Electronic structure and dynamics of torsion-locked photoactive yellow protein chromophores

Alice Henley, Matus E. Diveky, Anand M. Patel, Michael A. Parkes, James C. Anderson and Helen H. Fielding*

Department of Chemistry, University College London, 20 Gordon Street, London, WC1H 0AJ, UK

E-mail: h.h.fielding@ucl.ac.uk

---

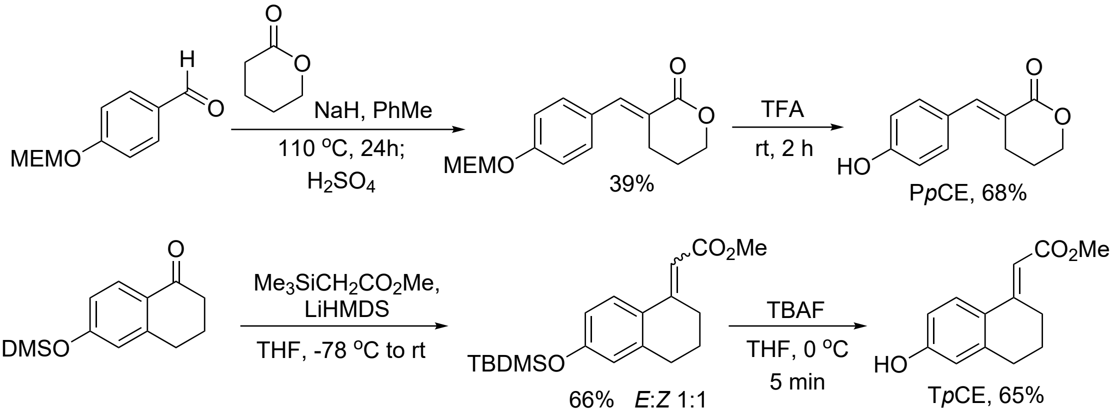
[*] For whom correspondence should be addressed

Supplementary Information

Preparation of P*p*CE and N*p*CE

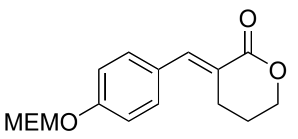
General experimental

Unless otherwise stated, all reactions were carried out under an atmosphere of nitrogen. All glassware was flame dried under a stream of nitrogen before use. Cooling to 0 °C was effected using an ice-water bath. Reactions were monitored by thin layer chromatography (TLC) using Polygram Sil G/UV$_{254}$ 0.25 mm silica gel precoated plastic plates with fluorescent indicator. Sheets were visualised using ultraviolet light (254 nm), ninhydrin or $KMnO_4$, as appropriate. Flash chromatography was performed using Fluorochem silica gel 60, 35-70 μ. The liquid phase was analytical grade 40-60 petroleum ether and ethyl acetate (EtOAc) unless otherwise stated. Removal of solvents (*in vacuo*) was achieved using a Vacuubrand diaphragm pump or house vacuum and Büchi rotary evaporators. All NMR data was collected using a Bruker AMX 300 MHz, Bruker AVANCE III 400 MHz, Bruker AVANCE 500 MHz or Bruker AVANCE III 600 MHz. Data was manipulated directly using Bruker XwinNMR (version 2.6) or TopSpin (version 2.1). Reference values for residual solvents were taken as $\delta = 7.27$ ($CDCl_3$) and 2.51 ppm (DMSO-*d6*) for ^{1}H NMR; $\delta = 77.16$ ppm ($CDCl_3$) for ^{13}C NMR. Multiplicities for coupled signals were denoted as: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet, br. = broad, apt. = apparent and dd = double doublet *etc.* Coupling constants (J) are given in Hz and are uncorrected. Where appropriate, COSY, DEPT, HMBC, HMQC and NOE experiments were carried out to aid assignment. Mass spectroscopy data was collected on a Thermo Finnigan Mat900xp (EI/CI) VG 70se (FAB) and Waters LCT Premier XE (ES) instruments. Infrared data was collected using a Perkin Elmer 1600 FTIR machine as a thin film unless otherwise stated. Elemental analysis was performed on an Exeter Analytical Inc. EA440 horizontal load analyser. Melting points are uncorrected and were recorded on a Stuart Scientific SMP3 system. Commercial solvents and reagents were used as supplied or purified in accordance with standard procedures, as described below. THF, Et_2O and toluene were obtained from solvent towers, where the degassed solvent was passed through a 7 micron filter under 4 bar pressure.



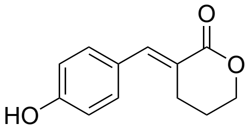
Scheme S1: MEM = $CH_3OCH_2CH_2OCH_2$–, TFA = CF_3CO_2H, LiHMDS = $LiN[Si(CH_3)_3]_2$, TBAF = $^n Bu_4NF$

(*E*)-3-(4-((2-methoxyethoxy)methoxy)benzylidene)tetrahydro-2*H*-pyran-2-one



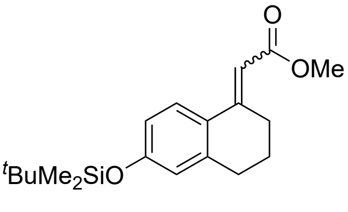
As shown in Scheme S1, NaH (0.15 g, 6.25 mmol) was added to a solution of MEM protected benzaldehyde[1] (1.02 g, 4.86 mmol) and *δ*-valerolactone (0.95 g, 9.49 mmol) in dry toluene (20 mL) and the solution was stirred for 5 min at -10 °C. The solution was stirred at rt for 1 h and then refluxed at 110 °C for 24 h. The solution was diluted with EtOAc (15 mL), quenched with H_2O (10 mL) and concentrated H_2SO_4 (0.70 mL) added dropwise. The mixture was diluted with EtOAc (1×20 mL), separated and the organic layer washed with sat. $NaHCO_3$ (20 mL), brine (20 mL), dried ($MgSO_4$) and concentrated *in vacuo*. Purification by flash chromatography (40% EtOAc/hexane) afforded MEM protected P*p*CE (0.38 g, 39%) as a waxy yellow solid; m.p. 39–40 °C; R_f = 0.26 (50% EtOAc/hexane); IR ν_{max} 2891 (C–H), 1697 (C=O), 1600 (C=C–C=O), 1510 (Ar), 1168 (C–O) cm^{-1}; ^{1}H NMR (600 MHz, $CDCl_3$) δ 7.88 (1H, br.d, J = 2.2, ArCH), 7.42 (2H, d, J = 8.7, 2H, ArH), 7.09 (2H, d, J = 8.8, ArH), 5.31 (2H, s, OCH_2OAr), 4.48 – 4.30 (2H, m, CO_2CH_2), 3.87 – 3.77 (2H, m, OCH_2CH$_2$O), 3.60 – 3.48 (2H, m, OCH$_2$CH_2O), 3.38 (3H, s, OCH_3), 2.88 (2H, td, J = 6.6, 2.3, CCH_2), 1.99 (2H, td, J = 11.1, 6.5, CCH$_2$CH_2); ^{13}C NMR (151 MHz, $CDCl_3$) δ 167.4 (OCO), 158.1 (CCO$_2$), 141.4 (ArCH), 132.2 (ArCH), 128.9 (ArC), 123.8 (ArC), 116.3 (ArCH), 93.3 (OCH$_2$OAr), 71.7 (MeOCH$_2$CH$_2$O), 68.7 (CO$_2$$CH_2$), 68.0 (MeOCH$_2$$CH_2$O), 59.2 (H$_3$$C$O), 26.1 ($CCH_2$), 23.1 (CCH$_2$$CH_2$); *m/z* (CI+) 601 (25%, 2M$^+$–NH$_3$), 584 (30%, 2M$^+$), 309 (55%, M$^+$–NH$_3$), 292 (100%, M$^+$); HRMS $C_{16}H_{20}O_5$ calcd. 292.1305, found 292.1306.

