8-3-2012

Expanding and Testing a Computational Method for Predicting the Ground State Reduction Potentials of Organic Molecules on the Basis of Empirical Correlation to Experiment

Eugene J. Lynch

Amy L. Speelman

Bryce A. Curry

Charles S. Murillo

Jason G. Gillmore
Hope College, gillmore@hope.edu

Recommended Citation

Lynch, Eugene J., Amy L. Speelman, Bryce A. Curry, Charles S. Murillo and Jason G. Gillmore. "Expanding and Testing a Computational Method for Predicting the Ground State Reduction Potentials of Organic Molecules on the Basis of Empirical Correlation to Experiment." Journal of Organic Chemistry 77, no. 15 (2012): 6423-6430. http://dx.doi.org/10.1021/jo300853k

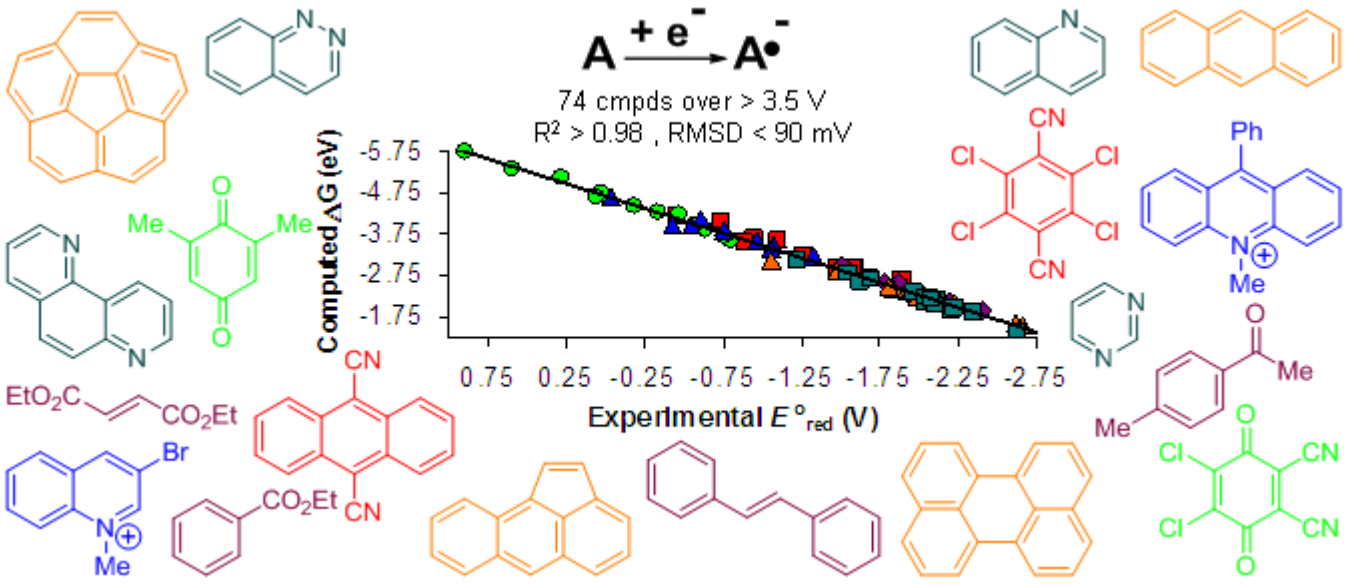
Expanding and testing a computational method for predicting the ground state reduction potentials of organic molecules on the basis of empirical correlation to experiment.

*Eugene J. Lynch, Amy L. Speelman, Bryce A. Curry, Charles S. Murillo, Jason G. Gillmore**

