Supporting Information

A DFT/MRCI Hamiltonian Parameterized Using Only *Ab Initio* Data: II. Core-Excited States

Teagan Shane Costain,[†] Simon P. Neville,[‡] Jibrael B. Rolston,[†] and Michael S. Schuurman[*,‡,†]

[†]*Department of Chemistry and Biomolecular Sciences, University of Ottawa, Ottawa, Canada*

[‡]*National Research Council Canada, 100 Sussex Dr., Ottawa, Canada, K1A 0R6*

E-mail: michael.schuurman@uottawa.ca

Contents

S1 Reference Data **S3**

S2 Parameter Sets **S16**

S3 Total CPU Times **S24**

S4 Simulation of NEXAFS and X-ray Photoelectron Spectra **S25**

S4.1 Ethyl Trifluoroacetate XPS S26

S4.2 Cytosine NEXAFS S27

S4.3 Thymine NEXAFS S29

S5 Molecular Geometries **S31**

S5.1 Training & Validation Data Sets S31

S5.2 Representative Applications S33

References **S40**

S1 Reference Data

The reference data is computed at the CVS-EOM-CCSDT level of theory with the aug-cc-pCVTZ basis set on non-hydrogen atoms, and the aug-cc-pVTZ basis set on hydrogen atoms. We denote this level of theory as simply aCTZ/CCSDT hereafter. These basis sets are likewise employed in the corresponding DFT/MRCI calculations. The training set consists of 9 molecules, and the validation set consists of 8 molecules. These are detailed in Table S1. The sets is chosen include 1s $\rightarrow \pi^*$ and σ^* transitions, as well as transitions from the 1s core orbital to Rydberg s- and p-type orbitals. This reference data includes singlet and triplet multiplicity excitation energies, as well as the ionization potential corresponding to the doublet core-ionized states for each K-edge. In Tables S2 and S3, the computed results for DFT/MRCI with the CVS-QE12 Hamiltonian are presented alongside the reference values. Table S2 contains the training set data; likewise, Table S3 contains the validation set data.

Table S1: Summary of Contents of Training (top) and Validation (bottom) Sets of Computed *ab initio* Reference Data.

Molecule	Formula	Transitions
Acetylene	C_2H_2	π^*, Ry
Ethylene	C_2H_4	π^*, Ry
Methane	CH_4	Ry
Water	H_2O	Ry
Ammonia	NH_3	Ry
Hydrogen Fluoride	HF	σ^*, Ry
Fluoroamine	NH_2F	σ^*, Ry
Azomethine	CH_2NH	π^*, Ry
Hypofluorous Acid	HOF	σ^*, Ry
Carbon Monoxide	CO	π^*, Ry
Dinitrogen	N_2	π^*, Ry
Formaldehyde	H_2CO	π^*, Ry
Hydrogen Cyanide	HCN	π^*, Ry
Hydrogen Isocyanide	HNC	π^*, Ry
Nitroxyl	HNO	π^*, Ry
Methyl Fluoride	CH_3F	σ^*, Ry
Ketene	C_2H_2O	π^*

Table S2: Training Set Reference ~~Data~~ Excitation and Ionization Energies Computed at the aCTZ/CCSDT Level of Theory and aCTZ/DFT/MRCI [CVS-QE12] Results. Energy in eV.

Molecule	State	Assignment	CCSDT	DFT/MRCI	ΔE
C_2H_2					
	$1\ ^1B_u$	C2 1s $\rightarrow \pi_1$*	286.06	285.28	-0.78
	$2\ ^1B_u$	C2 1s $\rightarrow \pi_2$*	286.06	285.28	-0.78
	$3\ ^1B_g$	C1 1s $\rightarrow \pi_1$*	286.11	285.33	-0.78
	$4\ ^1B_g$	C1 1s $\rightarrow \pi_2$*	286.11	285.33	-0.78
	$1\ ^3B_u$	C2 1s $\rightarrow \pi_1$*	285.52	285.00	-0.52
	$2\ ^3B_u$	C2 1s $\rightarrow \pi_2$*	285.52	285.00	-0.52
	$3\ ^3B_g$	C1 1s $\rightarrow \pi_1$*	285.61	285.07	-0.54
	$4\ ^3B_g$	C1 1s $\rightarrow \pi_2$*	285.61	285.07	-0.54
	$1\ ^2B_u$	C2 1s $\rightarrow \infty$	291.23	290.84	-0.39
	$2\ ^2A_g$	C1 1s $\rightarrow \infty$	291.31	290.91	-0.40
C_2H_4					
	$1\ ^1B_u$	C2 1s $\rightarrow \pi$*	285.06	284.82	-0.24
	$2\ ^1B_g$	C1 1s $\rightarrow \pi$*	285.07	284.83	-0.24
	$3\ ^1B_u$	C2 1s $\rightarrow$ 3s	287.36	286.77	-0.59
	$4\ ^1A_g$	C1 1s $\rightarrow$ 3s	287.40	286.80	-0.60
	$1\ ^3B_u$	C2 1s $\rightarrow \pi$*	284.62	284.58	-0.04
	$2\ ^3B_g$	C1 1s $\rightarrow \pi$*	284.67	284.61	-0.06
	$3\ ^3B_u$	C2 1s $\rightarrow$ 3s	287.12	286.67	-0.45
	$4\ ^3A_g$	C1 1s $\rightarrow$ 3s	287.15	286.70	-0.45
	$1\ ^2B_u$	C2 1s $\rightarrow \infty$	290.79	290.53	-0.26
	$2\ ^2A_g$	C1 1s $\rightarrow \infty$	290.84	290.56	-0.28
CH_4					
	$1\ ^1A_1$	C 1s $\rightarrow$ Ry 3s	286.85	286.64	-0.21
	$2\ ^1A_1$	C 1s $\rightarrow$ Ry 3p	288.18	287.92	-0.26
	$3\ ^1B_2$	C 1s $\rightarrow$ Ry 3p	288.18	287.92	-0.26
	$4\ ^1B_1$	C 1s $\rightarrow$ Ry 3p	288.18	287.92	-0.26
	$1\ ^3A_1$	C 1s $\rightarrow$ Ry 3s	286.47	286.31	-0.16
	$2\ ^3A_1$	C 1s $\rightarrow$ Ry 3p	287.97	287.79	-0.18
	$3\ ^3B_2$	C 1s $\rightarrow$ Ry 3p	287.97	287.79	-0.18
	$4\ ^3B_1$	C 1s $\rightarrow$ Ry 3p	287.97	287.79	-0.18
	$1\ ^2A_1$	C 1s $\rightarrow \infty$	290.84	290.81	-0.03
H_2O					
	$1\ ^1A_1$	O 1s $\rightarrow$ Ry 3s	534.12	535.27	1.15
	$2\ ^1B_1$	O 1s $\rightarrow$ Ry 3p	535.95	536.85	0.90
	$3\ ^1A_1$	O 1s $\rightarrow$ Ry 3p	537.65	538.66	1.01
	$1\ ^3A_1$	O 1s $\rightarrow$ Ry 3s	533.65	534.79	1.14
	$2\ ^3B_1$	O 1s $\rightarrow$ Ry 3p	535.57	536.48	0.91
	$3\ ^3A_1$	O 1s $\rightarrow$ Ry 3p	537.48	538.41	0.93
	$1\ ^2A_1$	O 1s $\rightarrow \infty$	539.78	539.96	0.18

Molecule	State	Assignment	CCSDT	DFT/MRCI	ΔE
NH_3					
	$1\,^1A'$	N 1s → Ry 3s	400.76	401.27	0.51
	$2\,^1A'$	N 1s → Ry 3p	402.44	402.83	0.39
	$3\,^1A''$	N 1s → Ry 3p	402.44	402.83	0.39
	$4\,^1A'$	N 1s → Ry 3p	403.57	404.04	0.47
	$1\,^3A'$	N 1s → Ry 3s	400.34	400.89	0.55
	$2\,^3A'$	N 1s → Ry 3p	402.15	402.63	0.48
	$3\,^3A''$	N 1s → Ry 3p	402.15	402.63	0.48
	$4\,^3A'$	N 1s → Ry 3p	403.43	403.92	0.49
	$1\,^2A'$	N 1s → ∞	405.58	405.63	0.05
HF					
	$1\,^1A_1$	F 1s → σ*	687.50	689.11	1.60
	$2\,^1A_1$	F 1s → Ry 3s	691.19	692.69	1.50
	$3\,^1B_1$	F 1s → Ry 3p	692.08	693.86	1.78
	$4\,^1B_2$	F 1s → Ry 3p	692.08	693.86	1.78
	$1\,^3A_1$	F 1s → σ*	686.99	688.61	1.62
	$2\,^3A_1$	F 1s → Ry 3s	691.03	692.44	1.41
	$3\,^3B_1$	F 1s → Ry 3p	691.92	693.61	1.69
	$4\,^3B_2$	F 1s → Ry 3p	691.92	693.61	1.69
	$1\,^2A_1$	F 1s → ∞	694.00	694.28	0.28
NH_2F					
	$1\,^1A'$	N 1s → σ*	402.15	401.96	-0.19
	$2\,^1A'$	N 1s → Ry 3s	403.51	403.81	0.30
	$3\,^1A''$	N 1s → Ry 3p	404.81	405.00	0.19
	$1\,^3A'$	N 1s → σ*	400.78	401.16	0.38
	$2\,^3A'$	N 1s → Ry 3s	403.06	403.47	0.41
	$3\,^3A''$	N 1s → Ry 3p	404.69	404.77	0.08
	$1\,^2A'$	N 1s → ∞	408.24	408.44	0.20
	$1\,^1A'$	F 1s → σ*	686.64	687.26	0.62
	$2\,^1A'$	F 1s → Ry 3s	688.92	689.99	1.07
	$3\,^1A''$	F 1s → Ry 3p	690.24	691.10	0.86
	$1\,^3A'$	F 1s → σ*	685.68	686.53	0.85
	$2\,^3A'$	F 1s → Ry 3s	688.89	689.79	0.90
	$3\,^3A''$	F 1s → Ry 3p	690.24	690.97	0.73
	$1\,^2A'$	F 1s → ∞	693.22	693.69	0.47
CH_2NH					
	$1\,^1A''$	C 1s → π*	285.65	285.33	-0.32
	$2\,^1A'$	C 1s → Ry 3s	288.75	288.29	-0.46
	$3\,^1A'$	C 1s → Ry $3p_z$	289.51	289.08	-0.43

Molecule	State	Assignment	CCSDT	DFT/MRCI	ΔE
CH_2NH					
	$1\,^3A''$	C 1s $\rightarrow \pi^*$	285.04	285.00	-0.04
	$2\,^3A'$	C 1s $\rightarrow$ Ry 3s	288.46	288.09	-0.37
	$3\,^3A'$	C 1s $\rightarrow$ Ry $3p_z$	289.33	288.94	-0.39
	$1\,^2A'$	C 1s $\rightarrow \infty$	292.75	292.66	-0.09
	$1\,^1A''$	N 1s $\rightarrow \pi^*$	398.57	398.66	0.09
	$2\,^1A'$	N 1s $\rightarrow$ Ry 3s	401.50	401.32	-0.18
	$3\,^1A'$	N 1s $\rightarrow$ Ry $3p_z$	402.48	402.22	-0.26
	$1\,^3A''$	N 1s $\rightarrow \pi^*$	398.12	398.37	0.25
	$2\,^3A'$	N 1s $\rightarrow$ Ry 3s	401.23	401.12	-0.11
	$3\,^3A'$	N 1s $\rightarrow$ Ry $3p_z$	402.40	402.13	-0.27
	$1\,^2A'$	N 1s $\rightarrow \infty$	405.51	405.43	-0.08
HOF					
	$1\,^1A'$	O 1s $\rightarrow \sigma^*$	532.50	531.77	-0.73
	$2\,^1A'$	O 1s $\rightarrow$ Ry 3s	537.20	537.37	0.17
	$1\,^3A'$	O 1s $\rightarrow \sigma^*$	531.80	531.05	-0.75
	$2\,^3A'$	O 1s $\rightarrow$ Ry 3s	536.71	537.00	0.29
	$1\,^2A'$	O 1s $\rightarrow \infty$	542.51	542.71	0.20
	$1\,^1A'$	F 1s $\rightarrow \sigma^*$	683.88	683.48	-0.40
	$2\,^1A'$	F 1s $\rightarrow$ Ry 3s	689.31	689.22	-0.09
	$1\,^3A'$	F 1s $\rightarrow \sigma^*$	683.03	682.86	-0.17
	$2\,^3A'$	F 1s $\rightarrow$ Ry 3s	689.29	688.87	-0.42
	$1\,^2A'$	F 1s $\rightarrow \infty$	694.26	693.82	-0.44

Table S3: Validation Set Reference ~~Data~~ Excitation and Ionization Energies Computed at the aCTZ/CCSDT Level of Theory and aCTZ/DFT/MRCI [CVS-QE12] Results. Energy in eV.

