JCTC Journal of Chemical Theory and Computation

Benchmark of Electronically Excited States for Semiempirical Methods: MNDO, AM1, PM3, OM1, OM2, OM3, INDO/S, and INDO/S2

Mario R. Silva-Junior and Walter Thiel*

Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr, Germany

Received January 15, 2010

Abstract: Semiempirical configuration interaction (CI) calculations with eight different Hamiltonians are reported for a recently proposed benchmark set of 28 medium-sized organic molecules. Vertical excitation energies and one-electron properties are computed using the same geometries as in our previous ab initio benchmark study on electronically excited states. The CI calculations for the standard methods (MNDO, AM1, PM3) and for the orthogonalization-corrected methods (OM1, OM2, OM3) include single, double, triple, and quadruple excitations (CISDTQ) using the graphical unitary group approach (GUGA) as implemented in the MNDO code. The CIS calculations for the established INDO/S method and the reparametrized INDO/S2 variant employ a modified version of the ZINDO program. As compared to the best theoretical reference data from the ab initio benchmark, all currently applied semiempirical methods tend to underestimate the vertical excitation energies, but the errors are much larger in the case of the standard methods (MNDO, AM1, PM3). Overall, the mean absolute deviations relative to the theoretical best estimates are lowest for OM3, and only slightly higher for OM1 and OM2 (in the range of 0.4−0.5 eV). INDO/S performs similar to OM2 for the vertical excitation energies of singlet states, but deteriorates considerably for triplet states. The INDO/S2 reparametrization for oxygen improves the results for low-lying singlet states of oxygen-containing compounds, but makes them worse for high-lying singlets as well as for triplets. The ab initio reference data for oscillator strengths and excited-state dipole moments are again best reproduced by the orthogonalization-corrected approaches (OM1, OM2, OM3), which thus emerge as the most favorable semiempirical methods overall for treating valence excited states of large organic chromophores.

1. Introduction

In recent years, there has been much progress in the research on electronically excited states. Elaborate experimental techniques are available to study photophysical processes in the nanosecond or femtosecond regime. Concomitantly, improved theoretical methods have been developed that allow realistic calculations on excited states and may thus provide guidance for the experimental work. On the ab initio side, MS-CASPT2 (multistate complete-active-space second-order perturbation theory)[1−4] and coupled cluster methods (CC2, CCSD, CC3)[5−7] are well established and offer high accuracy for small molecules. Time-dependent density functional theory (TD-DFT)[8] has become popular for calculations on medium-sized molecules, giving reasonable results for various (but not all) types of excited states at relatively low computational cost.[9,10] An alternative DFT-based method makes use of Kohn−Sham orbitals in a multireference configuration interaction (MRCI) framework, modified by incorporating five universal empirical parameters to alleviate problems with the double counting of dynamic electron correlation.[11]

* Corresponding author e-mail: thiel@mpi-muelheim.mpg.de.

Electronically Excited States

J. Chem. Theory Comput., Vol. 6, No. 5, 2010 **1547**

For a quantitative assessment of different theoretical approaches, reliable reference data are needed for benchmarking. Standard test sets are widely used for ground-state properties, for example, the G2 and G3 sets for thermochemistry.[12,13] We have recently introduced an ab initio benchmark set for electronically excited states of 28 medium-sized organic molecules with a total of 223 excitations.[14] On the basis of MS-CASPT2 and CC3 calculations on these molecules and of high-level ab initio data from the literature, we have proposed theoretical best estimates for the vertical excitation energies of 104 singlet and 63 triplet excited states. These reference data have been used to evaluate the performance of standard TD-DFT and DFT/MRCI approaches,[15] of the coupled cluster variant CCSDR(3) with noniterative triples corrections,[16] and of TD-DFT with a large number of different functionals[17,18] including double-hybrid functionals.[18]

Despite recent advances, the reliable description of electronically excited states in large molecules is still a challenging problem. Accurate ab initio methods such as MS-CASPT2 and CC3 are restricted to small molecules, and the computational cost for simpler treatments such as CC2 or DFT/MRCI still rises steeply with molecular size. TD-DFT is an attractive choice because of its computational efficiency and the availability of analytical gradients, but there are a number of well-documented problems of TD-DFT,[9,10] for example, with regard to change-transfer states[19] and singlet or triplet instabilities.[20] Moreover, the overall accuracy of TD-DFT is limited, with vertical excitation energies that typically show mean absolute deviations in the range of 0.3−0.5 eV from the theoretical best estimates in our benchmark set.[15] Given this situation, it seems worthwhile to explore the performance of semiempirical quantum-chemical methods for electronically excited states of large organic molecules.

Standard semiempirical molecular orbital (MO) methods such as MNDO,[21] AM1,[22] and PM3[23] are based on the NDDO (neglect of diatomic differential overlap) integral approximation and have been parametrized against ground-state properties, in particular heats of formation and geometries. They have been widely applied in computational studies of ground-state processes (for reviews, see, for example, refs 24−26). Applications to electronically excited states[27,28] are rare, however, mainly because these standard methods normally underestimate their energies strongly, as a result of the integral approximations and the ground-state parametrization. A straightforward remedy for this shortcoming would be a system-specific reparameterization[29] for a given application (see, for example, ref 30), which is, however, cumbersome in practice and also unsatisfactory from a conceptual point of view.

For the semiempirical calculation of vertical excitation energies, INDO/S (intermediate neglect of differential overlap for spectroscopy)[31,32] has been the method of choice for a long time. INDO/S describes excited states by CIS (configuration interaction with single excitations) and has been parametrized at this level. It has been widely used in studies of organic molecules[31,32] as well as transition metal complexes[33] and even lanthanides.[34] The INDO/S2 variant[35] is

a reparametrization designed to improve the results for oxygen-containing compounds. The lack of higher excitations in the INDO/S CI treatment effectively restricts applications to states dominated by single excitations. Another limitation is the focus on vertical processes: by its design, INDO/S targets spectroscopy rather than photochemistry, and it is thus not made for the exploration of excited-state potential energy surfaces (PES).

The orthogonalization-corrected OM*x* methods (OM1,[36,37] OM2,[38,39] and OM3[40]) employ the NDDO integral approximation, but go beyond the standard methods (MNDO, AM1, PM3) by including additional terms in the Fock matrix that represent Pauli exchange repulsions in an approximate manner. These terms effectively raise the energy of antibonding virtual MOs and of the associated excited states.[37,39] Therefore, one would expect an improved performance of the OM*x* methods not only for ground-state properties,[41] but also for excited-state properties, which had not been taken into account during the OM*x* parametrization. The applications published so far support this view, for example, the OM2 studies on the electronically excited states of butadiene,[42] retinal model systems,[43] and the rhodopsin chromophore.[44] In addition, OM2 predicts reasonable geometries for a set of 12 typical conical intersections[45] (as compared to ab initio reference data). Finally, OM2 has successfully been applied in excited-state surface-hopping dynamics calculations for several small molecules,[46] for all nucleobases,[47−49] and for retinal models.[50] These promising indications call for a more comprehensive assessment, with detailed comparisons to established semiempirical treatments.

In this Article, we present a systematic evaluation of the performance of the standard NDDO-based semiempirical methods (MNDO, AM1, PM3), the commonly used INDO-based approaches (INDO/S, INDO/S2), and the orthogonalization-corrected methods (OM1, OM2, OM3) for electronically excited states. Reference data are taken from our previous benchmark work and comprise theoretical best estimates as well as MS-CASPT2/TZVP and CC3/TZVP data.[14] This Article is structured as follows: Section 2 describes the computational methods used. Section 3 presents some general considerations on the current benchmarking. Sections 4 and 5 discuss the individual results for vertical excitation energies and one-electron properties, respectively. Section 6 is devoted to statistical evaluations, and section 7 offers a brief summary and outlook.

2. Computational Methods

All calculations were carried out at the optimized ground-state equilibrium geometries reported previously.[14,51] The standard semiempirical Hamiltonians with default parameters were used for MNDO, AM1, PM3, OM1, OM2, and OM3, as implemented in the current version of the MNDO99 code.[52] In the case of INDO/S, singlet states were computed using the default parameters, with $f_{\pi\pi} = 0.585$ and the Mataga−Nishimoto expression for the two-center two-electron repulsion integrals, while triplet states were treated using the recommended special parametrization and the Pariser−Parr formula for the Coulomb integrals.[32] The

Table 1. OM2 Results for the Two Lowest Singlet Excited States of the Linear Polyenes with k Double Bonds, for Different Types of CI Treatment[a]

threshold (%)	FCI	CISDTQ	MR-CISD[a]			
			90	85	80	75
$k = 3$						
2^1A_g (eV)	4.86	4.86	4.86	4.87	4.87	4.89
1^1B_u (eV)	5.33	5.33	5.33	5.33	5.33	5.32
CSFs (references)[b]	175	165	162 (10)	150 (7)	146 (6)	126 (5)
CPU time (s)[c]	0.03	0.03	0.03	0.03	0.03	0.02
$k = 4$						
2^1A_g (eV)	4.19	4.20	4.21	4.22	4.25	4.25
1^1B_u (eV)	4.79	4.79	4.78	4.78	4.78	4.78
CSFs (references)[b]	1764	1195	916 (12)	758 (9)	651 (6)	651 (6)
CPU time (s)[c]	0.35	0.25	0.22	0.17	0.14	0.15
$k = 5$						
2^1A_g (eV)	3.68	3.72	3.74	3.75	3.78	3.79
1^1B_u (eV)	4.40	4.41	4.39	4.38	4.37	4.38
CSFs (references)[b]	19 404	6601	4067 (16)	3079 (11)	2207 (7)	2003 (6)
CPU time (s)[c]	10.80	2.22	1.21	1.06	0.48	0.42
$k = 6$						
2^1A_g (eV)	3.32	3.41	3.41	3.43	3.48	3.50
1^1B_u (eV)	4.11	4.13	4.10	4.09	4.07	4.08
CSFs (references)[b]	226 512	28 278	13 254 (20)	8781 (13)	5931 (8)	4868 (6)
CPU time (s)[c]	446.09	17.90	6.98	4.05	2.15	1.52
$k = 7$						
2^1A_g (eV)	3.06	3.20	3.19	3.22	3.27	3.30
1^1B_u (eV)	3.75	3.87	3.88	3.87	3.85	3.85
CSFs (references)[b]	2 760 615	98 785	35 336 (26)	24 531 (16)	15 982 (9)	11 380 (7)
CPU time (s)[c]	144 070.91[d]	120.37	32.42	17.01	8.30	5.10

[a] See text. Active space composed of all π and π^* orbitals. Geometries optimized at the B3LYP/TZVP level. [b] Total number of configuration state functions in A_g and B_u symmetry (total number of reference configurations in MR-CISD given in parentheses). [c] Computation times refer to one AMD Opteron(tm) 845 2.8 GHz processor. [d] A semidirect algorithm is used, with all coupling coefficients being recomputed as needed.

INDO/S and INDO/S2 calculations were done with the ZINDO-MN program, version 1.2.[53]

Consistent with the underlying parametrization procedure, the INDO/S and INDO/S2 results were obtained at the CIS level. Following standard INDO/S conventions, the active space generally included the 10 highest occupied MOs and the 10 lowest unoccupied MOs, yielding a total of 101 configuration state functions (CSF).[31] In small molecules with less MOs, all occupied and unoccupied MOs were normally included, but high-lying states were treated with caution: when large spurious $\sigma \rightarrow \sigma^*$ contributions were encountered, the corresponding σ^* MOs were deleted from the active space. The excited-state dipole moments were computed using the recommended class IV charge model 2 (CM2).[35]

The NDDO-based semiempirical methods considered presently (MNDO, AM1, PM3, OM1, OM2, OM3) have been parametrized against ground-state reference data at the SCF (self-consistent-field) level, so that the effects of dynamic ground-state correlation should conceptually be taken into account in an average manner through the parametrization (and through the use of damped two-electron integrals). In electronically excited states, however, there are often static (near-degeneracy) correlation effects, which call for an explicit treatment also in a semiempirical framework, using a suitably chosen (small) active space. To be as unbiased as possible, we adopted a canonical active space with m electrons in n orbitals (mn) for each molecule, in analogy to our previous MS-CASPT2 benchmark study.[14]

This active space includes all occupied and unoccupied π-MOs in the case of $\pi \rightarrow \pi^*$ excitations, and in addition the occupied lone-pair MOs in the case of $n \rightarrow \pi^*$ excitations. For a given active space, the least biased correlation treatment is full CI, which, however, quickly becomes too expensive even at the semiempirical level. Therefore, our standard approach in the present benchmark was chosen to be the single-reference CISDTQ treatment, which includes all single, double, triple, and quadruple excitations relative to the closed-shell SCF determinant and which is expected to provide a balanced description of all relevant states (close to the full CI limit).

To check the performance and efficiency of different CI approaches, we performed test calculations on linear polyenes with k double bonds using the CI implementation in the MNDO99 code that is based on the graphical unitary group approach (GUGA).[54] Table 1 lists the OM2 results obtained from full CI (FCI), CISDTQ, and various MR-CISD treatments (multireference CI with single and double excitations), which are approximations to CISDTQ. It is obvious that the excitation energies from FCI calculations are reproduced very well by CISDTQ; the deviations (which increase with molecular size) are mostly much smaller than 0.1 eV. At the same time, the computational costs are much reduced for CISDTQ, for example, by a factor of 25 in the case of $k = 6$ where an in-core FCI treatment is still feasible; the much larger factor for $k = 7$ is caused by the switch to a less efficient semidirect FCI algorithm in the MNDO99 code (due to the large size of the CI matrix). In the MR-CISD

Electronically Excited States

J. Chem. Theory Comput., Vol. 6, No. 5, 2010 **1549**

Table 2. Statistical Results and Computation Times (s) for the Set of 222 Vertical Excitation Energies (in eV) Using Different Levels of Excitation in the CI Treatment for the OM1, OM2, and OM3 Methods[a]

threshold (%)	FCI	CISDTQ	MR-CISD[b]			
			90	85	80	75
OM1						
count[c]		222	222	222	222	222
mean		0.01	0.03	0.03	0.03	0.03
abs. mean		0.01	0.03	0.03	0.03	0.03
std. dev.		0.01	0.05	0.05	0.06	0.06
max. (+) dev.		0.04	0.27	0.28	0.30	0.28
max. (−) dev.			0.14	0.13	0.13	0.13
CPU time (s)[d]	513.84	54.13	20.04	16.61	15.53	13.96
OM2						
count[c]		222	222	222	222	222
mean		0.00	0.02	0.02	0.03	0.03
abs. mean		0.00	0.02	0.02	0.03	0.03
std. dev.		0.01	0.04	0.04	0.05	0.05
max. (+) dev.		0.03	0.16	0.18	0.19	0.19
max. (−) dev.			0.01	0.01	0.02	0.02
CPU time (s)[d]	510.65	52.53	18.09	15.74	14.88	13.62
OM3						
count[c]		222	222	222	222	222
mean		0.00	0.02	0.02	0.02	0.03
abs. mean		0.00	0.02	0.02	0.03	0.03
std. dev.		0.01	0.04	0.04	0.05	0.05
max. (+) dev.		0.04	0.16	0.17	0.16	0.21
max. (−) dev.			0.03	0.03	0.02	0.02
CPU time (s)[d]	485.64	52.42	17.62	15.30	14.02	13.42

[a] Results from a full CI within the selected active space are taken as reference. [b] See text. [c] Total number of states considered. [d] Computation times refer to one Intel Pentium 4-EMT64T 3.40 GHz processor.

calculations, the reference configurations are selected by an iterative procedure to fulfill the requirement that their combined weight in the CI wave function must exceed a given threshold value (for example, 90% or 85%); starting from a single-reference calculation, this is accomplished by adding the next most important reference configurations until this condition is satisfied (normally within one or two iterations). It is again evident that the MR-CISD results are essentially identical to the CISDTQ results and that they can be obtained at significantly less cost. This conclusion is corroborated by the corresponding data for all other benchmark molecules that are documented in the Supporting Information (see Tables S1−S6) and by the statistical data collected in Table 2. For the full set of vertical excitation energies in our benchmark, the mean absolute deviations from the OMx/FCI reference values amount to 0.00−0.01 eV for OMx/CISDTQ and to 0.02−0.03 eV for OMx/MR-CISD (thresholds 75−90%). In actual applications, it is thus perfectly legitimate to perform such semiempirical excited-state studies at the MR-CISD level; a threshold of 85% seems more than sufficient to ensure close reproduction of the CISDTQ and the FCI results. For the purposes of the present benchmark, we shall, however, rely on CISDTQ.