(*E*)-3-(4-hydroxybenzylidene)tetrahydro-2*H*-pyran-2-one (P*p*CE)



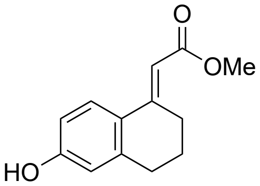
The MEM protected P*p*CE (0.19 g, 0.65 mmol) was stirred in TFA (6.35 mL) at rt for 2 h. The solution was quenched with $NaHCO_3$ (~50 mL) until it was neutralised (Litmus paper). The mixture was extracted with EtOAc (3 × 20 mL), washed with brine (20 mL), dried ($MgSO_4$) and concentrated *in vacuo*. Purification by flash chromatography (60% EtOAc/hexane) afforded **P*p*CE** (0.09 g, 68%) as white crystals; m.p. 170–172 °C; R_f = 0.26 (50% EtOAc/hexane); IR ν_{max} 3149 (O–H), 2960 (C–H), 1661 (C=O), 1572 (C=C–C=O), 1509 (Ar), 1172 (C–O) cm^{-1}; ^{1}H NMR (400 MHz, DMSO) δ 9.97 (1H, s, ArOH), 7.62 – 7.61 (1H, m, ArCH), 7.41 (2H, d, J = 8.7, ArH), 6.84 (2H, d, J = 8.7, ArH), 4.33 – 4.29 (2H, m, CO_2CH_2), 2.80 (2H, td, J = 6.6, 2.2, CCH_2), 1.89 – 1.86 (2H, m, CCH$_2$CH_2); ^{13}C NMR (151 MHz, DMSO) δ 166.6 (OCO), 158.7 (CCO2), 140.0 (ArCH), 132.5 (ArCH), 125.8 (ArC), 123.1 (ArC), 115.6 (ArCH), 68.1 (CO$_2$$CH_2$), 25.6 ($CCH_2$), 22.5 (CCH$_2$$CH_2$); *m/z* (CI+), 222 (100%, M$^+$–NH$_3$), 205 (35%, M$^+$–H); HRMS $C_{12}H_{12}O_3$+H calcd. 205.0859, found 205.0861.

Methyl (*E*/*Z*)-2-(6-((tert-butyldimethylsilyl)oxy)-3,4-dihydronapthalen-1(*2*H)-ylidene)acetate



A solution of LiHMDS (1 M in THF, 5.31 mL, 5.81 mmol) was added to methyl trimethylsilylacetate (0.99 mL, 6.03 mmol) in THF (6.80 ml) at -78 °C. After 30 min the TBDMS protected tetralone[2] (1.12 g, 4.02 mmol in 2 mL THF) was added drop wise over 10 min and the reaction warmed slowly to 0 °C and stirred for 3 h. The reaction was quenched with satd. aq. NH_4Cl (12 mL), diluted with water (50 mL) and extracted with EtOAc (3 × 60 mL), washed with brine (50 mL), dried (MgSO$_4$) and concentrated *in vacuo*. Purification by flash chromatography gave TBDMS protected N*p*CE (0.77 g, 2.32 mmol, 66%) as a 1:1 mixture of *E*/*Z* isomers isolated as a colourless oil; R$_f$ 0.58 (25% EtOAc/hexane); IR ν$_{max}$ 2930 (C–H), 1712 (C=O), 1595 (Ar), 1254 (Si–C), 1156 (Si–O) cm^{-1}. ^{1}H NMR signals for the separate isomers from a mixture enriched in one of them were assigned. The alkene stereochemistry was assigned by correlation of the diagnostic vinylic protons from N*p*CE: desired *E*-isomer ^{1}H NMR (600 MHz, CDCl$_3$) δ 7.55 (1H, d, J = 8.7 Hz, SiOCCHC*H*), 6.68 (1H, dd, J = 8.6, 2.4 Hz, SiOC*H*CH), 6.60 (1H, m, SiOCC*H*CCH$_2$), 6.24 (1H, s, C*H*CO$_2$CH$_3$), 3.73 (3H, s, C*H*$_3$), 3.23 – 3.13 (2H, m, SiOCHC(CH$_2$)$_2$C*H*$_2$), 2.74 (2H, t, J = 6.1 Hz, SiOCHCC*H*$_2$(CH$_2$)$_2$), 1.88 – 1.80 (2H, m, SiOCHCCH$_2$C*H*$_2$CH$_2$), 0.99 (9H, s, (C*H*$_3$)$_3$), 0.22 (6H, s, (C*H*$_3$)$_2$). *Z*-isomer δ 7.58 (1H, d, J = 8.6 Hz, SiOCCHC*H*), 6.63 (1H, dd, J = 8.6, 2.4 Hz, SiOCC*H*CH), 6.60 (1H, m, SiOCC*H*CCH$_2$), 5.71 (1H, s, C*H*CO$_2$CH$_3$), 3.69 (s, 3H, C*H*$_3$), 2.82 (t, J = 6.5 Hz, 2H, SiOCHCC*H*$_2$(CH$_2$)$_2$), 2.48 (dd, J = 6.7, 5.4 Hz, 2H, SiOCHC(CH$_2$)$_2$C*H*$_2$), 1.98 – 1.91 (m, 2H, SiOCHCCH$_2$C*H*$_2$CH$_2$), 0.98 (s, 9H, (C*H*$_3$)$_3$), 0.21 (s, 6H, (C*H*$_3$)$_2$). *Z*-isomer δ 7.58 (1H, d, J = 8.6 Hz, SiOCCHC*H*), 6.63 (1H, dd, J = 8.6, 2.4 Hz, SiOCC*H*CH), 6.60 (1H, m, SiOCC*H*CCH$_2$), 5.71 (1H, s, C*H*CO$_2$CH$_3$), 3.69 (3H, s, C*H*$_3$), 2.82 (2H, t, J = 6.5 Hz, SiOCHCC*H*$_2$(CH$_2$)$_2$), 2.48 (2H, dd, J = 6.7, 5.4 Hz, SiOCHC(CH$_2$)$_2$C*H*$_2$), 1.98 – 1.91 (2H, m, SiOCHCCH$_2$C*H*$_2$CH$_2$), 0.98 (9H, s, (C*H*$_3$)$_3$), 0.21 (6H, s, (C*H*$_3$)$_2$). ^{13}C NMR (151 MHz, CDCl$_3$) 1:1 mixture of *E*/*Z*-isomers δ 167.82 (*C*OOMe), 167.81(*C*OOMe), 157.25 (*E*-SiO*C*), 156.97 (*Z*-SiO*C*), 155.21 (*Z*-CH$_2$*C*CHCOOMe), 153.64 (*E*-CH$_2$*C*CHCOOMe), 142.31 (*E*-SiOCCH*C*CH$_2$), 141.13 (*Z*-SiOCCH*C*CH$_2$), 131.30 (*Z*-SiOCCH*C*H), 127.48 (*E*-CH$_2$*C*CHCOOMe), 126.46 (*E*-SiOCCH*C*H), 126.26 (*Z*-CH$_2$*C*CHCOOMe), 120.05 (SiOC*C*HCCH$_2$), 119.39 (SiOC*C*HCCH$_2$), 118.59 (*E*-SiOC*C*HCH), 116.89 (*Z*-SiOC*C*HCH), 112.27 (*Z*-*C*HCO$_2$CH$_3$), 109.97 (*E*-*C*HCO$_2$CH$_3$), 51.29 (*C*H$_3$), 51.07 (*C*H$_3$), 35.77 (*Z*-SiOCHC(CH$_2$)$_2$*C*H$_2$), 30.52 (*E*-SiOCHC*C*H$_2$(CH$_2$)$_2$), 29.71 (*Z*-SiOCHC*C*H$_2$(CH$_2$)$_2$), 28.29 (*E*-SiOCHC(CH$_2$)$_2$*C*H$_2$), 25.77 ((*C*H$_3$)$_3$), 25.76 ((*C*H$_3$)$_3$), 23.43 (*Z*-SiOCHCCH$_2$*C*H$_2$CH$_2$), 22.77 (*E*-SiOCHCCH$_2$*C*H$_2$CH$_2$), 18.33 (*C*(CH$_3$)$_3$), 18.31 (*C*(CH$_3$)$_3$), -4.22 (*C*H$_3$)$_2$), -4.23 (*C*H$_3$)$_2$). HRMS $[\mathrm{C_{19}H_{29}O_3Si}]^+$ calcd. 333.1860, found 333.1886.

