Electronic Supplementary Material (ESI) for Faraday Discussions.

Supplementary Information
for

Probing deactivation pathways of DNA nucleobases by two-dimensional electronic spectroscopy: first principles simulations

Artur Nenov,*[a] Javier Segarra-Martí,[b‡] Angelo Giussani,[b‡] Irene Conti, [b] Ivan Rivalta,[c] Elise Dumont,[c] Vishal K. Jaiswal,[b] Salvatore Flavio Altavilla,[b] Shaul Mukamel,[d] and Marco Garavelli*[b,c]

[a] Dipartimento di Chimica G. Ciamician, Universit di Bologna, Via F. Selmi 2, 40126 Bologna, Italy.
Tel: +39 051 2099495; E-mail: Artur.Nenov@unibo.it
[b] Dipartimento di Chimica G. Ciamician, Universit di Bologna, Via F. Selmi 2, 40126 Bologna, Italy.
[c] Laboratoire de Chimie, Ecole Normale Superieure de Lyon, 46 alle dItalie, 69364 Lyon, France.
[d] Department of Chemistry, University of California, Irvine, California 92697-2025, United States.
[‡] ‘These authors contributed equally to this work.’

1. π-orbitals of adenine

2. Reference calculation for adenine monomer in gas-phase

3. Cartesian coordinates (Angstroms) of all geometries used in the present work

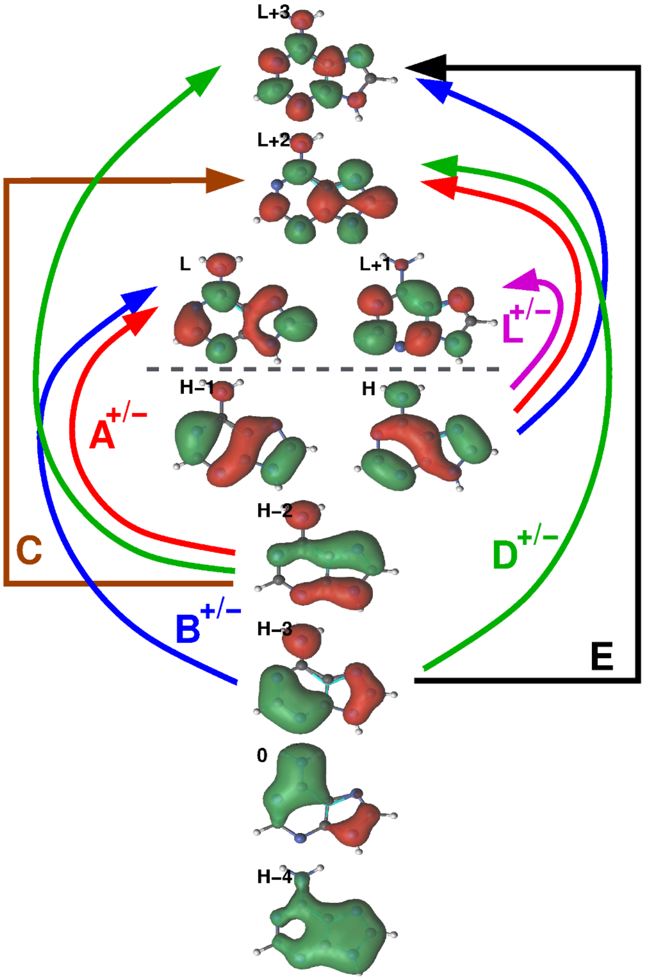

Fig. S1. List of the valence π-orbitals of adenine together with different single electron transitions which determine the low energy part of the absorption spectrum. Nearly degenerate transitions that are likely to mix are combined in groups denoted by a color code.