Department of Chemistry, Hope College, 35 E 12th St., Holland, MI 49423

gillmore@hope.edu

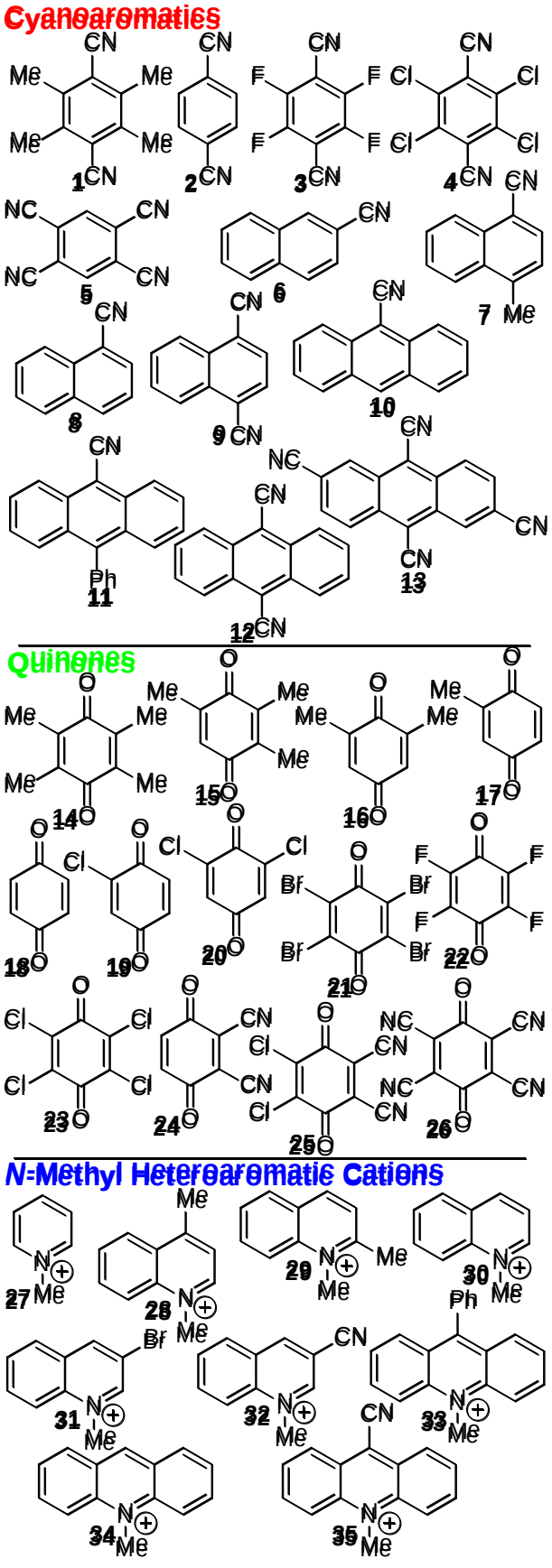
TOC Graphic / Graphical Abstract



ABSTRACT: A method for predicting the ground state reduction potentials of organic molecules based on the correlation of computed energy differences between the starting S_0 and one-electron reduced D_0 species with experimental reduction potentials in acetonitrile has been expanded to cover 3.5 V of potential range and 74 compounds across six broad families of molecules. Utilizing the Conductor-like Polarizable Continuum Model of implicit solvent allows a global correlation that is computationally efficient and of improved accuracy, with $R^2 > 0.98$ in all cases and root mean squared deviation errors < 90 mV (mean absolute deviations < 70 mV) for either B3LYP/6-311+G(d,p) or /6-31G(d) with appropriate choice of radii (UAKS or UA0). The correlations are proven robust across a wide range of structures and potentials, including four larger (27-28 heavy atoms) and more conformationally flexible photochromic molecules not used in calibrating the correlation. The method is also proven robust to a number of minor student 'mistakes' or methodological inconsistencies.

Introduction

The accurate and efficient calculation of oxidation and reduction potentials of molecules in solution, whether in a relative or an absolute sense, remains an area of active interest[1,2,3,4,5,6,7,8,9,10] due in part to the relevance of electron transfer reactions to solar energy conversion[6,9] and other reactions of materials or biological relevance. While the calculation of absolute potentials through a thermodynamic cycle is tractable for smaller molecules,[1,2,3,4] simpler calculations are preferable for the larger molecules often of interest to the aforementioned applications. Simpler comparisons of the energies of the molecule of interest and its corresponding one-electron reduced state can be made. Within a closely related family of molecules over a limited range of potentials, this energy difference (strictly an electron affinity calculation in the gas-phase) can directly correlate very well with the absolute reduction potential of the molecule in solution if a suitable implicit or explicit solvent model is used, though this must be corrected to a relative reduction potential of a reference electrode or secondary standard for direct



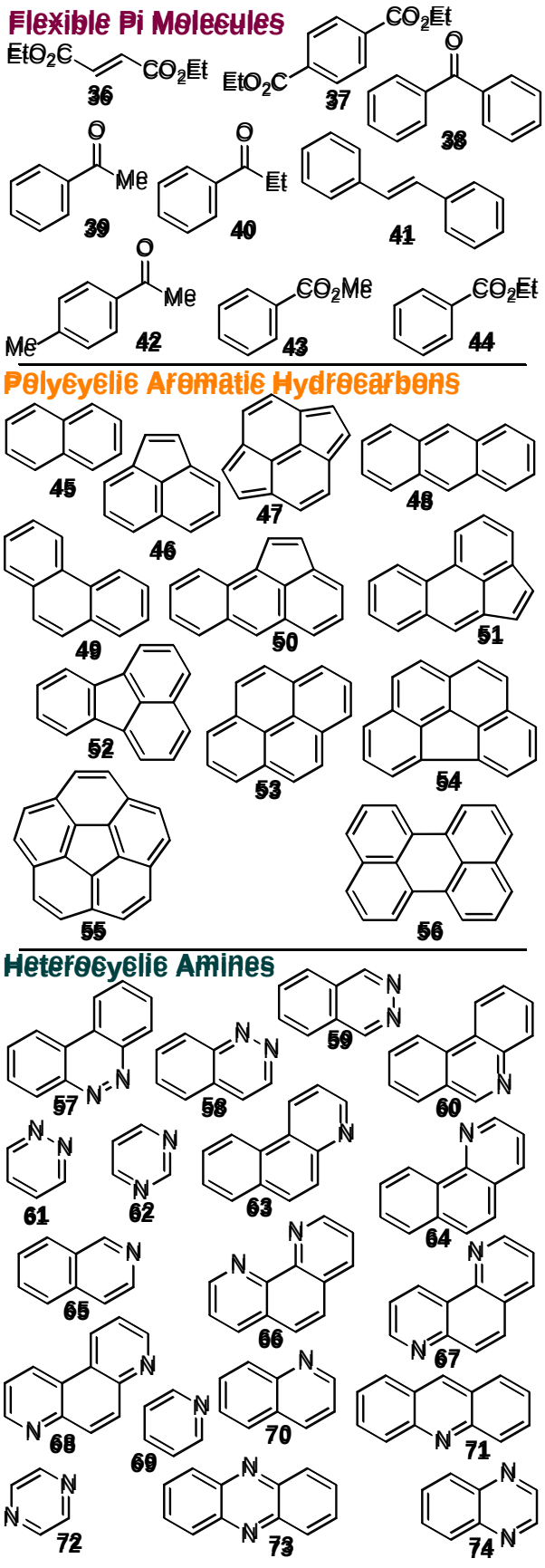
comparison to experiment. We have previously reported an alternative correlational method for predicting reduction potentials that covers a much more diverse range of structures and reduction potentials, either in the gas-phase, or in solution (as modeled with dielectric continuum solvent models.)

Gas phase correlations are, as we previously reported, only valid if one wishes to predict a compound within a well defined and closely related family of structural analogs for which experimental data is available to calibrate the correlation. Generally this is not the case for our research group. Moreover, we have demonstrated that gas phase correlations benefit more significantly from the diffuse functions of the 6-311+G(d,p) basis set than correlations with implicit solvent, where the 6-31G(d) basis set performs very well. The need to use the larger basis set more than offsets the time savings of neglecting solvent. Thus we have chosen to focus our efforts in using, testing, and improving our method of a global correlation using the CPCM (Conductor-like Polarizable Continuum Model) model with acetonitrile (our electrochemical solvent of choice and one for which a

large body of experimental reduction potential data is available in the literature.)

In the course of using our previously published method to predict reduction potentials for some photochromic molecules of interest to our experimental research program, we found that the energy calculations for a few structures, particularly those with the potential for intramolecular hydrogen-bonding, failed to converge when the CPCM solvent model was applied with the default (UA0) radii. Calculations on these structures did converge when the alternative UAKS radii parameters were used. Thus we began a systematic investigation, repeating our previously published correlations varying both basis set and CPCM radii parameters for our initial calibrant set, molecules **1-35**, spanning three families of conventional photooxidants.

Simultaneously, we sought to test (and ultimately expand) our global correlation with a more diverse range of structures covering a wider range of reduction potentials than we had previously reported. Thus we found experimental literature reduction potentials for 39 additional molecules (**36-74**) from three additional broad families, including molecules with more "flexible" pi systems, polycyclic aromatic hydrocarbons (PAHs), and heterocyclic amines, and proceeded to compute the necessary geometries and energies for these molecules as well.. Compounds **36-74** span a total of 1.6 V, including an additional 0.7 V beyond the 2.8 V window spanned by comounds **1-35** molecules we had previously reported.