Molecule	State	Assignment	CCSDT	DFT/MRCI	ΔE
CO					
	$1\ ^1B_2$	C 1s $\rightarrow \pi_1$*	287.66	286.58	-1.08
	$2\ ^1B_1$	C 1s $\rightarrow \pi_2$*	287.66	286.58	-1.08
	$3\ ^1A_1$	C 1s $\rightarrow$ Ry 3s	292.87	292.85	-0.02
	$4\ ^1B_2$	C 1s $\rightarrow$ Ry 3p	293.92	292.96	-0.96
	$5\ ^1B_1$	C 1s $\rightarrow$ Ry 3p	293.92	292.96	-0.96
	$1\ ^3B_2$	C 1s $\rightarrow \pi_1$*	286.25	285.98	-0.27
	$2\ ^3B_1$	C 1s $\rightarrow \pi_2$*	286.25	285.98	-0.27
	$3\ ^3A_1$	C 1s $\rightarrow$ Ry 3s	292.77	292.73	-0.04
	$4\ ^3B_2$	C 1s $\rightarrow$ Ry 3p	293.79	292.79	-1.00
	$5\ ^3B_1$	C 1s $\rightarrow$ Ry 3p	293.79	292.79	-1.00
	$1\ ^2A_1$	C 1s $\rightarrow \infty$	296.32	296.46	0.14
	$1\ ^1B_2$	O 1s $\rightarrow \pi_1$*	534.21	534.24	0.03
	$2\ ^1B_1$	O 1s $\rightarrow \pi_2$*	534.21	534.24	0.03
	$3\ ^1A_1$	O 1s $\rightarrow$ Ry 3s	538.78	539.60	0.82
	$4\ ^1B_2$	O 1s $\rightarrow$ Ry 3p	539.98	539.75	-0.23
	$5\ ^1B_1$	O 1s $\rightarrow$ Ry 3p	539.98	539.75	-0.23
	$1\ ^3B_2$	O 1s $\rightarrow \pi_1$*	533.78	533.91	0.13
	$2\ ^3B_1$	O 1s $\rightarrow \pi_2$*	533.78	533.91	0.13
	$3\ ^3A_1$	O 1s $\rightarrow$ Ry 3s	538.68	539.51	0.83
	$4\ ^3B_2$	O 1s $\rightarrow$ Ry 3p	539.90	539.58	-0.32
	$5\ ^3B_1$	O 1s $\rightarrow$ Ry 3p	539.90	539.67	-0.23
	$1\ ^2A_1$	O 1s $\rightarrow \infty$	542.33	541.18	-1.15
N_2					
	$1\ ^1B_{3u}$	N2 1s $\rightarrow \pi_1$*	401.22	400.74	-0.48
	$2\ ^1B_{2u}$	N2 1s $\rightarrow \pi_2$*	401.22	400.74	-0.48
	$3\ ^1B_{3g}$	N1 1s $\rightarrow \pi_1$*	401.27	400.78	-0.49
	$4\ ^1B_{2g}$	N1 1s $\rightarrow \pi_2$*	401.27	400.78	-0.49
	$5\ ^1A_g$	N2 1s $\rightarrow$ Ry 3s	406.79	405.98	-0.81
	$6\ ^1A_g$	N1 1s $\rightarrow$ Ry 3s	406.87	406.04	-0.83
	$1\ ^3B_{3u}$	N2 1s $\rightarrow \pi_1$*	400.34	400.20	-0.14
	$2\ ^3B_{2u}$	N2 1s $\rightarrow \pi_2$*	400.34	400.21	-0.13
	$3\ ^3B_{3g}$	N1 1s $\rightarrow \pi_1$*	400.43	400.26	-0.17
	$4\ ^3B_{2g}$	N1 1s $\rightarrow \pi_2$*	400.43	400.27	-0.16
	$5\ ^3A_g$	N2 1s $\rightarrow$ Ry 3s	406.69	405.83	-0.86
	$6\ ^3A_g$	N1 1s $\rightarrow$ Ry 3s	406.77	405.88	-0.89
	$1\ ^2B_{1u}$	N2 1s $\rightarrow \infty$	410.01	409.75	-0.26
	$2\ ^2A_g$	N1 1s $\rightarrow \infty$	409.90	409.69	-0.21

Molecule	State	Assignment	CCSDT	DFT/MRCI	ΔE
H_2CO					
	$1\,^1B_2$	C 1s $\rightarrow \pi^*$	286.26	285.89	-0.37
	$2\,^1A_1$	C 1s $\rightarrow$ Ry 3s	290.57	290.09	-0.48
	$3\,^1B_1$	C 1s $\rightarrow$ Ry 3p	291.57	291.00	-0.57
	$1\,^3B_2$	C 1s $\rightarrow \pi^*$	285.40	285.47	0.07
	$2\,^3A_1$	C 1s $\rightarrow$ Ry 3s	290.10	289.78	-0.32
	$3\,^3B_1$	C 1s $\rightarrow$ Ry 3p	291.32	290.86	-0.46
	$1\,^2A_1$	C 1s $\rightarrow \infty$	294.67	294.46	-0.21
	$1\,^1B_2$	O 1s $\rightarrow \pi^*$	530.93	531.29	0.36
	$2\,^1A_1$	O 1s $\rightarrow$ Ry 3s	535.44	535.12	-0.32
	$3\,^1A_1$	O 1s $\rightarrow$ Ry 3p	536.30	536.13	-0.17
	$1\,^3B_2$	O 1s $\rightarrow \pi^*$	530.52	530.98	0.46
	$2\,^3A_1$	O 1s $\rightarrow$ Ry 3s	535.44	535.03	-0.41
	$3\,^3A_1$	O 1s $\rightarrow$ Ry 3p	536.09	535.87	-0.22
	$1\,^2A_1$	O 1s $\rightarrow \infty$	539.21	538.96	-0.25
HCN					
	$1\,^1B_1$	C 1s $\rightarrow \pi_1^*$	286.82	285.84	-0.98
	$2\,^1B_2$	C 1s $\rightarrow \pi_2^*$	286.82	285.84	-0.98
	$3\,^1A_1$	C 1s $\rightarrow$ Ry 3s	289.53	288.67	-0.86
	$4\,^1A_1$	C 1s $\rightarrow$ Ry 3p	290.08	289.20	-0.88
	$1\,^3B_1$	C 1s $\rightarrow \pi_1^*$	286.17	285.52	-0.65
	$2\,^3B_2$	C 1s $\rightarrow \pi_2^*$	286.17	285.52	-0.65
	$3\,^3A_1$	C 1s $\rightarrow$ Ry 3s	289.00	288.34	-0.66
	$4\,^3A_1$	C 1s $\rightarrow$ Ry 3p	290.15	289.12	-1.03
	$1\,^2A_1$	C 1s $\rightarrow \infty$	293.42	293.37	-0.05
	$1\,^1B_1$	N 1s $\rightarrow \pi_1^*$	400.02	398.95	-1.07
	$2\,^1B_2$	N 1s $\rightarrow \pi_2^*$	400.02	398.99	-1.03
	$3\,^1A_1$	N 1s $\rightarrow$ Ry 3s	402.74	401.69	-1.05
	$4\,^1A_1$	N 1s $\rightarrow$ Ry 3p	403.22	402.33	-0.89
	$1\,^3B_1$	N 1s $\rightarrow \pi_1^*$	399.35	398.65	-0.70
	$2\,^3B_2$	N 1s $\rightarrow \pi_2^*$	399.35	398.72	-0.63
	$3\,^3A_1$	N 1s $\rightarrow$ Ry 3s	402.63	401.64	-0.99
	$4\,^3A_1$	N 1s $\rightarrow$ Ry 3p	403.09	402.14	-0.95
	$1\,^2A_1$	N 1s $\rightarrow \infty$	406.66	406.20	-0.46
HNC					
	$1\,^1B_1$	C 1s $\rightarrow \pi_1^*$	287.00	285.70	-1.30
	$2\,^1B_2$	C 1s $\rightarrow \pi_2^*$	287.00	285.70	-1.30
	$3\,^1A_1$	C 1s $\rightarrow$ Ry 3s	289.33	288.11	-1.22
	$4\,^1A_1$	C 1s $\rightarrow$ Ry 3p	290.92	289.70	-1.22

Molecule	State	Assignment	CCSDT	DFT/MRCI	ΔE
HNC					
	$1\,^3B_1$	C 1s $\rightarrow \pi_1$*	285.77	285.20	-0.57
	$2\,^3B_2$	C 1s $\rightarrow \pi_2$*	285.77	285.20	-0.57
	$3\,^3A_1$	C 1s $\rightarrow$ Ry 3s	289.30	288.07	-1.23
	$4\,^3A_1$	C 1s $\rightarrow$ Ry 3p	290.65	289.51	-1.14
	$1\,^2A_1$	C 1s $\rightarrow \infty$	293.64	293.35	-0.29
	$1\,^1B_1$	N 1s $\rightarrow \pi_1$*	400.93	400.76	-0.17
	$2\,^1B_2$	N 1s $\rightarrow \pi_2$*	400.93	400.76	-0.17
	$3\,^1A_1$	N 1s $\rightarrow$ Ry 3s	402.85	402.49	-0.36
	$4\,^1A_1$	N 1s $\rightarrow$ Ry 3p	404.33	404.20	-0.13
	$1\,^3B_1$	N 1s $\rightarrow \pi_1$*	400.58	400.47	-0.11
	$2\,^3B_2$	N 1s $\rightarrow \pi_2$*	400.58	400.47	-0.11
	$3\,^3A_1$	N 1s $\rightarrow$ Ry 3s	402.38	402.13	-0.25
	$4\,^3A_1$	N 1s $\rightarrow$ Ry 3p	404.36	404.09	-0.27
	$1\,^2A_1$	N 1s $\rightarrow \infty$	407.17	406.68	-0.49
HNO					
	$1\,^1A''$	O 1s $\rightarrow \pi$*	530.66	531.43	0.77
	$2\,^1A'$	O 1s $\rightarrow$ Ry 3s	537.01	537.01	-0.00
	$3\,^1A'$	O 1s $\rightarrow$ Ry 3p	538.68	538.60	-0.08
	$1\,^3A''$	O 1s $\rightarrow \pi$*	530.15	531.06	0.91
	$2\,^3A'$	O 1s $\rightarrow$ Ry 3s	537.01	536.89	-0.12
	$3\,^3A'$	O 1s $\rightarrow$ Ry 3p	538.30	538.16	-0.14
	$1\,^2A'$	O 1s $\rightarrow \infty$	541.47	541.68	0.21
	$1\,^1A''$	N 1s $\rightarrow \pi$*	399.27	399.62	0.35
	$2\,^1A'$	N 1s $\rightarrow$ Ry 3s	405.17	405.11	-0.06
	$3\,^1A'$	N 1s $\rightarrow$ Ry 3p	407.14	407.63	0.49
	$1\,^3A''$	N 1s $\rightarrow \pi$*	398.45	399.16	0.71
	$2\,^3A'$	N 1s $\rightarrow$ Ry 3s	404.62	404.77	0.15
	$3\,^3A'$	N 1s $\rightarrow$ Ry 3p	407.01	407.06	0.05
	$1\,^2A'$	N 1s $\rightarrow \infty$	410.09	410.52	0.43
CH_3F					
	$1\,^1A'$	C 1s $\rightarrow \sigma$*	289.42	289.15	-0.27
	$2\,^1A'$	C 1s $\rightarrow$ Ry 3s	289.66	289.44	-0.22
	$3\,^1A'$	C 1s $\rightarrow$ Ry 3p	290.68	290.40	-0.28
	$4\,^1A''$	C 1s $\rightarrow$ Ry 3p	290.68	290.40	-0.28
	$1\,^3A'$	C 1s $\rightarrow \sigma$*	288.82	288.85	0.03
	$2\,^3A'$	C 1s $\rightarrow$ Ry 3s	289.14	289.07	-0.07
	$3\,^3A'$	C 1s $\rightarrow$ Ry 3p	290.37	290.24	-0.13
	$4\,^3A''$	C 1s $\rightarrow$ Ry 3p	290.37	290.24	-0.13

Molecule	State	Assignment	CCSDT	DFT/MRCI	Δ E
CH_3F					
	$1\,^2A'$	C 1s $\rightarrow \infty$	293.70	293.94	0.24
	$1\,^1A'$	F 1s $\rightarrow \sigma^*$	688.22	689.55	1.33
	$2\,^1A'$	F 1s $\rightarrow$ Ry 3s	688.88	690.30	1.42
	$3\,^1A'$	F 1s $\rightarrow$ Ry 3p	689.90	691.32	1.42
	$4\,^1A''$	F 1s $\rightarrow$ Ry 3p	689.90	691.32	1.42
	$1\,^3A'$	F 1s $\rightarrow \sigma^*$	687.63	689.01	1.38
	$2\,^3A'$	F 1s $\rightarrow$ Ry 3s	688.79	690.18	1.39
	$3\,^3A'$	F 1s $\rightarrow$ Ry 3p	689.90	691.25	1.35
	$4\,^3A''$	F 1s $\rightarrow$ Ry 3p	689.90	691.25	1.35
	$1\,^2A'$	F 1s $\rightarrow \infty$	692.48	693.09	0.61
C_2H_2O					
	$1\,^1B_2$	C2 1s $\rightarrow \pi^*$	285.26	284.96	-0.30
	$2\,^1B_2$	C1 1s $\rightarrow \pi^*$	287.47	287.98	0.51
	$1\,^3B_2$	C2 1s $\rightarrow \pi^*$	285.46	284.92	-0.54
	$2\,^3B_2$	C1 1s $\rightarrow \pi^*$	286.23	286.52	0.29
	$1\,^2A_1$	C2 1s $\rightarrow \infty$	295.02	294.94	-0.08
	$2\,^2A_1$	C1 1s $\rightarrow \infty$	291.14	290.82	-0.32
	$1\,^1B_2$	O 1s $\rightarrow \pi^*$	532.92	533.24	0.32
	$1\,^3B_2$	O 1s $\rightarrow \pi^*$	532.52	532.97	0.45
	$1\,^2A_1$	O 1s $\rightarrow \infty$	540.26	540.08	-0.18

Table S4: Training Data Set Term Values Computed at the aCTZ/CCSDT Level of Theory and aCTZ/DFT/MRCI [CVS-QE12] Results. Energy in eV.

