Supporting Information for

The Coupling of Tautomerization to Hydration in the Transition State on the Pyrimidine Photohydration Reaction Path

Stefan Franzen[1], Bohdan Skalski[2], Libero Bartolotti[3], Bernard Delley[4]

[1] Department of Chemistry, North Carolina State University, Raleigh, NC 27695

[2] Faculty of Chemistry, Adam Mickiewicz University, Umultowska 89b, 61-614 Poznań, Poland

[3] Department of Chemistry, East Carolina University, Greenville, NC 27858, United States

[4] Paul-Scherer Institute, CH-5232, Villigen, Switzerland

1.0 Hydration of pyrimidines calculated using the NEB method

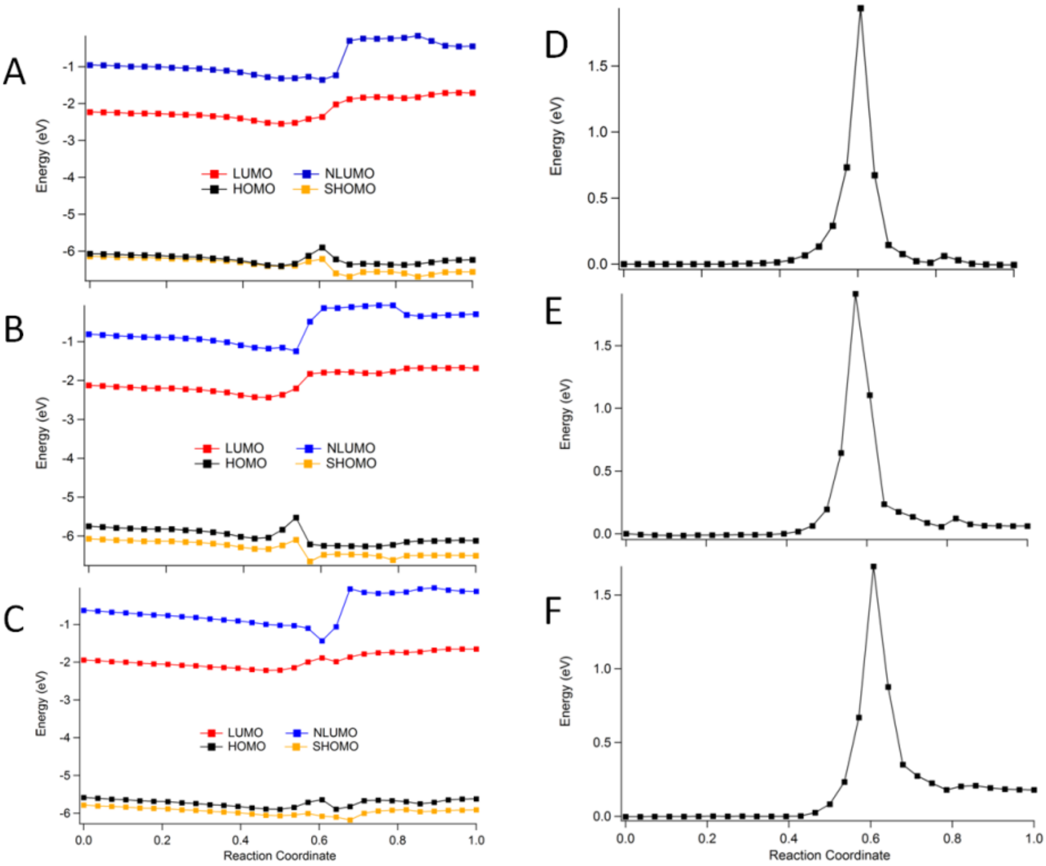

Figure S1. Nudged elastic band energies along the photohydration reaction coordinate for 1-methylpyrimidines. Molecular orbitals corresponding to the HOMO-1 (orange), HOMO (black), LUMO (red) and LUMO+1 (blue) are shown in the left panels A.) 1-MU B.) 1-MT and C.) 1-MC. The ground state energy along the reaction coordinate is shown in the right panels D.) 1-MU E.) 1-MT and F.) 1-MC.

1.1.1 1-methylcytosine NEB and DIMER calculations

Tables S1A, S1B and S1C provide the coordinates for the reactant, transition state and product for the hydration reaction of 1-methyl cytosine, respectively.

Table S1. 1-methylcytosine coordinates

Table S1A. Reactant (including H_2O)

1	H	1.982	-3.621	7.534
2	H	-2.666	-3.032	6.492
3	H	1.815	-1.854	7.823
4	C	1.869	-2.639	7.054
5	H	2.717	-2.455	6.388
6	N	0.662	-2.613	6.227
7	C	-0.549	-2.854	6.782
8	O	-0.503	0.043	8.122
9	C	0.787	-2.276	4.818
10	O	1.903	-2.047	4.361

11	N	-0.374	-2.232	4.091
12	C	-1.545	-2.486	4.655
13	N	-2.64	-2.465	3.836
14	C	-1.701	-2.803	6.043
15	H	0.068	0.5	7.482
16	H	-0.552	-3.086	7.847
17	H	-1.236	-0.29	7.578
18	H	-2.493	-2.093	2.902
19	H	-3.572	-2.41	4.23

Table S1B. Transition state coordinates

1	H	1.871	-3.476	7.742
2	H	-2.581	-3.003	6.527
3	H	2.079	-1.699	7.623
4	C	1.925	-2.625	7.05
5	H	2.752	-2.757	6.348
6	N	0.683	-2.54	6.28
7	C	-0.491	-2.547	6.931
8	O	-0.923	-0.776	7.756
9	C	0.775	-2.178	4.856
10	O	1.891	-1.981	4.396
11	N	-0.384	-2.121	4.137
12	C	-1.55	-2.305	4.733
13	N	-2.645	-2.36	3.928
14	C	-1.737	-2.404	6.17
15	H	-0.325	-0.1	7.378
16	H	-0.5	-3.002	7.918
17	H	-1.767	-1.256	6.761
18	H	-2.503	-2.122	2.95
19	H	-3.575	-2.264	4.319

Table S1C. Photohydrate product coordinates

1	H	1.659	-3.214	7.944
2	H	-1.892	-3.766	6.176
3	H	2.112	-1.539	7.486
4	C	1.811	-2.528	7.098
5	H	2.626	-2.894	6.465
6	N	0.597	-2.473	6.292
7	C	-0.645	-2.199	6.972
8	O	-0.88	-0.8	7.236
9	C	0.732	-2.275	4.903
10	O	1.844	-2.271	4.377
11	N	-0.431	-2.103	4.141

12	C	-1.593	-2.273	4.687
13	N	-2.706	-2.174	3.906
14	C	-1.815	-2.669	6.124
15	H	-0.095	-0.45	7.692
16	H	-0.624	-2.749	7.928
17	H	-2.745	-2.242	6.524
18	H	-2.565	-1.855	2.95
19	H	-3.62	-2.051	4.323

1.1.2 Cytosine vibrational wave numbers (cm^{-1})

Tables S2 gives the eigenvalues for the vibrational frequency calculation of the three structures given in Tables S2A, S2B and S2C, respectively. While the reactant and product have six zero eigenvalues as expected for a structure that is at the optimum geometry, the saddle point has two negative eigenvalues. A single negative eigenvalue would be expected along the reaction coordinate. Section 1.1.3 shows that the first eigenvalue corresponds to the bond breaking and reforming coordinate, while the second eigenvalue is a methyl rotation.

Table S2. 1-methylcytosine vibrational wave numbers (cm^{-1})

Reactant	Saddle	Product
0	-1401.7	0
0	-106.2	0
0	0	0
0	0	0
0	0	0
0	0	0
10.2	0	65.7
30	0	110.3
77.4	80.7	140
102.7	139.5	210.1
191	159.5	239.4
218.7	192.8	325.9
227.7	289.6	361
237.2	314.3	364
267.1	322.5	366.3
274.9	359.7	391
305.4	412.2	429.9
329.7	449.3	468.2
373.8	458.9	507.1
425.4	515	569.6
468.3	522.2	593.8
539.3	550.2	632.5
553.5	615.5	725.1

596.4	633.3	760.6
712.5	697.7	772
720.2	735.3	848.6
753.8	758.7	895.8
761.3	777.4	955.2
776.8	865.3	993.1
932.8	911.1	1043.9
962.8	1020.3	1073.3
1027	1034.1	1139.8
1051.9	1054.6	1142
1136.6	1116	1162.7
1151.8	1128.4	1204.5
1195.7	1163.8	1242.9
1234.3	1245.3	1268.1
1327	1300	1286.6
1362.7	1355.9	1359.9
1399.7	1383.4	1396.3
1473.5	1390.7	1403.7
1487.9	1433.3	1406.2
1489.7	1479.4	1432.6
1523.4	1480.8	1473.8
1592.2	1496.4	1494.4
1642.5	1572.9	1571.5
1663.9	1606.1	1647.8
1702.9	1712	1688.1
3015.2	1753.8	2998.4
3114	2996.1	3003.9
3163	3093.4	3022.5
3182.3	3157.3	3075.6
3208.3	3181.9	3081.9
3526.9	3199.6	3139.3
3661.1	3513	3517.8
3695.6	3651	3654
3812.2	3670.1	3725.3

1.1.3 1-methylcytosine hydrate eigenvector projections

In this section we address existence of two negative eigenvalues in the Hessian calculated at the transition state. There are two modes with negative eigenvalues. Mode 1, the first and largest negative eigenvalue is a HO-H bond breaking and C-H bond forming coordinate, which corresponds to the transition state of the model system, ethene + H_2O, system where the is a single negative eigenvalue at the transition state. The second negative eigenvalue in the cytosine system (mode 2) is shown to be a methyl rotation.

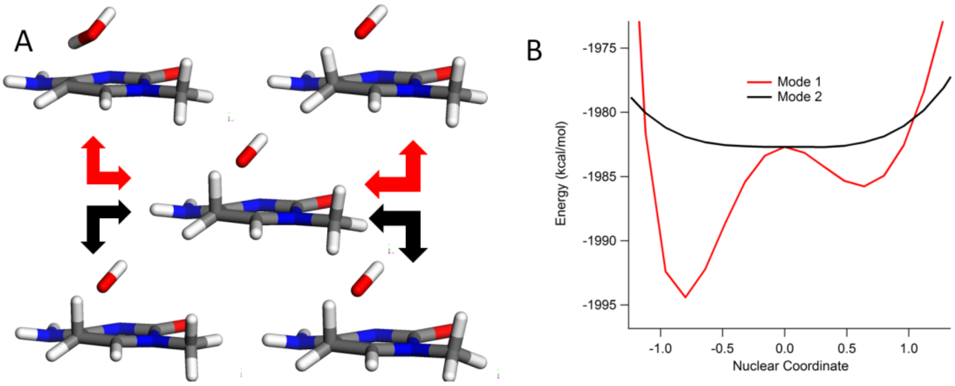

Figure S2. Eigenvector projections along the two negative eigenvalues calculated from the Hessian of the saddle point structure of 1-MC. A.) structures along each mode are shown using arrows that are colored to correspond to the potential surfaces. B.) potential energy surfaces for projections along mode 1 (red) and mode 2 (black). It is evident from inspection that mode 1 involves breaking of the O-H bond of H_2O and formation of the C-H bond at the C5 position of cytosine.

1.2.1 1-Methyluracil NEB and DIMER calculation

Tables S3A, S3B and S3C provide the coordinates for the reactant, transition state and product of the hydration of 1-methyluracil, respectively.

