Theoretical Modelling of Hydrogen Transfer Reactions

Róbert Izsák

Thesis submitted for the Degree of PhD

Theoretical Chemistry Group
School of Chemistry
Cardiff University

July 31, 2010

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Declaration

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed ..... (candidate) Date 31/07/10

Statement 1

This thesis is being submitted in partial fulfilment of the requirements for the degree of PhD.

Signed ..... (candidate) Date 31/07/10

Statement 2

This thesis is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references.

Signed ..... (candidate) Date 31/07/10

Statement 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ..... (candidate) Date 31/07/10

Acknowledgements

I would like to thank my supervisor, *Prof. Peter J. Knowles* for his help during my PhD studies, for his patient and always insightful, yet simple explanations of the most difficult topics. To *Dr. Milán Szőri*, coauthor of the paper that forms the second part of the thesis, for always finding plenty of time for discussion, reading through a significant portion of this thesis and his help with some of the figures. To *Prof. Béla Viskolcz*, for helping me through a number of practical difficulties, without which this thesis would certainly never have been written. To *Prof. Gyula Tasi*, the enthusiastic quantum chemist and my first teacher in the field for his memorable and inspiring classes. To *Prof. Imre G. Csizmadia*, who first introduced me to an international research group and helped me through the difficulties of my first publications. To *Dr. Marcelo F. Masman* for many discussions and for keeping the soul in me under circumstances nobody else could have done. And finally to *Ms. Bridgette Cooper* for many hours of conversation discussing among others each other's PhD topics, and *Dr. Norbert Fekete* for his friendly advice which although solely in an indirect manner but still greatly influenced the final form of this thesis.

"Chemistry is that branch of natural philosophy in which the greatest improvements have been and may be made: it is on that account that I have made it my peculiar study; but at the same time, I have not neglected the other branches of science. A man would make but a very sorry chemist if he attended to that department of human knowledge alone. If your wish is to become really a man of science, and not merely a petty experimentalist, I should advise you to apply to every branch of natural philosophy, including mathematics."

Waldman, professor of chemistry,
in Mary Shelley's Frankenstein (1818)

古池や　　*Furu ike ya*
蛙飛こむ　*kawazu tobikomu*
水のをと　　*mizu no oto*

A lonely pond in age-old stillness sleeps . . .
Apart, unstirred by sound or motion . . . till
Suddenly into it a lithe frog leaps.

Matsuo Bashō's Frog Haiku (1686),
translated by Curtis Hidden Page

Abstract

The first part of this thesis deals with some general aspects of hydrogen transfer reactions. Based on the idea of similarity between localized orbitals of functional groups in different molecules, an attempt is made to reflect this transferability in segments of the correlation energy belonging to the set of orbitals of a certain functional group. Various possibilities are examined for such partitioning. It turns out that localized orbitals are the best choice for this purpose since other transformations delocalize orbitals, and transferability is lost. Possibly, quantum chemical treatment would only be necessary with certain partitions, the rest could be transferred from previous calculations. Ideally, only contributions from "reacting orbitals", expected to carry most of the static correlation effects, would need to be calculated. For this reason, the transferability of segments of various compounds was checked. Alkanes were found to behave ideally, whereas molecules with large electrostatic differences, or delocalized electronic structure are harder cases allowing only a more qualitative use of the proposed method.

In the second half, the energetics of terminal and central OH-additions as well as allylic H-abstractions by OH in its reaction with propene was studied using several single and multireference ab initio techniques. Selection of the localized occupied orbitals forming the active space for multireference methods is discussed. Initial geometry optimizations and vibrational frequency analysis were carried out at the [5, 5]-CASPT2/cc-pVTZ level of theory. Multireference effects turned out to be negligible and the UCCSD(T)/cc-pVTZ model was chosen for final geometry optimizations and vibrational frequency analysis. Triples contributions are found to be very important, except for the π-complex, which has a UCCSD(T)/CBS relative enthalpy of -10.56 kJ/mol compared to infinitely separated propene + OH. The addition transition states are found to have relative enthalpies of -9.93 kJ/mol for the central and -9.84 kJ/mol for the terminal case. Allylic abstraction mechanisms, although lying significantly higher, still have only slightly positive barriers – a value of 3.21 kJ/mol for the direct and 1.67 kJ/mol for the consecutive case. Conventional transition state theory was used as a rough estimation for determining rate constants which turned out to agree well with experiment.

Contents

Notations

$\hat{O}$	operators corresponding (unless otherwise noted) to the expectation value O or $\langle\hat{O}\rangle$
Ψ	the exact or correlated wavefunction
$\Phi, \Phi_0, \Phi_i^b, \Phi_{ij}^{ab}, \ldots$	approximate determinant wavefunctions, with indices noting the ground (reference) and excited states with excitations from $\{i, j \ldots\}$ to $\{a, b \ldots\}$; corresponding to the set of coefficients (amplitudes) denoted by $\{c_i^a, \ldots\}$ or $\{t_i^a, \ldots\}$
ϕ_i	spin orbitals
φ_i	spatial orbitals
ϑ_i	Kohn-Sham orbitals
ψ_i	unspecified molecular orbitals
η_i	atomic orbitals
$F[\rho]$	F is a functional of the function ρ
$\mathbf{A}$	denotes matrices with elements A_{ij}
$\vec{v}, \underline{v}$	denote vectors with elements r_i
$\langle i\|\hat{O}\|j\rangle$	elements of the matrix representation of $\hat{O}$ in the basis $\{i, j \ldots\}$
$\langle i\|j\rangle$	overlap integral (scalar product) of orbitals (elements) labelled i, j
$\langle i\|\hat{h}\|j\rangle$	one electron integrals
$\langle ij\|kl\rangle$	two electron integrals, with $\langle ij\|ij\rangle$ being the Coulomb, $\langle ij\|ji\rangle$ the exchange integrals
$(i\|\hat{h}\|j), (ij\|kl)$	charge cloud notation of integrals, used with spatial orbitals only, the relation to the previous notation is $\langle i\|\hat{h}\|j\rangle = (i\|\hat{h}\|j)$ and $\langle ij\|kl\rangle = (ik\|jl)$
$\{\alpha_i\}$	a set of elements α_i

1 Introduction

Molecular electronic structure theory dates back to the foundation of quantum mechanics itself. The complexity of the relevant equations prohibited a direct solution, but approximate solutions were to follow soon (e.g. Hückel[1]). However, it was not until some 50 years ago that what is now referred to as quantum chemistry started off to become a theory suitable for quantitative predictions.[2] Perhaps the easiest comparison between theory and experiment can be made through reactions in gas phase, since the connection between calculated results and observable quantities can be estimated via statistical mechanics, and one does not need to worry about some complications that arise with condensed phase systems. Hydrogen transfer reactions are an important class of reactions, many of which take place in gas phase. Relevant atmospheric and industrially important reactions belong here. Solvation effects can also be considered, and various catalytic and biochemical reactions can be studied. In the first part of this thesis, these reactions are studied from a more general point of view to observe some general characteristics of how correlation energy may be partitioned on a spatial basis, which might be useful to consider for some methodological developments (e.g. a more "black box" like selection of active spaces). The second part is an application of theory to obtain the characteristics of some of the relevant species in the propene+OH system. These latter results have been published recently.[3] After this section, a short introduction of theoretical methods will follow. After that, each section has a more specifically related introductory subsection. In the remainder of this section, the aim is to select a few topics of more general interest to illustrate the importance of hydrogen transfer reactions.

