Theor Chim Acta (1995) 91: 91–108

Theoretica
Chimica Acta

A CASPT2 study of the valence and lowest Rydberg electronic states of benzene and phenol

Johan Lorentzon, Per-Åke Malmqvist, Markus Fülscher, Björn O. Roos
Department of Theoretical Chemistry, Chemical Centre, P.O.B. 124, S-221 00 Lund, Sweden

Received August 24, 1994/Accepted October 12, 1994

Summary. The valence excited states and the 3*s*, 3*p*, and 3*d* (united atom) Rydberg states of benzene and phenol have been obtained by the CASPT2 method, which computes a second-order perturbation correction to complete active space self-consistent field (CASSCF) energies. All non-zero dipole oscillator strengths are also computed, at the CASSCF level. For benzene, 16 singlet and 16 triplet states with excitation energies up to ca. 7.86 eV (63 400 cm^{-1}) are obtained. Of these, 12 singlet and three triplet energies are experimentally known well enough to allow meaningful comparison. The average error is around 0.1 eV. The highest of these singlet states ($2^1\mathrm{E}_{2g}$) is the highest valence $\pi\pi^*$ state predicted by elementary π-electron theory. Its energy is then considerably lower than has been suggested from laser flash experiments, but in perfect agreement with a reinterpretation of that experiment. For phenol, 27 singlet states are obtained, in the range 4.53–7.84 eV (63 300 cm^{-1}). Only the lowest has a well-known experimental energy, which agrees with the computed result within 0.03 eV. The ionization energy is in error by 0.05 eV.

Key words: C6H6 – C6H6O – Electronic excitation – *ab initio* – CASPT2 – Oscillator strength – Rydberg states

1 Introduction

The valence $\pi\pi^*$ excitation spectrum of benzene and of derivatives with similar π-electron structure has been studied since 60 years both experimentally and theoretically. Benzene is the parent compound for large branches of organic and physical chemistry. The benzene ring is also a common building block in larger molecules, where absorption and fluorescence spectroscopy can be used to show its presence, and to probe its environment. An important feature is the relative constancy of the valence $\pi\pi^*$ excitation spectrum, which allows a qualitative identification of spectral bands with corresponding bands in free benzene, while quantitative differences in detail may permit conclusions about substituents, ligands and medium interactions. Key properties of substituted benzenes are then excitation energies, transition moments and their directions, and electrostatic properties of the excited states. As an example, knowledge of the transition moment direction for a phenol band makes it possible to interprete fluorescence spectra from tyrosin in a protein in terms of orientation and dynamics. Experimentally, the transition dipole directions can be obtained by aligning the molecules in a crystal or stretched film, but typically *two* possible alternative results are obtained, with few experimental clues to which one is correct. Good calculations can aid the interpretation of such experiments.

There is also a theoretical and methodological interest. Simple parametrized semi-empirical methods are able to explain many features of the $\pi\pi^*$ spectra of benzene derivatives, and indeed most planar conjugated systems. It has proved harder to obtain accurate results by *ab initio* methods. In an earlier benzene study [1], we used the complete active space self-consistent field (CASSCF) method, followed by externally contracted configuration interaction (CCI). These calculations were extensive and the results are substantially improved, but one large excitation energy error of 0.7 eV (1^1B_{1u}) still remains. These results have been improved by multi-reference CI (MRCI) [2] and symmetry-adapted cluster CI (SAC-CI) [3]. Only with very large efforts has the error been reduced to 0.5 eV. This challenge has been one of the driving forces in the evolution and testing of new methods. We have developed recently a new *ab initio* method, called CASPT2, which appears to be well suited for this kind of problems. Our first attempt to apply this method to benzene was successful and provided some insight in the type of correlation effects involved [4]. The $\pi\pi^*$ valence states were obtained within 0.2 eV. Also, a number of single Rydberg states have been firmly assigned in benzene in the last decade.

For phenol, the spectrum seems to be much less known, apart from the lowest excited state. Our application of the CASPT2 method to the singlet states of benzene may thus be regarded as a test of the method, while the triplet states and the phenol calculations are primarily predictions.

2 Methods

All calculations were made by programs in the MOLCAS-2 package [5]. The basic approach is the same as has been used in a number of recent applications: a CASSCF calculation of the state(s) of interest is performed, and for each state the energy is refined by adding the dynamic correlation left out of the CASSCF. The dynamic correlation energy is computed by the complete active space perturbation theory to 2nd order method (CASPT2). Properties such as dipole moments, as well as transition moments, are computed on the CASSCF level for simplicity. These properties are usually not strongly affected by the dynamic correlation.

The CASSCF method is by now a well-known standard method of *ab initio* quantum chemistry. The occupied orbital set is subdivided into inactive orbitals, which are always fully occupied, and active orbitals, which have varying occupancy. A configuration interaction CI wave function is formed, which consists of all configuration functions that are possible, with prescribed total number of electrons, total spin, and possibly additional symmetry restrictions. The number of inactive and active orbitals of each symmetry type are specified, but their precise shape is determined by requiring the total energy to be stationary with respect to orbital as well as CI parameters. For any details, see Refs. [6, 7].

By proper selection of active orbitals, it is in principle, and often in practice, easy to ensure inclusion of all configurations that may be of importance to any of the states of interest, and (if needed) along any reaction or dissociation paths. On the other hand, it is rarely possible to include much dynamic correlation by simply increasing the number of active orbitals. The number of configurations increases proportional to the nth power of the size of the active space, where n is the number of active (correlated) electrons. A configuration count on the order to 10 000 is still fairly standard; with a modern work station, 500 000 configurations is a reasonable upper limit to a CASSCF calculation. Going to the limit by merely increasing the active space would include only a minor part of the dynamic correlation. Instead, a more comprehensive inclusion of all correlation effects is done in a subsequent calculation, which until recently was done by the Multi-Reference Singles and Doubles Configuration Interaction MR-SDCI method, or by some closely related modification. This step requires the selection of a subset of the CASSCF configurations to act as reference space. The quality of the results depends primarily on this selection, which usually has to be made as generous as computer resources allow. A recently developed alternative is the above-mentioned CASPT2 method, which uses the CASSCF wave function as a zeroth order, or root, function in a perturbation series. This method has given remarkably good results in a number of different spectroscopic applications, such as benzene [4], azabenzenes [8], polyenes [9] and furan [10]. For details on this method, see Refs. [11–13], and also a recent review article, Ref. [14].

Since semi-empirical calculations appear to handle the present type of problems rather well, while only correlating the valence pi electrons, one may wonder why a similar CASSCF, using orbital optimization and accurate integrals, would not do. Semi-empirical calculations are constructed by (a) assuming a simple approximation (such as SCF or, better, full valence CI) to the exact quantum-mechanical problem, (b) assuming that the relevant electrostatic integrals are either zero or can be expressed in terms of a smaller number of parameters and interatomic distances, and (c) choosing the parameter values to fit experimental data. A successful parameter fit has occasionally been taken to indicate that assumptions (a) and (b) express valid approximations to the exact problem; yet, it has been known for a long time that this is not so. The dynamic correlation left out by assumption (a) is large and variable; the individual integral estimates in (b) are invalid. As a consequence, attempts at *ab initio* calculations cannot rely on, e.g. σ–π separability, which may apply to the model-defining step of the semi-empirical methods. Earlier calculations on conjugated hydrocarbons as well as aromatic systems show that combinations of simultaneous $\sigma\sigma^*$ and $\pi\pi^*$ excitations are very important. Such excitations leave the number of π electrons unchanged, and they can be regarded as intramolecular dispersion effects which effectively lowers the electron–electron repulsion. Also, many excited states can be regarded as resonance states involving strongly ionic structures. Intramolecular dispersion stabilizes these states relative to the covalent ones, and must be fully accounted for in order to obtain correct relative energies. Inclusion of all the relevant excitations in an MRCI calculation is barely feasible for molecules of this size. On the other hand, CASPT2 includes these excitations size-consistently by perturbation theory, using a root function which need to correlate only the valence π electrons. A selection of published results obtained by established *ab initio* methods is presented in Table 1. The ionic states are evidently very troublesome.

