Electronic structure of the two isomers of the anionic form of *p*-coumaric acid chromophore

Dmitry Zuev,[1] Ksenia B. Bravaya,[1] T. Daniel Crawford,[2] Roland Lindh,[3] and Anna I. Krylov[1,a)]

[1]*Department of Chemistry, University of Southern California, Los Angeles, California 90089-0482, USA*
[2]*Department of Chemistry, Virginia Tech, Blacksburg, Virginia 24061, USA*
[3]*Department of Physical and Analytical Chemistry, Uppsala University, Uppsala SE-751 20, Sweden*

(Received 14 September 2010; accepted 25 October 2010; published online 20 January 2011)

A theoretical study of the electronic structure of the photoactive yellow protein (PYP) model chromophore, *p*ara-coumaric acid (*p*-CA), is presented. Electronically excited states of the phenolate and carboxylate isomers of the deprotonated *p*-CA are characterized by high-level *ab initio* methods including state-specific and multistate multireference pertrubation theory (SS-CASPT2, and MS-CASPT2), equation-of-motion coupled-cluster methods with single and double substitutions (EOM-CCSD) and with an approximate account of triple excitations (CC3). We found that the two isomers have distinctly different patterns of ionization and excitation energies. Their excitation energies differ by more than 1 eV, in contradiction to the experimental report [Rocha-Rinza *et al.,* J. Phys. Chem. A **113**, 9442 (2009)]. The calculations confirm metastable (autoionizing) character of the valence excited states of both phenolate and carboxylate isomers of *p*-CA^- in the gas phase. The type of resonance is different in the two forms. In the phenolate, the excited state lies above the detachment continuum (a shape resonance), whereas in the carboxylate the excited $\pi \rightarrow \pi^*$ state lies below the π-orbital ionization continuum, but is above the states derived from ionization from three other orbitals (Feshbach resonance). The computed oscillator strength of the bright electronic state in the phenolate is higher than in the carboxylate, in agreement with Hückel's model predictions. The analysis of photofragmentation channels shows that the most probable products for the methylated derivatives of the phenolate and carboxylate forms of *p*-CA^- are CH_3, CH_2O and CH_3, CH_2O, CO_2, respectively, thus suggesting an experimental probe that may discriminate between the two isomers.
 [doi:10.1063/1.3516211]

I. INTRODUCTION

Photoactive yellow protein (PYP) is a small bacterial photoreceptor found in Halorhodospira halophila,[1,2] which is responsible for the negative phototaxis of its host bacteria in response to blue light.[3] The PYP chromophore is a prosthetic group of *para*-coumaric acid (*p*-CA) linked to the protein by the thioester covalent bond.[4,5] Inside the protein in the dark-adapted state, the chromophore exists in the anionic (deprotonated) form.[4,5] Absorption of blue light [λ_{max} = 446 nm or 2.78 eV(Ref. 3)] triggers the chromophore's *trans–cis* photoisomerization, which, in turn, initiates the cascade of processes leading to signal transduction.[6,7]

To understand the complex photochemistry of photoactive proteins, it is desirable to distinguish between the intrinsic properties of the chromophore and effects due to chromophore–protein interactions, which may affect both the optical properties and the response to light absorption, e.g., dynamics on excited-state potential energy surfaces. Therefore, the properties of excited states of the unperturbed isolated chromophores are of particular interest. Several synthetic analogs of the PYP chromophore have been studied experimentally in gas phase.[8–11] Dynamics of the gas-phase photoisomerization of the 4-(3-oxobut-1-enyl)-phenolate anion (P^-) was studied using femtosecond mass-selection/electron detachment techniques.[11] This study demonstrated that the protein environment plays a twofold role, i.e., (i) impeding electron detachment from the chromophore and (ii) directing the photoreaction toward the *cis*-photoproduct.[11] Based on these experiments, the vertical detachment energy for the P^- chromophore was estimated to be 2.9 eV.[11]

The gas-phase absorption spectra of different model PYP chromophores including *p*-CA^-, methylated *p*-CA^- analogs, and *trans*-thiophenyl-*p*-coumarate (*p*-CT^-) have been measured using action spectroscopy coupled with ion-storage ring and electrospray techniques.[8–10] An interesting feature of these anionic systems is that the photoabsorption initiates two competing processes, i.e., photodetachment and fragmentation. Consequently, the photoabsorption bands obtained by action spectroscopy include signal from: detachment, excitation followed by fragmentation, and autoionization.

Determination of the relative cross sections of these processes is a challenging task and is still an open problem. Therefore, the interpretation of the action spectroscopy photoabsorption spectra requires caution. An attempt to distinguish between the detachment and excitation channels for the

a)Author to whom correspondence should be addressed. Electronic mail: krylov@usc.edu.

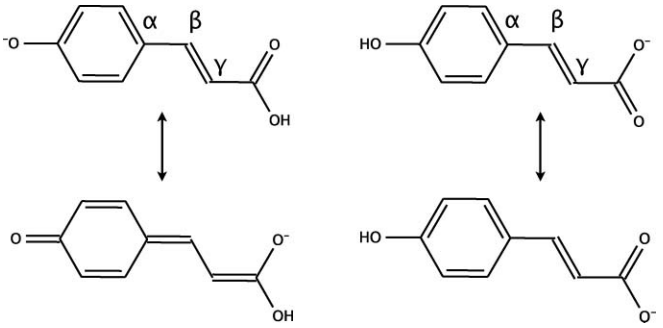

FIG. 1. The resonance structures of the model PYP chromophore, *p*-CA$^-$, in the phenolate (left) and carboxylate (right) forms.

anionic chromophore of the green fluorescent protein (GFP) has been reported by Frobes and Jockusch.[12] Although a fulfilling completely different physiological function, the GFP chromophore features remarkably similar electronic structure patterns, e.g., the lowest excited state of its anionic form is also an autoionizing resonance state.[13,14]

The *p*-CA anion serves as a minimal model of the PYP chromophore. Two *p*-CA$^-$ isomers that have been studied experimentally[10] are the phenolate and carboxylate forms shown in Fig. 1. The carboxylate is believed to be the more stable isomer in polar solvents,[10] whereas in the gas phase, the phenolate isomer is 13 kcal/mol lower in energy.[10] Their relative stability can be explained by different charge distribution patterns of the two isomers discussed in Sec. III A—the more localized charge on carboxylate can be solvated more efficiently, whereas in the gas phase the isomer with more delocalized excess charge becomes more stable.

In solution, isomerization between these two tautomers may occur. Since it is not clear whether gas-phase equilibrium conditions are reached in electrospray, methyl-substituted ethers of the *p*-CA$^-$ isomers have been used in the experiment.[10] Although the absorption spectra of the two isomers in solution are considerably different (4.40 and 3.49 eV for carboxylate and phenolate, respectively[10]), their gas-phase absorption maxima were reported to be identical (2.88 eV).

The absorption spectra of *p*-CA$^-$ have attracted considerable attention from theory for the following reasons: (i) it was found that the excited state of the phenolate form of the anion is a resonance state lying above the detachment continuum;[15,16] and (ii) there are large discrepancies in excitation energies computed by different quantum chemistry methods[10,17] as well as between theoretical and experimental absorption maxima.[10] Thus, *p*-CA$^-$ features a complex electronic structure characteristic of closed-shell anions and presents an interesting benchmark system for *ab initio* methods.

There are numerous theoretical studies on the absorption of the isolated PYP chromophores.[10,15–18] Gromov *et al.* have studied absorption of the gas-phase PYP chromophore in neutral[15] and anionic states.[18] However, when considering the anionic chromophore, the authors mainly focused on the phenolate, the biologically relevant form. In line with early evidence of the resonance character of the first excited state of the PYP chromophore based on the phenolate anion model,[16] Gromov *et al.* have demonstrated that the lowest excitation energy of a realistic chromophore (3.17–3.18 eV, EOM-CCSD) is indeed above the detachment continuum (2.51–2.90 eV, OVGF).[18,19] The conclusion was further supported by other theoretical results for *p*-CA$^-$.[10] Ma *et al.*[17] used many-body Green's function theory (MBGFT) to study excited states of the the carboxylate and phenolate forms of *p*-CA$^-$. They reported excitation energies of 2.95 and 4.37 eV for phenolate and carboxylate, respectively.[17] As a possible explanation of the experimental results, the authors suggested that the only form present in the gas phase is phenolate.[17] Since the methylated species were used in the experiment,[10] this conclusion seems questionable. Bochenkova and coworkers have reported excitation energies for the carboxylate and phenolate forms of *p*-CA$^-$ that are within 0.1 eV of the experimental peaks obtained with the augmented multiconfigurational quasidegenerate perturbation theory technique (aug-MCQDPT2).[10] At the same time, excitation energies computed using the approximate coupled-cluster doubles scheme (CC2, 4.79 and 3.10 eV for carboxylate and phenolate *p*-CA$^-$, respectively) (Ref. 10) are in agreement with the MBGFT results. Thus, either the absorption spectra of the deprotonated *p*-CA present a very challenging problem for modern quantum chemical methods or the experimental results require reinterpretation.

This study presents electronic structure calculations of the two *p*-CA$^-$ isomers. We characterized the lowest excited and ionized states of the chromophore using high-level *ab initio* methods including EOM-CCSD,[20–24] state-specific and multistate multireference perturbation theory (SS-CASPT2 and MS-CASPT2),[25] and a coupled-cluster method with an approximate account of triple excitations, CC3.[26] We rationalize the observed difference in the absorption spectra of the two molecules on the basis of the Hückel model. We also analyze the photodissociation pathways and suggest the formation of the CO_2 photofragment as an experimental probe for the carboxylate isomer of *p*-CA$^-$.

II. COMPUTATIONAL DETAILS

The equilibrium geometries of both isomers were optimized using MP2 with the aug-cc-pVDZ (Ref. 27) basis set. The resolution-of-the-identity (RI) technique[28–31] was employed. C_s symmetry was imposed during the geometry optimization. The following convergence thresholds were used in the optimization procedure: 1×10^{-6} hartree for the energy, 3×10^{-4} hartree/Å for the energy gradient, and 1.2×10^{-3} Å for displacements.