Table S3. 1-methyluracil coordinates

Table S3A. Reactant (including H_2O)

1	H	2.021	-1.107	1.48
2	H	-2.688	-0.528	0.511
3	H	1.728	0.628	1.873
4	C	1.856	-0.104	1.065
5	H	2.71	0.174	0.44
6	N	0.663	-0.1	0.213
7	C	-0.575	-0.332	0.761
8	O	0.259	2.711	2.206
9	C	0.816	0.172	-1.159
10	O	1.9	0.4	-1.685
11	N	-0.381	0.162	-1.863
12	H	-0.294	0.344	-2.862
13	C	-1.692	-0.082	-1.385
14	O	-2.648	-0.065	-2.157
15	C	-1.726	-0.337	0.041
16	H	0.513	3.283	1.462
17	H	-0.567	-0.514	1.836
18	H	-0.675	2.929	2.359

Table S3B. Transition state coordinates

1	H	1.939	-0.974	0.694
2	H	-2.595	-0.597	-0.432

3	H	1.977	0.819	0.727
4	C	1.926	-0.066	0.077
5	H	2.776	-0.07	-0.611
6	N	0.696	-0.03	-0.721
7	C	-0.508	-0.074	-0.085
8	O	-0.944	1.666	0.689
9	C	0.813	0.261	-2.11
10	O	1.899	0.398	-2.655
11	N	-0.393	0.356	-2.775
12	H	-0.316	0.464	-3.787
13	C	-1.695	0.063	-2.282
14	O	-2.626	-0.048	-3.07
15	C	-1.76	-0.015	-0.828
16	H	-0.379	2.357	0.289
17	H	-0.489	-0.533	0.901
18	H	-1.787	1.181	-0.239

Table S3C. Photohydrate product coordinates

1	H	1.683	-0.751	1.9
2	C	-1.84	-1.38	0.162
3	H	1.996	0.98	1.541
4	C	1.788	-0.008	1.097
5	H	2.638	-0.276	0.462
6	N	0.581	-0.005	0.278
7	C	-0.691	0.264	0.938
8	O	-0.947	1.662	1.127
9	C	0.73	0.214	-1.083
10	O	1.81	0.302	-1.657
11	N	-0.474	0.303	-1.805
12	H	-0.356	0.483	-2.801
13	C	-1.769	0.101	-1.344
14	O	-2.737	0.175	-2.087
15	C	-1.847	-0.281	0.117
16	H	-0.231	2.034	1.67
17	H	-0.646	-0.248	1.913
18	H	-2.802	0.073	0.52

1.2.2 1-methyluracil hydrate vibrational wave numbers (cm^{-1})

Table S4 gives the eigenvalues for the vibrational frequency calculation of the three sets of coordinates given in Tables S4A, S4B and S4C, respectively.

Table S4. 1-methyluracil hydrate vibrational wave numbers (cm^{-1})

Reactant	Saddle	Final
0	-1564.7	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
42.4	0	88.9
43.3	88.6	110.2
84.1	117.8	133.5
123	133.7	190.4
169.8	158.2	290.8
248.3	174.9	318.5
273.9	277.9	342.9
332	336.9	374.3
345.7	371.9	375.6
357.5	418	423.2
379.8	452.4	484
384.3	467.1	536
428.6	511.2	545.9
455	537.8	600.1
531.7	600.2	642.9
604.7	629.2	717.9
678.4	676.6	750.7
733.5	712.7	774.2
744.2	739.2	846.4
757.7	767.2	919.8
783.1	789.6	942.9
839.9	849.4	1004.3
954.3	931.6	1052.2
966.8	1009.9	1142.6
1006.5	1065.7	1151.9
1125.9	1088.1	1177
1153.8	1136.7	1193.4
1169.4	1175.2	1223.1
1210.8	1184.7	1254.7
1311.8	1302.2	1288.4
1343.4	1319.5	1323.4
1375	1369.2	1357.9
1413.1	1377.8	1382.5
1429.6	1394.7	1394
1478.7	1413.8	1415.3
1525.9	1446.8	1439.7

1638.6	1480.7	1480.5
1665.1	1493.2	1509
1738.9	1663.7	1712.4
1764	1735	1728.4
3039.2	1757.5	3016.2
3122	3033	3021.8
3162.3	3115.2	3027.6
3176	3140.5	3098.3
3225.6	3157.8	3103
3535.3	3172	3153
3705.1	3533.8	3522.1
3819.6	3652.1	3717.4

1.2.3 1-methyluracil hydrate eigenvector projections

There is a single negative eigenvalue in the Hessian of 1-methyluracil hydrate. The eigenvector projections shown in Figure 3 of the text verify that the coordinate corresponds to the bond breaking of an O-H bond on H_2O and simultaneous bond formation on the C5 carbon of uracil. These are analogous to the figures shown here for cytosine and thymine (Figures S2 and S3, respectively).

1.3.1 1-methylthymine NEB and DIMER calculation

Table S5. 1-methylthymine coordinates

Table S5A. Reactant coordinates (including H_2O)

1	H	2.105	-3.587	7.453
2	C	-2.997	-3.086	6.638
3	H	1.801	-1.855	7.847
4	C	1.938	-2.583	7.037
5	H	2.796	-2.302	6.419
6	N	0.754	-2.583	6.175
7	C	-0.491	-2.812	6.718
8	O	0.260	0.211	8.208
9	C	0.917	-2.317	4.809
10	O	2.002	-2.090	4.282
11	N	-0.279	-2.327	4.103
12	H	-0.191	-2.146	3.103
13	C	-1.587	-2.566	4.575
14	O	-2.540	-2.543	3.795
15	C	-1.654	-2.823	6.011
16	H	0.517	0.780	7.463
17	H	-0.481	-2.986	7.794
18	H	-0.673	0.431	8.360
19	H	-3.784	-2.858	5.909
20	H	-3.159	-2.467	7.534

21	H	-3.110	-4.142	6.930

Table S5B. Transition state coordinates

1	H	2.097	-3.361	7.669
2	C	-2.721	-3.548	6.581
3	H	2.095	-1.569	7.691
4	C	2.075	-2.459	7.046
5	H	2.937	-2.446	6.374
6	N	0.858	-2.463	6.228
7	C	-0.352	-2.624	6.828
8	O	-0.959	-0.93	7.655
9	C	0.977	-2.073	4.865
10	O	2.058	-1.809	4.357
11	N	-0.22	-2.035	4.181
12	H	-0.141	-1.843	3.182
13	C	-1.498	-2.477	4.624
14	O	-2.398	-2.623	3.804
15	C	-1.603	-2.65	6.074
16	H	-0.437	-0.177	7.312
17	H	-0.318	-3.103	7.805
18	H	-1.71	-1.464	6.659
19	H	-3.619	-3.376	5.976
20	H	-2.964	-3.33	7.63
21	H	-2.46	-4.615	6.496

Table S5C. Photohydrate product coordinates

1	H	1.639	-3.128	7.965
2	C	-1.966	-4.232	6.192
3	H	2.083	-1.456	7.486
4	C	1.808	-2.457	7.113
5	H	2.643	-2.828	6.511
6	N	0.609	-2.43	6.279
7	C	-0.673	-2.159	6.917
8	O	-0.923	-0.76	7.107
9	C	0.785	-2.203	4.925
10	O	1.876	-2.11	4.372
11	N	-0.406	-2.114	4.183
12	H	-0.276	-1.905	3.194
13	C	-1.709	-2.304	4.623
14	O	-2.662	-2.214	3.861
15	C	-1.832	-2.702	6.084
16	H	-0.217	-0.397	7.668
17	H	-0.645	-2.674	7.891

18	H	-2.751	-2.227	6.451
19	H	-2.795	-4.579	5.562
20	H	-2.174	-4.527	7.231
21	H	-1.043	-4.734	5.866

1.3.2 1-methylthymine hydrate vibrational wave numbers (cm^{-1})

Table S6 gives the eigenvalues for the vibrational frequency calculation of the three sets of coordinates given in Tables S6A, S6B and S6C, respectively.

Table S6. 1-methylthymine hydrate vibrational wave numbers (cm^{-1})

Initial	Saddle	Final
0	-1533.8	0
0	-65	0
0	0	0
0	0	0
0	0	0
0	0	0
65.3	0	65.3
91.6	0	91.6
109.3	84.5	109.3
172.8	104.1	172.8
206.6	132.5	206.6
251.7	164.1	251.7
273.3	172.1	273.3
311.4	229.8	311.4
346.8	263.6	346.8
363.2	284.7	363.2
372.7	348	372.7
420	381.8	420
438.5	408.5	438.5
507.8	433.3	507.8
548.9	462.8	548.9
600.4	490.1	600.4
649.1	552.2	649.1
703.8	615	703.8
726.9	636.4	726.9
740.4	675.1	740.4
777.1	694.5	777.1
863.6	734	863.6
928	764.3	928
974.6	779.3	974.6
1002.5	847	1002.5

1033.1	954.6	1033.1
1088.6	1012.3	1088.6
1103.3	1042.7	1103.3
1117.9	1055.5	1117.9
1152	1107.2	1152
1174.1	1110.6	1174.1
1235.4	1172.1	1235.4
1252.2	1202.8	1252.2
1264.8	1257.9	1264.8
1294	1314	1294
1339.4	1343.2	1339.4
1365.1	1351.4	1365.1
1372.7	1366.1	1372.7
1402.4	1399.2	1402.4
1409.6	1408.3	1409.6
1450.2	1438.9	1450.2
1466.7	1451.2	1466.7
1468.2	1465.3	1468.2
1510.6	1478.4	1510.6
1514.9	1502.7	1514.9
1709.1	1656.3	1709.1
1728.5	1735.6	1728.5
2996.5	1745.1	2996.5
3009.5	3004.7	3009.5
3025.9	3065.2	3025.9
3071.1	3078.2	3071.1
3102.7	3097.7	3102.7
3108	3148.6	3108
3139.9	3181.3	3139.9
3167.9	3192.5	3167.9
3519.9	3506	3519.9
3713.2	3675.6	3713.2

1.3.3 1-methylthymine hydrate eigenvector projections

Figure S3 shows the similarity of the two negative eigenvalues in the Hessian calculated at the transition state of the 1-MT to those in 1-MC. There are two modes with negative eigenvalues. Mode 1, the first and largest negative eigenvalue is a HO-H bond breaking and C-H bond forming coordinate, which corresponds to the transition state of the model system, ethene + H_2O, system where the is a single negative eigenvalue at the transition state. The second negative eigenvalue in the cytosine system (mode 2) is shown to be a methyl rotation. As shown in Figure 3 of the text 1-MU resembles the ethene case in having a single negative eigenvalue. As discussed in the text the methyl rotation can be projected out by finding a local maximum or minimum. This coordinate is not along the transition state and thus it does not interfere with the conclusions drawn here. However, if one wishes the calculate a rate constant using statistical mechanical methods, one must absolutely project out the methyl rotation to prevent it from given a spurious contribution to the barrier height.

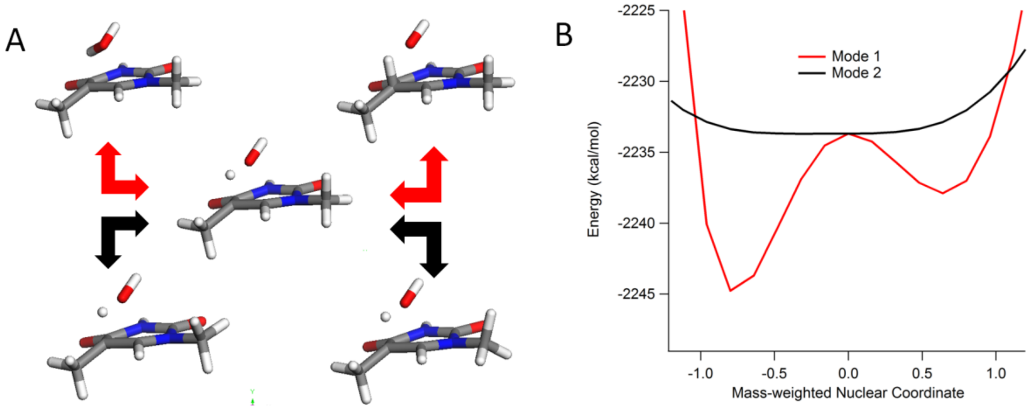

Figure S3.Eigenvector projections along the two negative eigenvalues calculated from the Hessian of the saddle point structure. A.) structures along each mode are shown using arrows that are colored to correspond to the potential surfaces. B.) potential energy surfaces for projections along mode 1 (red) and mode 2 (black). It is evident from inspection that mode 1 involves breaking of the O-H bond of H_2O and formation of the C-H bond at the C5 position of thymine.

The transition state was also validated using a calculation of the intrinsic reaction coordinate (IRC). The IRC trajectories are shown in Figure S4.