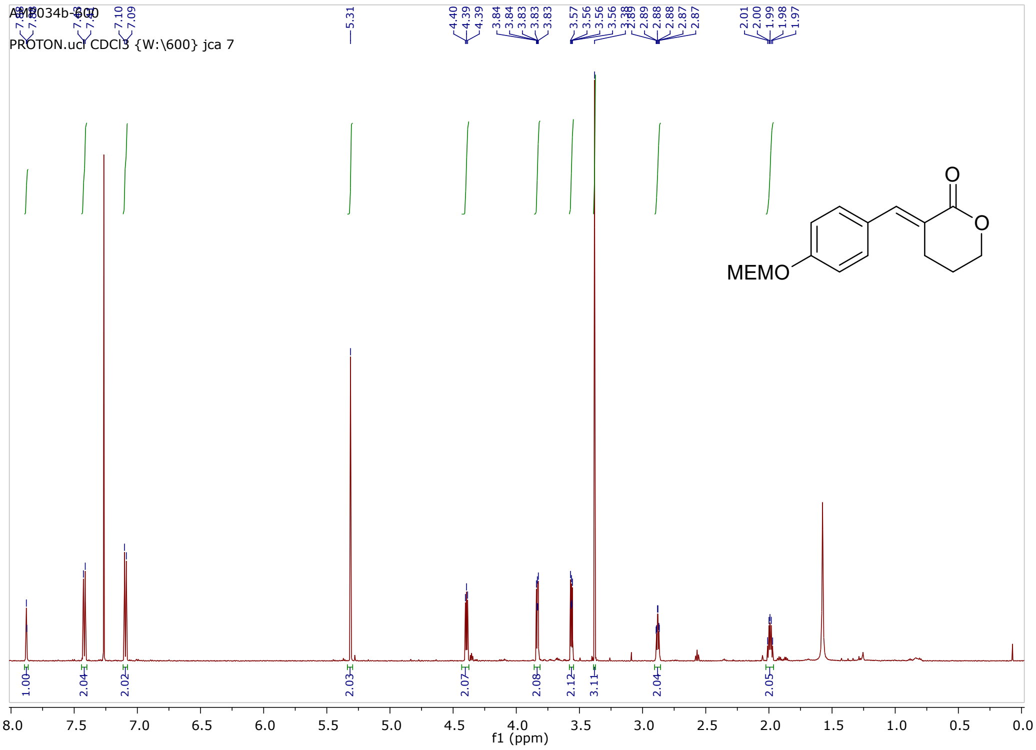
Methyl (*E*)-2-(6-hydroxy-3,4-dihydronaphthalen-1(2*H*)-ylidene)acetate (N*p*CE)

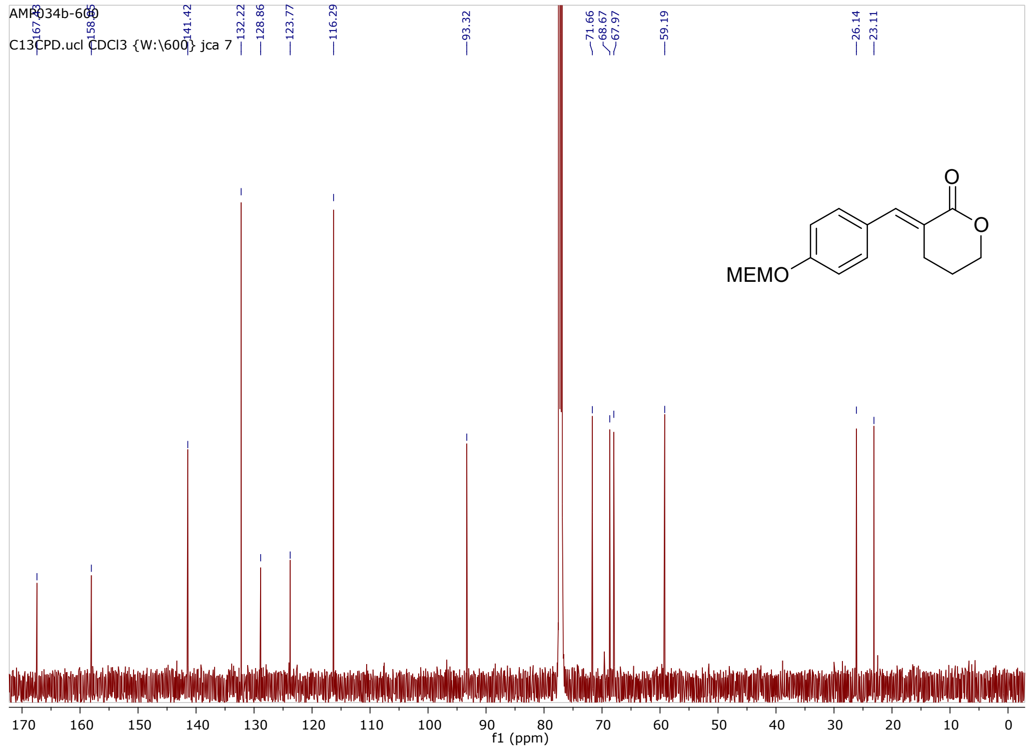


A solution of TBAF (1 M in THF, 1.89 mL, 1.89 mmol) was added to a 1:1 mixture of *E*/*Z*-isomers of TBDMS protected N*p*CE (0.52 g, 1.58 mmol) in THF (22 mL) at 0 °C and the solution left to stir for 5 min. The solution was then quenched with satd. aq. NH$_4$Cl (11 mL), diluted with water (20 mL) and extracted with EtOAc (3 × 30 mL), washed with brine (40 mL), dried (MgSO$_4$) and concentrated *in vacuo*. Purification by flash chromatography gave **N*p*CE** (0.23 g, 1.03 mmol, 65%) as white crystals; m.p. 174 °C ; R$_f$ 0.31 (25% EtOAc/hexane); IR ν$_{max}$ 3343 (O–H), 2919 (C–H), 1687 (C=O), 1586 (Ar) cm^{-1}; ^{1}H NMR (600 MHz, CDCl$_3$) δ 7.57 (1H, d, J = 8.5 Hz, HOCHC*H*), 6.69 (1H, d, J = 8.0 Hz, HOC*H*CH), 6.62 (1H, s, HOC*H*C), 6.23 (1H, s, C*H*CO$_2$Me), 3.74 (3H, s, CO$_2$C*H*$_3$), 3.23 – 3.13 (2H, m, CHCC*H*$_2$), 2.74 (2H, t, J = 5. Hz, CHCC*H*$_2$), 1.89 – 1.78 (2H, m, CH$_2$C*H*$_2$CH$_2$); ^{13}C NMR (151 MHz, CDCl$_3$) δ 168.0 (O*C*OMe), 157.2 (HC*C*CCH$_2$), 155.2 (HC*C*CH$_2$), 142.7 (HC*C*CH$_2$), 127.0 (HO*C*), 126.9 (HOCH*C*HC), 115.3 (HOC*C*HC), 114.1 (HO*C*HCHC), 109.8 (C*C*HCOOMe), 51.1 (COO*C*H$_3$), 30.5 (ArC*C*H$_2$CH$_2$CH$_2$), 28.3 (ArCCH$_2$CH$_2$*C*H$_2$), 22.7 (ArCCH$_2$*C*H$_2$CH$_2$); HRMS $[\mathrm{C_{11}H_{18}O_4} + \mathrm{NH_4}]^+$ calcd. 232.1543, found 232.1544.