Geometry optimizations of neutral radicals were performed with the long-range corrected ωB97X functional[32] using the aug-cc-pVDZ basis and with the IP-CISD (configuration interaction with single and double substitutions for ionized states) (Ref. 33) method with the 6-31+G(d,p) basis set. IP-CISD is an approximation to EOM-IP-CCSD (EOM-CCSD for ionized states),[20,34–38] which scales as N^5 and employs an uncorrelated Hartree–Fock determinant as a reference instead of the CCSD wave function. For comparison purposes and zero-point energy (ZPE) calculations, the structures of the closed-shell anions have been also

reoptimized with ωB97X/6-31+G(d,p). The absence of imaginary frequencies obtained by ωB97X/6-31+G(d,p) verify that the stationary points are indeed true minima, except the planar C_s structure for carboxylate, which was found to be a transition state with a single imaginary frequency corresponding to the rotation along the single bond at the bridge. This is probably caused by steric repulsion between the hydrogen atoms at the bridge C_β atom and the phenol ring. The structure reoptimized without the C_s symmetry constraint is nonplanar with a $C–C_\alpha–C_\beta–C_\gamma$ torsional angle of 14.5°. The energy difference between the two structures is minor [0.04 kcal/mol, ωB97X/6-31+G(d,p)]. We do not anticipate significant effects of slight nonplanarity on excitation energies and employ the planar geometry in all calculations. A higher torsional rigidity of the phenolate isomer can be explained by the two resonance structures (Fig. 1) resulting in the allylic character of the bridge moiety, and, consequently, partial double-bond character of $C_\alpha–C_\beta$.

Dissociation energies (D_e) for different fragmentation channels were computed with ωB97X/6-311+,+G(2df,2pd) as the difference between ground-state energies of the initial molecule and the dissociation products. ZPE corrections were calculated with ωB97X/6-31+G(d,p) at the geometries reoptimized at the same level of theory. To analyze conformational flexibility of the chromophore, we also performed *ab initio* molecular dynamics (AIMD) simulations using the B3LYP functional and the 6-31+G(d,p) basis set. The trajectories were propagated for 2 ps with the time-step of 0.5 fs. The grid used for all density functional theory (DFT) calculations contained 75 points in the Lebedev[39] radial grid and 302 points in the Euler–Maclaurin[40] angular grid.

Vertical excitation energies were computed using EOM-EE-CCSD (EOM-CCSD for excitation energies),[20–22,24] CC3,[26] and CASPT2.[25] Vertical detachment energies were computed by using Koopmans' theorem (i.e., negatives of orbital energies) and by EOM-IP-CCSD (Refs. 34, 35, and 37) with the 6-311+G(df,pd) basis.

The EOM-CCSD error bars are 0.1–0.3 eV for electronic states dominated by single excitations. Including triples reduces the errors to 0.01–0.02 eV.[41] In a recent benchmark study, Schreiber *et al.*[42] reported EOM-CCSD mean absolute and maximum errors of 0.12 and 0.23 eV, respectively. A recent study of uracil[43] demonstrated that even for well-behaved molecules inclusion of triple excitations and extending the basis set beyond augmented double-zeta can affect vertical excitations by as much as 0.3 eV. The CC3 method,[26] which is an iterative CC method with an approximate inclusion of triple excitations, has been shown to reduce the EOM-CCSD error bars down to 0.016 eV (maximum error) for singly excited states.[42]

The multireference calculations were performed with the complete active space self-consistent field (CASSCF) method[44] to account for near-degeneracies of different electronic configurations followed by multireference second-order perturbation theory (CASPT2) (Ref. 45) to include dynamical correlation. These calculations were performed with the 6-31G(d,p),[46] ANO-RCC-VDZP and ANO-RCC-VTZP (Refs. 47 and 48) basis sets. The ANO-RCC bases were used together with the Douglas–Kroll Hamiltonian[49] relativistic correction. For the first two rows of the periodic table, the ANO-RCC basis sets perform similarly to the ANO-L non-relativistic bases.

The active space was designed to include all π-orbitals antisymmetric with respect to the molecular plane for accurate description of the lowest $\pi \rightarrow \pi^*$ excitations. The resulting active space included 14 electrons in 12 orbitals. The state-averaged CASSCF (SA-CASSCF) approach was used with equal weight on the six lowest states. In subsequent CASPT2 and MS-CASPT2 (Ref. 25) calculations, the standard IPEA shift of 0.25 (Ref. 50) was employed and the 1s core orbitals of the second row elements were frozen in the calculations of the dynamical correlation correction. Oscillator strengths were computed using the complete active space state interaction (CASSI) algorithm.[51] The calculations employed the recently developed Cholesky decomposition (CD) methods to handle the two-electron integrals.[52,53] The calculations used the so-called atomic compact CD (acCD) auxiliary basis set[54,55] (generated with a CD threshold of 10^{-4}) along with the CD-CASSCF[56] and CD-CASPT2[57] implementations. In the CD-CASSCF implementation the local exchange approximation[58] was employed. The errors in excitation energies introduced by these approximations are less than 0.001 eV, as demonstrated in the recent benchmark study.[59]

The Cartesian geometries, relevant energies, and harmonic frequencies are provided in Supporting Materials.[60] The wave function analysis was performed using natural bond orbitals (NBO) procedure.[61,62] The EOM-CC and DFT calculations were performed with Q-Chem.[63] The CC3 and CASPT2 calculations were performed using PSI III[64] and MOLCAS,[65] respectively.

III. RESULTS AND DISCUSSION

A. Structures and charge distributions of *p*-CA$^-$

We considered several rotamers of the *p*-CA$^-$ anion derived by rotation along the phenolic C–O (carboxylate), C_γ-C(OOH) and C–O(H) (phenolate) bonds. All excitation and ionization energies for the phenolate are reported for the lowest energy structure with the C_γ-C(OOH) and C–O(H) bonds in anti and syn configurations, respectively. For the carboxylate, all results are obtained for the lowest energy syn-OH rotamer, except CASPT2 and CC3 excitation energies. CASPT2 and CC3 calculations were performed for the anti-OH rotamer, which is slightly higher in energy (by 0.06 and 0.55 kcal/mol at the RI-MP2/aug-cc-pVDZ and CCSD/6-31+G(d,p) levels, respectively; see Supporting Materials[60]).

The difference in the bond length alternation (BLA) patterns in the carboxylate and phenolate structures can be qualitatively explained by the analysis of leading resonance structures.[10] The two dominant resonance forms of the phenolate isomer are: (i) enol with the negative charge hosted by the –COOH fragment; and (ii) phenol with the charge localized on the phenolate oxygen. BLA at the bridging C–C bonds ($C_\alpha–C_\beta–C_\gamma$) for the phenol-like resonance structure corresponds to the single $C_\alpha–C_\beta$ and double $C_\beta–C_\gamma$ bonds (Fig. 1), whereas the enol-like resonance structure has an

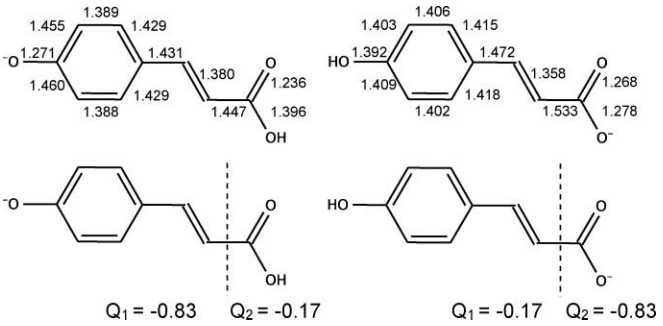

FIG. 2. Relevant geometric parameters (RI-MP2/aug-cc-pVDZ) and NBO charge distributions for the ground electronic state of the phenolate (left) and carboxylate (right) forms of *p*-CA^-.

opposite pattern. The two dominant resonance structures of carboxylate only affect the carboxylate moiety and do not disturb BLA in the bridge region. Thus, C_α–C_β is a single bond, and C_β–C_γ is a double bond.

The optimized geometries and the NBO charge distributions of both isomers are shown in Fig. 2. To quantify the charge distribution, we divide the molecule into two parts, as shown in Fig. 2. In the carboxylate, the excess charge is hosted by the CO_2 group, whereas in the phenolate the charge is located mainly on the phenol ring and the bridge and is more delocalized. This explains preferential deprotonation of the carboxylate in polar solutions, as solvation of the more compact carboxylate anion is more efficient. Although the charge distribution in phenolate does not reveal significant contribution of the enol resonance form, its signature can be found in the optimized geometrical parameters. The resonance analysis above predicts that the difference between the C_α–C_β and C_β–C_γ bond lengths should be larger in the carboxylate isomer, and this is indeed the case—the difference in the bond lengths is almost twice larger in carboxylate (0.114 Å) than in phenolate (0.051 Å) revealing notable contribution of the enol resonance structure. The latter effect is even more important for the anionic form of the GFP chromophore[14,66,67] where the two chromophore moieties (phenolic and imidazolone rings) are more similar.

B. *Ab initio* calculations of the electronically excited and ionized states of *p*-CA^-

As a closed-shell system, the *p*-CA^- anion is stable in the gas phase and has a relatively large vertical detachment energy (VDE). Detachment energies calculated by EOM-IP-CCSD are summarized in Table I. The computed VDE for the phenolate isomer agrees well with the experimental value of 2.9 eV reported for a similar model PYP chromophore.[11] Interestingly, Koopmans' theorem fails to predict the correct ordering of the ionized states in the carboxylate. According to Koopmans' theorem, the lowest ionization corresponds to electron removal from the highest occupied molecular orbital (HOMO) (Fig. 3), which is a π-like orbital delocalized over the phenol ring and the bridge, whereas the negative charge is mainly located on the carboxylate group and one could expect ionization from the latter moiety. Electron correlation changes the ordering of the ionized states and the lowest ionized state at the EOM-IP-CCSD level indeed corresponds to electron detachment from the carboxylate orbital (HOMO-3) formed by oxygen lone pairs (see Fig. 3). The differences between Koopmans and EOM-IP-CCSD VDEs range from 0.3–2.2 eV. Thus, correlation is required for both quantitative accuracy and for determining the correct state ordering. All ionized states have predominantly single-configurational Koopmans character (the leading R_1 amplitude is greater than 0.95).