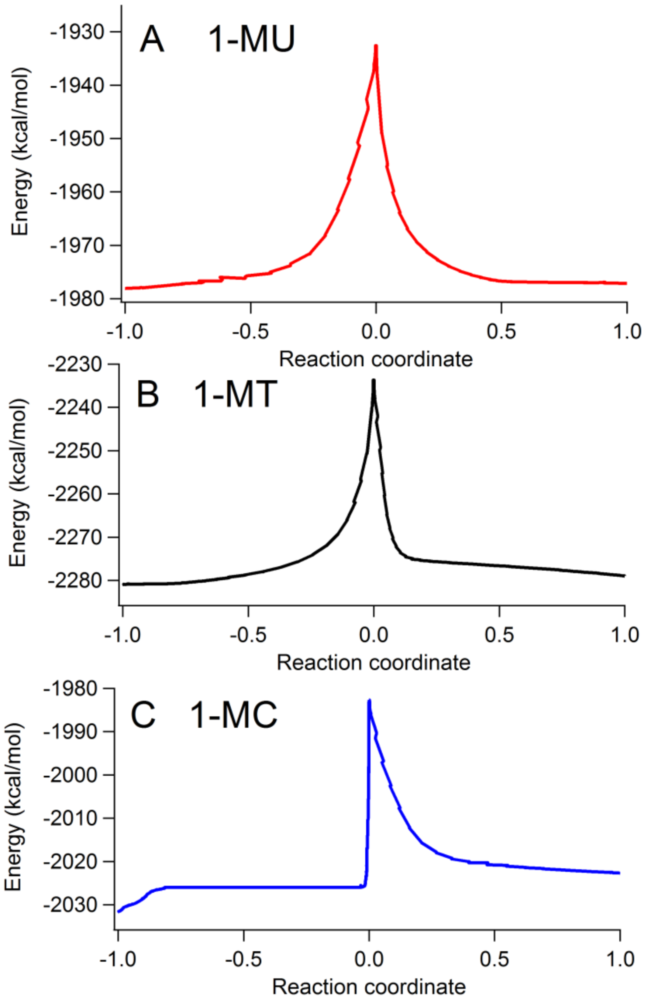

Figure S4. The energy profile of the IRC calculation for 1-methylcytosine, thymine and uracil.

The projected structures using the IRC method are compared to the starting structures for the geometry optimized reactant and product in Figures S5, S6 and S7 for cytosine, thymidine and uracil, respectively..

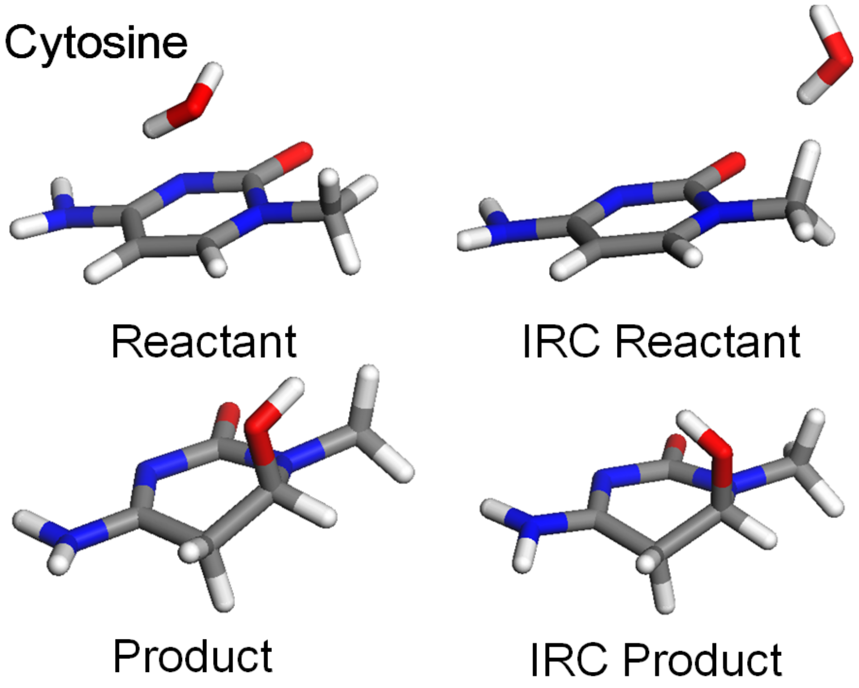

Figure S5. Comparison of the structures for 1-methylcytosine of the geometry optimized reactant and product (left) with the structures obtained using a IRC projection from the transition state (right).

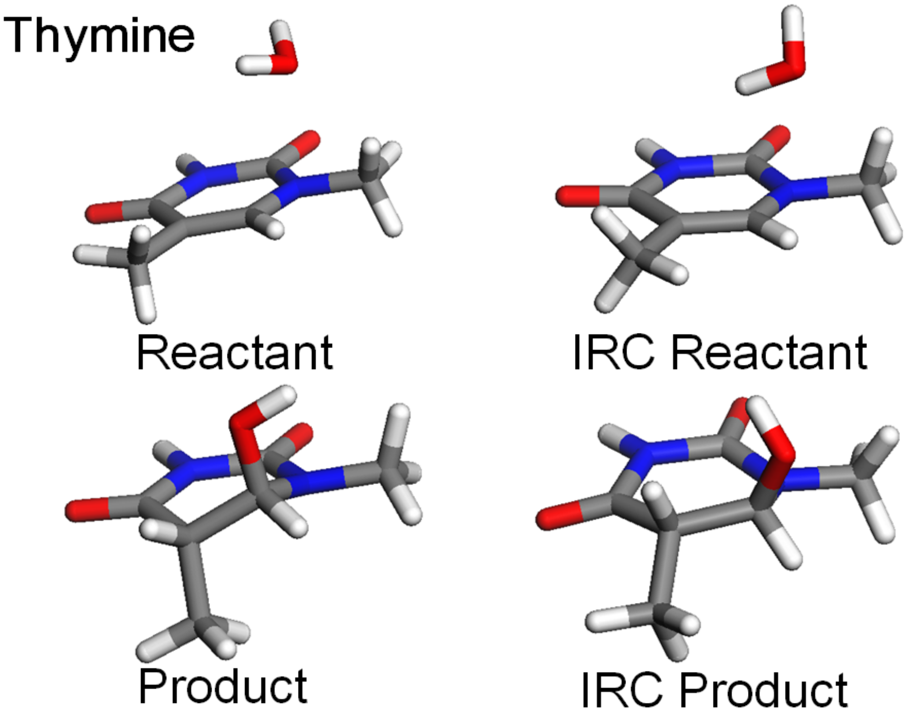

Figure S6. Comparison of the structures for 1-methylthymine of the geometry optimized reactant and product (left) with the structures obtained using a IRC projection from the transition state (right).

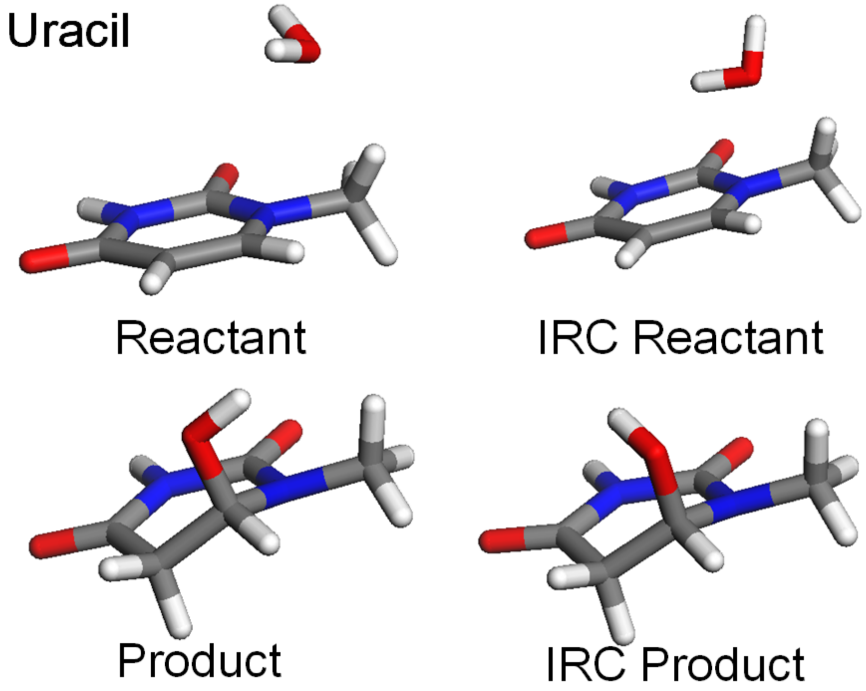

Figure S7. Comparison of the structures for 1-methyluracil of the geometry optimized reactant and product (left) with the structures obtained using a IRC projection from the transition state (right).

2.0 Hydration of pyrimidine tautomers calculated using the NEB method

In this section we report the coordinates and frequencies for the tautomers of the three 1-methylpyridimdines.

Table S7. 1-methyluracil(t) coordinates

Table S7A Reactant coordinates

1	H	1.9	-1.548	1.42
2	H	-2.597	-0.743	0.254
3	H	2.047	0.236	1.455
4	C	1.953	-0.642	0.802
5	H	2.809	-0.697	0.124
6	N	0.746	-0.538	-0.022
7	C	-0.474	-0.72	0.535
8	O	0.266	2.077	1.97
9	C	0.885	-0.156	-1.413
10	O	2.007	0.097	-1.85
11	N	-0.271	-0.101	-2.155
12	C	-1.433	-0.303	-1.588
13	O	-2.508	-0.198	-2.423
14	C	-1.621	-0.602	-0.208
15	H	0.376	2.357	1.045
16	H	-0.489	-0.947	1.6
17	H	-0.275	2.781	2.364
18	H	-3.32	-0.353	-1.91

Table S7B Transition state coordinates

1	H	2.024	-1.185	1.677
2	H	-2.339	-0.826	0.332
3	H	2.48	0.5	1.278
4	C	2.171	-0.468	0.859
5	H	2.95	-0.817	0.177
6	N	0.919	-0.335	0.113
7	C	-0.255	-0.35	0.769
8	O	-0.689	1.352	1.619
9	C	1.013	0.055	-1.298
10	O	2.127	0.261	-1.753
11	N	-0.151	0.126	-2.022
12	C	-1.296	-0.061	-1.429
13	O	-2.376	-0.045	-2.245
14	C	-1.501	-0.206	-0.004
15	H	-0.09	2.039	1.265
16	H	-0.267	-0.856	1.733
17	H	-1.539	0.936	0.631
18	H	-3.181	-0.12	-1.703

.

