Excited state charge maps and properties of molecules. Full CIS CNDOL benchmark for fast and reliable calculations

Luis A. Montero-Cabrera[†*], Ana L. Montero-Alejo[†], Edward Pyzer-Knapp[#], Alán Aspuru-Guzik[#], José M. García de la Vega[&], Mario Piris[+], Lourdes A. Díaz-Fernández[†], María E. Fuentes[‡], Carlos M. de Armas[†]

[†]Laboratorio de Química Computacional y Teórica, Departamento de Química Física, Universidad de La Habana, 10400 Havana, Cuba; [#]Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138, USA; [&]Departamento de Química Física Aplicada, Universidad Autónoma de Madrid, 28049 Madrid, Spain; [+]Kimika Fakultatea, Euskal Herriko Unibertsitatea (UPV/EHU), and Donostia International Physics Center (DIPC), 20018 Donostia, Spain and IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Spain, and [‡]Facultad de Ciencias Químicas, Universidad Autónoma de Chihuahua, Chihuahua 31100, Chihuahua, México.

ABSTRACT:

Complete Neglect of Differential Overlap taking into account L azimuthal quantum number of the atomic basis set (CNDOL) is both an *a priori* and approximate quantum mechanical Fockian which simplifies the raw Roothaan's HF formulation for the simulation of molecules. All multielectronic determinants corresponding to single excitations of the resulting SCF ground state monoelectronic wave function are variationally optimized after applying large-scale full singly-excited configuration interaction (FCIS) for modeling electron excitation properties. In this study the accuracy of CNDOL predicted excited state energies are benchmarked with respect to experimental data and compared to state of the art TD-DFT calculations. Excitons are expressed in terms of wave functions that comprise all valence electrons of the system for each different energy state giving multielectronic 3D density maps in place of the one electron nature of molecular orbitals. The binding energy of each excitation is evaluated by the Coulomb – exchange term in order to model exciton stability and localization. Graphics for mapping total charge distribution on the molecular geometry of ground and selected excited states, as well as the charge transfers occurring upon each transition are shown to be tools for a qualitative modeling and a better understanding of molecular excitation processes. It is shown that CNDOL can be used for building molecular wave functions comprising all valence electrons of relatively big polyatomic systems due to the superior speed of these calculations with respect to more sophisticated methods such as TD-DFT.

INTRODUCTION

Simulations of the light-matter interaction of molecular and complex polyatomic systems behavior toward light is a research question with fundamental and pragmatic implications. Theoretical investigations afford a platform for detailed investigation of these effects, and allow us to extend our studies to systems where the relevant experiment either has not or cannot be performed. Knowledge gained through these simulations can find practical applications, such as designing new photovoltaic materials. Simulations using quantum mechanical (QM) models provide independent information attending to the complexities of the systems at hand,[1] although they are often very computationally intensive.

Such QM models can be derived through many different approaches. Nowadays, the most frequently published studies in this area are those obtained with TD-DFT approach[2]. Kohn – Sham's orbitals are frequently used as molecular orbitals leading to interpretations for understanding electron state changes in the same way as those that were obtained by working with simple linear combination of atomic orbital (LCAO)[3].

The more "traditional" approach consists of calculating energies of different monoelectronic states of the system, i.e. the eigenvalues of each molecular orbital. In this case, the expected wavelength of an absorbed photon is deduced from the energy difference between such involved one-electron state eigenvalues, although including the energy of electronic interactions, as given by two-electron integrals of their corresponding molecular wave functions. It has been described for well-known Hartree – Fock´s (HF) Self Consistent Methods (SCF) methods[4-8]. A configuration interaction calculation performed on singly excited determinants (CIS), as obtained from the appropriate solution of a ground state SCF result, is a key in reaching the accuracy of multielectronic wave functions and it is still the object of active research[9-14].

Benchmarking of new, and even older methods which have found new purposes, is a common procedure[15-27]. This is a frequent and useful procedure in the case of DFT based methods for exploring the performance of new empirically or non-empirically implemented or corrected functionals to a given set of molecules. An "in depth" analysis of both the advantages and

drawbacks of employing TD-DFT methods to describe the excited states of relatively large systems has recently been published[28].

Some authors of this paper have developed an approach for obtaining a simplified Fockian called as Complete Neglect of Differential Overlap taking into account L azimuthal quantum number of the atomic basis set (CNDOL). The first paper checked the quality of limited CIS calculations for spectra of several hydrocarbons[29]. Reliability for excited state energy calculations with a hybrid CNDOL Fockian (as shown below) were further tested for a set of molecules based upon significant peaks of far UV gas phase experimental spectra[30]. Here, we show comprehensive benchmarking results of the original[29] approximate Fockian when the state determinants are optimized by a full CIS of valence state molecular orbitals. This benchmarking compares the success of CNDOL for reproducing well described experimental absorption spectra with high quality TD-DFT calculations. The molecules selected for testing were those inspired on the Fabian – Diaz – Seifert - Niehaus collection of compounds, with small variations[31]. For these molecules all apparent maxima and shoulders seen in the experimental data were considered, and available gas phase spectra are also added for comparisons. Comparisons on computing time are given, to highlight the potential utility of this method for high throughput virtual screening. Novel graphics of charge distribution maps corresponding to all valence electrons in several states are also shown. These graphics improve the toolbox for qualitative modeling and provides a better understanding of molecular excitation processes.

METHOD

The CNDOL approximate SCF-MO procedure is an improvement of the former PPP and CNDO/S[29] methods for modeling molecular spectroscopic phenomena. This kind of improvement of CNDO[32-34] was made by several other procedures[35-40] as derived from the INDO[41] and NDDO[34] Pople's original Fockians. Typically these were always attained by augmenting the number of terms and thus adding empirical parameters depending on *a posteriori* fittings. Thus, one of our enhancements resides in simplicity for the construction of Fock matrix's elements and leaving a few parameters to be evaluated by experimental values

from literature and integral formulas. Those parameters are thus *a priori* evaluated, and correspond to Slater exponents (ζ), ionization potentials (I_μ) and electron affinities (A_μ) of each basis atomic orbital. Together with the system's geometry, these parameters represent the only input data for each calculation.

Consequently, CNDOL follows the philosophy of preserving theoretical coherence for a better understanding and prediction of physical phenomena related to electronic excitations, even in an approximate framework. We have previously shown that this method provides electronic excitation energies and transition probabilities which are in fair accordance with experimental data for hydrocarbons[29] and a few peaks of selected molecules in gas phase[30]. Our procedure also gives theoretically consistent results for relatively large polyatomic systems[30, 42-44]. However, as it is desirable to use the method for predictions of novel compounds with many different elements[45], the benchmarking procedure should be extended to demonstrate utility in over a wide variety of molecules.

The CNDOL formalism can be established from the transformed Fock equations after the CNDO approximation [33, 34]. The corresponding matrix elements for CNDOL are:

$$F_{\mu\mu}^{CNDOL} = H_{\mu\mu} - \tfrac{1}{2} p_{\mu\mu}\gamma_{AA}^{ll} + \sum_{B}(P_{BB}^{l}\gamma_{AB}^{ll} + P_{BB}^{k}\gamma_{AB}^{lk}) \qquad (1)$$

where $\mu \in A$, and

$$F_{\mu\nu}^{CNDOL} = H_{\mu\nu} - \tfrac{1}{2} P_{\mu\nu}\gamma_{AB}^{lk} \qquad (2)$$

where $\mu \in A$ and $\nu \in B$

Here "*ll*" can be understood as "*ss*" or "*pp*" type orbitals and "*lk*" as "*sp*" or "*ps*"; $p_{\mu\mu}$ represents the density matrix element of the atomic orbital μ at center A; P^{l}_{AA} corresponds to the sum of all those $p_{\mu\mu}$ which have the same azimuthal quantum number "*l*" at center A and they act upon (or modify) the corresponding two electron integral values (γ^{ll} and γ^{lk}).

Admixing different terms of CNDO/1[33] and CNDO/2[32] appear in tables 1 and 2, respectively, to hold CNDOL considerations for building up to eight Fockians.

Table 1. CNDOL/1 Fockian denomination as determined by diagonal $H_{\mu\mu}$ one - electron matrix elements[a]

$$H_{\mu\mu} = U_{\mu\mu} - \sum_{B \neq A} V_{AB}^{l}$$

$U_{\mu\mu}$	$\sum_{B \neq A} V_{AB}^{l}$	
	$-\sum_{B \neq A}\left(Z_B^l \gamma_{AB}^{ll} + Z_B^k \gamma_{AB}^{lk}\right)$ Split core charge: **S**	$-\sum_{B \neq A} Z_B \Gamma_{AB}^{l}$ Complete core charge: **C**
$-I_\mu + \gamma_{AA}^{ll} - Z_A^l \gamma_{AA}^{ll} - Z_A^k \gamma_{AA}^{lk}$ Split core charge: **S**	**CNDOL/1SS**	**CNDOL/1SC**
$-I_\mu + \gamma_{AA}^{ll} - Z_A \gamma_{AA}^{ll}$ Complete core charge: **C**	**CNDOL/1CS**	**CNDOL/1CC**

a. I_μ and A_μ are the valence state ionization potential and electron affinity, respectively, of the μ AO's; Z is the effective nuclear charge of the atomic core; Γ_{AB}^{l} is the interaction term between an electron with azimuthal quantum number l on atom A with the core of atom B.

Table 2. CNDOL/2 Fockian denomination as determined by diagonal $H_{\mu\mu}$ one - electron matrix elements[a]

$$H_{\mu\mu} = U_{\mu\mu} - \sum_{B \neq A} V_{AB}^{l}$$

$U_{\mu\mu}$	$\sum_{B \neq A} V_{AB}^{l}$	
	$-\sum_{B \neq A}\left(Z_B^l \gamma_{AB}^{ll} + Z_B^k \gamma_{AB}^{lk}\right)$ Split core charge: **S**	$-\sum_{B \neq A} Z_B \Gamma_{AB}^{l}$ Complete core charge: **C**
$-\frac{1}{2}\left(I_\mu + A_\mu\right) + \frac{1}{2}\gamma_{AA}^{ll} - Z_A^l \gamma_{AA}^{ll} - Z_A^k \gamma_{AA}^{lk}$ Split core charge: **S**	**CNDOL/2SS**	**CNDOL/2SC**
$-\frac{1}{2}\left(I_\mu + A_\mu\right) + \frac{1}{2}\gamma_{AA}^{ll} - Z_A \gamma_{AA}^{ll}$ Complete core charge: **C**	**CNDOL/2CS**	**CNDOL/2CC**

a. I_μ and A_μ are the valence state ionization potential and electron affinity, respectively, of the μ AO's; Z is the effective nuclear charge of the atomic core; Γ_{AB}^{l} is the interaction term between an electron with azimuthal quantum number l on atom A with the core of atom B.

The CNDOL/2CC has been described in the first paper on this method[29] and is derived by modifying the original Pople's procedures to include a separation of basis orbitals according their azimuthal quantum number. A previous paper[30] successfully modeled the wave function and excitation energies of the human retinal binding pocket by using the CNDOL/1CS (previously named as CNDOL/21) Fockian. It was noted that CNDOL/2CC never converged for this system using the original technique. We will show how the use of a novel adiabatic

approach, which allows greater stability in the convergence of the SCF, was used to successfully perform calculations later in this study.