All calculations were done at the ground state equilibrium geometry. Thus, the excitation energies are vertical and uncorrected for zero-point vibration energy differences, and transition moments are not averaged over nuclear motion. For phenol, the out of plane OH torsion can possibly affect the results, but also this is not corrected for.

Table 1. Benzene valence excitation energies. Selected theoretical results

Method:	CIPSI	CI	CAS	MRCI	SAC	PT2	PT2	Exp[b]
Reference:	[40]	[41]	[1]	[2]	[3]	[4]	PW[a]	
1^1B_{2u}	5.31	5.00	4.9	4.73	5.25	4.70	4.84	4.90
1^1B_{1u}	8.12	7.63	7.4	6.87	6.60	6.10	6.30	6.20
1^1E_{1u}	8.59	8.34	7.8	7.80	7.47	7.06	7.03	6.94
2^1E_{2g}	8.33	8.33	8.1	8.39		7.77	7.90	7.8 ± 0.2
1^3B_{1u}	4.32	3.83	3.9		4.06	3.89	3.89	3.94
1^3E_{1u}	5.42	4.98	4.9		5.02	4.50	4.49	4.76
1^3B_{2u}	7.22	7.00	6.7		6.02	5.44	5.49	5.60
1^3E_{2g}	7.63	7.28	7.2			7.03	7.12	7.49 ± 0.25
Mean absolute errors (eV), *collected by VB type of state*								
^{1}Covalent[c]	0.41	0.10	0.0	0.17	0.35	0.20	0.06	
^{1}Ionic	1.79	1.41	1.03	0.76	0.47	0.11	0.10	
^{3}Covalent[c]	0.52	0.16	0.09		0.19	0.16	0.16	
^{3}Ionic	1.62	1.40	1.1		0.42	0.14	0.11	

[a] Present work.

[b] Vertical energies. Refs, see text

[c] The E_{2g} states excluded due to experimental uncertainty

For benzene, we used the same structure as in the MRCI calculations by Matos et al. which are close to experiment [15, 16]. Because of the high symmetry, a large number of transitions can occur only in combination with vibrational excitation, and we could not calculate their intensities in this work.

The phenol geometry was determined by Larsen from rotation spectroscopy [17]. The moments of inertia of a complete set of isotopic monosubstitutions determine, by the Kraitchman–Costain method, a set of so-called substitution coordinates. This method gives large errors for the near-zero coordinate values of any atoms close to a principal axis. This happens for atoms O, C_1, C_4 and H_4 along the *a* axis, and for the C_2 and C_6 atoms along the *b* axis. To stabilize the solution, Larsen applied three restrictions. His results are shown in Table 2.

Results from a theoretical study by Bock et al. [18] have been included in the table, and show that the restrictions assumed by Larsen are unsuitable. However, to our knowledge, the experimental structure has not been recomputed with more realistic assumptions, and we have used Larsens substitution geometry without further adjustment. Dipole and transition dipole moment directions are given as polar angles in the *a*, *b* plane of the inertial axes system. The atom positions in this coordinate system have been included in the table, to facilitate the proper interpretation of such data.

Benzene-like derivatives have four singlet excited valence states, or pairs of states, of $\pi\pi^*$ type. Below 9 eV, there are no $\sigma\pi^*$ excitations, and valence $\pi\sigma^*$, if they exist, are even higher up. The singlet valence states are, above the ground 1^1A_{1g} state, the 1^1B_{2u}, 1^1B_{1u}, 1^1E_{1u}, and 2^1E_{2g} states. The triplets are the 1^3B_{1u}, 1^3E_{1u}, 1^3B_{2u}, and 1^3E_{2g} states. The D_{6h} symmetry implied by this notation may be lowered in substituted benzenes, but it is still practical to use the same notation as long as the states retain a strong similarity to those of benzene. The first ionization level is the 1^2E_{1g} ion, and below it one expects and observes a number of Rydberg series, each of which can be conveniently labelled in a united-atom scheme, as *ns*, np_x, etc. Apart from the valence transition, we have thus also Rydberg $\pi\pi^*$, and $\pi\sigma^*$

Table 2. Experimental structure data

Atom	*a* (Å)	*b* (Å)	Atom	*a* (Å)	*b* (Å)
Phenol coordinates, principal axis system					
C_1	0.90985	0.00298			
C_2	0.21985	1.21097	H_2	0.76691	2.14870
C_3	– 1.17456	1.20345	H_3	– 1.71290	2.14377
C_4	– 1.87581	– 0.00294	H_4	– 2.95599	– 0.00452
C_5	– 1.16560	– 1.20407	H_5	– 1.70123	– 2.14602
C_6	0.22658	– 1.20882	H_6	0.78916	– 2.13227
O	2.28314	– 0.05461	H_7	2.62889	0.83821
Bond lengths (Å) *and angles* (°):					
	Used [17]	Theory [18]		Used [17]	Theory [18]
C_1C_2	1.3912	1.3955	$C_6C_1C_2$	120.85	120.42
C_2C_3	1.3944	1.3983	$C_1C_2C_3$	120.43	120.57
C_3C_4	1.3954	1.3942	$C_2C_3C_4$	120.48	120.52
C_4C_5	1.3954	1.3998	$C_3C_4C_5$	120.24	120.30
C_5C_6	1.3922	1.3926	$C_4C_5C_6$	120.79	120.69
C_6C_1	1.3912	1.3975	$C_5C_6C_1$	120.22	120.50
H_2C_2	1.0856	1.0841	$H_2C_2C_1$	120.01	120.22
H_3C_3	1.0835	1.0839	$H_3C_3C_4$	120.04	120.12
H_4C_4	1.0802	1.0828	$H_4C_4C_3$	120.25	120.35
H_5C_5	1.0836	1.0838	$H_5C_5C_4$	119.78	119.94
H_6C_6	1.0813	1.0820	$H_6C_6C_1$	119.23	128.86
OC_1	1.3745		OC_1C_6	117.01	
H_7O	0.9574		H_7OC_1	108.77	
$H_3 \ldots H_4$	2.4820	2.4852			
$H_5 \ldots H_4$	2.4820	2.4878			
Benzene coordinates					
C_1	0.00000	1.39502	H_1	0.00000	2.48004
C_2	1.20813	0.69751	H_2	2.14778	0.69751
Bond lengths (Å) (*all angles* 120°)					
	Used	Exp [15]		Used	Exp [15]
C_1C_2	1.3950	1.3965	C_1H_1	1.0850	1.0850

excitations. Finally, the open-shell singlet states also have a triplet companion at slightly lower energy.