A final remark in this section concerns the efficiency of semiempirical CI methods relative to TD-DFT. For a direct comparison, we have computed the three low-lying $^1B_2$ excited states of pyridine using the MNDO99 code for OM2/CISDTQ and the TURBOMOLE package (version 5.9.1)[55] for TDDFT-B3LYP/TZVP (starting from a previously converged SCF solution). The ratio of computation times is

1:1578 on a single-processor Intel Pentium 4-EMT64T (3.4 GHz), indicating a difference of about 3 orders of magnitude for a typical example from our benchmark suite. It is clear that the superior speed of the semiempirical CI methods will allow applications that are not feasible with TD-DFT or the even more costly ab initio methods, for example, calculations on larger chromophores or extended excited-state dynamics runs, provided that the accuracy of the semiempirical results is sufficient.

3. General Considerations

Geometries. As already mentioned, the presently used geometries are taken from our previous benchmark.[14] They represent equilibrium ground-state geometries optimized at the MP2/6-31G* level in a suitable point group (assuming the highest symmetry possible). For the sake of consistency, it is obviously advantageous to adopt a common set of geometries during benchmarking, but one may still wonder by how much the results would change upon reoptimizing the geometries at the semiempirical level. One would actually expect rather small changes because semiempirical methods generally yield realistic ground-state geometries for organic molecules (see ref 41 for corresponding statistical data). Test calculations on several of the benchmark molecules confirm this view. Table S7 (Supporting Information) lists the results for a typical example, two low-lying states of pyridine obtained at geometries optimized with MP2/6-31G*, AM1, and OM2. The computed AM1 and OM2 excitation energies vary by up to 0.1 eV for the different geometries, the (small) oscillator strengths appear to be rather insensitive, and the excited-state dipole moments show variations of around 0.1 D. Overall, these changes are small enough to justify the assumption that the qualitative conclusions of the present benchmark study will remain valid also when using geometries optimized at the semiempirical level.

States. Semiempirical methods employ a minimal basis of valence orbitals and can therefore not describe Rydberg states or states with substantial valence/Rydberg mixing properly. The currently used benchmark set was designed to include only valence excited states[14] and should thus be suitable for an assessment of semiempirical methods. One should keep in mind, however, that there is no clear-cut distinction between valence and other excited states in ab initio calculations, and there will be borderline cases especially for higher-lying states whose character may even change at different ab initio levels.[56] Despite these caveats, we present semiempirical results for all valence states considered previously,[14] while acknowledging that it may be easier for semiempirical methods to properly describe the low-lying valence states (say, below 6 eV).

Assignments. For each benchmark molecule, the electronically excited states were first classified according to point-group symmetry. Thereafter, within a given irreducible representation, one has to establish the proper correspondence between the states obtained in the ab initio reference calculations[14] and the present semiempirical calculations. This was accomplished by comparing the computed excited-state wave functions, along with the excitation energies, oscillator strengths, and excited-state dipole moments. Pro-

ceeding in this manner, a satisfactory mapping has been achieved in all cases. However, one specific problem should be pointed out. It is well-known[38,39] that the standard NDDO-based methods (MNDO, AM1, PM3) underestimate the gap between bonding and antibonding MOs, with the latter ones being too low in energy because of the symmetric splitting of bonding and antibonding levels. This creates special problems in alternant hydrocarbons and related molecules, where two singly excited configurations strongly contribute to two excited states, which qualitatively correspond to the plus and minus combination of these configurations (for example, HOMO → LUMO + 1 and HOMO − 1 → LUMO generating the L_b and B_b states in the polyenes). In the case of the standard NDDO-based methods, it is easier to populate a higher unoccupied orbital than to vacate the alternancy-related lower occupied orbital (as compared to ab initio methods). Therefore, the relative energies of the two corresponding singly excited configurations will differ, which will translate into different weights in the resulting CI wave functions of the corresponding pair of states. Such differences in the character of states have indeed been observed between standard NDDO-based and ab initio results, with the assignment based primarily on the character of the states in these cases.

4. Vertical Excitation Energies

From our previous study,[14] we have MS-CASPT2 reference data for 152 singlet states and 71 triplet states, and theoretical best estimates (TBE) for the vertical excitation energies of 104 singlet states and 63 triplet states. The full set of the computed semiempirical vertical excitation energies is given in Tables 3 and 4 along with the corresponding MS-CASPT2 and TBE results. INDO/S and INDO/S2 values are listed for all molecules, although they differ only for oxygen-containing compounds (INDO/S2 reparametrization for oxygen). For the OMx methods, the vertical excitation energies obtained from FCI, CISDTQ, and MR-CISD (thresholds 75−90%) calculations are documented in more detail in the Supporting Information (Tables S1−S6). A compilation of experimental data is available from one of our previous papers.[51]

In the following discussion of the results, we will focus on comparisons with the TBE values, but our qualitative conclusions remain valid also when considering CASPT2 or CC3 reference data.

Ethene, Butadiene, Hexatriene, and Octatetraene. The energy of the singlet $\pi\pi^*$ state of ethene (TBE 7.80 eV) is well reproduced by the OMx methods (errors of less than 0.1 eV), but strongly underestimated by the standard MNDO-type methods (by 1.2−1.6 eV) and somewhat overestimated by INDO/S (by 0.53 eV). The energy of the triplet $\pi\pi^*$ state of ethene (TBE 4.50 eV) is underestimated by all semiempirical methods, but to a varying extent (MNDO/AM1/PM3 by 1.5−1.9 eV, INDO/S by 1.3 eV, OMx by 0.34−0.43 eV).

Butadiene is the first member of the $C_{2n}H_{2n+2}$ polyene series. The excitation energy to the bright 1^1B_u $\pi\pi^*$ state (TBE 6.18 eV) is well predicted by the OMx methods (errors

of less than 0.1 eV) and also by INDO/S (too low by 0.21 eV), while the MNDO/AM1/PM3 values are again much too low (by 1.4−1.9 eV). More interesting is the dark 2^1A_g $\pi\pi^*$ state, which is mainly composed of two single excitations (HOMO → LUMO + 1 and HOMO − 1 → LUMO) and the double excitation HOMO $\Rightarrow$ LUMO (contributing about 33% to the CISDTQ wave function of any of the NDDO-based methods). While it has been recognized early on that this state becomes the lowest excited singlet for longer polyenes,[57,58] its position in butadiene has remained controversial for a long time. Relative to the currently adopted TBE of 6.55 eV, the excitation energies from MNDO/AM1/PM3 and from OMx are too low by 2.0−2.7 and 0.57−0.67 eV, respectively. The standard INDO/S CIS approach does not capture the double-excitation character of the 2^1A_g state and thus overestimates its energy by 0.32 eV; it has been pointed out[59] that inclusion of the doubly excited HOMO$\Rightarrow$LUMO configuration reduces the error significantly (to 0.06 eV). This is partly fortuitous, however, because other doubly excited configurations contribute another 20% to the OMx/CISDTQ wave functions such that this state is actually dominated by double excitations. The energies of the two lowest triplet $\pi\pi^*$ states of butadiene (1^1B_u and 2^1A_g) are again underestimated by all semiempirical methods, least so by OMx (errors of 0.30−0.59 eV).

In hexatriene and octatetraene, we focus on the state ordering of the two lowest singlet excited states, 2^1A_g and 1^1B_u. The energy gap between these states is minute in hexatriene (TBE 0.01 eV) and still small in octatetraene (TBE 0.19 eV), with 2^1A_g being lower. MNDO/AM1/PM3 and OMx overestimate these gaps by 1.0−1.4 and 0.47−0.70 eV, respectively; in the case of OMx, the 1^1B_u state is described quite well (too high by 0.04−0.26 eV), while the 2^1A_g state comes out too stable (by 0.23−0.38 eV). As expected, INDO/S CIS calculations give the wrong state ordering for these two small polyenes, with 2^1A_g more than 1.0 eV above 1^1B_u (including the doubly excited HOMO$\Rightarrow$LUMO configuration in the CI calculation does not recover the correct state ordering). The energies for the lowest 1^3B_u triplet state of hexatriene (TBE 2.40 eV) and octatetraene (TBE 2.20 eV) are well reproduced by OMx (errors of 0.11 eV or less), while those for the 1^3A_g triplet state (TBE 4.15 and 3.55 eV) are somewhat underestimated (by 0.31−0.42 eV). It should also be noted that all six NDDO-based methods predict realistic singlet−triplet energy differences (which may be important photochemically): the TBE differences $\Delta\Delta E$ (1^1B_u−1^3B_u) are reproduced to within 0.17−0.28 eV for hexatriene and 0.19−0.33 eV for octatetraene, and even better in the case of the A_g states (absolute deviations of at most 0.12 and 0.21 eV, respectively).

Cyclopropene, Cyclopentadiene, Norbornadiene, Benzene, and Naphthalene. The energies of the singlet and triplet excited states of cyclopropene are generally underestimated by all NDDO-based methods. For instance, in the case of the 1^1B_2 $\pi\pi^*$ state (TBE 7.06 eV), the deviations are about 1.5 eV for MNDO/AM1/PM3 and 0.47−0.67 eV for OMx. This state is rather diffuse at the CASSCF level[60] and may thus be problematic for methods using minimal basis sets.

Electronically Excited States

J. Chem. Theory Comput., Vol. 6, No. 5, 2010 **1551**

Table 3. Vertical Singlet Excitation Energies ΔE (eV) of All Evaluated Molecules As Compared to MS-CASPT2/TZVP Results and Theoretical Best Estimates (TBE)