The two electron integrals on a single center A are evaluated by means of the Pariser's[46] relation :

$$\gamma_{AA}^{ll} = I_{\mu} - A_{\mu} \tag{3}$$

A general formula has been used for two center – two electron integrals comprising empirical Mataga – Nishimoto's[47, 48] and Ohno's[49]:

$$\gamma_{AB}^{lk} = \left(a_{AB}^{lk\ 2} + cR_{AB} a_{AB}^{lk} + R_{AB}^{\ 2} \right)^{-1/2} \tag{4}$$

where $a_{AB}^{lk} = 2(\gamma_{AA}^{ll} + \gamma_{BB}^{kk})^{-1}$ and R_{AB} is the distance between centers, considered as points in space. It becomes the Mataga – Nishimoto's when $c = 2$, Ohno's when $c = 0$, and the so-called "modified Ohno's" approach used in this method when $c = 1$.

Core-electron potential terms for one center are evaluated by:

$$\Gamma_{\mu A} = \gamma_{AA}^{ll} = I_{\mu} - A_{\mu} \tag{5}$$

When the electron in the μ orbital on atom A interacts with the core of atom B, the two center core-electron potentials are calculated as:

$$\Gamma_{\mu B} = \left(\gamma_{AA}^{ll\ -2} + cR_{AB} \gamma_{AA}^{ll\ -1} + R_{AB}^{\ 2} \right)^{-1/2} \tag{6}$$

Non-diagonal elements of the monoelectronic matrix $H_{\mu\nu}$ ("resonance" integral terms) are constructed by multiplying the overlap integral by a factor representing the projection onto the bond axis[30, 50].

$$H_{\mu\nu} = -\tfrac{1}{2}(I_{\mu} + I_{\nu})\, f(r_{\mu\nu}, I_{\mu}, I_{\nu}) \text{, where } f(r_{\mu\nu}, I_{\mu}, I_{\nu}) = S_{\mu\nu} f'(x_{\mu}, y_{\mu}, z_{\mu}, x_{\nu}, y_{\nu}, z_{\nu}) \tag{7}$$

It must be remarked that only I_{μ}, A_{μ} and ζ Slater exponents for calculating $S_{\mu\nu}$ are parameters, taken from data in the literature[51, 52].

Managing large scale CIS and Full CIS

CIS routines are consistent with QM procedures for improving excited wave functions. They are thus taken as multielectronic states for any kind of Fockian, in the cases of either approximate or *ab initio* approaches. It is a very well-known variational procedure[53], although still requiring some development [10, 20, 22, 24], and represents the construction of an improved wave function of

the whole system as a linear combination of reference state functions. It is therefore clear that the most improvement is reached when the basis consists of as many as possible relevant reference states. This was evidenced in several previous works[44, 54-56] in which big systems provide a large basis and harbor very good results. Hence, it follows that the extension of our "active space", i.e. the selected reference wave functions of singly excited configuration determinants, should improve the quality and consistency of results for properties related to molecular excitations.

A CIS total wave function of a polyatomic system for a given ξ state is expressed as a linear combination of the selected Φ_σ basis single excited determinants:

$$\Psi_\xi^{CIS} = \sum_\sigma a_{\lambda\sigma}\Phi_\sigma \tag{8}$$

Then, $\mathbf{E}$ (a diagonal matrix of optimized molecular excitation energies) is routinely obtained by solving:

$$(\mathbf{H}-\mathbf{E})\mathbf{a} = 0 \tag{9}$$

where $\mathbf{a}$ is the matrix of CIS coefficients giving the participation of each Φ_σ singly excited configuration in the optimized states. The elements of $\mathbf{H}$ matrix are calculated from all the selected HF single excitation energies for any σ electron transition between $i \rightarrow k$ or $j \rightarrow l$ molecular orbitals, where k and l are empty for the ground state. Diagonal and off-diagonal of CIS matrix terms are:

$$\begin{aligned}
{}^1H_{\sigma\sigma} &= \varepsilon_k^{HF} - \varepsilon_i^{HF} + 2\langle ik|ki\rangle - \langle ik|ik\rangle \\
{}^1H_{\sigma\sigma'} &= 2\langle jk|li\rangle - \langle jk|il\rangle \\
{}^3H_{\sigma\sigma} &= \varepsilon_k^{HF} - \varepsilon_i^{HF} - \langle ik|ik\rangle \\
{}^3H_{\sigma\sigma'} &= -\langle jk|il\rangle
\end{aligned} \tag{10}$$

unless $i = j$ and $k = l$ [57]. <ik|ik> means two electron repulsion integrals regarding i and k MO's, and ε's are MO's eigenvalues.

After finding CNDOL MO's, all possible HF singly excited transition energies $H_{\sigma\sigma}$ of a given molecular system are evaluated and ordered according to their energy values. This constitutes our ($M = N_{filled\ orb.} \times N_{empty\ orb.}$)-dimensional "full active space"[54]. A CIS matrix of order M' is then

built with a selected number of the lowest energy terms, in order to size the active space according to the problem at hand and/or computational requirements. Then, $M' = M$ means "full CIS" (FCIS). Some authors of this paper made a previous[54] careful and critical discussion on the usefulness of FCIS for a case of fullerene cluster calculations and other similar large systems. It was concluded that whilst FCIS is not necessarily the best option in all cases, the small molecules selected for the present benchmarking made FCIS as a reasonable choice since the reduced single ζ SCF basis set intrinsic to these molecules means that increasing the number of reference configurations should improve the quality of the results. A recent reference benchmark paper for semiempirical methods[26] declared the elimination of certain "spurious" σ orbitals in the case of small molecules when a kind of full CIS was attempted in order to achieve good results. We avoided such procedure because we were unable to determine what orbitals were or were not "spurious".In practical terms, the computer program NDOL[58] selects the order of the active space to be either *n* (the number of basis functions or valence atomic orbitals used for calculating the HF procedure) or an integer multiple of this value up to the limiting FCIS value of M. Eventual truncations for selecting the active space are made at the higher energy SCF excited determinant eigenvalues, being meaningless what occupied and unoccupied MO's were involved, i.e. if *n* is chosen as the order of CIS matrix, the selected determinants are a subset comprising those *n* of lowest energy from the set of all possible excitations.

Electron density terms of each σ singly excited configuration can be obtained from SCF-MO coefficient matrices:

$$p_{\mu\nu,\sigma} = \sum_{i=1}^{N} n_{i,\sigma} c_{i\mu} c_{i\nu} \qquad (11)$$

where $n_{i,\sigma}$ is the occupation number of each i MO in the σ SCF configuration (0, 1 or 2 according to the number of electrons in each MO corresponding to any given σ determinant) and $c_{i\mu}$'s are the SCF-MO coefficients.

The electron density elements of each ξ CIS state are given by the square of the corresponding coefficients and densities of SCF basis configurations:

$$p_{\mu\nu,\xi} = \sum_{\sigma=1}^{M'} a_{\sigma\xi}^2 p_{\mu\nu}^{\sigma} \qquad (12)$$

This value can be described in atomic values by summing electron densities of all basis atomic orbitals on the same center A:

$$P_{A,\xi} = \sum_{\mu \in A} p_{\mu\mu,\xi} \tag{13}$$

By working with an orthogonal basis set, the trace of $\mathbf{p}$ is equal to the total number of electrons. Charges on each center are calculated by summing electrons which remain after subtracting the core positive charge:

$$Q_{A,\xi} = Z_A^{core} - P_{A,\xi} \tag{14}$$

This value is useful for representing the map of charge distribution of any polyatomic system corresponding to each electronic state ξ = 0, 1, 2, …, where ξ = 0 is the ground state. It is usual to represent them as S_0, S_1, S_2,… for singlets and T_1, T_2, … for triplets. Charge maps of such CIS multielectronic (i.e. comprising all valence electrons) wave functions can be illustrative on what happens to a polyatomic system when it becomes excited by the absorption of a photon.

An exciton is defined as a "charge – hole" pair quasi-particle in the framework of one electron wave function representation whose stability is related to the "binding" energy between the corresponding excited and ground states. Molecular orbital approaches therefore provide a simple view of excitons - an i one electron molecular function leaves a "hole" when excitation transfers its negative charge to the k virtual orbital. When describing many electron wave functions, as provided by ground state and excited determinants, exciton binding can be related to electron interaction terms that intervene in eq. (10) for transition energy calculations:[55]

$$^1E_{CE} = 2\langle ik|ki\rangle - \langle ik|ik\rangle \tag{15}$$
$$^3E_{CE} = -\langle ik|ik\rangle$$

These are the *Coulomb – exchange* energy terms for each HF excitation. In the case of a single determinant excitation, the value of <$ik|ik$> Coulombic (J's) and <$ik|ki$> exchange (K's) integrals can be considered as related to the expected binding between electrons and their left empty states (holes). Then, Coulomb-Exchange terms for each ξ CIS many electron excited state are modified by the $\mathbf{a}$ CIS transformation matrix:

$$\left[{}^{1}E_{CE}\right]_{\xi} = \sum_{\sigma=1}^{M'} a_{\sigma\xi}^{2}\left(-\langle ik|ik\rangle + 2\langle ik|ki\rangle\right)_{\sigma} \qquad (16)$$

$$\left[{}^{3}E_{CE}\right]_{\xi} = \sum_{\sigma=1}^{M'} a_{\sigma\xi}^{2}\left(-\langle ik|ik\rangle\right)_{\sigma}$$

RESULTS AND DISCUSSION

Nowadays, benchmarks are frequently published as comparisons between sets of theoretical results[18, 20, 21, 26, 59] on the grounds that the relevant experimental data are affected by molecular environmental fields and roto-vibrational structures of excited states. However, we feel that the statistical significance of comparisons toward series of experimental data validates the reliability of our theoretical model. Therefore, a set of relevant testing molecules were selected from a previous benchmarking paper[31], plus adding furfuraldehyde as other type of molecules with well determined experimental properties and interesting structures. Their formulas are illustrated in Fig. 1.

All molecular geometries were optimized at the MP2/4-31G(d,p) level of theory using the Gaussian 09[60] suite. Such geometries were taken as the ground state to calculate UV spectra using both TD-PBE0[61, 62] [6-31G(d,p)] by Gaussian 09[60] and FCIS with the CNDOL/2CC Fockian by the NDOL2014[58] computer program. In this case, valence state ionization potentials and electron affinities of all element atomic orbitals were taken from a classical Jaffé – Hinze paper[51] and Slater exponents for overlap integral calculations were those of Clementi and Raimondi[52].

CNDOL/2CC Fockian calculations have been observed to display SCF non-convergence problems for some molecules. This was overcome by the use of an adiabatic procedure where the first SCF iteration runs a density matrix obtained after diagonalizing the one electron $\mathbf{H}$ Hamiltonian and then proceeding to repeat diagonalizations of $\mathbf{F}$ with a progressive increment of $\mathbf{G}$ (the electron interaction term) up to 30 steps according to:

$\mathbf{F}(n) = \mathbf{H} + [\ln(n)/\ 3.401197]\ \mathbf{G}$

for $n = 1,30$.