Mol.[Atom]	Ref. State	State	CCSDT	DFT/MRCI	ΔE
C_2H_2 [C]					
	$1\,^1E$	$2\,^1E$	0.05	0.05	0.00
	$1\,^1E$	$1\,^3E$	-0.54	-0.28	0.26
	$1\,^3E$	$2\,^3E$	0.09	0.07	-0.02
	$1\,^1E$	$1\,^2B_u$	5.17	5.56	0.39
	$1\,^1E$	$1\,^2A_g$	5.25	5.63	0.38
C_2H_4 [C]					
	$1\,^1B_u$	$1\,^1B_g$	0.01	0.01	0.00
	$1\,^1B_u$	$2\,^1B_u$	2.30	1.96	-0.34
	$1\,^1B_u$	$1\,^1A_g$	2.34	1.99	-0.35
	$1\,^1B_u$	$1\,^3B_u$	-0.44	-0.23	0.21
	$1\,^3B_u$	$1\,^3B_g$	0.05	0.03	-0.02
	$1\,^3B_u$	$2\,^3B_u$	2.50	2.09	-0.41
	$1\,^3B_u$	$1\,^3A_g$	2.53	2.12	-0.41
	$1\,^1B_u$	$1\,^2B_u$	5.73	5.71	-0.02
	$1\,^1B_u$	$1\,^2A_g$	5.78	5.74	-0.04
CH_4 [C]					
	$1\,^1A_1$	$1\,^1T$	1.33	1.28	-0.05
	$1\,^1A_1$	$1\,^3A_1$	-0.38	-0.33	0.05
	$1\,^3A_1$	$1\,^3T$	1.50	1.48	-0.02
	$1\,^1A_1$	$1\,^2A_1$	3.99	4.16	0.17
H_2O [O]					
	$1\,^1A_1$	$1\,^1B_1$	1.83	1.58	-0.25
	$1\,^1A_1$	$1\,^3A_1$	-0.47	-0.48	-0.01
	$1\,^3A_1$	$1\,^3B_1$	1.92	1.69	-0.23
	$1\,^1A_1$	$1\,^2A_1$	5.66	4.68	-0.98
NH_3 [N]					
	$1\,^1A'$	$1\,^1E$	1.68	1.56	-0.12
	$1\,^1A'$	$1\,^1A'$	2.81	2.77	-0.04
	$1\,^1A'$	$2\,^3A'$	-0.42	-0.38	0.04
	$1\,^3A'$	$1\,^3E$	1.81	1.74	-0.07
	$1\,^3A'$	$2\,^3A'$	3.09	3.03	-0.06
	$1\,^1A'$	$1\,^2A'$	4.82	4.36	-0.46
HF [F]					
	$1\,^1A_1$	$2\,^1A_1$	3.69	3.58	-0.11
	$1\,^1A_1$	$1\,^1E$	4.58	4.75	0.17
	$1\,^1A_1$	$1\,^3A_1$	-0.51	-0.49	0.02
	$1\,^3A_1$	$2\,^3A_1$	4.04	3.83	-0.21
	$1\,^3A_1$	$1\,^3E$	4.93	5.00	0.07
	$1\,^1A_1$	$1\,^2A_1$	6.50	5.17	-1.33

Mol.[Atom]	Ref. State	State	CCSDT	DFT/MRCI	ΔE
NH_2F [N]					
	$1\,^1A'$	$2\,^1A'$	1.36	1.85	0.49
	$1\,^1A'$	$1\,^1A''$	2.66	3.04	0.38
	$1\,^1A'$	$1\,^3A'$	-1.37	-0.80	0.57
	$1\,^3A'$	$2\,^3A'$	2.28	2.31	0.03
	$1\,^3A'$	$1\,^3A''$	3.91	3.61	-0.30
	$1\,^1A'$	$1\,^2A'$	6.09	6.48	0.39
NH_2F [F]					
	$1\,^1A'$	$2\,^1A'$	2.28	2.73	0.45
	$1\,^1A'$	$1\,^1A''$	3.60	3.84	0.24
	$1\,^1A'$	$1\,^3A'$	-0.96	-0.73	0.23
	$1\,^3A'$	$2\,^3A'$	3.21	3.25	0.04
	$1\,^3A'$	$1\,^3A''$	4.56	4.43	-0.13
	$1\,^1A'$	$1\,^2A'$	6.58	6.43	-0.15
CH_2NH [C]					
	$1\,^1A''$	$1\,^1A'$	3.10	2.96	-0.14
	$1\,^1A''$	$2\,^1A'$	3.86	3.75	-0.11
	$1\,^1A''$	$1\,^3A''$	-0.61	-0.32	0.29
	$1\,^3A''$	$1\,^3A'$	3.42	3.08	-0.34
	$1\,^3A''$	$2\,^3A'$	4.26	3.93	-0.33
	$1\,^1A''$	$1\,^2A'$	7.10	7.33	0.23
CH_2NH [N]					
	$1\,^1A''$	$1\,^1A'$	2.93	2.67	-0.26
	$1\,^1A''$	$2\,^1A'$	3.91	3.56	-0.35
	$1\,^1A''$	$1\,^3A''$	-0.45	-0.29	0.16
	$1\,^3A''$	$1\,^3A'$	3.11	2.75	-0.36
	$1\,^3A''$	$2\,^3A'$	4.28	3.76	-0.52
	$1\,^1A''$	$1\,^2A'$	6.94	6.78	-0.16
HOF [O]					
	$1\,^1A'$	$2\,^1A'$	4.70	5.60	0.90
	$1\,^1A'$	$1\,^3A'$	-0.70	-0.73	-0.03
	$1\,^3A'$	$2\,^3A'$	4.91	5.95	1.04
	$1\,^1A'$	$1\,^2A'$	10.01	10.94	0.93
HOF [F]					
	$1\,^1A'$	$2\,^1A'$	5.43	5.75	0.32
	$1\,^1A'$	$1\,^3A'$	-0.85	-0.62	0.23
	$1\,^3A'$	$2\,^3A'$	6.26	6.01	-0.25
	$1\,^1A'$	$1\,^2A'$	10.38	10.34	-0.04

Table S5: Validation Data Set Term Values Computed at the aCTZ/CCSDT Level of Theory and aCTZ/DFT/MRCI [CVS-QE12] Results. Energy in eV.

Mol.[Atom]	Ref. State	State	CCSDT	DFT/MRCI	ΔE
CO [C]					
	$1\,^1E$	$1\,^1A_1$	5.21	6.26	1.05
	$1\,^1E$	$2\,^1E$	6.26	6.38	0.12
	$1\,^1E$	$1\,^3E$	-1.41	-0.60	0.81
	$1\,^3E$	$1\,^3A_1$	6.52	6.75	0.23
	$1\,^3E$	$2\,^3E$	7.54	6.81	-0.73
	$1\,^1E$	$1\,^2A_1$	8.66	9.88	1.22
CO [O]					
	$1\,^1E$	$1\,^1A_1$	4.57	5.36	0.79
	$1\,^1E$	$2\,^1E$	5.77	5.51	-0.26
	$1\,^1E$	$1\,^3E$	-0.43	-0.33	0.10
	$1\,^3E$	$1\,^3A_1$	4.90	5.60	0.70
	$1\,^1E$	$2\,^3E$	6.12	5.43	-0.69
	$1\,^1E$	$1\,^2A_1$	8.12	6.94	-1.18
N_2 [N]					
	$1\,^1E$	$2\,^1E$	0.05	0.04	-0.01
	$1\,^1E$	$1\,^1A_g$	5.57	5.24	-0.33
	$1\,^1E$	$2\,^1A_g$	5.65	5.30	-0.35
	$1\,^1E$	$1\,^3E$	-0.88	-0.54	0.34
	$1\,^3E$	$2\,^3E$	0.09	0.06	-0.03
	$1\,^3E$	$1\,^3A_g$	6.35	5.63	-0.72
	$1\,^3E$	$2\,^3A_g$	6.43	5.68	-0.75
	$1\,^1E$	$1\,^2B_{1u}$	8.68	8.95	0.27
	$1\,^1E$	$1\,^2A_g$	8.79	9.02	0.23
H_2CO [C]					
	$1\,^1B_2$	$1\,^1A_1$	4.31	4.20	-0.11
	$1\,^1B_2$	$1\,^1B_1$	5.31	5.11	-0.20
	$1\,^1B_2$	$1\,^3B_2$	-0.86	-0.42	0.44
	$1\,^3B_2$	$1\,^3A_1$	4.70	4.31	-0.39
	$1\,^3B_2$	$1\,^3B_1$	5.92	5.38	-0.54
	$1\,^1B_2$	$1\,^2A_1$	8.41	8.57	0.16
H_2CO [O]					
	$1\,^1B_2$	$1\,^1A_1$	4.51	3.83	-0.68
	$1\,^1B_2$	$1\,^1B_1$	5.37	4.84	-0.53
	$1\,^1B_2$	$1\,^3B_2$	-0.41	-0.31	0.10
	$1\,^3B_2$	$1\,^3A_1$	4.92	4.05	-0.87
	$1\,^3B_2$	$1\,^3B_1$	5.57	4.89	-0.68
	$1\,^1B_2$	$1\,^2A_1$	8.28	7.68	-0.60

Mol.[Atom]	Ref. State	State	CCSDT	DFT/MRCI	ΔE
HCN [C]					
	$1\,^1E$	$1\,^1A_1$	2.71	2.83	0.12
	$1\,^1E$	$2\,^1A_1$	3.26	3.36	0.10
	$1\,^1E$	$1\,^3E$	-0.65	-0.32	0.33
	$1\,^3E$	$1\,^3A_1$	2.83	2.82	-0.01
	$1\,^3E$	$2\,^3A_1$	3.98	3.60	-0.38
	$1\,^1E$	$1\,^2A_1$	6.60	7.53	0.93
HCN [N]					
	$1\,^1E$	$1\,^1A_1$	2.72	2.74	0.02
	$1\,^1E$	$2\,^1A_1$	3.20	3.38	0.18
	$1\,^1E$	$1\,^3E$	-0.67	-0.30	0.37
	$1\,^3E$	$1\,^3A_1$	3.28	2.99	-0.29
	$1\,^3E$	$2\,^3A_1$	3.74	3.49	-0.25
	$1\,^1E$	$1\,^2A_1$	6.64	7.25	0.61
HNC [C]					
	$1\,^1E$	$1\,^1A_1$	2.33	2.41	0.08
	$1\,^1E$	$2\,^1A_1$	3.92	4.01	0.09
	$1\,^1E$	$1\,^3E$	-1.23	-0.50	0.73
	$1\,^3E$	$1\,^3A_1$	3.53	2.88	-0.65
	$1\,^3E$	$2\,^3A_1$	4.88	4.31	-0.57
	$1\,^1E$	$1\,^2A_1$	6.64	7.65	1.01
HNC [N]					
	$1\,^1E$	$1\,^1A_1$	1.92	1.73	-0.19
	$1\,^1E$	$2\,^1A_1$	3.40	3.44	0.04
	$1\,^1E$	$1\,^3E$	-0.35	-0.29	0.06
	$1\,^3E$	$1\,^3A_1$	1.80	1.66	-0.14
	$1\,^3E$	$2\,^3A_1$	3.78	3.61	-0.17
	$1\,^1E$	$1\,^2A_1$	6.24	5.92	-0.32
HNO [O]					
	$1\,^1A''$	$1\,^1A'$	6.35	5.58	-0.77
	$1\,^1A''$	$2\,^1A'$	8.02	7.17	-0.85
	$1\,^1A''$	$1\,^3A''$	-0.51	-0.37	0.14
	$1\,^3A''$	$1\,^3A'$	6.86	5.83	-1.03
	$1\,^3A''$	$2\,^3A'$	8.15	7.10	-1.05
	$1\,^1A''$	$1\,^2A'$	10.81	10.25	-0.56
HNO [N]					
	$1\,^1A''$	$1\,^1A'$	5.90	5.49	-0.41
	$1\,^1A''$	$2\,^1A'$	7.87	8.01	0.14
	$1\,^1A''$	$1\,^3A''$	-0.82	-0.46	0.36
	$1\,^3A''$	$1\,^3A'$	6.17	5.62	-0.55
	$1\,^3A''$	$2\,^3A'$	8.56	7.90	-0.66
	$1\,^1A''$	$1\,^2A'$	10.82	10.90	0.08