TABLE I. Vertical detachment energies (VDE, eV) for the two *p*-CA^- isomers estimated by Koopmans theorem (KT, eV) and computed with EOM-IP-CCSD/6-311+G(df,pd)//RI-MP2/aug-cc-pVDZ. EOM-IP-CCSD/6-311+G(df,pd) adiabatic detachment energies (ADE, eV) computed using ωB97X/aug-cc-pVDZ optimized geometries of the neutrals for the first ionized state and IP-CISD/6-31+G(d,p) for the subsequent ones are also given.

Carboxylate				
Target state	Orbital	KT	VDE	ADE
$1^2A'$	n_1	6.08	3.91	3.54
$2^2A'$	n_2	6.01	4.07	–
$1^2A''$	π_1	5.54	4.18	3.99
$2^2A''$	π_2	4.72	4.75	4.57
Phenolate				
Target state	Orbital	KT	VDE	ADE
$1^2A''$	π_3	3.26	2.92	2.72
$1^2A'$	n_{ph}	6.68	4.54	–
$2^2A''$	π_2	5.84	5.60	–

The basis set effects on computed VDEs were analyzed by using several bases in the EOM-IP-CCSD calculations of the carboxylate isomer. VDEs for the first ionized state computed with 6-311+G(d,p), 6-311(2+,+)G(d,p), and 6-311+G(df,pd) are 3.80, 3.80, and 3.91 eV, respectively. Thus, the addition of diffuse functions has only minor effect on VDEs, whereas the effect of polarization is more pronounced.

The relaxation energies (VDE-AIE) are 0.4 eV and 0.2 eV for the first ionized state of carboxylate and phenolate, respectively. The comparison of the anionic and neutral states' geometries is presented in Fig. 4. The observed trends can be explained by the analysis of the corresponding molecular orbitals (MOs). The lowest ionized state of phenolate corresponds to electron removal from the conjugated π-system. The HOMO is bonding with respect to the three phenol carbon atoms, the C_α–C_β bridge and C_γ–C(OOH) bonds (Fig. 3). Thus, detachment from the HOMO leads to the increase in these bond lengths. In contrast, the HOMO has antibonding character with respect to phenolic O–C and C–O in the carboxyl group, and ionization results in bond length contraction (Fig. 4).

Lowest ionized state of the carboxylate corresponds to detachment from the HOMO-3, which is a σ-like orbital (Fig. 3). Electron detachment from the negatively charged COO group in the carboxylate reduces electron repulsion between the phenolic and carboxyl moieties thus reducing the C_γ–C(OO) bond length. Ionization also detunes the resonance in the carboxyl group; the BLA patterns in the carboxyl group of the ionized carboxylate and phenolate isomers are very similar (Fig. 4).

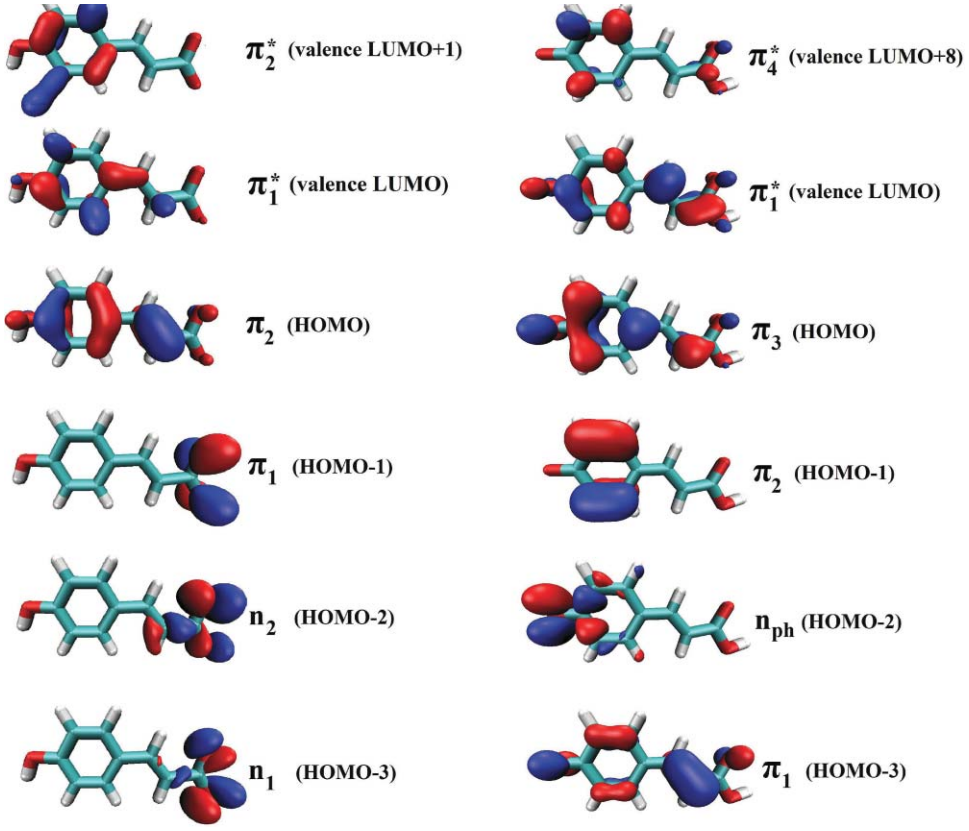

FIG. 3. Relevant MOs in the 6-31+G(d,p) basis. The π_3^* orbital for carboxylate and π_2^*, π_3^* orbitals for phenolate (not shown) are of diffuse character.

Despite relatively high detachment energy, both phenolate and carboxylate isomers do not support bound electronically excited singlet states, and the lowest valence excitation is embedded in the detachment continuum (Tables I–III). Such resonance states are common in molecular anions[13,68] and play an important role in dissociative electron attachment processes.[69,70] A complete description of the resonance excited states requires taking into account their interactions with the continuum. Several techniques can be used to tackle this problem, such as stabilization[71] and complex Hamiltonian methods including complex absorption potential[72,73] and complex-scaling methods.[74–77] In traditional electronic structure calculations, increasing the basis set results in low-lying excited states corresponding to excitation to diffuse orbitals approximating the continuum. These continuum-like states can mix with the metastable valence excited state (see, for example, Ref. 13) presenting an obstacle for obtaining converged (with respect to the one-electron basis set) results. By using bases with moderate diffuse character we enforce the localization of the excited state preventing its mixing with the diffuse continuumlike states, which can be thought of as approximate diabatization.

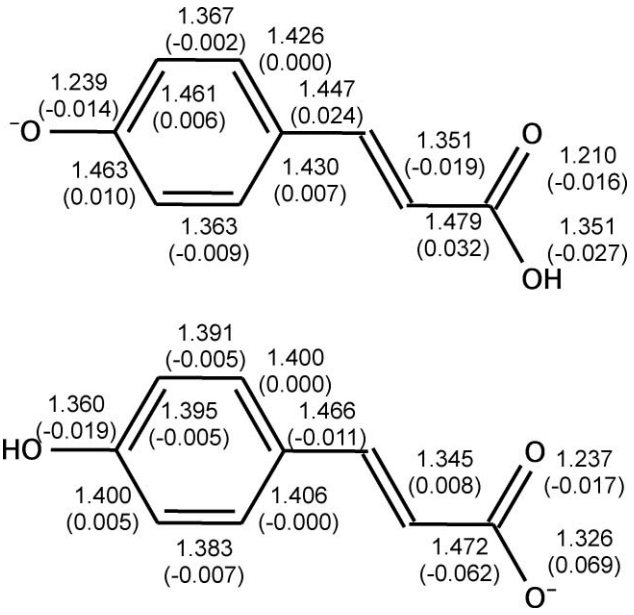

FIG. 4. Equilibrium structures of the ionized *p*-CA$^-$ optimized by ωB97X/aug-cc-pVDZ. Ionization-induced changes in the bond lengths, i.e., difference in the bond lengths between the neutral and anionic state optimized geometries (computed at the same level of theory) are given in parentheses.

In both isomers, the lowest bright state is a $\pi \rightarrow \pi^*$ transition. According to the EOM-EE-CCSD results, this state has mixed character in carboxylate, i.e., HOMO $\rightarrow$ valence lowest occupied molecular orbital, LUMO, ($R_1 = 0.40$) and HOMO $\rightarrow$ valence LUMO+1 ($R_1 = 0.39$). Tables II and III compare performance of different methods for excitation energies of the *p*-CA$^-$ chromophores. The phenolate isomer is relatively well behaved. The excited-state wave function is dominated by a single one-electron excitation ($\pi_3\pi_1^*$). Both single- (CC2, CC3, EOM-CCSD) and multireference approaches [multireference Møller-Plesset perturbation theory (MRMP2), MCQDPT2, CASPT2] as well as MBGFT yield similar S_0–S_1 excitation energies (Table II).

TABLE II. Vertical excitation energies (E_{ex}, eV) and oscillator strengths (f_i, in parenthesis) of the phenolate *p*-CA$^-$ isomer. The RI-MP2/aug-cc-pVDZ optimized geometries were used for the excitation energy calculations.

