Supporting Information for: Transient X-Ray Absorption Spectral Fingerprints of the S_1 Dark State in Uracil

Weijie Hua,[*,†,‡] Shaul Mukamel,[*,¶] and Yi Luo[§]

†*Department of Applied Physics, School of Science, Nanjing University of Science and Technology, 210094 Nanjing, China*

‡*Department of Theoretical Chemistry and Biology, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH Royal Institute of Technology, S-106 91 Stockholm, Sweden*

¶*Department of Chemistry and Department of Physics and Astronomy, University of California Irvine, Irvine, California 92697, United States*

§*Hefei National Laboratory for Physical Science at the Microscale, University of Science and Technology of China, 230026 Hefei, China*

E-mail: wjhua@njust.edu.cn; smukamel@uci.edu

This PDF file contains

Supplementary computational details 3

Figures S1-S17 5

Tables S1–S2 22

Supplementary computational details

By using a utility script of the Columbus package,[1–3] linear interpolations on Pulay's natural internal coordinates[4] were performed between four anchor structures i→ii→iii→iv [S_0 and S_2 minima, and S_2/S_1 and S_1/S_0 conical intersections (CoIns), respectively], optimized at the state-averaged (SA) complete-active-space self-consistent field (CASSCF) level[5,6] over 3 states with the 6-31G* basis set, taken from Nachtigallová et al.[7,8] The natural internal coordinates are combinations of bond stretching, bending, rocking, torsions, out-of-plane wagging, etc., and are especially useful to study the deformation of ring compounds. For uracil (12 atoms), there are 30 such internal coordinates (Table S2). About 100 geometries are interpolated between geometries i–iv to guarantee the smoothness of potential energy surfaces (PESs) and spectra. Valence and core excited states were respectively computed at the SA-CASSCF and SA-RASSCF levels with the aug-cc-pVDZ basis set[9] by using the Molpro program[10,11] (RASSCF, restricted-active-space self-consistent field). Here a larger basis set was used since the 6-31G* basis set is not suitable for core hole calculations, as it does not adequately describe the short-range asymptotic behavior of the core hole states.[12] This basis set change has negligible influence of the valence state PESs. For valence-state calculation of interpolated structures, initial guess was read from the adjacent geometry to keep the consistency of the active space and smoothness of the PESs. At each geometry, vertical core excitation was always performed after obtaining the converged valence state wavefunction.[13]

The steady-state XAS spectra were first computed. Calibration of core-excited state energies were done by comparing with gas-phase experiment of uracil by Feyer et al.[14] In the C, N, and O K-edge, energies of the simulated core-excited states were uniformly shifted by -4.0, -5.1, and -3.5 eV to match experiment.[14] All stick spectra were convoluted by Gaussian line shape with fixed half-width-at-half-maximum (HWHM) of 0.3 eV. A broader line width than core hole lifetime (C1s, 0.06 eV; N1s, 0.09 eV; O1s, 0.13 eV[15]) was used in order to represent the vibrionic coupling phenomenologically and to get better agreement with the

steady-state XAS experiment.[14]

The spectra and natural transition orbital (NTO) analysis[16,17] were simulated by using our in-house code MCNOX.[18] The NTOs provide a compact picture of a multi-electron state transition into one or a few pairs of particle to hole orbital transitions. For all analysis reported in this work, there is only one dominant pair, showing single-electron feature of the transitions. All graphical molecular orbitals (MOs) were generated using the Gabedit package[19] at contour threshold of 0.08.

Additional test calculations were performed to justify the optimized S_1 minimum (min S_1, labeled as structure v).[7] Following the same procedure, valence PESs were also computed at interpolated paths connecting i–v and v–ii. As a point test, RASSCF core excitation calculations were carried out only on geometry v and only at the O1s edge. Structures were superimposed to illustrate the structural changes, with graphics generated by VMD.[20]

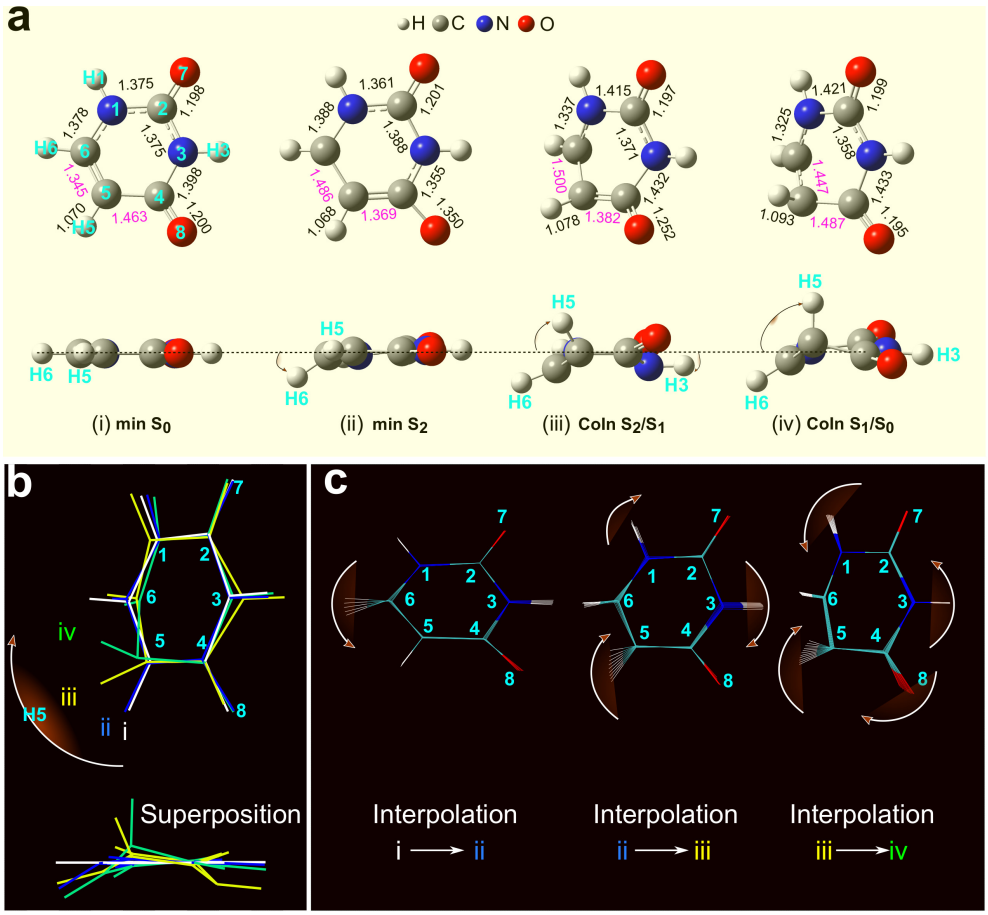

Figure S1. (a) Optimized minima and minimum-energy conical intersections. Top, top view. Selected bond lengths are labeled in Å, where significant structural changes (around C_5) are colored in magenta. Bottom, side view. Major deformations are indicated by arrows. (b) Superposition of the four anchor geometries (top, top view; bottom, side view). The deformation is global while the biggest changes happen in the deformation of the C_5–H_5 bond with respect to the molecular plane. (c) Illustration for the i–ii, ii–iii, iii–iv interpolations. Major deformations are indicated by arrows.