Mol.[Atom]	Ref. State	State	CCSDT	DFT/MRCI	ΔE
CH_3F [C]					
	$1\,^1A'$	$2\,^1A'$	0.24	0.29	0.05
	$1\,^1A'$	$1\,^1E$	1.26	1.25	-0.01
	$1\,^1A'$	$1\,^3A'$	-0.60	-0.30	0.30
	$1\,^3A'$	$2\,^3A'$	0.32	0.22	-0.10
	$1\,^3A'$	$1\,^3E$	1.55	1.39	-0.16
	$1\,^1A'$	$1\,^2A'$	4.28	4.79	0.51
CH_3F [F]					
	$1\,^1A'$	$2\,^1A'$	0.68	0.75	0.07
	$1\,^1A'$	$1\,^1E$	1.70	1.78	0.08
	$1\,^1A'$	$1\,^3A'$	-0.57	-0.54	0.03
	$1\,^3A'$	$2\,^3A'$	1.16	1.17	0.01
	$1\,^3A'$	$1\,^3E$	2.27	2.24	-0.03
	$1\,^1A'$	$1\,^2A'$	4.28	3.55	-0.73
C_2H_2O [C]					
	$1\,^1B_2$	$2\,^1B_2$	2.21	3.02	0.81
	$1\,^1B_2$	$1\,^3B_2$	0.20	-0.04	-0.24
	$1\,^3B_2$	$2\,^3B_2$	0.77	1.60	0.83
	$1\,^1B_2$	$1\,^2A_1$	5.88	5.87	-0.01
	$1\,^1B_2$	$2\,^2A_1$	9.76	9.99	0.23
C_2H_2O [O]					
	$1\,^1B_2$	$1\,^3B_2$	-0.40	-0.27	0.13
	$1\,^1B_2$	$1\,^2A_1$	7.34	6.83	-0.51

Table S6: Error Metrics of aCTZ/DFT/MRCI [CVS-QE12] Results for Term Values Computed at the aCTZ/CCSDT Level of Theory. Errors in eV.

Dataset	ME	STD DEV	MAE	RMSD
Training	0.00	0.38	0.26	0.38
Validation	-0.06	0.51	0.40	0.52

S2 Parameter Sets

The damping function applied to the off-diagonal matrix elements of the Hamiltonian has three parameters: d_1, d_2, and d_3. The form of this function is given in Eqn. 11 of the main text. The first parameter, d_1, corresponds to the default damping of off-diagonal matrix elements at a small value of $\Delta E_{\mathbf{ww'}}$. The latter two parameters, d_2 and d_3, affect how steeply the damping function approaches zero. Whereas the n in QEn corresponds to d_3, the choice of the integer exponent term in the DFT/MRCI Hamiltonian damping function, d_2 is an explicitly optimized parameter which is not otherwise strictly defined in the name of the Hamiltonian. As a part of the parameter optimization procedure, constraints were placed on d_2 to reduce the N_{root} dependence of the resultant parameter set for the DFT/MRCI CVS-QEn Hamiltonian. The full set of parameters yielded by applying each constraint is summarized in Table S7, corresponding to E_{sel} = 1 E_h. The constraints $d_2 \geq 2.303$ and $d_2 \geq 4.605$ correspond to respective damping function values of 0.10 and 0.01 at $\Delta E_{\mathbf{ww'}}$ = 1 E_h. The constraint $d_2 \geq 0.200$ corresponds to a damping function value close to 0.8 at this value of $\Delta E_{\mathbf{ww'}}$. Unconstrained optimizations were observed to push this value to zero to fit to outliers in the training data set, namely H_2O and HF. This, in turn, was observed to result in a strong N_{root} dependence, wherein the computed excitation energies are shifted up to 1.0 eV when an increasing number of roots (100 roots or more) were requested in a calculation. The computational cost of such calculations increases accordingly. Thus, a minimal constraint was set such that $d_2 \geq 0.200$.

The mean absolute errors (MAEs) for the parameter sets listed in Table S7 are tabulated in Tables S8 and S9 for the training set and validation set data, respectively. These data are presented graphically in Figure S1. To compare the N_{root} dependence of these parameter sets, a representative example was selected. The carbon NEXAFS of thymine was computed for each parameter set. These calculations employ the aug-cc-pCVDZ basis set. The energetic shift of the first root in the spectrum was determined by separately computing both $N = 1$ and $N = 40$ core excited-state roots. This value of ΔE, for the first carbon core-excited state root, is reported in Table S10, and depicted graphically in Figure S2, for each parameter set listed in Table S10. The spectra simulated using

Table S7: Complete Set of DFT/MRCI CVS-QE*n* Hamiltonian Parameters for E_{sel} = 1.0 Hartree.

Constraint	$p_C^{(v,v)}$	$p_C^{(c,v)}$	p_X	d_1	d_2	d_3	E_{sel}
$d_2 \geq 0.200$	0.425815	0.565179	0.254799	0.494512	0.200010	6.0	1.0
	0.425815	0.566093	0.254799	0.492991	0.200001	8.0	1.0
	0.425815	0.565268	0.254799	0.490723	0.200921	10.0	1.0
	0.425815	0.565711	0.254799	0.490446	0.200036	12.0	1.0
	0.425815	0.565546	0.254799	0.491028	0.200028	14.0	1.0
	0.425815	0.565868	0.254799	0.490210	0.200038	16.0	1.0
$d_2 \geq 2.303$	0.425815	0.539416	0.254799	0.553897	2.302585	6.0	1.0
	0.425815	0.548974	0.254799	0.540495	2.302619	8.0	1.0
	0.425815	0.558624	0.254799	0.534700	2.303182	10.0	1.0
	0.425815	0.564040	0.254799	0.529467	2.302585	12.0	1.0
	0.425815	0.567684	0.254799	0.522341	2.302903	14.0	1.0
	0.425815	0.569995	0.254799	0.518889	2.302585	16.0	1.0
$d_2 \geq 4.605$	0.425815	0.524422	0.254799	0.565402	4.605213	6.0	1.0
	0.425815	0.537032	0.254799	0.554308	4.605170	8.0	1.0
	0.425815	0.542700	0.254799	0.539983	4.605170	10.0	1.0
	0.425815	0.548460	0.254799	0.533383	4.613243	12.0	1.0
	0.425815	0.557102	0.254799	0.531182	4.605334	14.0	1.0
	0.425815	0.560389	0.254799	0.526050	4.609609	16.0	1.0

each of the parameter sets (using the latter computation of 40 core-excited state roots) are depicted across Figures S3, S4, and S5, with comparison to the experimental NEXAFS spectrum.[1] These spectra are simulated by Lorentzian convolution using a full-width half maximum (fwhm) of 0.40 eV.

Table S8: Mean Absolute Error of CVS-QE*n* Hamiltonian Parameter Sets Across the Training Dataset. Errors in eV.

	d_3					
d_2	6	8	10	12	14	16
≥ 0.200	0.30	0.31	0.32	0.33	0.33	0.34
≥ 2.303	0.65	0.58	0.55	0.53	0.51	0.49
≥ 4.605	0.77	0.69	0.63	0.58	0.55	0.54

The final parameter set selected in this work employs the constraint $d_2 \geq 2.303$, with an exponent value of d_3 = 12. This parameter set was determined to be balanced in its MAE across the training and validation data sets and in the observed N_{root} dependence for the selected example. The data points corresponding to this Hamiltonian, which we denote CVS-QE12, are indicated

Table S9: Mean Absolute Errors of CVS-QE*n* Hamiltonian Parameter Sets Across the Validation Dataset. Errors in eV.

	d_3					
d_2	6	8	10	12	14	16
≥ 0.200	0.46	0.47	0.47	0.48	0.49	0.49
≥ 2.303	0.59	0.55	0.54	0.54	0.51	0.51
≥ 4.605	0.66	0.62	0.57	0.54	0.54	0.53

Table S10: ΔE for the First Excited-State Root in Carbon K-shell Excitations of Thymine from $N = 1$ to $N = 40$ Roots. The ΔE values are tabulated corresponding to each Hamiltonian parameter set in Table S7. Energies in eV

	d_3					
d_2	6	8	10	12	14	16
≥ 0.200	-0.53	-0.54	-0.54	-0.56	-0.56	-0.57
≥ 2.303	-0.24	-0.30	-0.36	-0.40	-0.43	-0.46
≥ 4.605	-0.18	-0.20	-0.26	-0.30	-0.35	-0.38

and labeled in the plots of Figures S1 and S2. The parameters for CVS-QE12 are listed to the full six decimal places in Table S11, as determined by the optimization routine employed in this work. The DFT/MRCI Hamiltonian has canonically been presented in terms of two parameter sets, respectively corresponding to an E_{sel} cutoff value of either 1.0 or 0.8 E_h.[2] The smaller E_{sel} value of 0.8 E_h is intended to afford further computational savings. Here, the selected CVS-QE12 Hamiltonian corresponding to the standard cutoff of E_{sel} = 1 E_h was then additionally optimized in terms of a cutoff of E_{sel} = 0.8 E_h. This parameter set employs the appropriate $p_C^{(v,v)}$ and p_X terms from the corresponding valence Hamiltonian.[3] In terms of computational savings, this approach may nevertheless be superseded by the approximative methods *p*-DFT/MRCI and DFT/MRCI(2).[4,5]

Table S11: Final Set of DFT/MRCI Parameters for the CVS-QE12 Hamiltonian.

$p_C^{(v,v)}$	$p_C^{(c,v)}$	p_X	d_1	d_2	d_3	E_{sel}
0.425815	0.564040	0.254799	0.529467	2.302585	12.0	1.0
0.419210	0.560186	0.257655	0.545822	2.302585	12.0	0.8

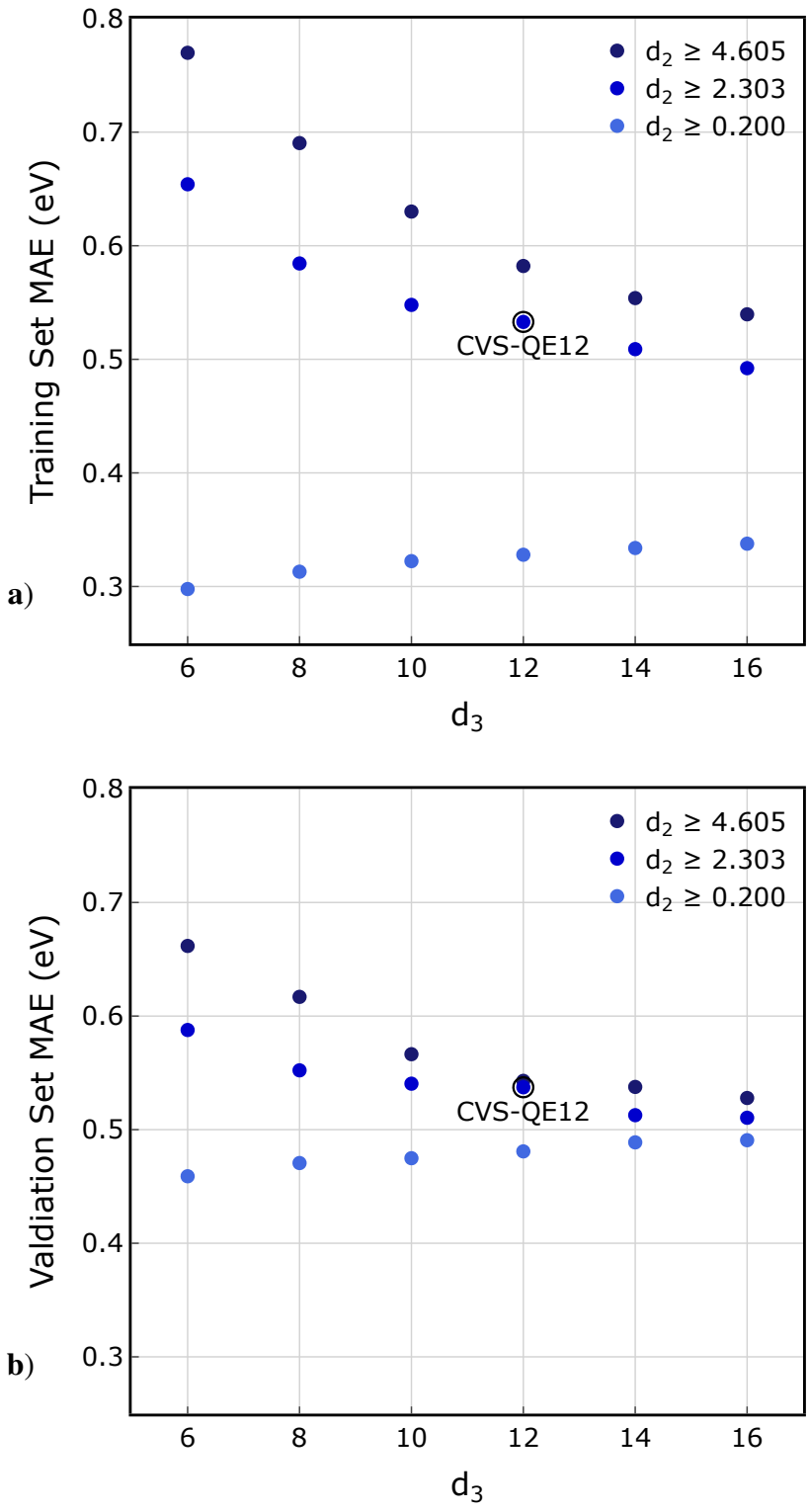

Figure S1: Training (a) and Validation (b) Set MAE of CVS-QE*n* Hamiltonian Parameter Sets. The series in the plots are grouped by the d_2 constraint applied, and errors are plotted as a function of the exponent d_3.