Most of the singlet transitions are forbidden in the full D_{6h} symmetry, and their counterparts in lower symmetry have rather small oscillator strength. They can still be observed by coupling to vibrational excitations of proper symmetry modes, and also by multi-photon techniques, such as multi-photon ionization (MPI). Nevertheless, some benzene states and most of the phenol ones have still not been unambiguously identified spectroscopically.

In *ab initio* calculations of the higher excited states, the Rydberg spectrum is a problem. Even if the calculation is for valence states, their proper description may require an extended basis set which produces intruders, e.g. extra states in the interesting energy range. Intruders complicate or prohibit the convergence of many *ab initio* procedures, in particular perturbation schemes. Also, states of interest may become distorted by a spurious mixing with intruders. One possible solution to this problem is to identify one or more orbitals as important only to intruders, and delete them from the orbital space in subsequent calculations. Alternatively, if one wants also to compute some Rydberg states, or if one suspects mixing between valence and Rydberg states as a genuine physical phenomenon (as in the ethene V-state [19]), one can use special basis functions which ensure a good description of the specific Rydberg states, without the excessive flexibility in the diffuse part which produces intruders. Well-described Rydberg states interact very weakly with other states, so their presence as intruders is usually not a problem in CASPT2.

The basis set consisted of our standard ANO basis sets [20, 21], augmented by a set of diffuse functions to be able to describe the Rydberg orbitals. We tried first a small conventional set of extra functions on carbon atoms for benzene, but this basis produces, in addition to a few well-described Rydberg orbitals, also a number of low-lying intruders. To alleviate the intruder problem, we prepared a special set of Rydberg-like ANO basis functions [22], placed at the molecular center of mass, in the following way. The cation state was computed by the CASSCF method (see below), using the standard basis together with the single-center set of primitive Gaussians, and a density matrix was prepared using the few lowest virtual orbitals of Rydberg type. The submatrix involving the Rydberg-type set was averaged over rotations, and its eigenvectors then describe pure *s*-, *p*- and *d*-type pseudo-natural orbitals. The eigenvalue, or pseudo-occupation number, of each orbital indicates its importance as a basis function. Thus, the MO coefficients of the orbitals with highest eigenvalue were used as contractions coefficients. For each angular momentum, eight primitive Gaussian were employed. In the present work, only one contracted function was used for each angular momentum component (see Table 3). The 3*s*, 3*p* and 3*d* Rydberg states are then well described, while the remaining Rydberg states are pushed up above the energy range considered here.

Our programs can handle only the D_{2h} point group and its subgroups, and the D_{6h} point group of benzene then presents notational as well as technical problems. The standard notation for axes and symmetry elements for planar D_{6h} and D_{2h} molecules are not compatible with each other [23]. In the benzene case we have choosen the *z* axis to be perpendicular to the molecular plane, and the *y* axis to pass through one of the carbon atoms. This is in accordance with the standard D_{6h} notation, but not with the D_{2h}, where by convention these axes would be the *x* and the *z*, resp. A different choice of axes will permute the B labels of the D_{2h} irreducible representations. The correlation between our D_{2h} and D_{6h} labels can be inferred from Table 4.

All the calculations follow essentially the same scheme: For each symmetry label of D_{2h}, a CASSCF calculation gives the CI expansion for each of a number of

Table 3. ANO Rydberg basis set for benzene and phenol

$\alpha(s)$	Coef.(*s*)	$\alpha(p)$	Coef.(*p*)	$\alpha(d)$	Coef.(*d*)
0.024624	0.45838001	0.042335	0.11559865	0.060540	0.02899452
0.011253	– 2.03790350	0.019254	– 0.49563462	0.027446	– 0.20250181
0.005858	– 1.97775801	0.009988	– 0.12633540	0.014204	– 0.26287310
0.003346	– 3.19520437	0.005689	– 0.48203244	0.008077	– 0.43382041
0.002048	3.72387448	0.003476	0.43944200	0.004927	0.01010840
0.001324	– 3.17700482	0.002242	– 0.40531187	0.003175	– 0.15865201
0.000893	1.70276130	0.001511	0.23333446	0.002137	0.08307267
0.000624	– 0.42140146	0.001055	– 0.06248348	0.001491	– 0.02436312

Table 4. Final benzene calculations

Symmetry	CASSCF states	States used for separate CASPT2 calc.
$^1\pi\pi^*$: *Active space* 00005242		
$^1A_g(D_{2h})$	Roots 1–4	$X^1A_{1g}, 2^1A_{1g}, 1^1E_{2g}, 2^1E_{2g}(D_{6h})$
$^1B_{3u}$	1–4	$1^1B_{2u}, 1^1E_{1u}, 2^1E_{1u}$
$^1B_{2u}$	1–4	$1^1B_{1u}, 1^1E_{1u}, 2^1E_{1u}$
$^1B_{1g}$	1–3	$1^1E_{2g}, 1^1A_{2g}, 2^1E_{2g}$
$^3\pi\pi^*$: *Active space* 00004242		
$^3A_g(D_{2h})$	Roots 1–3	$1^3E_{2g}, 2^3E_{2g}, 1^3A_{1g}(D_{6h})$
$^3B_{3u}$	1–3	$1^3E_{1u}, 1^3B_{2u}, 2^3E_{1u}$
$^3B_{2u}$	1–3	$1^3B_{1u}, 1^3E_{1u}, 2^3E_{1u}$
$^3B_{1g}$	1–3	$1^3A_{2g}, 2^3E_{2g}, 1^3A_{2g}$
$^{1,3}\pi\sigma^*$: *Active space* 31112121		
$B_{1u}(D_{2h})$	Roots 1–2	$1A_{2u}, 1E_{2u}$
B_{2g}	1–4	$1E_{1g}, 1B_{1g}, 2E_{1g}, 3E_{1g}$
B_{3g}	1–4	$1E_{1g}, 1B_{2g}, 2E_{1g}, 3E_{1g}$
A_u	1–2	$1A_{1u}, 1E_{1u}$

states with that symmetry, and orbitals such that the average energy of these states is stationary. Each of these CASSCF state functions is, in turn, used as the root function of a CASPT2 calculation. The orbital optimization can of course be modified: a single state, or the states to be averaged, can be specified by energy ordering, or as those with maximum projection onto a vector subspace of the CI coefficient space, and the average may be weighted. This can be used to improve stability or quality of results for particular states. However, when many states are studied, we prefer to find a simple general scheme. A suitable such scheme was found, and is illustrated in Table 4.

2.1 Benzene

For benzene, as for any molecule with higher than 2-fold rotation axes, special care must be taken to obtain proper symmetry states: the active orbital space must be invariant to symmetry operations, which usually means that these orbitals must be complete sets of basis functions for irreducible representations of the point group. When orbitals are optimized for the energy of a single component of a degenerate electronic state, or for an energy average which is not invariant to symmetry operations, the orbital updates must be restricted to prohibit symmetry breaking. In all our benzene calculations, some orbital symmetry restrictions have been enforced by prohibiting symmetry-breaking rotations among orbitals with different D_{6h} label but the same D_{2h} label. However, there is still some symmetry-breaking since the two components of *e* orbitals, in any specific calculation, are not completely equivalent, due to averaging over a set of states which is not invariant to symmetry operations. This effect is not very serious and can be neglected. There is a difference in energy, at most a few hundredths of an eV, between formally equivalent components of the E states, since they are the results of different sets of calculations.