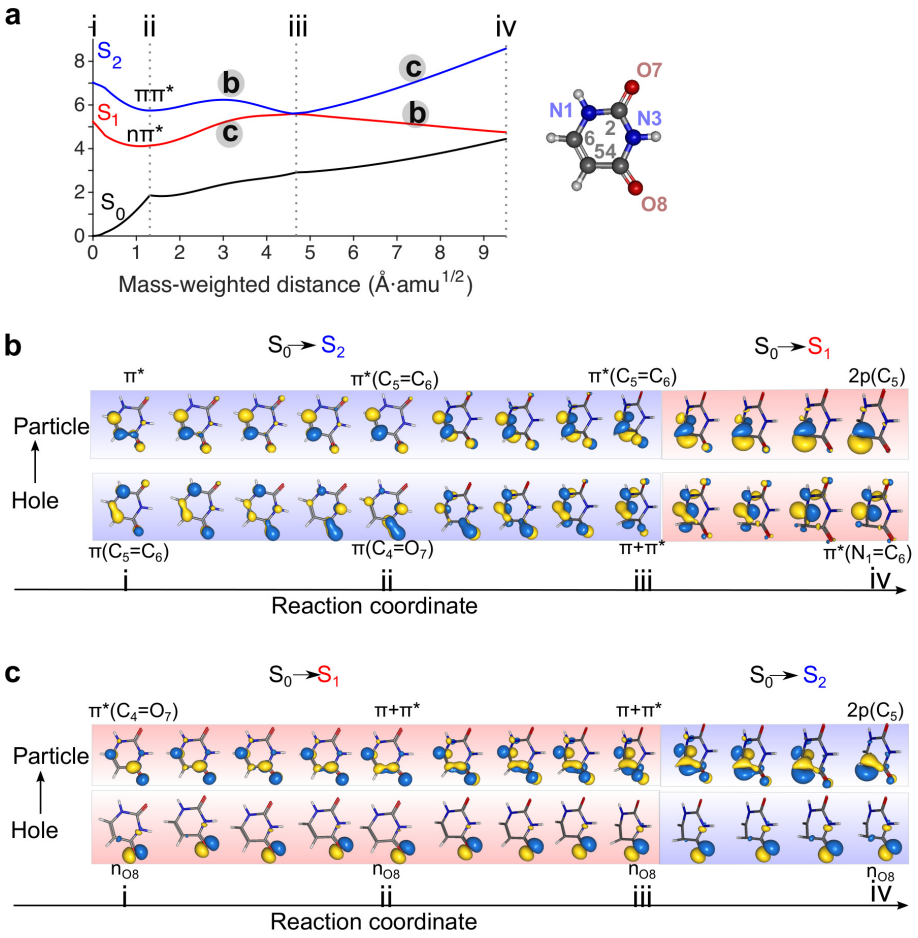

Figure S2. Evolution of valence excitations. (a) Recapture of potential energy surface of S_0-S_2 from Figure 1. (b-c) Evolution of state character as interpreted by NTOs from S_0. The two reaction coordinates are defined in panel a according to state character. Note the state index change near geometry iii. Contour isovalue=0.08 is used.

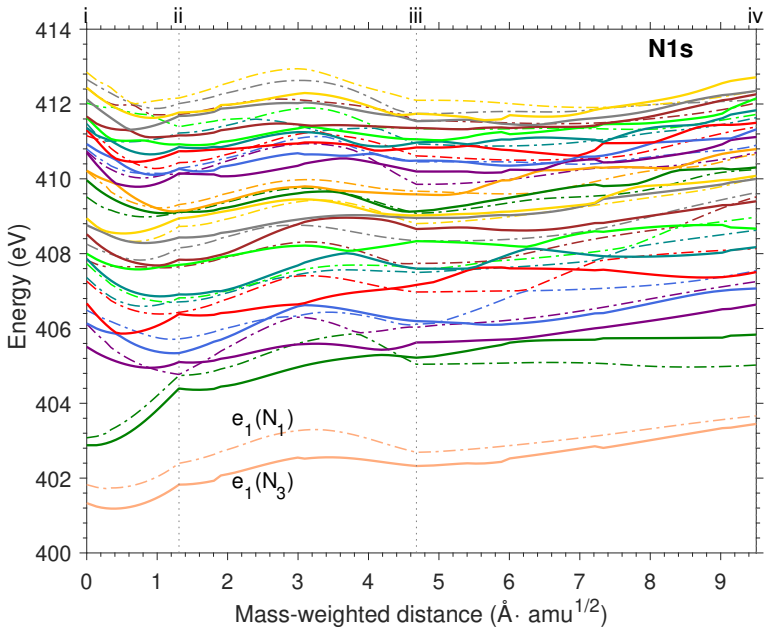

Figure S3. Computed vertical potential energy surfaces for 20 lowest N_1 1s and N_3 1s core excited states along the linear-interpolated internal coordinates. The lowest state of each set is labeled by e_1. All energies are uniformly shifted by -4.45 eV according to calibration to steady-state XAS experiment.[14]

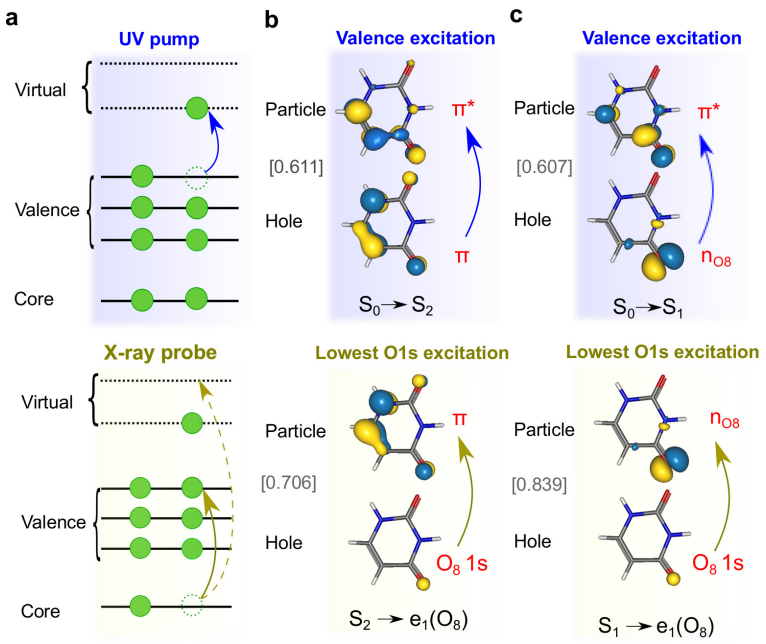

Figure S4. Core excitations from a valence-excited state. (a) Physical picture. Top: The UV pump initiates a valence excitation and leaves a hole. Bottom: The X-ray probe at time τ generates multiple core excitations, amongst the lowest core excitation (e_1) fills the valence hole (golden arrow). Gray dashed arrow denotes an ordinary higher excitation. (b-c) Computed NTOs[16,17] of uracil. Top: NTOs of valence transitions from S_0 to (b) $\pi\pi^*$ state S_2 and (c) $n\pi^*$ state S_1. Bottom: NTOs of core excitations from (b) S_2 and (c) S_1 to the lowest O1s core excitation state from O_8, $e_1(O_8)$. Calculations are on the ground state geometry (i.e., τ=0).