Table 3. CNDOL/2CC calculated electronic excitations compared with TD-PBE0[6-31G(d,p)] theoretical values and experimental bands in inert solvents and in gas phase.[a]

	CNDOL/2CC-full CIS excitations[b]				TD-DFT excitations[c]			Exp. UV/Vis maxima in inert solvents[d]		Exp. UV/Vis máxima in gas phase[e]	
Compound number and name	E [*f*]	λ	E_{CE}	t_{CPU}	E [*f*]	λ	t_{CPU}	E	λ	E	λ
1. Acrolein	***3.93 [0.00]***	***315***	***6.25***	***≤1***	***3.57 [0.00]***	***347***	***246***	***3.69***	***336 w***	***3.77[f]***	***329 w***
	5.59 [0.00]	***222***	***6.24***								
	5.97 [0.10]	***208***	***4.10***		***6.50 [0.39]***	***191***		***6.14***	***202 s***	***6.42[f]***	***193 s***
2. nitromethane	***2.32 [0.00]***	***534***	***8.44***	***≤1***			***160***				
	3.40 [0.00]	***365***	***8.38***					***3.6***	***340 sh***		
	4.73 [0.04]	***262***	***4.89***		***4.46 [0.00]***	***278***		***4.44***	***279 w***	***4.59[g]***	***270***
	6.43 [0.01]	***193***	***6.11***		***7.10 [0.15]***	***175***		***6.13***	***202 m***	***6.26[g]***	***198***
3. t-azomethane	***1.35 [0.00]***	***920***	***8.82***	***8***			***270***				
	3.86 [0.00]	***321***	***7.46***		***3.12 [0.00]***	***398***		***3.5***	***354 w***	***3.67[h]***	***338***
	4.78 [0.11]	***259***	***4.45***		***7.13 [0.00]***	***174***		***5.29***	***234 w***	***5.17[h]***	***240***
4. N, N, N', N' – tetramethyl-1,7-diamino-2,4,6-heptatriene	***3.40 [0.52]***	***365***	***3.15***	***92***	***3.23 [1.85]***	***384***	***4038***	***3.02***	***410 s***		

cation													
	3.72; 3.93; 4.77 [0.00]	334; 315; 261	4.68; 4.70; 4.35										
	5.10 [0.004]	243	3.16			4.64 [0.01]	267			4.84	256 m		
	5.37; 5.52 [0.00]	231; 225	4.23; 4.31										
	6.06 [0.005]	205	3.12			5.51 [0.01]	225			5.62	221 m		
5. diphenyl-methane cation	2.54 [0.00]	488	4.41	54				3357					
	3.54 [0.12]	351	3.54			3.02 [0.02]	411			2.80	442 s		
	4.08 [0.005]	304	3.62			3.07 [0.00]	404			4.08	304 m		
	4.10 [0.07]	302	3.86			3.40 [0.70]	364			4.20	295 m		
	4.31 [0.003]	288	3.53										
	4.40 [0.01]	282	3.75			4.63 [0.02]	268			4.65	267 m		
6. t-stilbene	3.96 [0.29]	313	3.22	71		4.09 [0.99]	303	3635		4.21[i]	315 s	4.00[j]	310
						4.75 [0.01]	261						
	4.08; 4.20; 4.93; 5.01; 5.09 [0.00]					4.75 [0.00]	261						
	5.13 [0.01]	242	3.13			5.35 [0.00]	232			5.36[k]	231 m		
	5.15; 5.15; 5.24; 5.29; 5.41 [0.00]					5.42 [0.00]	229						
	5.44 [0.13]	228	3.42			5.43 [0.19]	228			5.43[i]	224 m		
	5.38; 5.62 [0.00]												
	5.66 [0.01]	219	3.15										
	5.70; 5.71; 5.74; 5.75;					5.92 [0.00]	209						

	5.76 [0.00]												
	5.92 [0.57]	*209*	*3.46*										
	6.04 [0.57]	*205*	*3.32*			*6.56 [0.12]*	*189*			*6.15*[j]	*195 s*			
	6.10; 6.21; 6.27; 6.41; 6.48; 6.56; 6.60; 6.64; 6.70; 6.73; 6.80; 6.92; 6.94; 6.95 [0.00]	*203; 200; 198; 194; 191; 189; 188; 187; 185; 184; 183; 179; 179; 178*	*3.99; 3.80; 3.91; 3.91; 3.82; 3.81; 3.79; 3.05; 3.85; 3.87; 3.79; 3.37; 3.89; 3.45*			*6.56 [0.00]*	*189*						
	6.96 [0.09]	*178*	*3.49*			*6.60 [0.05]*	*188*			*6.94*[k]	*179 sh*		
	7.03; 7.05; 7.14; 7.18; 7.19; 7.20; 7.24; 7.25; 7.25; 7.29 [0.00]	*177; 176; 174; 173; 173; 172; 171; 171; 171; 170*	*3.92; 3.78; 3.50; 3.84; 3.40; 3.40; 3.81; 3.81; 3.40; 3.90*			*6.68 [0.00]*	*185*						
						6.75 [0.15]	*184*						
	7.35 [0.19]	*169*	*3.47*			*6.76 [0.17]*	*183*			*7.31*[k]	*169 vs*		
7. t-azobenzene	*1.04; 2.39 [0.00]*	*1190; 518*	*5.36; 4.59*	*71*				*2730*					
	2.68 [0.34]	*463*	*3.28*			*2.65 [0.00]*	*468*			*2.76*	*449 w*	*2.82*[l]	*440*
	3.63; 4.04 [0.00]	*341; 306*	*3.96; 3.27*										

	4.13 [0.01]	*301*	*3.37*			*3.81 [0.77]*	*326*			*3.92*	*316 s*		*4.12*[l]	*301*
						4.23 [0.05]	*293*							
	4.17; 4.19; 4.48; 5.01; 5.06 [0.00]	*298; 296; 277; 247; 245*	*3.27; 4.00; 5.08; 4.02; 4.00*			*4.23; 4.96 [0.00]*	*293; 250*							
	5.37 [0.002]	*231*	*3.21*			*5.47 [0.00]*	*227*			*5.41*	*229 s*		*5.64*[l]	*220*
8. naphthalene	*4.10 [0.00]*	*303*	*4.70*	*9*				*1820*		*3.87*	*320 vw*			
	4.17 [0.02]	*298*	*3.36*			*4.61 [0.00]*	*269*			*4.17*	*297 w*		*3.97*[m]	*313 w*
	4.73 [0.00]	*262*	*4.78*											
	4.75 [0.01]	*261*	*3.65*			*5.99 [0.00]*	*207*			*4.51*	*275 m*		*4.45*[m]	*278 m*
	4.85; 4.97 [0.00]	*256; 249*	*4.75; 4.65*											
	5.07 [0.01]	*245*	*3.87*											
	5.32 [0.00]	*233*	*4.84*											
	5.46 [0.002]	*227*	*3.68*			*6.20 [1.23]*	*200*			*5.61*	*221 s*		*5.89*[m]	*210 s*
9. pentacene	*2.07 [0.04]*	*600*	*2.44*	*965*		*1.94 [0.05]*	*639*	*10834*		*2.12*	*585 s*		*2.31*[n]	*536*
	2.24; 2.40; 3.13; 3.13 [0.00]	*554;517;397;397*	*3.71; 3.57; 2.74; 3.43*											
	3.37 [<0.01]	*368*	*3.05*			*3.06 [0.01]*	*405*			*3.06*	*404*			
	3.52 [0.00]	*353*	*3.32*											
	3.59 [0.05]	*346*	*3.00*			*3.31 [0.01]*	*375*			*3.54*	*350 m*			
	3.85; 3.96 [0.00]	*322; 313*	*3.41; 3.71*			*3.32 [0.00]*	*373*							
	4.05 [0.01]	*306*	*2.79*			*4.14 [0.00]*	*300*			*4.10*	*302 s*			

10. azulene	2.73 [0.01]	454	3.94	9	2.48 [0.01]	499	1574	2.1	606 w				
	3.38 [0.01]	367	3.04		3.77 [0.00]	329		3.8	326 m				
	3.95; 4.29 [0.00]	314; 289	5.02; 4.68										
	4.42 [0.18]	280	3.28					4.19	296 m				
	4.44; 4.52 [0.00]	279; 274	4.76; 4.89										
	4.52 [0.03]	274	3.56		4.80 [0.06]	258		4.51	275 s		4.70[o]	264 s	
11. heptalene	2.53; 3.24 [0.00]	490; 382	4.04; 4.55	29	2.50 [0.00]	497	2834	Long tail[p]					
	3.32 [0.05]	374	3.64		3.70 [0.12]	226		3.56	348 m				
	3.90 [0.03]	318	3.82		4.32 [0.004]	287							
	4.13;4.36 [0.00]	300; 284	4.06; 4.08										
	4.42 [0.05]	281	4.05										
	4.58; 4.62; 4.85; 5.02 [0.00]	271; 269; 256; 247	4.17; 4.01; 4.11; 4.14										
	5.12 [0.01]	242	4.08		4.96 [0.17]	250		4.9	253 s				
	5.15 [0.04]	241	4.19										
12. [18] annulene	1.76 [0.00]	706	2.15	455	1.96 [0.00]	632	10729	1.56[q]	795 w				
	1.85 [0.00]	670	2.02		2.31 [0.00]	537		2.76	448 s				
								2.94	422 m				
								3.04	408 m				
	3.47d [0.89]	358d	2.25d		3.59 [1.48]d	345d		3.36	369 vs				

	4.02d; 4.27; 4.28 [0.00]	309d; 291; 290	3.41d; 3.40; 3.41										
	4.28d [0.00]	290d	2.62d			4.17d [0.00]	297d			4.46	278 m		
	4.36d; 4.36d; 4.64;4.66d; 4.68 [0.00]	285d; 284d; 267; 266d; 265	3.38d; 3.41d; 3.37; 3.48d; 3.47										
	4.70d [0.00]	264d	2.71d			4.46d [0.00]				4.61	269 m		
13. benzaldehyde	3.96 [0.00]	313	5.20	2		3.68 [0.00]	337	873		3.69	336 w		
	4.95 [0.00]	250	4.95			4.90 [0.02]	253			4.46	278 m		4.51	275[r] m
	5.17 [0.01]	240	3.90			5.41[0.26]	229			5.16	240 s		5.34	232[r] s
	5.23 [0.00]	237	4.93										
	5.49 [0.01]	226	3.88										
	5.66 [0.00]	219	5.09										
	5.70 [0.01]	218	4.18										
	5.71; 5.98 [0.00]	217; 208	5.39; 5.31										
	6.00 [0.01]	207	4.11			5.90[0.00]	210						
	6.05; 6.34 [0.00]	205; 196	5.33; 5.11										
	6.47 [0.29]	192	4.24			6.54 [0.15]	190			6.2	200 s		6.36	195[r] s
14. nitrosobenzene	0.68 [0.00]	1816	5.83	2		1.46 [0.00]	847	845		1.65	752 w		
	3.53 [0.00]	352	5.73			4.14 [0.02]	300			4.05	306 m		4.22	294[s] m
	4.38 [0.09]	283	3.69			4.68 [0.24]	265			4.39	282 m		4.56	272[s] s
	4.80 [0.00]	259	5.34										
	4.92 [0.01]	252	3.74			4.81 [0.00]	256						4.71	263[s] s

	5.54; 5.58; 5.60 [0.00]	224; 222; 221	5.41; 5.47; 5.57										
	5.70 [0.02]	218	3.98			5.41 [0.001]	230					5.85	212[s] s
15. p-benzoquinone	2.25 [0.00]	552	5.24	≤1		2.46 [0.00]	503	776		2.71	458 w		
	3.18; 3.87 [0.00]	390; 320	5.43; 5.37			2.64 [0.00]	470						
	3.93 [0.00]	316	4.52			3.92 [0.00]	316			4.46	278 w		
	4.91 [0.00]	252	5.36										
	5.10 [0.10]	243	3.54			5.03 [0.00]	246			5.16	240 w		
16. o- benzoquinone	1.80 [0.00]	688	5.42	≤1		1.79 [0.00]	692	684		2.11	587 w		
						2.99 [0.00]	415						
	3.34 [0.01]	372	4.13			3.22 [0.05]	385			3.37	368 m		
17. pyridazine	2.68 [0.00]	463	6.31	≤1				673					
	3.60 [0.00]	345	5.99			3.71 [0.01]	334			3.65	340 w	3.30	375 w
	4.91 [0.03]	253	4.03			4.37 [0.00]	284			5.04	246 m	4.90	253 w
	5.11 [0.00]	243	5.95			5.61 [0.00]	221						
	5.25 [0.01]	236	4.17			5.80 [0.02]	214					6.2	200 w
	5.99; 6.50 [0.00]	207; 191	5.76; 6.22										
	6.54 [0.01]	190	4.70			6.31 [0.01]	197					6.39	194 m
	6.71 [0.00]	185	5.83			6.69 [0.00]	185						
	6.90 [0.02]	180	4.71										
	7.02 [0.03]	177	5.29			7.54 [0.40]	165					7.10	175 s