molecule	state	CASPT2[a]	TBE[b]	MNDO	AM1	PM3	OM1	OM2	OM3	INDO/S	INDO/S2
ethene	$1\,^1B_{1u}$ $(\pi \to \pi^*)$	8.54	7.80	6.18	6.63	6.63	7.82	7.78	7.85	8.33	8.33
E-butadiene	$1\,^1B_u$ $(\pi \to \pi^*)$	6.47	6.18	5.18	5.46	5.45	6.24	6.22	6.26	5.97[c]	5.97
	$2\,^1A_g$ $(\pi \to \pi^*)$	6.62	6.55	3.90	4.41	4.53	5.88	5.96	5.98	6.87[c]	6.87
hexatriene	$1\,^1B_u$ $(\pi \to \pi^*)$	5.31	5.10	4.59	4.78	4.77	5.35	5.33	5.36	4.86[c]	4.86
	$2\,^1A_g$ $(\pi \to \pi^*)$	5.42	5.09	3.23	3.63	3.74	4.79	4.86	4.86	5.90[c]	5.90
all-*E*-octatetraene	$2\,^1A_g$ $(\pi \to \pi^*)$	4.64	4.47	2.81	3.13	3.24	4.09	4.14	4.13	5.23[c]	5.23
	$1\,^1B_u$ $(\pi \to \pi^*)$	4.70	4.66	4.23	4.36	4.35	4.79	4.77	4.79	4.20[c]	4.20
cyclopropene	$1\,^1B_1$ $(\sigma \to \pi^*)$	6.76	6.76	5.22	5.35	5.67	6.33	5.75	5.93	6.92[d]	6.92
	$1\,^1B_2$ $(\pi \to \pi^*)$	7.06	7.06	5.58	5.54	5.58	6.59	6.42	6.39	6.90[d]	6.90
cyclopentadiene	$1\,^1B_2$ $(\pi \to \pi^*)$	5.51	5.55	4.37	4.61	4.68	5.14	5.07	5.09	5.03	5.03
	$2\,^1A_1$ $(\pi \to \pi^*)$	6.31	6.31	3.66	4.11	4.24	5.52	5.60	5.59	6.06	6.06
	$3\,^1A_1$ $(\pi \to \pi^*)$	8.52		6.25	6.51	6.52	7.63	7.47	7.53	7.72	7.72
norbornadiene	$1\,^1A_2$ $(\pi \to \pi^*)$	5.34	5.34	5.04	5.18	5.38	6.10	6.00	6.06	4.50	4.50
	$1\,^1B_2$ $(\pi \to \pi^*)$	6.11	6.11	5.83	5.99	6.11	6.65	6.34	6.46	5.54	5.54
	$2\,^1B_2$ $(\pi \to \pi^*)$	7.32		6.26	6.38	6.49	7.42	7.37	7.42	6.77	6.77
	$2\,^1A_2$ $(\pi \to \pi^*)$	7.45		6.27	6.33	6.45	6.92	6.66	6.74	6.90	6.90
benzene	$1\,^1B_{2u}$ $(\pi \to \pi^*)$	5.04	5.08	2.71	3.13	3.19	4.41	4.48	4.51	4.71	4.71
	$1\,^1B_{1u}$ $(\pi \to \pi^*)$	6.42	6.54	4.49	4.84	4.90	5.98	5.94	6.03	5.96	5.96
	$1\,^1E_{1u}$ $(\pi \to \pi^*)$	7.13	7.13	5.55	5.92	5.88	7.13	7.16	7.20	6.51	6.51
	$1\,^1E_{2g}$ $(\pi \to \pi^*)$	8.18	8.41	4.46	5.11	5.20	7.07	7.19	7.22	7.79	7.79
naphthalene	$1\,^1B_{3u}$ $(\pi \to \pi^*)$	4.24	4.24	2.35	2.68	2.75	3.76	3.81	3.84	3.92	3.92
	$1\,^1B_{2u}$ $(\pi \to \pi^*)$	4.77	4.77	3.84	4.06	4.09	4.85	4.83	4.87	4.50	4.50
	$2\,^1A_g$ $(\pi \to \pi^*)$	5.87	5.87	3.20	3.68	3.76	5.16	5.23	5.27	5.52	5.52
	$1\,^1B_{1g}$ $(\pi \to \pi^*)$	5.99	5.99	3.80	4.24	4.33	5.67	5.74	5.76	5.65	5.65
	$2\,^1B_{3u}$ $(\pi \to \pi^*)$	6.06	6.06	4.87	5.16	5.13	6.12	6.16	6.18	5.53	5.53
	$2\,^1B_{2u}$ $(\pi \to \pi^*)$	6.33	6.33	4.70	5.05	5.06	6.22	6.23	6.28	5.96	5.96
	$2\,^1B_{1g}$ $(\pi \to \pi^*)$	6.47	6.47	4.80	5.17	5.17	6.29	6.24	6.31	6.39	6.39
	$3\,^1A_g$ $(\pi \to \pi^*)$	6.67	6.67	3.85	4.36	4.44	5.94	6.03	6.05	6.79	6.79
	$3\,^1B_{3u}$ $(\pi \to \pi^*)$	7.74		4.27	4.84	4.95	6.68	6.80	6.83	7.34	7.34
	$3\,^1B_{2u}$ $(\pi \to \pi^*)$	8.17		6.09	6.57	6.57	8.09	8.12	8.18	7.85	7.82
furan	$1\,^1B_2$ $(\pi \to \pi^*)$	6.39	6.32	4.59	4.87	4.87	5.78	5.82	5.88	5.90	5.68
	$2\,^1A_1$ $(\pi \to \pi^*)$	6.50	6.57	3.53	3.96	4.04	5.39	5.43	5.51	5.81	5.74
	$3\,^1A_1$ $(\pi \to \pi^*)$	8.17	8.13	5.72	6.15	6.12	7.44	7.47	7.62	7.88	7.23
pyrrole	$2\,^1A_1$ $(\pi \to \pi^*)$	6.31	6.37	3.37	3.77	3.79	5.21	5.28	5.29	5.38	5.38
	$1\,^1B_2$ $(\pi \to \pi^*)$	6.57	6.57	4.42	4.65	4.55	5.77	5.86	5.94	5.16	5.16
	$3\,^1A_1$ $(\pi \to \pi^*)$	8.17	7.91	5.31	5.65	5.53	7.10	7.18	7.16	6.57	6.57
imidazole	$2\,^1A'$ $(\pi \to \pi^*)$	6.19	6.19	4.05	4.32	4.11	5.50	5.59	5.85	5.00	5.00
	$1\,^1A''$ $(n \to \pi^*)$	6.81	6.81	5.25	5.24	4.48	5.87	6.00	6.08	5.36	5.36
	$3\,^1A'$ $(\pi \to \pi^*)$	6.93	6.93	4.62	4.84	4.70	5.95	6.04	6.16	5.65	5.65
	$2\,^1A''$ $(n \to \pi^*)$	7.90		5.82	5.83	5.05	6.81	6.79	6.70	6.19	6.19
	$4\,^1A'$ $(\pi \to \pi^*)$	8.16		5.91	6.12	5.71	7.40	7.45	7.69	6.87	6.87
pyridine	$1\,^1B_2$ $(\pi \to \pi^*)$	5.02	4.85	3.01	3.38	3.35	4.56	4.65	4.83	4.76	4.76
	$1\,^1B_1$ $(n \to \pi^*)$	5.17	4.59	4.36	4.29	3.75	4.85	4.85	4.86	4.40	4.40
	$1\,^1A_2$ $(n \to \pi^*)$	5.51	5.11	4.43	4.34	3.96	5.17	5.06	4.84	5.42	5.42
	$2\,^1A_1$ $(\pi \to \pi^*)$	6.39	6.26	4.65	5.01	5.00	6.14	6.11	6.25	6.00	6.00
	$2\,^1B_2$ $(\pi \to \pi^*)$	7.27	7.27	5.85	6.21	6.03	7.39	7.48	7.62	6.75	6.75
	$3\,^1A_1$ $(\pi \to \pi^*)$	7.46	7.18	5.88	6.29	6.11	7.44	7.43	7.63	6.65	6.65
	$3\,^1B_2$ $(\pi \to \pi^*)$	8.60		4.76	5.33	5.32	7.09	7.20	7.44		
	$4\,^1A_1$ $(\pi \to \pi^*)$	8.69		5.31	5.82	5.63	7.53	7.69	8.09		
pyrazine	$1\,^1B_{3u}$ $(n \to \pi^*)$	4.21	3.95	3.77	3.56	3.29	3.90	3.81	4.04	3.75	3.75
	$1\,^1A_u$ $(n \to \pi^*)$	4.70	4.81	3.75	3.55	3.50	4.38	4.12	3.89	5.08	5.08
	$1\,^1B_{2u}$ $(\pi \to \pi^*)$	4.85	4.64	3.37	3.65	3.48	4.61	4.76	5.20	4.61	4.61
	$1\,^1B_{2g}$ $(n \to \pi^*)$	5.68	5.56	4.86	4.90	4.13	5.49	5.78	5.86	4.85	4.85
	$1\,^1B_{1g}$ $(n \to \pi^*)$	6.41	6.60	5.23	5.36	4.51	6.34	6.54	6.32	6.89	6.89
	$1\,^1B_{1u}$ $(\pi \to \pi^*)$	6.89	6.58	4.75	5.10	5.06	6.24	6.22	6.35	6.20	6.20
	$2\,^1B_{2u}$ $(\pi \to \pi^*)$	7.66	7.60	6.45	6.66	6.23	7.64	7.67	8.12	7.64	7.64
	$2\,^1B_{1u}$ $(\pi \to \pi^*)$	7.79	7.72	6.92	7.15	6.48	7.98	8.06	8.68	7.69	7.69
	$1\,^1B_{3g}$ $(\pi \to \pi^*)$	8.47		4.94	5.46	5.39	7.16	7.35	7.68	7.68	7.68
	$2\,^1A_g$ $(\pi \to \pi^*)$	8.61		6.04	6.46	6.00	7.89	8.07	8.95	9.31	9.31
pyrimidine	$1\,^1B_1$ $(n \to \pi^*)$	4.44	4.55	3.89	3.78	3.46	4.43	4.34	4.38	4.16	4.16
	$1\,^1A_2$ $(n \to \pi^*)$	4.80	4.91	4.03	3.90	3.66	4.73	4.54	4.40	4.50	4.50
	$1\,^1B_2$ $(\pi \to \pi^*)$	5.24	5.44	3.42	3.71	3.55	4.77	4.86	5.22	5.00	5.00
	$2\,^1A_1$ $(\pi \to \pi^*)$	6.63	6.95	4.92	5.29	5.14	6.37	6.36	6.62	6.39	6.39
	$3\,^1A_1$ $(\pi \to \pi^*)$	7.21		6.43	6.71	6.27	7.76	7.81	8.25	7.34	7.34
	$2\,^1B_2$ $(\pi \to \pi^*)$	7.64		6.37	6.72	6.33	7.63	7.72	8.00	6.92	6.92
	$3\,^1B_2$ $(\pi \to \pi^*)$	8.73		5.93	6.29	5.91	7.92	8.01	8.77	8.27	8.27
	$4\,^1A_1$ $(\pi \to \pi^*)$	9.19		5.33	5.79	5.61	7.37	7.51	8.00	8.13	8.13
pyridazine	$1\,^1B_1$ $(n \to \pi^*)$	3.78	3.78	4.10	4.13	3.15	4.01	4.37	4.14	3.79	3.79
	$1\,^1A_2$ $(n \to \pi^*)$	4.31	4.31	4.29	4.35	3.41	4.39	4.70	4.21	4.66	4.66
	$2\,^1A_1$ $(\pi \to \pi^*)$	5.18	5.18	3.25	3.56	3.41	4.63	4.74	5.08	4.95	4.95
	$2\,^1A_2$ $(n \to \pi^*)$	5.77	5.77	4.89	4.80	4.16	5.40	5.38	5.57	5.68	5.68
	$1\,^1B_2$ $(\pi \to \pi^*)$	6.13		4.97	5.32	5.09	6.34	6.29	6.60	6.30	6.30
	$2\,^1B_1$ $(n \to \pi^*)$	6.52		5.08	5.00	4.48	5.78	5.62	5.51	6.34	6.34
	$2\,^1B_2$ $(\pi \to \pi^*)$	7.29		6.01	6.32	6.03	7.29	7.32	7.76	7.04	7.04
	$3\,^1A_1$ $(\pi \to \pi^*)$	7.62		6.09	6.47	6.21	7.59	7.51	7.81	7.32	7.32
s-triazine	$1\,^1A_1''$ $(n \to \pi^*)$	4.60	4.60	4.22	4.00	3.74	4.80	4.51	4.57	4.74	4.74
	$1\,^1A_2''$ $(n \to \pi^*)$	4.66	4.66	3.94	3.76	3.53	4.48	4.24	4.28	4.61	4.61
	$1\,^1E''$ $(n \to \pi^*)$	4.70	4.70	4.08	3.89	3.61	4.66	4.40	4.45	4.44	4.44
	$1\,^1A_2'$ $(\pi \to \pi^*)$	5.79	5.79	3.94	4.13	3.79	5.06	5.12	5.69	5.45	5.45
	$2\,^1A_1'$ $(\pi \to \pi^*)$	7.25		5.51	5.80	5.34	6.74	6.78	7.33	6.90	6.90
	$1\,^1E'$ $(\pi \to \pi^*)$	7.50		6.93	7.16	6.56	7.72	7.88	8.47	7.57	7.57
	$2\,^1E''$ $(n \to \pi^*)$	7.71		6.13	6.00	5.73	7.48	7.01	7.17	7.00	7.00
	$2\,^1E'$ $(\pi \to \pi^*)$	8.99		6.26	6.54	6.04	8.26	8.28	9.03	8.95	8.95

Table 3. Continued

molecule	state	CASPT2[a]	TBE[b]	MNDO	AM1	PM3	OM1	OM2	OM3	INDO/S	INDO/S2
s-tetrazine	1^1B_{3u} (n → π*)	2.29	2.29	2.98	2.91	2.27	2.66	2.83	2.88	2.86	2.86
	1^1A_u (n → π*)	3.51	3.51	3.80	3.65	2.99	3.53	3.55	3.08	4.43	4.43
	1^1B_{1g} (n → π*)	4.73	4.73	5.09	5.58	3.69	5.08	6.15	6.49	4.38	4.38
	1^1B_{2u} (π → π*)	4.93	4.93	3.62	3.84	3.50	4.71	4.88	5.74	4.84	4.84
	1^1B_{2g} (n → π*)	5.20	5.20	4.42	4.41	3.87	5.02	5.33	6.11	4.94	4.94
	2^1A_u (n → π*)	5.50	5.50	4.99	4.78	3.80	4.89	4.65	5.06	5.59	5.59
	1^1B_{3g} (n → π*)[e]	5.86		6.54	6.41	5.07	5.93	6.28	5.88		
	2^1B_{2g} (n → π*)	6.06		5.82	6.02	4.42	5.97	6.73	6.64	6.64	6.64
	2^1B_{1g} (n → π*)	6.45		5.29	5.12	4.47	5.63	5.65	5.60	6.59	6.59
	3^1B_{1g} (n → π*)	6.73		5.56	5.39	4.92	6.29	6.39	6.72	7.64	7.64
	2^1B_{3u} (n → π*)	6.77		5.68	5.38	4.53	5.66	5.22	5.08	7.29	7.29
	1^1B_{1u} (π → π*)	6.94		5.70	5.91	5.29	6.75	6.78	7.23	6.64	6.64
	2^1B_{1u} (π → π*)	7.42		6.68	6.84	6.16	7.67	7.86	9.03	7.27	7.27
	2^1B_{2u} (π → π*)	8.14		7.16	7.32	6.63	8.24	8.16	9.04	8.18	8.18
	2^1B_{3g} (π → π*)	8.34		5.91	6.22	5.55	7.54	7.86	9.62	9.30	9.30
formaldehyde	1^1A_2 (n → π*)	3.99	3.88	3.21	3.07	2.87	3.71	3.55	3.59	3.62	4.09
	1^1B_1 (σ → π*)	9.14	9.10	8.46	8.68	8.63	9.30	7.93	9.01	10.95	11.36
	2^1A_1 (π → π*)	9.32	9.30	8.56	9.09	8.40	9.61	9.23	9.89	12.09	13.54
acetone	1^1A_2 (n → π*)	4.44	4.40	3.18	3.46	3.29	3.80	3.98	4.05	3.67	4.13
	1^1B_1 (σ → π*)	9.27	9.10	8.26	8.18	7.46	8.81	7.71	8.34	10.63	11.17
	2^1A_1 (π → π*)	9.31	9.40	7.97	7.86	8.28	8.33	8.08	8.51	10.87	11.79
p-benzoquinone	1^1B_{1g} (n → π*)	2.76	2.76	2.79	2.88	2.72	2.62	2.64	2.58	2.67	3.00
	1^1A_u (n → π*)	2.77	2.77	2.95	3.19	2.96	3.17	3.35	3.37	2.64	3.00
	1^1B_{3g} (π → π*)	4.26	4.26	4.27	4.43	4.43	4.82	4.62	4.68	4.75	4.79
	1^1B_{1u} (π → π*)	5.28	5.28	5.26	5.47	5.26	5.64	5.52	5.71	5.60	5.83
	1^1B_{3u} (n → π*)	5.64	5.64	4.42	4.70	4.68	5.24	5.34	5.31	5.56	5.59
	2^1B_{3g} (π → π*)	6.96	6.96	5.36	5.80	5.65	6.58	6.73	6.93	6.69	6.93
	2^1B_{1u} (π → π*)	7.92		6.28	6.66	6.65	7.81	7.79	7.87	7.41	7.45
formamide	$1^1A''$ (n → π*)	5.63	5.63	3.89	4.11	3.68	4.61	4.56	4.82	4.44	5.09
	$2^1A'$ (π → π*)	7.39	7.39	5.90	5.93	5.08	6.92	6.71	7.07	7.38	7.64
	$3^1A'$ (π → π*)	10.54		8.26	8.76	7.95	9.67	9.41	10.11	12.23	13.54
acetamide	$1^1A''$ (n → π*)	5.69	5.69	3.83	4.23	3.79	4.59	4.75	4.98	4.36	5.00
	$2^1A'$ (π → π*)	7.27	7.27	5.70	5.77	4.92	6.76	6.63	6.95	7.39	7.64
	$3^1A'$ (π → π*)	10.09		7.82	8.09	7.49	8.93	8.64	9.02	11.31	12.19
propanamide	$1^1A''$ (n → π*)	5.72	5.72	3.91	4.35	3.87	4.67	4.85	5.06	4.35	4.99
	$2^1A'$ (π → π*)	7.20	7.20	5.69	5.77	4.92	6.78	6.64	6.94	7.46	7.71
	$3^1A'$ (π → π*)	9.94		7.77	7.90	7.41	8.72	8.34	8.50	10.82	11.18
cytosine	$2^1A'$ (π → π*)	4.67	4.66	3.34	3.47	3.12	4.19	4.21	4.39	4.41	4.50
	$1^1A''$ (n → π*)	5.12	4.87	3.63	3.76	3.36	4.19	4.23	4.40	4.10	4.15
	$2^1A''$ (n → π*)	5.53	5.26	4.05	4.32	3.83	4.76	4.83	4.95	4.73	5.33
	$3^1A'$ (π → π*)	5.53	5.62	3.87	4.00	3.43	5.01	5.00	5.05	5.54	5.58
	$4^1A'$ (π → π*)	6.40		4.66	4.72	4.18	5.88	5.78	5.84	6.14	6.26
	$5^1A'$ (π → π*)	6.97		4.81	4.94	4.44	6.14	5.91	5.96	6.57	6.61
thymine	$1^1A''$ (n → π*)	4.95	4.82	3.77	4.18	3.63	4.34	4.52	4.68	4.05	4.63
	$2^1A'$ (π → π*)	5.06	5.20	4.02	4.09	3.63	4.97	4.81	4.81	5.07	5.17
	$3^1A'$ (π → π*)	6.15	6.27	4.75	4.97	4.23	5.78	5.56	5.65	6.04	6.22
	$2^1A''$ (n → π*)	6.38	6.16	4.42	4.96	4.21	5.32	5.47	5.69	4.94	5.72
	$4^1A'$ (π → π*)	6.53	6.53	4.93	4.85	4.37	5.91	5.73	5.90	6.65	6.94
	$3^1A''$ (n → π*)	6.85		4.95	5.41	4.91	6.07	6.08	6.08	6.50	6.78
	$5^1A'$ (π → π*)	7.43		5.67	5.78	5.08	6.71	6.50	6.73	7.35	7.67
	$4^1A''$ (n → π*)	7.43		5.62	6.01	5.24	6.48	6.36	6.45	7.05	7.50
uracil	$1^1A''$ (n → π*)	4.91	4.80	3.73	4.12	3.61	4.27	4.45	4.64	4.06	4.64
	$2^1A'$ (π → π*)	5.23	5.35	4.09	4.15	3.66	5.05	4.88	4.90	5.19	5.29
	$3^1A'$ (π → π*)	6.15	6.26	4.80	4.96	4.24	5.78	5.68	5.86	6.31	6.51
	$2^1A''$ (n → π*)	6.28	6.10	4.42	4.95	4.21	5.29	5.43	5.65	4.93	5.70
	$4^1A'$ (π → π*)	6.74	6.70	4.99	5.02	4.47	6.14	5.85	5.97	6.58	6.87
	$3^1A''$ (n → π*)	6.98	6.56	5.00	5.43	4.94	6.09	6.09	6.10	6.47	6.74
	$4^1A''$ (n → π*)	7.28		5.65	6.00	5.24	6.47	6.34	6.44	7.01	7.42
	$5^1A'$ (π → π*)	7.42		5.70	5.78	5.11	6.72	6.50	6.75	7.20	7.42
adenine	$1^1A''$ (n → π*)	5.19	5.12	4.08	3.99	3.81	4.81	4.60	4.62	4.55	4.55
	$2^1A'$ (π → π*)	5.20	5.25	3.08	3.21	2.95	4.19	4.23	4.33	4.33	4.33
	$3^1A'$ (π → π*)	5.29	5.25	3.73	3.89	3.54	4.83	4.79	4.90	4.61	4.61
	$2^1A''$ (n → π*)	5.96	5.75	4.36	4.32	4.09	5.10	5.05	5.16	4.69	4.69
	$4^1A'$ (π → π*)	6.34		4.21	4.38	4.03	5.61	6.03	5.80	5.87	5.87
	$5^1A'$ (π → π*)	6.64		4.62	4.75	4.17	6.00	5.91	6.14	5.36	5.36
	$6^1A'$ (π → π*)	6.87		4.79	4.92	4.45	6.02	6.50	6.27	6.17	6.17
	$7^1A'$ (π → π*)	7.56		5.14	5.25	4.69	6.49	6.64	6.74		

[a] SA-CASSCF/MS-CASPT2 results using the TZVP basis and MP2/6-31G* ground-state equilibrium geometries.[14] [b] Theoretical best estimates for vertical excitation energies. See ref 14 for details. [c] Including the doubly excited CSF (HOMO→LUMO) changes the energy of the $1^1B_u/2^1A_g$ states to 6.37/6.61 eV for butadiene, 5.14/5.29 eV for hexatriene, and 4.41/4.85 eV for octatetraene. [d] Using all orbitals in the active space yields 5.80/6.87 eV for $1^1B_1/1^1B_2$. [e] Double excitation.