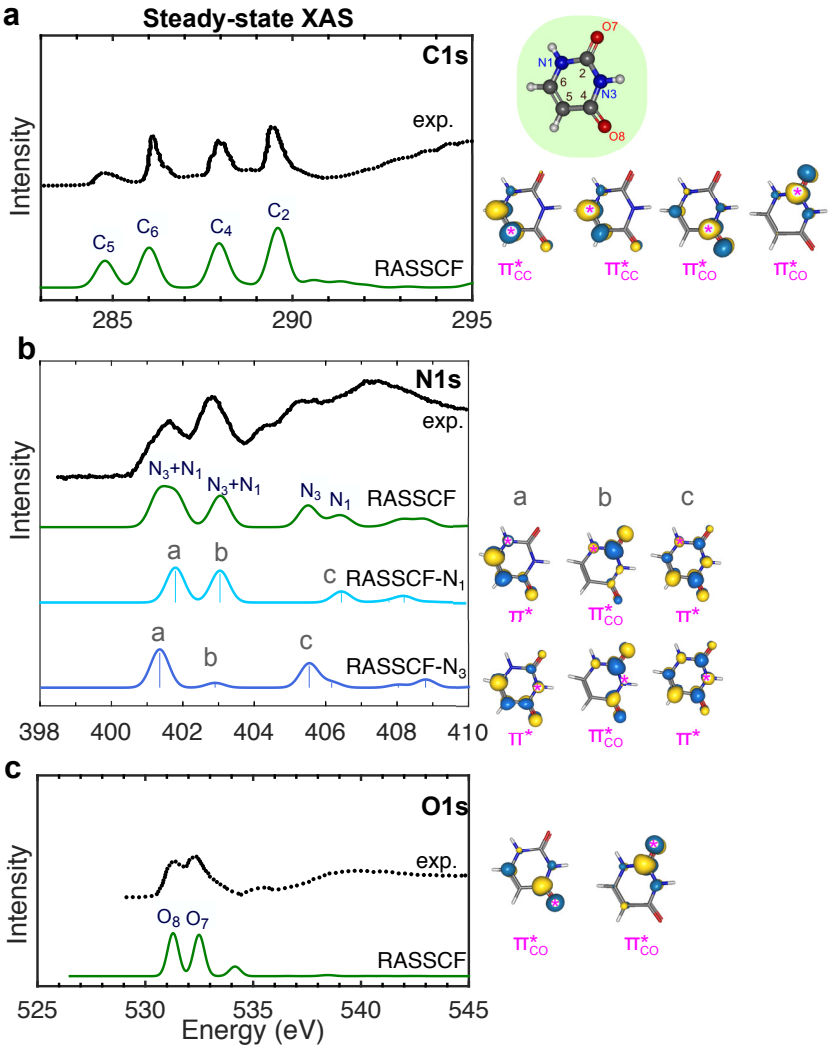

Figure S5. Steady-state XAS spectra. Simulated (a) C1s, (b) N1s, and (c) O1s XAS spectra of uracil at the ground state geometry by RASSCF compared with gas-phase experiment by Fayer et al.[14] Theoretical spectra are uniformly shifted by -4.0, -4.45, and -3.5 eV to match experiment. Each major peak is assigned as transition from 1s orbital to the lowest core excited state (e_1) of the manifold, and is interpreted by natural transition orbitals[16,17] as single-electron "hole" $\rightarrow$ "particle" transition within a NTO pair. The particle orbital is shown in the inset. Corresponding 1s hole orbital is always localized on the core-excited atom (the atom is labeled).

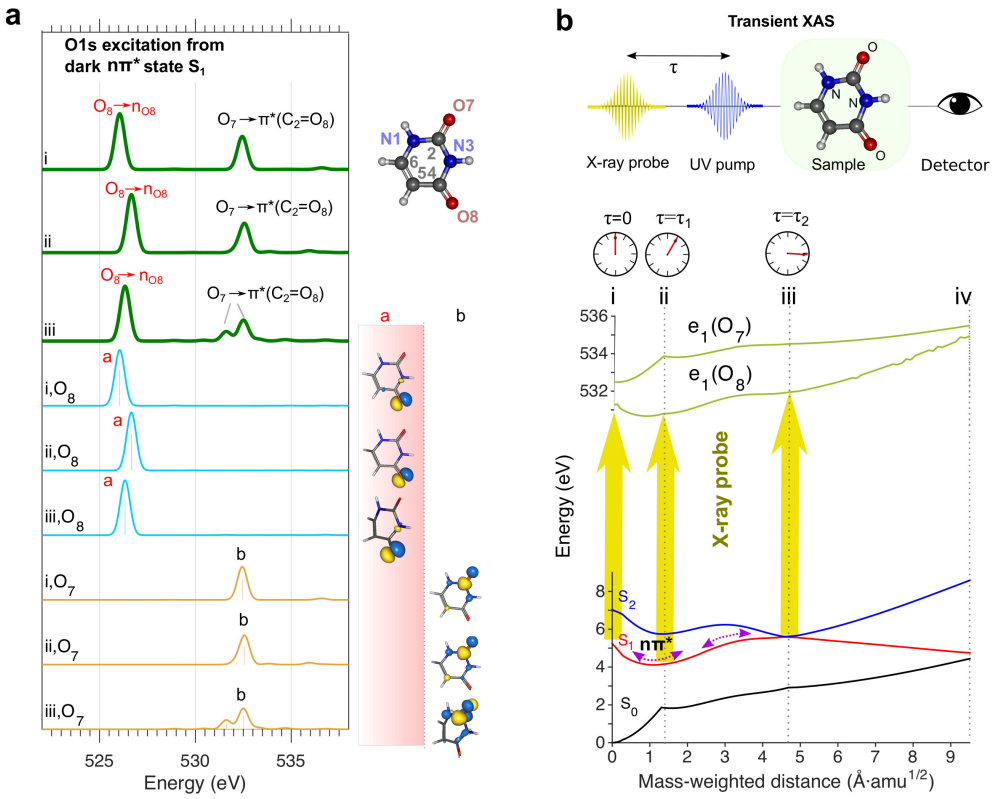

Figure S6. (a) Interpretation of O1s absorptions from dark state S_1 at geometries i, ii, and iii. Top, total spectra (top, thick green lines); bottom, atom-specific contributions. Major peaks are interpreted by NTOs with final orbitals depicted. Transitions to the O_8 lone pair orbitals are especially shaded with red color. (b) Schematic illustration of TXAS and the O1s core excitation from S_1. The two lowest core excited states, $e_1(O_8)$ and $e_1(O_7)$ have large oscillator strengths.