2. Reference calculation for adenine monomer in gas-phase

A reference calculation of adenine monomer in the gas-phase at a state-of-the-art SA-CASSCF/SS-CASPT2 level was performed to assess the validity of the approximations used in the main article. The protocol for the reference calculations has been first reported in ref. 1. The gas-phase geometry was optimized at single state CASSCF level with a full π-valence active space of twelve electrons in ten orbitals (i.e. CAS(12,10)) under C_S-symmetry with the generally contracted ANO-L basis set[2], whereas the following contraction scheme is adopted: C,O/[4s3p2d] and H/[2s1p]. It is well know that excited states described with the general ANO basis set are contaminated by spurious Rydberg character that can affect excited state energies and properties like transition dipole moments[3,4]. To account for this the basis sets was augmented with a set of 8s, 8p and 8d uncontracted Rydberg-type basis functions positioned at the center of nuclear charge of the lowest two cationic states[5,6] in all excited state calculations. We point out that we intentionally used a set of uncontracted Rydberg-type basis functions instead of the usually utilized 1s, 1p, 1d contracted Rydberg-type set[7] in order to properly describe Rydberg states up to 10.50 eV. The orbital exponents for the uncontracted Rydberg basis functions were obtained from ref. 8. The uneven description of electronic correlation for valence and Rydberg states at CASSCF level is the cause for valence-Rydberg mixing[7] which affects electronic properties of valence states. Well described Rydberg orbitals do not interact with the valence orbitals[5,9], therefore, the following scheme was applied to resolve valence-Rydberg orbital mixing: An ten electrons / nine orbitals active space (i.e. CAS(10,9)) comprising all valence orbitals besides the completely bonding orbital was constructed. The active space was augmented with a Rydberg orbital of A''-symmetry (the Rydberg orbitals of A' symmetry are not considered further), thus forming CAS(10,10). Next a state-averaged calculation was performed where it was paid attention that at least one of the optimized states had Rydberg character. Subsequently, the active space orbitals were localized using the Pipek procedure[10] in order to prevent spurious valence-Rydberg orbital mixing. Finally, the perfectly formed Rydberg orbital was deleted from the molecular orbital list. This procedure was performed 24 times until all Rydberg orbitals were removed and, consequently, no Rydberg states were found among the first 25 excited states. Once the valence orbitals were clean from Rydberg contaminations the lowest occupied orbital was added to the active space, thus, forming the full π-valence active space of twelve π-electrons in ten orbitals (i.e. CAS(12,10)) and the first 25 states with A'-symmetry were computed. To test the stability of the calculation the active space was systematically increased by four, eight and twelve additional extravalence secondary orbitals. The idea is that if the contribution of extravalence orbitals to the overall wavefunctions of the described states is negligible the CASPT2 excitation energies should not differ if the extravalence orbitals are treated variationally or perturbatively. If this is not true, than the orbitals must be included in the AS. A convergence was reached with twelve secondy orbtials, resulting in the active space RAS(0,0|12,10|2,12), where each of the three restricted active subspaces are characterized by the maximum number of simultaneously excited electrons and the number of orbitals. To reduce the computational effort only configurations with up to two excited electrons in these virtual orbitals were added to the list of configuration state functions. The contributions of the remaining configurations were treated perturbationally with the multiconfigurational counterpart of the Møller-Plesset method denoted as CASPT2/RASPT2[11]. All virtual orbitals were correlated in the perturbation procedure except for the 24 removed Rydberg orbitals with A''-symmetry. An IPEA shift of 0.0 a.u.[12] and an imaginary shift of 0.2 a.u.[13] were used. Transition dipole moments were calculated at SA-CASSCF/RASSCF level. The RICD approximation was used to speed up the calculation of two-electron integrals[14].

