

# CASSCF calculations for photoinduced processes in large molecules: Choosing when to use the RASSCF, ONIOM and MMVB approximations

Michael J. Bearpark<sup>a,\*</sup>, Francois Ogliaro<sup>b</sup>, Thom Vreven<sup>c</sup>, Martial Boggio-Pasqua<sup>a</sup>,  
Michael J. Frisch<sup>c</sup>, Susan M. Larkin<sup>a</sup>, Marc Morrison<sup>d</sup>, Michael A. Robb<sup>a</sup>

<sup>a</sup> Department of Chemistry, Imperial College London, London SW7 2AZ, UK

<sup>b</sup> Equipe de Chimie et Biochimie Théoriques, UMR 7565 - CNRS, Université Henri Poincaré, Nancy 1,  
BP 239, 54506 Vandoeuvre-lès-Nancy Cedex, France

<sup>c</sup> Gaussian, Inc., 340 Quinipiac St Bldg 40, Wallingford, CT 06492, USA

<sup>d</sup> Physical & Theoretical Chemistry Laboratory, South Parks Road, Oxford OX1 3QZ, UK

Received 13 March 2007; received in revised form 4 May 2007; accepted 5 May 2007

Available online 22 May 2007

## Abstract

In this article, we compare and contrast the RASSCF, ONIOM and MMVB electronic structure methods for calculating relaxation paths on potential energy surfaces of the excited states of large molecules, and for locating any resulting conical intersections at which nonadiabatic decay can take place. Each method is treated here as an approximation to CASSCF, which we choose as our reference level of theory, but which becomes prohibitively expensive computationally for large molecules. Both MMVB and ONIOM are hybrid computational methods – combining different levels of theory in an energy plus derivatives calculation at a particular molecular geometry – but they differ fundamentally in that MMVB is a hybrid-atom method, whereas ONIOM is a hybrid-molecule method. We explain this distinction through four representative applications: the photostability of pyracylene (studied with CASSCF, RASSCF, MMVB); large geometry changes in the singlet excited states of triangulene (studied with MMVB); a model for interstitial nickel defects in a synthetic diamond lattice (studied with ONIOM CAS:UFF); and the photochemical [4 + 4] cycloaddition of cyclohexadiene to naphthalene (studied with ONIOM CAS:MMVB). We show that each method is more appropriate for a particular type of photochemical problem. This article is part perspective, part review, and contains new results for three multi-state or photoinduced processes in complex systems.

© 2007 Elsevier B.V. All rights reserved.

**Keywords:** CASSCF; RASSCF; ONIOM; MMVB; Excited states; Hybrid method; QM/MM; Pyracylene; Triangulene; Diamond defect; Photochemical cycloaddition

## 1. Introduction

Reaction paths that cross multiple potential energy surfaces of small molecules have now been widely studied in detail, both experimentally and computationally. There are now so many examples that it would be misleading to single out any in particular, although their importance has been recognized [1], and applications range from atmospheric chemistry to materials and synthetic chemistry, and biochemistry. However, a challenge for computational chemistry at present is to study the same non-

adiabatic processes [2] in larger molecules [3], or molecules in a realistically modeled environment such as a protein [4], to the same degree of accuracy.

The reliable computational study of reaction paths and other processes from the excited states of molecules currently requires the use of methods for which the computational costs scale unfavorably with the size of the system. This is particularly the case for multi-reference descriptions, which are needed when surface crossings are involved, for example. A widely used method for studying such processes is complete active space SCF (CASSCF), which we consider as the reference method in this article. CASSCF gives a reliable first-order description of molecular excited states [5], with some weaknesses that are well understood, and which can be systematically corrected

\* Corresponding author. Tel.: +44 20 7594 5727.

E-mail address: [m.bearpark@imperial.ac.uk](mailto:m.bearpark@imperial.ac.uk) (M.J. Bearpark).

(at least in principle). Moreover, CASSCF has analytic energy gradients for excited-state geometry optimization and locating crossing points [6] (conical intersections [7–10]) between potential energy surfaces, which have proved to be important for understanding rapid radiationless deactivation of molecular excited states.

The purpose of this article is to compare and contrast computational methods developed in our research groups that behave like CASSCF, but with a significantly reduced computational cost. Because at a qualitative level (but preferably at a quantitative level as well) the methods must reproduce CASSCF, the theoretical foundation of each is directly related to that of CASSCF. We focus mainly on hybrid methods, which involve *partitioning* a molecule, such that different chemical regions are treated with different but appropriate levels of theory and accuracy. Typically in this approach, a small part of a molecule, such as a chromophore or reaction centre, is treated with a more accurate (but computationally more expensive method) than the rest, which avoids ‘wasting’ accuracy where it is unnecessary. Using a hybrid method for a series of calculations may save time overall, even if additional calibration steps are necessary. In fact, for many larger molecules, it may be the only way to make meaningful calculations possible, since CASSCF becomes prohibitively expensive quickly with increasing size of system and/or number of active orbitals (as we illustrate in more detail in Section 2.1).

Before continuing, we give a concise definition of a hybrid method and how it differs from a ‘conventional’ method. In the remainder of this article we will expand upon these definitions and explore the implications and subtleties in more detail.

- A conventional method treats all of a molecule at a single level of theory. Some conventional methods that we will discuss (such as CASSCF) are deliberately biased towards particular regions of a molecule (by, *e.g.* selection of active orbitals and electron configurations) but in general they need not be.
- A hybrid method treats a molecule at multiple levels of theory. The molecule is partitioned into different parts, and each part is calculated with a distinct computational method. The partition can be based on identifying a chromophore or reaction centre (*hybrid molecule*, Fig. 1, *e.g.* ONIOM [11–15], QM/MM [16–18]) or by treating some electron–electron interactions between some atoms differently (*hybrid atom*, Fig. 2, *e.g.* MMVB [19–24]). Hybrid methods are always biased towards a specific part of a molecule or system.

In addition to hybrid approaches to reduce the computational cost of CASSCF, we will also discuss the restricted active space SCF (RASSCF) method [25–27]. According to the definitions above, RASSCF is a conventional method, and reduces the cost of CASSCF computations by reducing the number of electron configurations that are considered.

In a very different type of hybrid method, the level of theory itself is separated into different contributions. These methods are referred to as *Compound Model Chemistries*, and examples are Pople et al.’s G [28–30], Petersson and co-workers’ CBS [31–34], and Martin et al.’s W [35,36] series of methods. These methods are in practice more similar to conventional methods

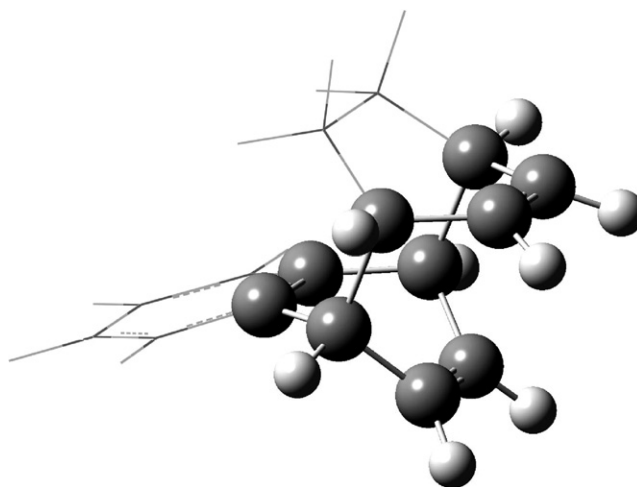


Fig. 1. The partitioning chosen here for modeling the [4+4] photochemical cycloaddition of cyclohexadiene to naphthalene using ONIOM. The model region (treated at a higher level of theory) is drawn as ball-and-stick; the remaining real system (treated at a lower level of theory) is drawn in wireframe. The doubly bonded  $S_1$  structure is shown.

than to hybrid methods, since the resulting level of theory is applied to the entire system. Using our terminology of *hybrid molecule* and *hybrid atom*, the G2 and CBS methods could be referred to as *hybrid theory* methods. Because compound model chemistries are not specifically designed to mimic CASSCF or for calculating excited states (although there are exceptions [37]), we will not discuss these methods any further here.

In Scheme 1, we summarize for reference the hybrid and conventional computational methods we will be concerned with in this article, and the relationships between them. On the vertical axis, we plot the accuracy of the *method*. It is important to realize that this does not necessarily correspond to the accuracy of the *calculation*: in the examples presented in this article, we will demonstrate that calculations using hybrid methods can give results that are very close to conventional methods, despite the

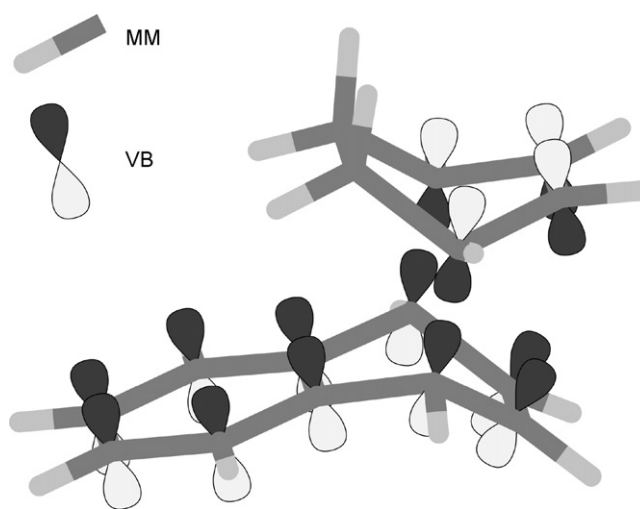
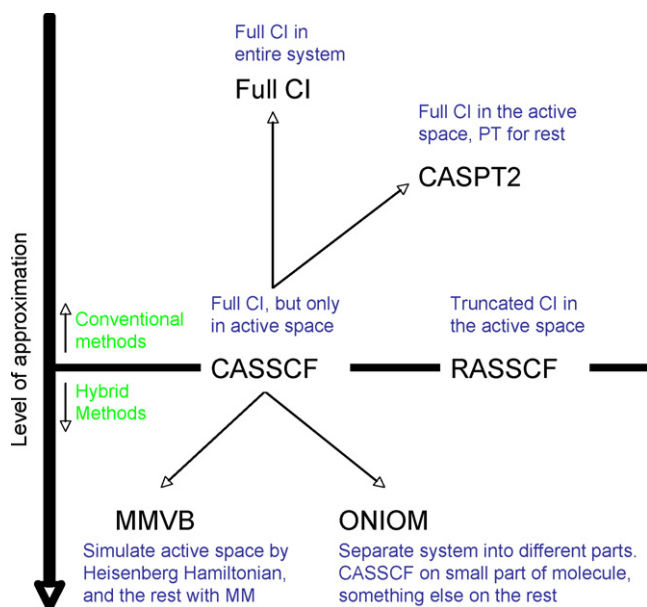


Fig. 2. The MMVB partitioning chosen here for modeling the [4+4] photochemical cycloaddition of cyclohexadiene to naphthalene. This corresponds to the ONIOM real system for this study (drawn in wireframe in Fig. 1).



Scheme 1.

use of a (much) less accurate level of theory for a large part of the system. Besides the methods mentioned in Scheme 1, there are various other methods available for the study of excited states of large molecules, but these are less appropriate for the types of chemistry that we are interested in. Therefore, we will not provide a comprehensive review of computational methods for excited states of large molecules: for a recent such review, focusing on non-hybrid methods, see for example [3].

In what follows, there will be two recurring themes: what *truncation* or approximations to CASSCF can be made? And how can this be translated into partitioning a molecule or system into different parts? Through the examples, we will show that the answers to these questions depend on the system that is being studied and the ‘chemical’ question being asked. We present three new excited-state applications in this article, each illustrating a different conceptual approach to partitioning a molecule, together with suggestions for where each is appropriate.

## 2. Methods

### 2.1. CASSCF as a reference for studying molecular excited states

We have recently reviewed the types of mechanistic information that can be obtained from calculations on molecular excited states, by treating conical intersections as both ‘reactive intermediates’ [38] having a well-defined electronic structure, and as the bridge geometrically between excited and ground state reaction pathways [39]. Further details are also available on the practical aspects of carrying out calculations such as those through which these concepts were developed [5].

To study excited-state reactivity, minima, transition structures, and conical intersections away from the vertical excitation (Franck–Condon) region of the ground state must be characterized. From this, it follows that the computational method used

must be able to describe both ground and excited states in a balanced way at both vertical and relaxed geometries, including regions where two (or more) electronic states are degenerate. Furthermore, our view is that analytic energy gradients and the ability to fully optimize crossing points are both essential for this, to avoid relying on limited potential energy surface scans which can introduce artificial constraints and lead to important mechanistic features being missed. Energy second derivatives are also desirable, both as a way of determining the nature of critical points found on a single potential energy surface, and also because of the potential insights that can now be gained [40–42] from characterizing critical points of conical intersection by evaluating second derivatives of the seam itself.

Clearly there is no single computational method that fulfills all of the above criteria at present, particularly for large molecules with many excited states. We choose CASSCF calculations as a compromise or intermediate reference point: they allow us to optimize excited-state geometries [6,43] and determine the electronic ‘character’ [44] of the states we obtain, and they also make it possible to run selected trajectory calculations to explore regions of extended surface crossing seams away from their minima. This aspect of our work is discussed more fully in [39,45]; recent results of dynamics studies and those of others [46–48] caution us against relying solely on static critical points determined on a potential energy surface for mechanistic information. CASSCF also provides a starting point for more accurate calculations, such as CASPT2 or MRCI [49–52].

Rather than being restricted to a single reference electronic configuration (typically with orbitals either doubly occupied or unoccupied), the orbitals in CASSCF are partitioned into core, active and virtual. The electrons occupying the active orbitals (grouped around the HOMO/LUMO) give rise to a linear expansion of multiple electronic configurations: in CASSCF, all allowed arrangements of active electrons in active orbitals are considered, and hence the resulting wavefunction can have fractional occupations (typically from 1.999 to 0.001) of these orbitals. The choice of active space is crucial to the success of the method, both practically (will the wavefunction converge?) and conceptually (what chemical process is being described?) Conceptually, the key to the work we describe is that the active space includes orbitals whose occupations change during a chemical process, such as excitation to a different state, or relaxation to a region of a particular potential energy surface with different electronic character. For example, in benzene [53–55], a CASSCF calculation would typically include 6 electrons in the  $6\pi$  orbitals with nodes above and below the plane of the molecule. This proves adequate to describe  $\pi \rightarrow \pi^*$  valence excited states and, *e.g.* search for the  $S_1/S_0$  conical intersection responsible for ‘channel 3’ decay. However, this active space is less able to describe excited states for which charge transfer is important, and inadequate to describe Rydberg excited states. Because the choice of active space depends on the process being described, we take the view that there is no single ‘correct’ active space for a particular molecule.

A key problem with CASSCF calculations is scaling with system size: there is a factorial dependence on both the number of active electrons and particularly on the number of active

orbitals generating many-electron configurations (full CI within the active space). This is much more severe than any dependence on the number of one-electron basis functions. The number of Slater determinants in a full configuration interaction (CI) calculation is:

$$n_{\text{Slater}} = \binom{M}{N_{\alpha}} \binom{M}{N_{\beta}} \quad (1)$$

where  $M$  is the number of active orbitals,  $N_{\alpha}$  and  $N_{\beta}$  the numbers of active electrons with  $\alpha$ - and  $\beta$ -spins, respectively, and the quantities in parentheses are binomial coefficients:

$$\binom{a}{b} = \frac{a!}{b!(a-b)!} \quad (2)$$

If spin-adapted configurations – also called configuration state functions (CSFs) – are used, then the number of configurations is given by the Weyl–Robinson formula [56]:

$$n_{\text{CSFs}} = \frac{2S+1}{M+1} \binom{M+1}{M-(N/2)-S} \binom{M+1}{(N/2)-S} \quad (3)$$

where  $S$  is the total spin, and  $N$  is the total number of active electrons.

Note that both  $n_{\text{Slater}}$  and  $n_{\text{CSFs}}$  correspond to a number of configurations for a given value of  $S_z$ . Of course, the presence of point-group symmetry could be used to reduce the number of Slater determinants and CSFs from these values.