For active spaces, a notation based on the irreducible representations of the D_{2h} subgroup is used. For each of the eight symmetry labels, in order according to

Table 4, the number of active orbitals is given in the order a_g, b_{3u}, b_{2u}, b_{1g}, b_{1u}, b_{2g}, b_{3g}, a_u. Minimal active spaces, comprising valence π orbitals and the Rydberg orbitals, are then labelled 0000 3231 for $\pi\pi^*$ and 3111 2121 for $\pi\sigma^*$ excited states.

As shown by Roos et al. [4], the minimal active space is not satisfactory for the $\pi\pi^*$ valence states, in particular for the 1^1B_{1u}, 1^1E_{1u} and 1^3B_{2u} states, which have an ionic character. We tried a somewhat larger active space, 0000 4242, which however is not identical to that by Roos et al. [4], since some of the orbitals will now be Rydberg orbitals. In exploratory calculations, which used an atom-centered segmented Rydberg basis set, this active space gave a severe perturbation of the 1^1E_{1u} state by an intruder. The CASPT2 quasi-natural orbital analysis shows this intruder to have a Rydberg type orbital populated, and we had the choice of either deleting this orbital, or of including the intruder state in the CASSCF calculation by increasing the active space. Both these approaches gave acceptable results. We decided to go ahead with the larger active space, 0000 5242, for all the singlet $\pi\pi^*$ states, in spite of the increased computer time, since this is a simpler procedure and also limits the risk for further complications later. This choice was kept when we later switched over to the single-center ANO Rydberg basis. For the triplet $\pi\pi^*$ states, we used the 0000 4242 active space. All the $\pi\sigma^*$ states are Rydberg states sharing the same $^2E_{1g}$ ion rest, and this simple and similar structure allowed us to use the minimal active space, 3111 2121, without problems.

In the subsequent CASPT2 calculations, the six core orbitals were left uncorrelated.

2.2 Phenol

For phenol, exploratory calculations with a smaller basis showed that, within the range of interest, there was no states with appreciable excitation from either the oxygen lone pair (O.l.p.) nor from the occupied sigma orbitals. A CASSCF calculation on the 14 lowest states gave a $\pi\pi^*$ spectrum very similar to benzene, with the E_{1u}-like ionic states pushed up above the E_{2g}-like states, and with two partial Rydberg series in the region near and below the "B_{1u}" state. A number of experiments were made with basis sets, and orbital generating procedures, to minimize the effects of crossings of interacting states. The procedures finally adopted were as follows. As already described, a single-center Rydberg basis was devised. Rydberg orbitals of phenol obtained with such contractions was compared to results using the uncontracted basis. The contraction errors in such quantities as orbital energy differences and second moments were negligible. Moreover, the contracted basis prepared for benzene turned out to give virtually the same results, and we decided to adopt this contraction as a common Rydberg basis for both molecules. The final calculations were subdivided into three groups: the valence $\pi\pi^*$ states, the Rydberg $\pi\pi^*$ states, and the $\pi\sigma^*$ states.

For the valence states, the orbitals and the CASSCF states were obtained by minimizing the average energies of the lowest seven A′ states. Three well-formed Rydberg orbitals had been obtained in earlier pilot studies, and were excluded from the basis, i.e. the optimized orbitals were kept orthogonal to these Rydberg orbitals. The 21 occupied sigma orbitals were inactive, and the active space consisted of nine π orbitals: the O.l.p, the six ring π, and two extravalence orbitals. These were similar to the benzene $2e_{2u}$ pair, and retained this character during the optimization procedure. It should be noted that such near-degenerate orbitals should always be treated in pairs, in order not to exaggerate artificially the symmetry breaking from D_{6h} symmetry. This resulted in 5292 configuration states in A′ symmetry. In the subsequent CASPT2, the seven core orbitals were frozen, resulting in 2 647 311 variables in the PT2 equation system. Also, the ground state was separately optimized to provide as good reference as possible for energies and transition moments, and also the two lowest ions were computed with the same CAS structure but one electron less.

The $\pi\pi^*$ Rydberg states were computed by starting out from the orbitals obtained in the ion calculation described above. First, a CASSCF calculation was performed, where the 21 occupied sigma orbitals were frozen, and the O.l.p. orbital inactive, with 11 π-orbitals active, for the average of the 12 lowest A′ states. This resulted in an optimized active orbital set containing the six valence π orbitals, three correlating orbitals, and three well-formed Rydberg orbitals. The six Rydberg states were identified, and for each the natural orbitals were extracted and used as a *state-specific* orbital set for the subsequent calculations. These consisted in a CASCI, i.e. a CASSCF without orbital reoptimization, using nine active orbitals just as for the valence states. The type of orbitals in the active space is now different, of course: each orbital set contains only that state-optimized Rydberg orbital which must be present, and there are minor variations of the correlating orbitals depending of which of the two ion rests the state contains. In this calculation, for simplicity, all the 21 sigma orbitals were frozen, giving 480 420 PT2 variables. The reference calculation of the ground state, and of the two ions, were repeated using orbitals prepared in a similar manner, and with the sigma orbitals frozen of course. The ionization potential is lowered by almost 0.35 eV, which shows that the effect of excluding the sigma orbitals from the correlation is substantial. Nevertheless, by using the ionization limits as reference for the Rydberg energies, this deficiency is minimized.

The $\pi\sigma^*$ Rydberg states were obtained from a CASSCF with the 21 occupied sigma orbitals, and also the O.l.p, inactive. The active space comprised six sigma orbitals and the six ring valence π orbitals, and the average energy of the lowest 12 states was minimized. The active sigma orbitals became, obviously, six Rydberg orbitals. The twelve states were each used as root state for subsequent CASPT2 calculations, without any reoptimization. In the CASPT2, the seven core orbitals were frozen, resulting in 3 847 312 PT2 parameters. A ground state calculation with the same conditions was done for reference.

3 Results

3.1 Benzene

The results for all benzene states are summarized in Tables 5 (singlets) and 6 (triplets). The valence $\pi\pi^*$ states are essentially in agreement with earlier CASPT2 calculations, but slightly improved. Some states are characterized as ionic, which means that the simplest valence-bond description of the state has a large weight of structures that contain ions. Such states have considerably more dynamic correlation, in particular $(\sigma\sigma^*\pi\pi^*)$ double excitations, than the ground state, and *ab initio* results can seriously overestimate the excitation energy.