Table S7C. Photohydrate product coordinates

1	H	1.944	-0.956	1.819
2	H	-1.588	-1.516	-0.058
3	H	2.432	0.7	1.333
4	C	2.117	-0.29	0.962
5	H	2.926	-0.681	0.339
6	N	0.903	-0.217	0.155
7	C	-0.346	0.007	0.842
8	O	-0.595	1.387	1.178
9	C	1.045	0.071	-1.212
10	O	2.156	0.117	-1.736
11	N	-0.121	0.281	-1.975
12	C	-1.266	0.064	-1.449
13	O	-2.354	0.246	-2.244
14	C	-1.506	-0.418	-0.046
15	H	0.178	1.718	1.667
16	H	-0.33	-0.598	1.766
17	H	-2.436	-0.009	0.376
18	H	-3.159	0.093	-1.717

Table S8. 1-methyluracil(t) vibrational wave numbers (cm^{-1})

Reactant	Saddle	Product
0	-1440.6	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
24.5	0	67.9
92.7	88.7	114.4
114	142.6	130.9
150.9	164.2	148.6
188.5	202.5	212.5
252.8	213.4	338.6
265.9	298.7	358.9
307.4	343.2	376.8
347.6	377.5	381.6
381.5	440.4	426.7
433.8	462.1	441.4
439.6	465.6	486
453	496.1	541.8
478.4	537.4	565.6
568.8	546	602.5
608.9	634.1	725.4
643.3	650.1	760.2
716	705.5	773.3
750.9	756.1	871.1
767.3	759.6	935.1
776.8	789.7	972.2
783.9	861.3	1015.7
966	925.2	1046.9
973.7	1010.7	1111
1064.3	1033.6	1126.3
1097.7	1112.4	1177.1
1150.5	1129	1183.6
1174	1139.9	1220.6
1210.3	1162.2	1250.4
1254.5	1225.7	1281.5
1328.7	1283.1	1293.1
1368.5	1336.8	1359.5
1419.7	1364.3	1369.3
1449.4	1403.3	1417.1
1505.9	1434.8	1428.5

1512	1445.4	1443.6
1583	1469.5	1463
1675.7	1491.3	1476.4
1684.2	1601.3	1685.7
1713.6	1699.6	1719.3
3048.9	1723.3	3009.4
3139.3	3037.2	3013
3164.9	3115	3033.6
3186.7	3140.1	3070.5
3202.4	3150.6	3088.8
3511.4	3189.2	3131.7
3709.7	3638.9	3703.7
3794.7	3675.4	3724.9

Table S9. 1-methylthymidine(t) coordinates

Table S9A Reactant coordinates

1	H	2.672	-1.344	1.095
2	C	-2.49	-0.744	0.31
3	H	2.288	0.335	1.598
4	C	2.444	-0.329	0.739
5	H	3.27	0.03	0.118
6	N	1.245	-0.331	-0.104
7	C	0.023	-0.531	0.449
8	O	0.76	2.373	1.982
9	C	1.39	-0.074	-1.519
10	O	2.504	0.186	-1.974
11	N	0.24	-0.131	-2.27
12	C	-0.915	-0.362	-1.703
13	O	-1.977	-0.399	-2.561
14	C	-1.128	-0.559	-0.3
15	H	1.126	2.712	1.148
16	H	0.008	-0.657	1.532
17	H	-0.11	2.801	2.045
18	H	-3.161	0.106	0.098
19	H	-2.42	-0.823	1.403
20	H	-2.987	-1.665	-0.04
21	H	-2.783	-0.593	-2.048

Table S9B. Transition state coordinates

1	H	2.526	-0.822	1.611
2	C	-2.167	-1.094	0.3
3	H	2.771	0.929	1.326
4	C	2.609	-0.045	0.843
5	H	3.449	-0.249	0.175
6	N	1.377	-0.022	0.049
7	C	0.189	-0.178	0.658
8	O	-0.446	1.469	1.551
9	C	1.472	0.447	-1.337
10	O	2.572	0.783	-1.747
11	N	0.325	0.445	-2.086
12	C	-0.811	0.103	-1.545
13	O	-1.855	0.049	-2.407
14	C	-1.056	-0.142	-0.134
15	H	0.094	2.228	1.252
16	H	0.2	-0.694	1.616
17	H	-1.196	0.985	0.492
18	H	-3.16	-0.757	-0.036
19	H	-2.224	-1.127	1.396
20	H	-2.011	-2.121	-0.067
21	H	-2.659	-0.187	-1.91

S9C. Photohydrate product coordinates

1	H	2.184	-0.535	1.858
2	C	-1.412	-1.624	0.022
3	H	2.657	1.117	1.35
4	C	2.355	0.118	0.991
5	H	3.172	-0.275	0.38
6	N	1.152	0.17	0.167
7	C	-0.11	0.415	0.821
8	O	-0.385	1.806	1.075
9	C	1.306	0.308	-1.221
10	O	2.414	0.238	-1.748
11	N	0.15	0.522	-1.997
12	C	-1.002	0.366	-1.462
13	O	-2.084	0.588	-2.256
14	C	-1.262	-0.095	-0.047
15	H	0.379	2.178	1.549
16	H	-0.096	-0.139	1.777
17	H	-2.177	0.388	0.333
18	H	-2.237	-1.972	-0.613
19	H	-1.623	-1.935	1.056

20	H	-0.489	-2.119	-0.311
21	H	-2.893	0.48	-1.724

Table S10. 1-methylthymine(t) vibrational wave numbers (cm^{-1})

Reactant	Saddle	Product
0	-1403	0
0	-162.2	0
0	0	0
0	0	0
0	0	0
0	0	0
31.3	0	60.9
53.5	0	98.1
85.3	11.1	132.6
91	63	181.8
138.2	107.2	191.2
174.1	189.1	217.8
226.3	208.6	230.6
243.8	221.4	263.4
260.2	286	278.4
267.8	301.3	348.3
315.3	356.3	375.8
334.9	383.9	412.7
342.7	430.9	444.9
387	447.7	494.7
409.5	458.8	514.6
458.7	502.4	562
459.3	518	591.7
476.1	549.3	673.4
529.8	622.2	723.9
633.2	659.9	760
673.3	669	769.1
739.5	746	877.9
750	748.5	923
772.7	773.1	974.8
858.1	854.7	999.8
972.2	909.4	1039.2
991.5	986.2	1096.4
1062.7	1004.9	1104
1067.8	1037.9	1117.2
1126.8	1100.7	1128.1
1136.7	1126.3	1162.9

1169	1146	1202.6
1224.6	1214.8	1231.4
1267	1221.1	1272.6
1334.4	1269.9	1301.4
1371	1326.6	1316.5
1384.9	1346.7	1352.1
1432.8	1380.3	1358.9
1454.6	1387.4	1386.3
1465.1	1403.2	1404.9
1481.6	1452.8	1436.9
1510.8	1460.2	1462.8
1517.9	1466.6	1474.4
1564.4	1479.1	1494.5
1630.7	1490.9	1505.1
1688	1604.3	1686.8
1715.4	1708.1	1723.3
2996	1786.7	2998.1
3036.8	2996.7	3012.5
3053.7	3041.7	3028.2
3100.8	3070.1	3041.2
3118.2	3076	3086.8
3138.6	3124.9	3104.8
3206.3	3132.3	3136.4
3701.7	3179.7	3143.6
3709	3674.1	3709.6
3822	3692.4	3715

Table S11. 1-methylcytosine (t) coordinates
Table S11A Reactant state

1	H	2.15	-1.638	1.283
2	H	-2.566	-0.935	0.441
3	H	1.859	0.04	1.873
4	C	1.978	-0.592	0.985
5	H	2.825	-0.245	0.386
6	N	0.779	-0.496	0.15
7	C	-0.462	-0.764	0.687
8	O	0.396	2.27	2.215
9	C	0.918	-0.122	-1.197
10	O	2.002	0.156	-1.71
11	N	-0.272	-0.086	-1.898
12	C	-1.574	-0.358	-1.437

13	N	-2.535	-0.254	-2.295
14	C	-1.612	-0.712	-0.031
15	H	0.736	2.554	1.349
16	H	-0.454	-1.012	1.748
17	H	-0.55	2.49	2.17
18	H	-3.428	-0.475	-1.841
19	H	-0.19	0.18	-2.877

Table S11B. Transition state coordinates

1	H	2.135	-1.326	1.68
2	H	-2.343	-0.807	0.55
3	H	2.313	0.456	1.629
4	C	2.185	-0.451	1.02
5	H	3.022	-0.55	0.325
6	N	0.948	-0.362	0.241
7	C	-0.246	-0.412	0.882
8	O	-0.642	1.325	1.82
9	C	1.058	0.026	-1.131
10	O	2.147	0.188	-1.664
11	N	-0.14	0.181	-1.78
12	C	-1.438	-0.083	-1.291
13	N	-2.38	-0.102	-2.168
14	C	-1.499	-0.227	0.167
15	H	-0.05	2.056	1.551
16	H	-0.233	-0.905	1.851
17	H	-1.476	0.912	0.82
18	H	-3.285	-0.276	-1.719
19	H	-0.074	0.357	-2.783

Table S11C. Photohydrate product coordinates

1	H	1.987	-0.936	1.987
2	H	-1.476	-1.602	0.058
3	H	2.425	0.726	1.468
4	C	2.141	-0.281	1.119
5	H	2.965	-0.662	0.509
6	N	0.926	-0.269	0.309
7	C	-0.34	-0.033	0.98
8	O	-0.591	1.351	1.277
9	C	1.078	0.043	-1.038
10	O	2.173	0.172	-1.589
11	N	-0.106	0.174	-1.761

12	C	-1.409	0.033	-1.303
13	N	-2.373	0.307	-2.105
14	C	-1.496	-0.502	0.107
15	H	0.118	1.668	1.863
16	H	-0.318	-0.617	1.917
17	H	-2.443	-0.194	0.567
18	H	-3.273	0.155	-1.637
19	H	0.013	0.469	-2.728

Table S12. 1-methylcytosine(t) vibrational wave numbers (cm^{-1})

Reactant	Saddle	Product
0	-1518.9	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
26.3	0	88.8
46.4	75.5	121.7
78.9	105.7	141.6
119.1	143.5	182.4
169.6	162.8	319.8
182.9	260.1	336.6
234.5	312.8	365.2
241.1	342.2	378.2
340.1	368.8	408
345.9	392.8	424.3
370.3	447.1	489.9
377.3	457.1	542.7
422.6	507.7	560.4
457.4	533	604.1
539.8	562.2	639.3
605.4	605.2	719.4
629.7	645.6	752.2
705.1	703.8	775.7
733.4	715	846.1
739.8	747.8	876.8
756.3	776	929
782.2	796.6	953.6
802.1	856.1	1029.4
920.8	930.4	1057.8
963	982.9	1105.4

1043.8	1023.9	1147.7
1118.9	1102.4	1164.7
1164.5	1133	1181.9
1192.5	1155.4	1207
1202.7	1175.8	1257.5
1278.1	1252.4	1287.9
1331.1	1315.6	1310.2
1364.9	1360.7	1350.4
1412.9	1374.9	1389.4
1448.8	1390.6	1409
1471.4	1414	1415.7
1504.3	1422	1432.9
1527.9	1472.4	1447.9
1619.6	1495.2	1484
1643	1510.1	1524.1
1704	1666.7	1706.2
1758.4	1677.3	1709.1
3036	1756	3006.8
3122.7	3041	3013.3
3183.9	3121.4	3031.9
3184.8	3145.7	3101.1
3242.9	3158.5	3118.7
3448.3	3187.9	3159.7
3549.4	3439.9	3448.8
3698.7	3526.4	3536.8
3818.4	3672.3	3724.3

4.0 Validation using hybrid density functional and CCSD methods

The result obtained using the PBE functional using the numerical basis set was compared to the two independent methods in order to validate the results. The first method selected was a hybrid density functional theory method. We used the M062X functional developed by Truhlar's group. The second method is the post-Hartree Fock method of the coupled cluster for singles and doubles. The results of this comparison for the 1-methylpyrimidines is present in Table S13. The total energy calculated using each method is presented for reactant transition state and product. Table S14 shows a further step in the analysis of these calculations, in which the energy was converted to eV and the reactant energy for each molecule was arbitrarily set equal to 0.

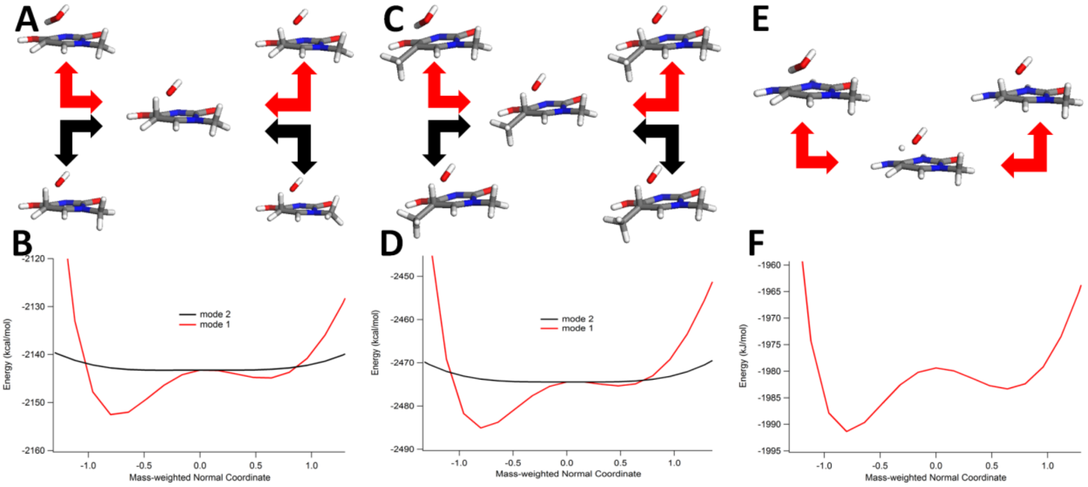

Figure S8. Depiction of the eigenvectors of the normal coordinates with negative eigenvectors for the tautomers of the 1-methylpyrimdines. A. 1-MU has two negative eigenvectors indicated by the red and black arrows. The red arrows show a HO-H bond breaking similar to 1-MU. The black coordinate is a methyl rotation. B. The double well potential corresponding to the projection along the eigenvector. C. 1-MT also has two negative eigenvectors indicated by the red and black arrows with same significance as in A, D. The corresponding potential energy surfaces for 1-MT. E. 1-MU has a single negative eigenvector. F. The potential energy surfaces for the corresponding eigenvector is a could well potential surface..