	Phenolate			
	Leading configurations			
Method	$\pi\pi^*$	$n\pi^*$	$\pi\pi^*$	$\pi\pi^*$
TDDFT/TZP[a]	3.24	–	–	–
CAM-B3LYP/aug-cc-pVTZ[b]	3.40	–	–	–
CC2/6-31G(d)[c,*]	2.89 (0.995)	–	4.04 (0.056)	–
CC2 /cc-pVDZ[d]	3.10 (1.234)	3.37 ($<10^{-3}$)	4.33 (0.081)	–
CC2/aug-cc-pVTZ[b]	3.10	–	–	–
MRMP2/(*p*-type) d-aug-cc-pVDZ[b]	2.79	–	–	–
MCQDPT2 /(*p*-type) d-aug-cc-pVDZ[b]	2.81	–	–	–
aug-MCQDPT2 / (*p*-type) d-aug-cc-pVDZ[b]	2.82 (0.99)	–	–	–
EOM-CCSD/6-31G(d)[c,f]	3.18 (1.016)	–	4.34 (0.042)	–
EOM-CCSD/cc-pVDZ [d,f]	3.17 (1.237)	3.98 ($<10^{-3}$)	4.49 (0.055)	–
MBGFT[e]	2.95	–	–	–
SS-CASPT2/ANO-RCC-VDZP	3.08	–	4.14	4.80
SS-CASPT2/ANO-RCC-VTZP	2.98	–	3.94	4.59
MS-CASPT2/6-31G(d,p)	3.17 (1.124)	–	4.23 (0.105)	4.90 (0.025)
MS-CASPT2/ANO-RCC-VDZP	3.07 (1.076)	–	4.12 (0.115)	4.82 (0.013)
MS-CASPT2/ANO-RCC-VTZP	2.96 (1.005)	–	3.93 (0.123)	4.64 (0.003)
EOM-EE-CCSD/6-31G(d,p)	3.35 (1.072)	–	4.53 (0.071)	–
EOM-EE-CCSD/6-31+G(d,p)	3.19 (1.06)	4.20 ($<10^{-3}$)	4.23 (0.1)	4.87 (0.01)
CC3/6-31+G(d,p)	2.98	–	–	–

[a]Reference 85.
[b]Reference 10.
[c]Reference 19.
[d]Reference 18.
[e]Non standard Gaussian basis set (total 690 functions), Ref. 17.
[f]p-CTM$^-$ anion.

The difference between MS-CASPT2 and SS-CASPT2 excitation energies is 0.02 eV. EOM-EE-CCSD provides a slightly higher excitation energy than MS-CASPT2 (0.2 eV). This trend agrees with the benchmark studies by Schreiber *et al.*[42] The difference is due to the absence of triple excitations in EOM-EE-CCSD, as the inclusion of triples at the CC3 level results in the decrease of excitation energy of the same magnitude (0.2 eV). Good agreement between MS-CASPT2/TZVP and CC3/TZVP for one-electron excitation dominated transition was also pointed out in Ref. 42.

TABLE III. Vertical excitation energies (E_{ex}, eV) and oscillator strengths (f_i, in parenthesis) of the carboxylate *p*CA$^-$ isomer. The RI-MP2/aug-cc-pVDZ optimized geometries were used for the excitation energy calculations.

	Carboxylate			
	Leading configurations			
Method	$n\pi^*$	$n\pi^*$	$\pi\pi^*$	$\pi\pi^*$
CAM-B3LYP/aug-cc-pVTZ[a]	–	–	4.70	–
CC2/aug-cc-pVTZ[a]	–	–	4.79	–
MRMP2/(p-type) d-aug-cc-pVDZ[a]	–	–	5.17	–
MCQDPT2/(p-type) d-aug-cc-pVDZ[a]	–	–	3.05	–
aug-MCQDPT2/(p-type) d-aug-cc-pVDZ[a]	–	–	2.85	–
MBGFT[b]	–	–	4.37	–
SS-CASPT2/ANO-RCC-VDZP	–	–	4.25	4.31
SS-CASPT2/ANO-RCC-VTZP	–	–	4.09	4.17
MS-CASPT2/6-31G(d,p)	–	–	4.21 (0.021)	4.36 (0.064)
MS-CASPT2/ANO-RCC-VDZP	–	–	4.19 (0.102)	4.28 (0.033)
MS-CASPT2/ANO-RCC-VTZP	–	–	3.88 (0.252)	4.11 (0.042)
EOM-EE-CCSD/6-31G(d,p)	–	–	4.70 (0.027)	4.89 (0.018)
EOM-EE-CCSD/6-31+G(d,p)	4.17 ($<10^{-3}$)	4.24 ($<10^{-3}$)	4.50 (0.030)	–
CC3/6-31+G(d,p)	–	–	4.21	–

[a]Reference 10.
[b]Non standard Gaussian basis set (total 690 functions), Ref. 17.

The excitation energy is sensitive to addition of diffuse basis functions to the basis set (0.16 eV). The extension of the basis from double- to triple-zeta quality results in the decrease of the MS-CASPT2 excitation energies by 0.11 eV (Table II). Our best estimate of the S_0–S_1 excitation energy of phenolate is 3.0 eV.

Carboxylate presents a more complex case. Previous quantum chemistry calculations yielded excitation energies (of the bright state) varying in the range of 2.85–4.79 eV. The discrepancies can be explained by the complex electronic structure of the molecule manifesting itself in a complete breakdown of the Koopmans picture, e.g., Hartree–Fock yields an incorrect order of the occupied MOs, as discussed above. Moreover, the two lowest valence virtual MOs are nearly generate (5.06 and 5.09 eV). While Koopmans theorem is rarely accurate, its failure in this system shows that correlation is essential for determining relative energies of the states with very different charge distributions (localized on the carboxylate moiety or delocalized in the phenolate π-system). Although CC and EOM-CC methods are invariant with respect to unitary transformations with active occupied and/or virtual orbital spaces (and are not very sensitive to small variations in occupied-virtual separation), breakdown of the Koopmans picture suggests that correlation is very important and one may need to go beyond double excitations.

According to the EOM-CCSD/6-31G(d,p) calculations, the vertical excitation energy for the bright state is 4.7 eV, and the wave function of the lowest excited state has multiconfigurational character dominated by two $\pi \rightarrow \pi^*$ transitions. As in the case of the phenolate, the addition of diffuse basis functions decreases the EOM-EE-CCSD excitation energy by 0.2 eV. The inclusion of triple excitations results in a decrease of 0.29 eV. We also note excellent agreement between CC3 and CASPT2. The difference between the SS- and MS-CASPT2 values is slightly larger than in the phenolate (0.2 eV). As discussed below, we consider the SS-CASPT2 value to be more reliable in this case. Thus, our best estimate of the excitation energy of the bright state in the carboxylate is 4.1 eV.

The (MS-)CASPT2/ANO-RCC-VTZP calculations of *p*-CA^- are, in most aspects, routine, and one can expect an accuracy of 0.1–0.2 eV.[42] However, in the analysis of the results, some care has to be taken owing to potential problems due to the resonance character of the excited states. In particular, the response to one-particle basis set expansions, the reference weight in the perturbational treatment of the dynamical correlation, i.e., the weight of the reference CASSCF wave function in the CASPT2 first-order solution, and the difference between SS- and MS-CASPT2 should be carefully monitored. First, let us establish that the active orbitals of the two species are correct and that no spurious diffuse orbitals contaminate the CASSCF expansion, which is a tell-tale sign of either the presence of Rydberg states or the electron detachment continuum. The active orbitals of both *p*-CA^- isomers were inspected visually and found to be of a valence π-like character. To further quantify the orbital character, we monitored the $\langle r^2 \rangle$ expectation value of the six SA-CASSCF states (for each isomer) and found that none of the CASSCF states developed diffuse character. Next, we inspected the reference weights in the state-specific CASPT2/ANO-RCC-VTZP calculations. For the phenolate form we found that the fifth root in SS-CASPT2 has a substantially lower reference weight as compared to the rest. In carboxylate we found two states with somewhat lower reference weights: the fourth and the sixth roots. Inspection on the weights in the SS-CASPT2 calculation with the ANO-RCC-VDZP basis reveals no roots with low reference weights (in both isomers). In carboxylate, the observed basis set dependence of the reference weight is further aggravated by the fact that the perturbational correction introduces significant diffuseness and significant mixing in the MS-CASPT2 procedure. The observed problematic behavior in the MS-CASPT2 calculations is likely due to low detachment energies of these species and, consequently, the resonance character of the excited states. As pointed out by Serrano-Andrés *et al.*,[78] the MS-CASPT2 procedure could be overestimating the mixing between reference states and in this case the SS-CASPT2 results could be more reliable. We also note that the basis set effects are smaller at the SS-CASPT2 level, e.g., as we go from the ANO-RCC-VDZP to the ANO-RCC-VTZP basis, the excitation energy of carboxylate decreases by 0.31 and 0.16 eV at the MS-CASPT2 and SS-CASPT2 levels, respectively. Hence, we consider the SS-CASPT2/ANO-RCC-VTZP results as our best estimate. Despite the observed problematic behavior, the differences between the SS- and MS-CASPT2 values are only 0.02 and 0.21 eV for phenolate and carboxylate, respectively. In the latter case, the deviation is due to mixing of four excited states at the MS-CASPT2 level. Moreover, as mentioned above, the perturbative treatment increases the diffuse character of the state. Therefore, it is not clear whether the large nondiagonal elements of the MS-CASPT2 effective Hamiltonian matrix are caused by the correlation-induced interaction of valence excited states or artificial interaction with the continuum. Therefore, we consider SS-CASPT2 results to be more reliable in the case of carboxylate.

Closely related to the CASPT2 and MS-CASPT2 multireference methods are MRMP2 and MCQDPT2, respectively. For phenolate, excitation energies computed with MRMP2 and CASPT2 differ by 0.3 eV. Multistate multireference MCQDPT2 and MS-CASPT2 methods also yield results within 0.3 eV, the MS-CASPT2 excitation energy being higher. The difference is likely due to the IPEA correction used in the CASPT2 and MS-CASPT2 calculations.