3.2 Benzene valence $^1\pi\pi^*$

- 1^1B_{2u} is dipole forbidden, but a weak band is seen which gets intensity by coupling i.e. to the ν_6 vibration mode. The excitation energy is 4.76 eV (origin) or

Table 5. Benzene singlet states

State	CAS	$\langle z^2 \rangle$	PT2	f_L	Exp.
Valence $^1\pi\pi^*$: *Active space* 00005242					
1^1B_{2u}	4.80	30.1	4.84		4.90 vert [24]
1^1B_{1u}	7.32	31.9	6.30		6.20 vert [24]
1^1E_{1u}	8.53	35.1	7.03	0.82	6.94 vert [27]
2^1E_{2g}	7.96	29.8	7.90		7.8 vert [28]?
Rydberg $^1\pi\pi^*$: *Active space* 00005242					
2^1E_{1u}	6.46	79.1	7.16	0.058	7.41 [27]
2^1A_{1g}	7.14	88.3	7.74		7.81? [33], assg A_{2g}
1^1E_{2g}	7.09	88.3	7.77		7.81 [30], assg $3R_g$
1^1A_{2g}	7.08	88.2	7.81		7.81 [33], see 2^1A_{1g}
Rydberg $^1\pi\sigma^*$: *Active space* 31112121					
1^1E_{1g}	6.26	46.2	6.38		6.334 [34]
1^1A_{2u}	6.66	44.7	6.86	0.052	6.932 [26]
1^1E_{2u}	6.74	45.1	6.91		6.953 [26]
1^1A_{1u}	6.82	45.4	6.99		—
1^1B_{2g}	7.33	64.9	7.58		7.460? [33], extrap.
1^1B_{1g}	7.29	65.2	7.58		7.460? [33], extrap.
2^1E_{1g}	7.33	63.5	7.57		7.535 [30], misassg E_{2g}
3^1E_{1g}	7.37	52.4	7.57		—

Table 6. Benzene triplet states

State	CAS	$\langle z^2 \rangle$	PT2	Exp.
Valence $^3\pi\pi^*$: *Active space* 00004242				
1^3B_{1u}	4.05	29.8	3.89	3.94 vert [35]
1^3E_{1u}	5.07	30.3	4.49	4.76 vert [35]
1^3B_{2u}	6.93	32.3	5.49	5.60 vert [35]
1^3E_{2g}	7.61	39.3	7.12	7.24–7.74 (See text).
Rydberg $^3\pi\pi^*$: *Active space* 00004242				
2^3E_{1u}	6.92	74.8	6.98	—
2^3E_{2g}	7.44	77.6	7.55	—
1^3A_{1g}	7.42	86.0	7.62	—
1^3A_{2g}	7.50	88.2	7.70	—
Rydberg $^3\pi\sigma^*$: *Active space* 31112121				
1^3E_{1g}	6.22	46.3	6.34	—
1^3A_{2u}	6.61	44.5	6.80	—
1^3A_{1u}	6.83	45.4	7.00	—
1^3E_{2u}	6.73	45.0	6.90	—
1^3B_{2g}	7.27	72.6	7.53	—
1^3B_{1g}	7.27	73.6	7.53	—
2^3E_{1g}	7.31	52.7	7.57	—
3^3E_{1g}	7.36	49.8	7.56	—

4.90 eV (vertical) [24]. It is a covalent state, and satisfactory results are obtained with all *ab initio* methods with modest effort. The CASSCF energy is quite acceptable.

- 1^1B_{1u} is also dipole forbidden. Lassettre et al. [25] concludes that two electronic bands contribute in the traditional "S_2" region at 6.2–6.5 eV, and places one at

6.20; the other as a band at 6.31, 6.41 ... eV. This is in good agreement with Hiraya and Shobatake [24], where the S_2 band of an absorption spectrum from a seeded He jet has a maximum at 6.20 eV (Origin: 6.03 eV), and with Johnson and Korenowski's 1^1E_{1g} Rydberg origin at 6.334 eV [26]. It is an ionic state, and the CASSCF energy is 1.1 eV too high.

- 1^1E_{1u} is a strong dipole allowed transition, with origin at 6.87 eV and vertical energy 6.94 eV [27, 24]. The experimental oscillator strength is in the region 0.6–1.1 [3]. The absolute extinction coefficient function obtained by Hiraya and Shabotake [24] can be integrated for good oscillator strengths, but only if it is known how to divide up the intensity among the bands. We have graphically integrated the data from their Fig. 1, and conclude that the combined strength of 1^1E_{1u} and 1^1A_{2u} is about 0.8. This is a very ionic state. The CASSCF energy is 1.6 eV too high, and is actually higher than the 2^1E_{2g}.
- 2^1E_{2g} is dipole forbidden. This is the highest valence state, expected from elementary MO theory, and has for a long time escaped observation. Nakashima et al. [28, 29] have reported laser flash experiments, which in cyclohexane solution give a rough but featured spectrum. The spectrum is sharpened and slightly shifted in gas phase. The peaks, if due to $S_n \leftarrow S_1$, implies S_n states at 7.0, 7.8, and 9.4 eV above the ground state. The first two peaks may very well be the 1^1E_{1u} and 1^1E_{2g} states. However, the third strong peak would be above the ionization limit or transitions to the conduction band, respectively. Of course, nothing prevents a high, expected valence state to become dissolved in a continuum, and still show its presence spectroscopically. However, this is a covalent state, which should not be very difficult to get right (with a multiconfiguration method – this state has ca. 30% contribution from double excitations). Our calculations have this state at 7.90 eV excitation energy with CASPT2, and 7.96 eV already at CASSCF level, and with expectation value $\sum_i z_i^2 = 29.8$ a.u., as expected for a valence state but in contrast to the value of 88.3 a.u. obtained for the Rydberg state 1^1E_{2g}, see below. Nakashima et al. may indeed have observed the $2^1E_{2g} \leftarrow 1^1B_{2u}$ transition, but as (part of?) the 7.8 eV peak. It should be considerably more intense than the transition to the Rydberg 1^1E_{2g}, so we tentatively accept an experimental value of 7.8 eV for 2^1E_{2g}. We are not in a position to explain the 9.4 eV peak. However, within the time window when the second energy quantum of 4.5 eV is absorbed, there are a large number of encounters between the initially excited benzene and unexcited molecules, which may form a complex with excitation energies in the range 4.5 eV.

In each case, the CASPT2 energy gives satisfactory results, in spite of the large differential correlation effects.

3.3 Benzene Rydberg $^1\pi\pi^*$ *and* $^1\pi\sigma^*$

The Rydberg states of benzene have been extensively studied by Whetten, Grubb and coworkers [30–33]. For those series that converge to a common ionic state, a convenient notation is to identify the states by a main quantum number and angular momentum label for the Rydberg orbital, even if they are not perfect angular eigenstates. The main features of the series are then:

1. The energy of configuration (ion)Ψ_{nlm} is $E = \text{I.P.} - 1/2(n - \delta_{nlm})^2$, where $n \geqslant 3$ and $n \geqslant l \geqslant m \geqslant 0$. For $m > 0$, the orbitals with $m = 1, 2, 4, 5, 7, 8, \ldots$ are

doubly degenerate, while this degeneracy is lifted for $m = 3, 6, \ldots$ which are actually slightly split pairs.
2. The quantum defects have values $-0.10 < \delta_{nlm} < 0.90$. For given lm, the value is fairly constant and goes to a limit when n grows.
3. With growing l, the quantum defect goes to zero because of diminishing penetration. The trend is that of a screened coulomb potential, i.e. $s < p < d \ldots$ However, the l levels are further split by the static interaction with the D_{6h} core.
4. Finally, the lowest ion state is the degenerate $^2E_{1g}$, so most levels will be further splitted by what Whetten et al. [31] calls the 'quartet' splitting, not referring to spin. This splitting decreases rapidly with l.