4.1 Total energies of the normal pyridimidine structures

Table S13. Energies of reactant, product, and transition states for photohydration reactions of cytosine, thymine and uracil in Hartrees calculated using the PBE functional and numerical basis set compared with the CCSD post-Hartree-Fock and the M062x hybrid functional methods.

Uracil	**PBE**	**M062x**	**CCSD**	**CCSD(T)**
Reactant	-530.190964	-530.5168923	-529.356629	-529.426799
Transition	-530.118434	-539.419430	-529.261053	-529.337867
Product	-530.191173	-530.514250	-529.365723	-529.434987
Thymine	**PBE**	**M062x**	**CCSD**	**CCSD(T)**
Reactant	-569.466777	-569.819171	-568.569669	-568.646189
Transition	-569.391151	-569.724268	-568.470884	-568.554487
Product	-569.464214	-569.8228759	-568.574610	-568.650935
Cytosine	**PBE**	**M062x**	**CCSD**	**CCSD(T)**
Reactant	-510.303271	-510.615184	-509.486056	-509.558471
Transition	-510.232672	-510.530432	-509.395245	-509.472828
Product	-510.296671	-510.616772	-509.489329	-509.559890

Table S14. Energies of reactant, product, and transition states for photohydration reactions of cytosine, thymine and uracil in eV calculated using the PBE functional and numerical basis set compared with post-Hartree-Fock CCSD and the M062x hybrid functional using DFT.

Uracil	**PBE**	**M062x**	**CCSD**	**CCSD(T)**
Reactant	0	0	0	0
Transition	1.9728152	2.650974	2.5996684	2.41895
Product	-0.0056856	-0.07187	-0.2473556	-0.222713
Thymine	**PBE**	**M062x**	**CCSD**	**CCSD(T)**
Reactant	0	0	0	0
Transition	2.0570228	2.58136	2.58136	2.494240
Product	0.0697092	-0.10077	-0.134392	-0.129146
Cytosine	**PBE**	**M062x**	**CCSD**	**CCSD(T)**
Reactant	0	0	0	0
Transition	1.9202916	2.30525	2.4700288	2.329489
Product	0.1795188	-0.0432	-0.0890288	-0.03859

4.2 Total energies of tautomers of pyridine structures

Table S15. Total energies of reactant, product, and transition states for photohydration reactions of tautomers of cytosine, thymine and uracil in Hartrees calculated using the PBE functional and numerical basis set compare to CCSD post-Hartree-Fock and the M062x hybrid functional methods.

Uracil	**PBE**	**M062x**	**CCSD**
Reactant	-530.160693	-530.476965	-529.324498
Transition	-530.106410	-530.391106	-529.232584
Product	-530.152615	-530.475661	-529.319870
Thymine	**PBE**	**M062x**	**CCSD**
Reactant	-569.437829	-569.783150	-568.534209
Transition	-569.363275	-569. 694717	-568.441182
Product	-569.425249	-569.781579	-568.530973
Cytosine	**PBE**	**M062x**	**CCSD**
Reactant	-510.298670	-510.612846	-509.485967
Transition	-510.227904	-510.527734	-509.394191
Product	-510.300518.	-510.623673	-509.497562

Table S16. Difference energies of reactant, product, and transition states for photohydration reactions of tautomers of cytosine, thymine and uracil in eV calculated using the PBE functional and numerical basis set compare to CCSD post-Hartree-Fock and the M062x hybrid functional methods.

Uracil	**PBE**	**M062x**	**CCSD**
Reactant	0	0	0
Transition	1.924835	2.3353168	2.5000152
Product	0.05026	0.0354208	0.125836
Thymine	**PBE**	**M062x**	**CCSD**
Reactant	0	0	0
Transition	2.02782	2.4053976	2.5303496
Product	0.3421272	0.0427512	0.0880344
Cytosine	**PBE**	**M062x**	**CCSD**
Reactant	0	0	0
Transition	1.8965112	2.3150352	2.4963048
Product	-0.07859	-0.294506	-0.315386

4.3 Application of the equation of motion CCSD method

The equation of motion (EOM) CCSD calculation was implemented in order to test whether the transition state was coincident with an intersection between the ground and excited state. This is an extremely expensive calculation and therefore it was only carried out for the transition state geometry. A benchmark for the single point calculation of the uracil hydrate required 11 seconds of CPU time, while the EOMCCSD calculation on the same system required 1.24 x 10^6 seconds. The results indicated that the transition state is not a conical intersection. There were no curve crossings at the transition state detected by the EOM-CCSD method. Similar results were obtained using TD-DFT methods, but these computations were not included since they add nothing new.

Table S17. Equation of motion CCSD energies (in Hartrees)

Uracil	**EOMCCSD**
Transition	-529.195391
Thymine	**EOMCCSD**
Transition	NA
Cytosine	**EOMCCSD**
Transition	-509.395246

4.3.1 EOM CCSD output for 1-methyluracil and H_2O

Excitation energies and oscillator strengths:

```
---------------------------------------------
Excited State   1:      Singlet-A      4.3235 eV   286.76 nm   f=?
Right Eigenvector
Alpha Singles Amplitudes
    I     SymI     A     SymA      Value
   38       1     39       1     0.225984
   38       1     40       1    -0.105999
   38       1     41       1    -0.116993
   38       1     42       1     0.111862
   38       1     43       1    -0.445559
   38       1     45       1     0.227136
   38       1     51       1    -0.108405
Beta   Singles Amplitudes
    I     SymI     A     SymA      Value
   38       1     39       1     0.225984
   38       1     40       1    -0.105999
   38       1     41       1    -0.116993
   38       1     42       1     0.111862
   38       1     43       1    -0.445559
   38       1     45       1     0.227136
   38       1     51       1    -0.108405
Total Energy, E(EOM-CCSD) =   -529.102165657

---------------------------------------------
Excited State   2:      Singlet-A      5.4788 eV   226.30 nm   f=?
Right Eigenvector
Alpha Singles Amplitudes
    I     SymI     A     SymA      Value
   35       1     43       1     0.115705
   35       1     46       1     0.106704
   35       1     48       1     0.212448
   35       1     50       1    -0.115204
   35       1     51       1    -0.288956
   36       1     51       1    -0.129293
   37       1     39       1     0.110276
   37       1     43       1    -0.225246
   38       1     40       1     0.105023
   38       1     46       1    -0.104083
   38       1     48       1    -0.110057
   38       1     51       1     0.160082
Beta   Singles Amplitudes
```

```
        I      SymI      A       SymA        Value
       35         1     43          1      0.115705
       35         1     46          1      0.106704
       35         1     48          1      0.212448
       35         1     50          1     -0.115204
       35         1     51          1     -0.288956
       36         1     51          1     -0.129293
       37         1     39          1      0.110276
       37         1     43          1     -0.225246
       38         1     40          1      0.105023
       38         1     46          1     -0.104083
       38         1     48          1     -0.110057
       38         1     51          1      0.160082

---------------------------------------------
Excited State    3:        Singlet-A        5.7074 eV   217.23 nm   f=?
Right Eigenvector
Alpha Singles Amplitudes
        I      SymI      A       SymA        Value
       35         1     43          1      0.121069
       35         1     48          1      0.176327
       35         1     51          1     -0.229303
       37         1     39          1     -0.189065
       37         1     43          1      0.345569
       37         1     45          1     -0.161769
Beta   Singles Amplitudes
        I      SymI      A       SymA        Value
       35         1     43          1      0.121069
       35         1     48          1      0.176327
       35         1     51          1     -0.229303
       37         1     39          1     -0.189065
       37         1     43          1      0.345569
       37         1     45          1     -0.161769

---------------------------------------------
Excited State    4:        Singlet-A        6.0757 eV   204.06 nm   f=?
Right Eigenvector
Alpha Singles Amplitudes
        I      SymI      A       SymA        Value
       34         1     43          1      0.114682
       37         1     42          1      0.104399
       37         1     43          1     -0.152964
       38         1     39          1     -0.345210
       38         1     40          1     -0.236109
       38         1     42          1      0.272216
```

```
      38        1     45        1      0.138648
      38        1     46        1      0.132581
 Beta   Singles Amplitudes
       I     SymI      A     SymA       Value
      34        1     43        1      0.114682
      37        1     42        1      0.104399
      37        1     43        1     -0.152964
      38        1     39        1     -0.345210
      38        1     40        1     -0.236109
      38        1     42        1      0.272216
      38        1     45        1      0.138648
      38        1     46        1      0.132581
```

4.3.2 EOM CCSD output for 1-methylcytosine and H_2O

Excitation energies and oscillator strengths:

```
 ---------------------------------------------
 Excited State    1:        Singlet-A        4.3946 eV   282.13 nm   f=?
 Right Eigenvector
 Alpha Singles Amplitudes
       I     SymI      A     SymA       Value
      36        1     43        1     -0.121975
      36        1     45        1      0.107351
      36        1     46        1      0.100929
      38        1     39        1      0.145434
      38        1     40        1      0.126320
      38        1     42        1     -0.188517
      38        1     43        1      0.281924
      38        1     44        1      0.167888
      38        1     45        1     -0.217603
      38        1     46        1     -0.168427
      38        1     47        1     -0.126280
      38        1     49        1     -0.101809
      38        1     53        1      0.160972
 Beta   Singles Amplitudes
       I     SymI      A     SymA       Value
      36        1     43        1     -0.121975
      36        1     45        1      0.107351
      36        1     46        1      0.100929
      38        1     39        1      0.145434
      38        1     40        1      0.126320
      38        1     42        1     -0.188517
      38        1     43        1      0.281924
      38        1     44        1      0.167888
```

```
     38        1      45        1     -0.217603
     38        1      46        1     -0.168427
     38        1      47        1     -0.126280
     38        1      49        1     -0.101809
     38        1      53        1      0.160972
Total Energy, E(EOM-CCSD) =    -509.233747612

---------------------------------------------
Excited State     2:         Singlet-A          5.1375 eV    241.33 nm    f=?
Right Eigenvector
Alpha Singles Amplitudes
       I      SymI      A      SymA        Value
     36        1      43        1     -0.134771
     37        1      42        1      0.122516
     37        1      43        1     -0.202173
     37        1      44        1     -0.120142
     37        1      45        1      0.167766
     38        1      41        1     -0.107864
     38        1      43        1      0.205936
     38        1      46        1      0.229490
     38        1      48        1     -0.119420
     38        1      49        1      0.152344
     38        1      50        1      0.117187
Beta    Singles Amplitudes
       I      SymI      A      SymA        Value
     36        1      43        1     -0.134771
     37        1      42        1      0.122516
     37        1      43        1     -0.202173
     37        1      44        1     -0.120142
     37        1      45        1      0.167766
     38        1      41        1     -0.107864
     38        1      43        1      0.205936
     38        1      46        1      0.229490
     38        1      48        1     -0.119420
     38        1      49        1      0.152344
     38        1      50        1      0.117187

---------------------------------------------
Excited State     3:         Singlet-A          5.3779 eV    230.54 nm    f=?
Right Eigenvector
Alpha Singles Amplitudes
       I      SymI      A      SymA        Value
     34        1      43        1      0.199500
     34        1      44        1      0.116577
     34        1      45        1     -0.140671
```

```
     34         1      48          1     -0.115502
     35         1      41          1     -0.110930
     35         1      43          1      0.316274
     35         1      44          1      0.163382
     35         1      45          1     -0.198939
     35         1      48          1     -0.188248
     38         1      43          1     -0.111467
     38         1      46          1     -0.103334
Beta    Singles Amplitudes
       I      SymI        A       SymA         Value
     34         1      43          1      0.199500
     34         1      44          1      0.116577
     34         1      45          1     -0.140671
     34         1      48          1     -0.115502
     35         1      41          1     -0.110930
     35         1      43          1      0.316274
     35         1      44          1      0.163382
     35         1      45          1     -0.198939
     35         1      48          1     -0.188248
     38         1      43          1     -0.111467
     38         1      46          1     -0.103334

---------------------------------------------
Excited State      4:           Singlet-A          5.6317 eV    220.15 nm    f=?
Right Eigenvector
Alpha Singles Amplitudes
       I      SymI        A       SymA         Value
     34         1      41          1     -0.110310
     34         1      43          1      0.317698
     34         1      44          1      0.165733
     34         1      45          1     -0.196085
     34         1      48          1     -0.192578
     35         1      43          1     -0.205688
     35         1      45          1      0.103308
     35         1      46          1     -0.128848
     35         1      48          1      0.133513
     37         1      43          1      0.135360
Beta    Singles Amplitudes
       I      SymI        A       SymA         Value
     34         1      41          1     -0.110310
     34         1      43          1      0.317698
     34         1      44          1      0.165733
     34         1      45          1     -0.196085
     34         1      48          1     -0.192578
     35         1      43          1     -0.205688
```