For carboxylate, in contrast, the excitation energies computed with the two sets of methods are rather different (see Table III). Note that 5.17 eV MRMP2 excitation energy corresponds to the sixth CASSCF root and the resulting MRMP2 state.[10] The CASPT2 excitation energy for the sixth excited state (5.32 eV, ANO-RCC-VDZP) is only slightly higher than the reported MRMP2 value. However, there is striking difference between the MS-CASPT2 and MCQDPT2 excitation energies.[10] The two methods differ by: (i) slightly different choices of the zero-order Hamiltonian; and (ii) using internally contracted versus noncontracted basis for the secondary space in MS-CASPT2 (Ref. 25) and MCQDPT2 (Ref. 79), respectively. Overall, one can expect similar performance of the two approaches. In addition to the above differences, more diffuse basis set was used in Ref. 10. The

nondiagonal elements of the MS-CASPT2/ANO-RCC-VDZP effective Hamiltonian result in strong mixing between the states; however, the perturbation-modified CAS solution of the lowest excited state is dominated by the lowest CASSCF excited root (the effective Hamiltonian matrix and eigenvectors are given in Supporting Materials.[60]). The mixing between the high-lying fifth and sixth (forth and sixth with CASSCF/ANO-RCC-VDZP) CASSCF states increases upon the increase of the basis set to ANO-RCC-VTZP and the lowest perturbation-modified CAS solution is represented by nearly equal contributions from the second, fifth, and sixth CASSCF roots. Therefore, the nondiagonal effective Hamiltonian matrix elements that couple the fifth and the sixth CASSCF/ANO-RCC-VTZP (fourth and sixth with ANO-RCC-VDZP) states are very sensitive to the basis set. The difference between the MS-CASPT2 and MCQDPT2 excitation energies could be due to even stronger mixing between these CASSCF states in a more diffuse basis set. Indeed, the MCQDPT2 perturbation-modified CAS solution for the lowest excited state is dominated by the sixth CASSCF state, and the corresponding excitation energy is 3.05 eV.[10]

The SS-CASPT2/ANO-RCC-VTZP values of the lowest excitation energies for the bright $\pi \rightarrow \pi^*$ transitions are 4.09 and 2.98 eV, for the carboxylate and phenolate forms, respectively. The CC3/6-31+G(d,p) values are 4.21 and 2.98 eV, the corresponding EOM-CCSD values are 4.50 and 3.19 eV.

Our best estimates of vertical excitation energies of the bright $\pi \rightarrow \pi^*$ state are 3.0 and 4.1 eV for the phenolate and carboxylate forms of p-CA$^-$, respectively. The higher excitation energy of the carboxylate can be explained by qualitative analysis of the electronic structure: in the carboxylate, one can expect excitation energy close to that of neutral phenol [4.5 eV (Ref. 80)], whereas in the phenolate, the phenol-like electronic structure is strongly perturbed by the residing negative charge resulting in the red-shifted absorption. A similar mechanism of the tuning of optical properties of biochromophores was found to be important in chemically initiated electron-exchange luminescence of luciferins.[81] The Hückel model provides an alternative explanation of the observed trend in excitation energies (see Sec. III C).

As mentioned above, an interesting feature of the both isomers is the resonance or near-resonance character of the lowest bright excited state. However, the two isomers show qualitatively different types of the resonance. As follows from Fig. 5, the S_1 state of phenolate is a shape resonance that lies above its own continuum (detachment from the HOMO). The energy diagram is different for carboxylate (Fig. 6), i.e., the S_1 state is above three ionized states corresponding to electron detachment from HOMO-1, HOMO-2, and HOMO-3 and 0.25 eV below its own continuum (detachment from HOMO). This is an example of a Feshbach resonance. One would expect longer lifetimes for this type of metastable state.

It should be noted that the converged energies of the resonance states may differ from the excitation energies reported here due to basis set effects and interactions with the continuum. Thus, more elaborate calculations are required to evaluate the positions of the resonances and the lifetimes of the metastable excited states in these complex systems.

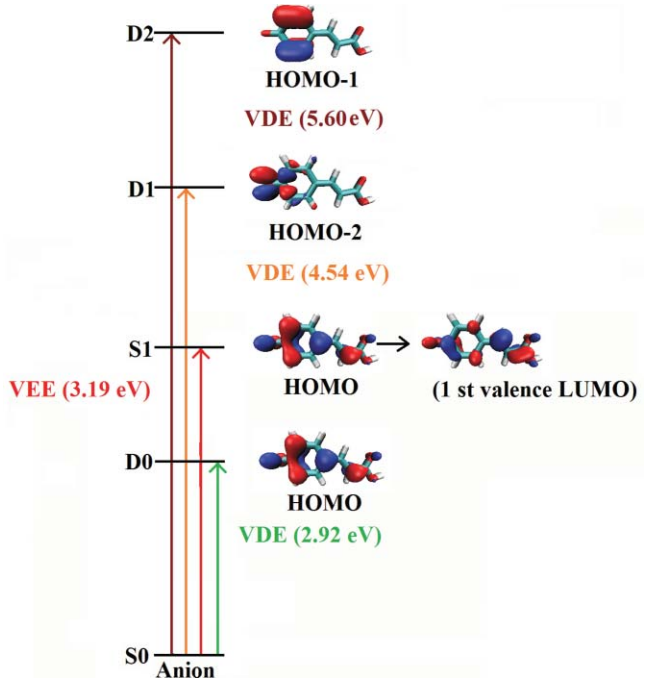

FIG. 5. Energy level diagram of the ionized and electronically excited states of the phenolate form. VEE: EOM-EE-CCSD/6-31+G(d,p); VDE: EOM-IP-CCSD/6-311+G(df,pd).

C. Molecular orbital framework

As in the case of GFP,[14] the trends in electronic properties of the isomers can be explained by a simple Hückel-like model. Although the resulting excitation and ionization energies cannot be considered as quantitative, this analysis provides a qualitative explanation of the observed differences between the two isomers.

As the electronic density redistribution mainly involves the bridge region for both isomers, the analysis is based on a model system consisting of the three bridge carbons. Assuming almost perfect resonance for the phenolate, which

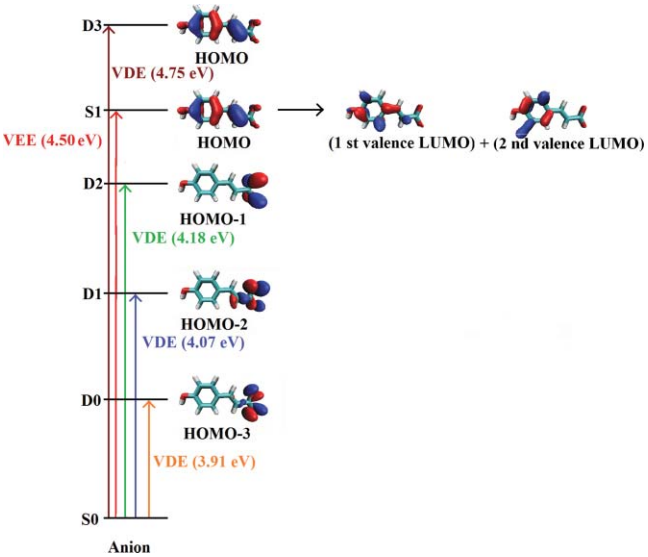

FIG. 6. Energy level diagram of the ionized and electronically excited states of the carboxylate form. VEE: EOM-EE-CCSD/6-31+G(d,p); VDE: EOM-IP-CCSD/6-311+G(df,pd).

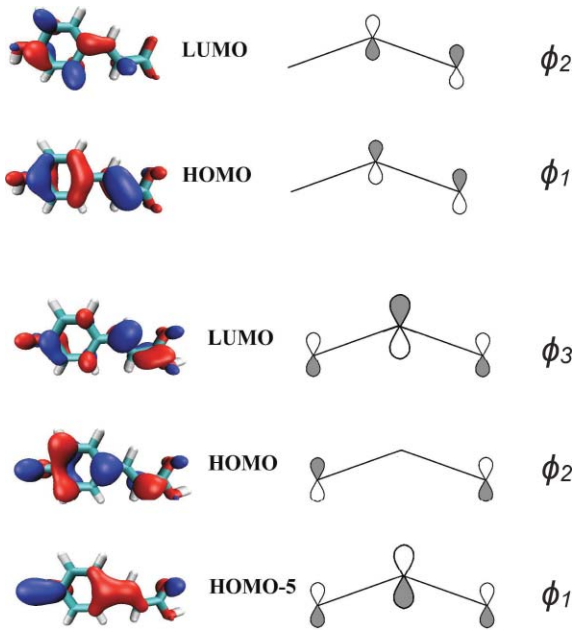

FIG. 7. Schematic representation of the Hückel model eigenfunctions (see text) and the corresponding MOs for the carboxylate (top) and the phenolate (bottom) *p*-CA^-.

makes the three atoms (C_α, C_β and C_γ) equivalent, the Hückel Hamiltonian is written as follows:

$$H = \begin{pmatrix} \varepsilon & \alpha & 0 \\ \alpha & \varepsilon & \alpha \\ 0 & \alpha & \varepsilon \end{pmatrix}, \tag{1}$$

where ϵ is an atomic *p*-orbital energy and α is a coupling matrix element between the two neighboring centers. This model is exactly equivalent to the Huckel's solution for the allyl radical. The diagonalization of this matrix yields the following eigenvalues:

$$E_1^{Ph} = \varepsilon + \sqrt{2}\alpha, \tag{2}$$

$$E_2^{Ph} = \varepsilon, \tag{3}$$

$$E_3^{Ph} = \varepsilon - \sqrt{2}\alpha. \tag{4}$$

The corresponding eigenfunctions $\{\phi_i\}_{i=1,3}$ are depicted in Fig. 7. The validity of Hückel's description is supported by the analysis of the MOs: the shapes of the HOMO-5, HOMO, and LUMO in the bridge C–C_γ region are indeed similar to the $\{\phi_i\}_{i=1,3}$ Hückel solutions. The corresponding detachment and excitation energies are $-(\varepsilon)$ and $-\sqrt{2}\alpha$, respectively.

For the carboxylate isomer, BLA in the bridge region is larger suggesting that the three carbons are no longer equivalent. Assuming $\varepsilon_\alpha << \varepsilon_\beta = \varepsilon_\gamma = \varepsilon$ and neglecting the coupling between C_α and C_β, we arrive to the following Hamiltonian:

$$h = \begin{pmatrix} \varepsilon' & 0 & 0 \\ 0 & \varepsilon & \alpha \\ 0 & \alpha & \varepsilon \end{pmatrix}, \tag{5}$$

which gives rise to the eigenvalues:

$$E_1^{\text{Carb}} = \varepsilon + \alpha \tag{6}$$

$$E_2^{\text{Carb}} = \varepsilon - \alpha. \tag{7}$$

The resulting detachment and excitation energies for the carboxylate are $(-\epsilon - \alpha)$ and -2α. Therefore, Hückel's model predicts higher excitation and detachment energies for the carboxylate isomer.

The model yields the following estimates of the transition dipole moments matrix elements for carboxylate and the phenolate: $\langle \pi|\mu|\pi^* \rangle = \frac{x_0}{2}$ and $\langle \pi|\mu|\pi^* \rangle = \frac{x_0}{\sqrt{2}}$, respectively, where x_0 is the average C–C bond length between the bridge atoms (see Appendix). This suggests that the oscillator strength for the phenolate form is higher than that of carboxylate, which is indeed confirmed by *ab initio* calculations (see Tables II and III). Additional details of the Hückel analysis are given in the Appendix.