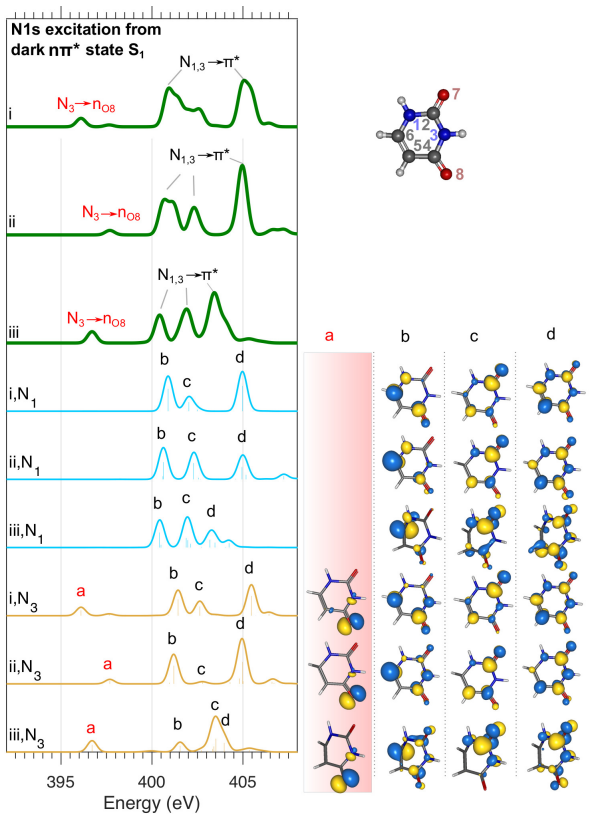

Figure S7. Same as Figure S6a to interpret the N1s absorptions from dark state S_1 at geometries i-iii.

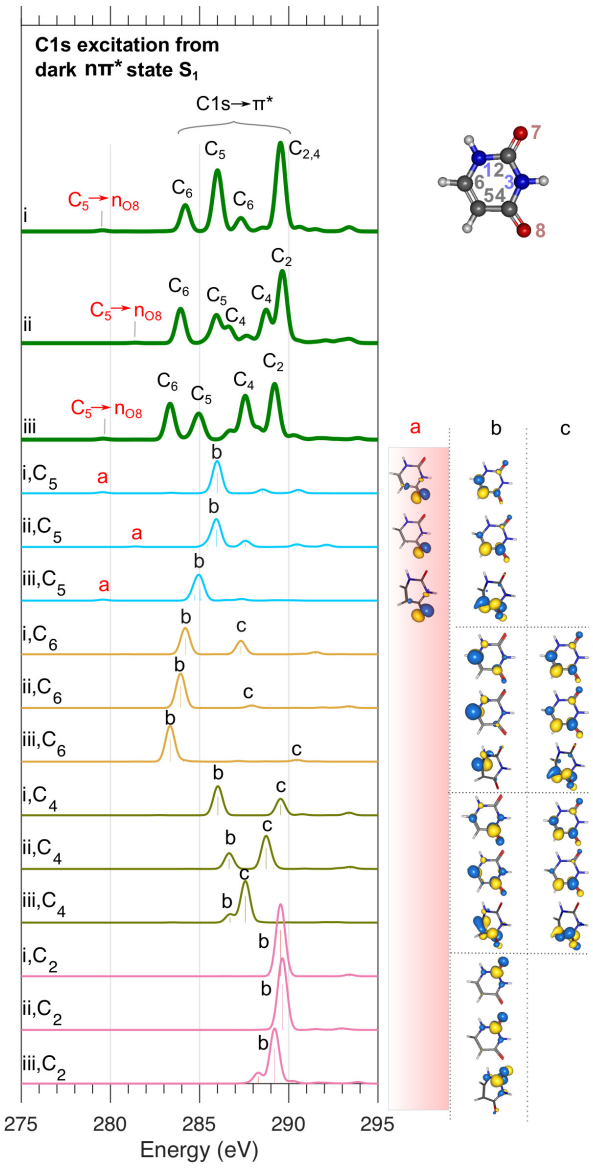

Figure S8. Same as Figure S6a to interpret the C1s absorptions from dark state S_1 at geometries i-iii.

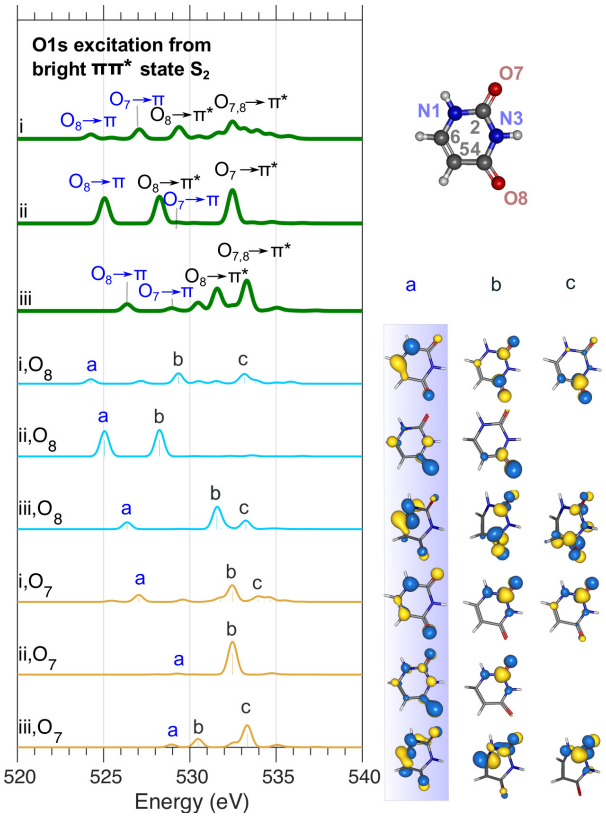

Figure S9. Same as Figure S6a to interpret the O1s absorptions from bright $\pi\pi^*$ state S_2 at geometries i-iii. Transitions filling the valence hole are especially shaded with blue color.