Cyclopentadiene has three valence excited singlet states, the lowest one being of B₂ symmetry (HOMO → LUMO transition) followed by two A₁ states (composed of combinations of HOMO − 1 → LUMO and HOMO → LUMO + 1 single excitations as well as double excitations). The semiempirical calculations again understimate the excitation energies, for example, with OM*x* by 0.41−0.46 eV for 1^1B_2 (TBE 5.55 eV) and by 0.71−0.79

eV for 2^1A_1 (TBE 6.31 eV); similar deviations are found for the corresponding triplet states. The INDO/S results are reasonable for the singlet states, but much too low for the triplet states (by 1.2−1.6 eV).

Norbornadiene with its two nonconjugated double bonds seems to be described reasonably well by all six NDDO-based methods, with the energies of the two lowest singlet excited states (1^1A_2, TBE 5.34 eV; 1^1B_2, TBE 6.11 eV) being

Electronically Excited States

J. Chem. Theory Comput., Vol. 6, No. 5, 2010 **1553**

Table 4. Vertical Triplet Excitation Energies ΔE (eV) of All Evaluated Molecules As Compared to MS-CASPT2/TZVP Results and Theoretical Best Estimates (TBE)

molecule	state	CASPT2[a]	TBE[b]	MNDO	AM1	PM3	OM1	OM2	OM3	INDO/S	INDO/S2
ethene	1^3B_{1u} ($\pi \to \pi^*$)	4.48	4.50	2.58	2.98	3.05	4.07	4.14	4.16	3.23	3.23
E-butadiene	1^3B_u ($\pi \to \pi^*$)	3.34	3.20	1.94	2.21	2.28	2.99	3.04	3.04	2.24	2.24
	1^3A_g ($\pi \to \pi^*$)	5.16	5.08	2.82	3.25	3.32	4.49	4.56	4.59	3.71	3.71
all-E-hexatriene	1^3B_u ($\pi \to \pi^*$)	2.71	2.40	1.61	1.81	1.88	2.41	2.46	2.45	2.65	2.65
	1^3A_g ($\pi \to \pi^*$)	4.31	4.15	2.40	2.75	2.81	3.73	3.80	3.81	4.39	4.39
all-E-octatetraene	1^3B_u ($\pi \to \pi^*$)	2.33	2.20	1.44	1.59	1.66	2.09	2.12	2.11	2.20	2.20
	1^3A_g ($\pi \to \pi^*$)	3.70	3.55	2.10	2.37	2.43	3.18	3.23	3.24	3.32	3.32
cyclopropene	1^3B_2 ($\pi \to \pi^*$)	4.35	4.34	2.53	2.68	2.76	3.77	3.80	3.72	3.18	3.18
	1^3B_1 ($\sigma \to \pi^*$)	6.51	6.62	4.89	5.02	5.23	6.01	5.48	5.65	7.25	7.25
cyclopentadiene	1^3B_2 ($\pi \to \pi^*$)	3.28	3.25	1.80	2.07	2.18	2.81	2.87	2.86	2.02	2.02
	1^3A_1 ($\pi \to \pi^*$)	5.11	5.09	2.67	3.05	3.14	4.23	4.30	4.31	3.46	3.46
norbornadiene	1^3A_2 ($\pi \to \pi^*$)	3.75	3.72	2.52	2.88	2.95	4.08	4.27	4.26	2.75	2.75
	1^3B_2 ($\pi \to \pi^*$)	4.22	4.16	2.45	2.81	2.84	3.89	4.10	4.07	3.15	3.15
benzene	1^3B_{1u} ($\pi \to \pi^*$)	4.17	4.15	2.13	2.50	2.56	3.66	3.74	3.76	3.72	3.72
	1^3E_{1u} ($\pi \to \pi^*$)	4.90	4.86	2.85	3.25	3.31	4.48	4.54	4.57	4.83	4.83
	1^3B_{2u} ($\pi \to \pi^*$)	5.76	5.88	4.42	4.73	4.72	5.79	5.80	5.85	5.46	5.46
	1^3E_{2g} ($\pi \to \pi^*$)	7.38	7.51	3.85	4.45	4.53	6.20	6.30	6.34	6.97	6.97
naphthalene	1^3B_{2u} ($\pi \to \pi^*$)	3.16	3.11	1.76	2.01	2.07	2.84	2.89	2.90	2.95	2.95
	1^3B_{3u} ($\pi \to \pi^*$)	4.25	4.18	2.50	2.83	2.88	3.87	3.91	3.94	4.15	4.15
	1^3B_{1g} ($\pi \to \pi^*$)	4.51	4.47	2.51	2.86	2.93	4.00	4.07	4.09	4.20	4.20
	2^3B_{2u} ($\pi \to \pi^*$)	4.68	4.64	2.65	3.02	3.07	4.22	4.28	4.31	4.61	4.61
	2^3B_{3u} ($\pi \to \pi^*$)	4.97	5.11	3.91	4.09	4.09	4.95	4.95	4.99	4.61	4.61
	1^3A_g ($\pi \to \pi^*$)	5.53	5.52	3.09	3.51	3.57	4.82	4.90	4.93	5.12	5.12
	2^3B_{1g} ($\pi \to \pi^*$)	6.21	6.48	4.98	5.33	5.32	6.46	6.49	6.52	7.33	7.33
	2^3A_g ($\pi \to \pi^*$)	6.38	6.47	4.84	5.23	5.23	6.50	6.51	6.58	6.39	6.39
	3^3A_g ($\pi \to \pi^*$)	6.59	6.79	3.61	4.11	4.19	5.68	5.76	5.79	7.32	7.32
	3^3B_{1g} ($\pi \to \pi^*$)	6.64	6.76	3.66	4.17	4.23	5.71	5.79	5.83	6.63	6.63
furan	1^3B_2 ($\pi \to \pi^*$)	4.18	4.17	2.18	2.47	2.50	3.40	3.50	3.53	2.79	2.95
	1^3A_1 ($\pi \to \pi^*$)	5.49	5.48	2.85	3.22	3.29	4.44	4.54	4.59	3.97	4.00
pyrrole	1^3B_2 ($\pi \to \pi^*$)	4.51	4.48	2.26	2.58	2.76	3.55	3.76	3.89	2.47	2.47
	1^3A_1 ($\pi \to \pi^*$)	5.52	5.51	2.88	3.23	3.23	4.46	4.59	4.64	3.81	3.81
imidazole	$1^3A'$ ($\pi \to \pi^*$)	4.65	4.69	2.46	2.77	2.89	3.70	3.95	4.06	3.04	3.04
	$2^3A'$ ($\pi \to \pi^*$)	5.74	5.79	3.59	3.80	3.60	4.77	4.95	5.27	4.32	4.32
	$1^3A''$ (n $\to \pi^*$)	6.36	6.37	4.89	4.84	4.14	5.43	5.60	5.70	5.42	5.42
	$3^3A'$ ($\pi \to \pi^*$)	6.44	6.55	4.49	4.60	4.12	5.79	5.69	5.99	5.59	5.59
	$4^3A'$ ($\pi \to \pi^*$)	7.44		4.87	4.95	4.48	6.28	6.20	6.37	7.12	7.12
	$2^3A''$ (n $\to \pi^*$)	7.51		5.50	5.48	4.83	6.50	6.47	6.33	7.48	7.48
pyridine	1^3A_1 ($\pi \to \pi^*$)	4.27	4.06	2.31	2.65	2.68	3.78	3.86	3.94	3.80	3.80
	1^3B_1 (n $\to \pi^*$)	4.57	4.25	3.96	3.86	3.37	4.37	4.48	4.47	4.02	4.02
	1^3B_2 ($\pi \to \pi^*$)	4.71	4.64	3.08	3.46	3.44	4.57	4.66	4.83	4.68	4.68
	2^3A_1 ($\pi \to \pi^*$)	5.03	4.91	3.22	3.58	3.50	4.67	4.74	4.97	4.93	4.93
	1^3A_2 (n $\to \pi^*$)	5.52	5.28	4.36	4.27	3.89	5.11	4.96	4.92	5.84	5.84
	2^3B_2 ($\pi \to \pi^*$)	6.03	6.08	4.79	5.08	4.92	6.02	5.97	6.17	6.02	6.02
	3^3A_1 ($\pi \to \pi^*$)	7.56		4.61	5.05	4.87	6.56	6.67	7.08		
	3^3B_2 ($\pi \to \pi^*$)	7.87		4.25	4.73	4.68	6.42	6.59	6.83		
s-tetrazine	1^3B_{3u} (n $\to \pi^*$)	1.61	1.89	2.43	2.36	1.88	2.08	2.35	2.36	2.30	2.30
	1^3A_u (n $\to \pi^*$)	3.28	3.52	3.55	3.38	2.80	3.31	3.31	2.87	4.32	4.32
	1^3B_{1g} (n $\to \pi^*$)	4.14	4.21	4.23	4.30	3.14	4.12	4.85	4.78	3.72	3.72
	1^3B_{1u} ($\pi \to \pi^*$)	4.37	4.33	2.95	3.11	2.78	3.90	4.07	4.74	3.54	3.54
	1^3B_{2u} ($\pi \to \pi^*$)	4.39	4.54	3.53	3.78	3.49	4.57	4.76	5.62	3.98	3.98
	1^3B_{2g} (n $\to \pi^*$)	4.94	4.93	4.15	4.13	3.62	4.73	5.04	5.66	4.91	4.91
	2^3A_u (n $\to \pi^*$)	5.04	5.03	4.69	4.46	3.67	4.67	4.40	4.79	5.89	5.89
	2^3B_{1u} ($\pi \to \pi^*$)	5.40	5.38	3.92	4.11	3.77	4.96	5.07	6.00	5.27	5.27
	1^3B_{3g} (n $\to \pi^*$)	5.57		5.24	5.53	4.86	5.75	6.00	5.53		
	2^3B_{2g} (n $\to \pi^*$)	5.97		5.65	5.83	4.32	5.84	6.50	6.43	7.05	7.05
	2^3B_{1g} (n $\to \pi^*$)	6.37		5.43	5.38	4.85	6.16	6.29	6.65	7.37	7.37
	2^3B_{3u} (n $\to \pi^*$)	6.54		5.55	5.24	4.35	5.49	5.09	4.95	7.87	7.87
	2^3B_{2u} ($\pi \to \pi^*$)	7.08		5.76	5.87	5.22	6.69	6.73	7.43	7.23	7.23
formaldehyde	1^3A_2 ($\pi \to \pi^*$)	3.58	3.50	2.92	2.74	2.57	3.40	3.23	3.24	3.14	3.68
	1^3A_1 ($\pi \to \pi^*$)	5.84	5.87	4.90	5.42	5.07	5.67	5.63	6.07	6.41	8.04
acetone	1^3A_2 (n $\to \pi^*$)	4.10	4.05	2.87	3.18	3.05	3.53	3.74	3.79	4.02	4.54
	1^3A_1 ($\pi \to \pi^*$)	6.04	6.03	4.50	4.97	4.65	5.33	5.45	5.79	8.52	10.33
p-benzoquinone	1^3B_{1g} (n $\to \pi^*$)	2.62	2.51	2.61	2.71	2.57	2.46	2.50	2.42	2.74	3.09
	1^3A_u (n $\to \pi^*$)	2.66	2.62	2.76	3.01	2.82	3.03	3.21	3.23	2.81	3.24
	1^3B_{1u} ($\pi \to \pi^*$)	2.99	2.96	2.18	2.45	2.41	2.75	2.79	2.91	3.48	3.80
	1^3B_{3g} ($\pi \to \pi^*$)	3.32	3.41	2.33	2.63	2.68	3.38	3.32	3.34	3.49	3.64
formamide	$1^3A''$ (n $\to \pi^*$)	5.40	5.36	3.68	3.87	3.49	4.39	4.34	4.58	4.00	4.69
	$1^3A'$ ($\pi \to \pi^*$)	5.58	5.74	4.02	4.16	3.54	4.89	4.77	4.98	5.21	5.93
acetamide	$1^3A''$ (n $\to \pi^*$)	5.41	5.42	3.61	4.00	3.62	4.38	4.54	4.76	4.99	5.69
	$1^3A'$ ($\pi \to \pi^*$)	5.63	5.88	3.86	4.11	3.49	4.86	4.86	5.07	6.61	7.31
propanamide	$1^3A''$ (n $\to \pi^*$)	5.45	5.45	3.68	4.11	3.69	4.44	4.61	4.81	5.00	5.70
	$1^3A'$ ($\pi \to \pi^*$)	5.80	5.90	3.87	4.12	3.50	4.87	4.87	5.07	6.61	7.30

[a] SA-CASSCF/MS-CASPT2 results using the TZVP basis and MP2/6-31G* ground-state equilibrium geometries.[14] [b] Theoretical best estimates for vertical excitation energies. See ref 14 for details.

bracketed by the semiempirical results (with individual deviations of a few tenths of an eV). Concerning the two lowest singlet B$_2$ states, MNDO/AM1/PM3 predicts a significantly higher oscillator strength for the lower one, while the OMx and ab initio calculations give the opposite trend, which is indicative of differences in the corresponding CI wave functions. The energy gap between these two singlet B$_2$ states is rather low in MNDO/AM1/PM3 (0.38−0.43 eV) as compared to OMx (0.77−1.03 eV) and the MS-CASPT2 reference value (1.21 eV).[14]

The electronic spectrum of benzene has often been studied theoretically and has served as a prototypical test case for many computational methods. There are four singlet valence $\pi\pi^*$ states in the 5–9 eV range. MNDO/AM1/PM3 underestimate their energies severely, by 1.2–3.9 eV. The OMx methods predict the first two singlet states (1^1B_{2u}, TBE 5.08 eV; 1^1B_{1u}, TBE 6.54 eV) too low by 0.51–0.67 eV, get the correct energy within 0.07 eV for the bright 1^1E_{1u} state (TBE 7.13 eV), and strongly underestimate the energy of the E_{2g} state (TBE 8.41 eV), which has significant double-excitation character, by 1.2–1.3 eV. The standard INDO/S CIS results for these $\pi\pi^*$ states suffer from $\sigma\sigma^*$ contamination, which reduces most of the excitation energies by about 0.5 eV; after excluding the corresponding σ^* MOs from the active space, the results are reasonable, with the excited singlet states in the right order and deviations of 0.37–0.62 eV from the TBE values. For the four triplet $\pi\pi^*$ states, the situation is similar as for the singlets, with an analogous performance of the different methods, which does not warrant further discussion.

In the case of naphthalene, our benchmark set includes 10 excited singlet states and 10 triplet states, which are all of $\pi\pi^*$ type. MNDO/AM1/PM3 again underestimate the excitation energies severely in general and give a partially incorrect state ordering. The OMx results are much closer to the TBE reference values, with a tendency to be too low by a few tenths of an eV, and normally produce the correct state ordering; the largest deviations from the TBE values occur for A_g states with significant double-excitation character where the OMx excitation energies are typically too low by about 0.6–0.7 eV. The INDO/S results appear to be of overall quality similar to those from OMx; the INDO/S energies for the A_g states are closer to the TBE values, which should be considered fortuitous because the INDO/S CIS calculations do not include double excitations.