35	1	45	1	0.103308
35	1	46	1	-0.128848
35	1	48	1	0.133513
37	1	43	1	0.135360

5.0 Crossover NEB calculation between two different tautomers

A NEB calculation was implemented with an initial structure consisting of the normal tautomer of each of the three pyrimidines and a final structure consisting of the tautomeric form. It appears that the algorithm misses the transition state for uracil and thymidine. The energies are reported below for sake of completeness. However, all comparisons of the two tautomeric forms of each pyrimidine in the manuscript are based on two parallel NEB trajectories, one for each tautomer. The crossover was not discussed since the significance of the calculation is not clear.

Table S18. Energies of reactant, product, and transition states for photohydration reactions for the reaction from the normal form to the tautomer of cytosine, thymine and uracil in Hartrees calculated using the PBE functional and numerical basis set compare to CCSD post-Hartree-Fock and the M062x hybrid functional methods.

Uracil	**PBE**	**M062x**	**CCSD**
Reactant	-530.190964	-530.516892	-529.356629
Transition	-530.166690	-530.401242	-529.242232
Product	-530.156534	-530.475661	-529.319870
Thymine	**PBE**	**M062x**	**CCSD**
Reactant	-569.466777	-569.812267	-568.569669
Transition	-569.425227	-569.782006	-568.
Product	-569.425249	-569.781579	-568.530973
Cytosine	**PBE**	**M062x**	**CCSD**
Reactant	-510.301404	-510.615184	-509.486056
Transition	-510.227994	-510.527513	-509.393706
Product	-510.300518	-510. 623673	-509.497562

6.0 Excited state geometry optimization at the transition state geometry

We conducted calculations of the excited state energy and geometry starting at the transition state geometry in order to ascertain whether there is a conical intersection that lies near the transition state, but is not precisely at the transition state geometry.

The optimization was constrained in the first excited state using the Occupation Fixed option in DMol^3. In this case the charge is disabled and the calculation type is set to "optimize". The result of geometry optimization was significant departure from the transition state geometry for all three pyrimidines. The water molecule is no longer in a bond forming geometry. Moreover, the ring distorts to a folded geometry that is similar to the geometry reported for the conical intersection from the excited state of uridine. The energy of the geometry optimized state is more than 25 kcal/mol higher than the transition state.

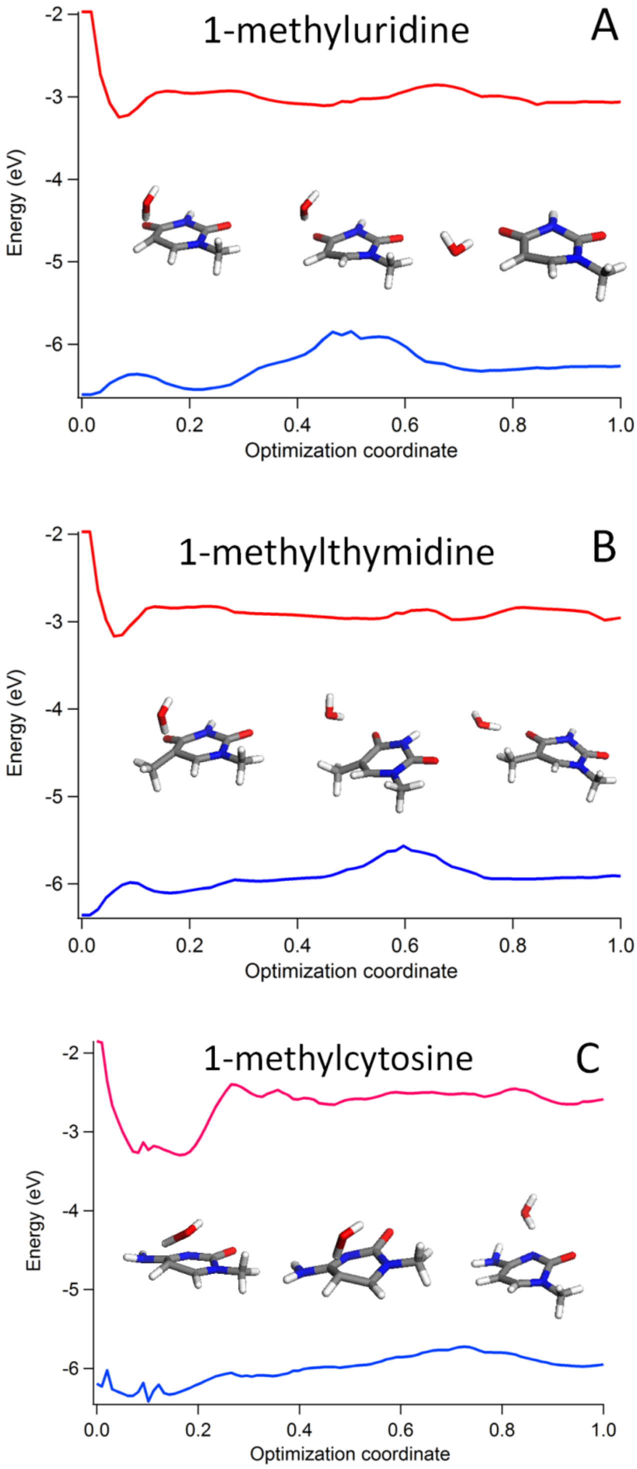

Figure S9. Energies of the HOMO and LUMO for the three pyrimidine hydrate models studied plotted from the transition state (0.0) to the optimized geometry coordinate (1.0) when constrained to find the lowest energy in the fixed occupation of the lowest excited state. The models are A. 1-MU, B. 1-MT and C. 1-MC.

In the case of uridine, this energy is approximately where the calculated CI is observed in the calculation of Ref. 14 (Matsika, *J. Phys. Chem. A*, 2004, 108, 7584-7590). Thus, these calculations show that there is no CI at the transition state geometry and the CI is located at a geometry, in which the H_2O molecule is dissociated from the pyridimine and in a folded geometry shown in Figure 5 of the text and the energies and selected structures along the trajectories for the optimization are plotted in Figure S9.

Table S19. Energies of transition states and the geometry optimized excited states starting at the transition geometry.

Uracil	**PBE Hartrees (kcal/mol)**
Optimized Excited Transition State	-530.063115
Transition State	-530.118434
Energy Gap	0.05531 (34.71)
Thymine	
Optimized Excited Transition State	-569.3478092
Transition State	-569.3911510
Energy Gap	0.04334 (27.20)
Cytosine	
Optimized Excited Transition State	-510.1816832
Transition State	-510.2326720
Energy Gap	0.0509 (31.99)

7.0 Ethene hydration calculated using the NEB method

The ethene hydration NEB calculation serves as a model system. Ground state NEB trajectory for ethene hydration is shown in Figure S10. The coordinates for each of the models, reactant, transition state and product are given Table S20. The total energies from the calculation using various methods are given Table S21. The barrier for hydrate formation is approximately ~2.0 eV or ~46.0 kcal/mol. This value is only slightly larger than the value for uracil or thymidine. The energies of the HOMO-1, HOMO and LUMO are shown in Figure S11. As seen in Tables 1 and S22, the formation of hydrates in the pyrimidines is slightly endothermic based on the PBE functional in DMol^3. The relative simplicity of ethene suggests it may be a good model for the transition state. In addition to reporting the coordinates for reactant, transition state and product (Tables S20A, S20B and S20C), the wave numbers for each state are reported in Table S21 in cm^{-1}.

The NEB for the model system of ethene + H_2O calculated using the hybrid-metal M06-2x functional and a 6-311+G(d,p) basis set implemented in Gaussian09 hadcomputational cost 9.3 times greater than the DMol^3 approach. The coupled cluster of singles and doubles (CCSD) method also implemented in Gaussian09 had a computational cost 1,800 times greater than the DMol^3 approach. We found that the EOMCCSD method required CPU time >26,000 times greater than the DMol^3 approach. Because of the scaling of these methods, the different factors are larger by a factor of 3-4 for the pyrimidines. For example, the EOMCCSD requires $>10^5$ times greater CPU for ethene and 1-MU where direct comparisons were made.

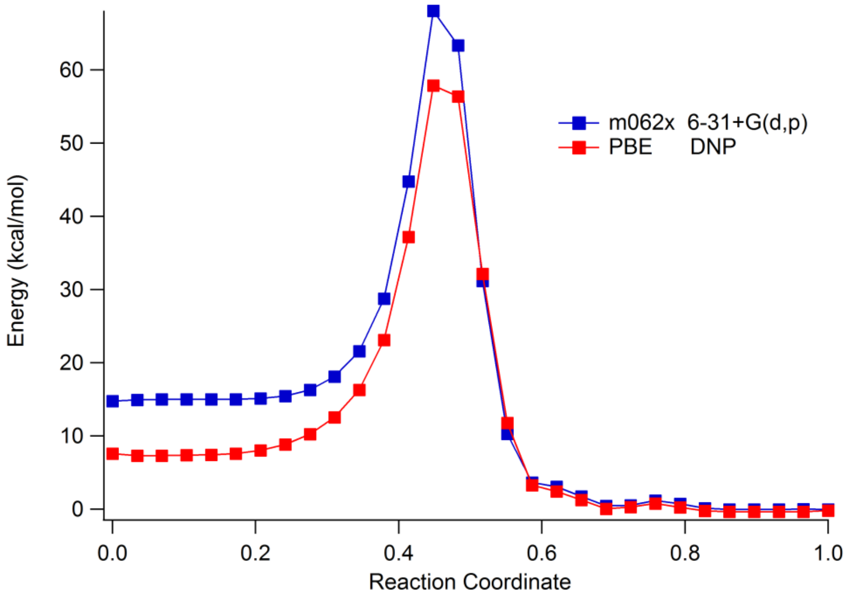

Figure S10. Energy of the NEB trajectory for each of 30 states calculated from reactant to product. The energy reported is based on the binding energy for each geometry calculated by two DFT methods. The first method is the numerical basis set (DNP) and the PBE functional calculated using DMol^3. The second method is hybrid functional M062x and calculation using Gaussian09.

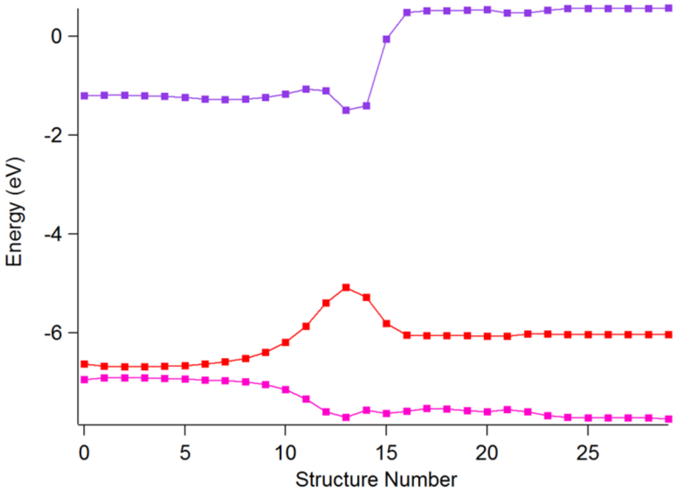

Figure S11. Energy of the HOMO-1, HOMO and LUMO for ethane for each of the 30 geometries used in the NEB calculation calculated using the numerical basis set (DNP) and the PBE functional calculated using DMol^3.