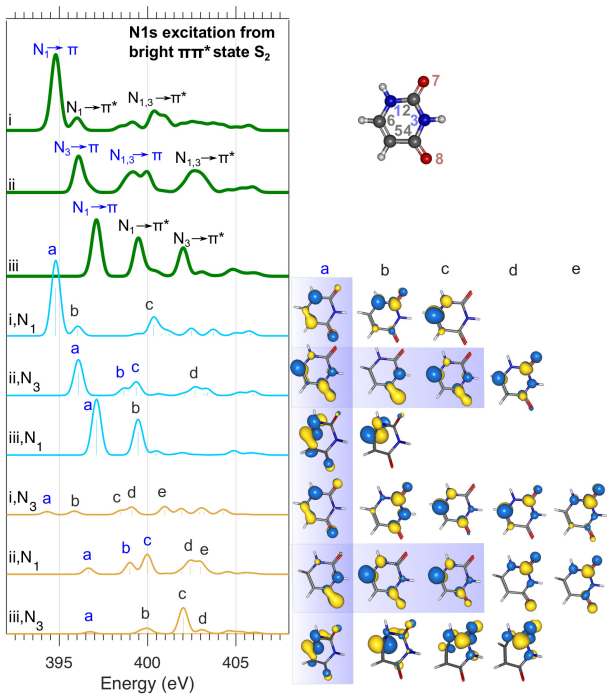

Figure S10. Same as Figure S6a to interpret the N1s absorptions from bright $\pi\pi^*$ state S_2 at geometries i-iii. Transitions filling the valence hole are especially shaded with blue color.

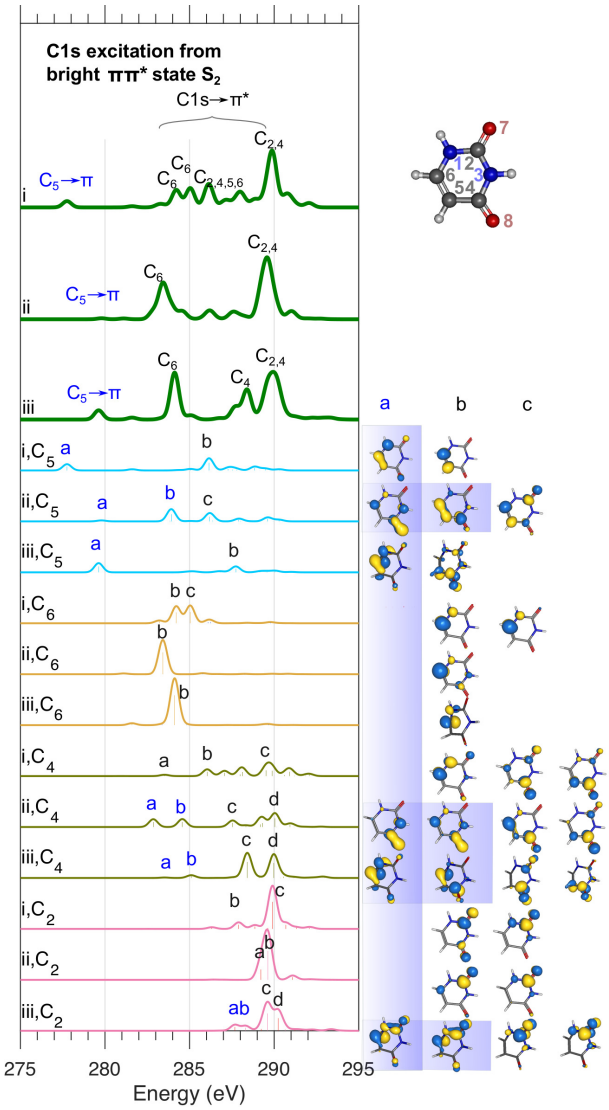

Figure S11. Same as Figure S6a to interpret the C1s absorptions from bright $\pi\pi^*$ state S_2 at geometries i-iii. Transitions filling the valence hole are especially shaded with blue color.

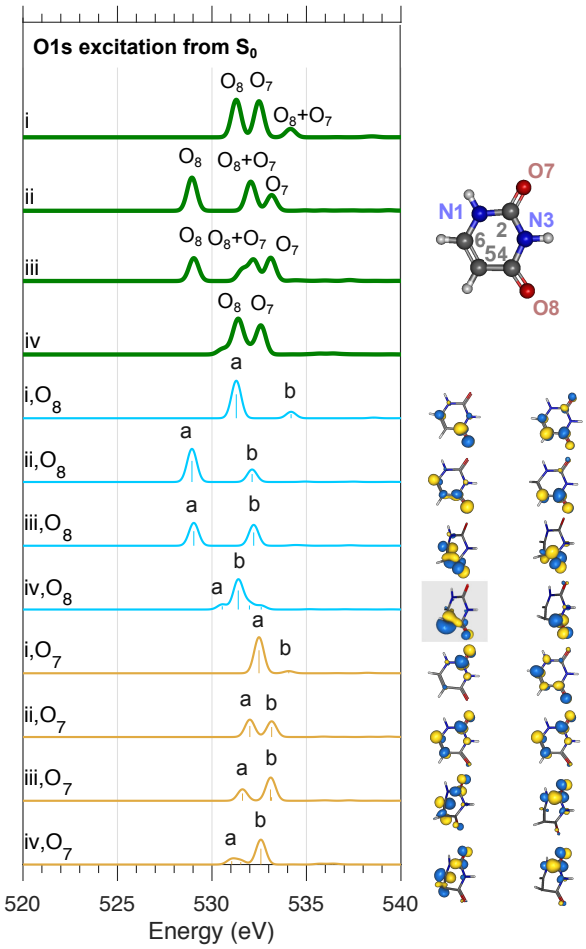

Figure S12. Same as Figure S6a to interpret the O1s absorptions from S_0 state at geometries i-iv. At geometry iv, transition to the 2p (C_5) orbital is especially shaded with gray color.

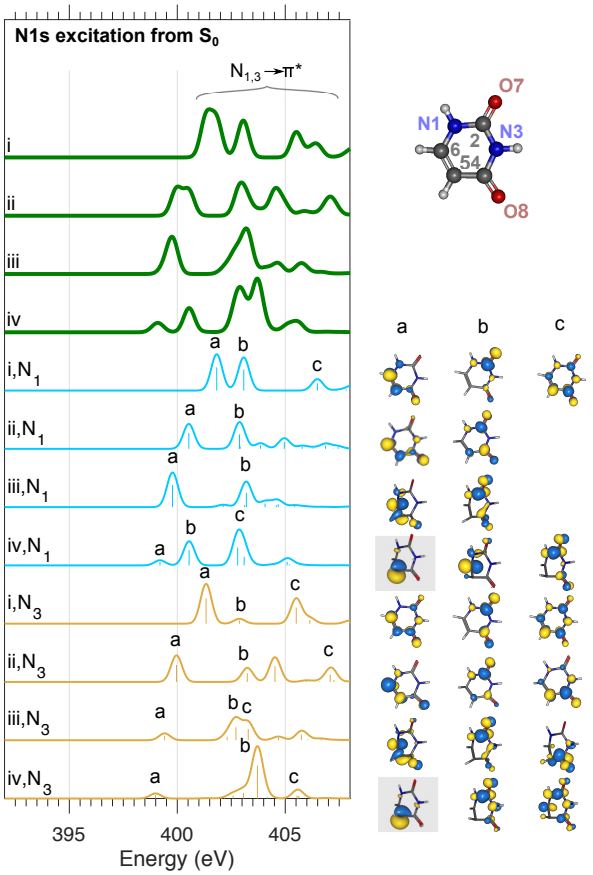

Figure S13. Same as Figure S6a to interpret the N1s absorptions from S_0 state at geometries i-iv. At geometry iv, transitions to the 2p (C_5) orbitals are especially shaded with gray color.