18. pyrimidine	2.39 [0.00]	520	6.45	≤1				727					
	3.69 [0.10]	336	4.08			4.05 [0.01]	306			4.16	298 w		3.85	322 w
	4.61 [0.00]	269	6.08										
	5.22 [0.00]	238	6.74			4.83 [0.00]	256			5.08	244 m		5.00	248 w
	5.52 [0.01]	225	3.99			5.56 [0.08]	223						6.42	193 m
	5.59; 5.76; 5.80 [0.00]	222; 215; 214	6.08; 5.81; 5.93			5.73 [0.00]	217							
	6.24 [0.01]	199	4.31			6.65 [0.00]	186						6.49	191 w
	6.45; 6.94 [0.00]	192; 179	5.40; 6.26											
	7.39 [0.00]	168	4.87			6.76 [0.06]	183						6.98	178 m
	7.42; 7.43; 7.69; 7.96 [0.00]	167; 167; 161; 156	5.71; 5.82; 6.29; 5.58			7.77 [0.00]	160							
	8.05 [0.24]	154	4.67			7.98 [0.35]	155						7.25	171 s
19. pyrazine	4.17 [0.00]	298	6.01	≤1		4.40 [0.01]	282	693		3.78	322 m		3.83	324 w
	5.20; 5.45 [0.00]	238; 227	5.76; 5.76											
	5.49 [0.03]	226	4.49			4.75 [0.00]	261			4.77	260 m		4.81	258 w
	6.15; 6.19 [0.00]	202; 200	6.32; 5.97			5.87 [0.00]	211							
	6.50 [0.01]	191	4.57			5.94 [0.03]	209			6.38	194 m		6.31	197 w
	6.58 [0.00]	189	5.86											
	6.60 [0.02]	188	5.22			6.17 [0.01]	201						6.7	185 w
	6.95 [0.01]	179	4.76			6.85 [0.04]	181						6.84	181 m
20. 1,2,4,5-tetrazine	0.80 [0.00]	1560	6.99	≤1				478						

	2.28 [0.00]	*545*	*6.49*			*2.28 [0.01]*	*545*			*2.31*	*539 w*		*2.25*	*551 w*
	3.36 [0.07]	*369*	*3.43*			*3.62 [0.00]*	*343*			*3.7*	*333 sh*			
	3.59; 4.11 [0.00]	*345; 302*	*6.50; 7.09*											
	4.78 [0.02]	*259*	*4.16*			*4.90 [0.00]*	*253*			*4.92*	*252 m*		*5.02*	*247 m*
21. porphirin	*1.52 [0.12]*	*814*	*1.86*	*1393*		*2.32 [0.00]*	*534*	*29035*		*2.21*[t]	*560 w*		*1.98*[t]	*628 w*
	1.71 [0.09]	*727*	*2.00*										*2.16*[t]	*575 w*
	2.36; 2.50; 2.63;2.89 [0.00]	*525; 496; 472; 429*	*3.32; 3.36; 3.34; 3.41*										*2.42*[t]	*512 sh*
	3.20 [0.00]	*388*	*2.44*			*2.49 [0.00]*	*499*			*2.54*[t]	*488 m*		*2.56*[t]	*484 m*
	3.48 [0.24]	*356*	*2.45*			*3.41 [0.50]*	*364*			*3.15*[t]	*394 s*		*3.33*[t]	*373 vs*
	3.49 [0.00]	*354*	*3.34*			*3.53 [0.00]*	*352*			*3.31*[t]	*375*			
	3.63 [0.15]	*342*	*2.56*			*3.58 [0.79]*	*346*			*3.50*[t]	*354*		*3.65*[t]	*340 sh*
	3.68 [0.00]	*337*	*3.29*											
	3.68 [0.00]	*337*	*3.32*			*3.75 [0.00]*	*331*							
	3.76; 3.81; 3.84; 3.86; 3.89; 3.90; 3.96; 3.98; 3.99; 4.07; 4.16; 4.23; 4.35; 4.51 [0.00]	*330; 326; 323; 321; 319; 318; 313; 311; 310; 305; 298; 293; 285; 275*	*3.26; 3.32; 3.31; 2.81; 3.36; 3.34; 3.28; 3.33; 3.33; 3.31; 3.25; 3.39; 3.34; 3.34*											
	4.70 [0.15]	*264*	*2.82*			*3.89 [0.37]*	*319*							
	4.70; 4.71	*264;*	*2.77;*											

	[0.00]	*264*	*2.69*										
	4.80 [0.20]	*258*	*2.82*			*3.96 [0.75]*	*313*			*4.17*[t]	*297*		*4.25*[t]	*292 s*
	4.81; 5.00; 5.09 [0.00]	*258; 248; 244*	*3.44; 3.48; 2.73*			*4.22; 4.25; 4.26; 4.30; 4.31; 4.37; 4.43 [0.00]*	*294; 292; 291; 289; 288; 284; 280*							
						4.49 [0.13]	*276*							
						4.61 [0.00]	*269*							
	5.17 [0.75]	*240*	*2.97*			*4.65 [0.12]*	*267*							
	5.17; 5.18; 5.25; 5.25 [0.00]	*240; 240; 236; 236*	*3.32; 2.67; 3.35; 3.25*			*4.84; 5.00; 5.35 [0.00]*	*256; 248; 232*							
	5.28 [1.24]	*235*	*2.85*			*5.41 [0.14]*	*229*			*5.39*[t]	*230*		*5.51*[t]	*225 s*
22. OO-trans-2-furaldehyde	*3.80 [0.00]*	*327*	*5.79*	*≤1*		*3.71 [0.00]*	*334*	*786*		*3.89*[u]	*318 w*		*3.50*[v]	*354 w*
	3.81 [0.06]	*326*	*3.99*			*5.09 [0.36]*	*244*			*4.65*[u]	*267 s*		*4.65*[v]	*267 s*
	3.93; 4.01 [0.00]	*316; 309*	*5.71; 5.87*											
	5.23 [0.02]	*237*	*4.38*			*5.98 [0.02]*	*207*			*5.50*[u]	*225 m*			

a. Energies in eV, wavelengths (λ) in nm, and *f* are calculated oscillator strengths. Intensities: s = strong, m = medium, w = weak, vw = very weak, sh = shoulder, f = diffuse. E_{CE} is the Coulomb exchange term for excitonic binding. t_{CPU} is the CPU time for runs calculating excitation energy, at fixed geometry, in seconds. Quantities smaller than 1 sec. have always been overscaled to 1.
b. Calculated results at the FCIS CNDOL/2CC||MP2|4-31g(d,p) level.

c. Calculated results at the TD-PBE0[6-31G(d,p)]|MP2[4-31G(d,p)] level as well behaved for UV spectra values according the conclusions of ref. [23].
d. Ref. [63] if otherwise is not quoted.
e. Ref. [64] if otherwise is not quoted.
f. Ref. [65].
g. Ref. [66].
h. Ref. [67].
i. Solution in heptane. See ref. [68].
j. Fluorescence excitation spectrum in Ar matrix at 15 K. See ref. [69].
k. Gas phase spectrum band reported in ref. [70].
l. See ref. [71].
m. Ref. [72].
n. Ref. [73].
o. Ref. [74].
p. Ref. [75].
q. Low energy absorption spectrum reported in frozen 3-methyl pentane at 77 K. See ref. [76]. Other maxima found in iso-octane solution reported in ref. [77].
r. Ref. [78].
s. Ref.[79].
t. Ref. [80]. Solution spectrum in CH_2Cl_2. All referred to band maxima (vertical transitions).
u. Ref. [81].
v. Ref. [82].

Our benchmark results are shown in Table 3. It should be observed that no omission was made of any output calculated value of either method, leaving some of the resulting theoretically forbidden transitions without an experimental counterpart. This is expected due to the very broad bands of experimental UV-Vis spectra and the approximate nature of both theoretical methods.

47 gas phase maxima calculated using CNDOL/2CC and 48 TD-DFT calculations were compared in this study. The number of compared maxima for solution absorption maxima were 74 and 71, respectively. A mean average error (MAE) of 0.27 eV of CNDOL/2CC results was found between solution maxima and 0.29 eV between those observed in the gas phase. The corresponding squared regression coefficients are $r_{soln}^2 = 0.9332$ and $r_{gas}^2 = 0.9356$, respectively. Values for our PBE0 TD-DFT calculations with the same set of experimental values and independent assignments were 0.30 eV ($r_{soln}^2 = 0.8964$) and 0.32 eV ($r_{gas}^2 = 0.8913$), respectively. A graphic comparison is illustrated in Fig. 2. It was observed that problems arise with assignments of t-azobenzene and naphthalene, with PBE0 TD-DFT predicted UV maxima at both solution and gas phases contributing to a poorer performance when using this level of theory.

Previous work using 500 compounds, with more than 700 calculated excited states as a benchmark determined that the best functional to use with TD-DFT in order to predict transition energies of singlet-excited states of organic molecules was such PBE0[61, 62] with a MAE of 0.22eV from a set of 614 excited states.[19] A more recent benchmark for UV-NIR spectra of dyes that could potentially be utilized in dye-sensitized solar cells gave a MAE of 0.18 eV by B2LYP TD-DFT calculations of 102 absorption maxima. Calculated transition energies of cationic cyanines, of anionic oxonols and various related dyes are much too high (MAE = 0.52 eV, 32 data) and correspondingly the absorption wavelengths are too low.[17] Our results can therefore be interpreted to have a comparable statistical confidence, even choosing a previous and fixed benchmarking sample, and removing no outliers from the comparison. Molecule 22, which was added in order to extend the sample, does not contribute to produce an improved fit, as can be easily observed in Table 3.

High throughput techniques have been shown to be effective in the design of molecules for photovoltaic applications[45, 83]. For these techniques, the CPU time expended for each

calculation is of critical importance. This value is overwhelmingly in favor of CNDOL/2CC results when calculated for processing the data set described in this paper. All calculations were performed in the same standalone computer to allow proper comparisons. CNDOL/2CC calculations resulted 25 times faster than TD-DFT as an average over all calculations. This means that our method allows both modeling of much larger systems, allowing up to nanoscopic dimensions, and also much larger libraries of molecules – an essential property in the realm of High Throughput Virtual Screening.

An inspection of Coulomb - exchange (E_{CE}) contributions to the assigned electron transition energies in Table 3 allows us to postulate some general rules. It becomes clear that traditional band assignment based on $n\pi^*$ and $\pi\pi^*$ like transitions remains non appropriate because we are dealing with many electron wave functions, some of them belonging to non – planar systems, and it means that the symmetry or quasi symmetry of electron clouds with respect to the molecular plane becomes distorted. However, E_{CE} serves as an index of locality and binding of excitons, as stated in a previous paper[55].