Finally, in engaging novice high school and two-year community college students in this computational endeavor, it became clear that occasional inconsistencies of method and parameters were occurring. While we faithfully repeated and corrected these anomalous calculations, we also seized these occurrences as opportunities to further test the robustness of our method, specifically developed for use by non-experts, to endure these minor variations.

Computational Details

The procedures we followed were analogous to those we previously reported. All calculations were carried out with the *Gaussian03* software package,[11] implemented through the *WebMO* graphical user interface.[12] Structures were drawn in the *WebMO* interface and preliminary optimizations were performed using that program's "comprehensive clean-up with mechanics" tool, prior to queuing gas-phase geometry optimizations on the cluster using Density Functional Theory on *Gaussian03* with the B3LYP hybrid functional[13,14,15] and either the MIDI![16,17] or 6-31G(d) basis set. The optimal geometry of the S_0 state was used as the starting point for the geometry optimization

of the corresponding D_0 state. Solution phase molecular energies of the S_0 and D_0 species were computed on *Gaussian03* with the B3LYP hybrid functional and either the 6-31G(d) or 6-311+G(d,p) basis sets. Implicit acetonitrile solvent ($\varepsilon = 36.64$) for these single-point energy calculations was implemented using the Conductor-like Polarizable Continuum Model (CPCM),[18,19] with either the default (in *Gaussian03*) UA0 or alternative UAKS radii parameters.

While it is strictly necessary to set the criterion SCF=tight with the larger basis set with diffuse functions, this criterion was inadvertently also used in all single-point energy calculations (except for the control experiment so noted). For consistency we also generally disabled symmetry in all calculations of both geometry and energy (as any time savings due to leaving symmetry enabled was minimal in our experience.)

Results & Discussion

We began this work by re-examining our previously reported correlations of the computed energy differences between the one-electron reduced D_0 and initial S_0 of compounds **1-35** with their literature reduction potentials at two different basis sets and with two different CPCM radii. Though the calculations were repeated from the beginning, the results for the UA0 radii are essentially identical to those we have reported previously. The correlations, along with four different measures of goodness of fit, are reported in Table 1. In addition to R^2 and root-mean-square deviation (RMSD) residuals which we have previously reported, we also include mean absolute deviation (MAD) residuals for comparison to other work in the field. Finally, in an ideal model, if the calculated D_0 - S_0 energy difference, essentially an electron affinity calculated in a dielectric continuum solvent model rather than in the gas phase, truly were "equivalent to" a reduction potential, the *x*-intercept (*-b/m*) should correspond to the reference potential of the reference electrode, and the slope should be unity. For the larger basis set, our

method does come close on both these metrics. However we find it both more useful and more precise to use the correlation rather than the calculated energy difference itself directly.

Table 1. Correlations of computed D_0 - S_0 energy difference with experimental reduction potential for compounds **1-35** varying basis set and CPCM radii.

Corr. #	B3LYP/	CPCM radii	slope *m* (eV/V)	*y*-intercept *b* (eV)	R^2	RMSD[a] residuals (V)	MAD[b] residuals (V)	*x*-intercept[c] (V)
1	6-31G(d)	UA0	-1.1220	-4.2335	0.9861	0.0835	0.0635	-3.7733
2	6-31G(d)	UAKS	-1.1513	-4.2606	0.9730	0.1170	0.1001	-3.7006
3	6-311+G(d,p)	UA0	-1.1214	-4.6193	0.9730	0.1171	0.0953	-4.1192
4	6-311+G(d,p)	UAKS	-1.1495	-4.6838	0.9846	0.0878	0.0705	-4.0746

[a]Root mean square deviation and [b]mean average deviation, taken from individual residuals for each compound as predicted by each trendline (as reported in the Supporting Information). [c]Computed *x*-intercept (= *-b/m*) corresponds to the reference electrode potential correction (SCE = NHE + 0.24 V = -4.12 V),[20] and any systematic inaccuracies of the computational model employed.

As we reported previously, the larger basis set with diffuse functions appears unnecessary to giving good predictive ability (indeed for the UA0 radii the errors are lower for 6-31G(d) than for 6-311+G(d,p)) from this method that includes implicit solvent. However it is clear that not including diffuse functions introduces an unknown systemic error that is accounted for in the correlation but is manifested in the deviation of the *x*-intercept away from the -4.12 V reference potential of the saturated calomel electrode (SCE) to which the experimental data is directly or indirectly referenced. This systemic deviation in intercept appears to be due to the overestimation of the energy of the anion when diffuse functions are not included (both S_0 and D_0 are more negative in 6-311+G(d,p) than in 6-31G(d), but the difference is larger for D_0 than S_0 and thus the difference between S_0 and D_0 increases with diffuse functions, shifting the intercept accordingly). Nevertheless, the method is equally predictive and more computationally efficient if the smaller basis set is used. Interestingly, it appears the default UA0 radii for CPCM pair best with the smaller 6-31G(d) basis set, while the alternative UAKS radii for

CPCM pair better with the larger 6-311+G(d,p) basis set, allowing equally good predictive abilities if the "appropriate" radii are chosen for each basis set.

Table 2. Experimental reduction potentials and calculated $D_0 - S_0$ energy differences for **36-74**.