Table 1. Vertical excitation energies (in eV), transition dipole moments (TDM) (in a.u.) out of the ground state (GS), the first (Lb) and the second (La) excited states to higher excited states S_N *calculated at SA-25-RASSCF(0,0|12,10|2,12)//SS-RASPT2/ANO-L[432,21]+Rydberg[8s8p8d] level for adenine monomer in the gas-phase. Labels according to Fig. 1. Region plotted in the spectra in Figs. 3.-5. of the main article is marked. → and => indicate single and double excitations.*

Label	vertical energy	TDM			main conf.	coeff.
		GS → S_N	Lb → S_N	La → S_N		
2A' (L_1^+) L_b	4.90	0.08	-	-	H→L+1 H-1→L	0.63 -0.54
3A' (L_2^+) L_a	5.13	1.35	0.50	-	H→L H-1→L+1	0.79 0.30
4A' (L_1^- + A_2^+)	6.24	1.00	0.50	0.70	H-1→L H→L+1 H-2→L	0.44 0.35 0.35
5A' (A_1^+)	6.37	0.15	0.20	1.20	H→L+2 H-2→L	0.53 0.36
6A' (A_2^+ - L_1^-)	6.37	1.50	0.33	0.90	H-2→L H→L+2 H-1→L	0.48 -0.34 -0.32
7A' (L_2^-)	6.84	1.60	0.44	0.28	H-1→L+1 H→L	0.65 -0.27
8A' (A_1^-)	7.03	1.00	0.55	0.56	H-1→L+2 H-2→L+1	0.59 0.33
9A' (A_2^-)	7.40	1.40	0.74	0.07	H-2→L+1 H-1→L+2	0.47 -0.28
10A' (B_1^+)	7.90	0.70	0.18	0.40	H-3→L	0.46
11A' ©	8.25	1.50	0.33	0.53	H-2→L+2	0.57
12A' (B_2^+)	8.53	0.80	0.72	0.13	H-3→L+1	0.53
13A'	8.76	0.30	0.46	1.15	H=>L	0.47
14A' (B_1^-)	9.12	0.50	0.48	0.26	0→L 0→L+1 H→L+3	0.34 -0.34 0.27
15A' (X)	9.19	0.45	0.16	0.31	0→L	0.48
16A' (X)	9.29	0.70	0.49	0.00	0→L+1	0.52
17A'	9.46	0.40	0.82	0.85	H=>L,L+1	0.58
18A'	10.06	0.45	0.29	0.32	H-3→L+2 H-1,H=>L+1,L+2 0→L+2 H-2→L+2	0.29 -0.25 0.23 0.20
19A'	10.25	0.45	0.19	1.10	H=>L H-1,H=>L+1,L+2 0=>L+2	0.28 -0.23 -0.22
20A' (D^+)	10.28	0.20	0.22	0.35	H=>L+1,L+2 H-3→L+2	0.38 0.33
21A'	10.36	0.30	0.19	0.56	H=>L+1 H-4→L	0.35 -0.35
22A' (X)	10.49	0.15	0.16	0.39	0→L+2	0.45
23A' (X)	10.53	0.45	0.52	0.39	0→L+2	0.26
24A' (E)	10.52	0.30	0.24	0.63	H-4→L H-4→L+2	0.35 -0.29

Table 2. Energy and excited state wavefunction comparison between the reference calculation (Table. 1) and a calculation at SA-25-CASSCF(4,4)/SS-CASPT2/ANO-L[321,2] level (used in the main article for the adenine dimer) for adenine monomer in the gas-phase. Labels according to Fig. 1.

	Energy		
Label	Reference	CAS(4,4)	main conf.
2A' (L_1^+) L_b	4.90	4.33	H→L+1 H-1→L
3A' (L_2^+) L_a	5.13	4.73	H→L H-1→L+1
10A' (B_1^+)	7.94	-	H-3→L
11A' ©	8.19	-	H-2→L+2
12A' (B_2^+)	8.53	-	H-3→L+1
13A'	8.83	8.99	H=>L
14A' (B_1^-)	9.12	-	0→L 0→L+1 H→L+3
15A' (X)	9.19	-	0→L
16A' (X)	9.29	-	0→L+1
17A'	9.54	8.73	H=>L,L+1
18A'	10.04	-	H-3→L+2 H-1,H=>L+1,L+2 0→L+2 H-2→L+2
19A'	10.22	9.46	H=>L H-1,H=>L+1,L+2 0→L+2
20A' (D^+)	10.31	-	H=>L+1,L+2 H-3→L+2
21A'	10.33	9.68	H=>L+1 H-4→L