To begin with, there is a wide range of experimental measurement techniques which could be used in studying such reactions. Chemical and photoionization mass spectrometry are such methods. Reactants of known concentration are mixed and then the mixture is

ionized, and measured by mass spectrometry.[4] For reactions involving unstable radical species, shock tube measurements are very useful. Here, the reaction takes place in a wave front generated by a sudden burst of a diaphragm between high and low pressure regions; the products can then be detected by various means.[5] Cavity ringdown spectroscopy is another important way of measuring atmospheric and combustion reactions. Here, a laser is reflected between two mirrors in a tube (the cavity), so that it in resonance with the mode of the cavity, and when it is turned off, an exponential decay of light intensity can be measured. If in a reaction a species can absorb the reflected light, the ringdown is faster, i.e. a faster exponential drop of light intensity is measured,[6] which can be used for various analytical purposes. Hydrogen exchange between halogens was in the focus of transition state spectroscopy for long.[7] In photodetachment spectroscopy, the species are ionized and mass selected, after which photoelectron detachment from XHX^- species is measured, and the spectra are found to have a fine structure belonging to long lived states of the neutral XHX species.[8]

Atmospheric chemistry, which deals with slow oxidative reactions taking place in the atmosphere, and the closely related field of rapid oxidative reactions covered by combustion chemistry, are both areas which attract considerable interest owing to their environmental and industrial importance. Both fields deal with systems involving chain reactions of radicals with competitive reaction channels, necessitating highly accurate experimental and theoretical treatment (a difference of half a kcal/mol in reaction barrier heights can change the whole kinetic profile of the system). Among these, various H abstraction reactions are very significant. The majority of atmospheric reactions is initialized by OH, which in turn mainly comes from the abstraction of an H from water by an O atom (a product of ozone photolysis).[9] In the followings, our main concern will be the various reactions of OH with hydrocarbons. Alkyl radicals are typically products of OH initialized reactions, and usually participate in further reactions resulting in various products, mainly oxygen containing species.[10] Alkoxy radicals play an important role among these. These radicals may go on dissociating to smaller fragments, or further H abstraction may occur by O_2 or in the form of intramolecular H abstraction (1,5 H-shift).[11] Another source of alkoxy radicals is H abstraction from the hydroxyl group of alcohols by OH. This and other possible H abstractions play an important role in the decomposition of alternative fuels like methanol or ethanol, but also in similar processes with alternative chlorofluoro-

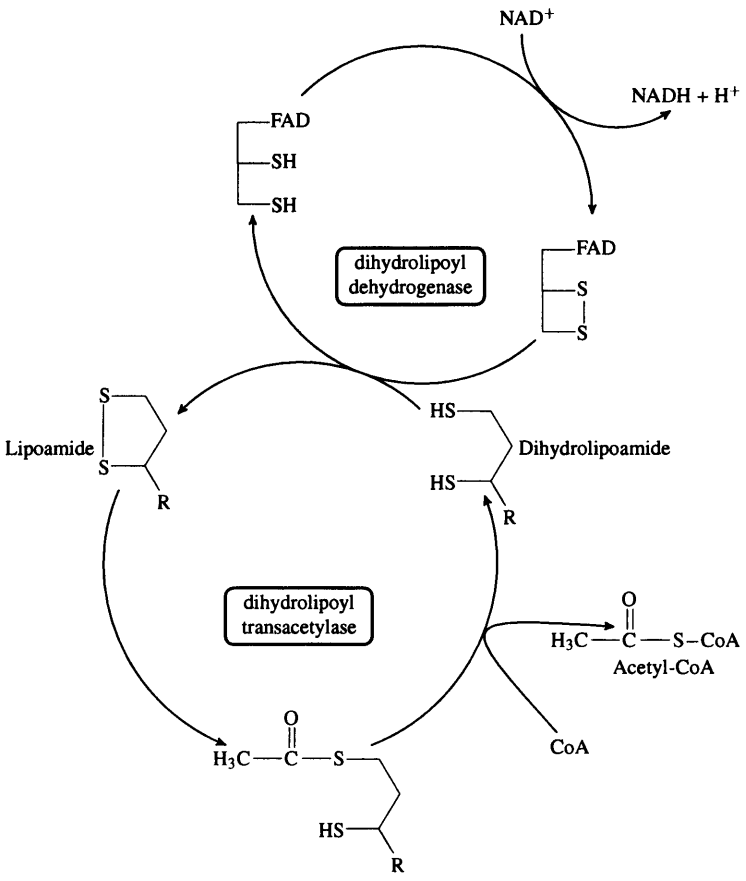

Figure 1: *Lipoamide as a redox catalyst in the formation of Acetyl-CoA*

carbons.[12] The dominant reaction with alkenes is OH addition to the double bond, but H abstraction may play a role at higher temperatures.[13,14] As far as aromatic compounds are concerned, the dominant reaction of benzene with OH is electrophilic addition to the benzene ring, and again, abstraction may have a role at higher temperatures.[15]

There are many important hydrogen transfer related biologically important reactions, like the one in Figure 1. Hydrogen transfer as a term may refer to the actual transfer of hydrogen, proton or hydride anion; all of which procedures are catalysed by enzymes in living systems. An example to hydride transfer could be the reduction of acetaldehyde to ethanol. Actual hydrogen transfer takes place in many important biological processes, e.g. ageing and oxidative stress.[16] The role of vitamins (antioxidants) can be understood

in terms of hydrogen abstractions when radicals (e.g. OH) are stabilized by abstracting H from them, and are, therefore, prevented from causing oxidative damage through processes such as lipid peroxidation.[17] Proton transfer reactions are among the most frequent ones in biology since acid-base reactions belong here, and biochemical reactions are highly pH sensitive. The enzymatic catalysis of proton transfer from carbon atoms is a much studied field of biochemistry.[18]

Finally, it is perhaps worth mentioning that the author of this thesis has a previous publication,[19] which is related to the more general field of H-transfer reactions. In Figure 1, a biological analogue is presented to the reactions studied in this work, which deals with the formation of X–H bonds from X–X bonds by double H abstraction from DNA. Model compounds (X = O, S, Se) are used to study possible mechanisms of a putative drug, where the two strands of DNA become tied together as a result. DFT studies were carried out with and without considering solvent effects. Docking and molecular dynamics studies helped identify the proper enantiomers that can form an initial complex with the DNA helix. An attempt is made to explore the role of various potential energy surfaces at the low level of HF/3-21G.

2 Methodology

2.1 Hartree-Fock Theory And Electron Correlation

Quantum chemistry is the discipline that describes the electronic structure of chemical species, and predicts values of chemically important physical quantities.[2] As it is based on quantum mechanics, its central notion is that of the *wavefunction*, that describes a system's physical properties yielding observables through the action of *operators*, and determines the possible energetic states the system may assume through the *Schrödinger equation*. Because of the vast computational demands, many approximations are necessary. Many times, rather than from necessity, they are introduced due to our ignorance of certain effects that are negligible for the accurate description of certain physical phenomena. Such effects typically neglected and only considered in special cases include relativistic effects, non-adiabatic corrections or spin-orbit coupling among others. In most of the cases, quantum chemists work in the nonrelativistic framework, further assuming that the motion of electrons is separated from that of the nuclei. In the Born-Oppenheimer approximation,[20,21] we neglect all linking terms, thus separating the nuclear kinetic term of the operator from the rest, and in general practice, only the electronic part is solved, and the nuclear one only approximated:

$$\hat{H}\Psi = E\Psi \tag{1}$$

Here $\hat{H}$ is the Hamilton operator, describing the kinetic and potential energies of the system, Ψ is the electronic wavefunction, E is the total electron energy, and (1) is the electronic Schrödinger equation (note that contribution from the nuclear-nuclear potential term is also calculated and added to the energy, but no further attention will be given to

this fact in the following discussion). The expectation value of the energy can then be obtained from the Rayleigh quotient:

$$E = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \tag{2}$$

where the denominator equals 1 if Ψ is normalized. The next step aims at separating the variables of n electrons described by Ψ. For this, one needs to replace the two electron Coulomb potential terms of electronic repulsion involved in $\hat{H}$, by an effective one electron potential which arises from assuming the form of the wavefunction as an antisymmetrized product of one electron functions (ϕ_i). This latter condition is to allow for the Pauli principle, and requires the wavefunction to change sign if the coordinates of two electrons are exchanged. With this we arrive at the *independent particle*, or *mean field* approximation, where the wavefunction is approximated by a single Slater determinant[22] (Φ_0) constructed from the one electron functions ϕ_i, that has the desired antisymmetric property. From (2), we know how the energy depends on the wavefunction, and so we are looking for a set of one electron functions, and a wavefunction constructed from them, that minimize the system's total energy ($E[\Phi_0] = E[\{\phi_i\}]$). This is because the *variational principle* states that the exact ground state energy is a lower limit to the energy expectation value belonging to any proper trial function, such as Φ_0, and the two are only equal if the trial function is the exact ground state wavefunction. Therefore the best approximate Φ_0 to the exact wavefunction is the minimum energy one. For this purpose, the following functional is introduced:

$$\mathscr{E}[\{\phi_i\}] = E[\{\phi_i\}] + \sum_{i,j=1}^{n} \varepsilon_{ij}(\delta_{ij} - \langle \phi_i | \phi_j \rangle) \qquad \delta\mathscr{E}[\{\phi_i\}] = 0 \tag{3}$$

where $\mathscr{E}[\{\phi_i\}]$ is the functional the variation of which with respect to ϕ_i we would like to disappear in the Lagrangian multiplier procedure, which is equivalent to minimizing $E[\{\phi_i\}]$ ($\delta E[\{\phi_i\}] = 0$) with the constraint of the orthonormality of ϕ_i. δ_{ij} is the Kronecker delta, and ε_{ij} are the Lagrangian multipliers. Note that the it is possible to directly minimize $\delta E[\{\phi_i\}]$ without applying the Lagrangian method if orthonormality is not required. This yields:

$$\hat{f}\phi_i = \sum_{j=1}^{n} \varepsilon_{ij}\phi_j \qquad i = 1, \ldots, n \tag{4}$$

Here $\hat{f}$ is the Fock operator, and (4) are the *Hartree-Fock equations*.[22–25] The Fock operator is defined as:

$$\hat{f} = \hat{h} + \sum_{i=1}^{n} (\hat{J}_i - \hat{K}_i) \tag{5}$$

where $\hat{h}$ is a one electron operator, describing the kinetic and nuclear-electron potential energy of a single electron, and $\hat{J}_i$ and $\hat{K}_i$ are effective one electron operators:

$$\hat{J}_i\phi_j(2) = \langle\phi_i(1)|\frac{1}{r_{12}}|\phi_i(1)\rangle\phi_j(2) \qquad \hat{K}_i\phi_j(2) = \langle\phi_i(1)|\frac{1}{r_{12}}|\phi_j(1)\rangle\phi_i(2) \tag{6}$$

where $\hat{J}_i$ is the Coulomb operator, the expectation value of which is analogous with inter-electronic repulsion in classical physics, whereas the exchange operator $\hat{K}_i$ does not have such an analogue, but arises from the antisymmetrised quantum mechanical treatment. It exchanges functions to the right from the operator in which r_{12} is the interelectronic distance. Electronic labels 1 and 2 are indicated to help identify which electron belongs to which orbital.

Without going into much detail here, it can be proven that $\hat{f}$ is invariant under unitary transformations,[26] and, therefore, ϕ_i can be transformed in a way that only diagonal ε_{ij} survive. The resulting equations are the *canonical Hartree-Fock equations*:

$$\hat{f}\phi_i = \varepsilon_i\phi_i \qquad i = 1, \ldots, n \tag{7}$$

where we dropped the second index from ε_{ii}, and kept the notation ϕ_i for the transformed orbitals. (7) are eigenvalue equations, where $\hat{f}$ depends on the eigenfunctions ϕ_i, which makes iterative solution from an initial guess necessary. Solutions of these equations are referred to as Hartree-Fock orbitals (ϕ_i) and corresponding orbital energies (ε_i). Now,

if ϕ_i is separated into a product of a spatial term (φ_i) and a spin term being either α or β, spin variables can be integrated, and we gain the spatial form of the equations. Since the Pauli principle allows the description of two electrons by the same spatial orbital (with different spin terms), it is straightforward to restrict these spatial orbitals to be the same, in which case we get the restricted Hartree-Fock formalism (RHF), which reduces the number of equations to half of those with spin-orbitals. If this restriction is not enforced, one gets equations for α and β separately; this is called the unrestricted Hartree-Fock (UHF) formalism.

If we use the orbitals thus gained, we can find the expression for the total electron energy we are looking for. Here we are going to present the energy expressions of the Hartree-Fock energy with one electron or spin orbitals (E_0^{HF}) and also the corresponding expression for the closed shell restricted Hartree-Fock case (E_0^{RHF}) with spatial orbitals:

$$E_0^{HF} = \sum_i^n \langle \phi_i | \hat{h} | \phi_i \rangle + \frac{1}{2} \sum_{ij}^n (\langle \phi_i \phi_j | \phi_i \phi_j \rangle - \langle \phi_i \phi_j | \phi_j \phi_i \rangle) \tag{8}$$

and

$$E_0^{RHF} = 2 \sum_i^{n/2} \langle \varphi_i | \hat{h} | \varphi_i \rangle + \sum_{ij}^{n/2} (2\langle \varphi_i \varphi_j | \varphi_i \varphi_j \rangle - \langle \varphi_i \varphi_j | \varphi_j \varphi_i \rangle) \tag{9}$$

where we have used one electron integrals $\langle i|\hat{h}|i\rangle$ and two electron integrals $\langle ij|ij\rangle$ (Coulomb integrals) and $\langle ij|ji\rangle$ (exchange integrals) with spin-orbital and spatial orbital basis. It is the number of two electron integrals that grows rapidly with the number of electrons/orbitals making calculations less and less feasible with larger systems.

If we now consider the spatial form of the Hartree-Fock equations, which are very similar to (7), it turns out that although they are simpler compared to the original n-electron Schrödinger equation (1), they are still a system of differential equations. In yet another approximation, the spatial functions φ_i are expanded with a finite number of known basis functions η_μ. In the RHF case:

$$\hat{f}\varphi_i = \varepsilon_i\varphi_i \qquad \varphi_i = \sum_{\mu=1}^{M} c_{\mu i}\eta_\mu \qquad i = 1, \ldots, M \tag{10}$$

Since M is finite, there will only be a finite number of solutions for φ_i. However, this results in a matrix equation:

$$\mathbf{fc} = \mathbf{Sc}\boldsymbol{\varepsilon} \qquad S_{\mu\nu} = \langle\eta_\mu|\eta_\nu\rangle \qquad f_{\mu\nu} = \langle\eta_\mu|\hat{f}|\eta_\nu\rangle \tag{11}$$

where $f_{\mu\nu}$ is an element of the Fock matrix $\mathbf{f}$, $\mathbf{S}$ is the overlap matrix, $\mathbf{c}$ contains the expansion coefficients of the finite basis and $\boldsymbol{\varepsilon}$ is a diagonal matrix of orbital energies. (11) is the generalized eigenvalue equation form of the *Hartree-Fock-Roothan-Hall equations*,[27,28] which, by various orthonormalization procedures (in practice, it is done by orthonormalizing the AO basis) may be transformed into an eigenvalue equation yielding

$$\mathbf{f}'\mathbf{c}' = \mathbf{c}'\boldsymbol{\varepsilon} \tag{12}$$

The advantage of (12) is that it is an algebraic matrix equation, which has very efficient computational algorithms to be solved with. On the other hand, the solutions depend on M, the number and also on the quality of basis functions chosen. There are many varieties of basis sets available, here we only mention Dunning's correlation consistent basis set series,[29–31] abbreviated as: (aug-)cc-pVXZ, where the size of the basis grows with X. Here VXZ refers to "valence X-tuple zeta" suggesting that X contracted basis functions (contracted from a number of Gaussian primitives) are used to represent each valence orbital. The prefix "cc-p" stands for "correlation consistent polarized" indicating the fact that more and more functions with higher angular momentum quantum numbers than minimally necessary are involved (polarized functuions). The prefix "aug-" prefix means "augmented" and indicates the inclusion of diffuse functions (very shallow Gaussian functions). These basis sets are constructed in a manner to converge systematically as X increases. For this reason it is possible to use them in extrapolations introduced later. As we have mentioned before, the solution has to be iterative, due to the dependence of the Fock operator on the solutions. This means that following some initial guess

at the orbitals, the Fock operator is constructed, and the above eigenvalue equation is solved, yielding an improved set of orbitals, that will be used to construct the operator in the next step, and this goes on, until the difference between the results of two consecutive steps goes below a predefined threshold value. This procedure is usually referred to as the *self consistent field* (SCF) algorithm.