For example, in a CAS(12, 12) singlet wavefunction, the numbers of Slater determinants and CSFs are 853,776 and 226,512, respectively. Similarly, in a CAS(14, 14) singlet wavefunction (see Sections 3 and 6), these numbers increase to 11,778,624 and 2,760,615, respectively. In a CAS(22, 22) singlet wavefunction (*c.f.* Section 4), these numbers become exceedingly large: the number of Slater determinants is  $5.0 \times 10^{11}$ , whereas the number of CSFs is  $8 \times 10^{10}$ . In a CAS(22, 22) triplet wavefunction, the number of CSFs becomes  $1.7 \times 10^{11}$ .

The practical limit for CASSCF is therefore currently around 16 active electrons while making use of symmetry, and these calculations are expensive, as one-electron (molecular orbital) and many-electron (configuration) expansion coefficients must be optimized to determine the wavefunction. Consequently, in the past, it has proved necessary to truncate systems to be studied. An example is our series of studies on various models for the retinal protonated Schiff base [57–59]. An alternative approach would be to truncate the active space, although the results of this for a  $\pi$  system can be unpredictable, due to instabilities [60–62]. For example, trying to represent a 12-electron  $\pi$  system with only 6 active  $\pi$  orbitals led to symmetry breaking for indacene [63].

## 2.2. RASSCF

In systems having a large number of active orbitals and electrons, CASSCF wavefunctions may contain an unmanageable number of electron configurations generated by the full CI expansion. In this context, the restricted active space SCF

method [26] can be used to reduce this number by restricting the excitations in the wavefunctions. This reduction is achieved by subdividing the active space into three categories: a set of orbitals with a limited number of vacancies (called the RAS1 space), a fully active orbital set (RAS2), and a set of orbitals with a limited number of electrons (RAS3). Thus, RASSCF wavefunctions use a truncated CI expansion (see Scheme 1), and because only a few holes and particles are allowed in RAS1 and RAS3, respectively, the use of RASSCF wavefunctions strongly reduces the number of electron configurations in the case of very large active spaces. However, a sensible choice for the important subspace RAS2 (kept as small as possible since full CI is used in that space) must be made to ensure a balanced description of all of the electronic states of interest.

The formulae for calculating the total number of electronic configurations for RASSCF are more complicated than those for CASSCF given in Section 2.1, so they are not reproduced here, but are discussed in detail in [64].

It is worth noting in passing that RASSCF can be used, not only to reduce the number of electron configurations, but also to increase it. Indeed, it is possible to use the full active space (CAS) as RAS2 (provided it is not too large), and to include further orbitals in RAS1 and RAS3. This enlargement of the active space is useful to include dynamic correlation effects that are missing at the CASSCF level [65], but this approach is outside the context of this study, and thus will not be developed further here.

## 2.3. Hybrid methods: truncation and partitioning

As outlined above, the advances in CASSCF allow us to calculate and investigate the relevant potential energy surfaces reliably, but the calculations are time-consuming and resource-hungry, and become more so disproportionately as the size of the target molecule increases. This problem exists in some form for most computational methods, and has been partly addressed for ground state calculations by, *e.g.* developments in linear scaling [66] density functional theory (DFT) for large molecules. Although these methods have been extended to compute excitation energies [67], they cannot yet describe excited-state potential energy surfaces in general [68]. Improved scaling schemes are not applicable to CASSCF, and no alternative (better scaling) method is available that can describe potential surfaces with the same accuracy and quality.

To reduce the computational cost of accurate calculations, one can follow either of the two traditional approaches: either use a reduced, lower level of theory to treat the full real system; or perform higher-level calculations on a truncated model representative of the real system [69]. In other words, we either truncate/approximate the computational method we would like to use, or truncate/approximate the molecule we wish to study at the desired level of theory. Both approaches can be unsatisfactory, and as we will show in this article, unnecessary. We will show how hybrid methods are often able to minimise errors by combining the best of both types of truncation.

Hybrid methods avoid using the same method for the whole molecule, and savings are made by treating different parts of

a molecule only as accurately as necessary. The molecule is divided up into a reactive model system, which is treated at a higher level of theory, while the remainder of the molecule is treated at a much cheaper lower level of theory. This ensures an appropriate level of accuracy for different regions of the system, and reduces the computational cost compared to the full calculation by limiting the expensive high-level calculation to the necessary model region. Fig. 1 shows an example of such a partition, for the cyclohexadiene + naphthalene cycloaddition described later in this article. Hybrid methods thus truncate both the level of theory and the size of the (high-level part of the) system, but by integrating the truncations into one calculation, attempt to reproduce the high-level calculation on the full system. Most hybrid methods combine quantum mechanical with molecular mechanical methods (QM/MM) [16–18,70,71]. The application of such methods to excited-state chemistry, however, is still much less understood than the application to ground states. In addition, hybrid QM/MM methods are not always the best choice, and we will consider QM/QM and other hybrid methods as well.

Hybrid methods rely on the concept that many reactions occur in a localized region of the molecule, while the remainder of the system plays a more minor (sometimes purely structural) role, but nevertheless its inclusion provides a more accurate description of the reactivity and the reaction mechanism. Well known examples of this (for ground electronic states) are the active sites of enzymes, which facilitate reactions by lowering activation energies, but the surrounding protein environment must still be considered since the tertiary structure of the folded protein influences the catalytic reactivity of the enzyme [72]. Although all hybrid methods are united by this fundamental concept, their formulation differs in some specific ways, which will be discussed in the following sections.

Extensive reviews of hybrid methods include those by Gao [73] and Sherwood [74]. As an example, the visual pigment rhodopsin has received much attention concerning its initial excitation and energy storage, which is facilitated by the surrounding protein. Although an excited-state description is necessary for the  $S_1$  state, there are at present few QM/MM studies employing a full *ab initio* treatment at the QM level because a high-level treatment of the chromophore is still computationally demanding, and many studies employ reduced chromophore models or protein environments. Nevertheless the *cis-trans* isomerization of the chromophore has been studied at the CASSCF and CASPT2//CASSCF [75–79] levels of theory and with ROKS [80] methods, where in each case force fields methods were used to describe the protein cavity. ONIOM(QM:MM) (to be discussed in more detail later) studies have also combined TDDFT methods with the protein modelling force field AMBER to investigate this system [81,82].

Although we choose CASSCF as a reference level of theory for the present, we want to note that the discussion of hybrid methods applied to excited states is more about different ways of partitioning a molecule for calculation than the particular reference level of theory chosen. Many of the same considerations would apply if the reference were chosen to be a more accurate [83] or more approximate [3,68,84] level of theory instead.

### 2.3.1. MMVB

The molecular mechanics with valence bond (MMVB) method is described in detail elsewhere [20,22–24]. The original aim [23] was to simulate CASSCF calculations for the ground and excited states of conjugated hydrocarbons, and to generate starting points for CASSCF geometry optimizations. MMVB has now outgrown this limited aim – for which the geometries were more than good enough – because the calculated relative energies were also often surprisingly good [19] and thus the method was usable in its own right. It is now implemented in a development version of Gaussian [85].

MMVB [22,23] is a hybrid quantum/molecular mechanics (QM/MM) method, which currently uses the MM2 potential [86,87] to describe a  $\sigma$ -bonded molecular framework. The active electrons – those involved in bond reorganization – are represented by a Heisenberg Hamiltonian [88,89]. For a particular choice of active space, the subset of electronic configurations with singly occupied localized orbitals is therefore used, and hence MMVB cannot properly represent ionic states, which are dominated by charge transfer terms, but works well for covalent excited states. MMVB can treat  $\pi$  systems with up to around 30 active electrons at present, and is many orders of magnitude faster than CASSCF.

MMVB differs from most other QM/MM methods in that the quantum and molecular mechanics parts are both present on certain (hybrid) atoms, and not separated into different (hybrid) QM and MM regions of a molecule. In other words, MMVB is a *hybrid-atom* method (Fig. 2). A VB carbon atom in MMVB has one free valence at present: some parameters in the potential for this atom were obtained from transformed CASSCF wavefunctions for small model systems [24]; some others were modified from the standard MM2 [87] force field. However, these modified parameters are interdependent, and geometry dependent.  $\sigma$ – $\pi$  separation is not enforced, and the interaction between  $\sigma$  and  $\pi$  electrons is treated at a similar level to CASSCF for the molecules studied here: the underlying  $\sigma$  framework can adapt to reorganization of the  $\pi$  system and is not fixed.

MMVB currently has a number of disadvantages. The first is that parameters must be generated for elements with ‘active’ orbitals, and we currently implement only  $sp^2/sp^3$  carbon atoms this way (although inactive sites within the sigma bonded framework can include any atom for which MM2 is defined, as in, *e.g.* ergosterol [90]). We also use a general set of parameters that is not fitted for any particular molecule (although we have experimented with this [24]) and which is assumed to be transferable between molecules, which introduces some additional approximations. This, on the other hand, avoids the need to construct the full CASSCF wavefunction for the system under investigation first, which would be required for parametrization/benchmarking purposes if the parameters were not transferable.

However, MMVB is based on the idea that any CASSCF wavefunction can be transformed into (*i.e.* projected onto [44,89]) a VB wavefunction via construction of an effective Hamiltonian [88]. Malrieu and co-workers recognized that the resulting VB integrals, Coulomb ( $Q$ ) and exchange ( $K$ ), had a distance and orientation dependence that could be fitted [91,92].

Our implementation [23] extended this original idea by fitting the exchange integrals additionally as a function of ‘hybridization’ ( $sp^2/sp^3$ ) of localized p orbitals (one per active site), and also by recognizing that much of the total Coulomb energy is provided by a standard molecular mechanics potential [93] (MM2 [86]). Consequently, MMVB can be used to describe the formation of new  $\sigma$ -bonds between ‘active’ sites.

The resulting method [22,23] is many orders of magnitude faster than the CASSCF method it was designed to simulate, because expensive integral evaluation has been completely replaced by lookup from analytic fitted functions. (In other words, we construct an effective Hamiltonian from parameters derived originally from CAS calculations on small prototype systems, rather than explicitly for the particular molecule being studied.) Because there is a Hamiltonian and hence an eigenvalue problem to solve, MMVB can describe excited states and bond reorganization using an ‘active space’ of orbitals chosen in the same way as for CASSCF. However, there is a significant speedup here as well, as in the VB representation there are fewer electron configurations to consider. Comparing to Section 2.1, the number of VB perfect pairing configurations (all active orbitals singly occupied) is given by:

$$n_{\text{Slater,MMVB}} = \frac{N!}{(N/2)!(N/2)!} \quad (4)$$

For 22 active electrons, there are 705,432 such configurations, many orders of magnitude less than for CASSCF. Furthermore, the necessary matrix elements are easier to evaluate [20].

In practice, MMVB can be thought of as a general modified MM force field method, with additional QM ‘springs’ and a new carbon (hybrid) atom type, which can treat excited states and bond reorganization, and can be used to optimize conical intersections. MMVB has been implemented in a development version of Gaussian [85], using an MM2 [87] force field based as closely as possible as that implemented in Tinker [94,95].

### 2.3.2. ONIOM

Morokuma and co-workers introduced the computational method ONIOM: “Our own  $N$ -layered Integrated molecular Orbital + molecular Mechanics method” [11,12,96,97]. This method generalizes the QM/MM idea of appropriate levels of theory and accuracy for different regions of a chemical system (Fig. 1). The main attraction of ONIOM is its generality. Unlike QM/MM methods, any number of levels of theory can be combined in ONIOM, although the current implementation is limited to three [98]. In addition, the lowest level of theory does not need to be molecular mechanics, allowing for QM/QM combinations. Of particular interest in the current context is that methods appropriate for excited states can be combined. Problems with determining the coupling between layers (inherent in other QM/MM methods that are additions rather than extrapolations [99]) are avoided, and energy derivatives are uniquely defined for the specific combination of levels of theory chosen.

The ONIOM(QM:MM) method can be regarded as a general QM/MM method, and has been used in a number of studies on excited states [14,82,100]. The excitation is then, however, forced to be localized in the QM region of the model. This restric-

tion is lifted when two or more QM methods are combined in ONIOM. Despite this attraction, the ONIOM(QM:QM) method has not yet been widely applied in excited-state calculations. To date, there are few published studies [14]; none of these have fully optimized a conical intersection between potential energy surfaces, and there are several reasons for this, which we partly address here:

- Which methods can (or should) be combined for a reliable description of excited states?
- Do the methods for all ONIOM layers have to be able to describe excited states, or will classical force fields (molecular mechanics) be sufficient for the outermost layer?
- Where should the disconnection between layers be made? Must all of the electronic excitation be described using the most accurate method?

ONIOM is not a ‘black box’ method—each of the (presently open) questions above requires a choice, and as Morokuma and co-workers have stated [11] there are no restrictions on making a bad choice. Though experience for ground states is accumulating, there are currently few guidelines for excited-state ONIOM calculations.

Before discussing the theory of ONIOM, we want to stress the difference between the *hybrid-atom* approach of MMVB and the *hybrid-molecule* approach of ONIOM (and most other QM/MM methods). This is illustrated with Figs. 1 and 2. Fig. 1 represents ONIOM for naphthalene + cyclohexadiene. An entire region of the system, represented by the ball-and-stick display style, is treated at the higher level of theory, while the remainder, represented by wireframe, is only treated at a lower level of theory. All the centers in either region, including the hydrogen atoms, are treated at the respective level of theory. In this particular example there are only four connections between the regions, but this can be as few as one (for example small organics [13]) to as many as tens or hundreds (for example in zeolite models [101]). Fig. 2 represents MMVB for the same system. Some centers only contribute to the potential through MM terms (represented by the tube display style), while others contribute through both MM terms and VB terms (represented by tube display and orbitals). In this case, the entire  $\pi$ -space is treated by VB. There are no specific places where there is a ‘cut’ between VB and MM: both levels of theory overlap and are applied to the same parts of the system. Which QM/MM approach, *hybrid-atom* or *hybrid-molecule*, is preferred, depends on the system and chemistry under investigation. We will comment on these aspects in the remaining sections. In fact, as we will show later, the approaches are not mutually exclusive, and can be applied in one calculation simultaneously.

The ONIOM energy is obtained through an extrapolation, and can be written as:

$$E^{\text{ONIOM}} = E_{\text{model}}^{\text{high}} + E_{\text{real}}^{\text{low}} - E_{\text{model}}^{\text{low}} \quad (5)$$

The *model* is a small fragment of the full *real* system, as illustrated in Figs. 1 and 9. The model represents the part of the system where the chemical process under investigation takes

place, and is ‘cut’ out of the real system. ONIOM assumes that the different parts of the system play different roles in the process, and can therefore be treated with different levels of theory, *high* and *low*. Through Eq. (5), ONIOM attempts to reproduce a calculation at the high level of theory for the full system. We define this calculation as the *target*. The absolute energy  $E^{\text{ONIOM}}$  is not expected to reproduce the absolute target energy  $E_{\text{real}}^{\text{high}}$ , but relative energies will be reproduced:

$$\Delta E^{\text{target}} = \Delta E_{\text{real}}^{\text{high}} \approx \Delta E^{\text{ONIOM}} \quad (6)$$

Examples of relative energies are the energy differences between critical points on a potential surface or between electronic states at a particular geometry.

From the equations above and Fig. 1, there are two interpretations of how ONIOM works:

- Adding substituent effects ( $E_{\text{real}}^{\text{low}} - E_{\text{model}}^{\text{low}}$ ) to a *high*-level calculation on the *model* system.
- Improving a *low*-level calculation on the *real* system in the region of the *model*, using ( $E_{\text{model}}^{\text{high}} - E_{\text{model}}^{\text{low}}$ ).