References

1. A. Nenov, I. Rivalata, S. Mukamel, M. Garavelli *Comput. Theor. Chem.* 2014, **1040**, 295.
2. P. O. Widmark, P. A. Malmqvist, B. O. Roos *Theor. Chim. Acta* 1990, **77 (5)**, 291–306.
3. M. P. Fulscher, L. Serrano Andres, B. O. Roos *J. Am. Chem. Soc.* 1997, **119 (26)**, 6168.
4. L. Serrano Andres, M. P. Fulscher *J. Am. Chem. Soc.* 1996, **118 (48)**, 12190.
5. J. Lorentzon, M. P. Fülscher, B. O. Roos J. Am. Chem. Soc. 1995, **117 (36)**, 9265.
6. L. Serrano-Andres, B. O. Roos *J. Am. Chem. Soc.* 1996, **118 (1)**, 185.
7. B.O. Roos, M.P. Fülscher, Malmqvist, P.-Å., Merchan, L.M. Serrano-Andres *in: Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*, S.R. Langhoff (Ed.), Kluwer Academic Publishers, Dordrecht, The Netherlands, 1995.
8. K. Kaufmann, W. Baumeister, M. Jungen *J. Phys. B-At. Mol. Opt.* 1989, **22 (14)**, 2223.
9. L. Serrano-Andres, M. Merchan, I. Nebotgil, R. Lindh, B. O. Roos *J. Chem. Phys.* 1993, **98 (4)**, 3151.
10. J. Pipek, P. G. Mezey *J. Chem. Phys.* 1989, **90 (9)**, 4916.
11. K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, K. Wolinski *J. Phys. Chem.* 1990, **94 (14)**, 5483.
12. G. Ghigo, B. O. Roos, P. A. Malmqvist Chem. Phys. Lett. 2004, **396 (1–3)**, 142.
13. N. Forsberg, P. A. Malmqvist *Chem. Phys. Lett.* 1997, **274 (1-3)**, 196.
14. F. Aquilante, L. De Vico, N. Ferre, G. Ghigo, P. A. Malmqvist, P. Neogrady, T. B. Pedersen, M. Pitonak, M. Reiher, B. O. Roos, L. Serrano-Andres, M. Urban, V. Veryazov, R. Lindh *J. Comput. Chem.* 2010, **31 (1)**, 224.

3. Cartesian coordinates (Angstroms) of all geometries used in the present work

GS

O	0.567568	-7.892113	-6.917897
C	0.892199	-6.592252	-6.490223
H	0.848925	-5.871551	-7.293304
N	-0.038266	-6.163289	-5.502806
C	-0.442840	-6.807172	-4.373048
H	-0.031304	-7.752965	-4.085471
N	-1.341369	-6.161927	-3.717261
C	-1.544639	-4.994910	-4.440483
C	-2.341124	-3.850426	-4.224758
N	-3.191789	-3.700812	-3.225922
H	-3.719837	-2.852369	-3.160556
H	-3.344520	-4.402575	-2.524642
N	-2.219845	-2.853263	-5.114827
C	-1.406586	-2.987147	-6.153691
H	-1.364069	-2.146097	-6.820005
N	-0.649886	-4.023759	-6.466840
C	-0.742578	-4.992685	-5.547731
O	3.070019	-2.298086	-7.380605
C	3.277320	-1.416800	-6.298991
H	3.058731	-0.393905	-6.559059
N	2.434635	-1.808798	-5.232378
C	2.487592	-3.003619	-4.575349
H	3.162719	-3.768430	-4.893608
N	1.681137	-3.090769	-3.589655
C	1.018820	-1.873756	-3.574342
C	0.012258	-1.326024	-2.731964
N	-0.544732	-2.005450	-1.754003
H	-1.234396	-1.575746	-1.167739
H	-0.277324	-2.955820	-1.558515
N	-0.370962	-0.082655	-2.991960
C	0.135899	0.578434	-4.035407
H	-0.256674	1.567654	-4.183427
N	1.033212	0.161805	-4.896117
C	1.461871	-1.077778	-4.592800
H	1.415063	-8.153432	-7.551588
H	1.922111	-6.687405	-6.146239
H	3.983420	-2.296672	-7.975413
H	4.332633	-1.533163	-6.052266