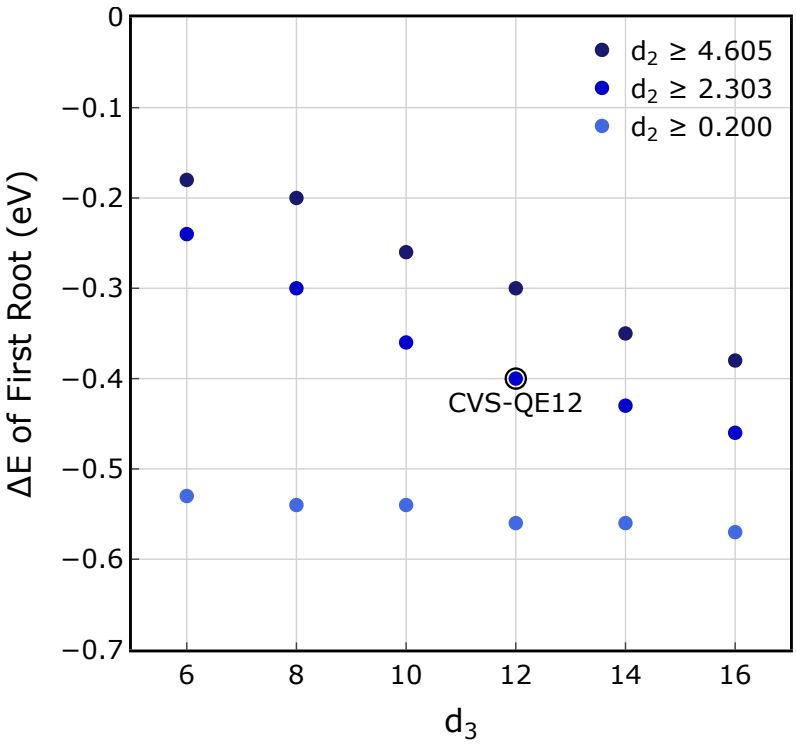

Figure S2: ΔE for the First Excited-State Root in Carbon K-shell Excitations of Thymine from $N = 1$ to $N = 40$ Roots. The ΔE values are tabulated corresponding to each CVS-QEn Hamiltonian parameter set in Table S7. The series in the plot are grouped by the d_2 constraint applied, and energy differences are plotted as a function of the exponent d_3.

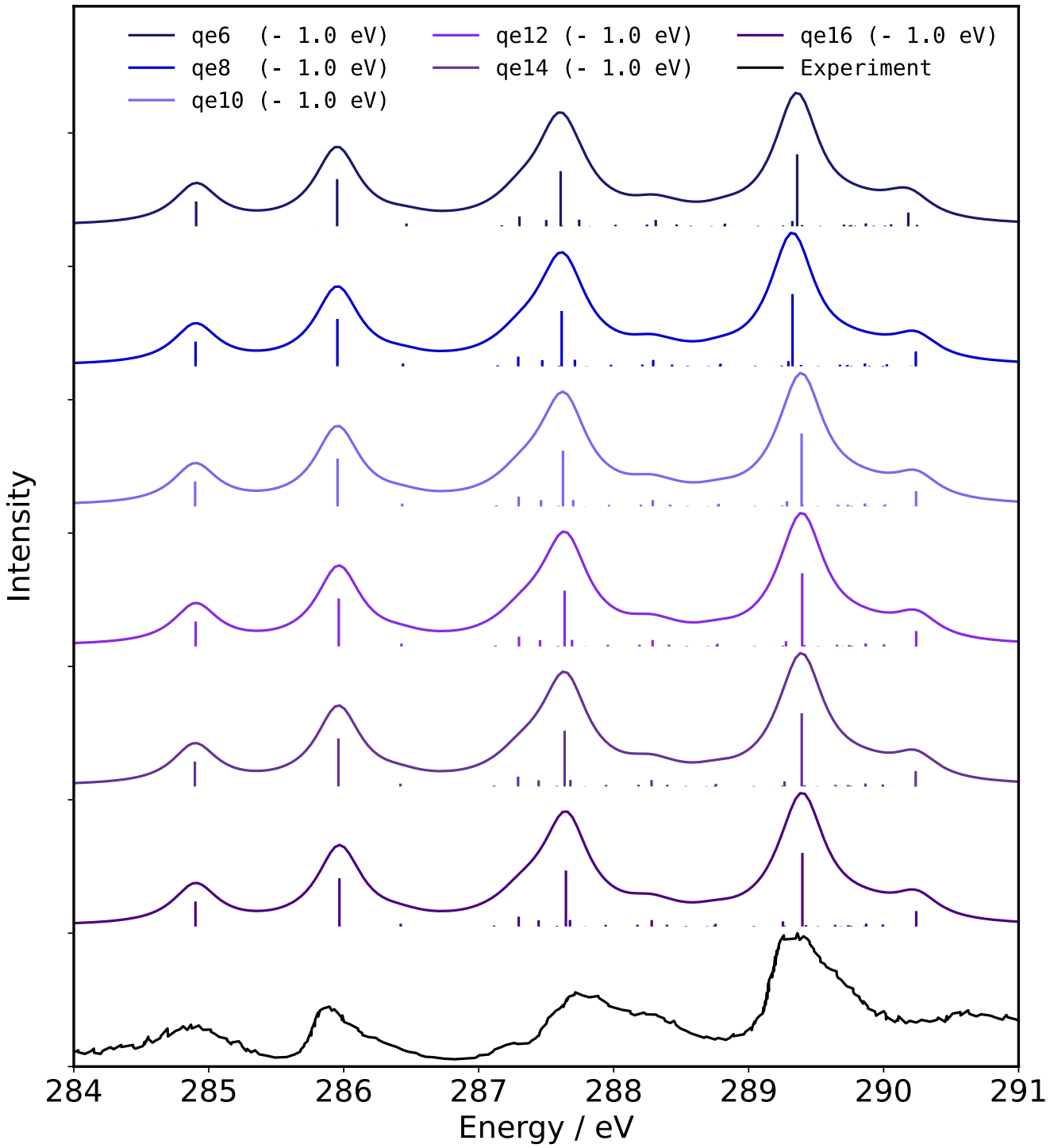

Figure S3: Carbon K-shell NEXAFS Spectrum of Thymine Simulated with CVS-QE*n* Hamiltonian Parameter Sets for the $d_2 \geq 0.200$ Constraint. Experimental spectrum (black) digitized from source.[1]

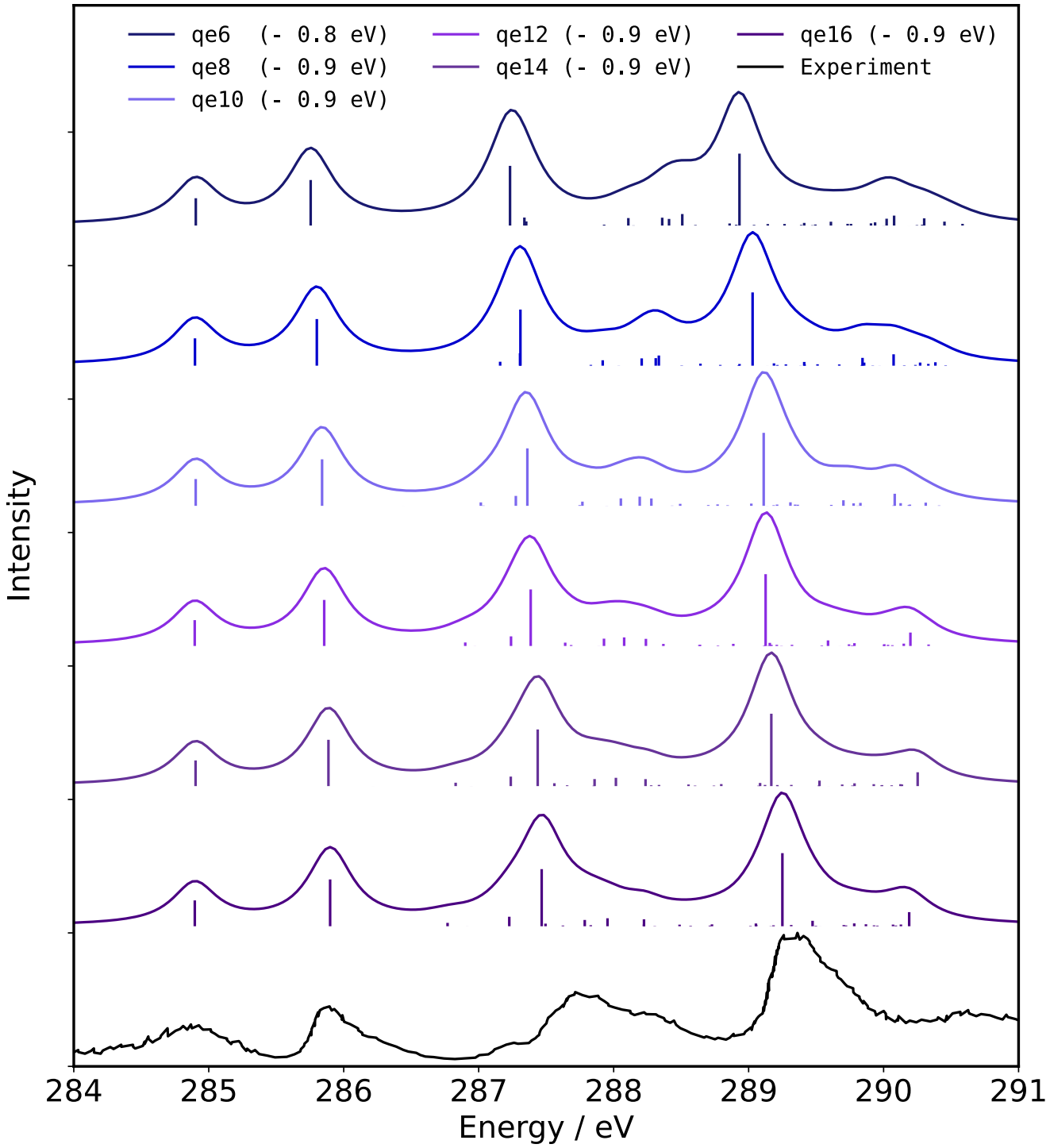

Figure S4: Carbon K-shell NEXAFS Spectrum of Thymine Simulated with CVS-QE*n* Hamiltonian Parameter Sets for the $d_2 \geq 2.303$ Constraint. Experimental spectrum (black) digitized from source.[1]