From Eq. (5), it is clear that ONIOM assumes that the effects of the substituents or changing fragment size, given by ( $E_{\text{real}}^{\text{low}} - E_{\text{model}}^{\text{low}}$ ), and level of theory, given by ( $E_{\text{model}}^{\text{high}} - E_{\text{model}}^{\text{low}}$ ), are separable. Based on our experience to date, we favour the first interpretation: ONIOM is useful in that it saves computational time and resources, by limiting expensive/slow accurate (*high*-level) calculations to a small molecule fragment, where they are essential. The surroundings can be described by much cheaper/faster (*low*-level) computational methods, which may give poor results on their own, yet have a beneficial effect on the fragment calculation. Even though three energy calculations are performed for a single ONIOM point in a two-level calculation, none is computationally as severe as the target calculation would be.

In many cases there is bonded interaction between the layers in ONIOM. This results in open valencies, which we saturate with (hydrogen) link atoms in the *model* system. All three terms in Eq. (5) then involve a chemically reasonable system, which allows the *low* level to be either QM or MM. This is in contrast with most other hybrid methods, which only allow QM/MM combinations. Since the geometry of the *model* is derived from the *real* system, the gradients and higher derivatives are well-defined, and can be obtained with equations similar to that for the energy:

$$\frac{\partial E^{\text{ONIOM}}}{\partial q} = \frac{\partial E_{\text{model}}^{\text{high}}}{\partial q} + \frac{\partial E_{\text{real}}^{\text{low}}}{\partial q} - \frac{\partial E_{\text{model}}^{\text{low}}}{\partial q} \quad (7)$$

When the ONIOM potential is used in geometry optimization or other ways to explore the potential surface, we use the integrated energy and derivatives from Eqs. (5) and (7) [15,102]. Because the individual wave functions are not directly coupled, Eq. (7) still involves three independent terms. The wavefunctions nevertheless indirectly affect each other through the geometry.

As mentioned above, ONIOM can be used to describe excited states and excitation energies. We can write the latter as the

difference of the ONIOM energies of the two states.

$$\begin{aligned} \Delta E^{\text{ONIOM}} &= E^{*,\text{ONIOM}} - E^{\text{ONIOM}} \\ &= (E_{\text{real}}^{*,\text{high}} + E_{\text{real}}^{*,\text{low}} - E_{\text{model}}^{*,\text{low}}) - (E_{\text{model}}^{\text{high}} + E_{\text{real}}^{\text{low}} - E_{\text{model}}^{\text{low}}) \\ &= (E_{\text{model}}^{*,\text{high}} - E_{\text{model}}^{\text{high}}) + (E_{\text{real}}^{*,\text{low}} - E_{\text{real}}^{\text{low}}) - (E_{\text{model}}^{*,\text{low}} - E_{\text{model}}^{\text{low}}) \\ &= \Delta E_{\text{model}}^{\text{high}} + \Delta E_{\text{real}}^{\text{low}} - \Delta E_{\text{model}}^{\text{low}} \end{aligned} \quad (8)$$

Eq. (8) shows that the ONIOM excitation energy is obtained from the excitation energies of the three sub-calculations, similar to Eq. (5) for the energy. One could argue that when the excitation is localized in the *high*-level region, the ONIOM excitation energy  $\Delta E^{\text{ONIOM}}$  can be approximated by  $\Delta E_{\text{model}}^{\text{high}}$ . This is, however, not necessarily possible, since the *low*-level region may interact differently with the two states that are involved. In the current work we always use the appropriate states for all the terms. In other work we investigate the approximation of ground-state-only methods for the low level of theory [100].

Recently we developed a Conical Intersection search algorithm for ONIOM [100], based on the standard conical intersection search method for CASSCF and MMVB [6]. In this case we assume that the difference between the states can be entirely described by the  $\Delta E_{\text{model}}^{\text{high}}$  term only. This modified gradient in the search algorithm can then be written as

$$\bar{\mathbf{g}}^{\text{ONIOM}} = 2(E_{\text{model}}^{*,\text{high}} - E_{\text{model}}^{\text{high}}) \frac{\mathbf{x}_{\text{model}}^{\text{high}}}{|\mathbf{x}_{\text{model}}^{\text{high}}|} + \mathbf{P} \frac{\partial \overline{E^{*,\text{ONIOM}}}}{\partial \mathbf{q}} \quad (9)$$

$\mathbf{x}$  is the vector that describes the gradient difference between the two states, and  $\mathbf{P}$  projects out the two degrees of freedom that lift the degeneracy. The bar indicates that the values are obtained with the assumption that the low level can be calculated for the ground state:

$$\overline{E^{*,\text{ONIOM}}} = E_{\text{model}}^{*,\text{high}} + E_{\text{real}}^{\text{low}} - E_{\text{model}}^{\text{low}} \quad (10)$$

Eq. (9) is used directly for studying the interstitial Ni defect in diamond (Section 5), as excitation is only present in the model region. We have started testing an approximation for systems such as cyclohexadiene + naphthalene (Section 6) in which all ONIOM sub-calculations involve excited states, but where  $\mathbf{P}$  in Eq. (9) and the branching space are evaluated for the model only. However, Section 6 focuses on the energy difference between two different excited-state minima, and no conical intersections are involved, so the issue of approximating  $\mathbf{P}$  does not arise here.

### 3. Pyracylene

Pyracylene is a polycyclic aromatic hydrocarbon known experimentally [103] to be photostable and nonfluorescent. We have recently presented [104] an extensive *ab initio* study of its excitation energies and ground and excited-state potential energy surfaces [103]. CASSCF [105] calculations showed that there is a readily accessible sloped [7]  $S_0/S_1$  CI (conical intersection, Figs. 3 and 4) close to a minimum on  $S_1$ , which leads

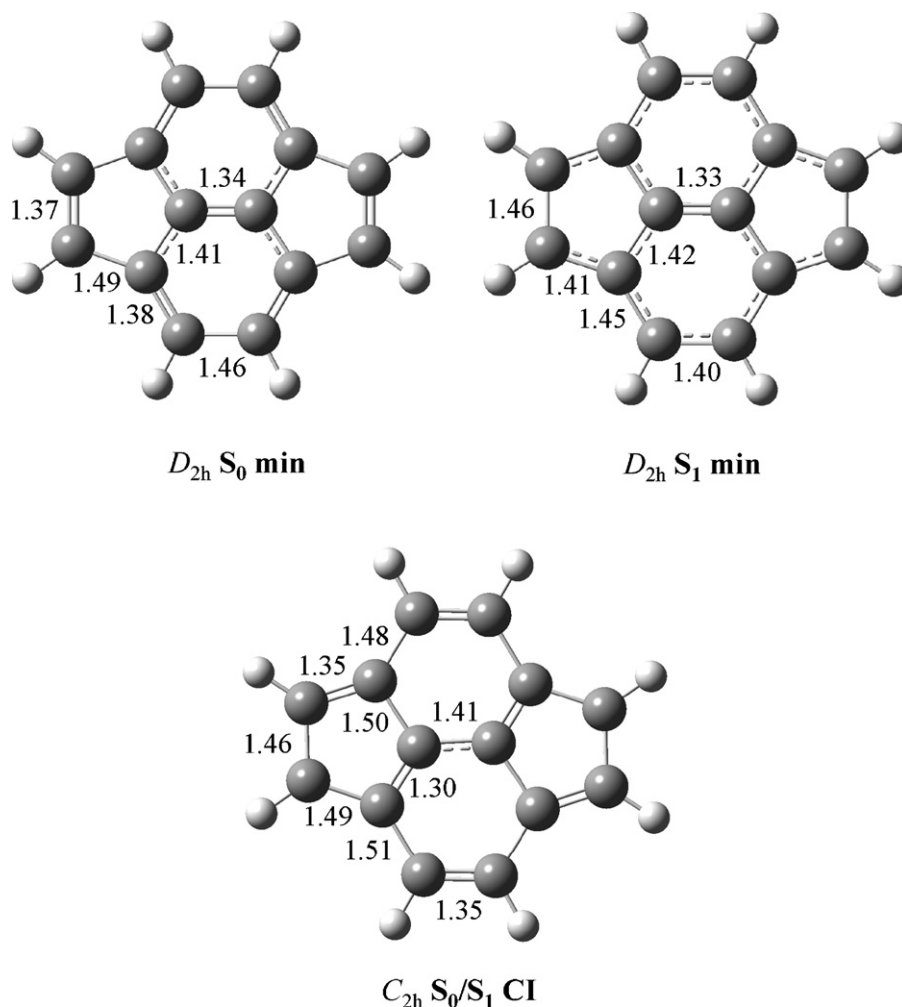


Fig. 3. CASSCF/4-31G optimized structures of pyracylene:  $D_{2h}$   $S_0$  minimum,  $D_{2h}$   $S_1$  minimum, and  $C_{2h}$   $S_0/S_1$  conical intersection (CI). All bond lengths are in angstroms. (RASSCF(14, 4 + 6 + 4) [2,2] bond lengths are not shown separately, as all but one are within 0.01 Å of the CASSCF value: the exception is 1.49 Å with CASSCF for the CI as shown; 1.47 Å with RASSCF).

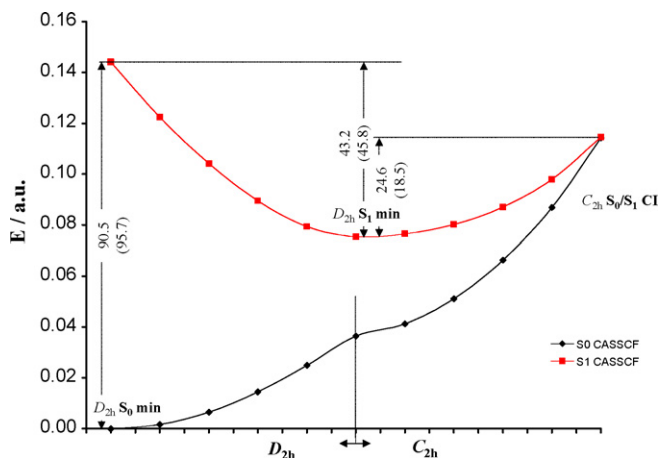


Fig. 4. CASSCF/4-31G potential energy profiles for the  $S_0$  and  $S_1$  states of pyracylene. Critical points are shown in Fig. 3. CASSCF relative energies are in kcal/mol (the corresponding RASSCF energies are shown in parentheses).

to ultrafast internal conversion and explains the observed photostability. This conical intersection has a similar chemical origin to the crossing previously located for  $S_0/S_1$  in azulene [106].

Pyracylene is also a prototype for a series of molecules: several authors have commented that the photophysical properties of larger polycyclic aromatic hydrocarbons containing 5-membered rings are very different to those without, particularly the characteristic lack of fluorescence [107–110] from  $S_1$ , ‘anomalous’ emission [111] from  $S_2$  and reduced resolution [112] of vibrational fine structure. Part of the reason for studying pyracylene is to benchmark approximations to CASSCF that could be used for studying these larger systems.

### 3.1. Methods: CASSCF and RASSCF

For pyracylene (Fig. 3), the CASSCF active space [105] is 14 electrons in 14 $\pi$  orbitals. All orbitals – including those that are not part of the active space – are fully optimized, so the  $\sigma$  orbitals can respond to changes in the  $\pi$  electron distribution.

Using the RASSCF method we can subdivide the active space into three parts, and restrict excitations from the first part and



into the third. The large reduction in electron configurations speeds up the calculation, although this speed-up is partially counter-balanced by the more complex and slower code that needs to be employed. RASSCF calculations [25–27] (using a well-defined subset of around ~1% of the CASSCF electron configurations) were shown to be able to reproduce CASSCF results satisfactorily for perylene, and are potentially useful therefore for larger systems where CASSCF is currently too expensive, when for example there is little or no symmetry to take advantage of.

Hybrid MMVB calculations were carried out for perylene in a previous study as well [21]. The calculations were benchmarked against CASSCF, in order to assess the reliability of this parameterized method for excited states of larger conjugated polycyclic aromatic hydrocarbons. The accessibility of the crossing was illustrated via MMVB dynamics simulations, and its topology (sloped in the classification introduced by Ruedenberg and co-workers [7]) accounts for the high photostability (Fig. 4). However, there were some questions over whether the  $S_1$  state obtained with MMVB was the state initially excited experimentally. (As discussed in Section 2.3.1, MMVB cannot properly describe ionic excited states, and perylene has several.) Nonetheless, MMVB has proved to be a very useful tool for generating initial geometries for CASSCF geometry optimizations [20,63,106,113], and an extensive CASSCF study of perylene only becomes feasible for us when using MMVB geometries as an initial guess. Although we will discuss some of the MMVB work on perylene here, we will focus on the CASSCF and RASSCF calculations.

ONIOM is not particularly suited for calculations on perylene. The entire molecule is effectively the chromophore for perylene, while for ONIOM, one needs to be able to identify a distinct ‘active part’ of the system, and there is no obvious place to truncate the molecule or CASSCF active space without problems.

We computed the ground ( $S_0$ ) and first three excited electronic states ( $S_1$ – $S_3$ ) of perylene with CASSCF by distributing the  $14\pi$  electrons in  $14\pi$  orbitals (14e, 14o), generating around  $10^6$ -electron configurations, as described in more detail in ref. [104]. The basis set used was 4-31G, partly because this was the one used in the original parameterization of MMVB, which we want to compare our results to, but also because it is flexible enough to describe valence states of small (planar) organic molecules.

The RASSCF approach [25–27] reduces the number of electron configurations from  $\sim 10^6$  with CASSCF to  $\sim 10^4$ . We started with the same  $14\pi$  electrons in  $14\pi$  orbitals active space, divided into three parts: orbitals with a limited number of vacancies (RAS1), fully active orbitals (RAS2), and orbitals with a limited number of electrons (RAS3). Both the size of RAS2 and the number of excitations from RAS1 and into RAS3 can be varied. From the outset, we limited excitations from the RAS1 space to singles and doubles only, and allowed only two electrons at most in RAS3. The size of RAS2 was determined by calibrating the RASSCF  $S_0$ – $S_1$  energy gap against the CASSCF value, in the region of the  $S_0/S_1$  conical intersection optimized with MMVB. Using six orbitals in RAS2 brings the

energy gap to 7.8 kcal/mol, just 0.4 kcal/mol away from the target CASSCF value of 8.2 kcal/mol. This level of calculation, denoted RASSCF(14, 4+6+4)[2,2], was then used to optimize the energies and geometries of different electronic states of perylene. Comparison with CASSCF orbital occupancies obtained for each excited state at other critical geometries (minima) was also made to ensure the partition is consistent for all of the geometries and pathways studied. MMVB starting geometries were used, so this study was both a test of the use of RASSCF to make CASSCF calculations more manageable, and a further test of the MMVB method. Using three methods in this way may seem like an unnecessary burden, but our experience is that the ability to study convergence of an energy or geometry in this way can be more valuable than having a single number, as in this way some estimate of errors – independent of the comparison with experiment – can be made.

### 3.2. Results and discussion

The first three singlet excited states ( $1B_{2u}$ ,  $1B_{3g}$ ,  $2A_g$ ) of perylene are very close in energy at the Franck–Condon geometry, all lying around 90 kcal/mol above the ground state. The optically active  $1B_{2u}$  state is marginally the lowest in energy, but relaxation along a  $D_{2h}$  coordinate (and strong vibronic coupling) leads almost immediately to a crossing with the  $2A_g$  excited state. This state is covalent, corresponding mainly to the  $(\text{HOMO})^2 \rightarrow (\text{LUMO})^2$  double excitation. The resulting  $S_1$  minimum energy structure ( $S_1$  min, Fig. 3) is quite different from the ground state minimum ( $S_0$  min, Fig. 3): there is a complete inversion of bond length alternation in the  $12\pi$ -electron periphery.

The adiabatic  $S_0 \rightarrow S_1$  ( $2A_g$ ) transition was measured at 650 nm (44 kcal/mol) [103] in good agreement with our CASSCF result (47.3 kcal/mol). However, while the  $S_1$  state is strongly stabilized at its minimum, the ground state is destabilized, resulting in a large decrease of the  $S_0/S_1$  energy gap to 24.4 kcal/mol. Along a  $C_{2h}$  coordinate, the two states eventually cross, and the resulting  $S_0/S_1$  conical intersection (Figs. 3 and 4) lies about 20 kcal/mol above the  $S_1$  minimum. This energy is well below the available energy in the system if vertically excited and the crossing is therefore energetically accessible. Furthermore, because the crossing is sloped, the energies of the two states are close some way below the fully optimized minimum crossing point.