Table S20. Ethene hydration coordinates

Table S20A. Reactant coordinates

1	C	-1.072	-0.136	0.500
2	H	-1.675	0.719	0.186
3	H	-1.388	-1.115	0.132
4	C	-0.00779	0.0064	1.300
5	H	0.306	0.985	1.672
6	H	0.591	-0.851	1.617
7	H	0.745	-0.091	-1.119

8	O	1.214	-0.208	-1.966
9	H	1.288	0.692	-2.322

Table S20B. Transition state coordiantes

1	C	-0.717	-0.123	-0.109
2	H	-1.241	0.754	-0.496
3	H	-1.149	-1.078	-0.409
4	C	0.125	-0.0188	1.031
5	H	0.178	0.942	1.549
6	H	0.279	-0.905	1.651
7	H	0.997	-0.041	-0.0899
8	O	0.724	-0.193	-1.329
9	H	0.805	0.663	-1.799

Table S20C. Photohydrate product coordinates

1	C	-0.289	-0.0627	-0.458
2	H	-1.231	0.503	-0.591
3	H	-0.455	-1.086	-0.842
4	C	0.0822	-0.113	1.011
5	H	0.232	0.901	1.406
6	H	-0.713	-0.600	1.594
7	H	1.015	-0.678	1.153
8	O	0.787	0.576	-1.164
9	H	0.571	0.559	-2.111

Table S21. Ethene photohydrate wave numbers (cm^{-1})

Initial	Saddle	Final
0	-1843.5	0
0	0	0
0	0	0
0	0	0
0	0	0
0	0	0
107.4	0	277.7
184.3	375	326.4
197.6	440.6	409.5
215.9	536.3	786.5
369.8	624.8	880.7
410.1	723.6	1009.9

879.1	821.1	1074.5
1008.9	860.9	1138.9
1019.9	987.7	1244.3
1097.3	1080.4	1265.6
1241.2	1206.1	1359.9
1359.5	1216.6	1412.1
1471.5	1376.5	1444.1
1634.5	1423.2	1467.8
1651.6	1492.2	1471.9
3154.7	1544.3	2938.9
3165.9	3090.6	3008.5
3244.7	3114.4	3018.7
3268.7	3180.2	3093.8
3652.8	3220.4	3126.7
3802.9	3651.6	3743.1

7.1 Comparison calculations and validation for ethene + H_2O

Table S22. Total energies of the reactant, transition state and product of the ethene + H_2O system in Hartrees.

Ethene	**PBE**	**M062x**	**CCSD**	**EOMCCSD**	**CCSD(T)**
Reactant	-154.889418	-154.989080	-154.658842	-154.658377	-154.675752
Transition	-154.815877	-154.896507	-154.555140	-154.555141	-154.578195
Product	-154.913455	-155.012620	-154.678573	-154.678455	-154.694915

7.2 Equation of Motion CCSD output for ethene + H_2O

The EOM CCSD transition properties are given for the reactant, saddle point and product of the ethene + H_2O system. The lowest root is 7.9 eV, 4.2 eV and 6.9 eV for the three respective structures. This shows that the ground state saddle point is not near a degeneracy or curve crossing with the excited state.

7.2.1 Reactant

Excitation energies and oscillator strengths:

```
 ---------------------------------------------
 Excited State   1:      Singlet-A      7.9032 eV  156.88 nm  f=?
 Right Eigenvector
 Alpha Singles Amplitudes
      I     SymI     A     SymA      Value
     13        1     14       1   -0.638943
     13        1     20       1    0.179651
     13        1     21       1    0.117731
 Beta   Singles Amplitudes
```

```
      I     SymI     A     SymA      Value
     13       1     14       1    -0.638943
     13       1     20       1     0.179651
     13       1     21       1     0.117731
 Total Energy, E(EOM-CCSD) =   -154.368406521

 ---------------------------------------------
 Excited State   2:        Singlet-A        8.0272 eV   154.45 nm   f=?
 Right Eigenvector
 Alpha Singles Amplitudes
      I     SymI     A     SymA      Value
     12       1     14       1    -0.331298
     12       1     18       1     0.376128
     12       1     20       1     0.183565
     12       1     21       1    -0.269355
     12       1     28       1    -0.107135
     13       1     16       1    -0.252019
 Beta  Singles Amplitudes
      I     SymI     A     SymA      Value
     12       1     14       1    -0.331298
     12       1     18       1     0.376128
     12       1     20       1     0.183565
     12       1     21       1    -0.269355
     12       1     28       1    -0.107135
     13       1     16       1    -0.252019

 ---------------------------------------------
 Excited State   3:        Singlet-A        8.2406 eV   150.45 nm   f=?
 Right Eigenvector
 Alpha Singles Amplitudes
      I     SymI     A     SymA      Value
     12       1     14       1     0.163816
     12       1     18       1    -0.188151
     12       1     21       1     0.147746
     13       1     16       1    -0.571379
     13       1     23       1     0.155019
 Beta  Singles Amplitudes
      I     SymI     A     SymA      Value
     12       1     14       1     0.163816
     12       1     18       1    -0.188151
     12       1     21       1     0.147746
     13       1     16       1    -0.571379
     13       1     23       1     0.155019
```

7.2.2 Saddle point

```
Excitation energies and oscillator strengths:

 ---------------------------------------------
 Excited State   1:      Singlet-A      4.2211 eV  293.72 nm  f=?
 Right Eigenvector
 Alpha Singles Amplitudes
     I     SymI    A    SymA     Value
    13       1    14      1    0.265289
    13       1    15      1   -0.204166
    13       1    16      1    0.167834
    13       1    17      1   -0.238498
    13       1    19      1    0.247557
    13       1    21      1    0.337476
    13       1    22      1   -0.198093
 Beta  Singles Amplitudes
     I     SymI    A    SymA     Value
    13       1    14      1    0.265289
    13       1    15      1   -0.204166
    13       1    16      1    0.167834
    13       1    17      1   -0.238498
    13       1    19      1    0.247557
    13       1    21      1    0.337476
    13       1    22      1   -0.198093
 Total Energy, E(EOM-CCSD) =  -154.400016237

 ---------------------------------------------
 Excited State   2:      Singlet-A      5.6076 eV  221.10 nm  f=?
 Right Eigenvector
 Alpha Singles Amplitudes
     I     SymI    A    SymA     Value
    13       1    14      1    0.477404
    13       1    15      1    0.360792
    13       1    16      1   -0.139642
    13       1    17      1    0.155125
    13       1    20      1    0.142680
 Beta  Singles Amplitudes
     I     SymI    A    SymA     Value
    13       1    14      1    0.477404
    13       1    15      1    0.360792
    13       1    16      1   -0.139642
    13       1    17      1    0.155125
    13       1    20      1    0.142680

 ---------------------------------------------
```

```
Excited State   3:      Singlet-A      6.5699 eV  188.72 nm   f=?
Right Eigenvector
Alpha Singles Amplitudes
     I     SymI     A     SymA      Value
    13       1     15       1     0.319110
    13       1     16       1     0.555092
    13       1     21       1    -0.117750
Beta  Singles Amplitudes
     I     SymI     A     SymA      Value
    13       1     15       1     0.319110
    13       1     16       1     0.555092
    13       1     21       1    -0.117750
```

7.2.3 Product

```
Excitation energies and oscillator strengths:

---------------------------------------------
Excited State   1:      Singlet-A      6.8761 eV  180.31 nm   f=?
Right Eigenvector
Alpha Singles Amplitudes
     I     SymI     A     SymA      Value
    11       1     14       1     0.142842
    13       1     14       1    -0.540257
    13       1     15       1     0.174109
    13       1     16       1     0.224938
    13       1     24       1     0.160516
Beta  Singles Amplitudes
     I     SymI     A     SymA      Value
    11       1     14       1     0.142842
    13       1     14       1    -0.540257
    13       1     15       1     0.174109
    13       1     16       1     0.224938
    13       1     24       1     0.160516
Total Energy, E(EOM-CCSD) =   -154.425881330

---------------------------------------------
Excited State   2:      Singlet-A      8.2684 eV  149.95 nm   f=?
Right Eigenvector
Alpha Singles Amplitudes
     I     SymI     A     SymA      Value
    13       1     14       1    -0.336507
    13       1     15       1    -0.281470
    13       1     16       1    -0.398904
```

```
      13          1       18          1     -0.187458
      13          1       20          1     -0.134165
      13          1       22          1     -0.104399
 Beta    Singles Amplitudes
       I        SymI       A        SymA        Value
      13          1       14          1     -0.336507
      13          1       15          1     -0.281470
      13          1       16          1     -0.398904
      13          1       18          1     -0.187458
      13          1       20          1     -0.134165
      13          1       22          1     -0.104399

 ---------------------------------------------
 Excited State     3:          Singlet-A          8.6320 eV    143.63 nm    f=?
 Right Eigenvector
 Alpha Singles Amplitudes
       I        SymI       A        SymA        Value
      13          1       15          1     -0.442262
      13          1       16          1       0.373041
      13          1       20          1     -0.292709
 Beta    Singles Amplitudes
       I        SymI       A        SymA        Value
      13          1       15          1     -0.442262
      13          1       16          1       0.373041
      13          1       20          1     -0.292709
```

8.0 Trans-thymine photohydration

The trans photohydration mechanism was studied using 1-methylthymine and H_2O. The transition state involved breaking a C-C bond in the ring to provide a route for the H atom to pass to the opposite side of the ring from the attacking H_2O. However, the barrier for this process was significantly higher than the other models studied. The NEB calculation for the photohydration reaction of thymine with trans addition to the C5-C6 π-bond is shown in Figure S12. The transition state is >100 kcal/mol higher in energy than in the reactant state. This is significantly larger than the barriers for the other reactions studied, which all have <50 kcal/mol based on the calculation using the PBE functional. Therefore, this pathway was discounted and it was concluded that it plays no role in the photochemistry of thymine.

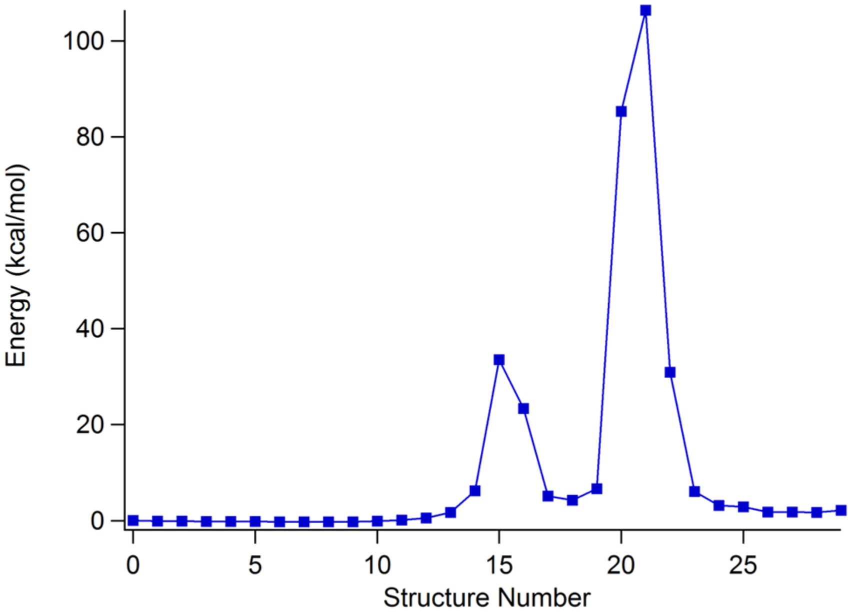

Figure S12. Ground state energy (S_0) for the trans addition of water in the 1-MT + H_2O system calculated using the NEB method.