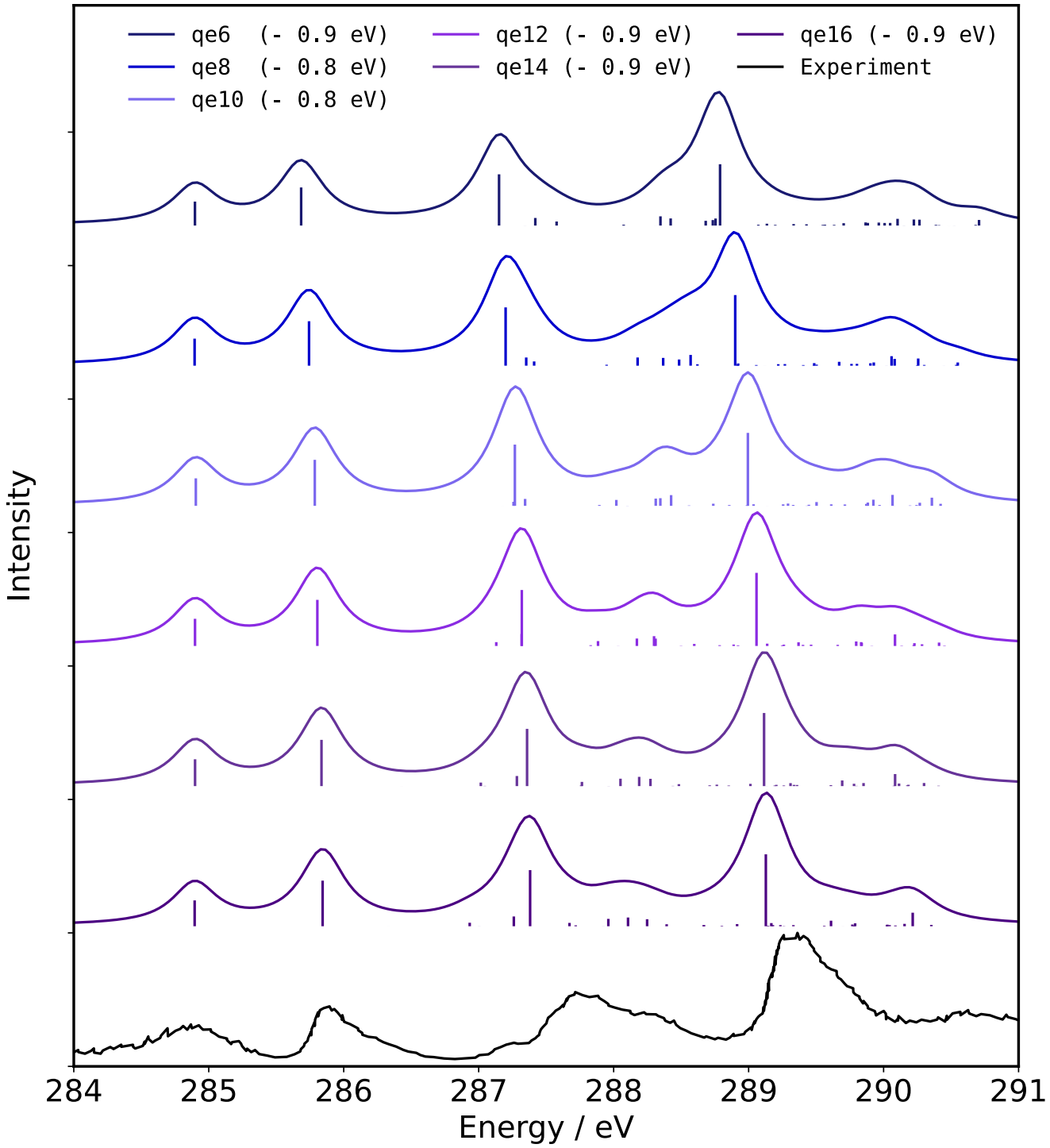

Figure S5: Carbon K-shell NEXAFS Spectrum of Thymine Simulated with CVS-QE*n* Hamiltonian Parameter Sets for the $d_2 \geq 4.605$ Constraint. Experimental spectrum (black) digitized from source.[1]

S3 Total CPU Times

In Table S12, the total CPU times are tabulated for the data in Table IV of the main text. The timings below include the KS-DFT computation. As before, these correspond to calculations run using a single thread of an Intel Xeon Gold 6130 processor. The aug-cc-pCVTZ basis set was used for ethyl trifluoroacetate, thymine, and allyl, and the aug-cc-pCVDZ basis was used for cytosine.

Table S12: Total CPU times (in seconds) of the DFT/MRCI, *p*-DFT/MRCI, and DFT/MRCI(2) calculations for the Representative Applications.

Ethyl Trifluoroacetate				
DFT/MRCI	*p*-DFT/MRCI	DFT/MRCI(2)	Atom	*N* Roots
2739	1975	1975	C	20
Cytosine				
DFT/MRCI	*p*-DFT/MRCI	DFT/MRCI(2)	Atom	*N* Roots
2734	353	276	C	50
574	114	110	N	20
7062	336	176	O	50
Thymine				
DFT/MRCI	*p*-DFT/MRCI	DFT/MRCI(2)	Structure	*N* Roots
1156	650	648	*gs*	4
1158	690	726	$n\pi^*$	4
1153	635	653	$\pi\pi^*$	4
Allyl				
DFT/MRCI	*p*-DFT/MRCI	DFT/MRCI(2)	Atom	*N* Roots
244	210	210	C	50

S4 Simulation of NEXAFS and X-ray Photoelectron Spectra

In this section, additional data is provided for each of the representative applications presented in the main text. In these applications, the carbon XPS of ethyl trifluoroacetate molecule, and the NEXAFS spectra for two nucleobases, cytosine and thymine, are computed at the DFT/MRCI level of theory and compared to experimental spectra. To compare DFT/MRCI to its comparative methods, *p*-DFT/MRCI, and DFT/MRCI(2), these spectra are recomputed using the CVS-QE12 Hamiltonian. Additional details of the calculations required to simulate the spectra of cytosine are also provided. For thymine, the ground and excited-state Oxygen K-shell NEXAFS spectra of thymine are computed using the three DFT/MRCI methods in the main text (using the aug-cc-pCVTZ/aug-cc-pVTZ basis sets). Here, a similar example is displayed for the set of ground-state NEXAFS spectra, for the carbon, nitrogen, and oxygen atoms. These are included here as an additional illustrative example. The C, N, and O XAS are simulated for each nucleobase using the aug-cc-pCVDZ basis set for all non-hydrogen atoms, and the aug-cc-pVDZ basis set for hydrogen atoms. In each figure, the vertical bars of the computed spectra represent the oscillator strengths, which are normalized to unity, and the lines represent the Lorentzian convolution, with a full width at half maximum (FWHM) of 0.60 eV. The simulated spectra are compared against the corresponding experimental spectrum, in black, which are digitized from their reference publications.[1,6,7] The computed spectra are appropriately red- or blue-shifted, as indicated in the legend of each plot. The peaks are typically shifted to align the first 1s $\rightarrow \pi^*$ transition to the first experimental peak. Similarly, the carbon K-edge XPS of ethyl trifluoroacetate is simulated using the DFT/MRCI methods with CVS-QE12 and the aug-cc-pCVTZ basis set on the carbon, oxygen, and fluorine atoms, and the aug-cc-pVTZ basis set on the hydrogen atoms. Using the ionization probabilities associated with each state, the spectrum is simulated using a Lorentzian convolution with a FWHM of 0.60 eV. The experimental XPS is once again displayed for reference.[8]

S4.1 Ethyl Trifluoroacetate XPS

In the main text, the CVS-QE12 and CVS-R2017 Hamiltonian are directly compared with the simulation of the carbon XPS of ethyl trifluoroacetate. In Figure S6, this spectrum is additionally simulated with the approximative methods *p*-DFT/MRCI and DFT/MRCI(2), for comparison. These results all employ the CVS-QE12 Hamiltonian, demonstrating that the approximative variations to the standard DFT/MRCI formulation do not degrade its good performance. These spectra are unshifted.

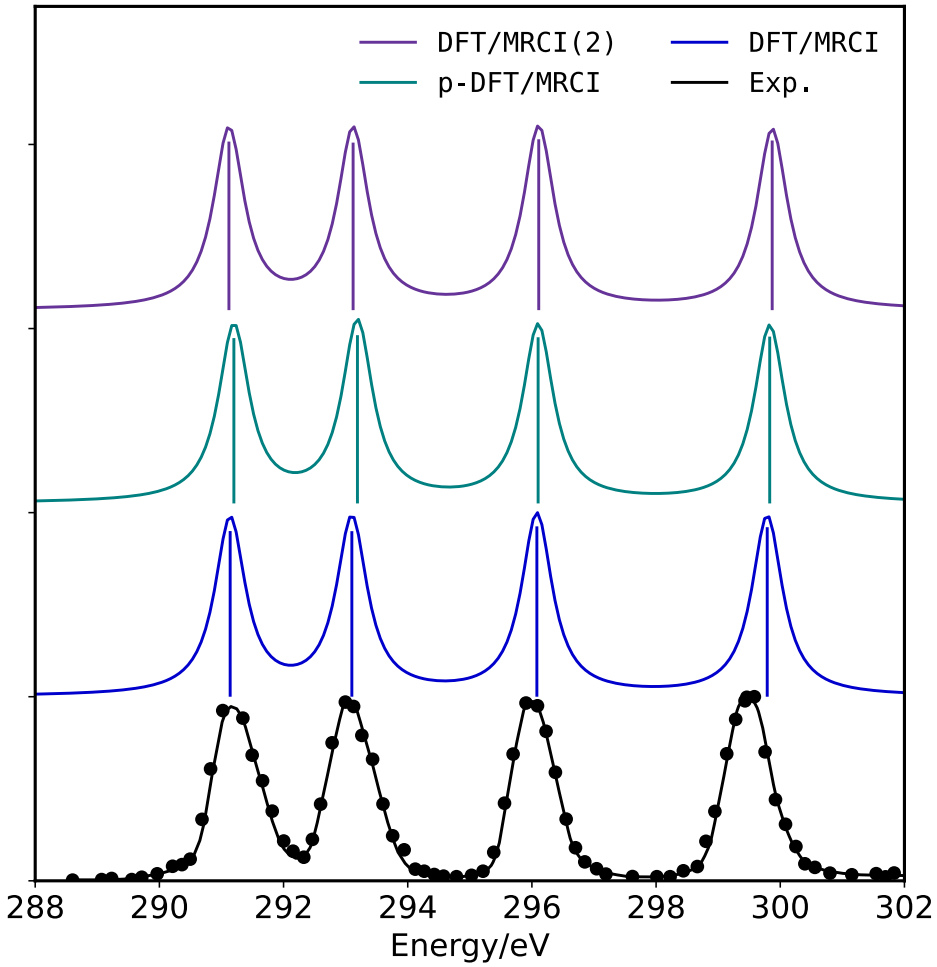

Figure S6: Carbon XPS of Ethyl Trifluoroacetate Simulated with DFT/MRCI(2) (purple), *p*-DFT/MRCI (teal), and DFT/MRCI (blue), using the CVS-QE12 Hamiltonian with Comparison to Experimental Gas-Phase X-Ray Photo-electron Spectrum (black). The simulated spectra consider only the anti-anti-conformer. Experimental data, shown with a line of best fit, are digitized from their source publication.[8]

S4.2 Cytosine NEXAFS

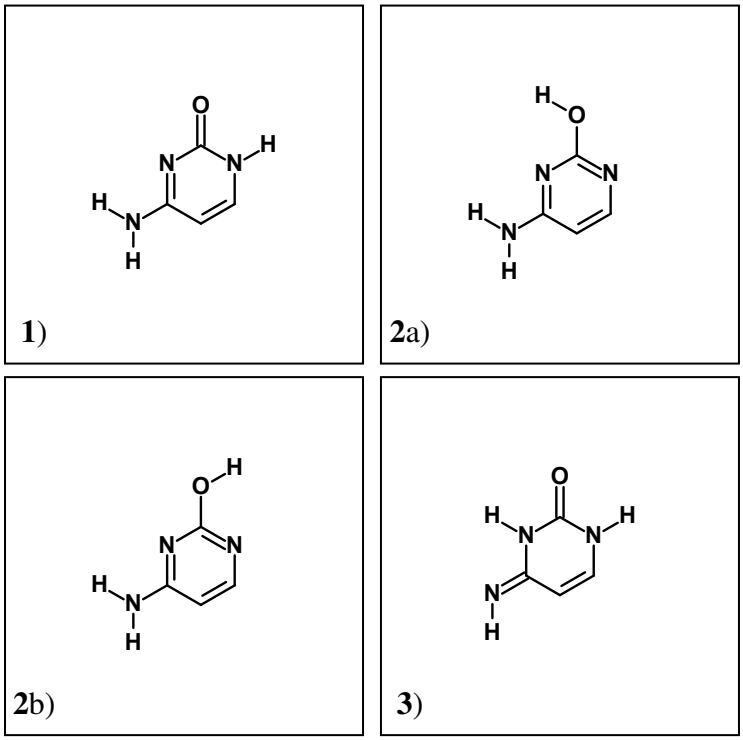

Figure S7: Four Tautomers of Cytosine used to simulate the NEXAFS Spectra. The structures and the naming convention used here are those employed by Feyer *et al.*[6,7]

The NEXAFS spectra for the carbon, nitrogen, and oxygen atoms of cytosine are simulated with CVS-QE12. In the main text, these spectra are compared to those simulated by the CVS-ADC(2)-x and fc-CVS-EOM-CCSD methods. For each spectrum, four structures of cytosine must be considered; the method for obtaining the experimental spectra required an evaporation temperature of 450K to obtain the gas-phase spectral data, and so the four lowest energy structures of cytosine were considered by Feyer *et al* in their theoretical analysis.[6,7] The same four structures are considered here. Their spectra are combined and weighted relative to one another using the Boltzmann Population Ratios (BPRs) computed for a temperature of 450K by Trygubenko *et al.*,[9] as cited by Feyer *et al.*[6] The four structures of cytosine, following the naming convention employed by Feyer *et al*, are illustrated in Figure S7. This procedure is also used to simulate these spectra with the approximative DFT/MRCI methods. A comparison between these methods for the resultant spectra is shown in Figure S8.

