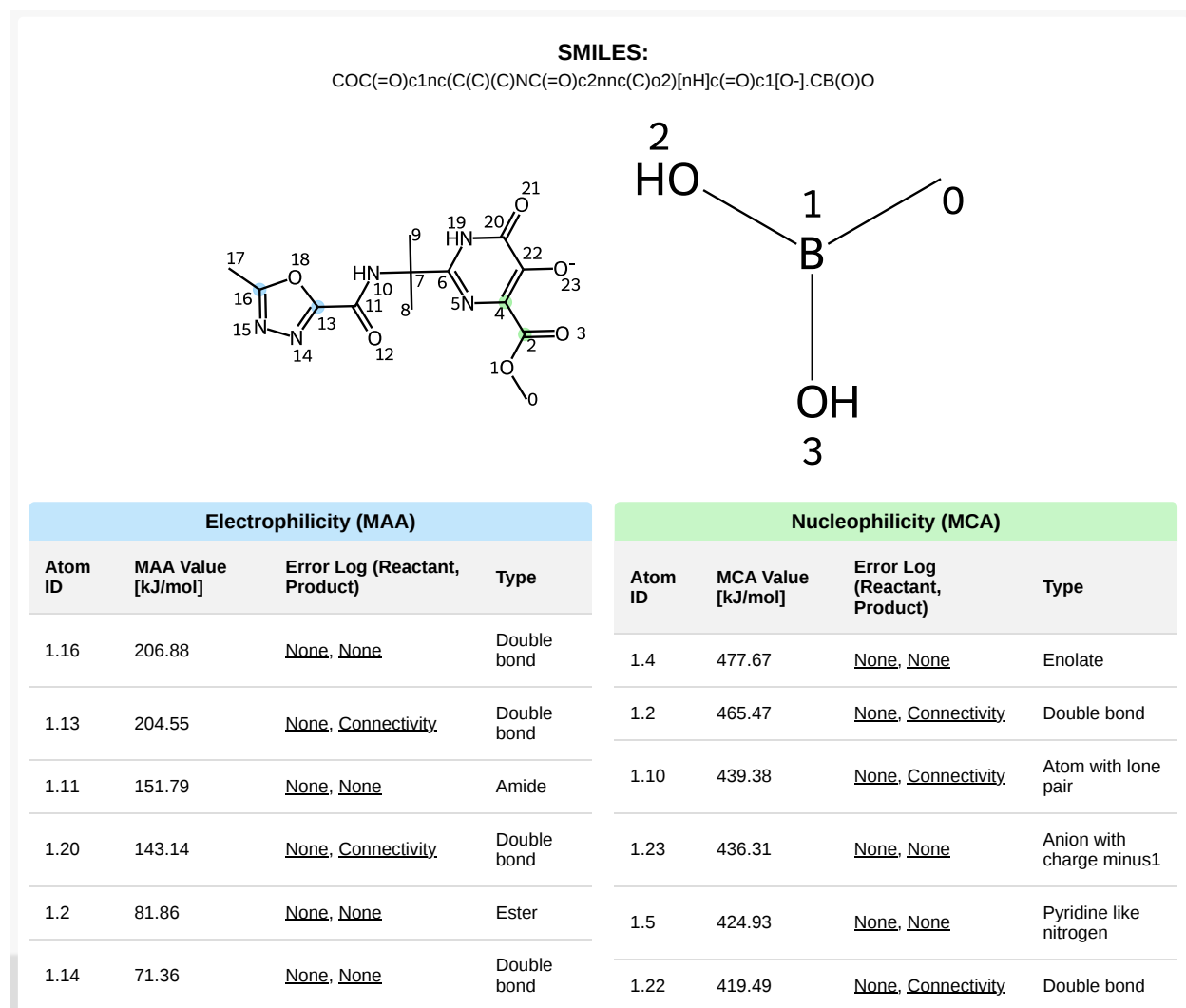| Nucleophilicity (MCA)			
Atom ID	MCA Value [kJ/mol]	Error Log (Reactant, Product)	Type
1.19	489.36	None, None	Nitranion
1.21	454.13	None, None	Atom with lone pair
1.10	449.73	None, Connectivity	Atom with lone pair
1.5	441.35	None, None	Pyridine like nitrogen
1.15	418.64	None, None	Pyridine like nitrogen
1.14	416.15	None, None	Pyridine like nitrogen
1.13	395.25	None, Connectivity	Double bond
1.22	376.07	None, None	Double bond