We expect that the results obtained from (12) will be better as M grows, the total electronic energy belonging to $M \to \infty$ is called the *Hartree-Fock limit.* Since it is difficult to reach this limit by direct calculations, many times an extrapolation is carried out, with for example the above mentioned bases. A popular formula for extrapolation is Feller's exponential formula:[32]

$$E_{HF}^{X} = E_{HF}^{\infty} + ae^{-bX} \tag{13}$$

where E_{HF}^{∞} is the Hartree-Fock limit, and E_{HF}^{X} is the HF energy belonging to the basis with X; a and b are parameters. Sometimes for small systems, numerical solutions of differential equations (7) may be used for benchmarking.[33]

In the derivation of Hartree-Fock theory many approximations have been made for which not even getting to the HF limit would account. In nonrelativistic quantum chemistry, electronic *correlation energy* (E_{corr}) is defined as the difference between the exact solution of (1) and the HF result:

$$E_{corr} = E_{exact} - E_{HF} \tag{14}$$

It should be noted that E_{corr} also depends on the basis set size, which is again, taken into consideration by extrapolation, typically with the following formula:[34]

$$E_{corr}^{X} = E_{corr}^{\infty} + \frac{\alpha}{X^3} \tag{15}$$

where X again refers to the Dunning basis sets, and α is a parameter.

The problem of giving a value for E_{exact}, however, still remains. Fortunately, based on HF but going beyond it, there is a number of approximations available for that purpose. These usually go back to the idea of *configuration interaction* method (CI), which looks for the exact wavefunction in the form of a linear combination of Slater determinants (configurations), which are excited from a common reference (Φ_0):

$$\Psi = \Phi_0 + \sum_{ia} t_i^a \Phi_i^a + \sum_{ijab} t_{ij}^{ab} \Phi_{ij}^{ab} + \sum_{ijkabc} t_{ijk}^{abc} \Phi_{ijk}^{abc} + \dots \tag{16}$$

where t_i^a, t_{ij}^{ab}, t_{ijk}^{abc} etc. are amplitudes (coefficients) belonging to singly, doubly, triply etc. excited configurations (Φ_i^a, Φ_{ij}^{ab}, Φ_{ijk}^{abc}). If all configurations for a given basis are considered, this method is referred to as *full CI* (FCI). This method gives highly accurate results, however, it is mainly used for benchmarking due to its prohibitive computational expense. More feasible approximations will be discussed in the next sections, but before that, some general features of electron correlation will be discussed.

Hartree-Fock is an independent particle theory, which assumes that an electron is only affected by the mean field of the others. However, it is to be expected that as the interelectronic distance decreases, the repulsion between two electrons would be larger than what the mean field estimates. Moreover, it can be shown that there is no depletion in the possibility density for two electrons when they are at or near the same place, which is unphysical. For electrons with the same spin this is excluded by the Pauli principle, which has been considered in the construction of Hartree-Fock theory. For all this, it can be said that the motion of electrons in HF is uncorrelated, giving rise to the difference with the exact solution, i.e. correlation energy. Since there is a singularity in the Hamiltonian operator at zero interelectronic distances, the wavefunction, or, rather, its derivative has to satisfy certain conditions at these places (cusp condition[35]). This requires the wavefunction to exhibit a specific shape (a Coulomb hole) at zero interelectronic distance. Related correlation effects are termed *short range correlation*, and the arising correlation energy is the *dynamic correlation energy*. The HF wavefunction does not have the desired shape. On the other hand, an expansion like (16) with enough terms may be able to reasonably approximate the cusp, although convergence may be slow. Since the Pauli principle keeps electrons with the same spin apart even in Hartree-Fock, it is also true

that the most significant correlation contributions come from electrons with opposing spins.

Another problem with HF is its failure to describe bond dissociation. In RHF, the electrons occupying the bonding orbital with different spin are forced to have the same spatial orbital. At a long interatomic distance, this gives rise to ionic terms in the wavefunction, where both electrons remain on one of the two dissociated orbitals; and there are terms where the two electrons sit on the two different dissociated species. The presence of the ionic terms increases the energy above the exact solution at long bond distances. If one uses UHF, where the restriction on spatial orbitals is dropped, the variation procedure will favour the two electrons at two different centres at long distances. However, as a consequence, the UHF wavefunction is not a spin eigenstate, and is a mixture of various spin states, introducing spin contamination. For this reason, it should be avoided whenever possible. The extent of spin contamination is measured by the difference between the expectation value of the spin squared operator with the UHF wavefunction and with the spin eigenstate we aim for. It should also be noted that spin contamination may be taken care of by constracting projection operators that remove the undesired components from the wavefunction. In the exact case, if an electron is localized on one centre, the other will prefer the other centre. In HF, since an electron sees only the mean field of the rest, assuming that one electron is localized, the other will be still found at both centres with the same probability, and therefore the probability of finding two electrons at one centre is not sufficiently decreased. Related correlation effects are called *long range correlation*, corresponding to *non-dynamic correlation energy*. If we consider a solution in the form of (16), we may arrive at the correct wavefunction with a linear combination of states where the ionic terms cancel. It is therefore very important to find a minimum set of configurations for these cases to ensure a good qualitative description.

Since crowding electrons together means a larger dynamic correlation, this type of correlation decreases when a chemical bond breaks. For similar reasons, if dynamic effects are important, HF usually underestimates binding. On the other hand, as it was seen before, HF overestimates binding, if non-dynamical correlation is important. This often leads to a cancellation of errors in HF, which is why often RHF agrees better with experiment than methods which only deal with non-dynamical effects properly. These two types of correlation are not always easily distinguished, but typically non-dynamic

effects are recovered by a CI expansion of a minimum set of important configurations, whereas dynamic correlation converges only slowly with the size of the CI expansion.[36] If dynamic correlation dominates the system (e.g. ground state equilibrium geometries), RHF is a qualitatively good description, and, therefore, it can be used as a reference in (16) to recover correlation energy. Such methods are termed *single reference* methods. If non-dynamic correlation is important (excited states, dissociation), and RHF fails, first a minimum CI should be calculated to deal with non-dynamic effects, and this wavefunction can then be used as a reference in (16) to recover dynamic correlation. Such procedures are called *multireference* methods. In the following sections these methods will be introduced to some extent. There are many variants, some requiring more computational resources than the other, but there are some points that can be checked with each of them:[37] (1) relation to FCI, (2) are they variational? (3) size consistency (additivity of energies of noninteracting subsystems).

2.2 Single Reference Methods

One of the simplest ways to account for correlation effects is offered by *perturbation theory.* Here, it is assumed that the exact Hamiltonian ($\hat{H}$) can be written as follows:

$$\hat{H} = \hat{H}_0 + \lambda\hat{W} \tag{17}$$

$\hat{H}_0$ is the Hartree-Fock Hamiltonian (sum of the Fock operators), $\hat{W}$ is the perturbation operator and λ is a small number called the perturbation parameter. If we substitute this into (1), and also expand E and Ψ in terms of λ, then, after multiplication terms with the same exponent in λ can be made equal, and from this, different orders of corrections may be gained. If with a reference of canonical Hartree-Fock orbitals $\hat{W}$ is chosen to be the difference of the exact Coulomb potential and the effective one electron potential of Hartree-Fock theory, then the result is called the Møller-Plesset perturbation series. Summing energetic terms up to the first order only yields the Hartree-Fock energy. The lowest order correction that actually improves the energy is that arising from the *second order Møller-Plesset perturbation theory*[38–40] (MP2), which is possibly the simplest of all wavefunction based electron correlation methods, where only double excitations

contribute to the energy expression:

$$E_{MP2} = \frac{1}{4}\sum_{ijrs} \frac{|\langle\phi_i\phi_j|\phi_r\phi_s\rangle - \langle\phi_i\phi_j|\phi_s\phi_r\rangle|^2}{\varepsilon_i + \varepsilon_j - \varepsilon_r - \varepsilon_s} \tag{18}$$

where i, j runs over occupied and r, s over virtual orbitals. MP2 is computationally the most feasible correlation method and is therefore in general used even for larger systems. It is also size consistent. However, it is not variational, and so it is not an upper bound to the exact energy (it may go below). It diverges if the HOMO-LUMO gap is not large enough (since the denominator of its energy expression includes differences between occupied and unoccupied orbital energies), and there may be other cases when its use is limited. Higher order perturbation theories (MP3, MP4,...) are also in general use.