D. Theory versus experiment

As mentioned above, recent gas-phase action spectroscopy measurements reported identical absorption maxima for the carboxylate and phenolate isomers of *p*-CA^-. As a source of gas-phase anions, the electrospray technique was used[8–10] and the ions were extracted from water–methanol solution. It has been demonstrated that for tyrosine, which also exists in the carboxylate and phenolate forms, the electrospray extraction from methanol–water solution results in the gas-phase mixture with relative abundances of the two isomers corresponding to the gas-phase equilibrium distribution.[82] In aqueous solution, deprotonation of carboxylate is preferable.[82] It was also shown that the relative populations are very sensitive to the presence of methanol in the solution. Thus, it is not clear which of the *p*-CA^- isomers is present in gas phase when the electrospray technique is used.[17,82] To achieve isomer specificity, the methylated analogs of *p*-CA were used in the experiment.[10] The comparison of the results obtained for nonmethylated species with the absorption spectra of the methylated compounds is justified because only a minor effect of methylation on excitation energies was reported for *p*-CA^- at the CC2 level of theory[10] and there is good agreement between EOM-CCSD and CC2 for this system (see Tables II and III).

The comparison of our calculations with the experimental spectra[10] is presented in Fig. 8. This experiment[10] did not distinguish between the detachment and excitation channels. Thus, both detachment and excitation can contribute to the experimental band. The computed SS-CASPT2/ANO-RCC-VTZP excitation and EOM-IP-CCSD/6-311+G(df,pd) detachment energies for the phenolate chromophore are 2.98 and 2.92 eV, and, therefore, both can be responsible for the band maximum. For carboxylate, the respective values are 4.09 and 3.91 eV. These values do not agree with the experimental peak, and the discrepancies are much larger than the anticipated error bars of these methods.

According to our calculations (and in agreement with previous studies), there are no transitions below 3.9 in the carboxylate (see Table III). We considered dipole-bound states[83] and isomerization in the gas phase as possible explanations of these discrepancies. The dipole bound states can exist in *p*-CA^- (dipole moment 4.82 D for carboxylate and 3.01 D for phenolate neutral radicals); however, they cannot account for the 1 eV difference.

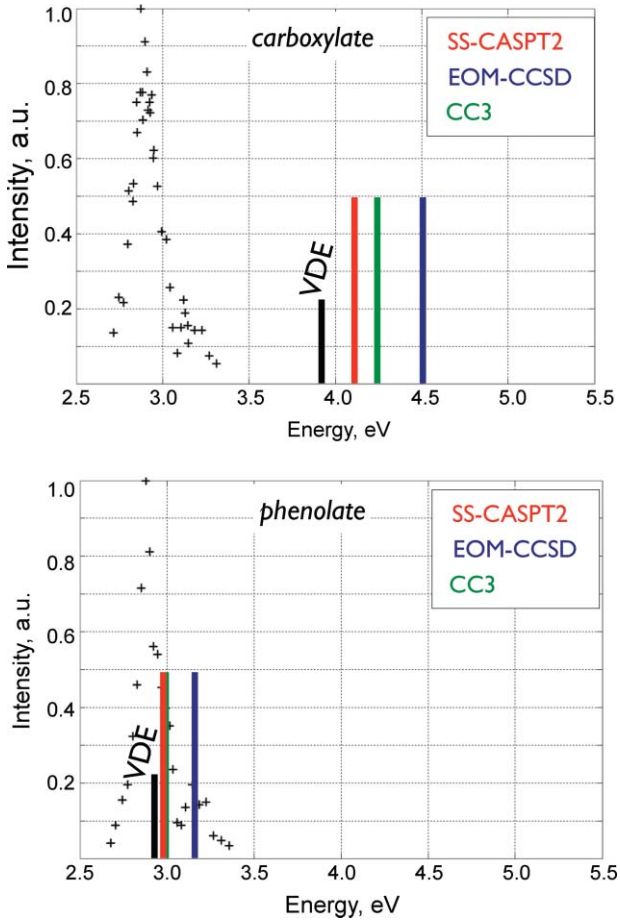

FIG. 8. The experimental spectra[10] and calculated VDE and VEE of the carboxylate (top) and phenolate (bottom) forms of *p*-CA$^-$. The experimental maximum is 2.88 eV (430 nm). The heights of the bars representing vertical excitation and ionization energies are arbitrary.

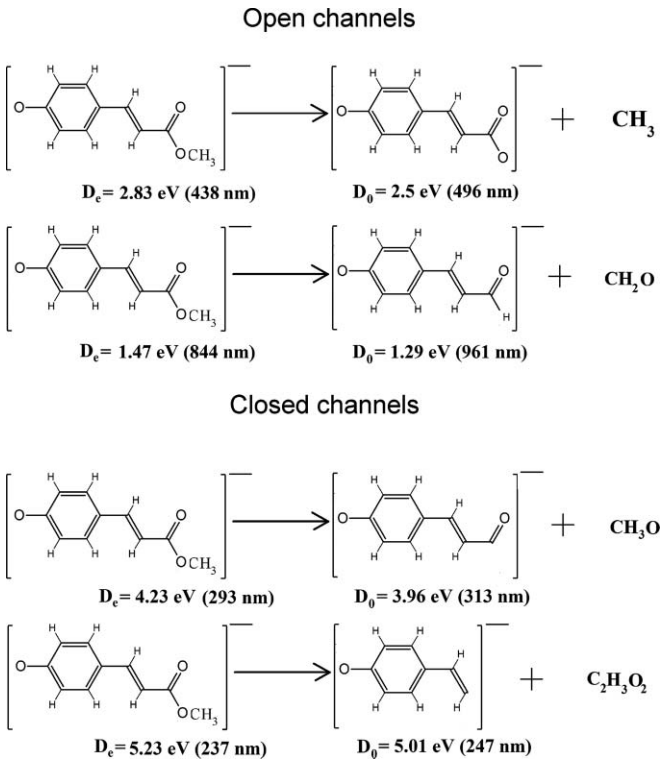

FIG. 9. Energetically allowed (top) and forbidden (bottom) fragmentation channels and the corresponding dissociation energies for the phenolate form of *p*-CA$^-$. D_0 and D_e denote dissociation energies computed with and without ZPE correction, respectively.

One of the possible explanations of the experimental results is a contamination of the carboxylate sample by the phenolate form. If this was the case, the experiment signal could be dominated by the phenolate owing to its much larger oscillator strength. The reported identical spectra for the two isomers[10] support this assumption. Since the higher energy region was not probed, the signal due to carboxylate was not observed.

The photoinduced fragmentation channels have been experimentally determined by registering the masses of the resulting charged fragments.[10] The 146 amu signal for the anionic photoproduct was reported for both carboxylate and phenolate (methylated *p*-CA$^-$). This channel was ascribed to the detachment of the neutral OCH_3 fragment. Note that the experimental mass resolution was ±2 amu.[10] The high yield of the neutral fragments other than ionized chromophore shows that the excitation to the S_1 state is an efficient channel, in addition to possible direct detachment for both forms of the chromophore. Thus, the experimental absorption bands for phenolate and carboxylate represent, at least partially, transition to the resonance electronically excited states.

To analyze the fragmentation pathways, we computed dissociation energies for methylated *p*-CA$^-$ in the phenolate and carboxylate forms (Figs. 9 and 10). The results for nonmethylated *p*-CA$^-$ are given in the supporting materials.[60] Our calculations show that the only energetically allowed channels for the phenolate correspond to the abstraction of CH_3 and CH_2O, as shown in Fig. 9. Within the experimental mass resolution, formaldehyde (CH_2O) is indistinguishable from the OCH_3 radical. Moreover, the mass of the anionic fragment resulting from the CH_3 detachment from *p*-CA$^-$ is 151 amu, which is ∼91% from the mass of the parent ion, whereas only the daughter ions with masses in the range of 20–80% of the parent ion mass are captured in this experimental setup.[9] Therefore, the predicted photofragmentation products for phenolate agree with the experimental data. For carboxylate, in addition to the detachment of the neutral CH_3 and CH_2O fragments, there is a low-energy fragmentation channel leading to CO_2 formation. This additional channel is specific for the carboxylate and can be used as an experimental probe to distinguish between the isomers. No evidence for CO_2 formation was reported in Ref. 10 raising a question about the nature of the absorbing species. Note that the production of CO_2 was reported as an efficient fragmentation channel in nonmethylated *p*-CA$^-$.[10]

Besides dissociation, photodetachment has been experimentally registered for the phenolate form,[10] which is consistent with computed detachment energies (Table I). Indeed, the experimental absorption maximum value (2.88 eV) is close to the calculated VDE (2.92 eV) suggesting an alternative decay route for the excited phenolate form.