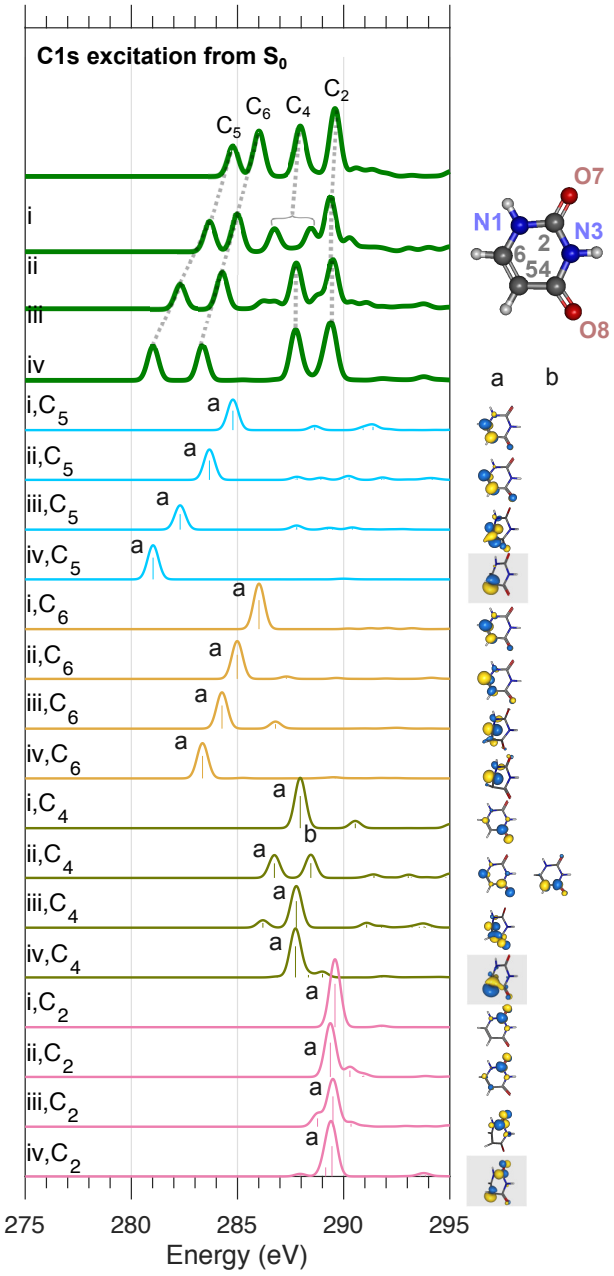

Figure S14. Same as Figure S6a to interpret the C1s absorptions from S_0 state at geometries i-iv. At geometry iv, transitions to the 2p (C_5) orbital are especially shaded with gray color.

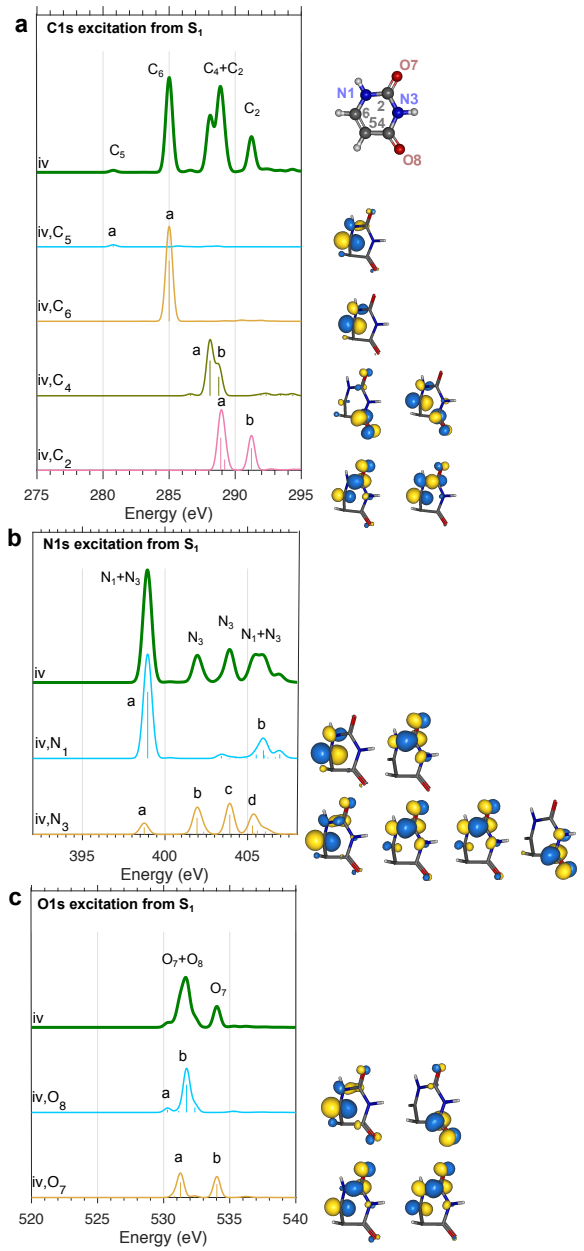

Figure S15. Same as Figure S6a to interpret the (a) C1s, (b) N1s and (c) O1s absorptions from S_1 state at geometry iv.