$L_a(5')^7$

O	0.535393	-7.849690	-6.993989
C	0.876556	-6.554109	-6.567384
H	0.891244	-5.836050	-7.374826
N	-0.080843	-6.095156	-5.625947
C	-0.570994	-6.748303	-4.531286
H	-0.243269	-7.736154	-4.282001
N	-1.434010	-6.056889	-3.857202
C	-1.508823	-4.853989	-4.503772
C	-2.320084	-3.658233	-4.327705
N	-3.186878	-3.578623	-3.358925
H	-3.723191	-2.719325	-3.268761
H	-3.360768	-4.333981	-2.697702
N	-2.154460	-2.629444	-5.124942
C	-1.233408	-2.722636	-6.135706
H	-1.235654	-1.943380	-6.860054
N	-0.510306	-3.919720	-6.509494
C	-0.660548	-4.849183	-5.656689
O	3.136796	-2.279141	-7.312051
C	3.353037	-1.400770	-6.232090
H	3.132201	-0.377076	-6.484005
N	2.518213	-1.802939	-5.161036
C	2.566208	-3.021582	-4.541363
H	3.229463	-3.781759	-4.889872
N	1.763760	-3.120197	-3.555066
C	1.110221	-1.904070	-3.506966
C	0.105067	-1.386568	-2.673009
N	-0.439056	-2.057532	-1.676779
H	-1.171280	-1.638131	-1.136008
H	-0.216955	-3.025567	-1.522983
N	-0.301447	-0.133430	-2.926857
C	0.206135	0.533417	-3.943849
H	-0.191029	1.521473	-4.089354
N	1.127785	0.142323	-4.795589
C	1.553712	-1.087725	-4.514703
H	1.407229	-8.140071	-7.580231
H	1.881622	-6.673603	-6.162832
H	4.053848	-2.306978	-7.900558
H	4.409052	-1.518514	-5.989050

$L_a(5')$[40]

O	0.589219	-7.736863	-7.071186
C	0.955529	-6.439863	-6.670624
H	0.965442	-5.739096	-7.493326
N	0.032057	-5.930942	-5.718678
C	-0.325686	-6.464038	-4.506806
H	0.030170	-7.424886	-4.190800
N	-1.096685	-5.698010	-3.812972
C	-1.240763	-4.561407	-4.558406
C	-2.159182	-3.452540	-4.453436
N	-3.035851	-3.384656	-3.478126
H	-3.663638	-2.593691	-3.441351
H	-3.135217	-4.103886	-2.769394
N	-2.138172	-2.482469	-5.372659
C	-1.019130	-2.559825	-6.216988
H	-0.870498	-1.720769	-6.859240
N	-0.509986	-3.802603	-6.718604
C	-0.534555	-4.679824	-5.785958
O	3.197373	-2.270102	-7.285892
C	3.443288	-1.387629	-6.213323
H	3.200600	-0.366549	-6.460193
N	2.658449	-1.792841	-5.107466
C	2.740931	-2.994150	-4.468910
H	3.425027	-3.742874	-4.804613
N	1.941778	-3.112169	-3.479646
C	1.248840	-1.913638	-3.444956
C	0.211107	-1.415535	-2.612041
N	-0.306101	-2.102922	-1.613234
H	-1.067031	-1.710087	-1.091608
H	-0.051706	-3.060935	-1.443623
N	-0.230075	-0.194290	-2.886453
C	0.260827	0.489978	-3.921160
H	-0.177396	1.457675	-4.077548
N	1.195305	0.122969	-4.762619
C	1.667797	-1.097073	-4.458921
H	1.443001	-8.047006	-7.673653
H	1.968022	-6.571091	-6.288871
H	4.084341	-2.260576	-7.919371
H	4.509578	-1.490637	-6.012040