Furan, Pyrrole, Imidazole, Pyridine, Pyrazine, Pyrimidine, Pyridazine, s-Triazine, and s-Tetrazine. Furan and pyrrole are isoelectronic five-ring heterocycles, which both have three singlet and two triplet $\pi\pi^*$ valence excited states. All semiempirical methods underestimate the corresponding excitation energies. The deviations from the TBE values are largest for MNDO/AM1/PM3 (1.5–3.0 eV) and still substantial for OMx (0.4–1.2 eV) and INDO/S (0.4–2.0 eV). For the first two close-lying excited singlets, the six NDDO-based methods give the reverse order for furan (like INDO/S) and the correct order for pyrrole (unlike INDO/S), while the sequence of the triplet states is always predicted correctly. INDO/S2 differs from INDO/S only in the oxygen parametrization and thus yields slightly different results for furan (surprisingly with slightly larger deviations from the TBE data).

The imidazole spectrum contains $n\pi^*$ and $\pi\pi^*$ valence transitions in the range between 6.0 and 8.5 eV.[14] All semiempirical methods underestimate the $\pi\pi^*$ excitation energies in a manner similar to that in furan and pyrrole. The energies of the lowest $n\pi^*$ state in the singlet manifold ($1^1A''$, TBE 6.81 eV) and in the triplet manifold ($1^3A''$, TBE 6.37 eV) are underestimated to an extent similar to those of the $\pi\pi^*$ states (MNDO/AM1/PM3 by 1.6–2.3 eV, OMx by

0.7–0.9 eV, INDO/S by 1.0–1.5 eV), so that OMx and INDO/S give the correct order of the singlet and triplet states.

Pyridine is the first of a series of azabenzenes in our benchmark set. The introduction of one nitrogen atom lowers the symmetry and splits degenerate levels, such that the four valence $\pi\pi^*$ states in benzene correlate with six $\pi\pi^*$ states of symmetry A_1 and B_2 in pyridine. TBE values are available for the lowest four of these states both in the singlet case (4.85, 6.26, 7.18, and 7.27 eV) and in the triplet case (4.06, 4.64, 4.91, and 6.08 eV). The OMx results scatter around these reference data (typically within 0.3 eV or less), and the INDO/S results are of similar quality. There are two additional $n\pi^*$ states in the singlet manifold (1^1B_1, TBE 4.59 eV; 1^1A_2, TBE 5.11 eV) and in the triplet manifold (1^3B_1, TBE 4.25 eV; 1^3A_2, TBE 5.28 eV) whose energies are again well reproduced by OMx (typically within 0.3 eV) and also by INDO/S, even though the splitting of these two states is underestimated by OMx and overestimated by INDO/S. Overall, however, both the OMx methods and the INDO/S perform quite well for pyridine.

Similar remarks apply to the azabenzenes with two ring nitrogen atoms, that is, pyrazine, pyrimidine, and pyridazine, which are represented in our benchmark set only through their singlet excited states. The available TBE values for the eight lowest singlets in pyrazine and for the four lowest singlets in the other two molecules are generally well reproduced by the OMx methods, with typical deviations of about 0.3 eV both for $\pi\pi^*$ and $n\pi^*$ transitions and with a state ordering that is generally analogous to the one obtained from the ab initio reference calculations. The INDO/S results are generally in the same ballpark as the OMx results. Focusing on problem cases, we note that the deviations from the TBE values are larger than usual for the OMx energies of the 1^1A_u $n\pi^*$ state in pyrazine (TBE 4.81 eV, OMx too low by 0.43–0.92 eV) and for the INDO/S energy of the 1^1B_{2g} $n\pi^*$ state in pyrazine (TBE 5.56 eV, INDO/S too low by 0.71 eV); as a consequence, the splitting between these two states is overestimated by OMx and underestimated by INDO/S (where their order is actually inverted).

In the case of s-triazine, we focus on the four lowest singlet transitions for which TBE values have been derived. The three lowest singlet states are of $n\pi^*$ type and almost degenerate, lying within 0.1 eV (TBE 4.60–4.70 eV). The semiempirical calculations also yield three close-lying $n\pi^*$ states, typically within 0.3 eV, which appear in a similar energy range with OMx (4.3–4.8 eV) and INDO/S (4.4–4.7 eV), although their order is generally different from that found at the ab initio level. The energy of the lowest dark $\pi\pi^*$ singlet state ($1^1A_2'$, TBE 5.79 eV) is underestimated by all semiempirical methods (e.g., OMx by 0.10–0.73 eV, INDO/S by 0.34 eV). For the higher transitions above 7 eV, there is fair agreement with the available MS-CASPT2/TZVP data, with deviations roughly as expected in this energy range.

s-Tetrazine has a large number of known $n\pi^*$ and $\pi\pi^*$ states, with TBE values being available for the lowest six (eight) excited states in the singlet (triplet) manifold. Without going into detail, we note that the OMx and INDO/S methods perform about as well as anticipated from the experience

Electronically Excited States

J. Chem. Theory Comput., Vol. 6, No. 5, 2010 **1555**

from the other azabenzenes. Given the number of close-lying states, it is inevitable, however, that there are sometimes differences in the state orderings and outliers for some of the computed excitation energies, particularly in the case of OM3.

Formaldehyde, Acetone, *p*-Benzoquinone, Formamide, Acetamide, and Propanamide. The lowest excited states of formaldehyde are $n\pi^*$ transitions, 1^3A_2 (TBE 3.50 eV) and 1^1A_2 (TBE 3.88 eV). Their energies are underestimated slightly by OMx (by 0.10−0.33 eV) and INDO/S (by 0.26−0.36 eV), and more so by MNDO/AM1/PM3 (by 0.58−1.01 eV). For the $\pi\pi^*$ triplet state (1^3A_1, TBE 5.87 eV), the computed energies are too low for MNDO/AM1/PM3 (by 0.45−0.97 eV), while they scatter around the TBE value for OMx (within 0.2 eV) and are too high for INDO/S (by 0.54 eV) and especially for INDO/S2 (by 2.2 eV). The two remaining valence excited singlet states in our benchmark set lie above 9 eV, 1^1B_1 ($\sigma \rightarrow \pi^*$ type, TBE 9.10 eV) and 2^1A_1 ($\pi \rightarrow \pi^*$ type, TBE 9.30 eV). Their MS-CASPT2/TZVP wave functions indicate significant contributions from higher excitations in the 1^1B_1 state and some notable Rydberg-valence mixing in the 2^1A_1 state, which should only be partially recovered in semiempirical calculations; given this situation, the errors for the six NDDO-based methods of 0.2−1.2 eV are not excessive, while the INDO/S energies are too high by 1.9−2.8 eV.

The electronic spectrum of acetone is qualitatively similar to that of formaldehyde, and the performance of various semiempirical methods is also similar for both molecules, with a slight increase in the deviations from the TBE values for the lowest three states below 6 eV.

In *p*-benzoquinone, the two lowest singlet excited states are of $n\pi^*$ type. They are almost degenerate (1^1B_{1g}, TBE 2.76 eV; 1^1A_u, TBE 2.77 eV). All semiempirical methods reproduce the energy of the lowest state quite well (within 0.18 eV or less), but except for INDO/S, they compute the second state too high and thus give a sizable gap between these two dark $n\pi^*$ states (MNDO/AM1/PM3 0.16−0.31 eV, OMx 0.55−0.79 eV). The energy of the third $n\pi^*$ singlet state (1^1B_{3u}, TBE 5.64 eV) is underestimated to a different extent (MNDO/AM1/PM3 by 1.0−1.2 eV, OMx by 0.30−0.40 eV, INDO/S by 0.08 eV). The positions of the four $\pi\pi^*$ singlet states are given by OMx and INDO/S with the expected accuracy of typically 0.4 eV; for example, the first bright B_{1u} transition (TBE 5.28 eV) responsible for the first strong peak in the spectrum is calculated somewhat too high with OMx and INDO/S (by 0.24−0.43 eV). The results for the four lowest triplet states are largely analogous and will thus not be discussed in detail, except for noting that the OMx methods reproduce the TBE values to within 0.2 eV for all of these states except 1^3A_u (TBE 2.62 eV, OMx too large by 0.41−0.61 eV).

Formamide, acetamide, and propanamide have analogous valence excited states. The lowest excited singlet is an $n\pi^*$ state ($1^1A''$, TBE 5.63−5.72 eV) followed by a $\pi\pi^*$ state ($2^1A'$, TBE 7.20−7.39 eV) and another high-lying $\pi\pi^*$ state ($2^1A'$, around 10 eV or higher). The triplet manifold begins with a $n\pi^*$ state ($1^3A''$, TBE 5.36−5.45 eV) followed by a nearby $\pi\pi^*$ state ($1^3A'$, TBE 5.74−5.90 eV). The energies

of the $n\pi^*$ singlet and triplet states in these primary amides are generally underestimated by all semiempirical methods (e.g., in OMx by 0.64−1.10 eV and in INDO/S by 0.43−1.43 eV). Likewise, the energies of the $\pi\pi^*$ triplet state are underestimated in OMx by similar amounts (0.71−1.03 eV) such that the OMx triplet−triplet gaps are realistic, whereas INDO/S gives gaps that are too large. Finally, for the first $\pi\pi^*$ singlet state, the OMx and INDO/S results seem reasonable (OMx too low by 0.26−0.51 eV, INDO/S within 0.26 eV).

Cytosine, Thymine, Uracil, and Adenine. The singlet excited states of these nucleobases complete our benchmark set. We focus in the discussion on the OMx results because the MNDO/AM1/PM3 excitation energies are generally much too low as usual, whereas the INDO/S results are mostly in the same range as the OMx results but appear less regular as compared to the TBE values.

The valence excited states of cytosine consist of four $\pi\pi^*$ (A′) states and two $n\pi^*$ (A″) states, which are all dominated by singly excited configurations. The OMx energies of these states are consistently lower than the available TBE values, typically by about 0.4−0.5 eV (with individual deviations in the range of 0.27−0.68 eV). Because these deviations are fairly uniform, the same state ordering is obtained from OMx as from the ab initio reference calculations. There is one caveat, however: according to the TBE values, the lowest excited state of cytosine is a $\pi\pi^*$ state at 4.66 eV followed by a nearby $n\pi^*$ state at 4.87 eV, whereas these two states are essentially degenerate at the OMx level (within 0.02 eV).

Thymine and uracil share the same heterocyclic ring structure and differ only in one methyl substituent, and hence their excited states are similar in character and can be discussed together. There are four $\pi\pi^*$ (A′) and four $n\pi^*$ (A″) valence excited singlet states. According to the available TBE values, the state ordering is $1^1A'' < 2^1A' < 2^1A'' < 3^1A'$ in both molecules, with an $n\pi^*$ state being lowest. The same state ordering is found in the OMx calculations, which generally underestimate the excitation energies in thymine and uracil, by 0.14−0.53 eV for the lowest two states around 5 eV and by 0.40−0.85 eV for the remaining three or four states with TBE values between 6−7 eV.

In adenine, the three lowest singlet states are close in energy: the first $n\pi^*$ state ($1^1A''$, TBE 5.12 eV) is followed by two essentially degenerate $\pi\pi^*$ states ($2^1A'$ and $3^1A'$, TBE 5.25 eV) and a second $n\pi^*$ state ($2^1A''$, TBE 5.75 eV). The OMx energies are generally smaller than these TBE values, as in the case of the other nucleobases, but the deviations are less uniform. The main difference is that the two $\pi\pi^*$ states are not degenerate, but show a substantial split in OMx (by 0.56−0.64 eV). This changes the state ordering such that the lowest state in the OMx calculations is a $\pi\pi^*$ state ($2^1A'$) followed by the close-lying $1^1A''$ and $3^1A'$ states. The energy of the fourth excited state ($2^1A''$) is underestimated in OMx to a similar extent as in the other nucleobases (by 0.59−0.70 eV).

5. One-Electron Properties

Most of the benchmarking activities for electronically excited states in the literature address excitation energies, although

one-electron properties such as oscillator strengths and excited-state dipole moments could also serve as a sensitive probe for the quality of computational methods. However, as compared to excitation energies, these properties are known to converge more slowly upon basis set extension,[14,15,61] and we have therefore not yet derived corresponding theoretical best estimates for our benchmark set. Available reference data include a range of published ab initio results as well as MS-CASPT2/TZVP and CC2/TZVP values calculated for the benchmark molecules in our previous work.[14] These data can be used to evaluate the performance of our semiempirical results, which were obtained using the standard active spaces and CI treatments described in section 2 (CISDTQ for NDDO-based methods, CIS for INDO/S). Different choices for the active space and the CI treatment would affect the semiempirical results, of course, but these changes are usually rather minor for reasonable alternative choices.

5.1. Oscillator Strengths. Table S8 (Supporting Information) lists the computed oscillator strengths for all optically active states in our benchmark. It contains ab initio values previously collated from the literature,[14] published CASPT2 data from the Roos group, and our own results from MS-CASPT2/TZVP and CC2/TZVP[14] as well as semiempirical (MNDO, AM1, PM3, OM1, OM2, OM3, INDO/S) calculations. Generally speaking, there is broad qualitative agreement between the different sets of results, with a proper distinction between strong, medium, and weak transitions. For the low-lying excited states, the majority of the semiempirical oscillator strengths are comparable in magnitude to the MS-CASPT2 and CC2 results, while larger deviations are sometimes found for high-lying bright states where the ab initio reference data also often show some scatter. As expected, n → π* transitions are normally weak and thus have low oscillator strengths both at the ab initio and at the semiempirical level. We refrain from detailed individual comparisons at this point and provide a more quantitative statistical evaluation in section 6.2.

5.2. Dipole Moments. Table S9 (Supporting Information) lists ground-state as well as excited-state dipole moments obtained from published CASPT2 work in the Roos group and from our own MS-CASPT2/TZVP,[14] RI-CC2/TZVP, and semiempirical calculations. Coupled-cluster CC2 results were determined using unrelaxed densities with the RICC2[62−65] program of the TURBOMOLE package. As pointed out before,[15] the two sets of CASPT2-based results in Table S9 often differ appreciably, and the CC2 dipole moments correlate only roughly with those from the published CASPT2 and from our own MS-CASPT2/TZVP calculations (with correlation coefficients of 0.8341 and 0.8705, respectively). This scatter in the reference data calls for some caution in the assessment of the semiempirical results.

For all semiempirical methods, the computed ground-state dipole moments agree reasonably well with the ab initio results. The mean absolute deviations are around 0.3 D and thus of similar magnitude as in previous comparisons with experiment.[41] The situation is much less satisfactory for the excited-state dipole moments. The MNDO, AM1, and PM3

results are reasonable for low-lying n → π* states (e.g., in pyridine with deviations of less than 0.2 D), but they often also lie far away from the range spanned by the three sets of reference data (see above), especially for high-lying states where deviations up to 7 D are encountered (e.g., in uracil and thymine); in these latter cases, the composition of the excited-state wave function is different from the ab initio reference. The OM*x* results also scatter around the range of the reference values, but to a lesser extent. Again, there are low-lying states that are well described (e.g., in cyclopentadiene and pyrrole), while there are outliers for high-lying states especially of the nucleobases (e.g., in thymine with deviations reaching 4 D). Finally, for technical reasons, INDO/S and INDO/S2 dipole moments have been computed only for singlet states; they differ because of the use of different C_{kk} and D_{kk} parameters in the CM2 approach.[35] Again the INDO/S and INDO/S2 results deteriorate for the high-lying states (e.g., in thymine with deviations of 8−9 D from MS-CASPT2/TZVP). A more quantitative assessment of the semiempirical excited-state dipole moments is given in section 6.2.

6. Statistics

In the following, we present a detailed statistical evaluation of the computed vertical excitation energies and one-electron properties for all semiempirical methods studied. As reference data, we mainly use theoretical best estimates of vertical excitation energies and MS-CASPT2/TZVP oscillator strengths and dipole moments.