Another factor that may affect the spectra is the presence of different rotamers and low-barrier hindered rotation

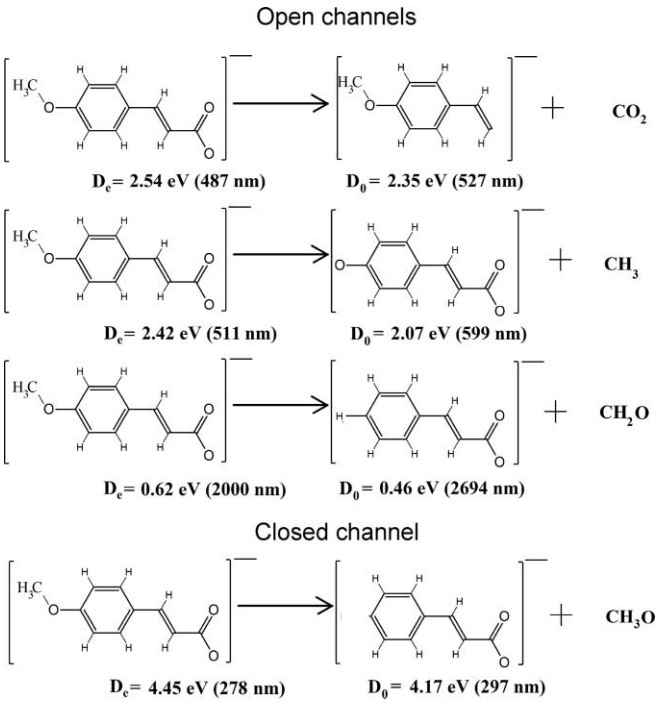

FIG. 10. Energetically allowed (top) and forbidden (bottom) fragmentation channels and the corresponding dissociation energies for the carboxylate form of *p*-CA$^-$. D_0 and D_e denote dissociation energies computed with and without ZPE correction, respectively.

along single bonds of the chromophore. To quantify the effect of the phenolic OH group rotation on the absorption spectra of carboxylate, we computed excitation energies of the anti- and syn- rotamers using EOM-EE-CCSD/6-31+G(d,p). We found that isomerization has only minor effect on vertical excitation energy (changes about 0.08 eV). We analyzed conformational flexibility using AIMD simulations. The average values of C–C_α–C_β–C_γ and C_β–C_γ–C(O)–O dihedral angles along 2 ps trajectories for the methylated carboxylate are 7.18° and 4.70° (methylated phenolate—9.03° and 5.95°), respectively (T = 298 K). Thus, the chromophore is rather inflexible and we anticipate no significant distortion of the conjugated π-system due to rotation along single bonds, and, consequently, no significant variations in excitation energy.

We suggest that the action spectra measured for the two methylated *p*-CA$^-$ isomers in Ref. 10 are due to the phenolate isomer (which probably contaminated the carboxylate sample) based on the following considerations: (i) there are no one-photon transitions (either excitation or ionization) below 3.8 eV in the carboxylate; (ii) the oscillator strength for the lowest bright state of the phenolate is three orders of magnitude higher than in the carboxylate. Thus, a small admixture of the phenolate may result in a relatively large signal; (iii) in the carboxylate there is an additional fragmentation channel leading to CO_2 production, which is not accessible in the phenolate; however, this product was not observed in the experiment;[10] (iv) in the gas phase, phenolate is lower in energy than carboxylate.[10] An alternative explanation might be due to two-photon absorption; however, it does not explain the striking similarity of the two spectra.

IV. CONCLUSIONS

We report an electronic structure study of the excited and ionized states of *p*-CA$^-$, a model PYP chromophore. We compare the optical properties (vertical excitation and detachment energies) of the two isomers, phenolate and carboxylate. In addition to the high-level calculations of the vertical excitation and detachment energies, we present a qualitative explanation of the observed differences. Our best estimates of vertical excitation energies of the two isomers are 3.0 and 4.1 eV (SS-CASPT2/ANO-RCC-VTZP) for phenolate and carboxylate, respectively. We note excellent agreement between SS-CASPT2 and CC3 (although the latter was employed with a modest basis set). The EOM-CCSD values are within 0.2 eV of CC3.

Our results do not support the experimental conclusion that the two isomers have identical absorption spectra.[10] As a possible explanation, we suggest contamination of the carboxylate sample by the phenolate isomer. Due to the large difference in oscillator strength of the two isomers (three orders of magnitude), a small admixture of phenolate may result in relatively large absorption. The absence of CO_2 fragments (which can be produced in by the carboxylate) in the experiment supports our suggestion that the recorded spectrum is due to phenolate. Our work suggests that the experiment should be revisited probing higher energies for the spectroscopic signature of the carboxylate isomer. The production of CO_2 at higher energies would confirm the presence of the carboxylate form.

ACKNOWLEDGMENTS

We are grateful to Prof. L. H. Andersen, Dr. A. V. Bochenkova and Dr. A. A. Granovsky for stimulating discussions. This work was conducted in the framework of the *iOpenShell* Center for Computational Studies of Electronic Structure and Spectroscopy of Open-Shell and Electronically Excited Species (iopenshell.usc.edu) supported by the National Science Foundation through the CRIF:CRF CHE-0625419+0624602+0625237 and CHE-0951634 (AIK) Grants. TDC was supported by the National Science Foundation through Grants CHE-0715185 and CHE-0741927 (CRIF:MU).

APPENDIX: HÜCKEL DESCRIPTION OF THE TRANSITION DIPOLE MOMENTS OF THE CARBOXYLATE AND PHENOLATE ISOMERS OF *p*-CA$^-$

As described in Sec. III C, we employ a Hückel-like model to compare the transition dipole moments of the carboxylate and phenolate. Our dipole moment analysis described below is very similar to that reported by Dewar and Longuet-Higgins.[84] We consider the three carbon atoms on the bridge assuming that the arrangement is linear. The respective Hamiltonians and eigenenergies for the two isomers are given in Sec. III C. The Hückel description of the MOs is shown in Fig. 7.

For the carboxylate, the model yields the following eigenstates:

$$\phi_1 = \frac{1}{\sqrt{2}}(p_\beta + p_\gamma), \tag{A1}$$

$$\phi_2 = \frac{1}{\sqrt{2}}(p_\gamma - p_\beta). \tag{A2}$$

And for the phenolate

$$\phi_1 = \frac{1}{2}(p_\alpha + \sqrt{2}p_\beta + p_\gamma), \tag{A3}$$

$$\phi_2 = \frac{1}{\sqrt{2}}(p_\gamma - p_\alpha), \tag{A4}$$

$$\phi_3 = \frac{1}{2}(p_\alpha - \sqrt{2}p_\beta + p_\gamma), \tag{A5}$$

where p_α, p_β, p_γ are the p-orbitals of the three carbon atoms C_α, C_β, C_γ, respectively. Note that the analysis in Ref. 14 used incorrect normalization of ϕ_1 and ϕ_3; however, the results were not affected by this mistake.

We consider a linear arrangement of the atoms such that p_β is located at $x = 0$, p_α at $x = -x_0$, p_γ at $x = x_0$, i.e., $\langle p_\beta|x|p_\beta\rangle = 0$, $\langle p_\alpha|x|p_\alpha\rangle = -x_0$, $\langle p_\gamma|x|p_\gamma\rangle = x_0$. We assume zero overlap between the orbitals centered on different atoms: $\langle p_\alpha|x|p_\beta\rangle = 0$, $\langle p_\alpha|x|p_\gamma\rangle = 0$, $\langle p_\beta|x|p_\gamma\rangle = 0$.

The transition dipole moment matrix element for carboxylate corresponding to the first excitation (HOMO → LUMO) is

$$\begin{aligned}\langle\phi_1|x|\phi_2\rangle &= \frac{1}{2}\langle p_\beta + p_\gamma|x|p_\gamma - p_\beta\rangle \\ &\approx \frac{1}{2}\langle p_\gamma|x|p_\gamma\rangle - \frac{1}{2}\langle p_\beta|x|p_\beta\rangle = \frac{1}{2}x_0\end{aligned} \tag{A6}$$

And for phenolate

$$\begin{aligned}\langle\phi_2|x|\phi_3\rangle &= \frac{1}{2\sqrt{2}}\langle -p_\alpha + p_\gamma|x|p_\alpha - \sqrt{2}p_\beta + p_\gamma\rangle \\ &\approx -\frac{1}{2\sqrt{2}}\langle p_\alpha|x|p_\alpha\rangle + \frac{1}{2\sqrt{2}}\langle p_\gamma|x|p_\gamma\rangle \\ &= \frac{1}{\sqrt{2}}x_0.\end{aligned} \tag{A7}$$

The first excited state ($\pi \rightarrow \pi^*$) for carboxylate corresponds to $\phi_1 \rightarrow \phi_2$ transition, whereas in phenolate the first excited state is described by $\phi_2 \rightarrow \phi_3$ excitation. The oscillator strength is proportional to $|\langle\phi_1|x|\phi_2\rangle|^2 = x_0^2/4$ and $|\langle\phi_2|x|\phi_3\rangle|^2 = x_0^2/2$ for carboxylate and phenolate, respectively. Therefore, the oscillator strength is expected to be higher for phenolate. These results are qualitatively consistent with our calculations; however, the difference in the magnitudes according to the calculations (see Tables II and III) for the isomers is much larger (30 times).