Significant differences among E_{CE} energies only appear in molecules that involve electronic states of different symmetry with respect to molecular planes. It can be observed that the largest E_{CE} term contributions, that reflect an exciton involving more intense electron interactions, mostly appear in small molecules and chromophores with heteroatoms presenting transitions that could be labeled as $n\pi^*$ type. They become even larger as the molecular dimensions become smaller. Those excitons that imply a collective charge rearrangement over extended π electron systems show lower E_{CE} energy values. This case can be understood as describing phenomena of an extended charge-hole quasiparticle interactions or less bound excitons. Qualitative results are in agreement with the aforementioned previous work where the behavior of charge distributions upon the excitation in molecular aggregates was described by means of E_{CE} values. Consequently, E_{CE} data is here presented as a tool for defining exciton related properties in molecular systems.

Charge distribution maps of the benchmark and Harvard's clean energy project (CEP) molecules can be built with the results of eq. (14) by plotting them against the molecular geometry with

Jmol, an open source project (v. 12.2.32). A *python* script for parsing the appropriate output values of NDOL2014 was written for this purpose.

A comparison of acrolein charge distribution maps is shown in Figure 3 to test the validity of CNDOL/2CC's charge distribution towards other values when reduced to atoms. Red to blue regions (negative to positive values) denote an accumulation and depletion of electron density, respectively, while green color represents those molecular regions where the partial charge is around zero. This comparison can only be approximate, because the orthogonal basis of CNDOL is hard to compare with Mulliken's charge reduced to heavy atoms, or Bader's "atoms in molecules" considerations. However, a qualitative similarity is clearly observed over all.

Fig. 4 shows Q_ξ charge distribution maps for the ground and the lowest three calculated excited states of benzaldehyde, together with differential maps for each excited state, $\Delta Q_\xi = Q_0 - Q_\xi$, where it is shown the charge variation of each of them with respect to the ground state. It was made with the purpose of visualizing the "dynamics" of electron cloud rearrangements upon excitation, which should be related with the E_{CE} gauge described above. The numerical meaning of the color scale is doubled for the ΔQ_ξ maps, although in this case it represents the density changes upon electronic excitation. Color scales are the same for all Q_ξ figures for the sake of facilitate comparisons.

The S_1 state involves a charge redistribution of the electron density only involving extreme zones of the map, while S_2 and S_3 states involve the whole molecule. It is reflected by the E_{CE} energies. Taking into account that here the π electron system is delocalized over the whole molecule, the length scale of the charge redistribution should be similar for these excited states. Consequently, insignificant differences could be expected among the E_{CE} values compared with other examples in the present benchmark (see Table 3). However, ΔQ_ξ maps clearly show that only the S_1 state imply a charge rearrangement, mostly over the carbonyl group. The other states (S_2 and S_3) involve changes mainly over the benzene chromophore, as expected.

Fig. 5 shows a map for the lowest two calculated excitations on a porphyrin molecule where electron cloud migrate to both sides of the macro ring according to the symmetry. The concentration of charge notably appears in the center of the ring, leaving a place for any cation

(as it occurs in chlorophyll with Mg^{2+}) to balance it in a very stable condition, even for excited states.

Finally, CEP's molecules are shown in Figs. 6 – 8. They are those reported in a recent paper[45] as among the most promising of more than 2.6 million screened as potential organic solar cell materials. Rules for establishing their abilities as solar light harvesters are described in the related papers. What we easily find after a simple observation of graphics and values for excitation energy and Coulomb – exchange term is that CEP77 seems very attractive since expected to absorb photons in the near IR, showing the lowest exciton binding and a dramatic charge displacement upon excitation from the donor to the Si containing portion of the molecule.

CONCLUSIONS

A significant set of molecules selected because of their difficulty for obtaining good reproduction of UV-Vis spectra with several methods was calculated by both FCIS CNDOL/2CC and PBE0 TD-DFT methods. The results display the reliability of the approximate quantum mechanical MO procedure for predicting excited state properties on the grounds of a well-designed theoretical framework for simplifications. This result show an important increase in the efficiency of the calculation of these properties, since this approximate method displays comparable results to the traditional, and more expensive, TD-DFT procedures. This has large implications for High Throughput Virtual Screening efforts, in which the ability to calculate properties for extremely large libraries of molecules within a reasonable time period is critical. In term of method development, this contribution shows advances which show CNDOL to be a reliable and useful tool with modest computational requirements for even significantly large molecular systems.

It is also shown that the FCIS scheme, i.e. a very large active CIS space, provided adequate excited state descriptions even in the case of small molecules, where the SCF basis set does not provide good molecular basis functions.

In addition, the possibility to use several Fockians as constructed from convenient combination of one electron matrix element terms combined with an appropriate adiabatic procedure for

reducing probabilities of non-convergence problems, offers a flexible method for affording and finally obtaining electronic states and density matrices of different kinds of molecular systems. The behavior of the Coulomb – exchange energy component of FCIS electronic transitions involving multielectronic wave functions together with calculations of density matrices of each excited state also provide a source of valuable data for modeling and understanding processes that cannot be perceived by experimental means.

Computer graphics of charge distribution of different electronic states are shown as a simple, yet informative, way for perceiving a qualitative behavior of charge displacements and location in either ground or excited states. It could be used as a tool for understanding the behaviors of molecules and nanoscopic systems upon excitation.

ACKNOWLEDGEMENTS

This work was mostly supported by the Universidad de La Habana, in Cuba, and the Universidad Autónoma de Madrid, in Spain, with the support of the project A1/035856/11 of the *Agencia Española de Cooperación Internacional para el Desarrollo*. AAG and EOPK are grateful to the Department of Energy for funding through Grant DE-SC0008733. LAMC is also indebted to the David Rockefeller Center for Latin American Studies through a Santander Scholarship supported in Harvard University in 2012. Several colleagues facilitated the completion of this work as Dr. Semion Saikin and Professor Jorge Domínguez (Harvard). This work has been inspired by the mentor and most important critics of LAMC, Prof. Jürgen Fabian, of TU Dresden.

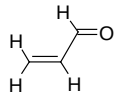
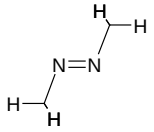
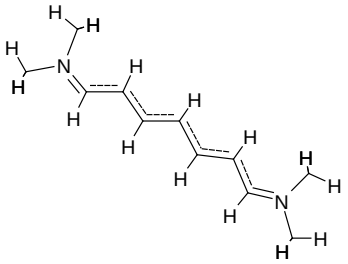
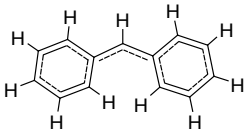
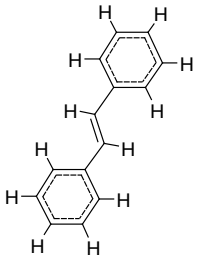
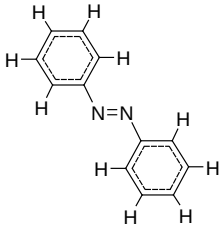
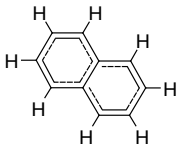
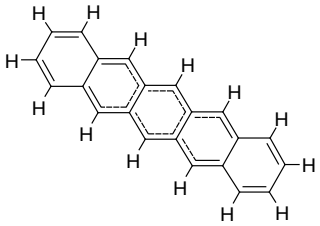
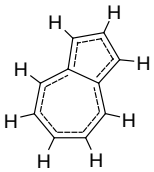
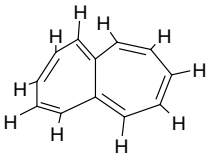
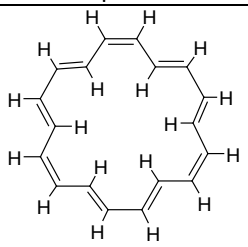
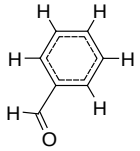
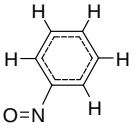
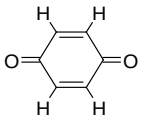
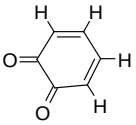
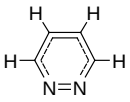
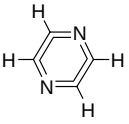
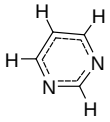
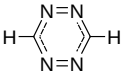
REFERENCES

1. Carter, E. A., *SCIENCE* **2008,** *321* (5890), 800-803.
2. Petersilka, M.; Gossmann, U. J.; Gross, E. K. U., *Phys. Rev. Lett.* **1996,** *76* (8), 1212-15.
3. Stowasser, R.; Hoffmann, R., *J. Am. Chem. Soc.* **1999,** *121* (14), 3414-3420.
4. Pariser, R.; Parr, R. G., *J. Chem. Phys.* **1953,** *21*, 466-71.
5. Pariser, R.; Parr, R. G., *J. Chem. Phys.* **1953,** *21*, 767-76.
6. Pople, J. A., *Trans. Faraday Soc.* **1953,** *49*, 1375-85.
7. Jaffe, H. H., *Acc. Chem. Res.* **1969,** *2* (5), 136-143.
8. Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J., *J. Phys. Chem.* **1992,** *96* (1), 135-149.
9. Li, S. L.; Marenich, A. V.; Xu, X.; Truhlar, D. G., *The Journal of Physical Chemistry Letters* **2013,** *5* (2), 322-328.
10. Mochizuki, Y.; Koikegami, S.; Amari, S.; Segawa, K.; Kitaura, K.; Nakano, T., *Chem. Phys. Lett.* **2005,** *406* (4-6), 283-288.
11. Dreuw, A.; Head-Gordon, M., *Chem. Revs.* **2005,** *105*, 4009-4037.
12. Sims, J. S.; Hagstrom, S. A., *J. Chem. Phys.* **2014,** *140* (22), 224312.
13. Li, R.; Sun, E.; Jin, M.; Xu, H.; Yan, B., *J. Phys. Chem. A* **2014,** *118* (14), 2629-2637.
14. Fukuda, R.; Ehara, M.; Cammi, R., *J. Chem. Phys.* **2014,** *140* (6), 064114.
15. Kronik, L.; Tkatchenko, A., *Acc. Chem. Res.* **2014,** *47* (11), 3208-3216.
16. Pastore, M.; Mosconi, E.; De Angelis, F.; Grätzel, M., *J. Phys.Chem. C* **2010,** *114* (15), 7205-7212.
17. Fabian, J., *Dyes and Pigments* **2010,** *84* (1), 36-53.
18. Sauer, S. P. A.; Schreiber, M.; Silva-Junior, M. R.; Thiel, W., *J. Chem. Theory Comput.* **2009,** *5* (3), 555-564.
19. Jacquemin, D.; Wathelet, V.; Perpète, E. A.; Adamo, C., *J. Chem. Theory Comput.* **2009,** *5* (9), 2420-2435.
20. Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W., *J. Chem. Phys.* **2008,** *129* (10), 104103-14.
21. Schreiber, M.; Silva-Junior, M. R.; Sauer, S. P. A.; Thiel, W., *J. Chem. Phys.* **2008,** *128* (13), 134110-25.
22. Rolik, Z.; Szabados, A.; Surjan, P. R., *J. Chem. Phys.* **2008,** *128* (14), 144101-11.
23. Jacquemin, D.; Perpete, E. A.; Scuseria, G. E.; Ciofini, I.; Adamo, C., *J. Chem. Theory Comput.* **2008,** *4* (1), 123-135.
24. Pieniazek, P. A.; Arnstein, S. A.; Bradforth, S. E.; Krylov, A. I.; Sherrill, C. D., *J. Chem. Phys.* **2007,** *127* (16), 164110-19.
25. Trani, F.; Scalmani, G.; Zheng, G.; Carnimeo, I.; Frisch, M. J.; Barone, V., *J. Chem. Theory Comput.* **2011,** *7*, 3304–3313.
26. Silva-Junior, M. R.; Thiel, W., *J. Chem. Theory Comput.* **2010,** *6* (5), 1546-1564.
27. Peverati, R.; Truhlar, D. G., *PCCP* **2012,** *14* (32), 11363-11370.
28. Körzdörfer, T.; Brédas, J.-L., *Acc. Chem. Res.* **2014,** *47* (11), 3284-3291.
29. Montero, L. A.; Alfonso, L.; Alvarez, J. R.; Perez, E., *Int. J. Quantum Chem.* **1990,** *37* (4), 465-83.
30. Montero-Cabrera, L. A.; Röhrig, U.; Padron-García, J. A.; Crespo-Otero, R.; Montero-Alejo, A. L.; García de la Vega, J. M.; Chergui, M.; Röthlisberger, U., *J. Chem. Phys.* **2007,** *127* (14), 145102.
31. Fabian, J.; Diaz, L.; Seifert, G.; Niehaus, T., *THEOCHEM* **2002,** *594*, 41-53.
32. Pople, J. A.; Segal, G. A., *J. Chem. Phys.* **1966,** *44* (9), 3289-96.