Comparing RASSCF and CASSCF optimized structures for perylene (Fig. 3) shows that there is excellent agreement between the two methods: the standard error on the bond lengths of the minima does not exceed 0.005 Å with a maximum error below 0.01 Å. The structure of the conical intersection is also well reproduced with a standard error on the bond lengths under 0.01 Å.

However, with RASSCF we also found a slightly distorted  $S_1$  structure with  $C_{2h}$  symmetry lying 1.2 kcal/mol below the  $D_{2h}$  minimum. Such a structure could not be located with CASSCF, and this weak localization of the bonds is probably due to the restricted number of electron configurations used. This is a small

perturbation in the topology of the  $S_1$  surface, but is a note of caution for future studies.

RASSCF and CASSCF energy differences were very similar at optimized structures. All adiabatic excitation energies were well reproduced – in particular the order of the electronic states was preserved – with the largest errors occurring for the vertical excitation energies, as the excited-state potential energy surfaces are steep in this region.

From an energy point of view, MMVB provides a qualitative description of the  $S_0$  and  $S_1$  potential energy surfaces, including the nonadiabatic reaction path leading to the  $S_0/S_1$  conical intersection. The main difference is that the crossing occurs at 24.6 kcal/mol above the  $S_1$  minimum at the CASSCF/4-31G level, instead of 7.9 kcal/mol with MMVB. (The crossing lies 14.3 kcal/mol above the  $S_1$  minimum using RASSCF/6-31G(d).) CASSCF may well overestimate the barrier to the crossing, as for cyclohexadiene/hexatriene photochemical interconversion [114]. CASPT2 or similar methods could improve the energetics, provided geometry changes at this level of theory for a system of this size can be taken into account.

#### 4. Triangulene

Triangulene is a prototype example of a non-Kekule hydrocarbon. With a  $C_3$  symmetry axis, a pair of degenerate nonbonding molecular orbitals and hence a triplet ground state [115] is expected. Clar and Stewart [116] first proposed the existence of triangulene over 50 years ago, but its reactivity meant that synthesis of derivatives was only achieved recently, following interest in high-spin ground states as a source of molecular magnetism.

##### 4.1. Methods: MMVB

With pyracylene, we discussed speeding up slow but feasible CASSCF calculations through RASSCF configuration selection. Because triangulene (Fig. 5) has a total of  $22\pi$  electrons – 8 more than pyracylene – analogous CASSCF calculations are not possible at present, because of the factorial dependence of the number of electronic configurations with the number of active orbitals (for the full active space, there would be  $\sim 10^{10}$  configurations).

Choosing only a subset of active space orbitals for an extended  $\pi$  system such as triangulene with CASSCF may lead to problems with symmetry breaking (*e.g.* indacene [63], *c.f.* [60–62]) and an active space that cannot represent the chosen states in an unbiased way over the required range of geometries. We have not yet explored using RASSCF for triangulene, as we would be doing so without being able to calibrate configuration selection against CASSCF.

On the other hand, MMVB is well suited for calculations on triangulene. There are only  $\sim 10^6$  configurations in the VB part of MMVB, because there are far fewer many-electron basis functions with  $S_z = 0$  where all of the orbitals are singly occupied (VB), compared to those where an orbital can be doubly, singly, or unoccupied (CASSCF). The VB ‘perfect-pairing’ basis functions are therefore a small subset of those that would be used in CASSCF. Furthermore, there are no orbitals to optimize for

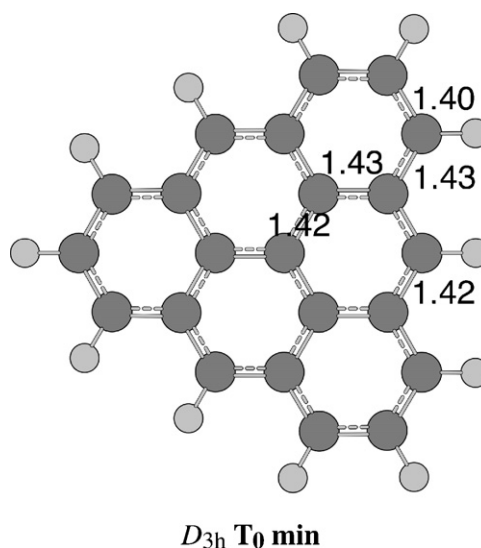


Fig. 5.  $D_{3h}$  triplet minimum structure of triangulene, computed with MMVB. All bond lengths are in angstroms.

MMVB, as the VB Coulomb and exchange parameters have been fitted, and VB parameters for carbon are available in MMVB. Hybrid MMVB calculations can therefore be carried out on molecules for which CASSCF calculations are currently impossible, even with high symmetry, but there is also a lack of optimized excited-state geometries and non-vertical excitation energies to benchmark MMVB against to date.

For triangulene, we can ask several questions that MMVB calculations are sufficient to answer: is the ground state a triplet spin state, as predicted? If so, does the minimum energy structure have maximum  $D_{3h}$  symmetry, or lower symmetry? And how much higher in energy than the triplet is the lowest singlet state? Does this singlet state also have  $D_{3h}$  symmetry? In these calculations, the range of geometry changes is small and in-plane, and the changes are due to reorganization of  $\pi$  electrons, and do not affect the number of  $\sigma$  bonds present in the molecule. While the MMVB results may not be to chemical accuracy, they may provide sufficient semi-quantitative mechanistic information to interpret previous experiments, and perhaps suggest new ones. Our results are summarized in the next subsection, but we mention several of our findings in the context of the technical detail used to obtain them in this section, in order to justify the choice of MMVB for this system.

For triangulene, the energy separation between the  $S_0$  and  $S_1$  states at the  $S_0$  minima/TS is predicted to be only about 12 kcal/mol. For CASSCF, this would almost certainly require state averaging of  $S_0$  and  $S_1$ , while optimizing with the lower  $S_0$  state gradient (leading to additional computational expense solving the CPMCSCF equations for orbital relaxation [43], even if only approximately). For MMVB, with ‘orbitals’ that are only treated implicitly (via interaction parameters), there is no need for state averaging, and this also speeds up conical intersection searches.

MMVB can be considered to be either a molecular mechanics method coupled to a model electronic eigenvalue problem that modifies MM force constants for particular atom types via

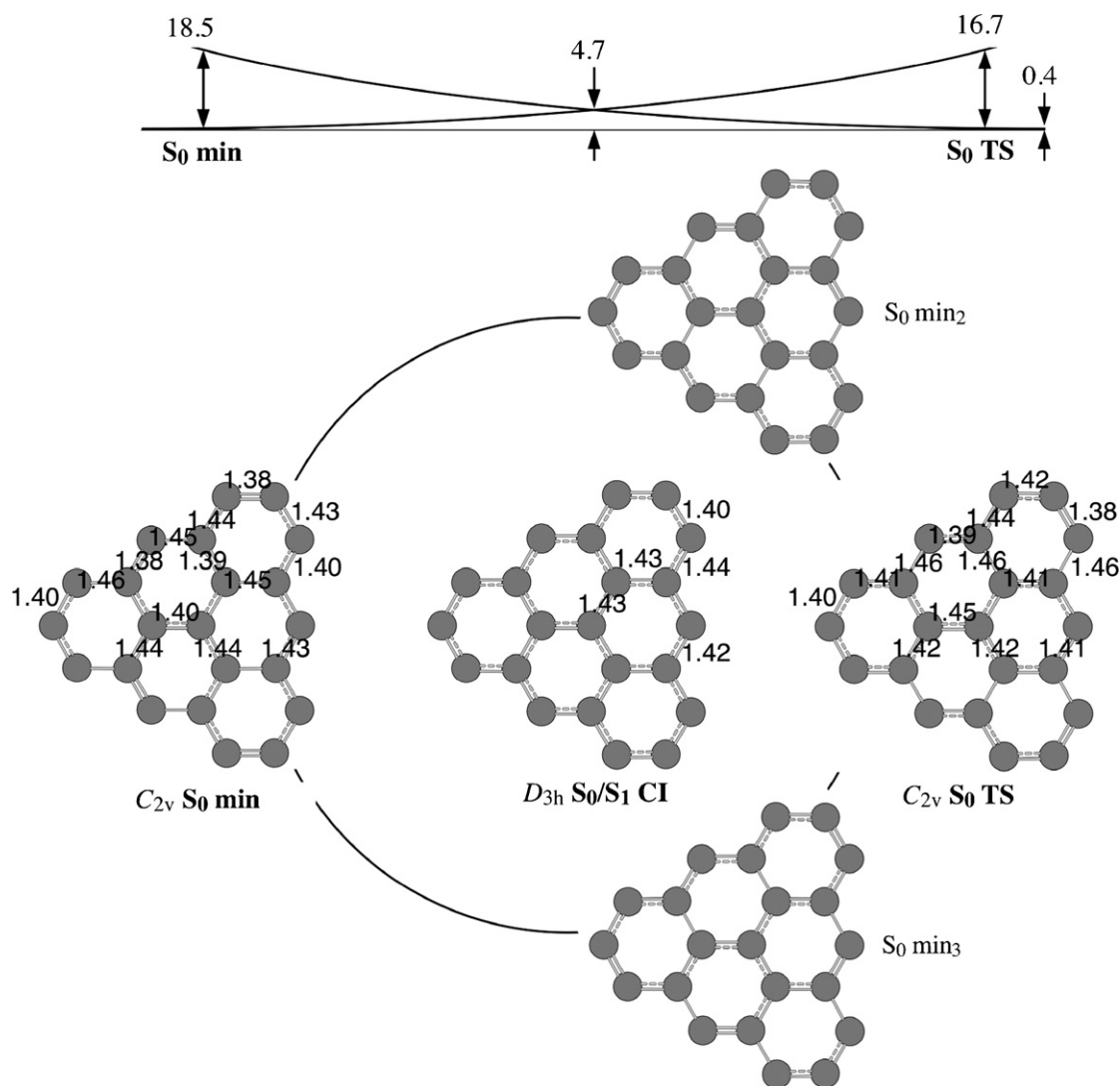


Fig. 6. The  $D_{3h}$   $S_0/S_1$  conical intersection of triangulene, and associated  $C_{2v}$  minima and transition structures (only 1 of 3 shown), together with their relative energies in kcal/mol. All bond lengths are in angstroms.

a computed electron density ( $P_{ij}$ ) [22,44], or as a CASSCF calculation where the eigenvalue problem has been transformed to a much more compact basis, and the integrals either evaluated as parameters of distance/orientation or treated implicitly by mechanics. MMVB can therefore describe reorganization of  $\pi$  electrons and associated geometry in different states, as well as new  $\sigma$  bonds forming (although in the present implementation we describe the first better than the second for technical reasons that are described elsewhere [20,21]). For triangulene, there is a clear partitioning; the connectivity of the  $\sigma$ -bonded framework is fixed, but bond lengths change in response to the different  $\pi$  electron distributions in different electronic states.

The most difficult part of the MMVB calculations on triangulene described here was the search for an  $S_0$  transition structure, although without MMVB, such a search could probably not have been carried out at all. The challenge for any transition structure search is to have a good guess for the starting geometry. Starting with an  $S_0$  minimum structure, we attempted linear interpolation/QST2 searches between two adjacent equivalent

minima (Fig. 6), which were unsuccessful. Instead, we chose to use the electronic structure of  $S_1$  at the  $S_0$  minimum geometry as a guide, which shows quite a different bonding pattern to  $S_0$  itself, and correlates diabatically with a ground state transition structure. Thus, we searched for a point on  $S_0$  having the  $S_1$  electronic structure; effectively following this state down through the crossing, continuing on  $S_0$  in the opposite direction to the pathway from the crossing to the  $S_0$  minimum. (This is along the gradient difference coordinate through the  $S_1/S_0$  conical intersection, which is computed as part of the intersection search. This correlation – and the possibility of following  $S_1$  at the  $S_0$  minimum directly through the crossing along the gradient difference vector between the two states – is characteristic of Jahn–Teller systems, e.g. [117].) MMVB provides a way to do this type of optimization by calculating and ‘fixing’ the electronic structure (via the density  $P_{ij}$  [22]) at a particular geometry, optimizing the geometry while keeping the density fixed, then recomputing the wavefunction/density matrix and repeating until the geometry optimization has converged. The key to

making this approximation work is to recognise when the bonding pattern sought is on  $S_1$ , and when states have swapped and the bonding pattern sought is on  $S_0$ . Also, this procedure appears to work because  $P_{ij}$  changes relatively slowly with geometry for a particular diabatic electronic state.

MMVB also has numerical frequencies (no analytical frequencies yet)—which are accurate enough to support the idea that the  $S_0$  TS found as described above is a real transition structure, by computing the transition vector ( $450i\text{ cm}^{-1}$ ), and following an IRC for 21 steps in the forward and reverse directions (although we have to use caution, as the surface is very flat, and the reaction path curved). A stepsize of  $0.005\text{ \AA}$  was used for the numerical frequencies.

Thus, in MMVB, we have an MM-based method, yet we have accurate numerical frequencies because of the new Gaussian implementation [85]. While the MMVB method is implemented on top of MM2 at the moment, it could in principle be built with any force field, by providing the appropriate new terms in that potential.

The reliability of MMVB for these calculations on triangulene is supported by our recent studies on vinylbiphenyl and 2-vinyl-1,3-terphenyl [45]. Here, CASSCF and MMVB calculations were carried out on a smaller model (vinylbiphenyl) to benchmark MMVB for this type of molecule, with MMVB then used to carry out calculations on the benzene-substituted 2-vinyl-1,3-terphenyl for which full CASSCF calculations are not yet possible. Agreement between CASSCF and MMVB calculations was not perfect, but sufficient to interpret the experiments.

The ONIOM method is not particularly suitable for calculations on triangulene, for the same reasons in the previous section on pyracylene: the whole molecule is effectively a chromophore.

#### 4.2. Results and discussion

Four critical points were optimized for triangulene with MMVB: a  $D_{3h}$  minimum on  $T_0$  (Fig. 5), a  $D_{3h}$   $S_0/S_1$  crossing, a  $C_{2v}$   $S_0$  minimum and a  $C_{2v}$   $S_0$  transition structure (TS) (Fig. 6). Two of these points were previously located [22]: the  $D_{3h}$   $T_0$  minimum and  $C_{2v}$   $S_0$  minimum.

The lowest energy critical point found for triangulene is the  $D_{3h}$  minimum on the triplet  $T_0$  potential energy surface (Fig. 5), *i.e.* triangulene has a triplet ground state, which is 23 kcal/mol below the  $S_0$  minimum. The triplet ground state was previously characterized experimentally [118,119], and also subsequently confirmed [120–123] and shown to have  $D_{3h}$  symmetry. We calculate the triplet state  $T_0$  to be the lowest in energy at all of the critical points optimized on the singlet  $S_0$  and  $S_1$  potential energy surfaces. (We use a convention here that the lowest singlet state is described as  $S_0$ , although we calculate this to be an excited state, higher in energy with respect to the triplet ground state  $T_0$ .)