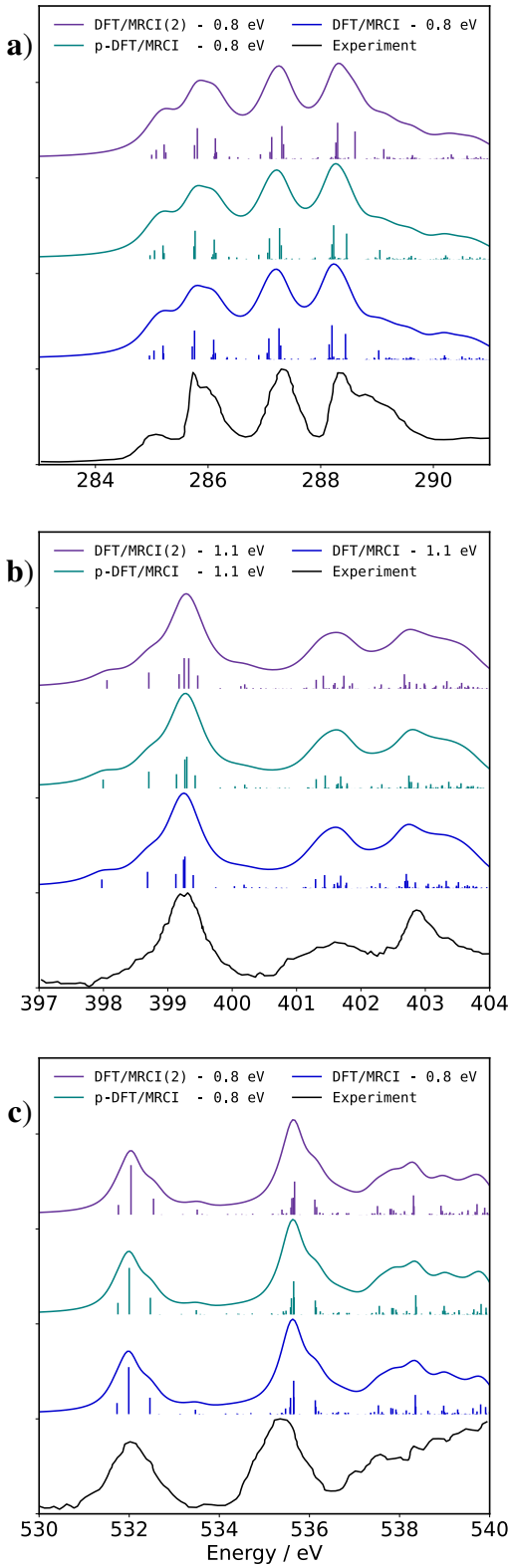

Figure S8: NEXAFS Spectrum of Carbon (**a**), Nitrogen (**b**), and Oxygen (**c**) K-shells of Cytosine Simulated with DFT/MRCI(2) (purple), p-DFT/MRCI (teal), and standard DFT/MRCI (blue), with the CVS-QE12 Hamiltonian. Experimental spectra (black) digitized from source.[6,7]

S4.3 Thymine NEXAFS

In the main text, the ground and excited-state oxygen K-shell NEXAFS spectra of thymine are simulated using the ground and excited-state geometries. These results include a comparison between DFT/MRCI and its approximative variants, *p*-DFT/MRCI, and DFT/MRCI(2); however, simulation of the complete NEXAFS spectra may better illustrate the agreement between each of the DFT/MRCI methods. Below, the ground-state NEXAFS spectra are computed for each of the carbon, nitrogen, and oxygen atoms using each of the DFT/MRCI methods with the CVS-QE12 Hamiltonian (see Fig. S9). The aug-cc-pCVDZ basis set is used here. As was the case for the previous applications, these examples show remarkable agreement across the three methods.

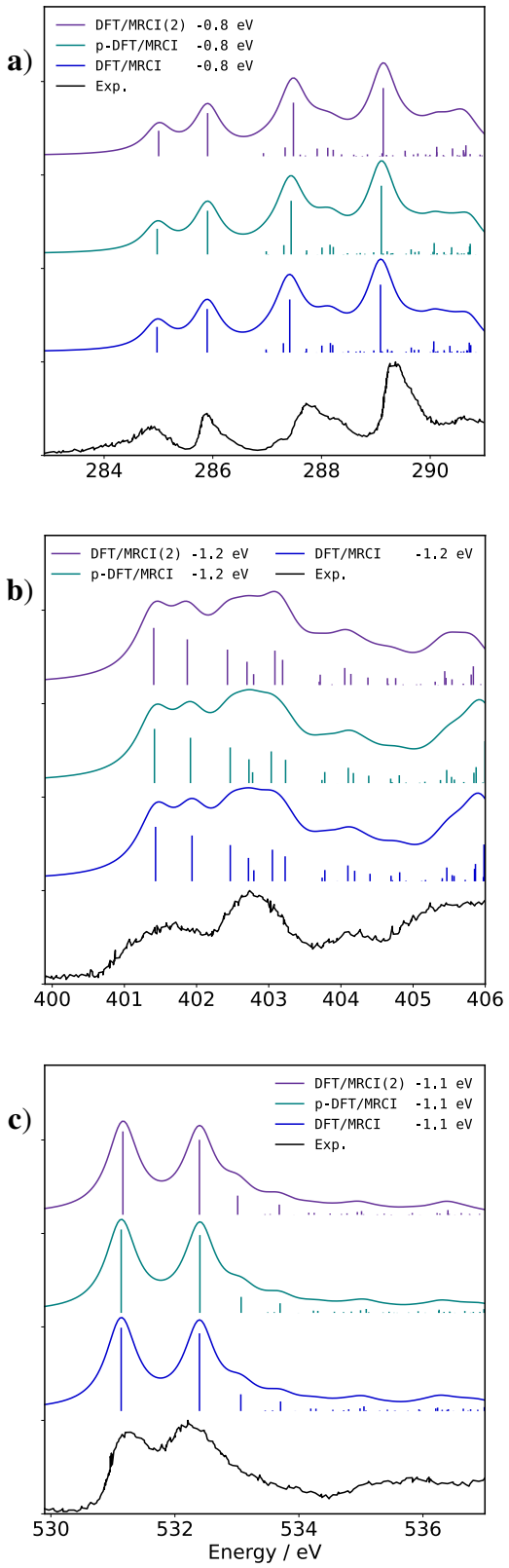

Figure S9: NEXAFS Spectrum of Carbon (**a**), Nitrogen (**b**), and Oxygen (**c**) K-shells of Thymine Simulated with DFT/MRCI(2) (purple), p-DFT/MRCI (teal), and standard DFT/MRCI (blue), with the CVS-QE12 Hamiltonian. Experimental spectra (black) digitized from source.[1]

S5 Molecular Geometries

S5.1 Training & Validation Data Sets

Table S13: Training Set Molecules

Acetylene (Å) from Ref. 10			
C	0.0000000	0.0000000	0.6035179
C	0.0000000	0.0000000	-0.6035179
H	0.0000000	0.0000000	1.6616643
H	0.0000000	0.0000000	-1.6616643
Ethylene (Å) from Ref. 10			
C	0.0000000	0.6669040	0.0000000
C	0.0000000	-0.6669040	0.0000000
H	0.0000000	1.2295220	0.9222906
H	0.0000000	-1.2295220	0.9222906
H	0.0000000	1.2295220	-0.9222906
H	0.0000000	-1.2295220	-0.9222906
Methane (a.u.) from Ref. 11			
C	0.0000000	0.0000000	0.0000000
H	0.0000000	-1.6802672	1.1881283
H	0.0000000	1.6802672	1.1881283
H	1.6802672	0.0000000	-1.1881283
H	-1.6802672	0.0000000	-1.1881283
Water (a.u.) from Ref. 11			
O	0.0000000	0.0000000	-0.1239094
H	0.0000000	1.4299373	0.9832658
H	0.0000000	-1.4299373	0.9832658
Ammonia (a.u.) from Ref. 11			
N	0.0000000	0.0000000	-0.1277339
H	0.0000000	-1.7699235	0.5915931
H	1.5327987	0.8849617	0.5915931
H	-1.5327987	0.8849617	0.5915931
Hydrogen Fluoride (a.u.) from Ref. 11			
F	0.0000000	0.0000000	0.0873178
H	0.0000000	0.0000000	-1.6460179
Fluoroamine (a.u.) from Ref. 11			
N	-0.1327913	-1.3902255	0.0000000
H	0.7312334	-2.0796879	-1.5632639
F	0.0202951	1.2453338	0.0000000
H	0.7312334	-2.0796879	1.5632639

Azomethine (a.u.) from Ref. 11			
H	-1.9941130	1.6102925	0.0000000
N	-1.2098856	-0.1539692	0.0000000
C	1.1968660	0.0335860	0.0000000
H	2.3196128	-1.6891573	0.0000000
H	2.2341950	1.8182637	0.0000000
Hypofluorous Acid (a.u.) from Ref. 11			
O	-1.4228157	-0.1186463	0.0000000
F	1.2907140	0.0107481	0.0000000
H	-1.7499948	1.6803922	0.0000000

Table S14: Validation Set Molecules

Carbon Monoxide (a.u.) from Ref. 11			
O	0.0000000	0.0000000	0.9139730
C	0.0000000	0.0000000	-1.2182430
Dinitrogen (Å) from Ref. 11			
N	0.0000000	0.0000000	0.5503900
N	0.0000000	0.0000000	-0.5503900
Formaldehyde (a.u.) from Ref. 11			
C	0.0000000	0.0000000	1.1442483
O	0.0000000	0.0000000	-1.1415137
H	0.0000000	1.7698527	2.2461407
H	0.0000000	-1.7698527	2.2461407
Hydrogen Cyanide (Å) from Ref. 11			
H	0.0000000	0.0000000	1.0680000
C	0.0000000	0.0000000	0.0000000
N	0.0000000	0.0000000	-1.1560000
Hydrogen Isocyanide (a.u.) from Ref. 12			
N	0.0000000	0.0000000	0.431366
C	0.0000000	0.0000000	-0.740682
H	0.0000000	0.0000000	1.424529
Nitroxyl (a.u.) from Ref. 11			
N	-1.1543258	-0.1614981	0.0000000
O	1.1319819	0.0357051	0.0000000
H	-1.9267677	1.6772444	0.0000000
Fluoromethane (a.u) from Ref. 11			
F	1.2127043	0.0000000	0.0000000
C	-1.3975536	0.0000000	0.0000000
H	-2.0733764	0.9730934	-1.6854471
H	-2.0733764	-1.9461867	0.0000000
H	-2.0733764	0.9730934	1.6854471

Ketene (Å) from Ref. 10			
C	0.0000000	0.0000000	-1.2954795
C	0.0000000	0.0000000	0.0185135
O	0.0000000	0.0000000	1.1835785
H	0.0000000	0.9389301	-1.8188138
H	0.0000000	-0.9389301	-1.8188138

S5.2 Representative Applications

Table S15: Ethyl Trifluoroacetate (Å), aug-cc-pVTZ/MP2

anti-anti Conformer			
C	1.357858	-0.218256	0.000000
F	2.400882	0.558853	0.000000
C	0.058629	0.610200	0.000000
O	0.060562	1.784722	0.000000
O	-0.976264	-0.176763	0.000000
C	-2.275883	0.425209	0.000000
C	-3.296073	-0.686593	0.000000
H	-4.291642	-0.259911	0.000000
H	-3.188619	-1.308775	-0.878808
H	-3.188619	-1.308775	0.878808
H	-2.360919	1.051580	0.875722
H	-2.360919	1.051580	-0.875722
F	1.409583	-0.984896	-1.062925
F	1.409583	-0.984896	1.062925

Table S16: Cytosine (Å), aug-cc-pVTZ/CCSD

Tautomer 1			
C	1.047109	-1.182624	-0.000096
C	-0.196676	-1.705671	0.000087
N	-1.276704	-0.887554	0.000259
C	-1.184372	0.517864	0.000736
N	0.081589	1.051680	0.000183
C	1.122714	0.257421	-0.000054
N	2.343661	0.839163	-0.002104
H	2.392646	1.840837	0.003224
H	3.183793	0.297212	0.005501
O	-2.208784	1.166824	-0.000610
H	-2.211634	-1.256381	0.000120
H	-0.392256	-2.768189	-0.000093
H	1.926396	-1.805693	-0.000714
Tautomer 2a			
C	1.133587	-0.260295	0.000000
C	1.108261	1.146011	0.000000
C	-0.138836	1.727919	0.000000
N	-1.289777	1.038757	0.000000
C	-1.130229	-0.273068	0.000000
N	0.005380	-0.972181	0.000000
O	-2.261946	-0.993106	0.000000
H	-1.994993	-1.918045	0.000000
H	-0.236827	2.806628	0.000000
H	2.013202	1.734006	0.000000
N	2.293962	-0.960228	0.000000
H	2.255186	-1.961049	0.000000
H	3.179851	-0.497138	0.000000
Tautomer 2b			
C	1.149196	-0.248218	0.000000
C	1.088399	1.159870	0.000000
C	-0.171036	1.709026	0.000000
N	-1.297868	0.979208	0.000000
C	-1.101116	-0.333460	0.000000
N	0.046917	-0.998498	0.000000
O	-2.196184	-1.107557	0.000000
H	-2.946519	-0.503487	0.000000
H	-0.302682	2.783849	0.000000
H	1.977788	1.771128	0.000000
N	2.328949	-0.913094	0.000000
H	2.314216	-1.914961	0.000000
H	3.202145	-0.426571	0.000000

Tautomer 3			
C	1.266040	-0.330491	0.000000
C	1.165048	1.126942	0.000000
C	-0.042971	1.707812	0.000000
N	-1.205514	0.967313	0.000000
C	-1.219279	-0.415212	0.000000
N	0.029648	-0.982883	0.000000
H	0.051052	-1.990106	0.000000
O	-2.247168	-1.055814	0.000000
H	-2.107202	1.407382	0.000000
H	-0.174573	2.779580	0.000000
H	2.066156	1.718204	0.000000
N	2.315358	-1.056510	0.000000
H	3.144112	-0.471187	0.000000

Table S17: Thymine (Å), aug-cc-pVDZ/CCSD(T) from Ref. 13.