If (16) is truncated after the double excitations, we get the *singles and doubles CI* (CISD) method.[41–43] The amplitudes must be determined by variationally minimizing the expectation value of the CISD energy, therefore the resulting energy is an upper bound of the exact solution. However, it is not size consistent, since the product wavefunction of subsystems should contain quadruple excitations but it is truncated after doubles. For this reason, this method is not recommended as it is, however, it is possible to define simple corrections which account for this error. These include the simple additive correction of Davidson,[44] to modifications of the iterative procedure like coupled pair functional (CPF) methods,[45] or simply including missing terms in quadratic CI (QCI).[46]

The *coupled cluster* (CC) method[47,48] is based on the following form of the correlated wavefunction:

$$\Psi = e^{\hat{T}}\Phi_0 \qquad \hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 \ldots \tag{19}$$

$\hat{T}$ is the excitation operator consisting of single ($\hat{T}_1$), double ($\hat{T}_2$) and higher excitations ($\hat{T}_3, \ldots$). If all excitations are included, (19) is equivalent to (16). However, for practical applications, $\hat{T}$ is truncated after double excitations, yielding the *singles and doubles coupled cluster* (CCSD) model. However, in contrast with CISD, CCSD is size consistent, as due to the exponential treatment, higher excitations appear in the form of products of

$\hat{T}_1$ and $\hat{T}_2$. Since higher excitations are involved, a variational evaluation is too expensive in most cases. Amplitudes are obtained from a nonlinear set of equations which one gets by projecting the Schrödinger equation to the reference, singly and doubly excited functions:

$$\langle\Phi_0|\hat{H} - E_{CCSD}|\Psi\rangle = 0 \qquad \langle\Phi_i^a|\hat{H} - E_{CCSD}|\Psi\rangle = 0 \qquad \langle\Phi_{ij}^{ab}|\hat{H} - E_{CCSD}|\Psi\rangle = 0 \quad (20)$$

Via perturbation theory, the effect of triples can also be included once one has the CCSD amplitudes. This results in the CCSD with perturbative triples (CCSD(T)[49]) method, which is considered to be the most accurate standard method for general use in computational chemistry. For open shell cases, the number of equations is three time as much as for closed shells. Additionally, without spin restriction in the CC equations, the unrestricted CCSD (UCCSD) model is not spin adapted, even if an RHF reference is used. However, a full spin adaptation is very complicated. It is simpler to devise a spin adaptation for CISD, and then the linear terms in CCSD can be spin adapted similarly. This is the partially spin adapted CCSD method, which we will refer to here simply as restricted CCSD (RCCSD[50]). It should also be noted that the difference between UCC and RCC methods decrease as more excitations are involved in $\hat{T}$ as they converge to the same FCI value.

2.3 Multireference Methods

In multireference methods, we rely on a reference wavefunction of many determinants, rather than a single one. It should be noted here, that a single determinant is not necessarily an eigenfunction of the spin squared operator (as opposed to the operators belonging to components of the spin angular momentum), however, this property can be ensured by certain linear combinations of determinants which are called Configuration State Functions (CSF). Two important special cases when determinants are eigenfunctions of the total spin are the closed shell singlet states and the high spin states. This often necessitates multideterminantal treatment as the simplest qualitatively correct approach excluding single reference methods in certain cases. If spin symmetry is needed, or

one simply wants to exploit the fact that the number of CSFs is smaller than the number of determinants (e.g. for an active space with 2 electrons in two orbitals, the number of determinants is 6, the number of CSFs is 3), then CSFs must be constructed.[51] However, since these functions are more complicated, there is an application for both approaches.

In *multiconfiguration self consistent field* (MCSCF) theory,[52–54] one seeks to deal with the non-dynamic electron correlation by constructing the minimal CI expression mentioned before:

$$\Psi = \sum_{I}^{K} c_I \Phi_I \tag{21}$$

In this method, the common set of orbitals used to construct Φ_I is optimized together with the CI coefficients (c_I). Orbitals that are occupied in any Φ_I are called internal, while the rest of the orbitals are called external orbitals corresponding to the occupied/unoccupied distinction in HF theory. Internal and external orbitals may mix during the optimization. A good choice of Φ_I is necessary for good performance. It is generally advisable to choose configurations where the valence orbitals of the molecule are occupied. A more general way of generating Φ_I is to select a subset of orbitals in the valence space, and perform an FCI in this subset yielding all possible corresponding Φ_I. This method is called the *Complete Active Space SCF* (CASSCF) method.[55,56] This expansion may contain more configurations than minimally necessary, however, due to special methods available for CASSCF, it may be faster than a general MCSCF. If the CAS spans the full valence space, then the wavefunction is granted to behave correctly at dissociation. Moreover, in this special case, the method is also size consistent. In most of the cases, however, only a smaller subspace can be chosen due to the very steep (factorial) scaling of the method with the number of active orbitals.

Multireference perturbation theory[57,58] relies on a similar expansion than what was seen in the single reference case. However, there is a number of differences. The reference function is now chosen as an MCSCF type wavefunction. The choice of zeroth order Hamiltonian is no longer as trivial as in the single reference case (i.e. the sum of one electron Fock operators), but it can be made.[59] Furthermore, the reference wavefunction is no longer in general an eigenfunction of $\hat{H}_0$, therefore the first order equations have to

be solved iteratively instead of an analytic solution as in the single reference case. We have seen while discussing MP2 theory that only double excitations from the reference wave function contribute to the energy expression. However, since the reference here is a more general one, single excitations will also appear in the energy expression of the second order. Second and third order perturbation theories (CASPT2 and CASPT3) are available for general use, as long as the underlying active space is feasible.

Multireference CI (MRCI) theory[60–62] for practical applications is terminated at the double excitations (MRCISD):

$$\Psi = \sum_{I} c_I \Phi_I + \sum_{Sa} c_a^S \Phi_S^a + \sum_{Pab} c_{ab}^P \Phi_P^{ab} \tag{22}$$

where Φ_I are all the configurations taken from the underlying MCSCF, i.e. the ones constituting the reference function, Φ_S^a and Φ_P^{ab} are all the configurations that one can possibly get from single and double excitations on Φ_I. The main bottleneck of the applicability of this method is the very rapid increase of the number of Φ_P^{ab}. If one rather than following (22) constructs the wavefunction taking the reference as a single entity, one gets the *internally contracted MRCI* (ICMRCI)[63] method. The obvious advantage is that in this case the number of excitations does not depend on the number of Φ_I. However, the excited configurations then become rather complicated, being combinations of CSFs (determinants) themselves making it harder to work with them. In practical applications, only the numerous Φ_P^{ab} configurations are contracted. The ICMRCI is a reasonable approximation to MRCI, with negligible errors, and significant improvement in applicability. The Davidson correction[44,64] and CPF methods[45,65] have their multireference equivalents to deal with the problem of size consistency.

As with single reference methods, one would perhaps anticipate the introduction of *multireference CC* (MRCC) methods.[66,67] A proper theoretical approach to these is very difficult to reach, and they have been considered very problematic until recently. However, there are some promising new implementations of this method, which may be able to change this. Unfortunately, with a singles and doubles truncation (MRCCSD), these methods are still very demanding to compute. Presently, they may be useful for very small systems.