On the  $S_0$  surface, the  $C_{2v}$  minimum previously located ( $S_0$  min, Fig. 6, left) has a diradical structure. This study developed from the search for a minimum on  $S_1$ , and a question over why the  $S_0$  minimum did not have  $D_{3h}$  framework symmetry. Both turn out to be inextricably linked: there is in fact no minimum where the gradient on  $S_1$  goes to zero. Instead,

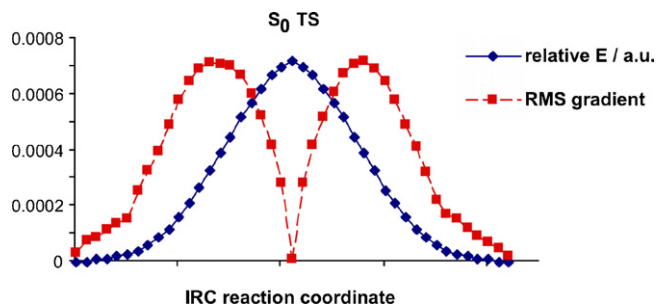


Fig. 7. IRC from a triangulene  $S_0$  TS, leading to equivalent minima as shown in Fig. 6. Energy above the minima in Hartrees and gradient are shown.

we find that a peaked  $S_0/S_1$  conical intersection with a  $D_{3h}$  geometry (Fig. 6, centre) is the lowest-energy point on  $S_1$ . In other words, the first singlet state of triangulene calculated with MMVB appears to show the Jahn–Teller effect, leading to lower-symmetry structures away from the degeneracy on the ground state. This explains the existence of a lower ( $C_{2v}$ ) symmetry minimum, but also suggests there should be an associated transition structure in the ‘moat’ around the symmetric structure, for interconverting equivalent minima.

Such an  $S_0$  transition structure was eventually located ( $S_0$  TS, Fig. 6, right), and an IRC calculation (Figs. 7 and 8) shows that it links two energetically equivalent but geometrically distinct minima ( $S_0$  min<sub>2</sub> and  $S_0$  min<sub>3</sub>) in forwards and reverse directions, as predicted. This supports the suggestion that the  $S_0/S_1$  crossing in triangulene is responsible for a Jahn–Teller stabilization. The transition structure proved difficult to locate as it has quite a different geometry to the minimum itself, but is  $<1$  kcal/mol higher in energy.

As we discussed in more detail in a recent review [38], when looking at a photochemical reactivity problem, it often proves

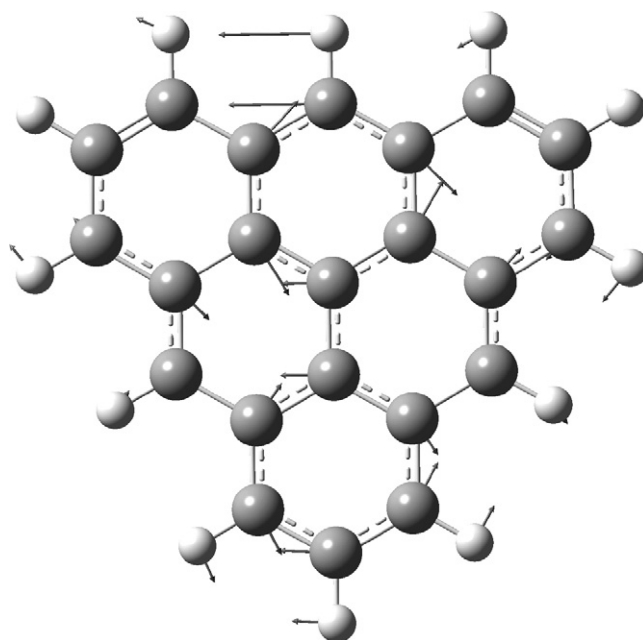


Fig. 8. The transition vector computed at a triangulene  $S_0$  transition structure with MMVB, via a numerical frequency calculation.

useful to be able to recognise both specific diabatic states – particular bonding patterns, which may be on excited or ground states depending on the molecular geometry – and adiabatic state labels, e.g.  $S_0$  and  $S_1$ . For MMVB, it is straightforward to do both [22]. For triangulene:  $S_0$  at  $S_0$  min correlates with  $S_1$  at the  $S_0$  TS structure and vice versa; the  $S_0$ – $S_1$  energy gap is around 17 kcal/mol in each case, and both states intersect at the crossing, which is around 5 kcal/mol above the  $C_{2v}$  minimum. The choice of horizontal axis in Fig. 6 is deliberate: from the crossing  $S_0/S_1$  CI, the gradient difference vector (computed as part of the crossing search, and which lifts the degeneracy to first order) leads in one direction towards  $S_0$  min, and in the opposite direction towards  $S_0$  TS. Thus, the minima and transition structures in triangulene are linked in two directions: curved around the crossing (via the transition vector) and straight through the crossing (via the gradient difference vector). This is a characteristic of Jahn–Teller systems.

Because the transition structure located is <1 kcal/mol above the  $S_0$  minimum, we predict that the singlet state of triangulene – if observable – would have a time-averaged  $D_{3h}$  structure, and be a case of the dynamic Jahn–Teller effect (c.f. [124]). The Jahn–Teller stabilisation – the difference between the minima/transition structures and the degeneracy – for triangulene (4.7 kcal/mol) is a little less than that determined for the cyclopentadienyl radical with CASSCF (6 kcal/mol) and MMVB (7.2 kcal/mol) [20,117].

## 5. Ni defect in diamond

Transition metals catalyze the high-pressure, high-temperature synthesis of diamond. Nickel has been shown to form point defects in natural and synthetic diamonds [125–130]. Mason et al. [131] have characterized an interstitial Ni defect in diamond, with transitions between ground and excited levels measured at  $\sim 1.4$  eV. Here we investigate computationally the geometries and energetics of the ground and excited states of an interstitial nickel defect in diamond. We find a Jahn–Teller active  $T_d$  structure, and two slightly distorted  $D_{2h}$  structures <0.5 kcal/mol lower in energy. This study is a starting point, to illustrate the methods chosen and the reasons for choosing them.

### 5.1. Methods: ONIOM

We study the electronic structure using a cluster of 78 carbon centers around a single nickel atom. Although the excitations are located primarily at the Ni site, we feel that a large cluster is required to mimic the solid-state Ni defect in diamond, in particular for the geometric constraints and relaxation, which will indirectly affect the excitation energies. Previous calculations have been reported [129], but not all [132,133] involved explicit geometry optimization [134]. The excitations can be modeled with a very small active space in CASSCF. Despite this small active space, the remainder of the cluster is so large that it renders full CASSCF calculations intractable. In fact, even ground state calculations using DFT or HF become expensive for such large clusters. If the number of electron configurations could be

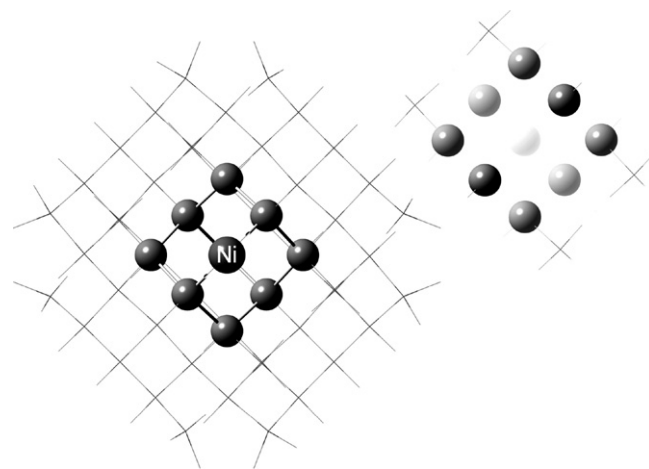


Fig. 9. The ONIOM partitioning chosen here for modeling an interstitial Ni defect in diamond. Left/Main: The model region is drawn as ball-and-stick; the remaining real system is drawn in wireframe. Not all atoms are shown: those in front of the Ni and others to an equivalent depth have been removed ('Z-clipped') to show the model system in context. Right/Inset: A close-up of the left-hand view using 'fog' to exaggerate depth, with the Ni atom hidden; emphasizing that this is an interstitial defect, not a substitutional one. This inset view is equivalent to Fig. 10 (right) rotated  $45^\circ$ , and with Ni and C7 hidden.

reduced using RASSCF, it would not speed up the calculation, since the CASSCF bottleneck is formed by the number of inactive orbitals. Also MMVB cannot be applied because at present it is not parameterized for metals.

ONIOM, on the other hand, is ideally suited for the study of such a process. The active part of the system can be clearly identified. In our ONIOM calculations, we assume that the effect of most of the cluster on the excitation is purely steric, which allows us to use MM methods in the low level.

The ONIOM *model* we used is shown in Figs. 9 and 10: an interstitial (not substitutional)  $Ni^+$  atom +  $C_{10}$  adamantane cage (the 16 hydrogen link-atoms are not shown). Interaction between the metal d orbitals (Fig. 11) and the adamantane cage splits the metal energy levels [135]. Here we assume that the amount of splitting only depends directly on the distance to the nearest 10 carbon atoms, which are therefore included in the QM region and model system. The full *real* system is necessary to describe the relaxation (and hence indirectly the excitation energies) of the model system realistically: without it, the adamantane cage will expand too much, affecting the calculated T–E state splitting.

As high-level (QM) method we used CASSCF with state averaging and the lanl2mb basis set. ROHF appears to give very good starting orbitals for the CAS. Our model is that we have  $Ni^+$  ( $3d^9$ ), which is found to be the most stable configuration (see for example [136]), so this is a calculation on states arising from splitting of a  $^2D$  term into triply degenerate T and doubly degenerate E levels in a tetrahedral (or approximately tetrahedral) environment. For the low level we used the UFF force field [137]. There are no specific UFF parameters for  $Ni^+$  ( $3d^9$ ), but the Ni atom appears in both the model and real (whole system) parts of the ONIOM equation, and any undefined terms therefore cancel. The Van der Waals interaction at the MM level is zeroed for nickel. Without zeroing, we observed unphysical symmetry

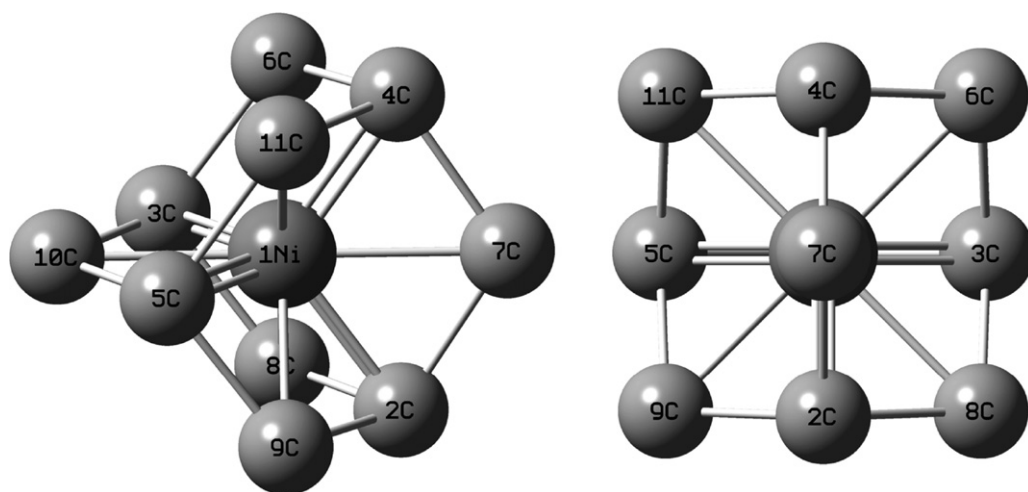


Fig. 10. Labeling for the C–C and C–Ni bondlengths for Ni in diamond, shown in Table 1.

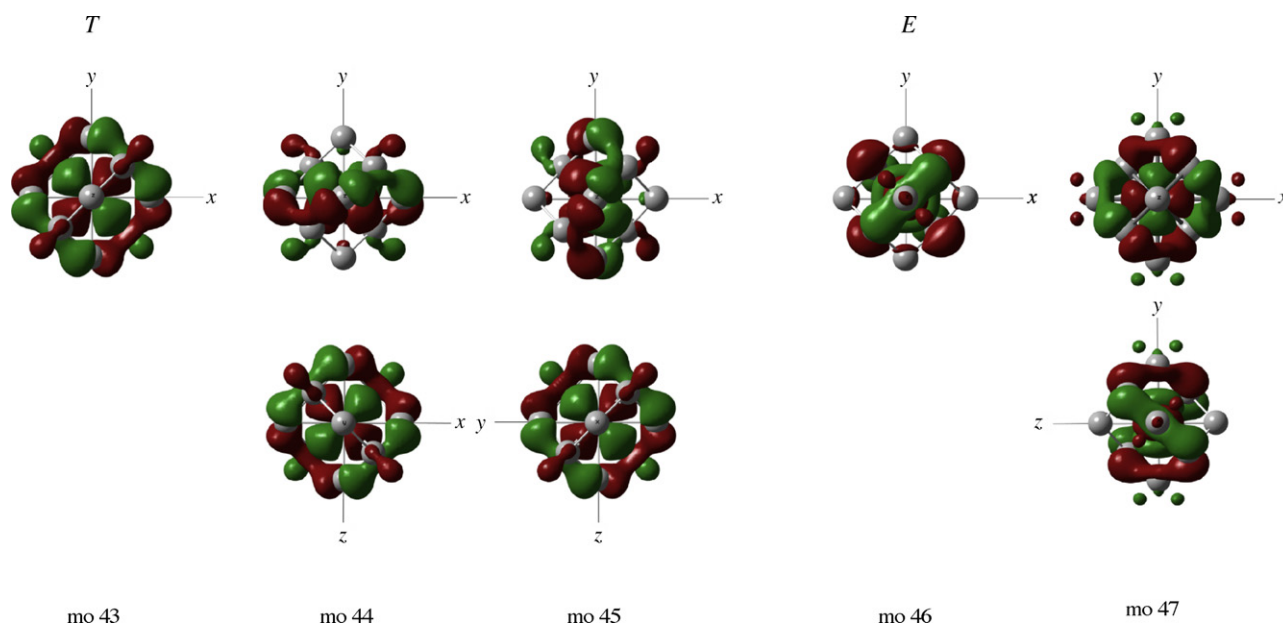


Fig. 11. The CAS(9, 5) active space for the model part of the ONIOM Ni in diamond calculations.

breaking effects. We did not include Coulomb interactions in the UFF contributions to ONIOM.

The critical points (conical intersection and associated minima) were optimized with state averaging over the first two states (equal weights) with an active space of 9 electrons in 5 orbitals.<sup>1</sup> The optimizations of the minima were done on the ground state. The CPMSCF contribution [43] was included to ensure that we have a reliable description of the Jahn–Teller distortion. The upper three states do not participate and do not need to

<sup>1</sup> We started the optimizations without symmetry, and had geometry optimization problems when using the (9, 5) active space, state-averaging over the first two states. This is most likely due to the doubly occupied orbitals in the active space, and the way the CASSCF code neglects the corresponding orbital rotations. Therefore we switched to the (3, 2) active space, and after geometry optimization with this active space, we recalculated energies and checked that forces were converged with the (9, 5) active space.

be included in the state averaging in the optimizations. The final energetics, reported in the tables, are obtained with state averaging over the first five states using equal (0.2) weights, but using geometries obtained with state averaging over the first two states.

## 5.2. Results and discussion

We show the bond lengths of the optimized structures in Table 1, with geometries labeled as in Fig. 10. Relative energies are given in Table 2.

The CASSCF active space used is shown in Fig. 11. Orbital energies are not defined for CASSCF; instead we rely on the state energies, and many-electron expansion coefficients to characterize the electronic states. This potentially leads to some confusion when comparing to the literature, much of which emphasizes the importance of the orbital energy levels. Here, the ground states have electronic configuration ( $t^6e^3$ ) which gives rise to a doubly

Table 1

C–C and Ni–C bond lengths for the three critical points located on the  $D_0$  and  $D_1$  potential energy surfaces of the Ni in diamond system, calculated with CASSCF/lanl2mb as part of an ONIOM calculation

R	MIN1 (Å)	MIN2 (Å)	CI (Å)
10.5	1.706	1.713	1.710
10.3	1.706	1.713	1.710
7.4	1.706	1.713	1.710
7.2	1.706	1.713	1.710
5.9	1.712	1.708	1.710
9.2	1.712	1.708	1.710
3.8	1.712	1.708	1.710
8.2	1.712	1.708	1.710
5.11	1.712	1.708	1.710
11.4	1.712	1.708	1.710
3.6	1.712	1.708	1.710
6.4	1.712	1.708	1.710
Ni.2	1.751	1.75	1.755
Ni.4	1.751	1.75	1.755
Ni.3	1.751	1.75	1.755
Ni.5	1.751	1.75	1.755
Ni.7	1.962	1.938	1.950
Ni.10	1.962	1.938	1.950
Ni.11	1.943	1.955	1.948
Ni.6	1.943	1.955	1.948
Ni.9	1.943	1.955	1.948
Ni.8	1.943	1.955	1.948

Atom numbering is shown in Fig. 10.

degenerate E ground state at  $T_d$  geometries, whereas the excited states have a ( $t^5e^4$ ) electronic configuration, giving rise to a triply degenerate T excited state.