L_a(3')[11]

O	0.484384	-7.709813	-7.130475
C	0.871234	-6.458767	-6.624369
H	0.914599	-5.695890	-7.387484
N	-0.065682	-6.019981	-5.646899
C	-0.474237	-6.668020	-4.519632
H	-0.084163	-7.627617	-4.250603
N	-1.343736	-6.003720	-3.844567
C	-1.525516	-4.831675	-4.556403
C	-2.292824	-3.664589	-4.324055
N	-3.073382	-3.494109	-3.274119
H	-3.651064	-2.675680	-3.234214
H	-3.280879	-4.235233	-2.627832
N	-2.178192	-2.678049	-5.210246
C	-1.379889	-2.818774	-6.262347
H	-1.339029	-1.974540	-6.925343
N	-0.641384	-3.855765	-6.588275
C	-0.743317	-4.838129	-5.672595
O	2.965748	-2.155347	-7.329711
C	3.180001	-1.249514	-6.268713
H	2.954186	-0.231016	-6.541011
N	2.346886	-1.631220	-5.192094
C	2.389572	-2.841679	-4.552500
H	3.055795	-3.611639	-4.880760
N	1.578965	-2.936148	-3.563199
C	0.914807	-1.743765	-3.524595
C	-0.055307	-1.189496	-2.639379
N	-0.530979	-1.873898	-1.611292
H	-1.264074	-1.463014	-1.045740
H	-0.311904	-2.859235	-1.486774
N	-0.494198	0.063174	-2.817954
C	-0.055399	0.748336	-3.975917
H	-0.248404	1.796384	-3.980008
N	0.965959	0.290901	-4.832221
C	1.369569	-0.889105	-4.566737
H	1.349067	-8.017877	-7.718283
H	1.871518	-6.640903	-6.231485
H	3.878891	-2.180126	-7.924400
H	4.235723	-1.362237	-6.022049

$L_a(3')^{40}$

O	0.505519	-7.677268	-7.180465
C	0.908172	-6.449174	-6.631201
H	0.990741	-5.668121	-7.372459
N	-0.040931	-6.011844	-5.666730
C	-0.442465	-6.649226	-4.529839
H	-0.037440	-7.598703	-4.244065
N	-1.322983	-5.991511	-3.865148
C	-1.525681	-4.836973	-4.597109
C	-2.314692	-3.683133	-4.377121
N	-3.091923	-3.522304	-3.324389
H	-3.682720	-2.714149	-3.276499
H	-3.252930	-4.256265	-2.657431
N	-2.227078	-2.708132	-5.278369
C	-1.421386	-2.844788	-6.325784
H	-1.396223	-2.007269	-6.997852
N	-0.658289	-3.868384	-6.637970
C	-0.742366	-4.843841	-5.712149
O	2.956323	-2.092605	-7.321539
C	3.180303	-1.200173	-6.250499
H	2.952840	-0.178500	-6.510485
N	2.351815	-1.594232	-5.174623
C	2.400309	-2.812307	-4.542551
H	3.099611	-3.562547	-4.849494
N	1.548193	-2.941594	-3.595273
C	0.826857	-1.779883	-3.593307
C	-0.055755	-1.199624	-2.617884
N	-0.424132	-1.872224	-1.548180
H	-1.097090	-1.458387	-0.920889
H	-0.218756	-2.862918	-1.454137
N	-0.463925	0.066336	-2.772301
C	-0.282161	0.552677	-4.075729
H	-0.717785	1.511903	-4.252944
N	0.878139	0.275838	-4.859284
C	1.308598	-0.897850	-4.609348
H	1.380249	-8.000573	-7.744752
H	1.893411	-6.667050	-6.218972
H	3.869837	-2.132391	-7.914842
H	4.238108	-1.314144	-6.013519