Figure S9: Calculated MAAs and MCAs for the second reaction step in the retrosynthetic route by Manifold (see Figure S4). The atom site 1.19 is deprotonated compared to Figure S6. The MAAs and MCAs are obtained at the r²SCAN-3c SMD(DMSO)//GFN1-xTB ALPB(DMSO) level of theory.

1.20	34.50	None, None	Double bond
1.12	-0.97	None, None	Double bond
1.21	-52.24	None, None	Double bond
1.15	-61.24	None, None	Double bond
1.5	-61.74	None, None	Double bond
1.3	-66.73	None, None	Double bond

1.20	373.46	None, Connectivity	Double bond
1.6	367.10	None, Connectivity	Double bond
1.2	351.21	None, Connectivity	Double bond
1.12	338.39	None, None	Amide
1.3	331.36	None, None	Ester
1.4	322.81	None, None	Double bond
1.23	317.60	None, Connectivity	Phenol
1.1	274.86	None, None	Atom with lone pair
2.2	269.87	None, None	Atom with lone pair
2.3	269.87	None, None	Atom with lone pair
1.16	262.42	None, None	Double bond
1.11	190.58	None, Connectivity	Double bond
1.18	174.49	None, None	Ether

Figure S9: Continued.

SMILES:
COC(=O)c1nc(C(C)(C)NC(=O)c2nnc(C)o2)[nH]c(=O)c1[O-].CB(O)O

Electrophilicity (MAA)			
Atom ID	MAA Value [kJ/mol]	Error Log (Reactant, Product)	Type
1.16	206.88	None, None	Double bond
1.13	204.55	None, Connectivity	Double bond
1.11	151.79	None, None	Amide
1.20	143.14	None, Connectivity	Double bond
1.2	81.86	None, None	Ester
1.14	71.36	None, None	Double bond

Nucleophilicity (MCA)			
Atom ID	MCA Value [kJ/mol]	Error Log (Reactant, Product)	Type
1.4	477.67	None, None	Enolate
1.2	465.47	None, Connectivity	Double bond
1.10	439.38	None, Connectivity	Atom with lone pair
1.23	436.31	None, None	Anion with charge minus1
1.5	424.93	None, None	Pyridine like nitrogen
1.22	419.49	None, Connectivity	Double bond

Figure S10: Calculated MAAs and MCAs for the second reaction step in the retrosynthetic route by Manifold (see Figure S4). The atom site 1.23 is deprotonated compared to Figure S6. The MAAs and MCAs are obtained at the r^{2}SCAN-3c SMD(DMSO)//GFN1-xTB ALPB(DMSO) level of theory.

1.22	53.00	None, None	Double bond
1.6	50.60	None, None	Double bond
1.4	34.98	None, None	Double bond
1.12	14.93	None, None	Double bond
1.15	-17.13	None, Connectivity	Double bond
1.5	-45.17	None, None	Double bond
1.21	-77.24	None, None	Double bond
1.3	-108.39	None, None	Double bond

1.15	418.47	None, None	Pyridine like nitrogen
1.6	415.25	None, None	Double bond
1.14	409.38	None, None	Pyridine like nitrogen
1.21	388.73	None, None	Atom with lone pair
1.3	381.45	None, None	Ester
1.20	349.43	None, Connectivity	Double bond
1.12	335.83	None, None	Amide
1.1	332.79	None, None	Atom with lone pair
1.13	317.89	None, None	Double bond
1.19	309.30	None, None	Atom with lone pair
1.16	280.77	None, None	Double bond
2.2	269.86	None, None	Atom with lone pair
2.3	269.86	None, None	Atom with lone pair
1.11	218.05	None, Connectivity	Double bond
1.18	174.46	None, None	Ether

Figure S10: Continued.