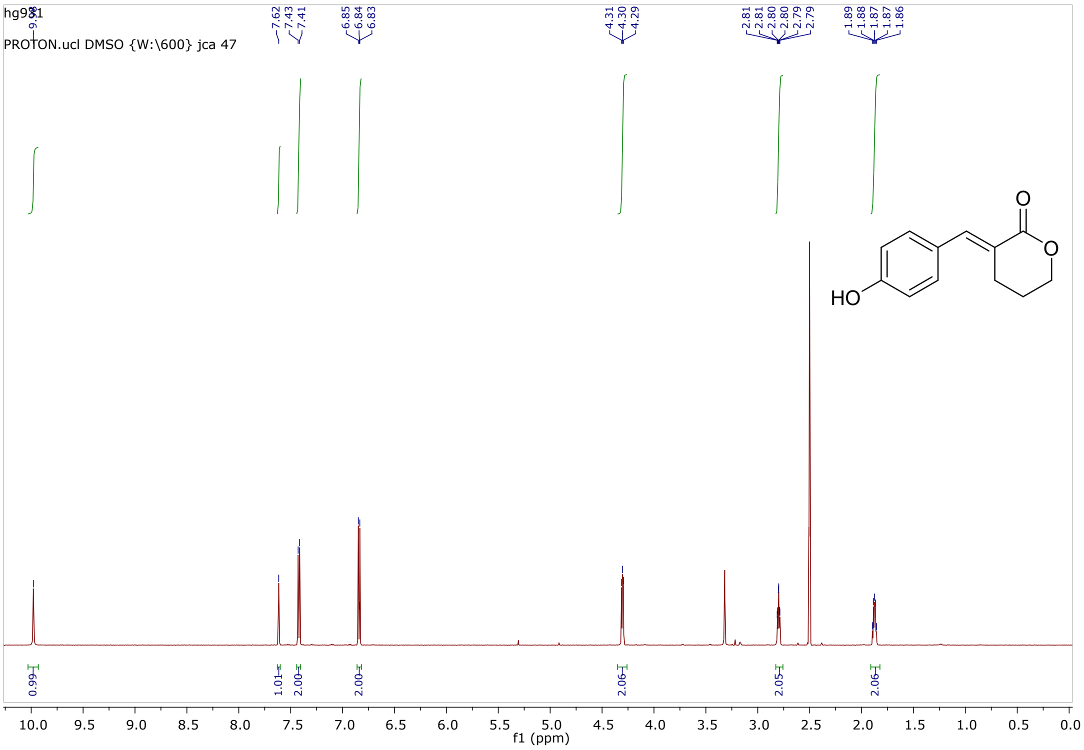
NMR spectra

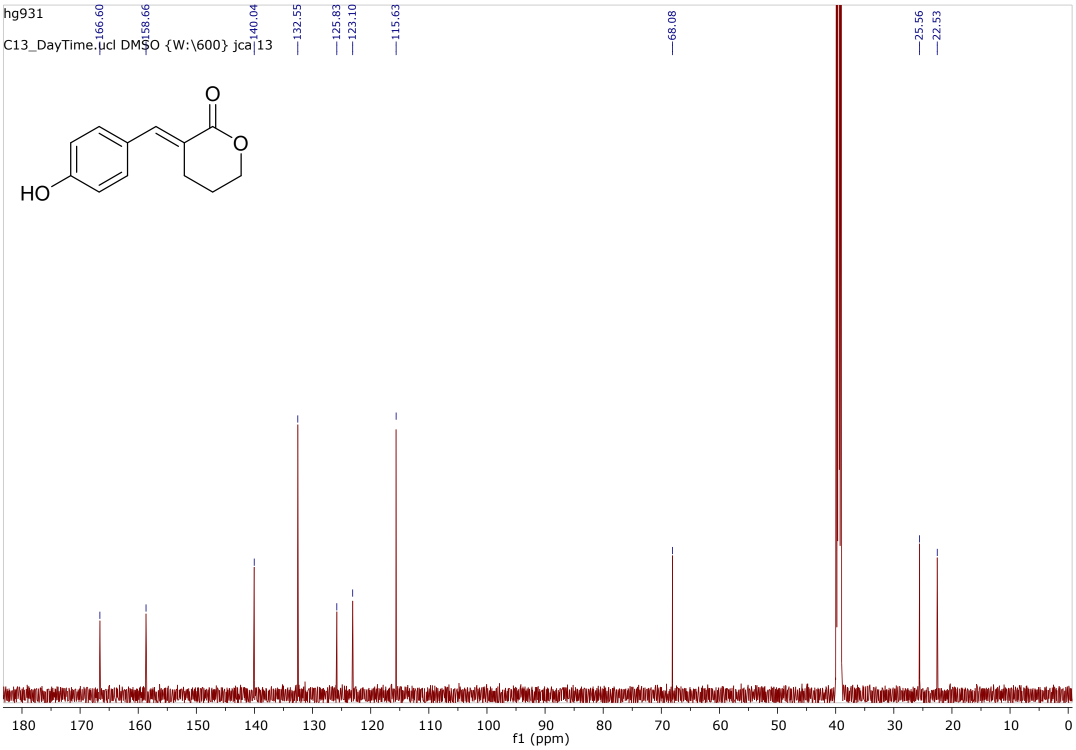
(*E*)-3-(4-((2-methoxyethoxy)methoxy)benzylidene)tetrahydro-2*H*-pyran-2-one



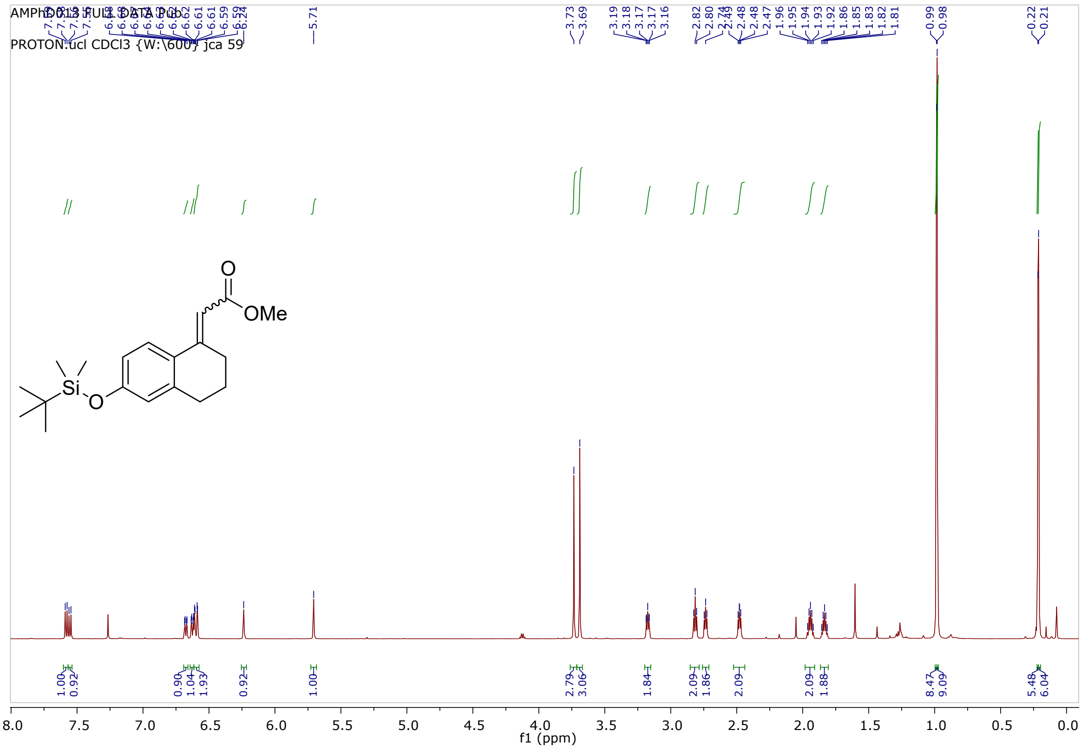


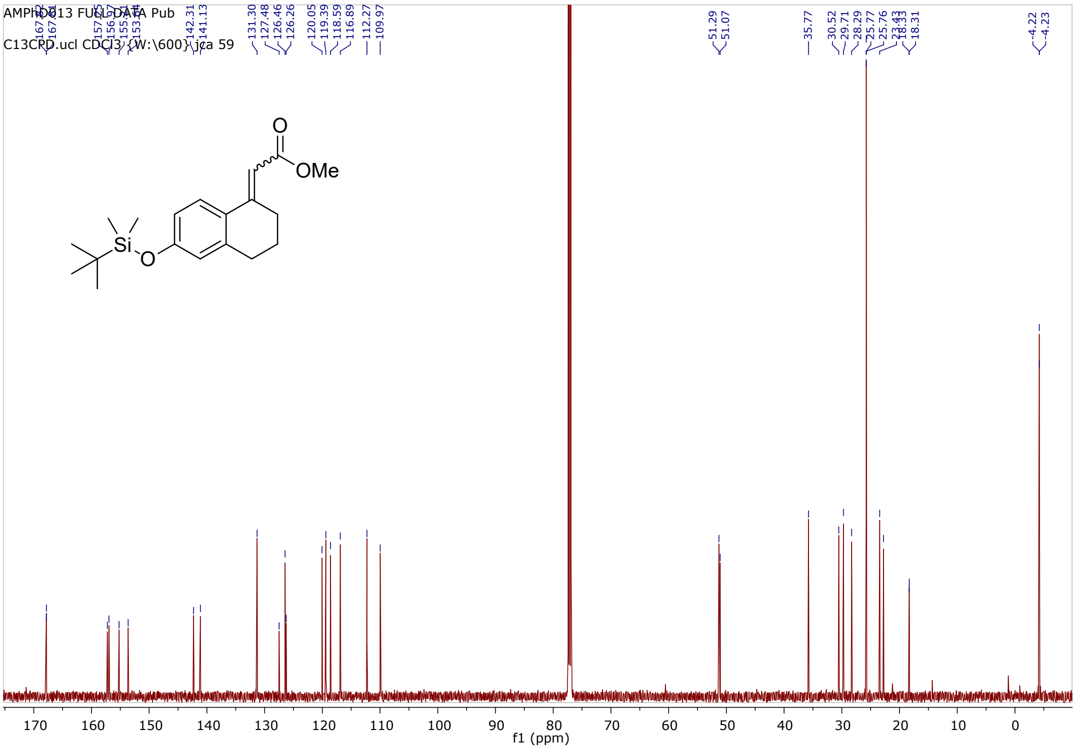
(*E*)-3-(4-hydroxybenzylidene)tetrahydro-2*H*-pyran-2-one (P*p*CE)