6.1. Vertical Excitation Energies. Table 5 summarizes statistical results for singlet excited states, both for the full set of 104 states and for three subsets (hydrocarbons, CHN compounds, CHO and CHNO compounds). The standard NDDO-based methods (MNDO, AM1, PM3) obviously underestimate the vertical excitation energies systematically and by a large margin, with mean absolute deviations (MAD) above 1.0 eV in each category. The overall MAD values are 1.35, 1.19, and 1.41 eV, respectively, with the worst performance being found in MNDO for hydrocarbons (MAD 1.68 eV) and in PM3 for CHN and oxygen-containing compounds (MAD 1.43−1.48 eV). The OM*x* methods (OM1, OM2, OM3) yield much better results: the vertical excitation energies are still mostly underestimated (overall on average by 0.34, 0.36, and 0.22 eV), with slightly higher mean absolute deviations (0.45, 0.50, and 0.45 eV). Looking at the three subsets separately, the OM*x* methods perform rather uniformly for hydrocarbons (MAD 0.40−0.42 eV) and also rather well for CHN compounds (MAD 0.38−0.46 eV), but somewhat larger errors occur for oxygen-containing compounds especially in OM1 (MAD 0.57 eV) and in OM2 (MAD 0.62 eV). Among the OM*x* methods, the OM3 errors appear to be most uniform, and the systematic underestimation of the energies seems least pronounced (with an overall mean error of −0.22 eV). However, one should not overemphasize this distinction from OM1 and OM2 because the three OM*x* methods perform quite similarly in general. Both INDO/S and its variant INDO/S2 with modified oxygen parameters yield an overall MAD value of 0.51 eV, comparable to OM2 (see above). They also tend to under-

Electronically Excited States

J. Chem. Theory Comput., Vol. 6, No. 5, 2010 **1557**

Table 5. Deviations of Semiempirical Vertical Excitation Energies (in eV) for Singlet States with Respect to Theoretical Best Estimates

	MNDO	AM1	PM3	OM1	OM2	OM3	INDO/S	INDO/S2
CH-Containing Molecules								
count[a]	25	25	25	25	25	25	25	25
mean	−1.68	−1.36	−1.29	−0.27	−0.29	−0.25	−0.20	−0.20
abs. mean	1.68	1.36	1.29	0.42	0.41	0.40	0.42	0.42
std. dev.	1.89	1.55	1.48	0.52	0.51	0.49	0.47	0.47
max. (+) dev.				0.04	0.76	0.66	0.72	0.81	0.81
max. (−) dev.	3.95	3.30	3.21	1.34	1.22	1.19	0.84	0.84
CHN-Containing Molecules								
count[a]	43	43	43	43	43	43	42	42
mean	−1.11	−1.01	−1.43	−0.27	−0.21	−0.09	−0.34	−0.34
abs. mean	1.22	1.13	1.43	0.38	0.45	0.46	0.48	0.48
std. dev.	1.41	1.27	1.53	0.49	0.54	0.57	0.63	0.63
max. (+) dev.	0.69	0.85		0.37	1.42	1.76	0.92	0.92
max. (−) dev.	3.00	2.60	2.58	1.16	1.09	1.08	1.45	1.45
CHO and CHNO-Containing Molecules								
count[a]	36	36	36	36	36	36	36	36
mean	−1.27	−1.09	−1.46	−0.47	−0.55	−0.35	−0.12	0.22
abs. mean	1.29	1.14	1.48	0.57	0.62	0.47	0.62	0.61
std. dev.	1.44	1.25	1.63	0.64	0.69	0.52	0.88	1.04
max. (+) dev.	0.18	0.42	0.19	0.56	0.58	0.60	2.79	4.24
max. (−) dev.	3.04	2.61	2.53	1.19	1.39	1.06	1.37	0.90
All Molecules								
count[a]	104	104	104	104	104	104	103	103
mean	−1.30	−1.12	−1.40	−0.34	−0.36	−0.22	−0.23	−0.11
abs. mean	1.35	1.19	1.41	0.45	0.50	0.45	0.51	0.51
std. dev.	1.55	1.34	1.55	0.55	0.59	0.54	0.70	0.77
max. (+) dev.	0.69	0.85	0.19	0.76	1.42	1.76	2.79	4.24
max. (−) dev.	3.95	3.30	3.21	1.34	1.39	1.19	1.45	1.45

[a] Total number of states considered.

Table 6. Deviations of Semiempirical Vertical Excitation Energies (in eV) for $n\pi^*$ and $\pi\pi^*$ Singlet States with Respect to Theoretical Best Estimates up to 6 eV

	MNDO	AM1	PM3	OM1	OM2	OM3	INDO/S	INDO/S2
$n\pi^*$								
count[a]	33	33	33	33	33	33	32	32
mean	−0.62	−0.61	−1.08	−0.26	−0.18	−0.14	−0.31	−0.14
abs. mean	0.77	0.80	1.10	0.38	0.46	0.43	0.48	0.35
std. dev.	0.92	0.89	1.20	0.50	0.56	0.54	0.61	0.44
max. (+) dev.	0.69	0.85	0.19	0.40	1.42	1.76	0.92	0.92
max. (−) dev.	1.86	1.52	1.95	1.19	1.07	0.92	1.37	1.06
$\pi\pi^*$								
count[a]	25	25	25	25	25	25	25	25
mean	−1.42	−1.18	−1.30	−0.27	−0.26	−0.11	−0.19	−0.17
abs. mean	1.42	1.20	1.32	0.44	0.40	0.39	0.38	0.38
std. dev.	1.59	1.34	1.48	0.50	0.46	0.45	0.45	0.46
max. (+) dev.	0.01	0.19	0.17	0.76	0.66	0.81	0.81	0.81
max. (−) dev.	2.67	2.19	2.30	1.06	1.02	0.92	0.92	0.92

[a] Total number of states considered.

estimate the vertical excitation energies (overall mean errors of −0.23 and −0.11 eV, respectively), but their results scatter more strongly than the OMx results (overall standard deviations of 0.70−0.77 eV as compared to 0.54−0.59 eV for OMx). Concerning the subsets, our evaluation confirms[66] that INDO/S tends to give somewhat higher errors for oxygen-containing compounds (MAD 0.62 eV), but the claimed improvement by the reparametrized INDO/S2 variant[35] is hardly seen in our data (MAD 0.61 eV).

The statistics in Table 5 include several singlet states above 6 eV for which TBE values are available. Because we expect in general that minimal-basis-set semiempirical calculations on valence excited states will become less appropriate the higher the energy, we have performed a second evaluation restricted to singlet states up to 6 eV, considering $\pi\pi^*$ and $n\pi^*$ states separately. The results are shown in Table 6. It is obvious that the restriction to energies below 6 eV improves the statistics especially for MNDO, AM1, and PM3, where the MAD values for $\pi\pi^*$ states remain in the 1.2−1.4 eV range while those for the $n\pi^*$ states amount to 0.8−1.1 eV. The improvement is less pronounced for the OMx and INDO/S methods, which treat the two types of excitation in a fairly balanced manner: for example, the MAD values in OMx range from 0.39−0.44 eV for $\pi\pi^*$ states and from 0.38−0.46 eV for $n\pi^*$ states; the corresponding values for INDO/S are 0.38 and 0.48 eV. Comparison of the INDO/S and INDO/S2 statistics indicates that the reparametrization for oxygen

Table 7. Deviations of Semiempirical Vertical Excitation Energies (in eV) for Triplet States with Respect to Theoretical Best Estimates

	MNDO	AM1	PM3	OM1	OM2	OM3	INDO/S	INDO/S2
CH-Containing Molecules								
count[a]	27	27	27	27	27	27	27	27
mean	−1.84	−1.50	−1.44	−0.42	−0.37	−0.35	−0.38	−0.38
abs. mean	1.84	1.50	1.44	0.45	0.42	0.41	0.57	0.57
std. dev.	1.96	1.61	1.55	0.55	0.54	0.52	0.74	0.74
max. (+) dev.				0.36	0.55	0.54	0.85	0.85
max. (−) dev.	3.66	3.06	2.98	1.31	1.21	1.17	1.63	1.63
CHN-Containing Molecules								
count[a]	20	20	20	20	20	20	20	20
mean	−1.24	−1.10	−1.44	−0.39	−0.27	−0.05	−0.43	−0.43
abs. mean	1.29	1.16	1.44	0.43	0.44	0.49	0.70	0.70
std. dev.	1.50	1.32	1.55	0.56	0.52	0.56	0.92	0.92
max. (+) dev.	0.54	0.47		0.19	0.64	1.08	0.86	0.86
max. (−) dev.	2.63	2.28	2.43	1.05	0.92	0.87	2.01	2.01
CHO and CHNO-Containing Molecules								
count[a]	16	16	16	16	16	16	16	16
mean	−1.35	−1.07	−1.34	−0.57	−0.53	−0.39	−0.04	0.60
abs. mean	1.38	1.15	1.37	0.62	0.60	0.49	0.72	1.02
std. dev.	1.54	1.29	1.56	0.73	0.70	0.57	0.96	1.44
max. (+) dev.	0.14	0.39	0.20	0.41	0.59	0.61	2.49	4.30
max. (−) dev.	2.63	2.26	2.40	1.04	1.03	0.89	1.54	1.48
All Molecules								
count[a]	63	63	63	63	63	63	63	63
mean	−1.52	−1.27	−1.41	−0.45	−0.38	−0.26	−0.31	−0.15
abs. mean	1.55	1.30	1.42	0.49	0.47	0.45	0.65	0.72
std. dev.	1.72	1.44	1.55	0.61	0.58	0.54	0.86	1.01
max. (+) dev.	0.54	0.47	0.20	0.41	0.64	1.08	2.49	4.30
max. (−) dev.	3.66	3.06	2.98	1.31	1.21	1.17	2.01	2.01

[a] Total number of states considered.

has not affected the $\pi\pi^*$ states much, but has improved the energies of the $n\pi^*$ states below 6 eV (MAD 0.35 eV).

Table 7 presents the statistical results for triplet excited states, again for all states available and for three subsets (hydrocarbons, CHN compounds, CHO and CHNO compounds). The total number of states with TBE values is smaller than in the singlet case (63 vs 104), but still large enough for meaningful evaluations. As compared to the singlets, the overall performance for triplets is essentially the same for OMx (MAD 0.45−0.49 eV vs 0.45−0.50 eV), but somewhat worse for MNDO/AM1/PM3 (MAD 1.30−1.55 eV vs 1.19−1.41 eV) and also for INDO/S and INDO/S2 (MAD 0.65−0.72 eV vs 0.51 eV), but the general trends remain the same. The semiempirical excitation energies again tend to be too low, even slightly more so than in the case of the singlets (see the overall mean errors). Considering the subsets, the largest mean absolute deviations occur for the hydrocarbons in MNDO/AM1/PM3 and for the oxygen-containing compounds for OM1, OM2, INDO/S, and INDO/S2 (again in analogy to the singlets). Somewhat surprisingly, we find for the oxygen-containing compounds that INDO/S2 performs worse than INDO/S for the triplet states (MAD 1.02 eV vs 0.72 eV) despite the reparametrization, while OM3 again appears to be most balanced and performs best (MAD 0.47 eV vs 0.57−0.62 eV for OM1 and OM2).

For many photophysical and photochemical processes, the energy difference between excited states is of crucial importance. Table 8 provides a statistical evaluation for three such differences that are particularly relevant, between the

two lowest singlets, between the two lowest triplets, and between the lowest singlet and triplet states. It is obvious that the OMx methods show by far the best performance, with mean absolute deviations from the TBE values of 0.40−0.43, 0.23−0.32, and 0.20−0.21 eV for the S_2−S_1, T_2−T_1, and S_1−T_1 energy gaps, respectively. The corresponding MAD values for the other semiempirical methods (MNDO, AM1, PM3, INDO/S, INDO/S2) are significantly higher and generally lie in the range between 0.4−0.6 eV for all three energy gaps considered. The associated mean errors indicate that these energy differences are normally underestimated in MNDO, AM1, and PM3 (on average by 0.3−0.4 eV for T_2−T_1 and by 0.1−0.2 eV otherwise), while they are generally overestimated in INDO/S and INDO/S2 (on average by 0.4−0.6 eV for T_2−T_1 and by 0.2−0.3 eV otherwise); by contrast, the OMx energy gaps are too low on average by only 0.1−0.2 eV for T_2−T_1 and scatter around the TBE values for S_2−S_1 and S_1−T_1 (on average within 0.1 eV). These data show that the OMx methods predict the energy sequence within the excited states much better than the other semiempirical methods studied.

The overall performance for all singlet and triplet states considered can be judged from the correlation plots shown in Figure 1. The OMx results are rather close to the ideal correlation line with unit slope and yield reasonably high correlation coefficients (r = 0.9391−0.9520). The MNDO energies are generally too low and scatter considerably (r = 0.8107), while the AM1 and PM3 energies are also too low but more regular (r = 0.8574−0.8838). The INDO/S and INDO/S2 data also show some scatter around the TBE values

Table 8. Deviations of Semiempirical S_2-S_1, T_2-T_1, and S_1-T_1 Energy Differences (in eV) with Respect to Theoretical Best Estimates

	MNDO	AM1	PM3	OM1	OM2	OM3	INDO/S	INDO/S2
S_2-S_1[a]								
count[b]	27	27	27	27	27	27	27	27
mean	−0.13	−0.14	−0.24	0.04	−0.11	−0.10	0.27	0.20
abs. mean	0.64	0.52	0.57	0.42	0.40	0.43	0.64	0.55
std. dev.	0.83	0.68	0.67	0.47	0.46	0.49	0.88	0.80
max. (+) dev.	1.35	1.14	1.02	0.89	0.70	0.79	2.26	2.34
max. (−) dev.	1.65	1.42	1.29	0.75	0.97	1.02	1.22	1.22
T_2-T_1[c]								
count[b]	19	19	19	19	19	19	19	19
mean	−0.30	−0.26	−0.37	−0.12	−0.17	−0.14	0.45	0.57
abs. mean	0.47	0.43	0.48	0.23	0.30	0.32	0.58	0.72
std. dev.	0.61	0.52	0.53	0.30	0.36	0.41	0.88	1.17
max. (+) dev.	1.46	1.02	0.50	0.46	0.60	0.70	2.52	3.81
max. (−) dev.	1.00	0.86	0.88	0.63	0.67	1.12	0.41	0.41
S_1-T_1[d]								
count[b]	20	20	20	20	20	20	20	20
mean	−0.05	−0.06	−0.10	0.04	−0.07	−0.04	0.28	0.25
abs. mean	0.44	0.38	0.38	0.20	0.21	0.21	0.61	0.59
std. dev.	0.61	0.52	0.47	0.25	0.26	0.25	0.76	0.75
max. (+) dev.	1.52	1.11	0.81	0.54	0.46	0.39	1.80	1.80
max. (−) dev.	1.07	1.07	0.86	0.31	0.55	0.50	0.92	0.98

[a] Energy difference between the two lowest singlet excited states. [b] Total number of molecules considered. [c] Energy difference between the two lowest triplet excited states. [d] Energy difference between the lowest singlet and triplet excited states.
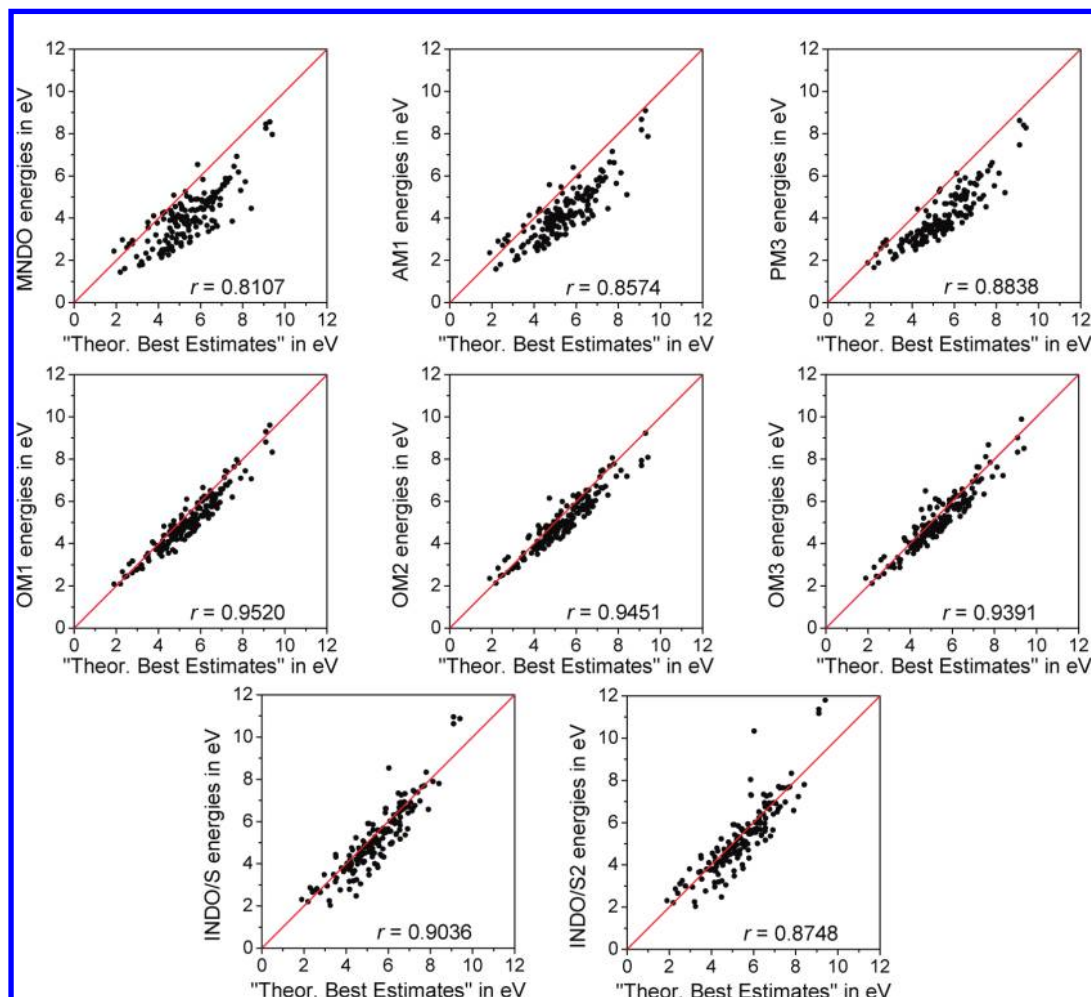
Figure 1. Correlation plots of vertical excitation energies for all states considered in this study using theoretical best estimates as reference data.