	Experimental		Calculated $D_0 - S_0$ Energy Difference (eV)			
	literature	lit	6-31G(d)		6-311+G(d,p)	
Cmpd	E^o_{red} (V)[a]	ref	UA0	UAKS	UA0	UAKS
36	-1.5	*21*	-2.345	-2.230	-2.733	-2.648
37	-1.78	*21*	-2.491	-2.466	-2.909	-2.933
38	-1.86	*21*	-3.045	-3.043	-3.552	-3.608
39	-2.1	*21*	-3.302	-3.303	-3.430	-3.648
40	-2.15	*21*	-3.595	-3.700	-3.941	-4.093
41	-2.2	*21*	-1.920	-1.833	-2.352	-2.309
42	-2.2	*21*	-1.990	-1.881	-2.419	-2.353
43	-2.22	*21*	-2.040	-1.946	-2.480	-2.432
44	-2.4	*21*	-2.806	-2.777	-3.201	-3.218
45	-2.66	*22*	-2.538	-2.461	-2.934	-2.898
46	-1.8	*22*	-2.571	-2.476	-2.961	-2.909
47	-1.06	*22*	-3.179	-3.147	-3.552	-3.562
48	-2.1	*22*	-3.745	-3.832	-4.076	-4.206
49	-2.62	*22*	-3.282	-3.121	-3.739	-3.595
50	-1.47	*22*	-3.387	-3.232	-3.867	-3.733
51	-1.81	*22*	-3.478	-3.330	-3.981	-3.861
52	-1.92	*22*	-3.570	-3.436	-4.090	-3.991
53	-2.22	*22*	-3.764	-3.656	-4.320	-4.254
54	-1.84	*22*	-3.893	-3.770	-4.368	-4.289
55	-1.99	*22*	-4.108	-3.993	-4.526	-4.456
56	-1.81	*22*	-4.329	-4.214	-4.691	-4.629
57	-1.554	*23*	-4.153	-4.069	-4.774	-4.742
58	-1.686	*23*	-4.367	-4.263	-4.706	-4.644
59	-1.976	*23*	-4.650	-4.662	-5.074	-5.125
60	-2.118	*23*	-4.927	-4.947	-5.278	-5.337
61	-2.12	*23*	-5.173	-5.301	-5.562	-5.729
62	-2.34	*23*	-2.659	-2.894	-3.002	-3.239
63	-2.14	*23*	-3.013	-3.170	-3.313	-3.478
64	-2.208	*23*	-2.971	-3.125	-3.252	-3.412
65	-2.22	*23*	-3.098	-3.301	-3.405	-3.615
66	-2.042	*23*[b]	-3.361	-3.557	-3.647	-3.859
67	-2.092	*23*	-3.545	-3.810	-3.831	-4.112
68	-2.044	*23*	-3.527	-3.673	-3.814	-3.974
69	-2.636	*23*[b]	-3.481	-3.662	-3.769	-3.961
70	-2.105	*23*	-4.184	-4.382	-4.458	-4.672
71	-1.62	*23*	-2.628	-2.486	-3.125	-3.013
72	-2.08	*23*	-2.279	-2.133	-2.736	-2.623
73	-1.227	*23*	-2.100	-2.068	-2.596	-2.615
74	-1.702	*23*	-1.804	-1.612	-2.387	-2.232

[a]vs. SCE in CH_3CN (or corrected to vs. SCE according to Refs. 20 and 24 for those not so reported in the primary reference). [b]reconfirmed experimentally in our laboratory

Next to further test our method we computed D_0 - S_0 energy differences (Table 2) for thirty-nine additional compounds (**36-74**) of known experimental reduction potential, falling largely outside the range of potentials spanned by compounds **1-35** used to create the four different correlations (Table 1), at the four different combinations of basis set and CPCM radii. Applying the appropriate correlation from Table 1, we were able to predict the reduction potentials of compounds **36-74** and compare them to the experimental data available in the literature. The results of these predictions (notably made on compounds not included in the correlation itself) are reported in Table 3. As expected, in all cases the R^2 of the correlation of experimental reduction with the value predicted from the energy difference on the basis of a correlation not including these molecules is greater than for the correlation reported in Table 1, though still reasonably good (0.94-0.96). Likewise the RMSD and MAD residuals are generally larger (except, oddly, in the case of 6-311+G(d,p) with the UA0 radii) but still respectable. Importantly, there does seem to be some evidence here that the larger basis set can perhaps achieve greater predictive results when the compound one is computing is not a member of the calibrant set (as is typically the case when actually applying our methodology to unknown molecules.) Interestingly the UA0 radii continue to pair well with 6-31G(d), but which radii are better for the larger 6-311+G(d,p) basis set is less clear in this test.

Table 3. Using correlations 1-4 from Table 1 (based on compounds **1-35**) to predict reduction potentials of compounds **36-74**.

B3LYP/	CPCM radii	R^2	RMSD[a] residuals (V)	MAD[b] residuals (V)
6-31G(d)	UA0	0.9402	0.1324	0.1032
6-31G(d)	UAKS	0.9514	0.1918	0.1769
6-311+G(d,p)	UA0	0.9423	0.0842	0.0633
6-311+G(d,p)	UAKS	0.9570	0.1154	0.0987

[a]Root mean square deviation and [b]mean average deviation, taken from individual residuals for each compound as predicted by each trendline.

New correlations based only upon molecules **36-74** are possible and are reported in the supporting information. However the most valuable way to utilize the new data regarding these three additional families of molecules is to incorporate them into a single larger correlation of all 74 compounds. These correlations (5-8) are reported in Table 4. Clearly the expansion of the correlations to all 74 compounds yields meaningful gains in R^2 and RMSD and MAD residuals in three out of four cases (all but 6-31G(d) with UA0 radii), improving predictive ability and giving increased confidence to predictions of reduction potentials outside our initial window of potentials of the first 35 calibrant molecules. However it is also clear that the correlations do move slightly further from a slope of unity and an *x*-intercept equal to the -4.12 V reference potential of SCE. The model is thus accommodating other systemic errors of the method in the slope and intercept terms of the correlation. Far from being disheartening, we find this encouraging that while the computed "electron affinity" in a dielectric continuum implicit solvent model is not itself a perfect analog for solution reduction potential, it correlates extremely well and has good predictive abilities over a wide range of potentials with systemic errors largely accounted for in the fitting parameters of slope and intercept.

Table 4. Correlations of computed D_0 - S_0 energy difference with experimental reduction potential for compounds **1-74** varying basis set and CPCM radii.

Corr. #	B3LYP/	CPCM radii	slope *m* (eV/V)	*y*-intercept *b* (eV)	R^2	RMSD[a] residuals (V)	MAD[b] residuals (V)	*x*-intercept[c] (V)
5	6-31G(d)	UA0	-1.1751	-4.2519	0.9887	0.0880	0.0662	-3.6184
6	6-31G(d)	UAKS	-1.2461	-4.2887	0.9840	0.1050	0.0825	-3.4416
7	6-311+G(d,p)	UA0	-1.1242	-4.6167	0.9852	0.1007	0.0796	-4.1068
8	6-311+G(d,p)	UAKS	-1.1930	-4.6903	0.9889	0.0829	0.0629	-3.9314

[a]Root mean square deviation and [b]mean average deviation, taken from individual residuals for each compound as predicted by each trendline (as reported in the Supporting Information). [c]Computed *x*-intercept (= *-b/m*) corresponds to the reference electrode potential correction (SCE = NHE + 0.24 V = -4.12 V), and any systematic inaccuracies of the computational model employed.