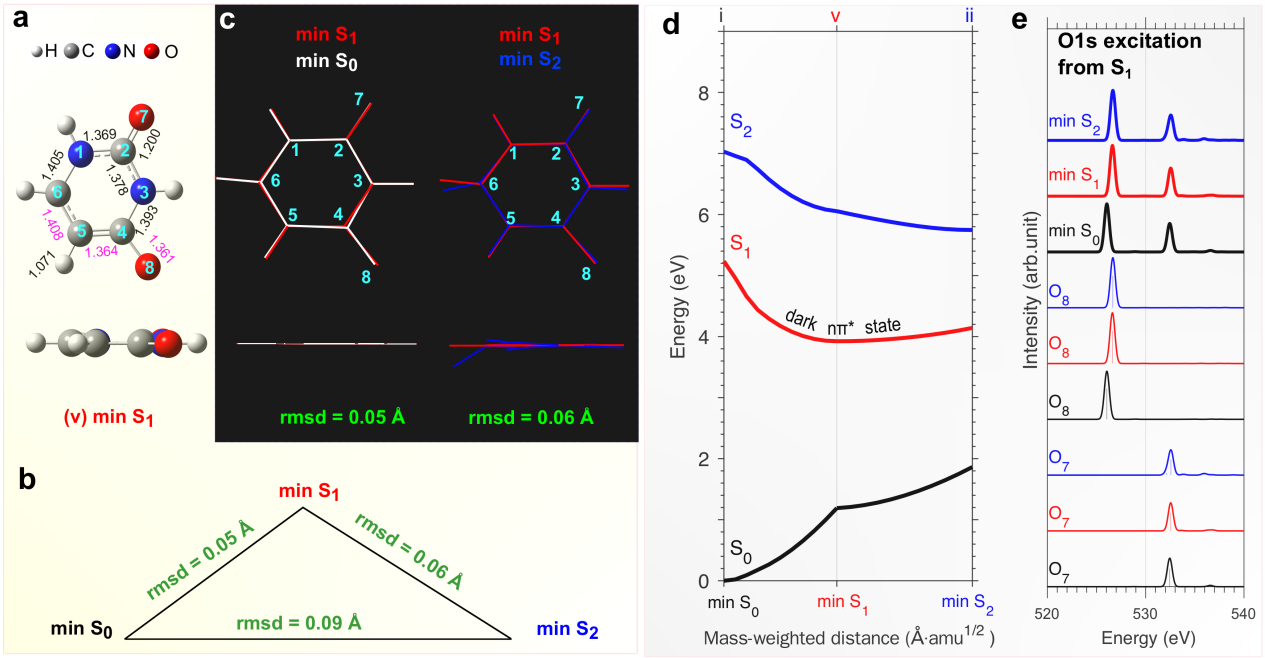

Figure S16. Additional justification for the S_1 minima (min S_1, denoted as geometry v). (a) Optimized structure. Top, top view; bottom, side view. Selected bond lengths are labeled in Å. (b) Schematic illustration of the structural difference of min S_0, min S_1 and min S_2 in terms of the root-mean-squared distances (rmsd). (c) Superimposed structures of min S_1 with min S_0 (left) or min S_2 (right). Top, top view; bottom, side view. (d) CASSCF valence PESs along another linear interpolated internal coordinates (LIIC) connecting min S_0 and min S_2 via min S_1. (f) Comparison of simulated O1s absorptions from dark state S_1 at the three minima. Top, total spectra (thick lines); bottom, atom-specific contributions from O_8 and O_7 (thin lines). The three total spectra are similar, which shows that excited state absorptions from the dark $n\pi^*$ state S_1 is less sensitive to structures.

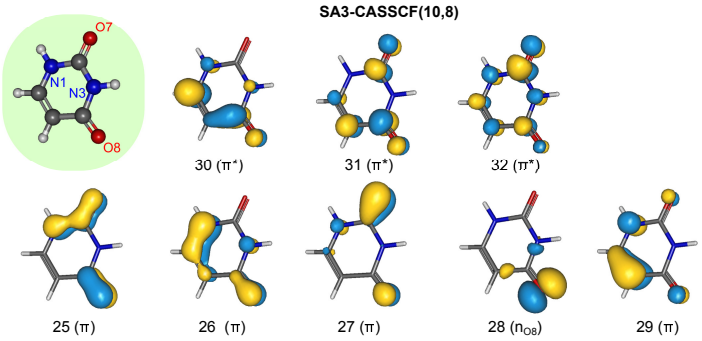

Figure S17. CASSCF(10, 8) active space. Active space for uracil at the optimized ground state (min S_0) at the CASSCF level (10 electrons in 8 orbitals) averaged over 3 lowest states S_0–S_2.

Table S1. Cartesian coordinates (aligned, in Å) of anchor structures used.

i	(min S_0)			ii	(min S_2)		
C	-0.880497	-0.001460	-0.920817	C	-0.871166	-0.005059	-0.914306
C	0.228394	-0.001460	1.305006	C	0.282036	-0.000618	1.232143
C	1.514814	0.004984	0.608334	C	1.519029	-0.078069	0.650061
C	1.535060	0.007608	-0.736754	C	1.587378	0.027097	-0.830498
N	-0.880497	-0.001460	0.453828	N	-0.840575	0.018865	0.473339
N	0.375309	0.005359	-1.481593	N	0.363057	0.068080	-1.482237
O	-1.882652	-0.006002	-1.577641	O	-1.904816	-0.051370	-1.524876
O	0.091176	-0.006405	2.497617	O	-0.033837	0.022239	2.544352
H	-1.779645	-0.009010	0.885316	H	-1.722118	0.056455	0.939648
H	2.411511	0.006173	1.192763	H	2.406213	-0.139102	1.242198
H	2.445087	0.011685	-1.302851	H	2.366382	0.592757	-1.312882
H	0.404548	0.005077	-2.475286	H	0.344754	0.126085	-2.477125
iii	(CoIn S_2/S_1)			iv	(CoIn S_1/S_0)		
C	-0.866570	-0.002161	-0.856336	C	-0.876830	-0.032274	-0.890777
C	0.281459	0.049116	1.270136	C	0.260095	-0.006308	1.325951
C	1.497393	-0.126604	0.638339	C	1.483556	-0.381581	0.569033
C	1.496880	0.228724	-0.818814	C	1.480099	0.244513	-0.735659
N	-0.789089	0.482805	0.423286	N	-0.881208	0.073641	0.463095
N	0.398806	-0.176331	-1.464873	N	0.402979	0.101935	-1.493465
O	-1.867453	-0.200016	-1.482408	O	-1.846732	-0.119441	-1.590055
O	-0.050320	-0.254368	2.438650	O	0.079147	0.120681	2.499856
H	-1.676263	0.580480	0.870603	H	-1.781977	0.163051	0.885102
H	2.056878	-1.009866	0.901164	H	1.359404	-1.452714	0.387921
H	2.358797	0.505085	-1.391471	H	2.295373	0.801784	-1.167509
H	0.359028	-0.319060	-2.455497	H	0.373389	0.352904	-2.462873
v	(min S_1)						
C	-0.881535	-0.001501	-0.918584				
C	0.276456	-0.001371	1.219919				
C	1.511011	0.005005	0.638855				
C	1.573661	0.008277	-0.767696				
N	-0.890662	-0.004492	0.459251				
N	0.364974	0.004794	-1.484217				
O	-1.896018	-0.004340	-1.560426				
O	0.043220	-0.005207	2.560877				
H	-1.785967	-0.009166	0.891655				
H	2.400459	0.007324	1.235752				
H	2.476629	0.013264	-1.336159				
H	0.370751	0.006975	-2.477739				

Table S2. Sample Columbus output (file icoordtyp) for 30 natural internal coordinates of uracil at geometry i. The atomic orders are consistent with Table S1.