($r = 0.8748−0.9036$). Similar separate correlation plots for the singlet and triplet states are presented in the Supporting Information (Figures S1 and S2).

Figure 2 provides a visual summary in the form of a histogram plot for the deviations of the computed vertical excitation energies from the TBE values of all states below
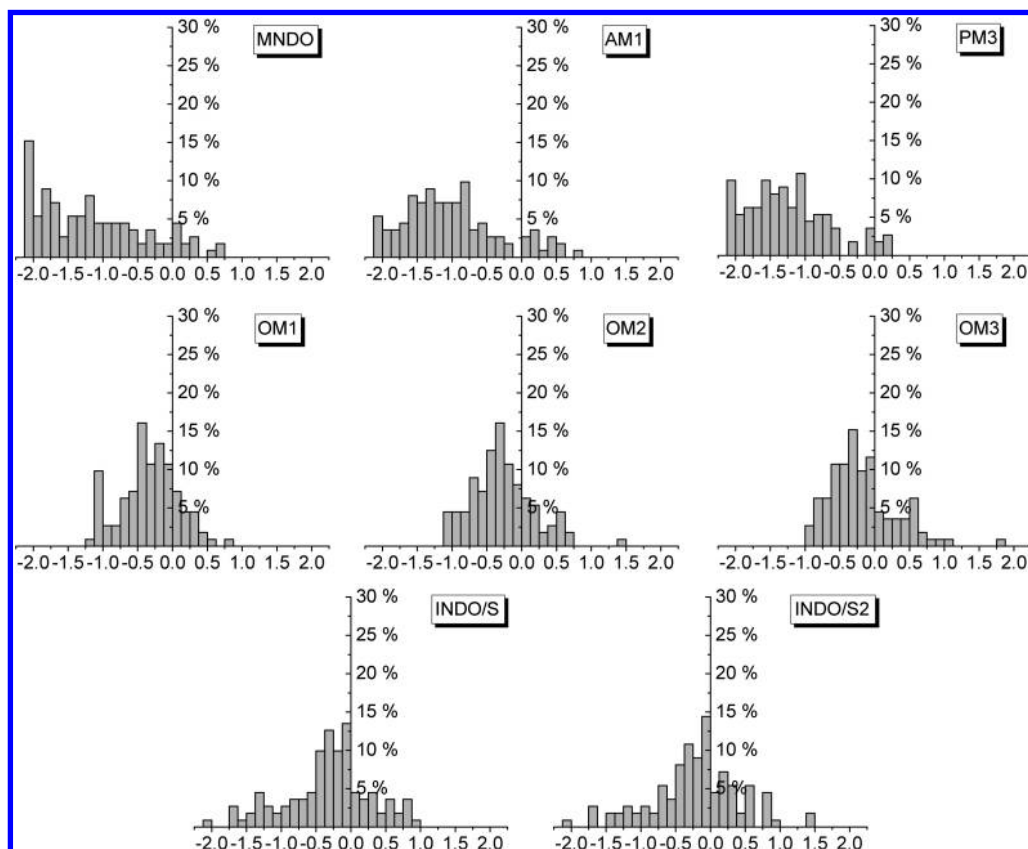
Figure 2. Histogram plots for the deviations of the calculated vertical excitation energies from the theoretical best estimates for all states below 6 eV considered in this study. In the case of MNDO, AM1, and PM3, the left-hand column collects all states whose energies are underestimated by more than 2 eV.

6 eV. Analogous histograms are given in the Supporting Information (Figures S3 and S4) separately for the nπ* and ππ* states. On the basis of these histograms and the statistical data presented in this section, it is possible to rank the semiempirical methods investigated here according to their overall performance to describe excitation energies. The OM*x* methods are clearly the best choice, with minor differences in the performance of the three variants, with a slight edge of OM2 and OM3 over OM1. INDO/S and INDO/S2 also perform reasonably well on average, specially for low-lying states, but show a considerably wider error distribution. Finally, the MNDO, AM1, and PM3 methods that are well established and widely used for ground states are least suitable for electronically excited states in their standard ground-state parametrization.

Among the previously studied DFT-based methods,[15] TD-BP86 and TD-BHLYP show mean absolute deviations from the theoretical best estimates similar to those of the OM*x* methods, while TD-B3LYP and especially DFT/MRCI perform better (MAD for singlets, TD-B3LYP 0.27 eV, DFT/MRCI 0.22 eV, OM*x* 0.45−0.50 eV; triplets, TD-B3LYP 0.45 eV, DFT/MRCI 0.25 eV, OM*x* 0.45−0.49 eV). This is not unexpected in view of the computational costs, which are higher by about 3 orders of magnitude for TD-B3LYP (see section 2) and by even more in the case of DFT/MRCI. The DFT-based results have also been evaluated against other ab initio reference data.[15] The corresponding statistical evaluations for the semiempirical methods are documented

in the Supporting Information (Tables S10−S19) but will not be discussed here.

6.2. One-Electron Properties. As mentioned before, we have not yet established theoretical best estimates for the oscillator strengths and state dipole moments in our benchmark set. Therefore, we use three sets of ab initio reference data: the comparisons with our own MS-CASPT2/TZVP results[14] are presented here, while those with the published CASPT2 data mainly from the Roos group and with our own CC2/TZVP data[15] are given in the Supporting Information (Figures S5−S8, Tables S20 and S21). The qualitative conclusions from these comparisons are the same for each set of reference data.

Table 9 summarizes the statistical evaluation of the calculated oscillator strengths for all dipole-allowed transitions, and Figure 3 shows the corresponding correlation plots against the MS-CASPT2/TZVP reference data. Visual inspection of these plots indicates that the OM*x* oscillator strengths scatter around the ideal correlation line with unit slope and correlate reasonably well with the reference data ($r = 0.9032−0.9425$). The MNDO/AM1/PM3 oscillator strengths tend to be too low and correlate less well ($r = 0.8690−0.9235$), while INDO/S and INDOS/2 tend to overestimate the reference values and have low correlation coefficients ($r = 0.8781−0.8882$). These visual impressions are corroborated by the statistical results. For the whole set of dipole-allowed transitions considered, the mean absolute deviations from the MS-CASPT2/TZVP oscillator strengths

Electronically Excited States

J. Chem. Theory Comput., Vol. 6, No. 5, 2010 **1561**

Table 9. Deviations of Semiempirical Oscillator Strengths (in Length Representation) of Dipole-Allowed States with Respect to ab Initio Reference Data

	MNDO	AM1	PM3	OM1	OM2	OM3	INDO/S	INDO/S2
Published CASPT2 Results[a]								
count[b]	100	100	100	100	100	100	98	98
mean	−0.085	−0.058	−0.066	0.014	0.006	0.027	0.075	0.070
abs. mean	0.114	0.100	0.102	0.081	0.075	0.102	0.148	0.150
std. dev.	0.203	0.184	0.182	0.146	0.119	0.181	0.222	0.226
max. (+) dev.	0.511	0.564	0.406	0.309	0.307	0.831	0.707	0.707
max. (−) dev.	0.870	0.844	0.831	0.830	0.432	0.822	0.627	0.750
MS-CASPT2/TZVP Results[a]								
count[b]	100	100	100	100	100	100	98	98
mean	−0.100	−0.073	−0.081	−0.001	−0.009	0.011	0.058	0.053
abs. mean	0.131	0.113	0.105	0.086	0.077	0.094	0.134	0.137
std. dev.	0.197	0.166	0.155	0.134	0.112	0.152	0.203	0.208
max. (+) dev.	0.346	0.342	0.314	0.421	0.393	0.609	0.659	0.659
max. (−) dev.	0.711	0.529	0.516	0.515	0.320	0.507	0.671	0.794
CC2/TZVP Results[a]								
count[b]	100	100	100	100	100	100	98	98
mean	−0.035	−0.008	−0.015	0.064	0.057	0.077	0.124	0.120
abs. mean	0.092	0.087	0.084	0.102	0.082	0.118	0.151	0.150
std. dev.	0.171	0.153	0.146	0.158	0.130	0.189	0.229	0.230
max. (+) dev.	0.490	0.543	0.385	0.440	0.387	0.810	0.845	0.845
max. (−) dev.	0.795	0.670	0.670	0.475	0.312	0.467	0.414	0.537

[a] See ref 14. [b] Total number of states considered.
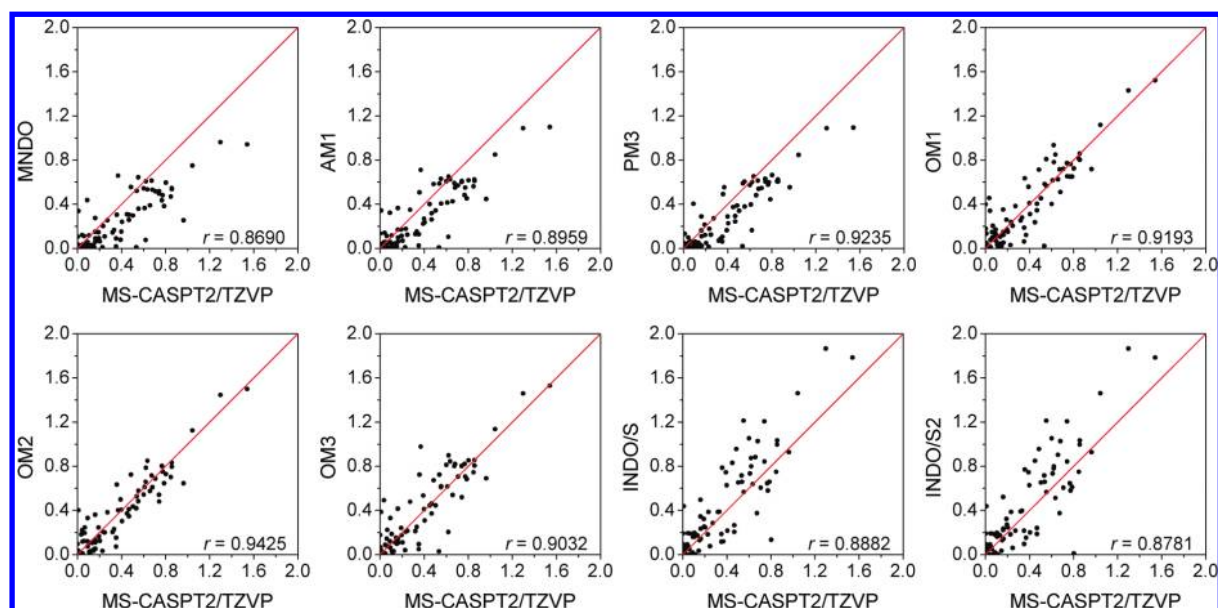
Figure 3. Correlation plots of oscillator strengths for all dipole-allowed transitions using MS-CASPT2/TZVP results as reference data.

are 0.077−0.094 for OM*x*, 0.105−0.131 for MNDO/AM1/ PM3, and 0.134−0.137 for INDO/S and INDO/S2. The overall mean deviations confirm that the OM*x* oscillator strengths are on average close to the MS-CASPT2/TZVP reference data (within 0.01), while the results from MNDO/ AM1/PM3 are mostly too low (on average by 0.07−0.10) and those from INDO/S and INDO/S2 mostly too high (on average by 0.05−0.06).

Table 10 presents the statistical evaluation of all nonzero state dipole moments in our benchmark set, and Figure 4 shows the corresponding correlation plots against the MS-CASPT2/TZVP reference data. It is evident that the semiempirical results often differ rather strongly from the ab initio data. The performance of OM*x* (MAD 0.68−0.71 D) and MNDO/AM1/PM3 (MAD 0.66−0.76 D) is similar in this

case and better than that of INDO/S and INDO/S2 (MAD 1.08−1.19 D). The overall mean errors suggest that state dipole moments tend to be underestimated by OM*x* (on average by 0.10−0.23 D) and somewhat more by MNDO/ AM1/PM3 (on average by 0.27−0.37 D), while INDO/S and INDO/S2 tend to overestimate them (on average by 0.40−0.49 D). These trends are also apparent from the correlation plots.