‘CI’ is the  $T_d$  symmetric system. The first two states ( $D_0$  and  $D_1$ ) are degenerate, and the upper three states are degenerate ( $D_2/D_3/D_4$ ) as predicted. The splitting between these two groups of states (Table 3) is 27.7 kcal/mol or about 1.2 eV, which is consistent with the experimentally observed splitting of 1.4 eV. However, this geometry is not a true minimum in the full space, and was optimized with CASSCF and Opt=Conical. For the ground E states: the XY, XZ, and YZ d orbitals are always completely filled, and the ( $X_2-Y_2$ ) and  $Z_2$  d orbitals are partially filled. With partially filled degenerate orbitals and high symmetry, we expect this to be a Jahn–Teller system.

Using the CASSCF code, we explicitly calculate the directions that lift the 2-fold degeneracy of the lowest states at the unstable  $T_d$  geometry: at this point, the system can relax, lowering the energy. Two lower-energy structures were found, both having  $D_{2d}$  symmetry: ‘MIN1’ and ‘MIN2’. The XY, XZ, and

YZ orbitals are again completely filled, but the ground state of MIN2 has two electrons in ( $X_2-Y_2$ ) and one in  $Z_2$ , while in MIN1 there is one-electron in ( $X_2-Y_2$ ) and two electrons in  $Z_2$ .

The energy gain from the Jahn–Teller distortion is very little: 0.29 kcal/mol for MIN1, and 0.23 kcal/mol for MIN2, and the  $D_0$  and  $D_1$  states have split apart by just over 1 kcal/mol in each case (Table 2). Both changes are an order of magnitude smaller than the splitting of levels induced by the tetrahedral environment: hence the Jahn–Teller distortion represents a small perturbation on this. The small energy change on  $T_d \rightarrow D_{2d}$  relaxation is consistent with the small geometry changes shown in Table 1: only one Ni–C bond length changes by more than 1%. The splitting between  $D_1$  and  $D_2$  – which we correlate with the experimental energy splitting – is consequently almost unaffected, whichever geometry (Table 1) we consider (Table 3).

The  $D_{2d}$  group has an E representation (as well as  $A_1$ ,  $A_2$ ,  $B_1$  and  $B_2$ ), which allows it to have doubly degenerate states. However, the double degeneracy present in  $T_d$  symmetry splits to give  $A_1 + B_1$  when the symmetry is lowered to  $D_{2d}$ . Therefore, states that are doubly degenerate in  $T_d$  would be expected not to be degenerate in  $D_{2d}$ , and cannot give the E representation that the  $D_{2d}$  group has. This is consistent with the splitting of the  $D_0$  and  $D_1$  energy levels shown in Table 3.

In the literature there does not appear to be consensus on the symmetry and Jahn–Teller effect in this system, or even that this interstitial  $Ni^+$  is the defect [132,134] absorbing at 1.4 eV (32 kcal/mol). We are concerned that from experiment the structure is determined [138] to be  $C_{3v}$ , while we obtain  $D_{2d}$  (which is nevertheless consistent with Orgel’s predictions for a  $d^9$  tetrahedral system in ‘Introduction to Transition Metal Chemistry’ [139]). With improved level of theory, our results may change though, which would show how important careful consideration of the type of calculation and choice of model is. The Jahn–Teller stabilization we calculate is very small in this case, about a quarter of a kcal/mol, and we cannot predict if changing the level of theory would make this stabilization larger, smaller, or even disappear. Examples of improvements to the level of theory are: (1) enlarging the QM region in ONIOM; (2) increase the size of the basis set in CASSCF (split valence/polarization) or level of theory (CASPT2); (3) higher level of theory for the low level region (HF); (4) allow the excitation to extend into the low level region (e.g. CIS for low level). However, even with this minimal representation, our calculated splitting is closer to the experimental

Table 2

$D_0$  and  $D_1$  energies at the three critical points located on the  $D_0$  and  $D_1$  potential energy surfaces of the Ni in diamond system, calculated with CASSCF/lanl2mb as part of an ONIOM calculation

Geometry	Symmetry	ONIOM <sup>a</sup> (Hartrees)	$D_0$ (0.5) <sup>b</sup> (Hartrees)	$D_1$ (0.5) <sup>b</sup> (Hartrees)	$\Delta E$ on $D_0$ (kcal/mol)	$\Delta E$ ( $D_0-D_1$ ) (kcal/mol)
MIN1	$D_{2d}$	–550.165479	–550.879449	–550.877542	0.0	1.20
MIN2	$D_{2d}$	–550.165390	–550.879266	–550.877584	0.06	1.06
CI	$T_d$	–550.165017	–550.878694	–550.878694	0.29	0.00

Geometries are shown in Fig. 10/Table 1.

<sup>a</sup> Extrapolated ONIOM  $D_0$  energy.

<sup>b</sup> Model calculation only. Numbers in parentheses indicate CASSCF state averaging coefficients.

Table 3  
Excitation energies for the Ni in diamond system obtained by state averaging five states, at the geometries optimized by state averaging two states shown in Fig. 10/Table 1

Geometry	Symmetry	$D_0$ (0.2) <sup>a</sup> (Hartrees)	$D_1$ (0.2) (Hartrees)	$D_2$ (0.2) (Hartrees)	$D_3$ (0.2) (Hartrees)	$D_4$ (0.2) (Hartrees)	$\Delta E$ ( $D_1$ – $D_2$ ) (eV)
MIN1	$D_{2d}$	–550.879307	–550.877428	–550.834802	–550.834102	–550.834102	1.16
MIN2	$D_{2d}$	–550.879123	–550.877469	–550.834492	–550.834492	–550.833796	1.17
CI	$T_d$	–550.878566	–550.878566	–550.834489	–550.834489	–550.834489	1.20

<sup>a</sup> Numbers in parentheses indicate CASSCF state averaging coefficients.

value than a value of 0.8 eV for Ni<sup>+</sup> interstitial in adamantane without a fully optimized geometry [129].

## 6. [4 + 4] Photochemical cycloaddition of cyclohexadiene and naphthalene

As an example of the study of excited-state reactivity for extended conjugated systems with hybrid methods, we look at the [4 + 4] photochemical cycloaddition of cyclohexadiene to naphthalene, for which we calculate the relative stabilities of the singly bonded and doubly bonded minima in the  $S_1$  electronic state (Figs. 1 and 12). We have previously studied the photochemical cycloaddition of butadiene + butadiene using CASSCF [140] and MMVB dynamics [141]. Here we combine the two methods using ONIOM. This study is also a starting point (*c.f.* Section 5), to illustrate the combination of methods chosen and the reasons for choosing them. [4 + 4] Photocycloadditions to larger aromatic molecules have been studied experimentally: see for example [142–144]. Part of this is in connection with photochromism, such as that exhibited by the dimerization of anthracene, which cannot be studied with CASSCF at present (28 active electrons).

### 6.1. Methods: CASSCF, MMVB, and ONIOM(CASSCF:MMVB)

In the full CASSCF calculations on this system (which we carried out for calibration) the active space consists initially of 14 electrons distributed in 14 $\pi$  orbitals, which becomes 12 $\pi$  orbitals plus 2 $\sigma$  orbitals (one bonding and one anti-bonding) for the singly bonded structure, and 10 $\pi$  orbitals plus 4 $\sigma$  orbitals (two bonding and two anti-bonding) for the doubly bonded structure. This active space results in about 6 million configurations, and despite current programs being able to handle this (the same sized active space was used for pyracylene, Fig. 3), it is by no means a routine calculation.

This is a situation where the use of RASSCF to reduce the size of a full CAS(14, 14) calculation could be problematic, as there are different numbers of  $\pi$  and  $\sigma$  orbitals at the two geometries studied. This would mean a different selection of orbitals making up the RAS subspaces for each isomer, and while this would probably not affect the calculated relative energies too much, it means that the  $S_1$  potential energy surface linking the two structures computed with RASSCF would be discontinuous.

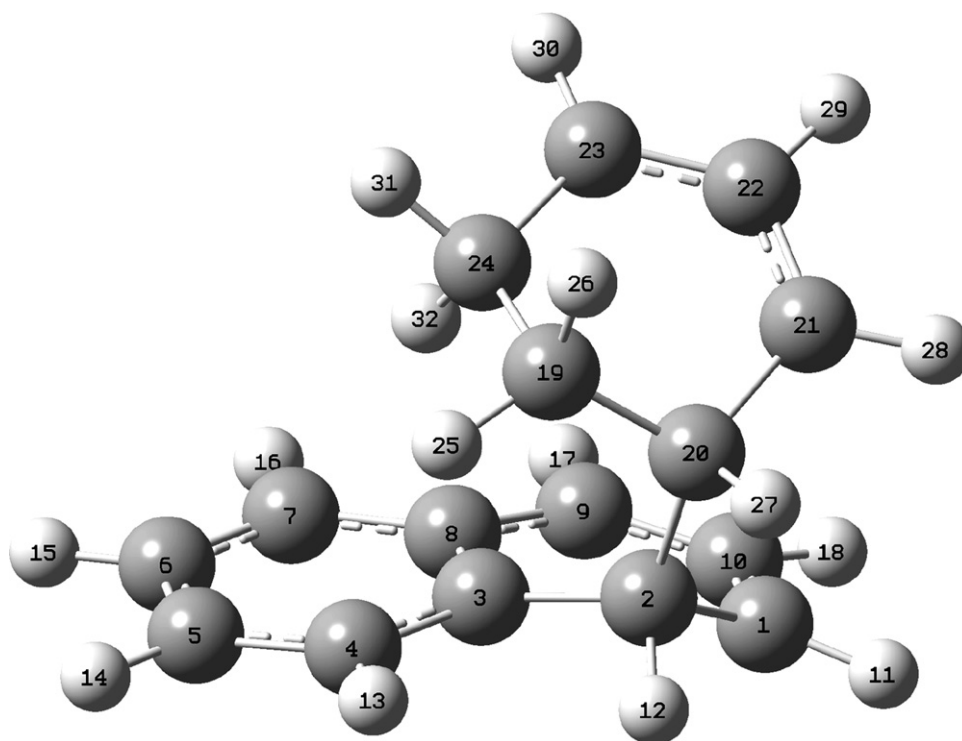


Fig. 12. Labeling for the C–C bondlengths for CHD + naphthalene, shown in Table 5. The singly bonded  $S_1$  structure is shown.



MMVB can be applied to this system, as VB parameters for carbon are available. The changes are primarily due to reorganization of  $\pi$  electrons, but as Fig. 2 shows when compared with Figs. 1 and 12, we are also using  $\pi$  orbitals to form one or two new  $\sigma$  bonds, which we can describe with MMVB using the VB parameters developed for model systems. This emphasises that we can use MMVB to describe reorganization of a  $\pi$  system, as well as new  $\sigma$  bond formation: we are not restricted to a fixed connectivity, as for conventional MM2. However, as we will show later, MMVB does not reproduce the CASSCF data accurately.

At first sight, the naphthalene + cyclohexadiene addition also appears to be suitable for study using ONIOM, since the bond breaking and formation takes place in a localized part of the system. However, the conjugated system, and therefore possibly the excitation, extends over the entire system. In the ONIOM(CAS:MM) calculations on Ni-defects in diamond, we assumed that the excitation is fully localized: the change in wave function or density upon excitation is entirely located in the active region, and we did not assign a charge distribution to the spectator (inactive) region. The latter implies that changes in electrostatic interaction between the active and inactive regions are neglected. These approximations do not always hold. For example, the charge density often changes dramatically upon excitation. In a polar environment such as water, the electrostatic interaction between the layers then needs to be taken into account. Another example where the assumption above may fail is the excitation in conjugated systems. Although the excitation may still occur primarily in a small part of the conjugated system, the ‘tail’ of the excitation may extend well into the remainder. In previous work, we investigated the effect of the size of the conjugated system for a series of retinal protonated Schiff base (PSB) models. Although the excitation is localized in the vicinity of the PSB, the topology of the excited-state changes significantly with the size of the model [57–59]. Also, the behavior of ONIOM was investigated for this system [14].

Here we report ONIOM calculations on the naphthalene and cyclohexadiene addition that do not force the excitation to be localized in the model system. To achieve this, we use MMVB as the low level in ONIOM, with CASSCF being the high-level method. In other words, we integrate the *hybrid-atom* approach with the *hybrid-molecule* approach in a single calculation. Each of the three ONIOM sub-calculations is an excited-state calculation here, giving an integrated ONIOM  $S_1$  energy.

In Fig. 1, we show the ONIOM partitioning we used. We included only what is initially the  $\pi$ -space of the part of the sys-

tem where the reaction takes place in the high-level region, which will describe the formation of the new  $\sigma$  bond(s). The models for cyclohexadiene and naphthalene are then butadiene and benzene, respectively. We assume that the excitation takes place in this part of the system. However, it is clear that the extended conjugated system, including the remainder of the naphthalene, will stabilize the excited state, for which we use a MMVB description. Effectively, this calculation involves three levels of theory: CASSCF, VB, and MM, by combining the hybrid-molecule and hybrid-atom methods. The MMVB approach combines a VB description of the  $\pi$ -space and resulting new  $\sigma$  bonds with an MM description of the remaining  $\sigma$ -space. On top of that, ONIOM(CAS:MMVB) is used to extrapolate to the CASSCF level of theory for a small region of the system. Although the CASSCF contribution in ONIOM still involves an active space of 10 electrons in 10 orbitals, the number of configurations is reduced dramatically compared to the target.

The ONIOM(CAS:MMVB) geometry optimizations were carried out without using symmetry constraints for the singly bonded minimum, and with  $C_s$  symmetry for the doubly bonded minimum. The CASSCF excited-state stand-alone calculations and those in ONIOM were performed with the orbitals fully optimized on the  $S_1$  state (no state averaging). The 4-31G basis set was used throughout, as this is a first test of combining CASSCF and MMVB with ONIOM, and 4-31G was the basis used for the original MMVB parameterisation on model systems [23]. We first optimized the geometries at the MMVB level of theory. The optimized MMVB geometries were then used to start the ONIOM geometry optimization, from which the optimized geometries were in turn used to start the full CASSCF geometry optimizations. This step-wise execution of geometry optimization proved to be highly efficient, as the initial explorations of the  $S_1$  potential energy surface could be carried out very rapidly.

For these MMVB calculations, the delocalization/bridging correction [24] was switched off, as with this correction enabled, the structures studied here (Figs. 1 and 12) – with one or two new  $\sigma$  bonds between cyclohexadiene and naphthalene – could not be located using MMVB on its own. Instead, attempts to optimize both structures led to a cage compound with *four* new  $\sigma$  bonds. Similar problems have been observed before when studying the prototype [4 + 4] photocycloaddition of butadiene + butadiene with MMVB [141], and also reflect problems with parameterizing the part of the VB potential needed to stabilize new  $\sigma$ -bonds [20]. Parameterised hybrid methods such as MMVB can be orders of magnitude faster than conventional methods, but

Table 4

Energies of and energy difference (kcal/mol) between the singly bonded (Fig. 12) and doubly bonded (Fig. 1) cyclohexadiene + naphthalene structures in the  $S_1$  state at various levels of theory

Method	Singly bonded (Hartrees)	Doubly bonded (Hartrees)	Difference (kcal/mol)
Full CASSCF (target)	–614.3180	–614.2649	33.3
ONIOM	–385.5618	–385.4943	42.4
MMVB	–3.3134	–3.2024	69.6
CASSCF model-only	–385.0316	–384.9188	70.8

The singly bonded structure is consistently the more stable one.