Additional material for the covalent inhibitor reactivity predictions:
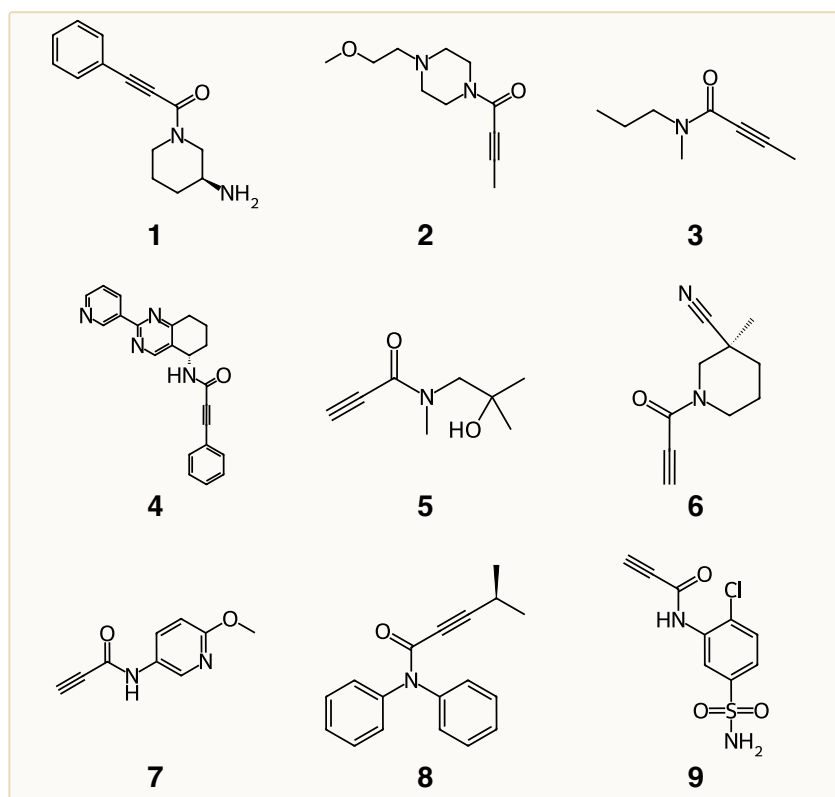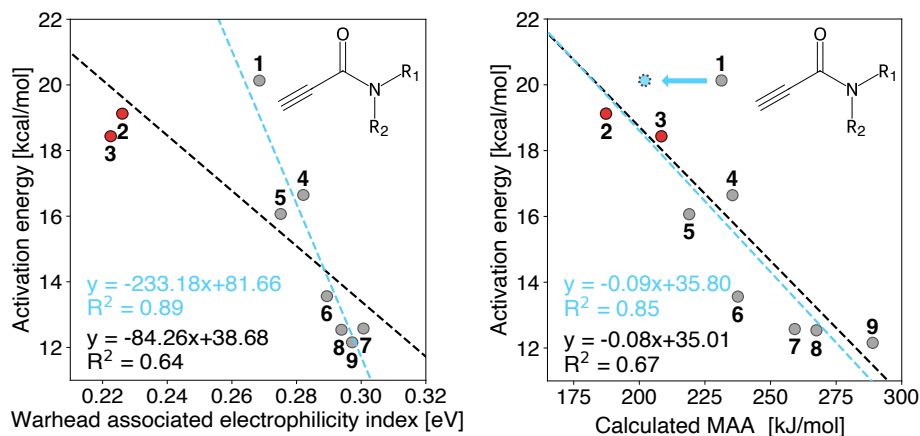