1	C3–C2	STRE	sc=.9200
2	C4=C3	STRE	sc=.8600
3	N5–C1	STRE	sc=.9000
4	N5–C2	STRE	sc=.9000
5	N6–C1	STRE	sc=.9000
6	N6–C4	STRE	sc=.9000
7	O7=C1	STRE	sc=.8260
8	O8=C2	STRE	sc=.8260
9	H9–N5	STRE	sc=.9000
10	H10–C3	STRE	sc=.9000
11	H11–C4	STRE	sc=.9000
12	H12–N6	STRE	sc=.9000
13	6-membered ring	BEND	sc=.7900
14	6-membered ring	BEND	sc=.7900
15	6-membered ring	BEND	sc=.7900
16	6-membered ring	TORS	sc=.9600
17	6-membered ring	TORS	sc=.9600
18	6-membered ring	TORS	sc=.9600
19	ring sXY C1–O7	ROCK	sc=.8000
20	ring sXY C1–O7	OUT	sc=.7000
21	ring sXY C2–O8	ROCK	sc=.8000
22	ring sXY C2–O8	OUT	sc=.7000
23	ring sXY C3–H10	ROCK	sc=.8000
24	ring sXY C3–H10	OUT	sc=.7000
25	ring sXY C4–H11	ROCK	sc=.8000
26	ring sXY C4–H11	OUT	sc=.7000
27	ring sXY N5–H9	ROCK	sc=.8000
28	ring sXY N5–H9	OUT	sc=.7000
29	ring sXY N6–H12	ROCK	sc=.8000
30	ring sXY N6–H12	OUT	sc=.7000

Supplementary References and Notes

(1) Lischka, H.; Müller, T.; Szalay, P. G.; Shavitt, I.; Pitzer, R. M.; Shepard, R. Columbus–a program system for advanced multireference theory calculations. *WIREs Comput. Mol. Sci.* **2011**, *1*, 191.

(2) Lischka, H.; Shepard, R.; Pitzer, R. M.; Shavitt, I.; Dallos, M.; Müller, T.; Szalay, P. G.; Seth, M.; Kedziora, G. S.; Yabushita, S. et al. High-level multireference methods in the quantum-chemistry program system COLUMBUS: Analytic MR-CISD and MR-AQCC gradients and MR-AQCC-LRT for excited states, GUGA spin–orbit CI and parallel CI density. *Phys. Chem. Chem. Phys.* **2001**, *3*, 664.

(3) Lischka, H.; Shepard, R.; Shavitt, I.; Pitzer, R. M.; Dallos, M.; Müller, T.; Szalay, P. G.; Brown, F. B.; Ahlrichs, R.; Böhm, H. J. et al. COLUMBUS, an ab initio electronic structure program, release 7.0. 2015.

(4) Pulay, P.; Fogarasi, G.; Pang, F.; Boggs, J. E. Systematic ab initio gradient calculation of molecular geometries, force constants, and dipole moment derivatives. *J. Am. Chem. Soc.* **1979**, *101*, 2550.

(5) Werner, H.; Knowles, P. J. A second order multiconfiguration SCF procedure with optimum convergence. *J. Chem. Phys.* **1985**, *82*, 5053–5063.

(6) Knowles, P. J.; Werner, H.-J. An efficient second-order MC SCF method for long configuration expansions. *Chem. Phys. Lett.* **1985**, *115*, 259–267.

(7) Nachtigallová, D.; Aquino, A. J. A.; Szymczak, J. J.; Barbatti, M.; Hobza, P.; Lischka, H. Nonadiabatic Dynamics of Uracil: Population Split among Different Decay Mechanisms. *J. Phys. Chem. A* **2011**, *115*, 5247–5255.

(8) A mirror image of CoIn S_1/S_0 coordinates was used, which well reproduced the valence PESs of ref. 7 (otherwise the interpolated PESs will have additional barriers). A geom-

etry and its mirror image have the same electronic structure and both geometries have been reported earlier (see, e.g., refs. 21,22 for the geometry used in this work).

(9) Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

(10) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M.; Celani, P.; Korona, T.; Lindh, R.; Mitrushenkov, A.; Rauhut, G. et al. MOLPRO, version 2012.1, a package of ab initio programs. 2012.

(11) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. Molpro: a general-purpose quantum chemistry program package. *WIREs Comput. Mol. Sci.* **2012**, *2*, 242–253.

(12) Hua, W.; Oesterling, S.; Biggs, J. D.; Zhang, Y.; Ando, H.; de Vivie-Riedle, R.; Fingerhut, B. P.; Mukamel, S. Monitoring conical intersections in the ring opening of furan by attosecond stimulated X-ray Raman spectroscopy. *Struct. Dyn.* **2016**, *3*, 023601.

(13) Zhang, S. B.; Kimberg, V.; Rohringer, N. Nonlinear resonant Auger spectroscopy in CO using an x-ray pump-control scheme. *Phys. Rev. A* **2016**, *94*, 063413.

(14) Feyer, V.; Plekan, O.; Richter, R.; Coreno, M.; de Simone, M.; Prince, K. C.; Trofimov, A. B.; Zaytseva, I. L.; Schirmer, J. Tautomerism in Cytosine and Uracil: A Theoretical and Experimental X-ray Absorption and Resonant Auger Study. *J. Phys. Chem. A* **2010**, *114*, 10270–10276.

(15) Zschornack, G. H. *Handbook of X-Ray Data*; Springer: New York, 2007.

(16) Martin, R. Natural transition orbitals. *J. Chem. Phys.* **2003**, *118*, 4775.

(17) Malmqvist, P. Å.; Veryazov, V. The binatural orbitals of electronic transitions. *Mol. Phys.* **2012**, *110*, 2455–2464.

(18) Hua, W.; Luo, Y.; Mukamel, S. Multi-Configurational methods for ultrafast NOlinear X-ray spectra (MCNOX), version 1.0. 2019.

(19) Allouche, A.-R. Gabedit–A graphical user interface for computational chemistry softwares. *J. Comput. Chem.* **2011**, *32*, 174–182.

(20) Humphrey, W.; Dalke, A.; Schulten, K. VMD – Visual Molecular Dynamics. *J. Mol. Graph.* **1996**, *14*, 33–38.

(21) Matsika, S. Radiationless Decay of Excited States of Uracil through Conical Intersections. *J. Phys. Chem. A* **2004**, *108*, 7584–7590.

(22) Merchán, M.; González-Luque, R.; Climent, T.; Serrano-Andrés, L.; Rodríguez, E.; Reguero, M.; Peláez, D. Unified Model for the Ultrafast Decay of Pyrimidine Nucleobases. *J. Phys. Chem. B* **2006**, *110*, 26471–26476.