Summarizing the statistical evaluations for the oscillator strengths and state dipole moments, the OM*x* methods show the best overall performance for these one-electron properties, without any pronounced preference for one of the three variants. To put the OM*x* results into perspective, we note that rather similar deviations from the ab initio reference data have also been found for DFT-based methods, both for oscillator strengths (MAD: TD-B3LYP 0.113, DFT/MRCI

Table 10. Deviations of Semiempirical State Dipole Moments (in D) with Respect to ab Initio Reference Data[a]

	MNDO	AM1	PM3	OM1	OM2	OM3	INDO/S	INDO/S2
Published CASPT2 Results[b]								
count[c]	142	142	142	142	142	142	104	104
mean	−0.61	−0.58	−0.51	−0.47	−0.11	−0.34	0.11	0.20
abs. mean	0.93	1.03	1.02	0.94	0.86	0.92	1.25	1.35
std. dev.	1.44	1.53	1.54	1.44	1.38	1.43	2.16	2.26
max. (+) dev.	2.37	3.81	2.98	5.22	5.75	5.49	9.73	9.92
max. (−) dev.	7.17	7.28	7.21	6.58	5.82	6.10	4.92	5.08
MS-CASPT2/TZVP Results[b]								
count[c]	142	142	142	142	142	142	104	104
mean	−0.37	−0.34	−0.27	−0.23	−0.13	−0.10	0.40	0.49
abs. mean	0.66	0.76	0.72	0.68	0.68	0.71	1.08	1.19
std. dev.	0.91	1.05	1.00	1.00	1.07	1.07	2.03	2.11
max. (+) dev.	3.52	3.73	3.98	5.97	5.64	5.38	10.01	9.81
max. (−) dev.	4.91	4.60	4.91	4.25	4.68	4.37	4.90	4.30
RI-CC2/TZVP Results[b]								
count[c]	140	140	140	140	140	140	103	103
mean	−0.16	−0.13	−0.07	−0.04	0.33	0.09	0.66	0.75
abs. mean	0.83	0.88	0.80	0.74	0.80	0.81	1.16	1.26
std. dev.	1.11	1.19	1.12	1.15	1.25	1.22	2.18	2.32
max. (+) dev.	4.25	4.53	4.57	6.81	5.35	5.01	10.62	9.75
max. (−) dev.	2.93	3.44	3.78	1.85	4.90	4.59	1.98	2.13

[a] Ground and excited states. Only singlet states in the case of INDO/S and INDO/S2. [b] See ref 14. [c] Total number of states considered.
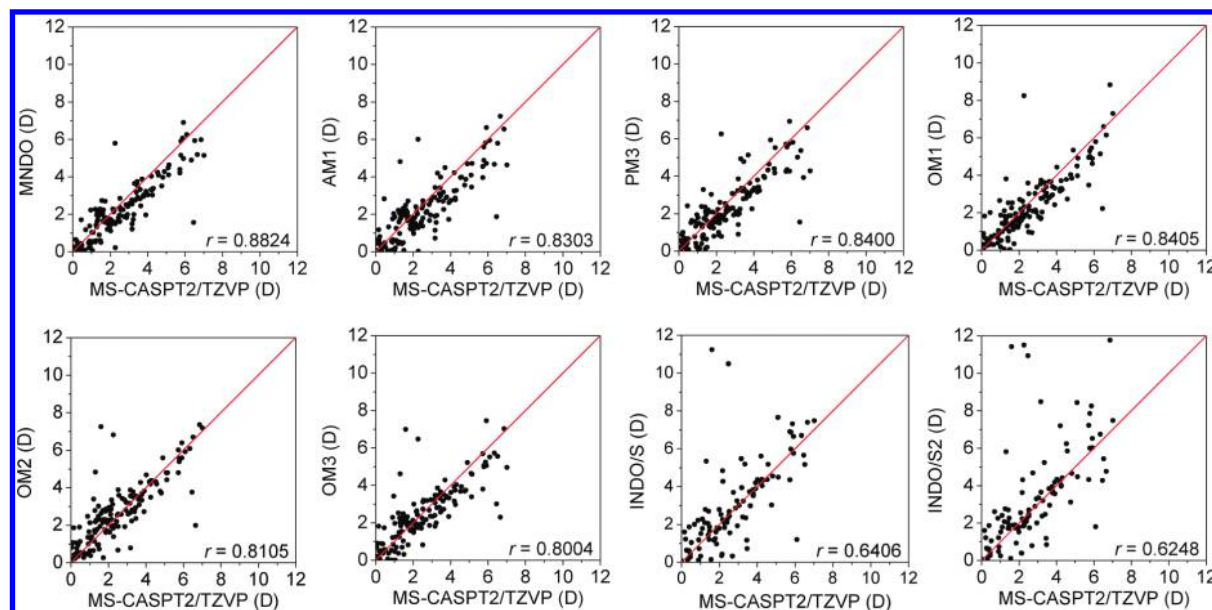
Figure 4. Correlation plots of nonzero state dipole moments (in D) for all states considered, using MS-CASPT2/TZVP results as reference data.

0.069, OM*x* 0.077−0.094) and for state dipole moments (MAD: TD-B3LYP 0.59 D, DFT/MRCI 0.58 D, OM*x* 0.68−0.71 D).[15]

7. Conclusions

We have performed a comprehensive validation study for eight semiempirical methods to assess their performance in describing vertical excitation energies and one-electron properties of electronically excited states. The semiempirical results were compared against ab initio reference data for a recently developed benchmark set of 28 medium-sized organic molecules, focusing on the published theoretical best estimates for excitation energies and MS-CASPT2/TZVP oscillator strengths and excited-state dipole moments.[14] The benchmark set includes only valence excited states because

Rydberg states cannot be described by semiempirical methods due to the use of a minimal basis set.

The standard ground-state methods MNDO, AM1, and PM3 strongly underestimate the reference excitation energies, typically by more than 1 eV, and they often give a wrong order of the excited states because the errors are not uniform. These methods thus seem unsuitable for excited-state work unless one is prepared to undertake a system-specific reparametrization.

The INDO/S method has been parametrized for spectroscopic purposes and has often been applied to compute electronic spectra at the semiempirical level. In the current benchmark, it performs reasonably well, with typical errors of 0.5 eV for excited singlets and 0.6−0.7 eV for triplet states. By design, INDO/S does not properly account for

Electronically Excited States

J. Chem. Theory Comput., Vol. 6, No. 5, 2010 **1563**

states with significant double-excitation character, and it is thus not surprising that the INDO/S errors are smaller for low-lying than for high-lying singlet states. The INDO/S2 variant has been reparametrized for oxygen,[35] which indeed improves the results for low-lying singlet states of oxygen-containing compounds, but turns out to be detrimental for the high-lying singlets and the triplet states of our benchmark molecules. Hence, INDO/S2 cannot be recommended as a general-purpose alternative to INDO/S.

The NDDO-based orthogonalization-corrected methods OM1, OM2, and OM3 show the best overall performance in the present benchmark. They give vertical excitation energies with typical errors of 0.4−0.5 eV both for valence excited singlet and for triplet states, and they predict the order of the excited states and the gaps between the low-lying states more reliably than the other semiempirical methods. The OM*x*/CISDTQ wave functions for valence excited states normally show a qualitatively similar composition to the corresponding ab initio CASSCF and CASPT2 wave functions (with regard to the leading configurations). The OM*x* methods provide reasonable one-electron properties (oscillator strengths and excited-state dipole moments), which are overall somewhat closer to the ab initio reference data than those obtained from the other semiempirical methods. The performance of the three OM*x* variants is generally quite similar, although OM2 and OM3 would seem to have a slight edge over OM1 in an overall assessment.

Conceptually, the OM*x* methods are superior to INDO/S because of the use of the more refined NDDO integral approximation, and to all other semiempirical methods considered here because of the explicit inclusion of corrections that mimic Pauli exchange repulsion and thus cause an unsymmetric splitting of bonding and antibonding levels. These advances are believed to contribute to the good OM*x* performance for electronically excited states, which is achieved despite the fact that the adjustable OM*x* parameters have been determined by calibrating purely against ground-state reference data. One may anticipate that further improvements are possible through a general-purpose reparametrization of the OM*x* methods that includes ground-state and excited-state reference data in a balanced manner. In such reparametrization work as well as in other applications, it seems justified to replace the CISDTQ treatment employed here by a more cost-efficient MR-CISD approach, which yields essentially the same results provided that the reference configurations dominate the CI wave functions (e.g., by requiring their combined weight to exceed 85%).

In summary, at the semiempirical level, the OM*x* methods appear to be the best choice for studying excited-state phenomena in large organic chromophores. While the results from the current benchmark support such applications using the existing ground-state parametrization, further improvements in the OM*x* description of electronically excited states may be expected from a balanced reparametrization.

Supporting Information Available: Additional numerical results for excitation energies and one-electron properties (Tables S1−S9), additional correlation and histogram plots (Figures S1−S8), and additional statistical evaluations (Tables S10−S22). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483–5488.

(2) Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218–1226.

(3) Roos, B. O.; Andersson, K.; Fülscher, M. P.; Malmqvist, P.-Å.; Serrano-Andrés, L.; Pierloot, K.; Merchán, M. Multiconfigurational Perturbation Theory: Applications in Electronic Spectroscopy. In *Advances in Chemical Physics: New Methods in Computational Quantum Mechanics*; Prigogine, I., Rice, S. A., Eds.; John Wiley & Sons: New York, 1996; Vol. 93.

(4) Finley, J.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. *Chem. Phys. Lett.* **1998**, *288*, 299–306.

(5) Christiansen, O.; Koch, H.; Jørgensen, P. *Chem. Phys. Lett.* **1995**, *243*, 409–418.

(6) Christiansen, O.; Koch, H.; Jørgensen, P. *J. Chem. Phys.* **1995**, *103*, 7429–7441.

(7) Koch, H.; Christiansen, O.; Jørgensen, P.; de Meras, A. M. S.; Helgaker, T. *J. Chem. Phys.* **1997**, *106*, 1808–1818.

(8) Runge, E.; Gross, E. K. U. *Phys. Rev. Lett.* **1984**, *52*, 997–1000.

(9) Dreuw, A.; Head-Gordon, M. *Chem. Rev.* **2005**, *105*, 4009–4037.

(10) Casida, M. E. *J. Mol. Struct. (THEOCHEM)* **2009**, *914*, 3–18.

(11) Grimme, S.; Waletzke, M. *J. Chem. Phys.* **1999**, *111*, 5645–5655.

(12) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **1997**, *106*, 1063–1079.

(13) Curtiss, L. A.; Raghavachari, K.; Redfern, P. C.; Pople, J. A. *J. Chem. Phys.* **2000**, *112*, 7374–7383.

(14) Schreiber, M.; Silva-Junior, M. R.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2008**, *128*, 134110/1–25.

(15) Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2008**, *129*, 104103/1–14.

(16) Sauer, S. P. A.; Schreiber, M.; Silva-Junior, M. R.; Thiel, W. *J. Chem. Theory Comput.* **2009**, *5*, 555–564.

(17) Jacquemin, D.; Wathelet, V.; Perpete, E. A.; Adamo, C. *J. Chem. Theory Comput.* **2009**, *5*, 2420–2435.

(18) Goerigk, L.; Moellmann, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4611–4620.

(19) Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126*, 4007–4016.

(20) Cordova, F.; Doriol, L. J.; Ipatov, A.; Casida, M. E.; Filippi, C.; Vela, A. *J. Chem. Phys.* **2007**, *127*, 164111/1–18.

(21) Dewar, M. J. S.; Thiel, W. *J. Am. Chem. Soc.* **1977**, *99*, 4899–4907.

(22) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. *J. Am. Chem. Soc.* **1985**, *107*, 3902–3909.

(23) Stewart, J. J. P. *J. Comput. Chem.* **1989**, *10*, 209–220.

(24) Thiel, W. Semiempirical Methods. In *Handbook of Molecular Physics and Quantum Chemistry*; Wilson, S., Ed.; Wiley: Chicester, UK, 2003; Vol. 2.

(25) Thiel, W. Semiempirical Quantum-Chemical Methods in Computational Chemistry. In *Theory and Applications of Computational Chemistry: The First 40 Years*; Dykstra, C. E., Kim, K., Frenking, G. E., Eds.; Elsevier: Amsterdam, 2005.

(26) Bredow, T.; Jug, K. *Theor. Chim. Acta* **2005**, *113*, 1–14.

(27) Thiel, W. *J. Am. Chem. Soc.* **1981**, *103*, 1425–1431.

(28) Dewar, M. J. S.; Fox, M. A.; Campbell, K. A.; Chen, C. C.; Friedheim, J. E.; Holloway, M. K.; Kim, S. C.; Liescheski, P. B.; Pakiari, A. M.; Tien, T. P.; Zoebisch, E. G. *J. Comput. Chem.* **1984**, *5*, 480–485.

(29) Gonzalez-Lafont, A.; Truong, T. N.; Truhlar, D. G. *J. Phys. Chem.* **1991**, *95*, 4618–4627.

(30) Creatini, L.; Cusati, G. G.; Persico, M. *Chem. Phys.* **2008**, *347*, 492–502.

(31) Ridley, J.; Zerner, M. *Theor. Chim. Acta* **1973**, *32*, 111–134.

(32) Ridley, J. E.; Zerner, M. C. *Theor. Chim. Acta* **1976**, *42*, 223–236.

(33) Zerner, M. C.; Loew, G. H.; Kirchner, R. F.; Muellerwesterhoff, U. T. *J. Am. Chem. Soc.* **1980**, *102*, 589–599.

(34) Kotzian, M.; Rösch, N.; Zerner, M. C. *Theor. Chim. Acta* **1992**, *81*, 201–222.

(35) Li, J.; Williams, B.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Phys.* **1999**, *110*, 724–733.

(36) Kolb, M. Thesis, Universität Wuppertal, 1991.

(37) Kolb, M.; Thiel, W. *J. Comput. Chem.* **1993**, *14*, 775–789.

(38) Weber, W. Thesis, Universität Zürich, 1996.

(39) Weber, W.; Thiel, W. *Theor. Chim. Acta* **2000**, *103*, 495–506.

(40) Scholten, M. Thesis, Universität Düsseldorf, 2003.

(41) Otte, N.; Scholten, M.; Thiel, W. *J. Phys. Chem. A* **2007**, *111*, 5751–5755.

(42) Strodel, P.; Tavan, P. *J. Chem. Phys.* **2002**, *117*, 4677–4683.

(43) Wanko, M.; Hoffmann, M.; Strodel, P.; Koslowski, A.; Thiel, W.; Neese, F.; Frauenheim, T.; Elstner, M. *J. Phys. Chem. B* **2005**, *109*, 3606–3615.

(44) Hoffmann, M.; Wanko, M.; Strodel, P.; König, P. H.; Frauenheim, T.; Schulten, K.; Thiel, W.; Tajkhorshid, E.; Elstner, M. *J. Am. Chem. Soc.* **2006**, *128*, 10808–10818.

(45) Keal, T. W.; Koslowski, A.; Thiel, W. *Theor. Chim. Acta* **2007**, *118*, 837–844.

(46) Fabiano, E.; Keal, T. W.; Thiel, W. *Chem. Phys.* **2008**, *349*, 334–347.

(47) Fabiano, E.; Thiel, W. *J. Phys. Chem. A* **2008**, *112*, 6859–6863.

(48) Lan, Z.; Fabiano, E.; Thiel, W. *J. Phys. Chem. B* **2009**, *113*, 3548–3555.

(49) Lan, Z.; Fabiano, E.; Thiel, W. *ChemPhysChem* **2009**, *10*, 1225–1229.

(50) Keal, T. W.; Wanko, M.; Thiel, W. *Theor. Chim. Acta* **2009**, *123*, 145–156.

(51) Schreiber, M.; Silva-Junior, M. R.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2008**, *128*, 134110/1–25; Supporting Information deposited as EPAPS Document No. E-JCPSA6-128-032811. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.

(52) Thiel, W. *MNDO99, version 6.1*; Mülheim an der Ruhr, Max-Planck-Institut für Kohlenforschung: Germany, 2007.

(53) Zerner, M. C.; Ridley, J. E.; Bacon, A. D.; Head, J. D.; Edwards, W. D.; McKelvey, J.; Cuberson, J. C.; Knappe, P.; Cory, M. G.; Weiner, B.; Baker, J. D.; Parkinson, W. A.; Kannis, D.; Yu, J.; Roesch, N.; Kotzian, M.; Karelson, T. T. M. M.; Zheng, X.; Pearl, G.; Broo, A.; Cullen, K. A. J. M.; Li, J.; Hawkins, G. D.; Thompson, J. D.; Kelly, C. P.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. *ZINDO-MN, version 1.2*; Quantum Theory Project, University of Florida, Gainesville, and Department of Chemistry, University of Minnesota, Minneapolis: Florida/Minnesota, 2005.

(54) Koslowski, A.; Beck, M. E.; Thiel, W. *J. Comput. Chem.* **2003**, *24*, 714–726.

(55) *TURBOMOLE V5.9.1 2007*, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989–2007, TURBOMOLE GmbH, since 2007; available from http://www.turbomole.com.

(56) Silva-Junior, M. R.; Sauer, S. P. A.; Schreiber, M.; Thiel, W. *Mol. Phys.* **2010**, *108*, 453–465.

(57) Schulten, K.; Karplus, M. *Chem. Phys. Lett.* **1972**, *14*, 305–309.

(58) Hudson, B. S.; Kohler, B. E. *Chem. Phys. Lett.* **1972**, *14*, 299–304.

(59) Head, J. D. *Int. J. Quantum Chem.* **2003**, *95*, 580–592.

(60) González-Luque, R.; Merchán, M.; Roos, B. O. *Z. Phys. D: At., Mol. Clusters* **1996**, *36*, 311–316.

(61) Miura, M.; Aoki, Y.; Champagne, B. *J. Chem. Phys.* **2007**, *127*, 084103/1–16.

(62) Hättig, C.; Weigend, F. *J. Chem. Phys.* **2000**, *113*, 5154–5161.

(63) Hättig, C.; Hald, K. *Phys. Chem. Chem. Phys.* **2002**, *4*, 2111–2118.

(64) Hättig, C.; Köhn, A.; Hald, K. *J. Chem. Phys.* **2002**, *116*, 5401–5410.

(65) Hättig, C.; Köhn, A. *J. Chem. Phys.* **2002**, *117*, 6939–6951.

(66) Voityuk, A. A.; Zerner, M. C.; Rösch, N. *J. Phys. Chem. A* **1999**, *103*, 4553–4559.