Figure S11: Covalent inhibitor reactivity prediction for 9 different propargylamides. The calculated MAAs are obtained at the r²SCAN-3c SMD(DMSO)//GFN1-xTB ALPB(DMSO) level of theory and otherwise ωB97XD/cc-pVDZ CPCM(H$_2$O). The blue regression lines are defined in the main text. The red dots are considered outliers in the work of Hermann et al.[13]
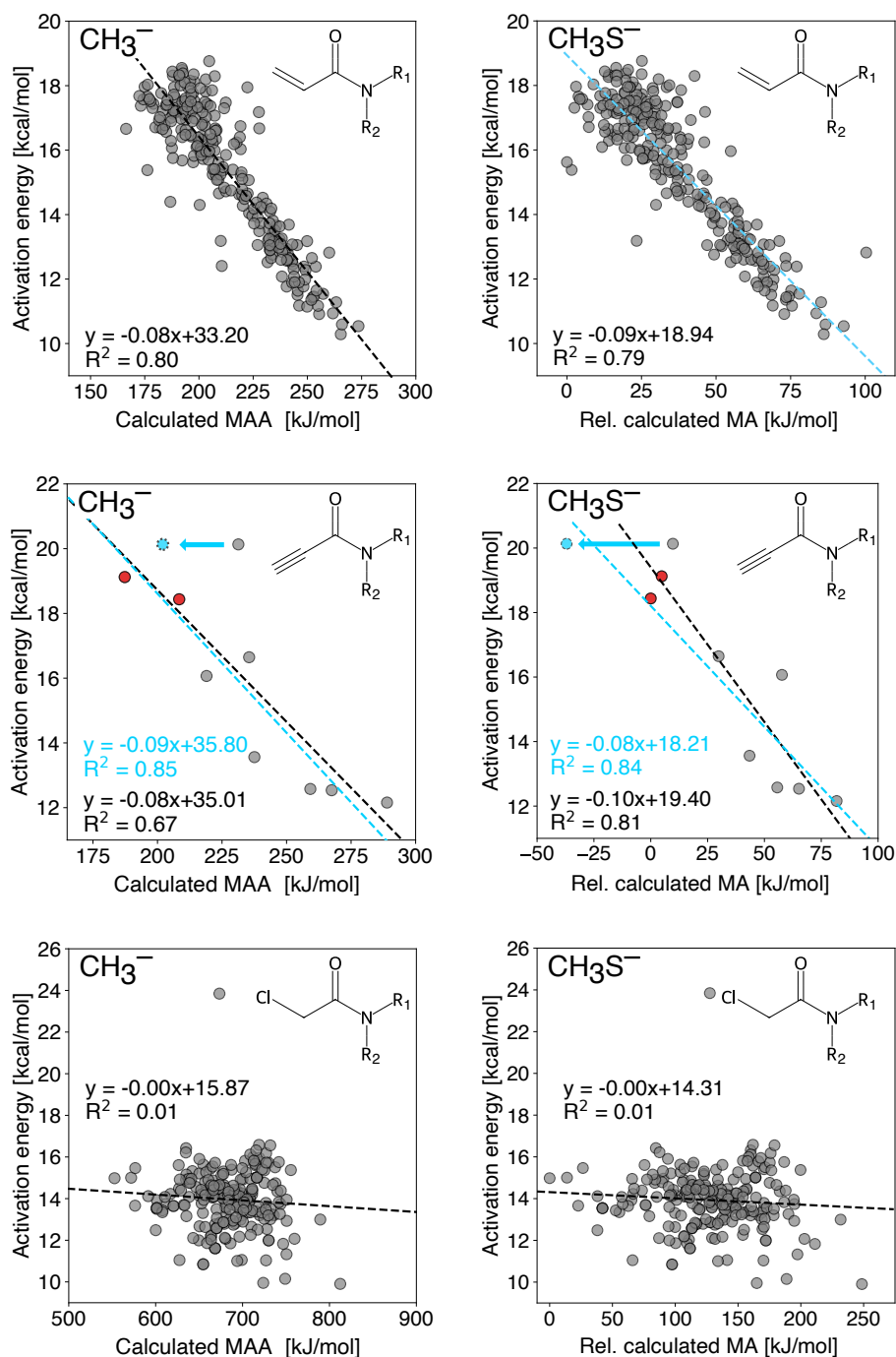
Figure S12: Covalent inhibitor reactivity prediction for various acrylamides (top), propargylamides (middle), and 2-chloroacetamides (bottom). The calculated MAAs (left) and relative methanethiolate affinities (MAs) (right) are obtained at the r²SCAN-3c SMD(DMSO)//GFN1-xTB ALPB(DMSO) level of theory. The black regression lines consider all points, whereas the blue regression lines are defined in the main text. The red dots are considered outliers in the work of Hermann et al.[13]