Min$_{\mathrm{CT}(3'\rightarrow 5')}$

O	0.586992	-7.885147	-7.060866
C	0.887776	-6.560751	-6.690852
H	0.807312	-5.868998	-7.516183
N	-0.013673	-6.116174	-5.688014
C	-0.491415	-6.797462	-4.605557
H	-0.213656	-7.810650	-4.419129
N	-1.280026	-6.088958	-3.880201
C	-1.332012	-4.848111	-4.510211
C	-1.978741	-3.605574	-4.206800
N	-2.788250	-3.455318	-3.127200
H	-3.406137	-2.669410	-3.155791
H	-3.197329	-4.272398	-2.716256
N	-1.720511	-2.567504	-4.959769
C	-0.815991	-2.666338	-6.026228
H	-0.898860	-1.891880	-6.747400
N	-0.261138	-3.854689	-6.449987
C	-0.535653	-4.849261	-5.627608
O	3.033745	-2.437769	-7.371941
C	3.198497	-1.527958	-6.317179
H	2.960892	-0.513139	-6.602079
N	2.351064	-1.931453	-5.249823
C	2.363316	-3.155453	-4.657064
H	3.015749	-3.928175	-5.013791
N	1.568579	-3.259423	-3.631744
C	0.909032	-2.096132	-3.589630
C	-0.063378	-1.533097	-2.687633
N	-0.533510	-2.169166	-1.666426
H	-1.266135	-1.740050	-1.108897
H	-0.369973	-3.170025	-1.556553
N	-0.406300	-0.243733	-2.922683
C	0.089534	0.409970	-3.926672
H	-0.264586	1.421812	-4.047355
N	1.014009	0.009866	-4.824490
C	1.376127	-1.218999	-4.628855
H	1.433974	-8.151858	-7.692994
H	1.923178	-6.616419	-6.354783
H	3.938533	-2.384529	-7.977436
H	4.252747	-1.619495	-6.055879

Min$_{\mathrm{CT}(5'\rightarrow 3')}$

O	0.467416	-7.626314	-6.904865
C	0.919877	-6.370315	-6.469202
H	0.955288	-5.637033	-7.262422
N	0.037134	-5.860221	-5.467428
C	-0.306500	-6.437913	-4.264994
H	0.129269	-7.370867	-3.957644
N	-1.169658	-5.760840	-3.578889
C	-1.404447	-4.653378	-4.317452
C	-2.278786	-3.510852	-4.140231
N	-3.052004	-3.366634	-3.122281
H	-3.602112	-2.518675	-3.033896
H	-3.072482	-4.023924	-2.347056
N	-2.257245	-2.590980	-5.122841
C	-1.445912	-2.720285	-6.126902
H	-1.396975	-1.883981	-6.807135
N	-0.659385	-3.782397	-6.456098
C	-0.639626	-4.689155	-5.535153
O	2.941617	-2.363628	-7.349828
C	3.076580	-1.451115	-6.267037
H	2.831226	-0.442238	-6.558054
N	2.200347	-1.847555	-5.247664
C	2.282960	-3.023200	-4.530534
H	2.996999	-3.768636	-4.797608
N	1.455101	-3.092001	-3.568277
C	0.728503	-1.906584	-3.613321
C	-0.114037	-1.272048	-2.667982
N	-0.539682	-1.924806	-1.572438
H	-1.093888	-1.418004	-0.914682
H	-0.163711	-2.815940	-1.317565
N	-0.522262	-0.055922	-2.887297
C	-0.171327	0.589630	-4.060514
H	-0.341426	1.640264	-4.083627
N	0.854888	0.122768	-4.951171
C	1.213638	-1.075381	-4.639381
H	1.285786	-7.972969	-7.535895
H	1.931698	-6.577334	-6.120689
H	3.878060	-2.350954	-7.907513
H	4.126173	-1.523445	-5.982044