33. Pople, J. A.; Segal, G. A., *J. Chem. Phys.* **1965,** *43* (10;Pt. 2), S136-S149, discussion, S150-S151.
34. Pople, J. A.; Santry, D. P.; Segal, G. A., *J. Chem. Phys.* **1965,** *43* (10;Pt. 2), S129-S135.
35. Voityuk, A. A.; Zerner, M. C.; Roesch, N., *J. Phys. Chem. A* **1999,** *103* (23), 4553-4559.
36. Rossi, I.; Truhlar, D. G., *Chem. Phys. Lett.* **1995,** *233* (3), 231-236.
37. Thiel, W., *Theor. Chim. Acta* **1981,** *59*, 191-208.
38. Dewar, M. J. S.; Thiel, W., *Theor. Chem. Acc.* **1977,** *46* (2), 89-104.
39. Stewart, J. J. P., *J. Comput. Chem.* **1989,** *10* (2), 221-264.
40. Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P., *J. Am. Chem. Soc.* **1985,** *107* (13), 3902-3909.
41. Pople, J. A.; Beveridge, D. L.; Dobosh, P. A., *J. Chem. Phys.* **1967,** *47* (6), 2026-33.
42. Pérez Badell, Y.; Montero, L. A.; Perez, C., *THEOCHEM* **2006,** *769* (1-3), 77-82.
43. Montero, L. A.; Diaz, L. A.; Castillo, N., *Chem. Phys. Lett.* **2002,** *364* (1,2), 176-179.
44. Fuentes, M. E.; Peña, B.; Contreras, C.; Montero, A. L.; Chianelli, R.; Alvarado, M.; Olivas, R.; Rodríguez, L. M.; Camacho, H.; Montero-Cabrera, L. A., *Int. J. Quantum Chem.* **2008,** *108*, 1664–1673.
45. Hachmann, J.; Olivares-Amaya, R.; Jinich, A.; Appleton, A. L.; Blood-Forsythe, M. A.; Seress, L. R.; Román-Salgado, C.; Trepte, K.; Atahan-Evrenk, S.; Er, S.; Shrestha, S.; Mondal, R.; Sokolov, A.; Bao, Z.; Aspuru-Guzik, A., *Energy & Environmental Science* **2013,** *7* (2), 698-704.
46. Pariser, R., *J. Chem. Phys.* **1953,** *21*, 568-9.
47. Nishimoto, K.; Mataga, N., *Z. Phys. Chem. (Munich)* **1957,** *12*, 335-8.
48. Mataga, N.; Nishimoto, K., *Z. Phys. Chem. (Munich)* **1957,** *13*, 140-157.
49. Ohno, K., *Theor. Chim. Acta* **1964,** *2* (3), 219-227.
50. Wolfsberg, M.; Helmholz, L., *J. Chem. Phys.* **1952,** *20* (5), 837-843.
51. Hinze, J.; Jaffe, H. H., *J. Am. Chem. Soc.* **1962,** *84*, 540-6.
52. Clementi, E.; Raimondi, D. L., *J. Chem. Phys.* **1963,** *38*, 2686-9.
53. Hylleraas, E., *Z. Physik* **1928,** *48*, 469-494.
54. Montero-Alejo, A. L.; Menendez-Proupin, E.; Fuentes, M. E.; Delgado, A.; Montforts, F. P.; Montero-Cabrera, L. A.; Garcia de la Vega, J. M., *PCCP* **2012,** *14* (37), 13058-13066.
55. Montero-Alejo, A. L.; Fuentes, M. E.; Montero, L. A.; de la Vega, J. M. G., *Chem. Phys. Lett.* **2011,** *502* (4-6), 271-276.
56. Montero-Alejo, A. L.; Fuentes, M. E.; Menéndez-Proupin, E.; Orellana, W.; Bunge, C. F.; Montero, L. A.; García de la Vega, J. M., *Phys. Rev. B* **2010,** *81* (23), 235409.
57. Pople, J. A., *Proc. Phys. Soc. London* **1955,** *68A*, 81-9.
58. Montero Cabrera, L. A.; Montero Alejo, A. L.; Bunge Molina, C.; González Santana, S.; Crespo Otero, R.; Mora Díez, N.; Delgado Gran, A. *NDOL2014: A computer program for calculation of approximate SCF-CI electron excitations and excited state properties of molecules*, 6.81; Universidad de La Habana and Universidad Autónoma de Madrid: Havana, 2014.
59. Voityuk, A. A., *J. Chem. Theory Comput.* **2014,** *10* (11), 4950-4958.
60. 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.; J. A. Montgomery, J.; 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, J. M.; 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, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, , Revision A.01; Gaussian, Inc.: Wallingford CT, 2009.
61. Adamo, C.; Scuseria, G. E.; Barone, V., *J. Chem. Phys.* **1999,** *111* (7), 2889-2899.
62. Adamo, C.; Barone, V., *J. Chem. Phys.* **1999,** *110* (13), 6158-6170.
63. D.M.S., *UV Atlas of Organic Compounds*. Butterworths – Verlag Chemie: Weinheim – London, 1966 – 67.
64. Herzberg, G., *Molecular Spectra and Molecular Structure. III Electronic Spectra and Electronic Structure of Polyatomic Molecules*. Van Nostrand Reinhold Company: New York, 1966; p 745.
65. Walsh, A. D., *Trans. Faraday Soc.* **1945,** *41*, 498-505.
66. Nagakura, S., *Mol. Phys.* **1960,** *3* (2), 152-162.
67. Calvert, J. G.; Pitts, J. N., *Photochemistry*. John Wiley & Sons, Inc.: New York, 1966; p 899.
68. Suzuki, H., *Bull. Chem. Soc. Jpn.* **1960,** *33* (3), 379-388.
69. Gudipati, M. S.; Maus, M.; Daverkausen, J.; Hohlneicher, G., *Chem. Phys.* **1995,** *192* (1), 37-47.
70. Champagne, B. B.; Pfanstiel, J. F.; Plusquellic, D. F.; Pratt, D. W.; Van Herpen, W. M.; Meerts, W. L., *J. Phys. Chem.* **1990,** *94* (1), 6-8.
71. Andersson, J.-Å.; Petterson, R.; Tegnér, L., *Journal of Photochemistry* **1982,** *20* (1), 17-32.
72. George, G. A.; Morris, G. C., *J. Mol. Spectrosc.* **1968,** *26* (1), 67-71
73. Heinecke, E.; Hartmann, D.; Müller, R.; Hese, A., *J. Chem. Phys.* **1998,** *109* (3), 906-911.
74. Brouwer, L. D.; Hippler, H.; Lindemann, L.; Troe, J., *J. Phys. Chem.* **1985,** *89* (21), 4608-4612.
75. Dauben, H. J.; Bertelli, D. J., *J. Am. Chem. Soc.* **1961,** *83* (22), 4659-4660.
76. Baumann, H.; Oth, J. F. M., *Helv. Chim. Acta* **1982,** *65* (6), 1885-1893.
77. Sondheimer, F.; Wolovsky, R.; Amiel, Y., *J. Am. Chem. Soc.* **1962,** *84* (2), 274-284.
78. Kimura, K.; Nagakura, S., *Theor. Chim. Acta* **1965,** *3*, 164-173.
79. Tabei, K.; Nagakura, S., *Bull. Chem. Soc. Jpn.* **1965,** *38* (6), 965-971.
80. Edwards, L.; Dolphin, D. H.; Gouterman, M.; Adler, A. D., *J. Mol. Spectrosc.* **1971,** *38* (1), 16-32.
81. Montero-Cabrera, L. A unified approach of orbital resonance integrals for semiempirical calculations in the zero differential overlap molecular orbital approximation. Technische Universität Dresden, Dresden, 1980.
82. Zwarich, R.; Rabinowitz, I., *J. Chem. Phys.* **1975,** *63* (11), 4565-4577.
83. Hachmann, J.; Olivares-Amaya, R.; Atahan-Evrenk, S.; Amador-Bedolla, C.; Sánchez-Carrera, R. S.; Gold-Parker, A.; Vogt, L.; Brockway, A. M.; Aspuru-Guzik, A. n., *The Journal of Physical Chemistry Letters* **2011,** *2* (17), 2241-2251.
84. Wiberg, K. B.; Rosenberg, R. E.; Rablen, P. R., *J. Am. Chem. Soc.* **1991,** *113* (8), 2890-2898.

Figure 1. Selected set of benchmark molecules

			
1.- acrolein	2.- nitromethane	3.- t-azomethane	4.- N, N, N', N' – tetramethyl-1,7-diamino-2,4,6-heptatriene cation
			
5.- diphenyl-methane cation	6.- t-stilbene	7.- t-azobenzene	8.- naphthalene
			
9.- pentacene	10.- azulene	11.- heptalene	12.- [18] annulene
			
13.- benzaldehyde	14.- nitrosobenzene	15.- p-benzoquinone	16.- o- benzoquinone
			
17.- pyridazine	18.- pyrimidine	19.- pyrazine	20.- 1,2,4,5-tetrazine

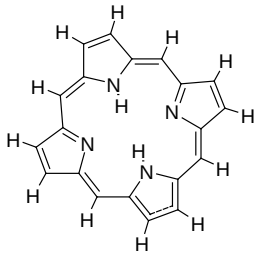
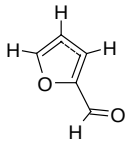
	
21.- porphirin	22.- OO-trans-2-furaldehyde

Figure 2. Plot and linear fit of calculated excitation energies according the match in Table 3 for reported UV spectra of benchmarking molecules.