Figure 1 contains the plots from which correlations 1-8 in Tables 1 and 4 were constructed. Graphically it is easy to see the greater divergence that occurs for calibrations based on UAKS radii than those based on UA0 radii when the correlation is expanded from calibrants **1-35** to include all molecules **1-74**, and the greater range of potentials these molecules span.

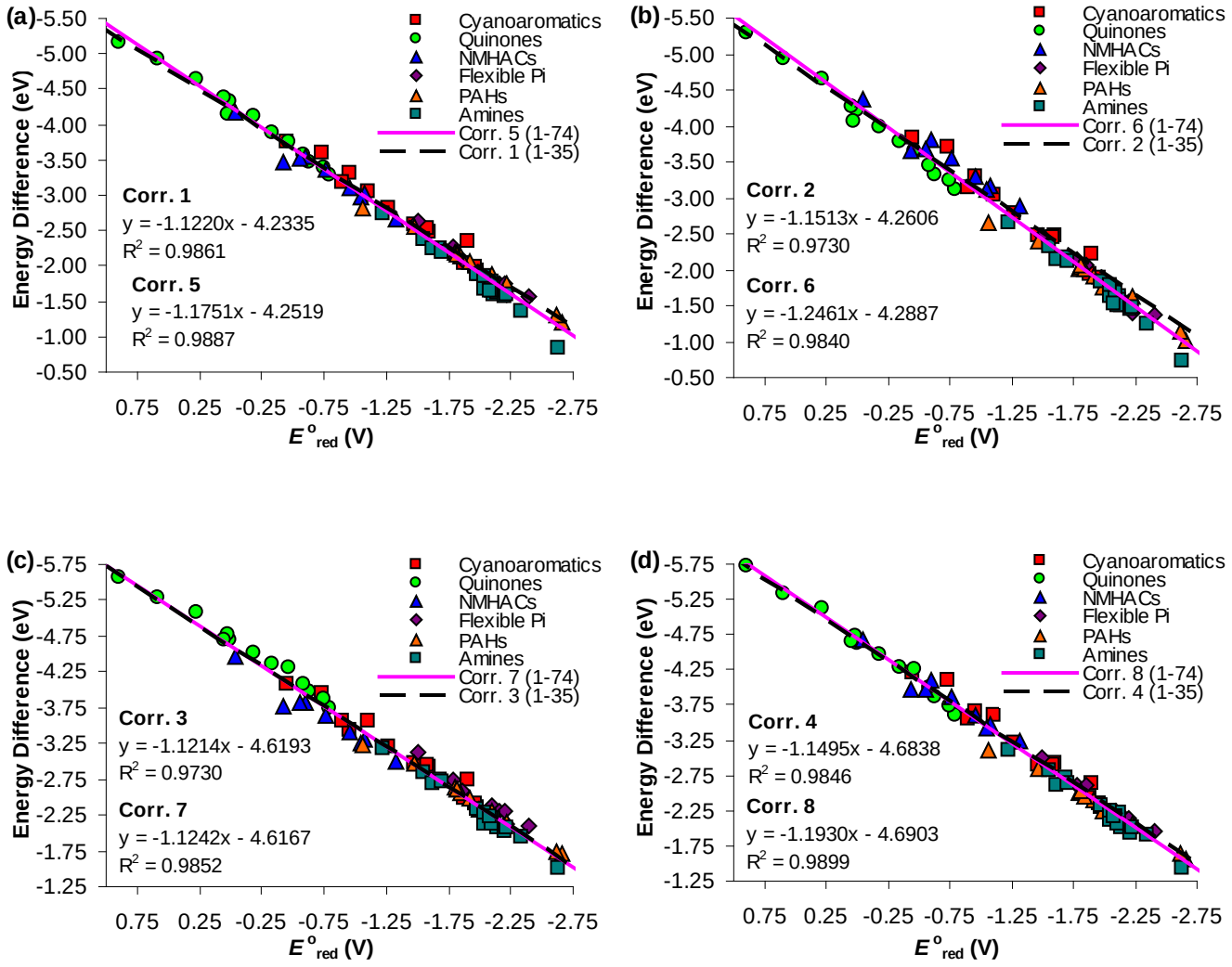
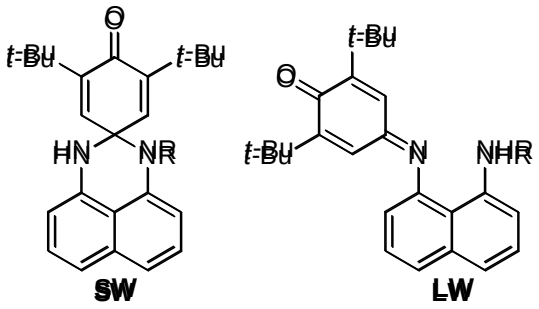


Figure 1. Plots of computed D0 – S0 energy differences vs. experimental reduction potential, varying basis set and CPCM radii. **(a)** 6-31G(d)/UA0; **(b)** 6-31G(d)/UAKS; **(c)** 6-311+G(d,p)/UA0; **(d)** 6-311+G(d,p)/UAKS. Dashed black lines represent correlations based on the initial calibrant set **1-35**; solid pink lines represent correlations based on all molecules **1-74**. Correlation #s match those used in the text and Tables 1 & 3.

PSHDs (**75** R=H; **76** R=Me)



We have applied our method to the prediction of the reduction potentials of several larger and more complicated photochromic molecules of interest to our research group, perimidinespirohexadienones (PSHDs, **75-76)**[25,26] prototypical among them. The data in Table 5 indicate that all eight correlations do a fairly good job predicting the reduction potentials of these four photochrome structures, with RMSD errors on these 10 compounds less than the RMSD errors of the original correlations. In some cases the predicted errors are smaller than the errors in the experimental data. In this small sample set the correlations (1-4) based on the initial 35 calibrants outperform those based on all 74 calibrant molecules (correlations 5-8), but this trend does not hold when we examine additional photochromes which we will report in the future.[27] These first four photochromes provide proof of principle that these calculations can be useful in guiding us to more reducible analogs of the PSHDs as synthetic targets; we have also found them vital to understanding the more complicated systems we are now investigating.