Table 5  
Bond lengths (Å) for the singly and doubly bonded naphthalene + CHD adducts

	Singly bonded				Doubly bonded			
	Model	MMVB	CAS	ONIOM	Model	MMVB	CAS	ONIOM
C2–C20 (C9–C23)	1.58	1.67	1.62	1.63	1.68	1.68	1.65	1.65
C1–C2	1.51	1.56	1.51	1.52	1.51	1.58	1.51	1.52
C1–C10	1.44	1.44	1.42	1.40	1.34	1.33	1.34	1.31
C2–C3	1.51	1.55	1.52	1.52	1.49	1.56	1.51	1.52
<i>C3–C4</i>	–	1.40	1.38	1.38	–	1.44	1.44	1.43
C3–C8	1.44	1.48	1.47	1.45	1.54	1.43	1.44	1.42
<i>C4–C5</i>	–	1.43	1.42	1.44	–	1.45	1.43	1.45
<i>C5–C6</i>	–	1.45	1.43	1.45	–	1.45	1.43	1.45
<i>C6–C7 (C4–C5)</i>	–	1.38	1.38	1.38	–	–	–	–
<i>C7–C8 (C3–C4)</i>	–	1.47	1.45	1.45	–	–	–	–
C8–C9 (C2–C3)	1.40	1.43	1.40	1.44	–	–	–	–
C9–C10 (C1–C2)	1.40	1.40	1.40	1.39	–	–	–	–
C1–C2	1.51	1.56	1.51	1.52	1.51	1.58	1.51	1.52
<i>C19–C20</i>	–	1.56	1.55	1.52	–	1.56	1.55	1.53
<i>C19–C24</i>	–	1.55	1.54	1.57	–	1.64	1.55	1.65
C20–C21	1.50	1.56	1.50	1.50	1.48	1.55	1.51	1.50
C21–C22	1.40	1.41	1.39	1.39	1.54	1.34	1.34	1.33
C22–C23 (C20–C21)	1.38	1.39	1.39	1.39	–	–	–	–
<i>C23–C24 (C19–C20)</i>	–	1.54	1.50	1.51	–	–	–	–

See Fig. 12 for labeling (which shows the singly bonded structure) and Fig. 1 for the ONIOM partitioning used (which shows the doubly bonded structure). The three (horizontal) segments of the table correspond to the geometrical parameters between the fragments, within the naphthalene fragment, and within the CHD fragment, respectively. Only the unique bond lengths are given for the symmetric doubly bonded structure, with the equivalent bond indicated in parenthesis in the first column. The entries in italic in the first column indicate those bonds that are described by MMVB in the ONIOM calculations.

depend upon reliable parameters, which are being reinvestigated at the time of writing.

## 6.2. Results and discussion

We show the results of the ONIOM and conventional calculations in Table 4. We first consider the traditional approaches. When we reduce the size of the system (model-only), but retain the accuracy (CASSCF) of the calculation, we find an error of about 37 kcal/mol compared to the target. When we lower the level of theory (MMVB), but consider the entire system, we find an error of about 36 kcal/mol. It is clear that neither approach is satisfactory in this case. However, when we combine both approaches in the ONIOM hybrid method, we reduce the error to 9 kcal/mol. Although this is still somewhat larger than we typically observe in ONIOM studies, it is much smaller than in the traditional approaches, and might be reduced further by considering different partitions. Further progress in understanding this difference may come from analysing how much of the  $S_1$  excitation for the full cyclohexadiene + naphthalene system is in the naphthalene itself [22], as this may differ for the two structures we have considered. As can be seen from Table 4, the singly bonded structure is more stable at each of the levels of theory, indicating that the CHD moiety is involved in the excitation. However, it is not possible to identify – from the energies alone – which of the two structures causes the error in the MMVB and model-only calculations, since the absolute energies for the different levels of theory cannot be compared directly.

In Table 5 we show the bond lengths of the naphthalene + CHD system optimized at the various levels of theory.

We see that the bond between the two fragments is consistently long for all the methods. This is an indication that the excitation is indeed localized in this part of the system, and is treated correctly in all cases. Generally, the ONIOM values for bonds that are in the MMVB layer resemble the values of the conventional MMVB calculation, and the ONIOM values for bonds that are in the CASSCF layer resemble those of the conventional (target) CASSCF calculation on the full system. This is again an indication of correct behaviour in the ONIOM calculation. Further, we see that most of the bonds do not differ too much from the target (full CASSCF) values, with a few exceptions that are mostly in the model-only and MMVB calculations. This is expected, since these are the most approximate levels of theory, and also had the largest error in the prediction of the relative energies discussed above. Also the C19–C24 bond in the doubly bonded ONIOM calculation is too large, but this bond is in the MMVB layer and therefore ‘inherits’ the error from the MMVB level of theory. It is not clear why MMVB and ONIOM describe this bond so much better for the singly bonded system.

## 7. Conclusion

In studying a photochemical reaction or multi-state process, we would like to be able to apply the CASSCF level of theory (Section 2.1), including all electrons/orbitals from the reacting system that undergo any significant reorganization in the active space. However, these calculations become prohibitively expensive for large molecules and large numbers of active orbitals. In this paper, we considered three alternative strategies

for making these large calculations practical, and determining key observables such as relative energies and geometries:

- Use RASSCF (Section 2.2; truncation) to identify a central part of the CASSCF active space that is most important, reducing the total number of electron configurations that need to be treated relative to CASSCF.
- Treat some parts of the molecule less accurately than others using the ONIOM layered method (Section 2.3.2; partition—hybrid molecule). Relative to CASSCF, this reduces the number of integrals to be computed involving orbitals outside the active space, as well as potentially reducing the size of the active space.
- Parametrize a less expensive model that can then be applied to the full system and all of the active electrons, such as MMVB (Section 2.3.1; partition—hybrid atom). Here, the integral calculation is avoided completely, and the number of configurations reduced with respect to CASSCF.

Alternative strategies not covered here include: using a faster conventional method such as TDDFT [3,68] (and references cited therein), although this is a different method from CASSCF, with its own advantages and problems; or parameterize the integrals for a CASSCF-like method, which still leaves a large eigenvalue problem to solve for many active electrons [46,84,145–147].

For the three approximations to CASSCF discussed in this article, the examples chosen suggest the following guidelines.

When the entire molecule is effectively the chromophore or reaction centre, hybrid-molecule approaches such as ONIOM may not be appropriate, as there is no well-defined ‘active part’ of a molecule. Typical examples of this situation are provided by the family of polycyclic aromatic hydrocarbon (PAH) compounds, including the pyracylene (Section 3) and triangulene (Section 4) molecules that we have studied here. For middle-sized 14–20 electron active spaces, both MMVB and RASSCF could be used, but RAS is certainly the method that will provide the more precise results, and was therefore the natural choice to study the 14-electron pyracylene system. From a 20-electron large active space and beyond, MMVB is the only way to undertake a computational study (in contrast to doing a single very expensive calculation) and therefore was employed to study the 22-electron triangulene species. However, for the cyclohexadiene + naphthalene study, MMVB on its own was found to have problems, and ONIOM to be useful.

When the system can be partitioned into a small enough active region and a spectator region, then the natural approach is a hybrid-molecule formalism such as ONIOM. Typically, a full CASSCF is carried out on the active but smaller part of the molecule, whereas the remainder is treated with a lower level of theory. This was the case in our study of the geometries and energetics of the ground and excited states of an interstitial nickel defect in diamond (Section 5). Here, the lower level of theory was used to model a steric effect on a chromophore/active site, indirectly affecting the splitting between ground and excited states.

Between these truncation and partitioning approaches, we experimented with a third way that might be a good general alternative when the excitation is essentially (but not fully) localized in a small part of the system. In such cases, it might still be advantageous to perform full CAS calculations on the region of the chromophore where the excitation is mostly localized, and treat the excitation tail with a lower-level method that necessarily involves a truncated CI expansion. The effectiveness of this approach was demonstrated by ONIOM(CAS:MMVB) calculations on the 14-electron photochemical cycloaddition of cyclohexadiene and naphthalene (Section 6).

A question remains concerning the usefulness of the methods discussed in this article for the future, with computational speeds increasing and computer architectures changing. Increased performance depends on being able to take advantage of parallelism, rather than relying on faster individual processors. This is possible with a large CAS active space, where the generation of electron configurations and matrix elements – which can be parallelised and scales well over many processors – dominates the integral evaluation and linear algebra. Consequently, prototype full CAS and RAS calculations on larger molecules will get faster, but the range of molecules accessible with the hybrid methods we have discussed will also increase. In other words, hybrid methods will not go out of date quickly; instead we predict that they will become more useful, despite the need to explicitly identify the active site and combination of methods required.

Our hope is that this will contribute to the development of well-defined, systematic and reproducible approximations (*i.e.* a model chemistry) for large-scale excited-state calculations in chemistry, biology and materials science.

## Acknowledgements

We thank Dr. Martin Paterson (Heriot Watt University) for earlier assistance with the Ni in diamond CASSCF calculations, and Prof. Alison Mainwood (King’s College London) for suggesting this problem. We also thank Kate Hall for comments on this article. The work described here was supported in part by EPSRC via grants GR/R75236/02 (SL) and GR/S94704/01 (MBP). CASSCF calculations for pyracylene were run using the EPSRC National Service for Computational Chemistry Software ([www.nscs.ac.uk](http://www.nscs.ac.uk)).

## References

- [1] A.H. Zewail, *J. Phys. Chem. A* 104 (2000) 5660–5694.
- [2] J.C. Tully, *Faraday Discuss.* 127 (2004) 463–466.
- [3] A. Dreuw, *Chem. Phys. Chem.* 7 (2006) 2259–2274.
- [4] M. Garavelli, *Theor. Chem. Acc.* 116 (2006) 87–105.
- [5] L. Blancafort, F. Ogliaro, M. Olivucci, M.A. Robb, M.J. Bearpark, A. Sinicropi, in: A.G. Kutateladze (Ed.), *Computational Methods in Photochemistry*, CRC Press, Boca Raton, 2005, pp. 31–110.
- [6] M.J. Bearpark, M.A. Robb, H.B. Schlegel, *Chem. Phys. Lett.* 223 (1994) 269–274.
- [7] G.J. Atchity, S.S. Xantheas, K. Ruedenberg, *J. Chem. Phys.* 95 (1991) 1862–1876.
- [8] D.R. Yarkony, *J. Phys. Chem. A* 105 (2001) 6277–6293.
- [9] D.R. Yarkony, *Acc. Chem. Res.* 31 (1998) 511–518.