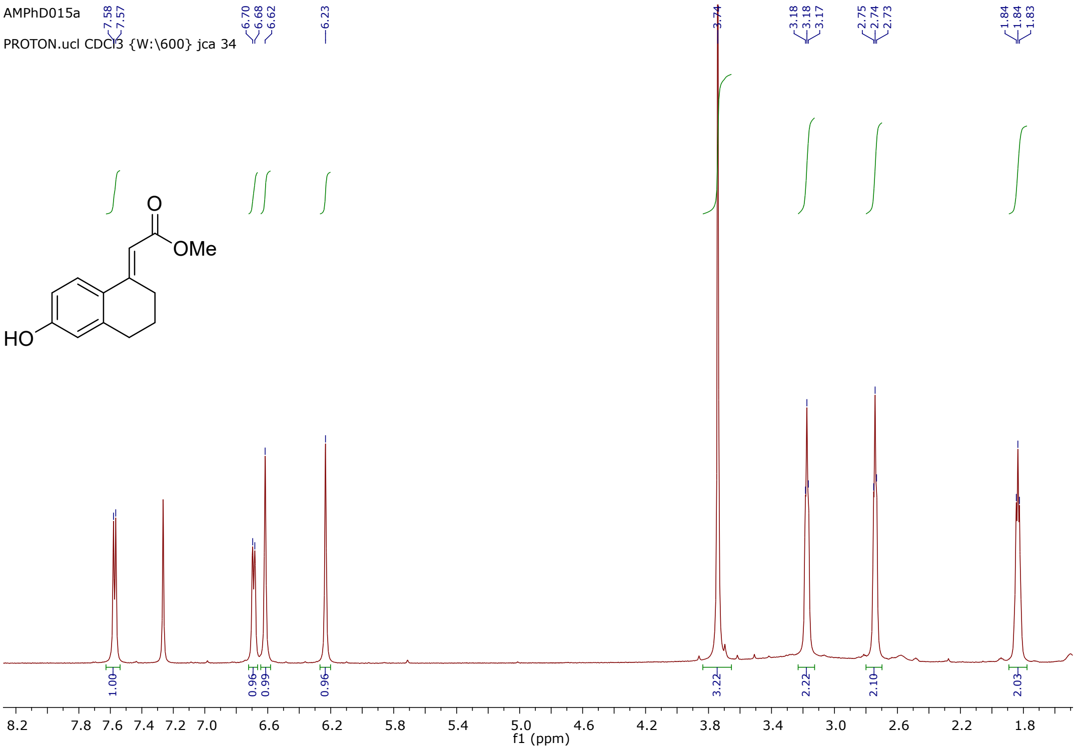


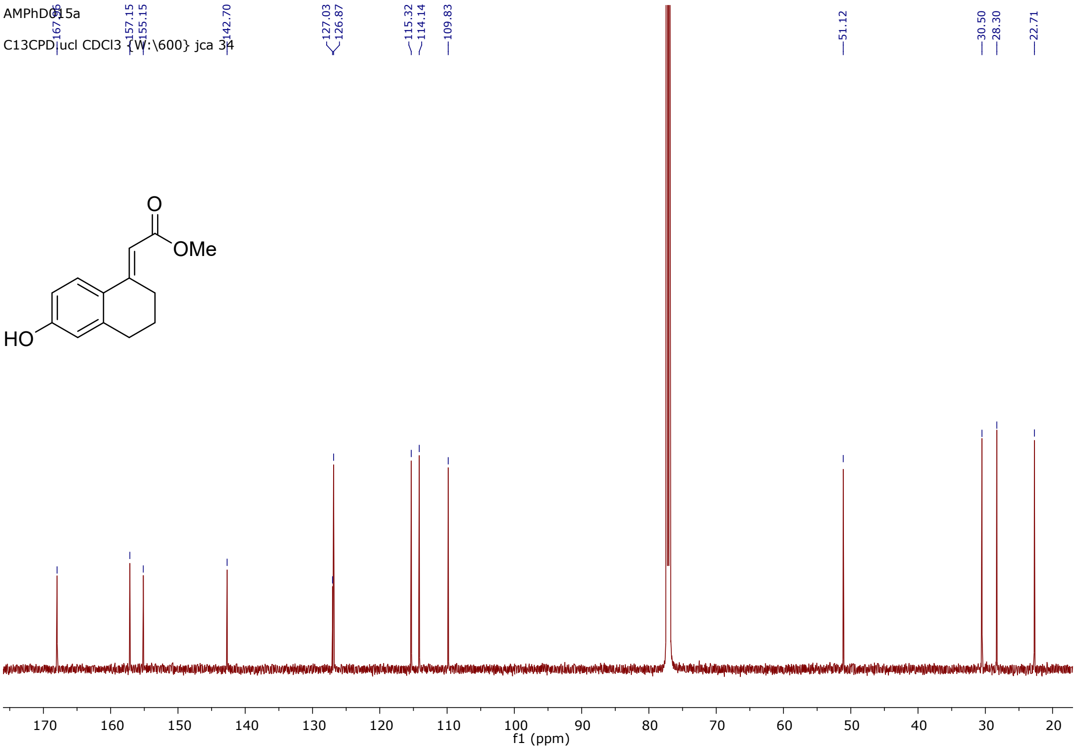
Methyl (*E*/*Z*)-2-(6-((tert-butyldimethylsilyl)oxy)-3,4-dihydronapthalen-1(*2*H)-ylidene)acetate





Methyl (*E*)-2-(6-hydroxy-3,4-dihydronaphthalen-1(2*H*)-ylidene)acetate (N*p*CE)





VMI images

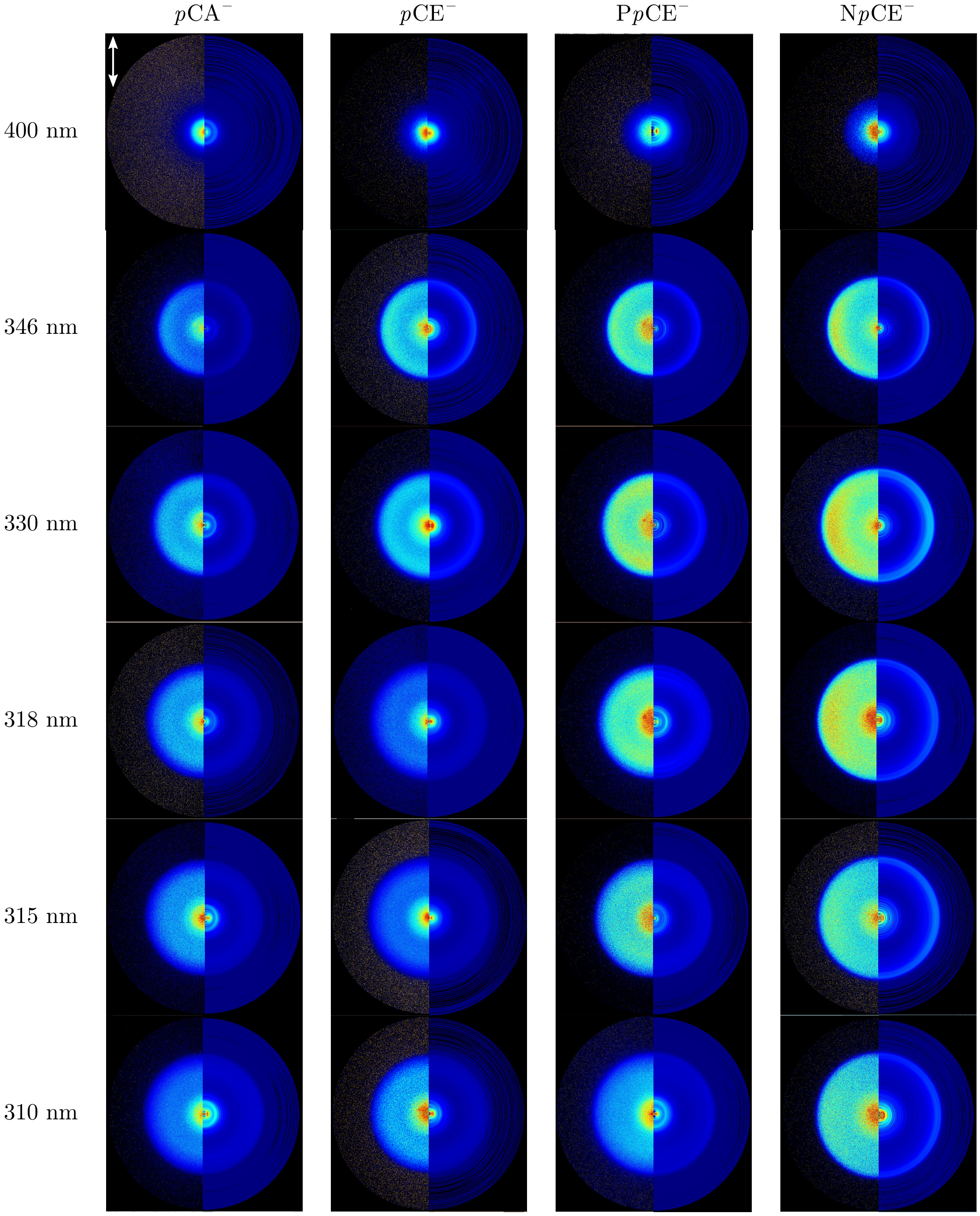
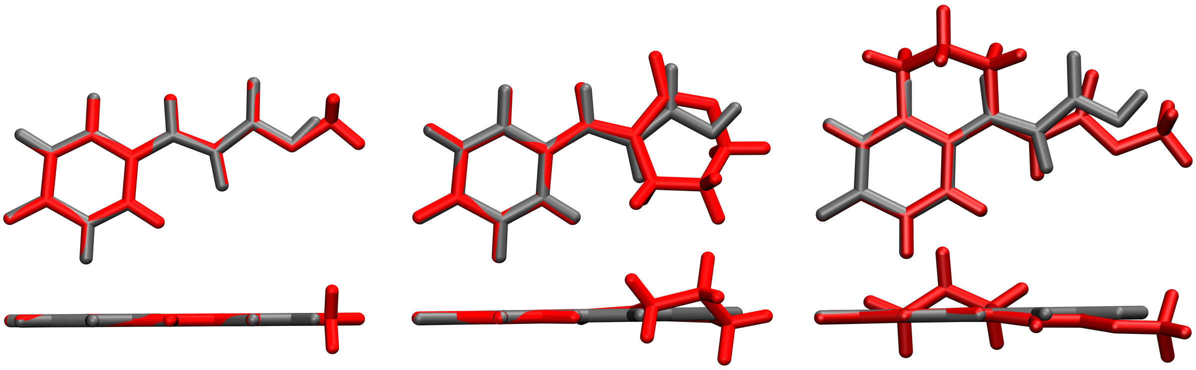