The preparation of **75-78** in their short wavelength form (**SW**) and characterization of their photogenerated long wavelength form (**pLW**) was recently reported by our group. The electrochemistry of these molecules and the proof of structure of the differential electrogenerated long wavelength isomers (**eLW**) of the quinazolinespirohexadienones (**77-78**) will be reported in a forthcoming manuscript. However we include the information on their experimental reduction potentials hereto demonstrate the excellent agreement to our computational predictions (Table 5), particularly for such complicated molecules outside the structural motifs of our calibrant molecules **1-74**. Moreover, it was our initial calculations of this sort that first helped us realize the differential photochromic and electrochromic ring opening of **77-78** and assign tentative structures to the **pLW** and **eLW** isomers of each.

Table 5. Comparison of experimental reduction potentials of photochromes **75-76** to those predicted using correlations 1-8.

Cmpd	Experimental	Predicted E^{o}_{red} (V vs. SCE), based on correlation #[b]							
	E_{red} (V vs SCE)[a]	Corr 1 ±0.084	Corr 2 ±0.117	Corr 3 ±0.117	Corr 4 ±0.088	Corr 5 ±0.088	Corr 6 ±0.105	Corr 7 ±0.101	Corr 8 ±0.083
75 SW	-1.744 ± 0.020	-1.836	-1.813	-1.835	-1.846	-1.759	-1.698	-1.830	-1.784
75 LW	-0.939 ± 0.028	-0.947	-0.946	-0.980	-1.012	-0.877	-0.897	-0.950	-0.980
76 SW	-1.689 ± 0.013	-1.683	-1.664	-1.706	-1.720	-1.607	-1.560	-1.695	-1.663
76 LW	-0.919 ± 0.028	-0.935	-0.935	-0.972	-1.004	-0.856	-0.886	-0.931	-0.973
RMSD[c]	*0.023*	*0.047*	*0.015*	*0.035*	*0.024*	*0.123*	*0.043*	*0.101*	*0.038*

[a]Reversible ground state reduction potential E^{o}_{red} for LW isomers and half-peak potential $E^{1/2}_{red}$ for irreversible reduction of SW isomers, in dry acetonitrile containing 0.1 M tetrabutylammonium hexafluorophosphate electrolyte, normalized to ferrocene/ferrocenium, and corrected to vs. SCE (Refs. 20 and 24); [b]± root mean square deviation error of the correlation (from Tables 1 & 2) indicated beneath each correlation #; [c]root mean square deviation of the experimental (from error bars) or predicted (relative to experimental) reduction potentials for these four photochromic structures.

As these photochromes are the largest molecules we have examined with this computational methodology, they prove the best test for the computational efficiency of our method. Gas-phase geometries of S_0 or D_0 calculated at B3LYP/6-31G(d) on a single 2.60 GHz AMD Opteron-252 processor (with 8 GB RAM and 250 GB HD) took 12-32 hours, while MIDI! geometries were typically slightly faster (8-30 hours). CPCM single point energies took 7-22 hours at 6-311+G(d,p) and only about one hour (40-100 minutes) at 6-31G(d), for either UA0 or UAKS radii. UAKS generally took about 10% longer. At 6-311+G(d,p), D_0 states generally took about two times longer (ca. 20 hrs) than the corresponding S_0 state of the same molecule (ca. 10 hrs); the difference was slightly less pronounced at 6-31G(d). All calculations on calibrants **1-74** were considerably faster due to their smaller size, generally similar to the times reported in our previous work.

The LW isomers of **75-76** and related photochromes also have a readily accessible second reduction potential which we have also determined experimentally. Attempts to use the computed energy difference between the doubly-reduced dianion (S_0 or T_0) and the singly-reduced D_0 anion radical and the with any of the 8 correlations we developed and calibrated for predicting first reduction potentials (from D_0 anion radical and S_0 neutral computed energy differences) unfortunately but unsurprisingly fail to yield an accurate prediction of these second reduction potentials. Presumably a new correlation of second reduction potentials could be developed, if a sufficient range of second reduction potentials could be obtained experimentally to calibrate a correlation, though we have no immediate plans to pursue this further.

Finally, we have made several different attempts to test and demonstrate the robustness of our model to the minor variations in exact methodology that can easily creep into computational protocols over time, particularly when training novice undergraduate, community college, and high school students in their execution. One such opportunity was found when we inadvertently used the *Gaussian* convergence keyword SCF=tight for all our single point energies regardless of basis set throughout this study, when this is only necessary for the 6-311+G(d,p) basis set with diffuse functions. We went back and repeated one entire correlation (see Supporting Information) of molecules **1-35** without this keyword for the 6-31G(d)/CPCM(UA0) correlation. The slopes, x- and y-intercepts, and R2 value for this correlation all changed by less than 0.1 %, while RMSD and MAD increased from by less than 1 %. The maximum change in any given prediction was 4.1 mV and the average change was 1.0 mV. Moreover, if the SCF=tight energy differences were used to predict reduction potentials in the correlation created without this criterion, the RMSD and MAD errors both remained unchanged to the fourth decimal place (at 0.0835 V and 0.0635 V, respectively.) Likewise, if the energy differences found without SCF=tight were used in the correlation created with the SCF=tight criterion (Corr. 1), the RMSD and MAD errors again remained unchanged at the fourth decimal place (at 0.0840 V and 0.0640 V, respectively.) The

maximal changes in any given predicted reduction potential in either of these uses of data across correlations was 1.5 mV and the average change was only 0.5 mV. Clearly the method is robust with respect to this casual mistake in convergence criteria in *Gaussian*.

Next, we found that at times we had been inconsistent in our application of symmetry (it was our default to disable symmetry) in our geometry optimizations and single point energies. We therefore reexamined several molecules of a range of sizes and symmetries with and without disabling symmetry for gas-phase geometry optimizations. Then we computed the single point energies (with CPCM solvent using both UA0 or UAKS radii with both basis sets of interest) on all these geometries, again both with and without disabling symmetry. We then determined the maximum impact any of the resulting differences made by symmetry could have on a predicted reduction potential. The molecules we studied most exhaustively were dicyanobenzene **2**, aceanthrylene **50**, and a new photochrome we will report in the future. For highly symmetric **2**, the largest computed energy difference in this experiment was 28 neV, corresponding to a difference in predicted reduction potential of less than 44 nV. For less symmetric **50**, the largest computed energy difference was 3.5 μeV for a difference in predicted reduction potential of less than 3.0 μV. Finally for a highly unsymmetric photochromic LW, the largest computed energy difference was 4.9 meV for a difference in predicted reduction potential of less than 3.0 mV. It is not clear that disabled symmetry is even responsible for these differences, as it is possible that slightly different minima were found for each geometry. Rather these results set an upper bound on the error introduced by the application of symmetry. These errors are at least one and potentially several orders of magnitude less than the error of our method, and therefore of no practical consequence.