2.4 Density Functional Theory and Composite Models

So far, pure ab initio methods have been discussed; methods that rely on a few basic physical constants and principles, and a hierarchy of approximations. There is, however, another hierarchy of methods that arises from parametrization. Even ab initio methods require numerically optimized parameters, it is enough to think of basis set parametrization, or perhaps the much simpler problem of basis set extrapolation. However, these are not based on empiria in the sense that the parameters are not introduced to improve agreement with experiment. In the early days of quantum chemistry, many efforts were made to make theoretical calculations more feasible, to enable them to deal with chemically interesting, larger species. Since the bottleneck of theoretical calculations is computing the rapidly increasing two electron integrals, it is not surprising that many of these attempts aimed at avoiding calculating these. In *semiempirical methods*,[68] a hierarchy was built up based on the number of integrals that were not computed, but were taken from fitted parameters that were optimized to properly reproduce certain experimental data (or sometimes ab initio calculations). Models like AM1 or PM3 at the cost of some fitted parameters perhaps lose something of their general applicability, as it might turn out that they have difficulties with systems excluded from parametrization (e.g. systems with d type orbitals in these cases), but in exchange they can be used for much larger systems than ab initio methods. Perhaps the extreme of parametrization is *molecular mechanics*,[69] when the quantum treatment is replaced with the ball and spring model of classical physics requiring parameters for every single type of bond stretch, bending of angles, etc. Their applicability is very much reduced to certain classes of molecules for which they were parametrized, but at such a price they are the only methods suitable for treating macromolecular systems, like those of proteins. In recent times, other developments have been made with the introduction of approximate models of *density functional theory* (DFT), which need only a few empirical parameters, if any, and are relatively accurate, and can be applied to much larger systems than high accuracy ab initio models.

The central notion of DFT is the *electron probability density* ($\rho(\vec{r})$) that gives the probability of finding an electron at a spatial point defined by $\vec{r}$. Integrating the density over the full space yields the total number of electrons. Instead of using a multivariable

wavefunction, DFT tries to calculate molecular properties from the electron density, that depends only on three spatial coordinates. The theoretical basis of DFT is the two theorems[70] proven by Hohenberg and Kohn. The first Hohenberg-Kohn theorem states that the external potential ($\hat{v}(\vec{r})$) is determined within an additive constant by the electron density. The external potential (external to the system of electrons) can be the one electron nuclear-electron interaction potential ($\hat{v}_{ne}(\vec{r})$), describing the attraction between one electron and all the nuclei, but it is not restricted to that. Once the external potential is determined, the Hamiltonian, and hence the wavefunction and the system's energy are also determined. This allows for determining physical properties using the density instead of the wavefunction. The second theorem (combined with some later developments proposed by Levy) states that the trial density that minimises the total energy is the exact ground state density. This is the variational principle of DFT. Here, we have considered the energy as a functional of the density.

What remains to be seen is how to find the density without first constructing the wavefunction, and how to find the energy once we have a density, i.e. what is the form of the energy functional. To answer these questions, Kohn and Sham devised a procedure[71] that will be introduced here. First, they considered a reference system of noninteracting electrons, for which a determinant is an exact wavefunction, and the electronic Schrödinger equation separates into equations of one electron quantities while still remaining exact. The Hamiltonian would then contain the kinetic energy and the nuclear-electronic attraction potential operators (the latter as external potential). The idea of the Kohn-Sham procedure is that one should look for a system of interest, where the kinetic energy, easily calculated in the reference, could be exact. For this reason, the total electronic energy of a molecular system can be written as a functional of density as follows:

$$E[\rho] = T_s[\rho] + V_{ne}[\rho] + J[\rho] + E_{xc}[\rho] \qquad \rho(\vec{r}) = \sum_i |\vartheta_i|^2 \tag{23}$$

where $T_s[\rho]$ is the kinetic energy functional of the noninteracting reference system, $V_{ne}[\rho]$ is the nuclear-electronic attraction potential functional, and $J[\rho]$ is that of the Coulomb electronic repulsion. These are easily calculated. The last term, $E_{xc}[\rho]$ is the *exchange-correlation functional*, which incorporates the difference between the kinetic energy term

of the molecular and noninteracting system (expected to be small), and that between the Coulombic, and the exact electron repulsion term. The contributions from the last three terms in (23) are considered to arise from the effective external potential for the molecular system in the Kohn-Sham procedure. Using the fact that the electron density can be given as a sum of orbital densities from ϑ_i spatial one electron functions, and that these are also introduced exactly in the reference system, the energy can be minimized with respect to these with the constraint of orthonormality:

$$\mathscr{E}[\{\vartheta_i\}] = E[\{\vartheta_i\}] + \sum_{i,j=1}^{n} \varepsilon_{ij}(\delta_{ij} - \langle\vartheta_i|\vartheta_j\rangle) \qquad \delta\mathscr{E}[\{\vartheta_i\}] = 0 \tag{24}$$

which is very similar to eq. (3) of Hartree-Fock theory. We may proceed analogously,[72,73] gaining the DFT equivalent of the Hartree-Fock equations, the *Kohn-Sham equations*:[71]

$$(\hat{t}_s + \hat{v}_{eff})\vartheta_i = \varepsilon_i\vartheta_i \tag{25}$$

where $\hat{t}_s$ is the one electron kinetic energy operator of the reference system, and $\hat{v}_{eff}$ is the effective external one electron potential mentioned before. Once the functions ϑ_i are available, ρ can be constructed, and the energy of the system can be obtained. The Kohn-Sham equations correspond to an independent particle system, but unlike HF theory, there is no direct physical meaning of ϑ_i and ε_i. Nevertheless, (25) are exact in principle since correlation is involved through the exchange-correlation potential. However, the main problem of DFT is that $E_{xc}[\rho]$ is unknown, hence practical DFT calculations are only approximations. There are numerous ways of estimating the exchange-correlation term, and it would be much more convenient to deal with correlation energy via that functional than constructing tedious wavefunction methods. This is more obvious if we consider that due to this formulation, correlation effects would be considered in the iterative SCF solution of (25), without having to set up separate calculations for that purpose. But since there is no systematic way of improving approximations in DFT, as there are in wavefuction theory, one can only rely on various model functions, some of which containing empirical parameters to calculate E_{xc}, and cannot be sure to what extent correlation is actually dealt with in DFT. Similarly to ab initio methods, DFT also utilizes

basis set expansion of the Kohn-Sham equations. As a consequence of approximate exchange-correlation functionals, KS methods are no longer variational, but they are size consistent.

In practical methods, the exchange-correlation functional is usually separated into $E_{xc} = E_x + E_c$, into terms for treating exchange and correlation effects, that are then modelled separately, using numerical calculations based on ab initio results, or empirical fitting to yield better results compared to experiment, at which point DFT becomes similar to semiempirical methods. For E_x, often the exchange term of the HF model is used for modelling. In the local density approximation (LDA), it is required that E_{xc} satisfies conditions valid for the uniform electron gas, an analytic description of which was provided by Dirac. The local spin density approximation (LSDA) allows the densities for different spins to be separately optimised (which would not be needed if an exact theory was available), creating the equivalent of UHF theory in DFT. Local theories assume the slow variation of the density. As this is very often not the case, in the generalized gradient approximation (GGA), an attempt is made to adapt DFT to greater variation of the density by involving an explicit dependence on the graidents of ρ in E_{xc}. Finally, hybrid functionals mix the exchange term of HF theory with exchange and correlation terms of GGA functionals, where the mixing coefficients are also fitted. All these approximations apply numerous functionals providing a large scale of DFT methods. The two models, B3LYP[74] and BH&HLYP[75] that will later be used in this thesis are hybrid functional methods.

Finally, a few words about *composite model chemistries* should be said here. These are referred to as composite, as they use several steps of low computational cost to approximate one expensive method. A typical setup is

$$E_{HL,LB} = E_{LL,LB} - E_{LL,SB} + E_{HL,SB} \tag{26}$$

where a high level (subscript HL) method with a large basis (subscript LB) is approximated by performing the high level calculation at a small basis (SB) and then adding the difference of a low level (LL) calculation with large and small bases. That is to say the basis set effect is considered additive in this case. Other effects often consid-

ered additive include: relativistic, Born-Oppenheimer and spin-orbital effects. Basis set extrapolation is also often used, and so is fitting of some correction function against experiment to yield better calculated properties. DFT methods and scaled frequencies are utilized widely within composite models. Various families of these exist, including the Gaussian,[76] Weizmann[77] and CBS[78] methods for thermochemical calculations of larger species, or the HEAT[79] method developed for accurate calculations on small molecules. A variant, G3MP2B3[80] mentioned in later sections belongs to the family of Gaussian methods, and aims at the approximation of QCISD(T) level with large basis via eq. (26) using MP2 low level calculations, and B3LYP optimized geometries and frequencies for thermochemistry (see the following section). A modification of this method, G3MP2BH&H is also proposed and used in some of our calculations.