- [10] W. Domcke, D.R. Yarkony, H. Köppel (Eds.), *Conical intersections: Electronic Structure, Dynamics and Spectroscopy*, World Scientific, Singapore, 2004.
- [11] S. Dapprich, I. Komaromi, K.S. Byun, K. Morokuma, M. Frisch, *J. Mol. Struct. Theochem.* 462 (1999) 1–21.
- [12] T. Vreven, K.S. Byun, I. Komaromi, S. Dapprich, J.A. Montgomery, K. Morokuma, M.J. Frisch, *J. Chem. Theor. Comput.* 2 (2006) 815–826.
- [13] T. Vreven, K. Morokuma, *J. Comput. Chem.* 21 (2000) 1419–1432.
- [14] T. Vreven, K. Morokuma, *J. Chem. Phys.* 113 (2000) 2969–2975.
- [15] T. Vreven, K. Morokuma, O. Farkas, H.B. Schlegel, M.J. Frisch, *J. Comput. Chem.* 24 (2003) 760–769.
- [16] A. Warshel, M. Levitt, *J. Mol. Biol.* 103 (1976) 227–249.
- [17] M.J. Field, P.A. Bash, M. Karplus, *J. Comput. Chem.* 11 (1990) 700–733.
- [18] U.C. Singh, P.A. Kollman, *J. Comput. Chem.* 7 (1986) 718–730.
- [19] M.J. Bearpark, F. Bernardi, M. Olivucci, M.A. Robb, *J. Phys. Chem. A* 101 (1997) 8395–8401.
- [20] M.J. Bearpark, M. Boggio-Pasqua, *Theor. Chem. Acc.* 110 (2003) 105–114.
- [21] M.J. Bearpark, M. Boggio-Pasqua, M.A. Robb, F. Ogliaro, *Theor. Chem. Acc.* 116 (2006) 670–682.
- [22] M.J. Bearpark, M.A. Robb, F. Bernardi, M. Olivucci, *Chem. Phys. Lett.* 217 (1994) 513–519.
- [23] F. Bernardi, M. Olivucci, M.A. Robb, *J. Am. Chem. Soc.* 114 (1992) 1606–1616.
- [24] M. Garavelli, F. Ruggeri, F. Ogliaro, M.J. Bearpark, F. Bernardi, M. Olivucci, M.A. Robb, *J. Comput. Chem.* 24 (2003) 1357–1363.
- [25] M. Klene, M.A. Robb, L. Blancafort, M.J. Frisch, *J. Chem. Phys.* 119 (2003) 713–728.
- [26] P.A. Malmqvist, A. Rendell, B.O. Roos, *J. Phys. Chem.* 94 (1990) 5477–5482.
- [27] J. Olsen, B.O. Roos, P. Jorgensen, H.J.A. Jensen, *J. Chem. Phys.* 89 (1988) 2185–2192.
- [28] L.A. Curtiss, K. Raghavachari, P.C. Redfern, V. Rassolov, J.A. Pople, *J. Chem. Phys.* 109 (1998) 7764–7776.
- [29] L.A. Curtiss, K. Raghavachari, G.W. Trucks, J.A. Pople, *J. Chem. Phys.* 94 (1991) 7221–7230.
- [30] J.A. Pople, M. Head-Gordon, D.J. Fox, K. Raghavachari, L.A. Curtiss, *J. Chem. Phys.* 90 (1989) 5622–5629.
- [31] J.A. Montgomery, M.J. Frisch, J.W. Ochterski, G.A. Petersson, *J. Chem. Phys.* 112 (2000) 6532–6542.
- [32] J.A. Montgomery, M.J. Frisch, J.W. Ochterski, G.A. Petersson, *J. Chem. Phys.* 110 (1999) 2822–2827.
- [33] J.W. Ochterski, G.A. Petersson, J.A. Montgomery, *J. Chem. Phys.* 104 (1996) 2598–2619.
- [34] J.A. Montgomery, J.W. Ochterski, G.A. Petersson, *J. Chem. Phys.* 101 (1994) 5900–5909.
- [35] S. Parthiban, J.M.L. Martin, *J. Chem. Phys.* 114 (2001) 6014–6029.
- [36] J.M.L. Martin, G. de Oliveira, *J. Chem. Phys.* 111 (1999) 1843–1856.
- [37] G.A. Petersson, D.K. Malick, M.J. Frisch, M. Braunstein, *J. Chem. Phys.* (2006) 125.
- [38] M.J. Bearpark, M.A. Robb, in: M.S. Platz, R.A. Moss, M. Jones Jr. (Eds.), *Reviews of Reactive Intermediate Chemistry*, Wiley, 2007, pp. 379–414.
- [39] G.A. Worth, M.A. Robb, M.J. Bearpark, in: M. Olivucci (Ed.), *Computational Photochemistry*, Elsevier, Amsterdam, 2005, pp. 171–190.
- [40] M.J. Paterson, M.J. Bearpark, M.A. Robb, L. Blancafort, G.A. Worth, *PCCP* 7 (2005) 2100–2115.
- [41] F. Sicilia, M.J. Bearpark, L. Blancafort, M.A. Robb, *Theor. Chem. Acc.* 118 (2007) 241–251.
- [42] F. Sicilia, M.J. Bearpark, M.A. Robb, L. Blancafort, *J. Phys. Chem. A* 111 (2007) 2182–2192.
- [43] N. Yamamoto, T. Vreven, M.A. Robb, M.J. Frisch, H.B. Schlegel, *Chem. Phys. Lett.* 250 (1996) 373–378.
- [44] L. Blancafort, P. Celani, M.J. Bearpark, M.A. Robb, *Theor. Chem. Acc.* 110 (2003) 92–99.
- [45] M. Boggio-Pasqua, M.J. Bearpark, F. Ogliaro, M.A. Robb, *J. Am. Chem. Soc.* 128 (2006) 10533–10540.
- [46] T. Martinez, *J. Acc. Chem. Res.* 39 (2006) 119–126.
- [47] B.K. Carpenter, *Chem. Soc. Rev.* 35 (2006) 736–747.
- [48] B.K. Carpenter, *Angew. Chem. Int. Ed.* 37 (1998) 3341–3350.
- [49] P. Celani, H.J. Werner, *J. Chem. Phys.* 112 (2000) 5546–5557.
- [50] B.O. Roos, *Acc. Chem. Res.* 32 (1999) 137–144.
- [51] J. Finley, P.A. Malmqvist, B.O. Roos, L. Serrano-Andres, *Chem. Phys. Lett.* 288 (1998) 299–306.
- [52] P.J. Knowles, H. Werner, *J. Theor. Chim. Acta* 84 (1992) 95–103.
- [53] J.M.O. Matos, B.O. Roos, P.A. Malmqvist, *J. Chem. Phys.* 86 (1987) 1458–1466.
- [54] B.O. Roos, K. Andersson, M.P. Fulscher, *Chem. Phys. Lett.* 192 (1992) 5–13.
- [55] I.J. Palmer, I.N. Ragazos, F. Bernardi, M. Olivucci, M.A. Robb, *J. Am. Chem. Soc.* 115 (1993) 673–682.
- [56] R. Pauncz, *The Symmetric Group in Quantum Chemistry*, CRC, Boca Raton, 1995.
- [57] M. Garavelli, T. Vreven, P. Celani, F. Bernardi, M.A. Robb, M. Olivucci, *J. Am. Chem. Soc.* 120 (1998) 1285–1288.
- [58] T. Vreven, F. Bernardi, M. Garavelli, M. Olivucci, M.A. Robb, H.B. Schlegel, *J. Am. Chem. Soc.* 119 (1997) 12687–12688.
- [59] M. Garavelli, P. Celani, F. Bernardi, M.A. Robb, M. Olivucci, *J. Am. Chem. Soc.* 119 (1997) 6891–6901.
- [60] S. Capponi, N. Guihery, J.P. Malrieu, B. Miguel, D. Poilblanc, *Chem. Phys. Lett.* 255 (1996) 238–243.
- [61] N. Guihery, D. Maynau, J.P. Malrieu, *Chem. Phys. Lett.* 248 (1996) 199–206.
- [62] N. Guihery, J.P. Malrieu, D. Maynau, K. Handrick, *Int. J. Quant. Chem.* 61 (1997) 45–54.
- [63] M.J. Bearpark, P. Celani, F. Jolibois, M. Olivucci, M.A. Robb, F. Bernardi, *Mol. Phys.* 96 (1999) 645–652.
- [64] M. Klene, Ph.D. Thesis. King’s College London, 2002.
- [65] M. Boggio-Pasqua, M.J. Bearpark, M. Klene, M.A. Robb, *J. Chem. Phys.* 120 (2004) 7849–7860.
- [66] G.E. Scuseria, *J. Phys. Chem. A* 103 (1999) 4782–4790.
- [67] R. Bauernschmitt, R. Ahlrichs, *Chem. Phys. Lett.* 256 (1996) 454–464.
- [68] A. Dreuw, M. Head-Gordon, *Chem. Rev.* 105 (2005) 4009–4037.
- [69] K. Morokuma, D.G. Musaev, T. Vreven, H. Basch, M. Torrent, D. Khoroshun, V. IBM, *J. Res. Dev.* 45 (2001) 367–395.
- [70] H. Lin, D.G. Truhlar, *Theor. Chem. Acc.* 117 (2007) 185–199.
- [71] J. Gao, M.A. Thompson (Eds.), *Combined Quantum Mechanical and Molecular Mechanical Methods*, ACS, Washington, 1998.
- [72] S.E. Butcher, J.M. Burke, *Biochemistry* 33 (1994) 992–999.
- [73] J. Gao, in: K.B. Lipkowitz, D.B. Boyd (Eds.), *Reviews in Computational Chemistry*, 7, VCH Publishers, New York, 1995, pp. 119–185.
- [74] P. Sherwood, in: J. Grotendorst (Ed.), *Modern Methods and Algorithms of Quantum Chemistry Proceedings*, vol. 3, second ed., John von Neumann Institute for Computing, Jülich, 2000, pp. 285–305.
- [75] T. Andrniow, N. Ferre, M. Olivucci, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 101, 2004, pp. 17908–17913.
- [76] N. Ferre, M. Olivucci, *J. Am. Chem. Soc.* 125 (2003) 6868–6869.
- [77] A. Migani, A. Sinicropi, N. Ferre, A. Cembran, M. Garavelli, M. Olivucci, *Faraday Discuss.* 127 (2004) 179–191.
- [78] P.B. Coto, A. Sinicropi, L. De Vico, N. Ferre, M. Olivucci, *Mol. Phys.* 104 (2006) 983–991.
- [79] P.B. Coto, A. Strambi, N. Ferre, M. Olivucci, *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, 2006, pp. 17154–17159.
- [80] U.F. Rohrig, L. Guidoni, A. Laio, I. Frank, U. Rothlisberger, *J. Am. Chem. Soc.* 126 (2004) 15328–15329.
- [81] J.A. Gascon, V.S. Batista, *Biophys. J.* 87 (2004) 2931–2941.
- [82] T. Vreven, K. Morokuma, *Theor. Chem. Acc.* 109 (2003) 125–132.
- [83] L. Serrano-Andres, M. Merchan, R. Lindh, *J. Chem. Phys.* (2005) 122.
- [84] R. Izzo, M. Klessinger, *J. Comput. Chem.* 21 (2000) 52–62.
- [85] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery, T. Vreven, G. Scalmani, K.N. Kudin, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, X. Li, H.P. Hratchian, J.E. Peralta, A.F. Izmaylov, E. Brothers, V. Staroverov,

- R. Kobayashi, J. Normand, J.C. Burant, J.M. Millam, M. Klene, J.E. Knox, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, W. Chen, M.W. Wong, J.A. Pople, Gaussian Development Version Revision E.05, Gaussian Inc., Wallingford CT, 2006.
- [86] N.L. Allinger, in: V.I. Gold, D. Bethell (Eds.), *Advances in Physical Organic Chemistry*, 13, Academic Press, New York, 1976, pp. 1–82.
- [87] N.L. Allinger, *J. Am. Chem. Soc.* 99 (1977) 8127–8134.
- [88] P. Durand, J.P. Malrieu, *Advances in Chemical Physics*, 67, Wiley, New York, 1987, pp. 321–412.
- [89] F. Bernardi, M. Olivucci, J.J.W. McDouall, M.A. Robb, *J. Chem. Phys.* 89 (1988) 6365–6375.
- [90] F. Bernardi, M. Olivucci, I.N. Ragazos, M.A. Robb, *J. Am. Chem. Soc.* 114 (1992) 8211–8220.
- [91] M. Said, D. Maynau, J.P. Malrieu, *J. Am. Chem. Soc.* 106 (1984) 580–587.
- [92] M. Said, D. Maynau, J.P. Malrieu, M.A.G. Bach, *J. Am. Chem. Soc.* 106 (1984) 571–579.
- [93] G. Treboux, D. Maynau, J.P. Malrieu, *J. Phys. Chem.* 99 (1995) 6417–6423.
- [94] J.W. Ponder, *Tinker 4.2*, Washington, 2004.
- [95] J.W. Ponder, F.M. Richards, *J. Comput. Chem.* 8 (1987) 1016–1024.
- [96] S. Humbel, S. Sieber, K. Morokuma, *J. Chem. Phys.* 105 (1996) 1959–1967.
- [97] M. Svensson, S. Humbel, R.D.J. Froese, T. Matsubara, S. Sieber, K. Morokuma, *J. Phys. Chem.* 100 (1996) 19357–19363.
- [98] T. Vreven, K. Morokuma, *J. Phys. Chem. A* 106 (2002) 6167–6170.
- [99] T. Vreven, K. Morokuma, in: D.C. Spellmeyer (Ed.), *Annual Reports in Computational Chemistry*, vol. 2, Elsevier, Amsterdam, 2006, pp. 37–53.
- [100] S. Larkin, M.J. Bearpark, T. Vreven, in preparation.
- [101] J.T. Fermann, T. Moniz, O. Kiowski, T.J. McIntire, S.M. Auerbach, T. Vreven, M.J. Frisch, *J. Chem. Theor. Comput.* 1 (2005) 1232–1239.
- [102] T. Vreven, M.J. Frisch, K.N. Kudin, H.B. Schlegel, K. Morokuma, *Mol. Phys.* 104 (2006) 701–714.
- [103] B. Freiermuth, S. Gerber, A. Riesen, J. Wirz, M. Zehnder, *J. Am. Chem. Soc.* 112 (1990) 738–744.
- [104] M. Boggio-Pasqua, M.A. Robb, M.J. Bearpark, *J. Phys. Chem. A* 109 (2005) 8849–8856.
- [105] B.O. Roos, *Advances in Chemical Physics*, vol. 69, Wiley, 1987, p. 399.
- [106] M.J. Bearpark, F. Bernardi, S. Clifford, M. Olivucci, M.A. Robb, B.R. Smith, T. Vreven, *J. Am. Chem. Soc.* 118 (1996) 169–175.
- [107] M. Sarobe, S. Flink, L.W. Jenneskens, B.L.A. Vanpoecke, J.W. Zwikker, *J. Chem. Soc. Chem. Commun.* (1995) 2415–2416.
- [108] M. Sarobe, J.D. Snoeijer, L.W. Jenneskens, M.Q. Slagt, J.W. Zwikker, *Tetrahedron Lett.* 36 (1995) 8489–8492.
- [109] R. Dunsbach, R. Schmidt, *J. Photochem. Photobiol. A Chem.* 83 (1994) 7–13.
- [110] F. Schael, H.G. Lohmannsroben, *J. Photochem. Photobiol. A Chem.* 69 (1992) 27–32.
- [111] C. Gooijer, I. Kozin, N.H. Velthorst, M. Sarobe, L.W. Jenneskens, E. Vlietstra, *J. Spectrochim. Acta Part A—Mol. Biomol. Spectrosc.* 54 (1998) 1443–1449.
- [112] N.D. Marsh, C.J. Mikolajczak, M. Wornat, *J. Spectrochim. Acta Part A—Mol. Biomol. Spectrosc.* 56 (2000) 1499–1511.
- [113] M. Boggio-Pasqua, M. Ravaglia, M.J. Bearpark, M. Garavelli, M.A. Robb, *J. Phys. Chem. A* 107 (2003) 11139–11152.
- [114] P. Celani, S. Ottani, M. Olivucci, F. Bernardi, M.A. Robb, *J. Am. Chem. Soc.* 116 (1994) 10141–10151.
- [115] J.A. Berson, *Acc. Chem. Res.* 30 (1997) 238–244.
- [116] E. Clar, D.G. Stewart, *J. Am. Chem. Soc.* 75 (1953) 2667–2672.
- [117] M.J. Bearpark, M.A. Robb, N. Yamamoto, *J. Spectrochim. Acta Part A—Mol. Biomol. Spectrosc.* 55 (1999) 639–646.
- [118] G. Allinson, R.J. Bushby, J.L. Paillaud, D. Oduwale, K. Sales, *J. Am. Chem. Soc.* 115 (1993) 2062–2064.
- [119] G. Allinson, R.J. Bushby, J.L. Paillaud, *J. Mater. Sci. Mater. Electron.* 5 (1994) 67–74.
- [120] G. Allinson, R.J. Bushby, J.L. Paillaud, M. Thorntonpett, *J. Chem. Soc. Perkin Trans. 1* (1995) 385–390.
- [121] G. Allinson, R.J. Bushby, M.V. Jesudason, J.L. Paillaud, N. Taylor, *J. Chem. Soc. Perkin Trans. 2* (1997) 147–156.
- [122] K. Fukui, J. Inoue, T. Kubo, S. Nakazawa, T. Aoki, Y. Morita, K. Yamamoto, K. Sato, D. Shiomi, K. Nakasuji, T. Takui, *Synth. Met.* 121 (2001) 1824–1825.
- [123] J. Inoue, K. Fukui, T. Kubo, S. Nakazawa, K. Sato, D. Shiomi, Y. Morita, K. Yamamoto, T. Takui, K. Nakasuji, *J. Am. Chem. Soc.* 123 (2001) 12702–12703.
- [124] L. Blancafort, W. Adam, D. Gonzalez, M. Olivucci, T. Vreven, M.A. Robb, *J. Am. Chem. Soc.* 121 (1999) 10583–10590.
- [125] L. Allers, A. Mainwood, *Diamond Relat. Mater.* 7 (1998) 261–265.
- [126] I. Kiflawi, H. Kanda, A. Mainwood, *Diamond Relat. Mater.* 7 (1998) 327–332.
- [127] A. Mainwood, *Diamond Relat. Mater.* 8 (1999) 1560–1564.
- [128] K. Johnston, A. Mainwood, *Diamond Relat. Mater.* 11 (2002) 631–634.
- [129] M. Watkins, A. Mainwood, *J. Phys. Condens. Matter* 15 (2003) S2913–S2927.
- [130] K. Johnston, A. Mainwood, A.T. Collins, G. Davies, D. Twitchen, M. Newton, J.M. Baker, *Diamond Relat. Mater.* 9 (2000) 424–427.
- [131] P.W. Mason, F.S. Ham, G.D. Watkins, *Phys. Rev. B* 60 (1999) 5417–5428.
- [132] J.P. Goss, P.R. Briddon, R. Jones, S. Oberg, *J. Phys. Condens. Matter* 16 (2004) 4567–4578.
- [133] J.E. Lowther, *Phys. Rev. B* 51 (1995) 91–96.
- [134] R. Larico, W.V.M. Machado, L.V.C. Assali, J.F. Justo, *Diamond Relat. Mater.* 14 (2005) 380–382.
- [135] B.N. Figgis, M.A. Hitchman, *Ligand Field Theory and its Applications*, Wiley-VCH, New York, 2000.
- [136] E. Constantino, A. Rimola, L. Rodriguez-Santiago, M. Sodupe, *New J. Chem.* 29 (2005) 1585–1593.
- [137] A.K. Rappe, C.J. Casewit, K.S. Colwell, W.A. Goddard, W.M. Skiff, *J. Am. Chem. Soc.* 114 (1992) 10024–10035.
- [138] K. Iakubovskii, G. Davies, *Phys. Rev. B* (2004) 70.
- [139] L.E. Orgel, *An Introduction to Transition Metal Chemistry: Ligand Field Theory*, Methuen, London, 1960.
- [140] M.J. Bearpark, M. Deumal, M.A. Robb, T. Vreven, N. Yamamoto, M. Olivucci, F. Bernardi, *J. Am. Chem. Soc.* 119 (1997) 709–718.
- [141] M. Deumal, M.J. Bearpark, B.R. Smith, M. Olivucci, F. Bernardi, M.A. Robb, *J. Org. Chem.* 63 (1998) 4594–4600.
- [142] J. McCullough, *J. Chem. Rev.* 87 (1987) 811–860.
- [143] N.C. Yang, J. Masnovi, W. Chiang, T. Wang, H. Shou, D.D.H. Yang, *Tetrahedron* 37 (1981) 3285–3300.
- [144] N.C. Yang, J. Libman, M.F. Savitzky, *J. Am. Chem. Soc.* 94 (1972) 9226–9227.
- [145] A. Toniolo, A.L. Thompson, T. Martinez, *J. Chem. Phys.* 304 (2004) 133–145.
- [146] M. Hoffmann, M. Wanko, P. Strodel, P.H. Konig, T. Frauenheim, K. Schulten, W. Thiel, E. Tajkhorshid, M. Elstner, *J. Am. Chem. Soc.* 128 (2006) 10808–10818.
- [147] M. Wanko, M. Hoffmann, P. Strodet, A. Koslowski, W. Thiel, F. Neese, T. Frauenheim, M. Elstner, *J. Phys. Chem. B* 109 (2005) 3606–3615.