Other variables that occurred in our calculations were the application of different basis sets to the geometry optimizations. Our group has at times used MIDI! and 6-31G(d) interchangeably for gas phase geometry optimizations. A similar analysis of the few molecules for which we had CPCM single

point energies computed at the same level of theory for both MIDI! and 6-31G(d) geometries showed a maximum computed energy difference of 7 meV and a maximum difference in predicted reduction potential of 2 mV, more than a factor of 40 less than our current error bars. Thus we believe it is safe to use our correlations and our methods on any set of good reasonable gas phase geometries regardless of specific basis set used. Similarly small differences were found regardless of whether *Gaussian03* or *QChem 3.0*[28] were used for the gas-phase geometry optimizations.

During the course of our work, all of which was conducted using *Gaussian03*, we did work briefly on another cluster running a different revision (B.05 vs. D.01) of the software. We found that the difference in any given energy was often nil, but occasionally as high as 1.5 meV. The maximum difference in predicted reduction potential was only 0.2 mV. We wanted to confirm that it was acceptable to move this project to *Gaussian09* in the future. Thus we also compared *Gaussian03* (revision D.01) to *Gaussian09* (revision A.02)[29] and found even smaller energy differences of at most 1.1 meV, though in this case the largest difference in predicted reduction potential was of 0.3 mV. We find it curious that different revisions or editions most often gave absolutely no difference in energy all the way to the final decimal place (nHartree = 27 neV) as should be expected, but sometimes varied by up to a millielectronVolt even for identical input files, with no obvious pattern, presumably due to finite convergence criteria. Nevertheless this level of consistency is within our error bars by well more than an order of magnitude, and thus our method and correlation should be robust to version and revision of Gaussian used, at least with respect to those specifically tested.

Finally, the failure of some of our photochromes (analogs of **75-76** with an OH capable of intramolecular hydrogen-bonding to an imine N) to converge when performing CPCM calculations with the default UA0 radii was a motivating factor in establishing the UAKS correlations (which despite also being a United Atom Topological Model like UA0, without explicit spheres for hydrogen,[11,29]

fortuitously converged our phototochromes.) Now that we have those data, we can compare the errors that arise when the UA0 correlation is applied to a UAKS energy difference, and vice versa. When the UA0 radii correlation is applied to the energy differences computed for compounds **1-74** using UAKS radii the RMSD error increases from 0.0880 V to 0.1333 V (a 51% increase) for 6-31G(d) and from 0.1007 V to 0.1035 V (a 3% increase) for 6-311+G(d,p). Similarly, when our new UAKS correlation is applied to predict reduction potentials from the energy differences computed for **1-74** using UA0 radii the RMSD error increases from 0.1050 to 0.1080 V (a 3% increase) for 6-31G(d) and from 0.0829 V to 0.1079 V for 6-311+G(d,p) (a 30% increase). While the MAD errors are (by definition) lower than the RMSD errors, the relative increases in MAD are greater than in RMSD in all cases. In all these misapplications of correlation based on CPCM radii the largest individual change in predicted reduction potential is 142 mV, while the average change is 54 mV. These are smaller than the corresponding largest residual and mean average deviation respectively for most of our methods. Thus, while certainly not optimal, if radii parameters need to be adjusted to allow a calculation to converge for a given molecule of interest, it is likely still possible to apply our correlations without establishing a full new correlation for the new radii parameters.

Conclusions

We have tested and improved our previous correlation and demonstrated its robustness and utility for computationally predicting the first ground state reduction potentials of a wide range of organic molecules spanning over 3.5 V of the potential window across several families of conjugated organic molecules with a variety of functional groups, including larger, more complicated and more flexible molecules than were used to form the calibration. Furthermore the method is robust and tolerant to a range of modest 'mistakes' and variations in computational methodology, and thus suitable for non-expert users. Good global correlations over a wide potential window are now available for CPCM

implemented with either the default UA0 or alternative UAKS radii parameters, with the B3LYP functional using either the 6-31G(d) or 6-311+G(d,p) basis sets.

We have demonstrated that either of these radii models for CPCM work well in our application, and to date we have been able to get all molecules of interest to our research group (included herein or not) to converge and give good results with at least one of these two radii models. UA0 uses the United Atom Topological Model applied on atomic radii of the UFF force field for heavy atoms, while hydrogens are enclosed in the sphere of the heavy atom to which they are bonded.[11,29] UAKS meanwhile uses the same United Atom Topological Model applied on atomic radii optimized for the PBE0 hybrid functional[30] (PBE1PBE in *Gaussian09*) and the 6-31G(d) basis set.[11,29] In neither case are explicit spheres for hydrogen atoms specified, nor have we done so on an individual basis. The convergence of some of our intramolecularly hydrogen-bonded photochrome species with UAKS radii that failed to converge with UA0 is likely fortuitous – others may find one or the other radii better suited to their needs. However explicit spheres for all or individual hydrogens are an option, and may be particularly worthwhile in molecules where a hydrogen may be too far from any heavy atom, or too close to more than one. As we move to *Gaussian09* in the future we will likely explore using the UFF radii, which is now the default for CPCM and which does include explicit hydrogens. A useful discussion of radii in CPCM calculations was recently published,[31] after this currently reported work was completed. This may prove helpful to us or others using similar methodology in the future.

Finally, we maintain that correlational or "relative" methods such as ours, while perhaps less elegant, maintain an advantage over "absolute" methods in that systematic deviations in the computational approach can be accounted for in slope and intercept terms. Nor is correction of experimental data (obtained relative to an electrode) to an absolute scale required. Thus we feel our methods are

complementary to others in the literature and provide a useful alternative to computationally novice and experienced users alike.