Franck-Condon point minimum			
C	-1.6332	0.0426	-0.0001
C	0.7570	0.8384	-0.0001
C	1.1979	-0.5713	0.0000
C	0.2400	-1.5432	0.0000
C	2.6817	-0.8510	0.0002
N	-0.6415	1.0215	-0.0007
N	-1.1219	-1.2509	-0.0002
O	-2.8353	0.2919	0.0005
O	1.5118	1.8105	0.0003
H	-0.9734	1.9835	-0.0002
H	0.4869	-2.6087	0.0000
H	-1.8153	-1.9901	0.0005
H	3.1589	-0.4028	-0.8884
H	3.1587	-0.4025	0.8887
H	2.8721	-1.9376	0.0005

Table S18: Thymine (Å), aug-cc-pVDZ/CCSD from Ref. 13.

$n\pi^*$ minimum			
C	1.6756	0.0108	0.0230
C	-0.6393	0.8270	-0.0428
C	-1.1709	-0.4819	-0.0168
C	-0.2637	-1.5185	0.0106
C	-2.6680	-0.6951	-0.0181
N	0.7453	1.0515	0.1200
N	1.1227	-1.2492	0.0249
O	2.8776	0.2283	-0.0350
O	-1.3720	1.9491	0.1211
H	1.1236	1.9641	-0.1066
H	-0.5521	-2.5678	-0.0154
H	1.7941	-2.0025	-0.0250
H	-3.1289	-0.2542	0.8817
H	-3.1341	-0.2281	-0.9015
H	-2.8981	-1.7710	-0.0329
$\pi\pi^*$ saddle-point			
N	0.6451	-1.0681	-0.0013
H	0.9407	-2.0373	-0.0009
C	1.6033	-0.1081	0.0001
O	2.8197	-0.2668	0.0006
N	1.1090	1.2319	0.0005
H	1.8635	1.9158	0.0010
C	-0.1892	1.6224	-0.0010
H	-0.4105	2.6884	-0.0014
C	-0.7824	-0.8153	-0.0001
O	-1.5291	-1.8220	0.0005
C	-1.1881	0.5501	-0.0002
C	-2.6474	0.8865	0.0006
H	-2.8068	1.9771	-0.0001
H	-3.1465	0.4513	0.8852
H	-3.1475	0.4502	-0.8830

Table S19: Neutral and Cationic Radicals (Å), Experimental geometry, from Ref. 14 (sourced from Ref. 15).

CH_3			
C	0.0000	0.0000	0.0000
H	1.0790	0.0000	0.0000
H	-0.5395	-0.9344	0.0000
H	-0.5395	0.9344	0.0000
CH_3CH_2			
C	-0.0118	-0.6932	0.0000
C	-0.0118	0.7930	0.0000
H	1.0059	-1.0974	0.0000
H	-0.5089	-1.0936	0.8816
H	-0.5089	-1.0936	-0.8816
H	0.0769	1.3430	-0.9217
H	0.0769	1.3430	0.9217
t-Butyl			
C	0.0000	0.0000	0.1941
C	0.0000	1.4748	-0.0201
C	1.2773	-0.7374	-0.0201
C	-1.2773	-0.7374	-0.0201
H	0.0000	1.7206	-1.0909
H	1.4901	-0.8603	-1.0909
H	-1.4901	-0.8603	-1.0909
H	-0.8843	1.9428	0.4116
H	0.8843	1.9428	0.4116
H	2.1247	-0.2056	0.4116
H	1.2404	-1.7372	0.4116
H	-1.2404	-1.7372	0.4116
H	-2.1247	-0.2056	0.4116
Allyl			
C	0.0000	0.0000	0.0000
H	0.0000	1.0690	0.0000
C	-1.2643	-0.6638	0.0000
C	1.2643	-0.6638	0.0000
H	-2.1768	-0.1068	0.0000
H	2.1768	-0.1068	0.0000
H	-1.2643	-1.7328	0.0000
H	1.2643	-1.7328	0.0000
CO^+			
C	0.0000	0.0000	0.0000
O	0.0000	0.0000	1.1152

NH_2			
N	0.0000	0.0000	0.0000
H	0.0000	0.8036	0.6347
H	0.0000	-0.8036	0.6347
N_2^+			
N	0.00000	0.00000	0.55821
N	0.00000	0.00000	-0.55821
OH			
H	0.0000	0.0000	0.0000
O	0.0000	0.0000	0.9697
HO_2			
O	0.0553	-0.6132	0.0000
O	0.0553	0.7198	0.0000
H	-0.8853	-0.8528	0.0000

Table S20: Neutral and Cationic Radicals (Å), cc-pVTZ/MP2, from Ref. 14.

i-Propyl			
C	-0.0168860854	0.5472979108	0.0000000000
C	-0.0095830121	-0.1988128213	1.2865253572
C	-0.0095830121	-0.1988128213	-1.2865253572
H	0.2685769360	1.5881742923	0.0000000000
H	-0.7587444280	-0.9922619543	1.2836398162
H	-0.7587444281	-0.9922619543	-1.2836398162
H	0.9567932274	-0.6859518581	1.4643668216
H	-0.2045112126	0.4507905322	2.1369563305
H	0.9567932274	-0.6859518581	-1.4643668216
H	-0.2045112125	0.4507905322	-2.1369563305
CH_2Br			
Br	0.000000000	0.000000000	0.000000000
C	1.836138000	0.000000000	0.000000000
H	2.340495690	-0.946867198	0.000000000
H	2.340495690	0.946867198	0.000000000
CH_2Cl			
Cl	0.000000000	0.000000000	0.000000000
C	1.693600000	0.000000000	0.000000000
H	2.195099781	-0.948199572	0.000000000
H	2.195099781	0.948199572	0.000000000

NH_3^+			
N	0.000000000	0.000000000	0.000000000
H	1.019162000	0.000000000	0.000000000
H	-0.509581616	0.882619827	0.000000000
H	-0.509581616	-0.882619819	0.000115381

Table S21: Cationic Benzene Radical (Å), cc-pVTZ/EOM-IP-CCSD, from Ref. 16.

Bz^+			
C	-1.2392908131	-2.1390691708	-0.0000000000
C	-0.7225974533	-1.1894682957	-0.0000000000
C	0.7225974533	-1.1894682957	-0.0000000000
H	1.2392908131	-2.1390691708	-0.0000000000
C	1.4341864916	-0.0000000000	-0.0000000000
H	2.5128824541	-0.0000000000	-0.0000000000
C	0.7225974533	1.1894682957	0.0000000000
H	1.2392908131	2.1390691708	0.0000000000
C	-0.7225974533	1.1894682957	0.0000000000
H	-1.2392908131	2.1390691708	0.0000000000
C	-1.4341864916	0.0000000000	0.0000000000
H	-2.5128824541	0.0000000000	0.0000000000

Table S22: Anionic Superoxide Radical (Å), Experimental, from Ref. 15.

O_2^-			
O	0.0000	0.0000	0.0000
O	0.0000	0.0000	1.3500

References

(1) Plekan, O.; Feyer, V.; Richter, R.; Coreno, M.; De Simone, M.; Prince, K.; Trofimov, A.; Gromov, E.; Zaytseva, I.; Schirmer, J. A theoretical and experimental study of the near edge X-ray absorption fine structure (NEXAFS) and X-ray photoelectron spectra (XPS) of nucleobases: Thymine and adenine. *Chemical Physics* **2008**, *347*, 360–375.

(2) Grimme, S.; Waletzke, M. A combination of Kohn–Sham density functional theory and multi-reference configuration interaction methods. *The Journal of chemical physics* **1999**, *111*, 5645–5655.

(3) Costain, T. S.; Ogden, V.; Neville, S. P.; Schuurman, M. S. A DFT/MRCI Hamiltonian parameterized using only ab initio data: I. valence excited states. *The Journal of Chemical Physics* **2024**, *160*, 224106.

(4) Neville, S. P.; Schuurman, M. S. Removing the deadwood from DFT/MRCI wave functions: The p-DFT/MRCI method. *Journal of Chemical Theory and Computation* **2021**, *17*, 7657–7665.

(5) Neville, S. P.; Schuurman, M. S. A perturbative approximation to DFT/MRCI: DFT/MRCI (2). *The Journal of Chemical Physics* **2022**, *157*, 164103.

(6) Feyer, V.; Plekan, O.; Richter, R.; Coreno, M.; Vall-Llosera, G.; Prince, K. C.; Trofimov, A. B.; Zaytseva, I. L.; Moskovskaya, T. E.; Gromov, E. V.; others Tautomerism in cytosine and uracil: An experimental and theoretical core level spectroscopic study. *The Journal of Physical Chemistry A* **2009**, *113*, 5736–5742.

(7) 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. *The Journal of Physical Chemistry A* **2010**, *114*, 10270–10276.

(8) Gelius, U.; Basilier, E.; Svensson, S.; Bergmark, T.; Siegbahn, K. A high resolution ESCA instrument with X-ray monochromator for gases and solids. *Journal of Electron Spectroscopy and Related Phenomena* **1973**, *2*, 405–434.

(9) Trygubenko, S. A.; Bogdan, T. V.; Rueda, M.; Orozco, M.; Luque, F. J.; Šponer, J.; Slavíček, P.; Hobza, P. Correlated ab initio study of nucleic acid bases and their tautomers in the gas phase, in a microhydrated environment and in aqueous solution Part 1. Cytosine. *Physical Chemistry Chemical Physics* **2002**, *4*, 4192–4203.

(10) Loos, P.-F.; Scemama, A.; Blondel, A.; Garniron, Y.; Caffarel, M.; Jacquemin, D. A mountaineering strategy to excited states: Highly accurate reference energies and benchmarks. *Journal of chemical theory and computation* **2018**, *14*, 4360–4379.

(11) Simons, M.; Matthews, D. A. Transition-potential coupled cluster. *The Journal of Chemical Physics* **2021**, *154*, 014106.

(12) Fransson, T.; Brumboiu, I. E.; Vidal, M. L.; Norman, P.; Coriani, S.; Dreuw, A. XABOOM: An X-ray absorption benchmark of organic molecules based on carbon, nitrogen, and oxygen 1s$\rightarrow \pi^*$ transitions. *Journal of Chemical Theory and Computation* **2021**, *17*, 1618–1637.

(13) Wolf, T.; Myhre, R. H.; Cryan, J.; Coriani, S.; Squibb, R.; Battistoni, A.; Berrah, N.; Bostedt, C.; Bucksbaum, P.; Coslovich, G.; others Probing ultrafast $\pi\pi^*$/nπ^* internal conversion in organic chromophores via K-edge resonant absorption. *Nature communications* **2017**, *8*, 29.

(14) Hait, D.; Haugen, E. A.; Yang, Z.; Oosterbaan, K. J.; Leone, S. R.; Head-Gordon, M. Accurate prediction of core-level spectra of radicals at density functional theory cost via square gradient minimization and recoupling of mixed configurations. *The Journal of chemical physics* **2020**, *153*.

(15) Johnson III, R. CCCBDB computational chemistry comparison and benchmark database. *NIST Standard Reference Database* **1999**,

(16) Vidal, M. L.; Epshtein, M.; Scutelnic, V.; Yang, Z.; Xue, T.; Leone, S. R.; Krylov, A. I.; Coriani, S. Interplay of open-shell spin-coupling and Jahn–Teller distortion in benzene radical cation probed by x-ray spectroscopy. *The Journal of Physical Chemistry A* **2020**, *124*, 9532–9541.