[1] T. E. Meyer, Biochim. Biophys. Acta **806**, 175 (1985).
[2] T. E. Meyer, E. Yakali, M. A. Cusanovich, and G. Tollint, Biochem. J. **26**, 418 (1987).
[3] W. W. Sprenger, W. D. Hoff, J. P. Armitage, and K. J. Hellingwer, J. Bacteriol. **175**, 3096 (1993).
[4] W. D. Hoff, P. Düx, K. Hård, B. Devreese, I. M. Nugteren-Roodzant, W. Crielaard, R. Boelens, R. Kaptein, J. van Beeuman, and K. J. Hellingwerf, Biochemistry **33**, 13959 (1994).
[5] M. Baca, G. E. O. Borgstahl, M. Boissinot, P. M. Burke, D. R. Williams, K. A. Slater, and E. D. Getzoff, Biochemistry **33**, 14369 (1994).
[6] R. Kort, H. Vonk, X. Xu, W. D. Hoff, W. Crielaard, and K. J. Hellingwerf, FEBS Lett. **382**, 73 (1996).
[7] L. J. G. W. van Wilderen, M. A. van der Horst, I. H. M. van Stokkum, K. J. Hellingwerf, R. van Grondelle, and M. L. Groot, Proc. Natl. Acad. Sci. U.S.A. **103**, 15050 (2006).
[8] I. B. Nielsen, S. Boye-Peronne, M. O. A. El Ghazaly, M. B. Kristensen, S. Brondsted Nielsen, and L. H. Andersen, Biophys. J. **89**, 2597 (2005).
[9] L. Lammich, J. Rajput, and L. H. Andersen, Phys. Rev. E **78**, 051916 (2008).
[10] T. Rocha-Rinza, O. Christiansen, J. Rajput, A. Gopalan, D. B. Rahbek, L. H. Andersen, A. V. Bochenkova, A. A. Granovsky, K. B. Bravaya, A. V. Nemukhin, K. L. Christiansen, and M. B. Nielsen, J. Phys. Chem. A **113**, 9442 (2009).
[11] I.-R. Lee, W. Lee, and A. H. Zewail, Proc. Natl. Acad. Sci. U.S.A. **103**, 258 (2006).
[12] M. W. Forbes and R. A. Jockusch, J. Am. Chem. Soc. **131**, 17038 (2009).
[13] E. Epifanovsky, I. Polyakov, B. L. Grigorenko, A. V. Nemukhin, and A. I. Krylov, J. Chem. Theory Comput. **5**, 1895 (2009).
[14] E. Epifanovsky, I. Polyakov, B. L. Grigorenko, A. V. Nemukhin, and A. I. Krylov, J. Chem. Phys. **132**, 115104 (2010).
[15] E. V. Gromov, I. Burghardt, H. Köppel, and L. S. Cederbaum, J. Phys. Chem. A **109**, 4623 (2005).
[16] Z. He, C. H. Martin, R. Birge, and K. F. Freed, J. Phys. Chem. A **104**, 2939 (2000).
[17] Y. Ma, M. Rohlfing, and C. Molteni, J. Chem. Theory Comput. **6**, 257 (2010).
[18] E. Gromov, I. Burghardt, J. Hynes, H. Köppel, and L. Cederbaum, Photochem. Photobiol. **190**, 241 (2007).
[19] E. V. Gromov, I. Burghardt, H. Köppel, and L. S. Cederbaum, J. Am. Chem. Soc. **129**, 6798 (2007).
[20] D. Sinha, D. Mukhopadhyay, and D. Mukherjee, Chem. Phys. Lett. **129**, 369 (1986).
[21] J. F. Stanton and R. J. Bartlett, J. Chem. Phys. **98**, 7029 (1993).
[22] S. V. Levchenko and A. I. Krylov, J. Chem. Phys. **120**, 175 (2004).
[23] J. Simons, *Encyclopedia of Computational Chemistry*, chapter Equation of Motion (EOM) Methods for Computing Electron Affinities Wiley, New York, 1998.
[24] A. I. Krylov, Annu. Rev. Phys. Chem. **59**, 433 (2008).
[25] J. Finley, P.-Å. Malmqvist, B. O. Roos, and L. Serrano-Andrés, Chem. Phys. Lett. **288**, 299 (1998).
[26] H. Koch, O. Christiansen, P. Jørgensen, A. M. S. de Meras, and T. Helgaker, J. Chem. Phys. **106**, 1808 (1997).
[27] T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).
[28] M. W. Feyereisen, G. Fitzergald, and A. Komornicki, Chem. Phys. Lett. **208**, 359 (1993).
[29] O. Vahtras, J. Almof, and M. W. Feyereisen, Chem. Phys. Lett. **213**, 514 (1993).
[30] A. Komornicki and G. Fitzergald, J. Chem. Phys. **98**, 1399 (1993).
[31] D. E. Bernhold and R. J. Harrison, Chem. Phys. Lett. **250**, 477 (1996).
[32] J.-D. Chai and M. Head-Gordon, J. Chem. Phys. **128**, 084106 (2008).
[33] A. A. Golubeva, P. A. Pieniazek, and A. I. Krylov, J. Chem. Phys. **130**, 124113 (2009).
[34] S. Pal, M. Rittby, R. J. Bartlett, D. Sinha, and D. Mukherjee, Chem. Phys. Lett. **137**, 273 (1987).
[35] J. F. Stanton and J. Gauss, J. Chem. Phys. **101**, 8938 (1994).
[36] M. Kamiya and S. Hirata, J. Chem. Phys. **125**, 074111 (2006).
[37] P. A. Pieniazek, S. A. Arnstein, S. E. Bradforth, A. I. Krylov, and C. D. Sherrill, J. Chem. Phys. **127**, 164110 (2007).
[38] P. A. Pieniazek, S. E. Bradforth, and A. I. Krylov, J. Chem. Phys. **129**, 074104 (2008).
[39] V. I. Lebedev, Zh. Vychisl. Mat. Mat. Fiz. **15**, 48 (1975).
[40] W. C. Murray, N. C. Handy, and G. J. Laming, Mol. Phys. **78**, 997 (1993).
[41] H. Larsen, K. Hald, J. Olsen, and P. Jørgensen, J. Chem. Phys. **115**, 3015 (2001).
[42] M. Schreiber, M. R. Silva-Junior, S. P. A. Sauer, and W. Thiel, J. Chem. Phys. **128**, 134110 (2008).
[43] E. Epifanovsky, K. Kowalski, P.-D. Fan, M. Valiev, S. Matsika, and A. I. Krylov, J. Phys. Chem. A **112**, 9983 (2008).
[44] B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn, Chem. Phys. **48**, 157 (1980).
[45] K. Andersson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski, J. Phys. Chem. **94**, 5483 (1990).

[46]P. C. Hariharan and J. A. Pople, Theor. Chim. Acta **28**, 213 (1973).
[47]P. O. Widmark, P.-Å. Malmqvist, and B. Roos, Theor. Chim. Acta **77**, 291 (1990).
[48]B. O. Roos, R. Lindh, P.-Å. Malmqvist, V. Veryazov, and P. O. Widmark, J. Phys. Chem. **108**, 2851 (2004).
[49]M. Douglas and N. M. Kroll, Ann. Phys. (N.Y.) **82**, 89 (1974).
[50]G. Ghigo, B. O. Roos, and P.-Å. Malmqvist, Chem. Phys. Lett. **396**, 142 (2004).
[51]P.-Å. Malmqvist and B.O. Roos, Chem. Phys. Lett. **155**, 189 (1989).
[52]N. H. F. Beebe and J. Linderberg, Int. J. Quantum Chem. **12**, 683 (1977).
[53]F. Aquilante, T. B. Pedersen, and R. Lindh, Theor. Chem. Acc. **124**, 1 (2009).
[54]F. Aquilante, L. Gagliardi, T. B. Pedesen, and R. Lindh, J. Chem. Phys. **130**, 154107 (2009).
[55]F. Aquilante, R. Lindh, and T. B. Pedersen, J. Chem. Phys. **127**, 114107 (2007).
[56]F. Aquilante, T. B. Pedersen, R. Lindh, B. O. Roos, A. S. de Merás, and H. Koch, J. Chem. Phys. **129**, 024113 (2008).
[57]F. Aquilante, P.-Å. Malmqvist, T. B. Pedersen, A. Ghosh, and B. O. Roos, J. Theor. Comput. Chem. **4**, 694 (2008).
[58]F. Aquilante, T. B. Pedersen, and R. Lindh, J. Chem. Phys. **126**, 194106 (2007).
[59]J. Boström, M. G. Delcey, F. Aquilante, L. Serrano-Andrés, T. B. Pedersen, and R. Lindh, J. Theor. Comput. Chem. **6**, 747 (2010).
[60]See supplementary material at http://dx.doi.org/10.1063/1.3516211 for the Cartesian geometries, relevant energies, and harmonic frequencies.
[61]F. Weinhold and C. R. Landis, Chem. Educ. Res. Pract. **2**, 91 (2001).
[62]E. D. Glendening, J. K. Badenhoop, A. E. Reed, J. E. Carpenter, J. A. Bohmann, C. M. Morales, and F. Weinhold, NBO 5.0., Theoretical Chemistry Institute, University of Wisconsin, Madison, WI, 2001.
[63]Y. Shao, L. F. Molnar, Y. Jung *et al.*, Phys. Chem. Chem. Phys. **8**, 3172 (2006).
[64]T. D. Crawford, C. D. Sherrill, E. F. Valeev, J. T. Fermann, R. A. King, M. L. Leininger, S. T. Brown, C. L. Janssen, E. T. Seidl, J. P. Kenny, and W. D. Allen, J. Comput. Chem. **28**, 1610 (2007).
[65]F. Aquilante, L. de Vico, N. Ferré, G. P-Å Malmqvist, P. Neogrády, T. B. Pedersen, M. Pitonăk, M. Reiher, B. Roos, L. Serrano-Andrés, M. Urban, V. Veryazov, and R. Lindh, J. Comput. Chem. **31**, 224 (2010).
[66]S. Olsen and S. C. Smith, J. Am. Chem. Soc. **130**, 8677 (2008).
[67]I. Polyakov, E. Epifanovsky, B. L. Grigorenko, A. I. Krylov, and A. V. Nemukhin, J. Chem. Theory Comput. **5**, 1907 (2009).
[68]J. Simons, J. Phys. Chem. A **112**, 6401 (2008).
[69]J. Simons, Acc. Chem. Res. **39**, 772 (2006).
[70]M. Sobczyk and J. Simons, J. Phys. Chem. B **110**, 7519 (2006).
[71]A. U. Hazi and H. S. Taylor, Phys. Rev. A **1**, 1109 (1970).
[72]U. V. Riss and H.-D. Meyer, J. Phys. B **26**, 4503 (1993).
[73]G. Jolicard and E. J. Austin, Chem. Phys. Lett. **121**, 106 (1985).
[74]J. Aguilar and J. M. Combes, Commun. Math. Phys. **22**, 269 (1971).
[75]E. Balsev and J. M. Combes, Commun. Math. Phys. **22**, 280 (1971).
[76]B. Simon, Commun. Math. Phys. **27**, 1 (1972).
[77]W. P. Reinhardt, Annu. Rev. Phys. Chem. **33**, 223 (1983).
[78]L. Serrano-Andrés, M. Merchán, and R. Lindh, J. Chem. Phys. **122**, 104107 (2005).
[79]H. Nakano, J. Chem. Phys. **99**, 7983 (1993).
[80]S. J. Martinez III, J. C. Alfano, and D. H. Levy, J. Mol. Spectrosc. **152**, 80 (1992).
[81]Y.-J. Liu, L. De Vico, and R. Lindh, J. Photochem. Photobiol. A **194**, 261 (2008).
[82]Z. Tian and S. R. Kass, J. Am. Chem. Soc. **130**, 10842 (2008).
[83]K. D. Jordan and F. Wang, Annu. Rev. Phys. Chem. **54**, 367 (2003).
[84]M. J. S. Dewar and H. C. Longuet-Higgins, Proc. Phys. Soc., London, Sect. A **67**, 795 (1954).
[85]A. Sergi, M. Grüning, M. Ferrario, and F. Buda, J. Phys. Chem. B **105**, 4386 (2001).