2.5 Thermodynamic and Kinetic Properties

So far, the focus has been on various ways of calculating the electronic energy of a molecular system. However, in practice, calculating derivatives of the energy with respect to nuclear coordinates is equally important. Gradients are used for minimizing the energy as a function of nuclear positions, i.e. in geometry optimizations. Second derivatives are used to ensure that stationary points obtained by optimizing the geometry, are indeed minima, or first order saddle points (transition states), etc. By assuming a harmonic behaviour around the vicinity of the minimum energy point, one can interpret second derivatives as force constants belonging to the harmonic vibrations of the molecular system. Although the assumption of harmonicity is in many cases an oversimplification, since it is much harder to include anharmonicity into calculations, the harmonic oscillator model is the most common way of calculating molecular vibrations. As mentioned in previous sections, the nuclear Schrödinger equation is rarely solved, so the vibrational energy calculated this way, is basically an approximate way to consider nuclear motion. Since vibrational motion also has a contribution at 0 K, when other types of motion of the molecular system do not, the corresponding correction to the energy is termed *zero point vibrational energy* (ZPVE or ZPE). Calculating vibrations is also essential if one wants to predict various thermodynamic and kinetic properties. For this, we need to consider two things. Firstly, that the measured quantities correspond to a manifold of molecules,

rather than an isolated one, for which we have calculated results; and secondly, that these quantities need to be calculated at various temperatures, and therefore we are looking for *thermal corrections* for our calculated total energy.

Statistical mechanics makes the required connection between molecular calculations and macroscopic quantities, and enables us to calculate the various thermal corrections to get values for the latter for a temperature T. The statistical properties of a system in thermodynamic equilibrium are described by the *partition function* (Q). In an ideal gas of molecules, the partition function, containing translational, rotational, vibrational and electronic contributions, is the following:

$$Q = Q_t Q_r Q_v Q_e = \left[\left(\frac{2\pi mkT}{h^2}\right)^{\frac{3}{2}} V\right]\left[\frac{1}{\sigma}\sqrt{\left(\frac{kT}{h}\right)^3 \frac{\pi}{ABC}}\right]\left[\prod_i \frac{1}{1-e^{-\frac{h\nu_i}{kT}}}\right]\left[\sum_i g_i e^{-\frac{\varepsilon_i}{kT}}\right] \tag{27}$$

where the rotational part needs to be replaced with $kT/\sigma hB$ if the molecule is linear. Here, k and h are the Boltzmann and Planck constants, respectively. The other quantities include: the molar volume of the ideal gas (V), the molecular mass (m), the external symmetry number (σ), rotational constants (A,B,C), harmonic vibrational frequencies (ν_i). In the last term, g_i is the degeneracy of the ith electronic state, and if more than one states are considered, then ε_i is the relative energy of the ith one to the lowest lying state. Various functions of thermochemical interest can now be calculated using (27), like the entropy (S), the internal energy (U), the enthalpy (H), the Helmholtz free energy (A), the Gibbs free energy (G) and so on. S and H can be derived from Q as follows:

$$S = RT\left(\frac{\partial lnQ}{\partial T}\right)_V + RlnQ \qquad H = RT^2\left(\frac{\partial lnQ}{\partial T}\right)_V + RTV\left(\frac{\partial lnQ}{\partial V}\right)_T \tag{28}$$

From these, G can be determined, and there are similar expressions for U and A. Subscripts refer to variables that are constant with respect to derivation. These, or derivative quantities can now be compared to experimental values.

Transition state theory (TST) assumes that reacting chemical species go through a critical configuration, the transition state, for the reaction to occur. In practice, these are

determined by opitimizing first order saddle points of the potential energy surface (while with reactants minimum structures are located). This theory assumes that there is a chemical equilibrium between reactants and the transition state species and that the rate of a reaction depends on the concentration of the latter and the frequency of its decomposition to products. With the assumption that the decomposition frequency is identical with the imaginary frequency of the saddle point structure (a frequency that describes a vibration of the transition state that is becoming the translational motion of the separating reactants) and that the concentration of the transition state structure follows the Boltzmann distribution with respect to its relative energetics, the following expression is obtained for the rate constant:

$$k_{TST}(T) = \kappa \frac{kT}{h} K^{\circ} e^{-\frac{\Delta G^{\ddagger}}{RT}} \tag{29}$$

Here $\Delta G^{\ddagger}$ is the Gibbs free energy difference of reactants and transition states, and is a quasi-thermodynamic quantity, since the imaginary frequency is ignored in the transition state free energy. K° is there merely as a unit factor, it is unity for unimolecular, and the reciprocal of the standard concentration unit for bimolecular reactions. The tunnelling coefficient κ is usually ignored in initial estimations. A simple way for estimating κ is via Wigner's correction depending only on the imaginary frequency, and there are many other, more complicated ways of calculating it. Although (29) is the simplest approximation among rate theories, and there are more sophisticated models available, nevertheless, in a good number of cases, it yields a good estimation for the rate constant.

3 Hydrogen Transfer Reactions. A General Approach

3.1 General Theoretical Issues

As a result of their importance in a variety of fields, hydrogen transfer reactions are well studied both experimentally and theoretically, making comparisons possible. These reactions are favoured model systems for the study of chemical reactivity. In their simplest forms they serve as models for reactive scattering and dynamics,[81] or models for heavy + light-heavy atom reaction systems, e.g. symmetric H exchange between halogen atoms and hydrogen halides.[82,83] They have served as useful model systems for testing standard theoretical approaches,[84,85] and many attempts have been made to describe similar systems with simple yet reliable approximations.[86,87]

This type of reactions represents a good number of challenges.[88] As a result of an open shell calculation in these cases, one may easily end up with a manifold of adiabatic states. In a case like X + H–X reactions,[83,88] where X is a halogen, the problem arises as X approaches H–X. If a collinear geometry is considered, Σ and Π states may be distinguished based on the spatial symmetry of the whole system. In this case, the H-abstraction is energetically favoured in the Σ state, since the SOMO is oriented towards the H on H–X, whereas at larger distances the Π state is favoured due to more favourable multipole interactions. Hence there is a crossing of these two surfaces at some point. In the corresponding planar case (C_s symmetry) one can distinguish between states having the SOMO in the molecular plane (two orbitals of A' symmetry) or out of it (one orbital of A'' symmetry), and an avoided crossing occurs. The presence of multiple coupled potential energy surfaces makes an ab initio investigation very challenging. It should also be noted that calculating BSSE corrections is also not trivial[89] especially in a multireference case where the monomers have different active spaces compared to the dimer, although there

exist some attempts to carry this out.[89,90] The easiest solution is not to perform a BSSE correction, in which case large basis sets must be used.

For the case of Cl + HCl, Dobbyn et.al.[83] draws attention to the fact that RHF reference functions should be treated with caution, since symmetry breaking appears. Instead, they use state averaged MCSCF orbitals even for their single reference methodology, which may be justified as the wavefunction is mainly dominated by a single configuration. Fox and Schlegel[84] discusses various anomalies in optimizing the FHF transition state. The primary target of their paper is to show that artefacts in MP2 optimizations are due to the lack of consistent treatment of single excitations in MP2 theory. Earlier theoretical studies found the FHF transition state linear at the HF level,[91] whereas MP2 predicts a bent structure with C_s symmetry, and a minimum at the C_{2v} bent structure. Whereas the collinear structure can be explained by the lack of correlation treatment, and the related underestimation of interactions between lone pairs, the C_s symmetry transition states are explained as "residuals of avoided crossings between the bond making/breaking state and two broken symmetry, hydrogen bonding states." The latter are formally single excitations, and it is found that a proper inclusion of singles is able to treat this avoided crossing correctly. This is in line with the recommendation of CCSD(T) methodology whenever possible,[92,93] and that spin contamination has a negligible effect on the geometries of radical transition states compared to triples contributions,[92,93] although it should be noted that the use of RHF orbitals is highly recommended[94] even when the CC treatment is not spin restricted (UCC). In the work, of Luth and Scheiner,[85] the use of localized orbitals is recommended for constructing active spaces for hydrogen abstraction cases. In their work they rely on symmetry constraints besides localization to choose active orbitals. For the virtuals, the choice of which is less unambiguous, they recommend trying to find orbitals which are likely to be linked only with the H transfer, and not with other weak interactions.