Experimental Section

Compounds 75 and 76 (in their SW form) were prepared as previously reported.[25,26] Reduction potentials of the SW and LW isomers of each were determined in at least triplicate by cyclic voltammetry with a glassy carbon working electrode, platinum wire counter electrode, and a non-aqueous $Ag/AgNO_3$ reference electrode, in argon-purged solutions of dry HPLC grade acetonitrile containing 0.1 M tetrabutylammonium hexafluorophosphate electrolyte and 1-3 mM analyte. Results were normalized to ferrocene/ferrocenium by back-to-back experiments, and then corrected to vs. SCE.[20,24] LW isomers could be obtained by photolyzing the SW solution under argon with either the 405 nm line of a 350 W mercury arc lamp, or electrogenerated by repeated CV scans of the SW solution. The reported potentials are the reversible ground state reduction potentials (E^{o}_{red}) for LW isomers and the half-peak potentials ($E^{½}_{red}$) for irreversible reduction of SW isomers,

Acknowledgement. This work was funded by an NSF CAREER award (CHE-0952768) and a Camille & Henry Dreyfus Foundation Start-up Award (SU-04-040). EL (home institution College of the Canyons, Santa Clarita, CA) was supported by an NSF REU SITE award (CHE-0851194) to Hope College. High school students BC and CL (Holland High School, Holland, MI) were supported by the Research Experiences Across Cultures at Hope (REACH) program funded by a Howard Hughes Medical Institute grant to Hope College. Computations were conducted on the Midwest Undergraduate Computational Chemistry Consortium (MU3C) cluster, supported by NSF MRI grants CHE-0520704 and CHE-1039925, housed in the Hope College Computational Science & Modeling Laboratory. Support from CSM Laboratory director Prof. Brent P. Krueger and staff Mr. Paul Van Allsburg is also gratefully acknowledged. Several helpful reviewer comments are also gratefully acknowledged.

Supporting Information. Three Microsoft Excel workbooks (the first for correlations based on **1-35** alone, the second for correlations based on **36-74** alone, and the third for correlations based on all calibrant molecules **1-74**), each with separate tabbed worksheets for each basis set and radii combination, provide additional graphs and correlations by family of molecules as well as all correlations reported herein including misapplication of data to the "wrong" correlation. This includes complete data of the individual computed S_0 and D_0 energies for each molecule at each level of theory, along with the resulting energy differences and the literature reduction potentials to which they were correlated. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

[1] Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*, 2nd ed.; John Wiley & Sons, Ltd.: Chichester, West Sussex, England, 2004.

[2] Winget, P.; Weber, E. J.; Cramer, C. J.; Truhlar, D. G. *Phys. Chem. Chem. Phys.* **2000**, *2*, 1231.

[3] Baik, M.-H.; Friesner, R. A. *J. Phys. Chem. A* **2002**, *106*, 7407.

[4] Schmidt am Busch, M.; Knapp, E.-W. *J. Am. Chem. Soc.* **2005**, *127*, 15730.

[5] Speelman, A. L.; Gillmore, J. G. *J. Phys. Chem. A* **2008**, *112*, 5684.

[6] Paoprasert, P.; Laaser, J. E.; Xiong, W.; Franking, R. A.; Hamers, R. J.; Zanni, M.T.; Schmidt, J.R.; Gopalan, P. *J. Phys. Chem. C* **2010,** *114,* 9898.

[7] Tugsuz, T. *J. Phys. Chem. B* **2010,** *114*, 17092.

[8] Davis, A. P.; Fry, A. J. *J. Phys. Chem. A* **2010,** *114,* 12299.

[9] Häussler, M.; King, S. P.; Eng, M. P.; Haque, S.A.; Bilic, A.; Watkins, S. E.; Wilson, G. J.; Chen, M.; Scully, A. D. *J. Photochem. Photobiol. A* **2011**, *220*, 102.

[10] Wang, L.-P.; Van Voorhis, T. *J. Chem. Theory Comput.* **2012**, *8*, 610.

[11] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, Revision D.01, Gaussian, Inc.: Wallingford, CT, 2004.

[12] Schmidt, J.R.; Polik, W.F. *WebMO Pro*, v. 9.1; *WebMO Enterprise*, v. 10.1; WebMO LLC: Holland, MI, USA, 2009, 2010; available from http://www.webmo.net (last accessed April 2012).

[13] Becke, A. D. *J. Chem. Phys.* **1996**, *104*, 1040.

[14] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.

[15] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.

[16] Easton, R. E.; Giesen, D. J.; Welch, A.; Cramer, C. J.; Truhlar, D. G. *Theor. Chim. Acta* **1996**, *93*, 281.

[17] Li, J.; Cramer, C. J.; Truhlar, D. G. *Theor. Chem. Acc.* **1998**, *99*, 192.

[18] Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998,** *102*, 1995.

[19] Cossi, M.;Rega, N.;Scalmani, G.; Barone, V. *J. Comp. Chem.*, **2003**, 24, 669.

[20] Bard, A. J.; Faulkner, L. R. *Electrochemical Methods: Fundamentals and Applications*, 2nd ed.; John Wiley & Sons: New York, 2000.

[21] Fukuzimi, S.; Koumitsu, S.; Hironaka, K.; Tanaka, T. *J. Am. Chem. Soc.* **1987,** *109*, 305.

[22] Koper, C.; Sarobe, M.; Jenneskens, L. W. *Phys. Chem. Chem. Phys.* **2004,** *6*, 319.

[23] Millefiori, S. *J. Heterocycl. Chem.* **1970**, *7*, 145.

[24] Pavlishchuk, V. V.; Addison, A. W. *Inorg. Chim. Acta* **2000**, *298*, 97.

[25] Minkin, V. I.; Komissarov, V. N.; Kharlanov, V. A. Perimidinespirocyclohexadienones, in: J. C. Crano and R. J. Guglielmetti (Eds.) *Organic Photochromic and Thermochromic Compounds, Vol. 1*, Plenum Press: New York, 1999, pp. 315-340.

[26] Moerdyk, J. P.; Speelman, A. L.; Kuper, K. E.; Heiberger, B. R.; TerLouw, R. P.; Zeller, D. J.: Radler, A. J.; Gillmore, J. G. *J. Photochem. Photobiol. A* **2009**, *205*, 84.

[27] Sluiter, K. B.; Moerdyk, J. P.; Speelman, A. L.; Lynch, E. J.; Gillmore, J. G. Manuscript in preparation.

[28] Shao, Y.; Fusti-Molnar, L.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; Distasio, R. A., Jr.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.,; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S.R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khalliulin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y.M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock III, H. L.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer III, H. F.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys*. **2006**, *8*, 3172.

[29] Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, Jr., J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, N. J.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, Revision A.02, Gaussian, Inc., Wallingford CT, 2009.

[30] Adamo, C.; Barone, V. *J. Chem. Phys.* **1999,** *110*, 6158.

[31] Tao, J.-Y.; Mu, W.-H.; Chass, G. A.; Tang, T.-H.; Fang, D.-C. *Int. J. Quantum Chem.* **2012**, DOI: 10.1002/qua.24065