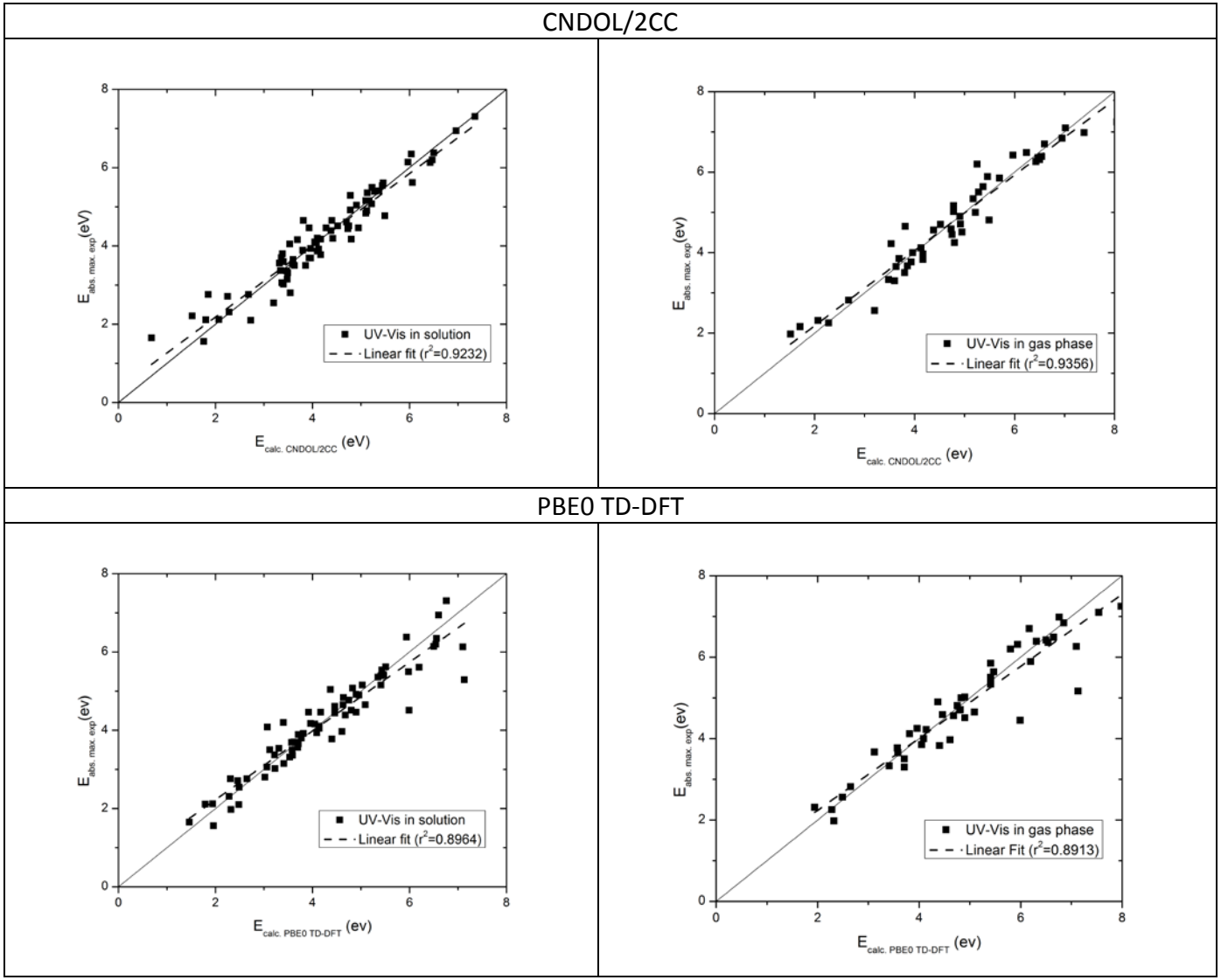
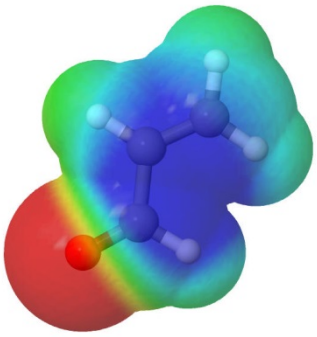
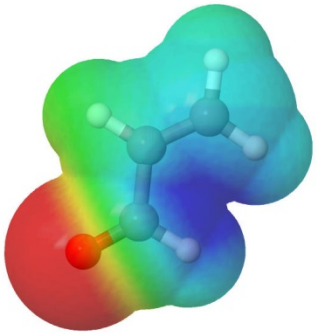
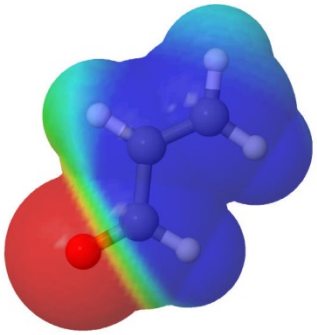


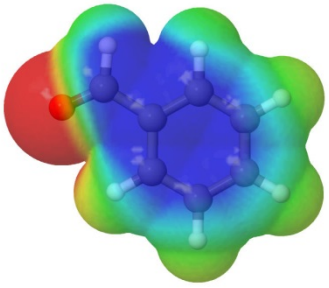
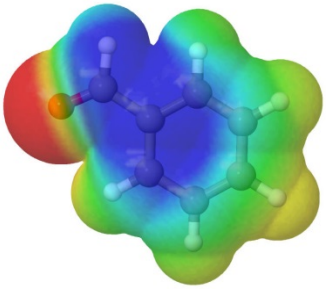
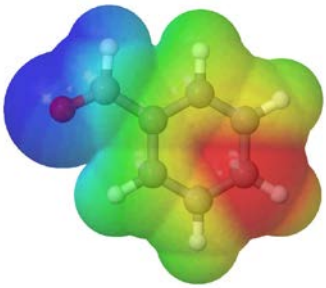
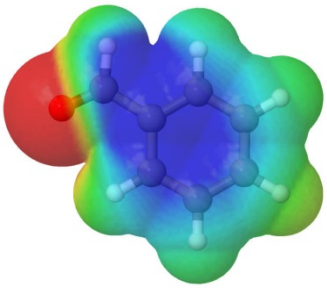
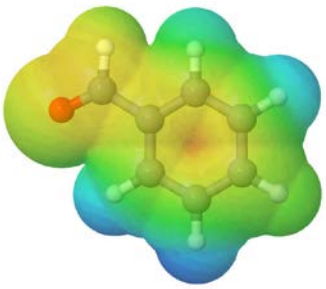
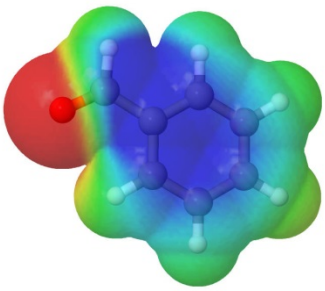
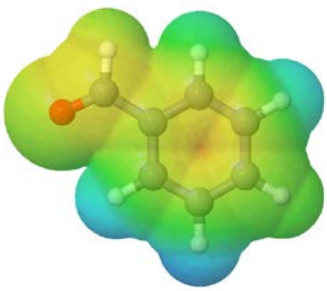
Figure 3. Electron charge maps of acrolein ground state (Q_{S0}) according different calculation approaches[a]

CNDOL/2CC[b]	
SCF-MO 4-31G(d,p), Mulliken's reduced charges to heavy atoms	
SCF-MO 6-31G(d,p)|4-31G(d,p) Atom charges according Bader's bond critical point along each bond[84]	

a.- Red and blue regions denote a negative and positive density (accumulation and depletion of electron density charge) respectively, and green regions represent a near to zero electron charge. All Q maps are plotted using the same isosurface contour value.

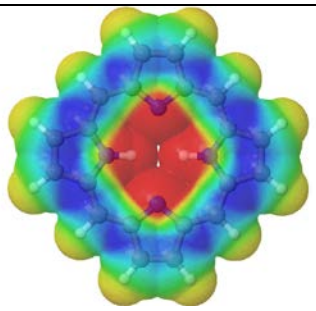
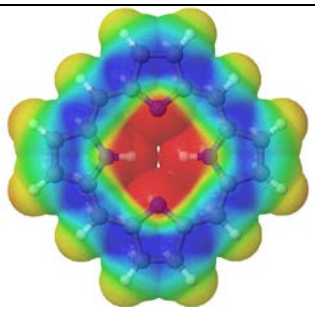
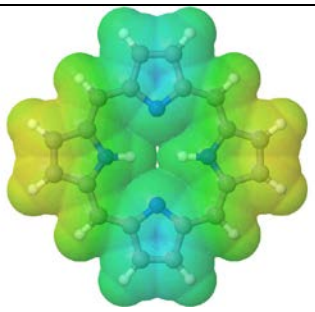
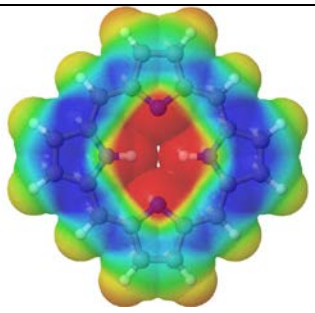
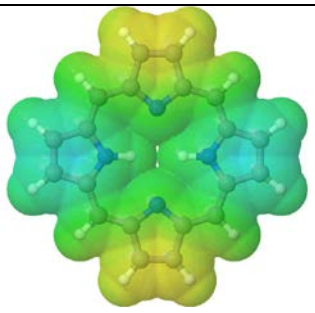
b.- Electron charge on A atom is calculated as $Q_A = Z^{core}_A - P_A$ where Z^{core}_A is the nuclear charge.

Figure 4. Electron charge maps of lowest energy states of benzaldehyde as calculated with the FCIS CNDOL/2CC procedure[a]

		
	Q_{S0}	
S_1, E_{CE} = 5.20 ev		
	Q_{S1}	ΔQ_{S1}
S_2, E_{CE} = 4.95 ev		
	Q_{S2}	ΔQ_{S2}
S_3, E_{CE} = 3.90 ev		
	Q_{S3}	ΔQ_{S3}

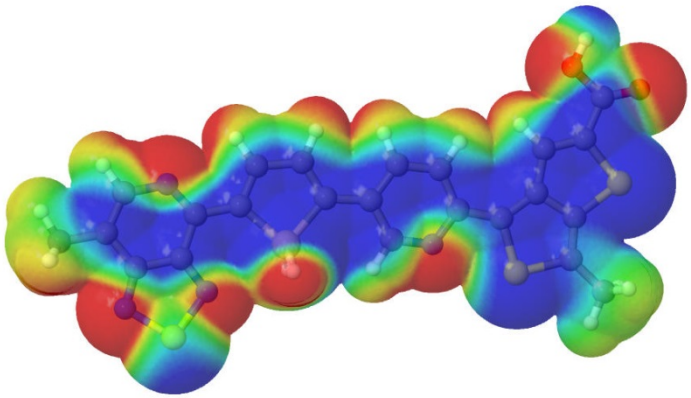
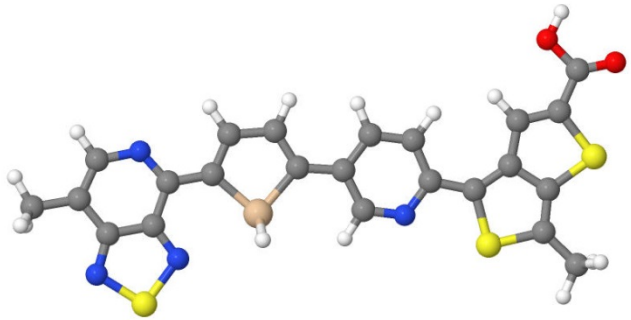
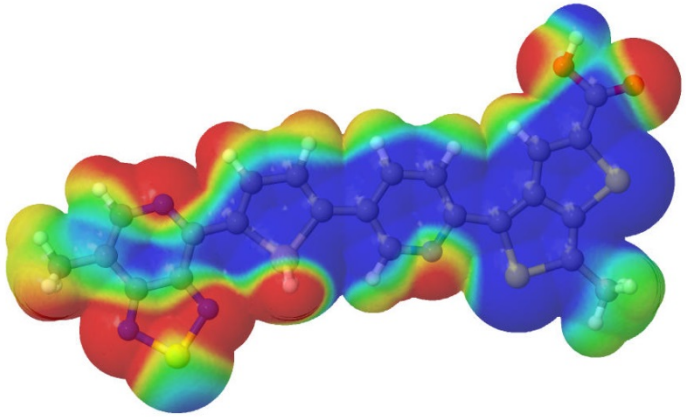
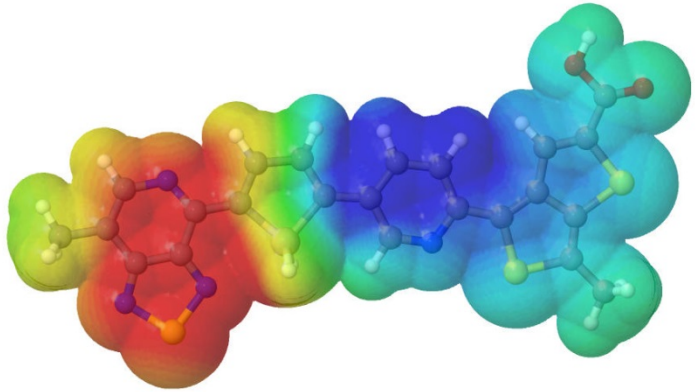
a.- Electron charge on each "A" atom is calculated as $Q_A = Z^{core}_A - P_A$ where Z^{core}_A is the core charge (eq. 14 in text). $\Delta Q_{Sn} = Q_{S0} - Q_{Sn}$ to illustrate charge evolution upon excitation. Ground state (Q_{s0}) and first (Q_{s1}), second (Q_{S2}) and third (Q_{S3}) excited states charge maps are represented by colors. Red and blue regions denote a negative and positive density (accumulation and depletion of electron density charge) respectively, and green regions represent a near to zero electron charge. All Q maps are plotted using the same isosurface contour value. ΔQ color contrast has been doubled with respect to Q maps to enhance the observation of differences. The Coulomb exchange term (E_{CE}) is shown for each excited state.