Figure S1: Photoelectron images recorded for each chromophore (left halves: raw VMI images; right halves: pBasex inverted images). In all images the polarisation axis of the laser is parallel to the double headed arrow (top left).

Computational

Structural comparisons



(a) pCE$^-$ (b) PpCE$^-$ (c) NpCE$^-$

Figure S2: Optimised anion structures of pCE$^-$, PpCE$^-$ and NpCE$^-$ (each shown in red) superimposed on the optimised pCA$^-$ structure (grey) with the positions of the carbon atoms in the phenol ring overlapped. Comparisons are viewed from above (top) and the side (bottom).

Benchmarking of vertical excitation energy calculations

Method	Basis set	S$_1$ ($1^1\pi\pi^*$)	S$_2$ ($1^1n\pi^*$)	S$_3$ ($2^1\pi\pi^*$)
TD-CAM-B3LYP	6-311++G(3df,3pd)	3.45 (0.91)	4.10 (0.00)	4.33 (0.11)
EOM-EE-CCSD^a	aug-cc-pVDZ	3.15 (1.06)	4.02 (0.00)	4.23 (0.10)
ADC(2)	cc-pVDZ	3.17 (1.01)	3.41 (0.00)	4.34 (0.09)
	aug-cc-pVDZ	2.96 (0.96)	3.33 (0.00)	4.05 (0.11)
	6-31+G**	2.96 (1.00)	3.35 (0.00)	4.13 (0.09)
	6-311+G**	2.96 (0.99)	3.33 (0.00)	4.11 (0.09)
SA-CASSCF(14,12)-PT2	cc-pVDZ	2.96	3.65	3.82
Experiment		2.89	-	-

Table S1: Vertical excitation energies (VEEs) of pCA$^-$ calculated using various methods compared with SA-CASSCF(14,12)-PT2/cc-pVDZ[3] calculated values and experimental values.[4] Oscillator strengths are shown in brackets.

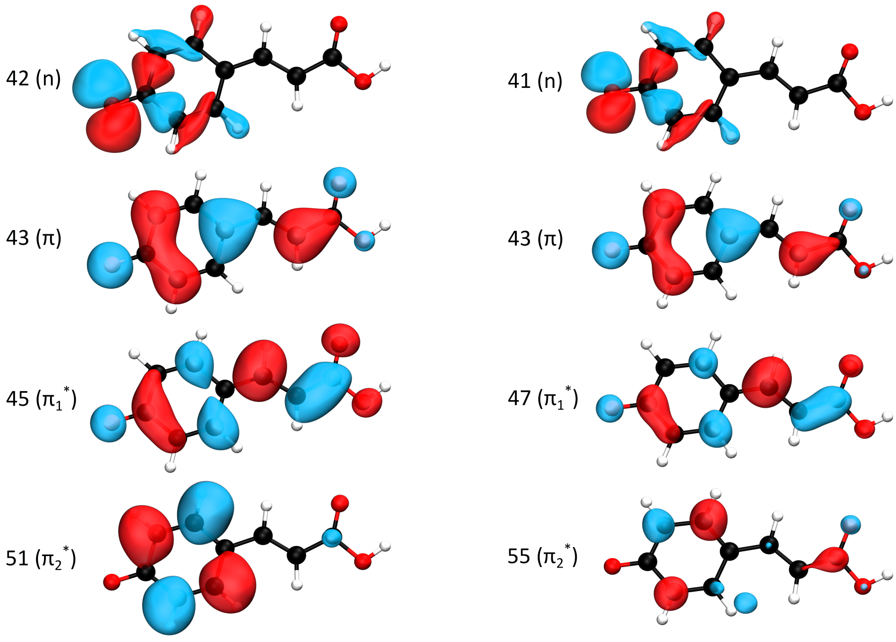
---

[a] EOM-EE-CCSD[5] calculations were performed within the Q-Chem software package.

TD-CAM-B3LYP and ADC(2) excited state calculations

method	state	VEE / eV	main contribution(s)		weighting	f
TD-CAMB3LYP	S_1	3.447	43→45	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_1^*$	0.956	0.910
	S_2	4.098	42→45	$\mathrm{n}\rightarrow\pi_1^*$	0.811	0.000
	S_3	4.333	43→51	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_2^*$	0.900	0.102
ADC(2)	S_1	2.959	43→47	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_1^*$	0.814	1.000
	S_2	3.346	41→47	$\mathrm{n}\rightarrow\pi_1^*$	0.739	0.000
	S_3	4.134	43→55	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_2^*$	0.501	0.091

Table S2: Excited states calculated with TD-CAMB3LYP/6-311++G(3df,3pd) and ADC(2)/6-31+G** for pCA$^-$. The vertical excitation energies (VEEs), main orbital contributions, weighting (c^2) of each contribution and oscillator strengths (f) for each transition are listed.



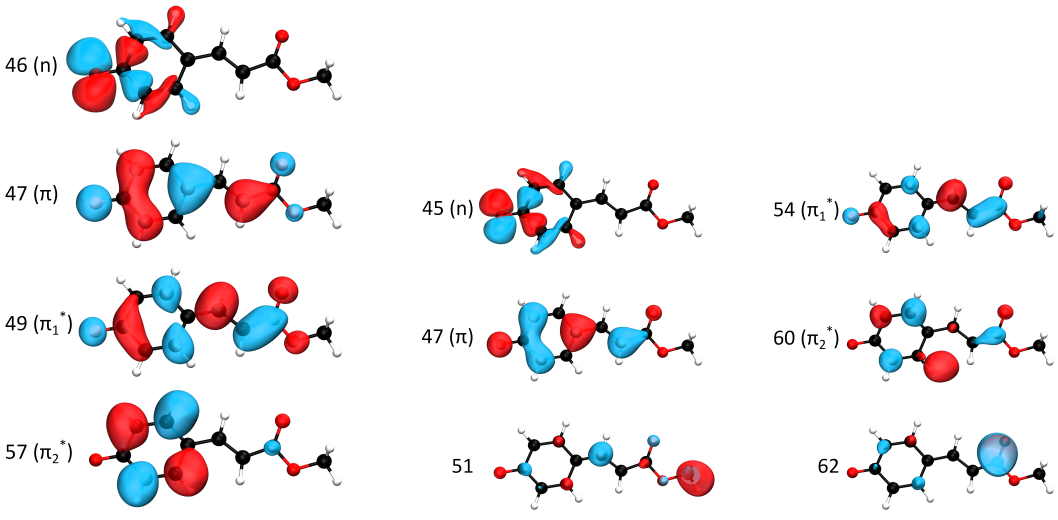
(a) TD-CAMB3LYP

(b) ADC(2)

Figure S3: Relevant orbitals for pCA$^-$ as listed in Table S2.

method	state	VEE / eV	main contribution(s)		weighting	f
TD-CAMB3LYP	S_1	3.434	47→49	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_1^*$	0.954	0.949
	S_2	4.124	46→49	$n\rightarrow\pi_1^*$	0.832	0.000
	S_3	4.311	47→57	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_2^*$	0.894	0.110
ADC(2)	S_1	2.957	47→54	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_1^*$	0.527	1.021
			47→51		0.378	
	S_2	3.368	45→54	$n\rightarrow\pi_1^*$	0.484	0.000
			45→51		0.250	
	S_3	4.101	47→60	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_2^*$	0.476	0.097
			47→62		0.098	

Table S3: Excited states calculated with TD-CAMB3LYP/6-311++G(3df,3pd) and ADC(2)/6-31+G** for pCE$^-$. The vertical excitation energies (VEEs), main orbital contributions, weighting (c^2) of each contribution and oscillator strengths (f) for each transition are listed.