- 2^1E_{1u} is the dipole allowed $3p_0$ state, with origin at 7.414 eV [27]. Note that Kitao and Nakatsuji [3] quotes an energy of 7.19 eV, referring to Whetten et al. [31]. However, we cannot find such a statement in that article, and since the transition is well known since 1956 we believe there has been a mistake. The CASPT2 energy for this state is then 0.25 eV too low, which is one of the largest errors encountered in this study. The calculated oscillator strength is 0.058. This is in general agreement with the measured extinction data [24], but we are not able to determine with any accuracy how the absorption intensity in this region should be divided up.
- 2^1A_{1g}, 1^1E_{2g}, and 1^1A_{2g} are terms of the two-photon allowed $3d_1$ configuration. One origin, at 7.807 eV, was assigned by Grubb et al. [33] to be 1^1A_{2g}. They observe the 1^1E_{2g} state with an origin 7.805 eV, which seems to be identical to the "$3R_g$" origin at 7.808 eV reported by Whetten, Fu and Grant [30] two years earlier.
- 1^1E_{1g} Observed by Johnson [34] by 2-photon absorption, with origin 6.334 eV.
- 1^1A_{2u} This is the traditional Wilkinson 2R-A band, the dipole allowed A_{2u} component of the $3p_1$ configuration. Whetten et al. [30] places the origin at 6.928 eV, Johnson and Korenowski at 6.932 [26].
- 1^1E_{2u} This component of the $3p_1$ configuration is only 3-photon allowed. Johnson and Korenowski [26] were able to separate it from the A_{2u} component by 3-photon ionization spectroscopy, and give the origin as 6.953 eV.
- 1^1A_{1u} This is a third component of the $3p_1$ configuration, and should be very close to the other two. It is 3-photon allowed, but appears to be yet unobserved.
- 1^1B_{2g} and 1^1B_{1g} These two states belong to the $3d_2$ configuration. They are allowed only as four-photon transitions, but nevertheless Grubb et al. have observed a number of higher members of the nd_2 Rydberg series, and offer a common extrapolated value of 7.460 eV for the $n = 3$ origins.
- 2^1E_{1g} is the single term of the $3d_0$ configuration. It is two-photon allowed. It has been observed by Whetten et al. [31] (but misassigned "E_{2g}") with origin at 7.535 eV.
- 3^1E_{1g} is the third term of the $3d_2$ configuration. It is two-photon allowed, but has not yet been observed, and thus our ordering of the two last states is tentative.

Since all the Rydberg states in this study have a common ion rest, it is expected that they can all be obtained with similar accuracy by any decent ab initio method. However, this turns out to be not quite true. The CASSCF errors for the $\pi\pi^*$ Rydberg states are all about -0.7 eV, excepting the 2^1E_{1u} state which appears to have an atypically large error both in CASSCF and CASPT2. The $\pi\sigma^*$ states have errors of -0.1 to -0.2 eV, which includes the unknown difference between vertical energies and zero-point corrected adiabatic energies. All the CASPT2 energies (except 2^1E_{1u}) are within 0.1 eV, so there is clearly a systematic difference in dynamic correlation between the $\pi\pi^*$ and $\pi\sigma^*$ Rydberg states.

3.4 Benzene triplet states

In contrast to singlet states, very little is known about triplets in general. The four valence states are known in the literature. Doering [35] gives convincing arguments for the following vertical energies: 1^3B_{1u}(3.94 eV), 1^3E_{1u}(4.63 eV), and 1^3B_{2u}(5.60 eV). By adding this 1^3B_{1u} energy to the 3.5–4.0 eV excitation energy from 1^3B_{2u} to 1^3E_{2g} (Ref. [29]) in solvent, corrected by the solvent shift (− 0.2 eV) of $1^3E_{2g} \leftarrow 1^3B_{2u}$(Ref. [29]), we obtain the somewhat wide range of 7.24–7.74 eV vertical energy for $1^3E_{2g} \leftarrow \tilde{X}$.

The ionic $^3B_{2u}$ state has the typically large CASSCF error of 1.33 eV. The other states are covalent, and the errors are much smaller. The CASPT2 errors are acceptable if we assume the correct experimental $^3E_{2g}$ energy to lie at the lower end of the range. The largest CASPT2 error is then − 0.27 eV for the $^3E_{1u}$ state.

The Rydberg triplets seems to be unknown. In the recent paper by Minaev et al. [36], a value for the $3p_1$ term 1^3A_{2u} is presented as experimental in a table. However, the reference is to the paper by Johnson and Korenowski [26] which is not possible, and the table entry is identical to that for another state. We conclude that they have made a mistake.

Table 7 presents the quantum defects computed from our $(3s, 3p, 3d)$ energies and from experimental data. Our computed vertical I.P., 9.21 eV, and the experimental adiabatic I.P. 9.246 eV [32], was used. It also includes the *average* quantum defects obtained by fitting the Rydberg formula to many members of each series, as used by Grubb et al. [33] as label of the series.

3.5 Phenol

The results for all phenol singlet states are summarized in Table 8. In contrast to benzene singlets, not much seems to be known about the phenol spectrum.

The vacuum UV singlet spectrum of a monosubstituted benzene is traditionally subdivided into three bands, denoted ^1L_b at about 2600 Å, the ^1L_a at about 2050 Å, and the ^1B band at ca. 1850 Å, see e.g. [37] and references therein. The lowest singlet, the ^1L_b band, for phenol is well known experimentally, and a late

Table 7. Benzene Rydberg states

Transition	δ_3(PT2)	δ_3(Exp)	δ_{av}(Exp)
$^1E_{1g}, 3s \leftarrow e_{1g}$	0.81	0.84	0.76
$^1E_{1u}, 3p_0$	0.42	0.27	0.16
$^1A_{1u}, 3p_1$	0.52	—	0.46
$^1A_{2u}, 3p_1$	0.59	0.58	
$^1E_{2u}, 3p_1$	0.57	0.56	
$^1E_{1g}, 3d_0$	0.12	0.18	0.05
$^1A_{1g}, 3d_1$	− 0.04	− 0.08	− 0.08
$^1A_{2g}, 3d_1$	− 0.11	− 0.08	
$^1E_{2g}, 3d_1$	− 0.07	− 0.08	
$^1B_{1g}, 3d_2$	0.11	—	0.24
$^1B_{2g}, 3d_2$	0.12	—	
$^1E_{1g}, 3d_2$	0.12	—	