9.0 Validation of the PBE functional using a numerical basis set.

We conducted comparison calculations on H_2O in order to determine the quality of the PBE functional compared to other reference methods. As elsewhere in this study we have chosen the M062x hybrid functional and the CCSD calculation as references for comparison. The vibrational frequencies, bond lengths/angles and dipole moments are compared in Tables S23, S25 and S26, respectively. The percent error in the normal mode calculation is given Table S24. Of the methods used, CCSD is the most accurate for frequencies and for structure. However, the PBE method is comparable to M062x and CISD in the calculation of frequencies and the angle. Only the bond length is less accurate. Although PBE is less accurate than these other methods, the % error in the bond length is 1.34% (Table S25). Finally, the dipole moment of water was compared. PBE gives by far the best agreement with the gas phase dipole moment for H_2O (Table S26).

Table S23. Comparison of vibrational wave numbers (cm^{-1}) for the three normal modes of H_2O calculated using various methods.

Method	PBE	M062X	CCSD	CISD	Expmt
Bend	1630.6	1604.9	1658	1670.3	1654
Sym. Stretch	3718.3	3906.3	3897	3949	3825
Asym. Stretch	3833.7	4013	3999	4046.1	3935

Table S24. Percent error for the three calculated normal mode of H_2O.

Method	PBE	M062X	CCSD	CISD

Bend	1.415	2.969	0.242	0.985
Sym. Stretch	2.790	2.125	1.882	3.242
Asym. Stretch	2.574	1.982	1.626	2.823
Average	2.260	2.359	1.250	2.350

Table S25. Calculated bond lengths and angles using various methods.

Method	PBE	M062x	CCSD	CISD	Experiment
δ (Å)	0.970	0.959	0.958	0.955	0.9572
%Error in δ	1.34	0.19	0.084	0.23	- - -
θ (degrees)	103.7	105.0	103.7	103.9	104.5
%Error in θ	0.77	0.48	0.77	0.57	- - - -

Table S26. Calculated dipole moment using various methods.

Method	PBE	M062x	CCSD	CISD	Experiment
μ (Debye)	1.905	2.172	2.352	2.246	1.86
%Error	2.42	16.8	26.5	20.8	- - -

10.0 Mulliken and Electrostatic Potential Fitting Charge Analysis

In the text we presented an analysis of the change the nuclear charge in the transition state relative to the reactant and product states based electrostatic potential (ESP) fitting. The Mulliken charge is another method for partitioning charge. Since the assignment of nuclear charge is not unique, it is worthwhile to compare the conclusion of different methods in order ensure that the conclusion reached is robust. Indeed, the analysis using the Mulliken charge set leads one to the same conclusion as the ESP fitting, namely that there is a polarization of the π-bond between C5 and C6 in all of the pyrimidines.

The Mulliken charges for the initial, saddle point and final structures are presented in Tables S27, S28 and S29 for 1-MU, 1-MT and 1-MC, respectively. The polarization of the C5-C6 bond is observed for each of the species studied. The C5 carbon acquires a significant negative charge in the transition state relative to the initial state. The shift in Mulliken charge is -0.225, -0.263 and -0.188 for 1-MU, 1-MT and 1-MC, respectively. The C6 carbon is shifted to more positive values of the Mulliken charge in the transition state by 0.094, 0.117 and 0.090 for 1-MU, 1-MT and 1-MC, respectively. In all cases, the shift in C6 is about half as large as the shift in C5. Aside from the hydrogen atoms H5 and H6, the next largest shift in Mulliken charge is on C4. However, this is less than 0.040 in all cases.

The electrostatic potential fitted charge sets for the three pyrimidines are given in Tables S30 - S32. Similar trends are observed to those obtained using Mulliken charge.

Table S27. Mulliken charges for 1-MU

Atom	Saddle	Initial	Final	Type	I - S	F - S
H (1)	0.119	0.125	0.114	Hm	0.006	-0.005
H (2)	0.102	0.139	0.143	Hm	0.037	0.041
H (2)	0.161	0.167	0.161	Hm	0.006	0
C (4)	-0.221	-0.201	-0.19	Cm	0.02	0.031
N (5)	-0.297	-0.293	-0.354	N1	0.004	-0.057
C (6)	0.542	0.55	0.56	C2	0.008	0.018
O (7)	-0.458	-0.437	-0.458	O2	0.021	0
N (8)	-0.324	-0.325	-0.307	N3	-0.001	0.017
H (9)	0.213	0.216	0.211	H3	0.003	-0.002
C (10)	0.417	0.455	0.445	C4	0.038	0.028
O (11)	-0.436	-0.423	-0.412	O4	0.013	0.024
C (12)	-0.197	-0.422	-0.285	C5	-0.225	-0.088
H (13)	0.154	0.126	0.094	H5	-0.028	-0.060
C (14)	0.114	0.208	0.244	C6	0.094	0.130
H (15)	0.099	0.13	0.088	H6	0.031	-0.011
O (16)	-0.539	-0.586	-0.481	Ow	-0.047	0.058
H (17)	0.278	0.263	0.269	Hw_6	-0.015	-0.009
H (18)	0.274	0.309	0.158	Hw_5	0.035	-0.116

Table S28. Mulliken charges for 1-MT

H (1)	0.117	0.124	0.111	Hm	0.007	-0.006
H (2)	0.151	0.126	0.097	Hm	-0.025	-0.054
H (3)	0.16	0.166	0.157	Hm	0.006	-0.003
C (4)	-0.22	-0.199	-0.186	Cm	0.021	0.034
N (5)	-0.293	-0.287	-0.357	N1	0.006	-0.064
C (6)	0.534	0.546	0.557	C2	0.012	0.023
O (7)	-0.464	-0.438	-0.459	O2	0.026	0.005
N (8)	-0.319	-0.322	-0.307	N3	-0.003	0.012
H (9)	0.211	0.215	0.209	H3	0.004	-0.002
C (10)	0.424	0.467	0.466	C4	0.043	0.042
O (11)	-0.447	-0.432	-0.416	O4	0.015	0.031
C (12)	-0.083	-0.346	-0.231	C5	-0.263	-0.148
C (13)	0.092	0.209	0.267	C6	0.117	0.175
H (14)	0.092	0.126	0.082	H6	0.034	-0.01
C (15)	-0.33	-0.289	-0.285	C5m	0.041	0.045

H (16)	0.145	0.141	0.135	H5m	-0.004	-0.01
H (17)	0.102	0.111	0.109	H5m	0.009	0.007
H (18)	0.114	0.102	0.115	H5m	-0.012	-0.006
O (19)	-0.539	-0.595	-0.48	Ow	-0.056	0.059
H (20)	0.278	0.259	0.268	Hw_6	-0.019	-0.01
H (21)	0.274	0.313	0.146	Hw_5	0.039	-0.128

Table S29. Mulliken charges for 1-MC

Atom	Saddle	Initial	Final	Type	I - S	F - S
H (1)	0.114	0.112	0.101	Hm	-0.002	-0.013
H (2)	0.127	0.118	0.085	Hm	-0.009	-0.042
H (3)	0.166	0.163	0.157	Hm	-0.003	-0.009
C (4)	-0.21	-0.19	-0.184	Cm	0.02	0.026
N (5)	-0.297	-0.288	-0.362	N1	0.009	-0.065
C (6)	0.465	0.474	0.492	C2	0.009	0.027
O (7)	-0.545	-0.617	-0.497	O2	-0.072	0.048
N (8)	-0.329	-0.328	-0.309	N3	0.001	0.02
C (9)	0.273	0.309	0.277	C4	0.036	0.004
C (10)	-0.228	-0.416	-0.275	C5	-0.188	-0.047
H (11)	0.08	0.119	0.144	H5	0.039	0.064
C (12)	0.117	0.207	0.259	C6	0.090	0.142
H (13)	0.277	0.259	0.268	H6	-0.018	-0.009
N (14)	-0.336	-0.32	-0.322	Nam	0.016	0.014
H (15)	0.207	0.21	0.209	Ham	0.003	0.002
H (16)	0.196	0.203	0.2	Ham	0.007	0.004
O (17)	-0.455	-0.427	-0.448	Ow	0.028	0.007
H (18)	0.101	0.118	0.077	Hw_6	0.017	-0.024
H (19)	0.277	0.295	0.128	Hw_5	0.018	-0.149

Table S30. Electrostatic potential fitted charges for 1-MU

H	0.167	0.152	0.179	Hm	0.015	0.027
H	0.224	0.263	0.174	H5	-0.039	-0.089
H	0.239	0.175	0.182	Hm	0.064	0.007
C	-0.591	-0.446	-0.556	Cm	-0.145	-0.11
H	0.222	0.209	0.226	Hm	0.013	0.017
N	0.147	-0.156	-0.205	N1	0.303	-0.049
C	-0.086	0.336	0.415	C6	-0.422	0.079
O	-0.760	-0.730	-0.576	Ow	-0.03	0.154
C	0.598	0.723	0.720	C2	-0.125	-0.003
O	-0.530	-0.527	-0.532	O2	-0.003	-0.005
N	-0.624	-0.634	-0.726	N3	0.010	-0.092
H	0.372	0.364	0.384	H3	0.008	0.02

C	0.713	0.699	0.833	C4	0.014	0.134
O	-0.538	-0.506	-0.543	O4	-0.032	-0.037
C	-0.511	-0.801	-0.628	C5	0.290	0.173
H	0.373	0.386	0.381	Hw_6	-0.013	-0.005
H	0.202	0.156	0.052	H6	0.046	-0.104
H	0.381	0.336	0.220	Hw_5	0.045	-0.116

Table S31. Electrostatic potential fitted charges for 1-MT

H	0.237	0.136	0.172	Hm	0.101	0.036
H	-0.889	0.260	-0.565	H5	-1.149	-0.825
H	0.317	0.163	0.18	Hm	0.154	0.017
C	-0.859	-0.404	-0.557	Cm	-0.455	-0.153
H	0.281	0.194	0.219	Hm	0.087	0.025
N	0.329	-0.273	-0.088	N1	0.602	0.185
C	-0.499	0.432	0.189	C6	-0.931	-0.243
O	-0.823	-0.491	-0.625	Ow	-0.332	-0.134
C	0.568	0.703	0.628	C2	-0.135	-0.075
O	-0.537	-0.525	-0.524	O2	-0.012	0.001
N	-0.670	-0.664	-0.613	N3	-0.006	0.051
H	0.406	0.419	0.365	H3	-0.013	-0.054
C	0.530	0.464	0.660	C4	0.066	0.196
O	-0.509	-0.381	-0.526	O4	-0.128	-0.145
C	0.260	-0.674	-0.037	C5	0.934	0.637
H	0.403	0.390	0.417	Hw_5	0.013	0.027
H	0.299	0.103	0.090	Hw_6	0.196	-0.013
C	0.409	-0.307	0.126	C5m	0.716	0.433
H	0.257	0.090	0.182	H5m	0.167	0.092
H	0.243	0.153	0.163	H5m	0.09	0.01
H	0.249	0.211	0.144	H5m	0.038	-0.067

Table S32. Electrostatic potential fitted charges for 1-MC

H	0.165	0.154	0.161	Hm	0.011	0.007
H	0.245	0.271	0.171	H5	-0.026	-0.1
H	0.212	0.183	0.169	Hm	0.029	-0.014
C	-0.564	-0.474	-0.517	Cm	-0.09	-0.043
H	0.215	0.21	0.209	Hm	0.005	-0.001
N	0.085	-0.22	-0.277	N1	0.305	-0.057
C	-0.028	0.331	0.367	C6	-0.359	0.036
O	-0.724	-0.747	-0.627	Ow	0.023	0.12
C	0.766	0.923	0.862	C2	-0.157	-0.061
O	-0.579	-0.564	-0.556	O2	-0.015	0.008

N	-0.746	-0.825	-0.811	N3	0.079	0.014
C	0.739	0.841	0.893	C4	-0.102	0.052
N	-0.790	-0.888	-0.929	Nam	0.098	-0.041
C	-0.619	-0.857	-0.616	C5	0.238	0.241
H	0.352	0.375	0.406	Hw_6	-0.023	0.031
H	0.177	0.158	0.065	H6	0.019	-0.093
H	0.364	0.312	0.212	Hw_5	0.052	-0.1
H	0.374	0.419	0.421	Ham	-0.045	0.002
H	0.354	0.398	0.395	Ham	-0.044	-0.003