Figure 5. Electron charge maps of lowest energy states of porphirin as calculated with the FCIS CNDOL/2CC procedure[a]

		
	Q_{S0}	
S_1, E_{CE} = 1.86 eV		
	Q_{S1}	ΔQ_{S1}
S_2, E_{CE} = 2.00 eV		
	Q_{S2}	ΔQ_{S2}

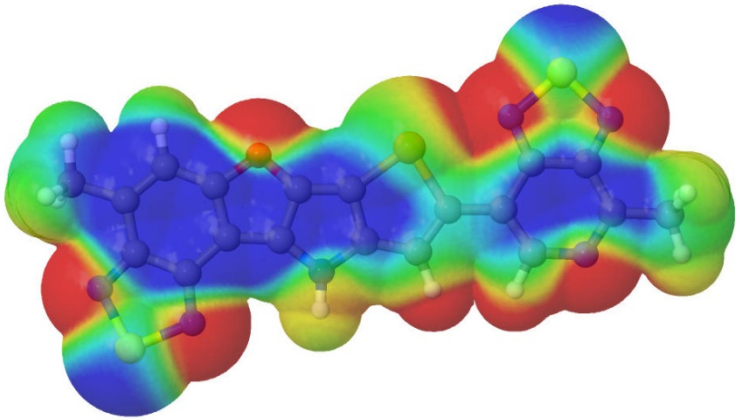
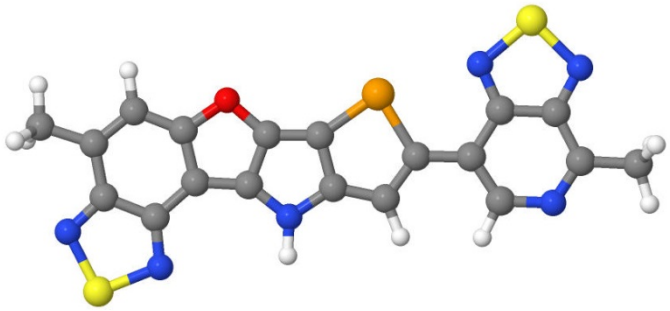
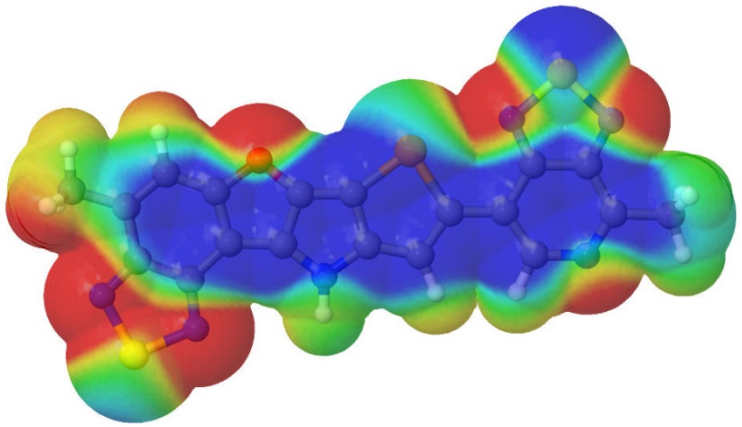
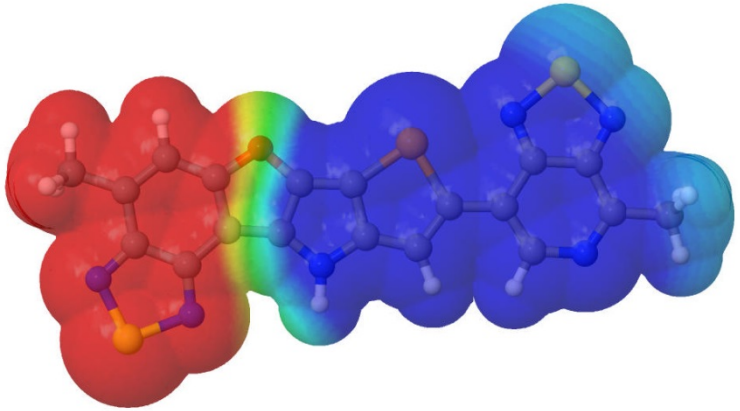
a.- Electron charge on each "A" atom is calculated as $Q_A = Z^{core}_A - P_A$ where Z^{core}_A is the core charge (eq. 14 in text). $\Delta Q_{Sn} = Q_{S0} - Q_{Sn}$ to illustrate charge evolution upon excitation. Ground state (Q_{s0}) and first (Q_{s1}), and second (Q_{S2}) excited states charge maps are represented by colors. Red and blue regions denote a negative and positive density (accumulation and depletion of electron density charge) respectively, and green regions represent a near to zero electron charge. All density Q maps are plotted using the same isosurface contour value. ΔQ color contrast has been doubled with respect to Q maps to enhance the observation of differences. The Coulomb exchange term (E_{CE}) is shown for each excited state

Figure 6. Electron charge maps of CEP1 ground and lowest energy excited state as calculated with the FCIS CNDOL/2CC procedure[a]

S_0	 Q_{S0}	
S_1 = 1.49 eV; E_{CE} = 3.77 eV	 Q_{S1}	 ΔQ_{S1}

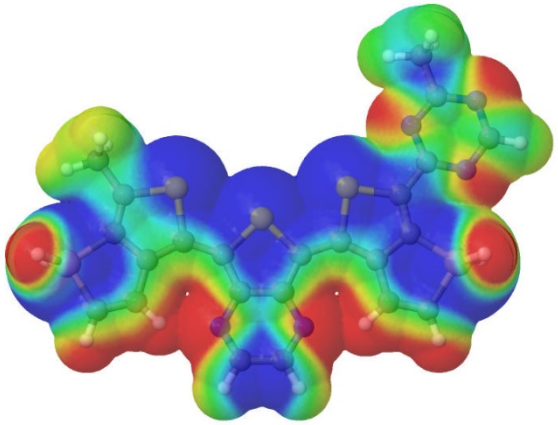
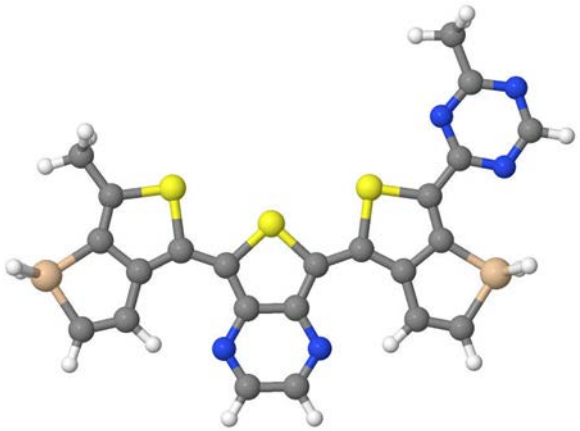
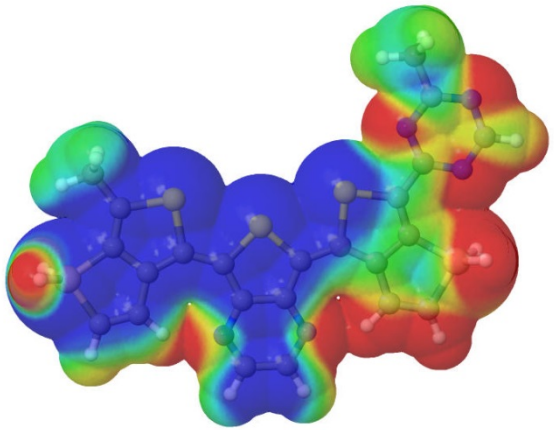
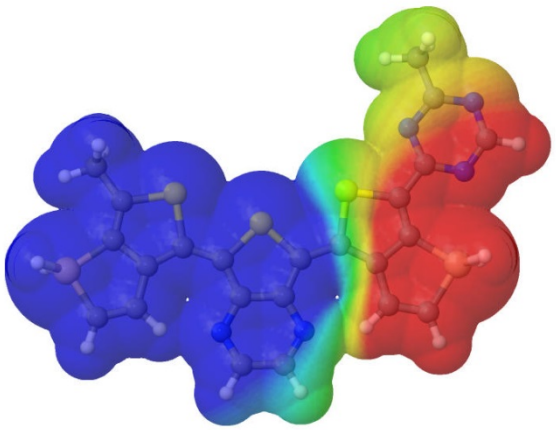
a.- Electron charge on A atom is calculated as $Q_A = Z^{core}_A - P_A$ where Z^{core}_A is the nuclear charge. $\Delta Q_{Sn} = Q_{S0} - Q_{Sn}$ to illustrate charge evolution upon excitation. Element color code: H (white); C (grey); N (blue); O (red); S (sulfur); Si (rose); Se (orange)

Figure 7. Electron charge maps of CEP5 ground and lowest energy excited state as calculated with the FCIS CNDOL/2CC procedure[a]

S_0	 Jmol Q_{S0}	 Jmol
S_1 = 1.10 eV; E_{CE} = 2.94 eV	 Jmol Q_{S1}	 Jmol ΔQ_{S1}

a.- Electron charge on A atom is calculated as $Q_A = Z^{core}_A - P_A$ where Z^{core}_A is the nuclear charge. $\Delta Q_{Sn} = Q_{S0} - Q_{Sn}$ to illustrate charge evolution upon excitation. Element color code: H (white); C (grey); N (blue); O (red); S (sulfur); Si (rose); Se (orange)

Figure 8. Electron charge maps of CEP77 ground and lowest energy excited state as calculated with the FCIS CNDOL/2CC procedure[a]

S_0	 Jmol Q_{S0}	 Jmol
S_1 = 1.09 eV; E_{CE} = 2.66 eV	 Jmol Q_{S1}	 Jmol ΔQ_{S1}

a.- Electron charge on A atom is calculated as $Q_A = Z^{core}_A - P_A$ where Z^{core}_A is the nuclear charge. $\Delta Q_{Sn} = Q_{S0} - Q_{Sn}$ to illustrate charge evolution upon excitation. Element color code: H (white); C (grey); N (blue); O (red); S (sulfur); Si (rose); Se (orange)