Table 8. Phenol singlet states

State	CAS	$\langle z^2 \rangle$	Dipole mom.	f_L^a	PT2	Exp.
Valence states						
$\tilde{X}$	0.00	33.8	(− 0.20, 0.54)			
“$^1B_{2u}$”	4.78	34.2	(− 0.20, 0.56)	0.007(− 86)	4.53	4.51, $f = 0.02$
“$^1B_{1u}$”	7.38	35.5	(+ 0.47,0.39)	0.005(16)	5.80	5.77, $f = 0.13$
“$^1E_{1u}$”	8.76	35.7	(+ 0.24, 0.47)	0.779(− 22)	6.50	6.66, $f = 1.1$
“$^1E_{1u}$”	8.73	36.0	(+ 0.03, 0.46)	0.681(61)	6.56	
“$^1E_{2g}$”	7.78	34.6	(+ 0.09, 0.54)	0.108(− 83)	7.14	
“$^1E_{2g}$”	8.14	34.5	(− 0.43, 0.59)	0.013(8)	7.72	
Ions						
“$^2E_{1g}$”($4\pi^-$)	7.49	30.9	(− 0.42, 0.56)		8.56	8.51
“$^2E_{1g}$”($3\pi^-$)	8.16	30.7	(− 0.84, 0.53)		9.35	
$\pi\pi^*$ *Rydberg states*						
$4\pi^{-1}3p$	6.33	94.6		0.011(− 4)	6.31	
$3\pi^{-1}3p$	6.87	96.4		0.049(− 82)	6.85	
$4\pi^{-1}3d$	6.94	93.7		0.021(50)	6.92	
$4\pi^{-1}3d$	6.97	93.4		0.007(− 17)	6.95	
$3\pi^{-1}3d$	7.48	101.3		0.0002(− 69)	7.53	
$3\pi^{-1}3d$	7.56	101.9		0.002(81)	7.60	
Ions, computed same way as the $\pi\pi^*$ *Rydberg states*						
“$^2E_{1g}$($4\pi^{-1}$)	7.69	30.9	(− 0.42, 0.56)		8.21	8.51
“$^2E_{1g}$($3\pi^{-1}$)	8.36	30.7	(− 0.84, 0.53)		8.87	
$\pi\sigma^*$ *Rydberg states*						
$4\pi^{-1}3s$	5.96	56.1		0.0002	5.76	
$4\pi^{-1}3p$	6.33	53.2		0.0100	6.20	
$4\pi^{-1}3p$	6.51	55.7		0.0048	6.41	
$3\pi^{-1}3s$	6.52	56.1		0.0000	6.42	
$3\pi^{-1}3p$	6.89	53.1		0.0186	6.88	
$4\pi^{-1}3d$	6.94	63.3		0.0011	6.75	
$4\pi^{-1}3d$	6.97	80.8		0.0001	6.95	
$3\pi^{-1}3p$	7.08	56.8		0.0040	7.07	
$4\pi^{-1}3d$	7.09	75.6		0.0015	7.10	
$3\pi^{-1}3d$	7.50	68.0		0.0000	7.57	
$3\pi^{-1}3d$	7.52	79.1		0.0002	7.62	
$3\pi^{-1}3d$	7.69	74.1		0.0000	7.84	

[a] Angular direction in parenthesis. See methods section

reference gives an origin at 36 348.7 cm^{-1} or 4.507 eV [38]. Kimura and Nagakura [39] report the ^1L_a band at 5.82 eV (vertical), and they claim the two components of the ^1B band to have maxima at 6.70 and 6.93 eV, respectively. However, their published spectrum makes an unconvincing case for this decomposition. There seems no reason for the assumed splitting of the ^1B band, and the values quoted would give the spectrum a different appearance from the figure in their article. For the higher excited states, we prefer therefore values obtained as follows: According to Petruska [37], the substituent shifts are fairly well established to be − 1740, − 2100, and − 1700 cm^{-1} for ^1L_b, ^1L_a, and ^1B, respectively. These values, combined with the well-known 0–0 origins of benzene at 4.76, 6.03, and 6.87 eV, gives estimated phenol band origins at 4.54, 5.77, and 6.66 eV. The first agrees well with the known origin at 4.507 eV. The third is then the unresolved components of the "E_{1u}". A small splitting and approximately equal intensities is in general agreement with most substituted benzenes. We have found no literature data on the Rydberg states of phenol.

As expected, the valence states are similar to the benzene states. We note the well-behaved covalent "B_{2u}" state, which is acceptable already on the CASSCF level even with state-averaged orbitals. The ionic "B_{1u}" state is 1.6 eV too high in CASSCF, and the "E_{1u}" states are 2.1 eV too high. These errors are similar to, but larger, than for benzene. The "E_{2g}" states are covalent, but are still not very accurately obtained by CASSCF, presumably since these high states get considerably mixed by a multitude of "exotic" higher states. The higher component has CASSCF and CASPT2 energies similar to those of benzene, while the lower component is sufficiently close to the higher component of the "E_{1u}" states to acquire a fairly large transition moment. This interaction evidently counteracts the static splitting of the "E_{1u}" states.

As mentioned earlier, the final $\pi\pi^*$ Rydberg calculations were performed with the σ orbitals uncorrelated. The dynamic correlation contributions from CASPT2 are therefore much smaller in this case. The I.P. is increased by dynamic correlation by 0.51 eV. In the series of calculations for valence $\pi\pi^*$ excitations, the σ orbitals were not frozen, and the dynamic correlation pushed up the I.P. by 1.06 eV. A closer analysis shows that excitations of type $\sigma \rightarrow \sigma^*$ coupled to excitations within the active space contributes -1.44 eV, while double excitations of type $\sigma\pi \rightarrow \sigma^*\pi^*$ contributes $+2.04$ eV, to the I.P. These large, but partially cancelling, contributions are lacking in the $\pi\pi^*$ Rydberg calculations.

We have to choose between using the ground state or the ion limits as reference for the Rydberg energies. In Table 8, all the phenol states are displayed, and we have chosen, conventionally, to enter the excitation energies from the ground state. However, the quantum defects presented in Table 9 are obtained from

Table 9. Phenol Rydberg states. Ionization energies and quantum defects

Transition	IP(PT2)	n^*	δ_3
$A''3s \leftarrow 4\pi$	2.792	2.21	0.79
$A'3p\pi \leftarrow 4\pi$	1.898	2.68	0.32
$A''3p\sigma \leftarrow 4\pi$	2.352	2.41	0.59
$A''3p\sigma \leftarrow 4\pi$	2.147	2.52	0.48
$A'3d\pi \leftarrow 4\pi$	1.283	3.26	−0.26
$A'3d\pi \leftarrow 4\pi$	1.260	3.29	−0.29
$A''3d\sigma \leftarrow 4\pi$	1.809	2.74	0.26
$A''3d\sigma \leftarrow 4\pi$	1.609	2.91	0.09
$A''3d\sigma \leftarrow 4\pi$	1.452	3.06	−0.06
$A''3s \leftarrow 3\pi$	2.937	2.15	0.85
$A'3p\pi \leftarrow 3\pi$	2.027	2.59	0.41
$A''3p\sigma \leftarrow 3\pi$	2.468	2.35	0.65
$A''3p\sigma \leftarrow 3\pi$	2.282	2.44	0.56
$A'3d\pi \leftarrow 3\pi$	1.342	3.18	−0.18
$A'3d\pi \leftarrow 3\pi$	1.271	3.27	−0.27
$A''3d\sigma \leftarrow 3\pi$	1.781	2.76	0.24
$A''3d\sigma \leftarrow 3\pi$	1.729	2.81	0.19
$A''3d\sigma \leftarrow 3\pi$	1.508	3.00	0.00

the computed ionization energies. These numbers are much more reliable, since the Rydberg states have essentially the same ion cores. Similarly, more reliable excitation energies for the Rydberg states can be obtained by shifting those of Table 8 to reproduce the experimental ionization limits.

Comparing Tables 9 and 7, it is easy to see e.g. how the four $3p\sigma \leftarrow 3,4\pi$ transitions closely correspond to the $3p_1 \leftarrow e_{1g}A_{1u} + A_{2u} + E_{2u}$ transitions of benzene, but it is harder to make the proper correspondence between the phenol $3d\sigma$ and benzene $3d_0, 3d_2$ orbitals. The lower symmetry makes the m quantum number useless, and we use only the letters σ and π to show the symmetry with respect to reflection in the molecular plane.