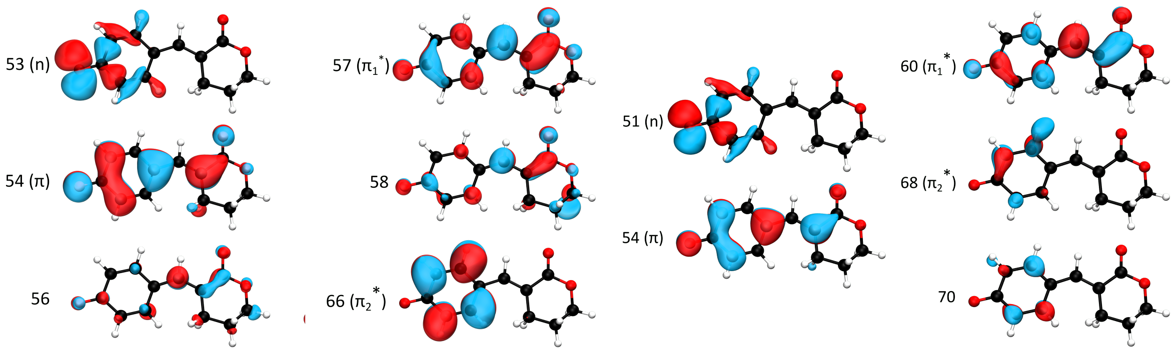


(a) TD-CAMB3LYP (b) ADC(2)

Figure S4: Relevant orbitals for pCE$^-$ as listed in Table S3.

method	state	VEE / eV	main contribution(s)		weighting	f
TD-CAMB3LYP	S_1	3.488	54→57	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_1^*$	0.594	0.952
			54→58		0.199	
			54→56		0.162	
	S_2	4.151	53→57	$\mathrm{n}\rightarrow\pi_1^*$	0.493	0.000
			53→58		0.188	
			53→56		0.135	
	S_3	4.340	54→66	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_2^*$	0.760	0.056
ADC(2)	S_1	3.019	54→60	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_1^*$	0.611	1.032
	S_2	3.389	51→60	$\mathrm{n}\rightarrow\pi_1^*$	0.671	0.000
	S_3	4.103	54→68	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_2^*$	0.242	0.061
			54→70		0.204	

Table S4: Excited states calculated with TD-CAMB3LYP/6-311++G(3df,3pd) and ADC(2)/6-31+G** for P*p*CE$^-$. The vertical excitation energies (VEEs), main orbital contributions, weighting (c^2) of each contribution and oscillator strengths (f) for each transition are listed.



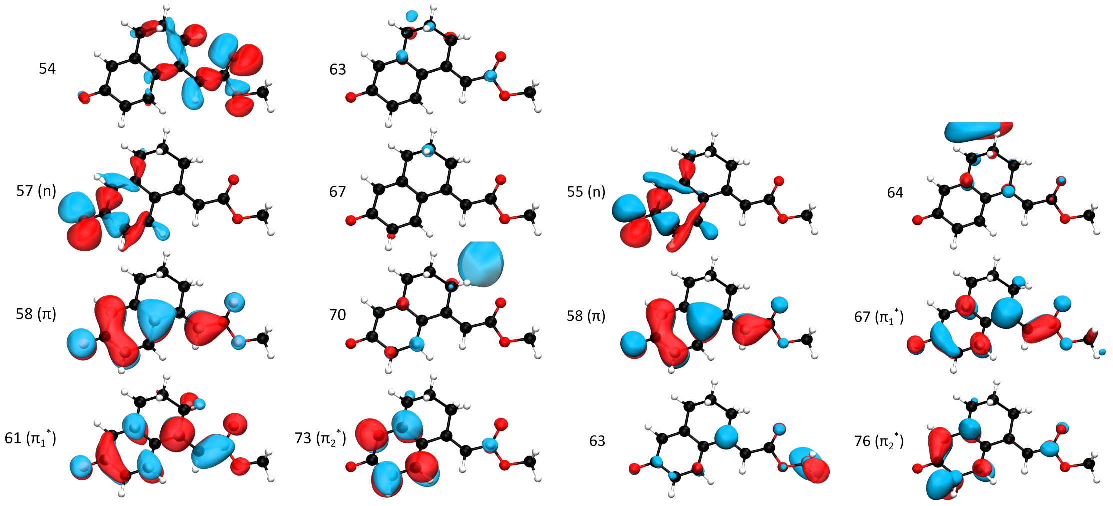
(a) TD-CAMB3LYP

(b) ADC(2)

Figure S5: Relevant orbitals for P*p*CE$^-$ as listed in Table S4.

method	state	VEE / eV	main contribution(s)		weighting	f
TD-CAMB3LYP	S_1	3.356	58→61	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_1^*$	0.876	0.812
	S_2	4.103	57→61	$\mathrm{n}\rightarrow\pi_1^*$	0.790	0.000
	S_3	4.402	58→73	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_2^*$	0.320	0.026
			58→63		0.177	
			58→70		0.120	
			54→67		0.082	
ADC(2)	S_1	2.822	58→67	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_1^*$	0.492	0.885
			58→63		0.216	
	S_2	3.258	55→67	$\mathrm{n}\rightarrow\pi_1^*$	0.464	0.000
			55→63		0.180	
	S_3	3.936	58→64	$\pi_{(\mathrm{HOMO})}\rightarrow\pi_2^*$	0.173	0.065
			58→76		0.137	

Table S5: Excited states calculated with TD-CAMB3LYP/6-311++G(3df,3pd) and ADC(2)/6-31+G** for N*p*CE$^-$. The vertical excitation energies (VEEs), main orbital contributions, weighting (c^2) of each contribution and oscillator strengths (f) for each transition are listed.



(a) TD-CAMB3LYP (b) ADC(2)

Figure S6: Relevant orbitals for N*p*CE$^-$ as listed in Table S5.

Coordinates of optimised structures

atom	anion coordinates /Å			neutral coordinates /Å		
	x	y	z	x	y	z
C	1.018412	-1.064838	-0.000039	0.997621	-1.075030	0.000084
C	0.430369	0.231492	-0.000022	0.442588	0.239254	0.000066
C	1.334833	1.328807	-0.000018	1.332794	1.349107	0.000018
C	2.692494	1.168728	0.000034	2.684188	1.180238	-0.000057
C	3.312270	-0.141529	0.000177	3.273448	-0.150401	-0.000178
C	2.370662	-1.249878	0.000010	2.341620	-1.271767	0.000010
H	0.368707	-1.932636	-0.000095	0.334390	-1.928733	0.000162
H	0.915051	2.330359	-0.000063	0.909348	2.345539	0.000050
H	3.355428	2.025466	0.000004	3.365451	2.019676	-0.000071
H	2.795100	-2.246788	-0.000033	2.772152	-2.263568	0.000044
O	4.551509	-0.311296	-0.000031	4.500182	-0.322802	-0.000064
C	-0.965424	0.468561	-0.000063	-0.969302	0.492960	0.000141
H	-1.261813	1.514094	-0.000059	-1.272483	1.534202	0.000142
C	-2.003486	-0.426630	-0.000124	-1.964757	-0.420544	0.000230
H	-1.841023	-1.494463	-0.000134	-1.787680	-1.485794	0.000238
C	-3.365566	0.015795	-0.000170	-3.370926	0.014994	0.000346
O	-4.263617	-1.041407	0.000126	-4.215052	-1.044908	-0.000263
H	-5.130141	-0.616203	0.000281	-5.115512	-0.690566	-0.000636
O	-3.806480	1.157343	0.000077	-3.761294	1.157258	-0.000159