4 Conclusions

The benzene states are well enough known to serve as a good check of the accuracy of the CASPT2 method. All energies are in almost perfect agreement with experiment, where these are reliably known. The energy of the highest valence $\pi\pi^*$ transition is in perfect agreement with laser flash absorbtion spectra, with a reassignment. Peaks in the spectrum are assumed to show secondary excitations from 1^1B_{2u} to states at 7.8 and 9.2 eV above the ground state [29]. The highest valence state (2^1E_{2g}) has been assigned to the 9.2 eV peak, which is much too high. We assume it to be at 7.8 eV, together with a weaker Rydberg state at slightly lower energy. The computed energies are vertical electronic transition energies. They have been compared to vertical data when possible, else to origins. In the latter case, the error due to relaxation and zero-point vibration energy is unknown, but may be of the same size as our average error, ca. 0.1 eV. The largest error occurs for the allowed Rydberg transition 2^1E_{1u}, which is 0.25 eV too low. This state interacts with the valence 1^1E_{1u} state, which comes out too high. It is tempting to see this as an example of a phenomenon observed in other cases as well: the interaction between a valence and a Rydberg state is erroneously described on CASSCF level, and the internal contraction of the CASPT2 wave function then becomes a serious restriction, deteriorating the description of both states. Experimental literature values for the oscillator strength of the most intense transition, E_{1u}, are scattered in the range 0.6–1.05; our own preferred value is 0.80, obtained from a graphical integration of experimental absolute extinction data, but this figure then includes the $A_{2u}\pi\sigma^*$ Rydberg band, which on the other hand includes very sharp peaks and has been imperfectly integrated. Our calculation gives $f = 0.82$, or 0.87 if the Rydberg transition is added. To the high-energy side, there is an uncertainty in the shape of the tail of the intense band. Thus, the 2^1E_{1u} Rydberg band may be given any value in a broad range, including the calculated 0.058. The computed intensities are thus in as good agreement to experiment as the experimental data permit.

The phenol results are more uncertain. The lowest excited singlet is unproblematic. For the higher states, Kimura and Nagakura [39] make the most definite statement, according to which our computed strongest lines would be in serious error, 0.20 and 0.37 eV too low. However, the quality and even the appearance of their published spectra have motivated us to discard their data and to use instead the substituent shifts to Petruska [37]. There are not enough experimental data of enough precision to properly assess the performance of the CASPT2 method in the phenol case. In most respects, there is no particular reason to believe it to be inferior to the benzene, except that the benzene results appear to be exceptionally good. An important exception is of course the energies of the $\pi\pi^*$ Rydberg states, which were performed with uncorrelated σ orbitals. We believe, partially on physical grounds, and partly from inspecting the size of various dynamic correlation energy contributions, that the defect in these energies can be almost compensated by simply using the computed ionization levels as references, i.e. for Rydberg states, the computed ionization energies should be regarded as the primary obtained quantities rather than the excitation energies above the ground state.

The work described in this paper has taken place during a prolonged period of time. The approach is experimentative, but in the end, the results can be boiled down to a not overly complicated set of rules. Good, specially prepared Rydberg basis sets minimize the interference between valence and Rydberg states. Accurate results can be obtained from state-average calculations, and are then not particularly complicated to handle. A broader account of the performance of the CASSCF/CASPT2 method can be found in a recent review article [14].

Acknowledgements. The research reported in this communication has been supported by a grant from the Swedish Natural Science Research Council (NFR), and also by IBM Sweden under a joint study contract.

References

1. Matos JMO, Roos BO, Malmqvist P-Å (1987) J Chem Phys 86:1458
2. Palmer MH, Walker IC (1989) Chem Phys 133:113
3. Kitao O, Nakatsuji H (1987) J Chem Phys 87:1169
4. Roos BO, Andersson K, Fülscher MP (1992) Chem Phys Lett 192:5
5. Andersson K, Blomberg MRA, Fülscher MP, Kellö V, Lindh R, Malmqvist P-Å, Noga J, Olsen J, Roos BO, Sadlej AJ, Siegbahn PEM, Urban M (1992) Widmark P-O MOLCAS Version 2 User's Guide. Dept. of Theor. Chem., Chem. Center, Univ. of Lund, Lund
6. Roos BO, Taylor PR, Siegbahn PEM (1980) Chem Phys 48:157
7. Roos BO (1980) Int J Quantum Chem S14:175
8. Fülscher MP, Andersson K, Roos BO (1992) J Phys Chem 96:9204
9. Serrano-Andrés L, Merchán M, Nebot-Gil I, Lindh R, Roos BO (1993) J Chem Phys 98:3151
10. Serrano-Andrés L, Merchán M, Nebot-Gil I, Roos BO, Fülscher MP (1993) J Am Chem Soc 115:6184
11. Wolinski K, Pulay P (1989) J Chem Phys 90:3647
12. Andersson K, Malmqvist P-Å, Roos BO, Sadlej AJ, Wolinski K (1990) J Phys Chem 94:5483
13. Andersson K, Malmqvist P-Å, Roos BO (1992) J Chem Phys 96:1218
14. Roos BO, Fülscher MP, Malmqvist P-Å, Serrano-Andrés PL, Merchán M (1994) Theoretical studies of the electronic spectra of organic molecules. In: Langhoff SR (ed) Quantum mechanical electron structure calculations with chemical accuracy. Kluwer, Dordrecht
15. Tamagawa K, Iijima T, Kumura M (1976) J Mol Struct 30:243
16. Stoichev BP (1954) Can J Phys 32:339
17. Larsen NW (1979) J Mol Struct 51:175
18. Bock CW, Trachtman M, George P (1986) J Mol Struct (Theochem) 139:63
19. Lindh R, Roos BO (1989) Int J Quantum Chem 53:813
20. Widmark P-O, Malmqvist P-Å, Roos BO (1990) Theor Chim Acta 77:291
21. Almlöf J, Taylor PR (1987) J Chem Phys 86:4070
22. Lorentzson J, Malmqvist P-Å, Roos BO (1994) to be published
23. Mulliken RS (1955) J Chem Phys 11:1997
24. Hiraya A, Shobatake K (1991) J Chem Phys 94:7700
25. Lassettre EN, Skerbele A, Dillon MA, Roos KJ (1967) J Chem Phys 48:5066
26. Johnson PM, Korenowski GM (1983) Chem Phys Lett 97:53
27. Wilkinson, PG (1956) Can J Phys 34:596
28. Nakashima N, Sumitani M, Ohmine I, Yoshihara K (1980) J Chem Phys 72:2226

29. Nakashima N, Inoue H, Sumitani M, Yoshihara K (1980) J Chem Phys 73:5620
30. Whetten RL, Fu K-J, Grant ER (1983) J Chem Phys 79:2620
31. Whetten RL, Grubb SG, Otis CE, Albrecht AC, Grant ER (1985) J Chem Phys 82:1115
32. Grubb SG, Whetten RL, Albrecht AC, Grant ER (1984) Chem Phys Lett 108:420
33. Grubb SG, Otis CE, Whetten RL, Grant ER, Albrecht AC (1985) J Chem Phys 82:1135
34. Johnson PM (1976) J Chem Phys 64:4143
35. Doering JP (1969) J Chem Phys 51:2866
36. Minaev BF, Knuts S, Ågren H, Vahtras O (1993) Chem Phys 175:245
37. Petruska J (1961) J Chem Phys 34:1120
38. Martinez III SJ, Alfano JC, Levy DH (1992) J Mol Spectrosc 152:80
39. Kimura K, Nagakura S (1965) Mol Phys 9:117
40. Rancurel P, Huron B, Praud L, Malrieu JP, Berthier G (1976) J Mol Spectros 60:259
41. Hay PJ, Shavitt I (1973) Chem Phys Lett 22:33