Table S6: Cartesian coordinates of the anionic and neutral radical forms of deprotonated pCA optimised using B3LYP/6-311++G(3df,3pd).

atom	anion coordinates /Å			neutral coordinates /Å		
	x	y	z	x	y	z
O	5.040363	-0.421536	-0.000737	4.934763	-0.415475	-0.000628
C	3.807044	-0.209403	-0.000578	3.713587	-0.204842	0.000348
C	2.827955	-1.284227	-0.000534	2.746950	-1.296053	-0.000166
H	3.217527	-2.295342	-0.000655	3.146079	-2.300968	-0.000458
C	1.482221	-1.052785	-0.000351	1.409623	-1.057445	-0.000035
H	0.803334	-1.898061	-0.000328	0.720256	-1.890303	-0.000152
C	0.938949	0.262120	-0.000186	0.894959	0.273096	0.000219
C	1.880055	1.327220	-0.000223	1.820092	1.354066	0.000100
H	1.494985	2.342753	-0.000100	1.427931	2.363337	0.000079
C	3.232121	1.120653	-0.000404	3.165636	1.143047	-0.000031
H	3.923575	1.954610	-0.000424	3.872525	1.961090	-0.000230
C	-0.449818	0.547014	0.000006	-0.508664	0.570412	0.000350
H	-0.709146	1.602371	0.000115	-0.779359	1.620604	0.000452
C	-1.516524	-0.311322	0.000066	-1.531701	-0.311594	0.000342
H	-1.390724	-1.384198	-0.000025	-1.386422	-1.381795	0.000240
C	-2.866735	0.177323	0.000262	-2.927611	0.163775	0.000458
O	-3.263149	1.332272	0.000588	-3.279053	1.319820	0.000887
O	-3.777136	-0.869366	0.000573	-3.785242	-0.877072	0.000832
C	-5.142836	-0.485479	0.001035	-5.180067	-0.539599	0.001382
H	-5.715318	-1.411236	0.001222	-5.710313	-1.486266	0.001623
H	-5.389427	0.105727	0.884306	-5.430097	0.040398	0.887661
H	-5.390018	0.105737	-0.882064	-5.430802	0.040344	-0.884733

Table S7: Cartesian coordinates of the anionic and neutral radical forms of deprotonated *p*CE optimised using B3LYP/6-311++G(3df,3pd).

atom	anion coordinates /Å			neutral coordinates /Å		
	x	y	z	x	y	z
C	3.512426	1.100380	-0.049477	3.502828	1.099786	-0.156930
C	2.178088	1.390385	-0.053311	2.177769	1.401511	-0.125727
C	1.158571	0.392942	0.005570	1.172244	0.399455	0.045449
C	1.627792	-0.952052	0.094465	1.608315	-0.947434	0.230676
C	2.959202	-1.263160	0.106850	2.929791	-1.273189	0.215701
C	4.004682	-0.260539	0.026791	3.963684	-0.269802	0.006903
H	4.253084	1.889454	-0.100122	4.257793	1.860477	-0.299670
H	1.860894	2.427091	-0.108685	1.854012	2.427623	-0.245041
H	0.918053	-1.761358	0.173716	0.884425	-1.724334	0.411857
H	3.281456	-2.294981	0.183938	3.262628	-2.291046	0.366229
O	5.222259	-0.548198	0.030887	5.166535	-0.567976	-0.019176
C	-0.191419	0.832920	-0.011936	-0.188975	0.845796	0.040765
H	-0.294410	1.912603	-0.001009	-0.306583	1.922488	0.061477
C	-1.406354	0.183549	-0.040829	-1.370087	0.168191	-0.008520
C	-2.584904	1.044476	-0.005074	-2.591646	1.045220	0.041444
O	-2.598793	2.258277	-0.077620	-2.544873	2.247752	-0.019993
C	-3.928693	-0.941593	0.413073	-3.940845	-0.963960	0.314467
H	-3.709763	-1.100980	1.476151	-3.801647	-1.209405	1.371261
H	-4.972944	-1.205916	0.241160	-4.973379	-1.177394	0.048457
C	-2.980079	-1.747951	-0.446133	-2.953500	-1.707565	-0.554281
H	-3.118628	-2.816637	-0.263974	-3.105220	-2.782638	-0.452352
H	-3.215203	-1.558474	-1.496640	-3.129823	-1.449515	-1.600511
O	-3.833828	0.447194	0.111548	-3.804373	0.458251	0.150871
C	-1.549974	-1.320393	-0.129932	-1.535701	-1.325035	-0.137761
H	-1.242205	-1.794182	0.810843	-1.311871	-1.800846	0.823736
H	-0.873480	-1.708592	-0.894251	-0.811915	-1.717464	-0.852168

Table S8: Cartesian coordinates of the anionic and neutral radical forms of deprotonated P*p*CE optimised using B3LYP/6-311++G(3df,3pd).

atom	anion coordinates /Å			neutral coordinates /Å		
	x	y	z	x	y	z
C	-1.219226	-1.510111	0.164766	-1.203935	-1.503055	0.328132
C	-0.868200	-0.135789	0.016059	-0.877619	-0.135663	0.076557
C	-1.963912	0.772699	-0.137207	-1.946623	0.777111	-0.204497
C	-3.262173	0.328396	-0.151006	-3.232921	0.318491	-0.272403
C	-3.634216	-1.060792	-0.016931	-3.581338	-1.071549	-0.052803
C	-2.507080	-1.957822	0.152541	-2.480236	-1.963309	0.271415
H	-0.430767	-2.235928	0.314077	-0.415375	-2.189024	0.599698
H	-4.074297	1.038092	-0.266632	-4.051314	0.992626	-0.489802
H	-2.729937	-3.010204	0.282479	-2.722693	-2.997191	0.474525
O	-4.821684	-1.457871	-0.032506	-4.753665	-1.473667	-0.121392
C	-1.681236	2.246884	-0.304145	-1.632790	2.235698	-0.415461
H	-1.520635	2.473706	-1.365847	-1.346221	2.404622	-1.459224
H	-2.552905	2.824447	0.009680	-2.527844	2.830693	-0.236130
C	-0.434473	2.643743	0.474188	-0.484582	2.655534	0.495603
H	-0.211862	3.706481	0.343531	-0.257605	3.714313	0.369087
H	-0.608851	2.480779	1.541893	-0.780816	2.515233	1.538336
C	0.755234	1.809942	0.012228	0.758015	1.829210	0.182492
H	1.639120	2.016946	0.609719	1.544802	2.009063	0.911788
H	1.031097	2.109628	-1.005899	1.185109	2.156343	-0.769653
C	0.496072	0.314106	0.006393	0.501646	0.337328	0.093895
C	1.570660	-0.555016	-0.028212	1.526781	-0.552471	-0.017060
H	1.389765	-1.616658	-0.075868	1.324339	-1.605463	-0.130166
C	2.961017	-0.212200	-0.028893	2.956578	-0.219165	-0.041381
O	3.515995	0.878824	-0.026014	3.458634	0.881879	-0.002543
O	3.737267	-1.368968	-0.042378	3.690723	-1.354772	-0.123875
C	5.138614	-1.158638	-0.058414	5.112307	-1.178000	-0.162451
H	5.590714	-2.148859	-0.072883	5.530366	-2.177571	-0.224879
H	5.448301	-0.595478	-0.940579	5.400358	-0.588254	-1.030936
H	5.471138	-0.611247	0.825308	5.459645	-0.673871	0.737610

Table S9: Cartesian coordinates of the anionic and neutral radical forms of deprotonated N*p*CE optimised using B3LYP/6-311++G(3df,3pd).

References

[1] H. Konno, S. Aoyama, K. Nosaka and K. Akaji, *Synthesis*, 2007, **2007**, 3666–3672.

[2] X.-D. Jiang, J. Zhang, X. Shao and W. Zhao, *Org. Biomol. Chem.*, 2012, **10**, 1966–1968.

[3] F. Fernández García-Prieto, M. A. Aguilar, I. Fdez. Galván, A. Muñoz-Losa, F. J. Olivares del Valle, M. L. Sánchez and M. E. Martín, *J. Phys. Chem. A*, 2015, **119**, 5504–5514.

[4] 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. Brøndsted Nielsen, *J. Phys. Chem. A*, 2009, **113**, 9442–9449.

[5] A. I. Krylov, *Annu. Rev. Phys. Chem.*, 2008, **59**, 433–462.