3.2 Localization

As we saw during the derivation of Hartree-Fock theory, these equations are invariant with respect to unitary transformations of the orbitals. Although it is most convenient

to work with the canonical form during the SCF procedure, there may be cases, when this degree of freedom can be exploited. The canonical orbitals are delocalized over many centres, whereas the traditional chemical way of thinking works with orbitals centred on one or two, or not too much more atoms, i.e. with various bonds between two atoms, or lone pairs, etc. Using unitary transformations, it is possible to *localize* these orbitals, that is to confine them to a relatively small space, and also to reflect similarities on the orbital level between chemically similar groups in different species. One may choose a two electron operator ($\hat{\mathcal{O}}$), the expectation value of which can then be optimized with respect to the transformation. However, there is no unique choice of this operator, and consequently, many localization procedures exist. Some of the possible choices are:

$$\hat{\mathcal{O}}_{\text{Boys}} = r_{12}^2 \qquad \hat{\mathcal{O}}_{\text{ER}} = \frac{1}{r_{12}} \qquad \langle \hat{\mathcal{O}}_{\text{PM}} \rangle = \sum_A |\rho_A|^2 \tag{30}$$

The *Boys* procedure[95] minimizes the expectation value of the square distance of two electrons, or equivalently maximizes the distance between orbital centroids, making the orbitals as compact as possible. The *Edmiston-Ruedenberg* localization[96] maximizes the expectation value of the inverse distance between electrons, i.e. the Coulomb repulsion. The *Pipek-Mezey* localization[97] maximizes the sum of square Mulliken charges (ρ_A) which corresponds to the number of electrons associated with an atomic centre (A). These methods yield mostly similar results. The Pipek-Mezey procedure should be used if the σ/π symmetry is to be preserved during the transformation.[97,98]

Local correlation methods[99–101] use localized molecular orbitals from which only excitations to virtual orbitals close to the occupied orbitals are considered. This means that each electron/electron pair (single, double excitations) has a different (local) domain of virtual orbitals smaller than the complete space, resulting in the decrease of computational efforts. These methods are based on the assumption that the large scaling of correlation methods is unphysical in a sense that they consider excitations that are unimportant because of the spatial distance of the orbitals from which and to which the electrons are excited. The localization of occupied orbitals is usually without any problem, the Pipek-Mezey method being preferred due to the already discussed advantages. However, localizing the virtual orbitals is more problematic, and instead, the idea of Pu-

lay was to use a set of projected atomic orbitals (PAOs). These orbitals are obtained from atomic orbitals by projection to the occupied space, and thus, they are orthogonal to the occupied orbitals, but not to each other. This means that under the orthogonality constraint, PAOs are the most similar transformed orbitals to the atomic orbitals, and they share the desired property of the latter of being localized. This enables us to determine the domains of virtuals necessary for these methods. Despite some problems like the fact that with significant changes in the geometry, the relevant domains may change, the big advantage of these methods is its nearly linear scaling allowing correlation methods to be used with larger systems. Methods like MP2 or CCSD(T) have their local equivalent (LMP2, LCCSD(T)). A further advantage would be the automatic BSSE correction because of the use of local domains. Unfortunately, these methods are only available for closed shell systems at present, therefore they cannot be used with systems being considered here.

3.3 Advantages of Localization. Transferability

A defining theme in chemistry is localization and transferability: the concept, for example, that the chemistry of functional groups is only weakly affected by the neighbouring molecular structure, and is therefore largely transferable between different molecules. This theme is reflected in some strategies for first-principles computation of molecular structure, suitable for use when the particular chemical transformation or property is known to be spatially localized. For example, QM/MM[102] and other spatial partitioning schemes allow the cost of an electronic structure computation to be nearly independent of the size of the outer spatial region, depending only on the size and complexity of the active site. Similarly, the complete active space multiconfiguration self-consistent field (CASSCF)[55,56] method is formally factorial scaling in the number of electrons, but this exponential growth can be avoided in some circumstances by choosing an active orbital space that is local to one region of the molecule, on the assumption that the principal electron correlation effects on quantities such as reaction barrier heights arise predominantly from these orbitals. In this part of the thesis, we explore the effect of such orbital-based partitioning, explicitly analysing the contributions to barrier heights from the correlation effects within and between different orbital spaces. The ultimate purpose of this analysis

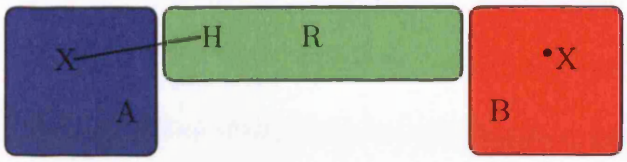

Figure 2: *Scheme for partitioning correlation energy*

is twofold. It can lead to practical algorithms for electronic structure computations with strongly reduced effort, in particular for reactions between polyatomic molecules with large residues that do not directly affect the electronics of the reaction. It can also lead to simple models for understanding the characteristics of potential energy surfaces in terms of transferable quantities for molecular fragments.

An example of a method that embodies orbital, rather than spatial, partitioning is the CIPT2 ansatz.[103] An active orbital space, obtained from a prior CASSCF calculation, defines a multireference configuration interaction (MRCI) ansatz,[60] which is then modified such that excitations involving one or two orbitals outside this active space are treated more simply using second-order perturbation theory, rather than variationally. Here, we follow a similar approach, but focusing in detail on a simple prototype class of bond-breaking chemical reaction, the transfer of hydrogen from a closed-shell hydride to a radical. We attempt to partition the correlation energy contribution to the reaction barrier height into a part that arises from the active space only, parts that correlate active orbitals with spectators on other parts of the reactant molecules, and pure spectator contributions, both intramolecular and intermolecular.

To make our envisaged partitioning well defined, we use the fact that localized orbitals are confined to a smaller part of the molecule. Such a partitioning would have numerous advantages. One of them is that since the localized orbital structure of the same functional group in different environments is expected to be nearly the same in different molecules,[104] the correlation contribution arising from this group is expected to be nearly identical, therefore, once calculated in one species, it would be *transferable* to another that contains the same functional group. This gives us a powerful tool to check our results.

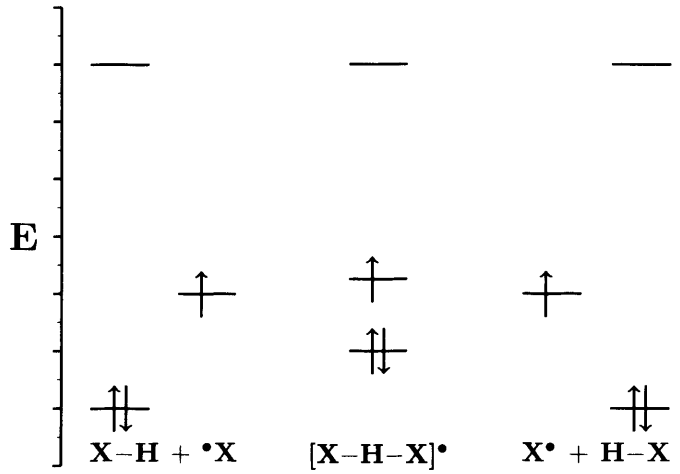

Figure 3: *Reacting orbitals (R space) at different reaction coordinates. The orbitals shown are the breaking/forming bond orbital, labelled R, and the SOMO labelled r in subsequent chapters*

3.4 Partitioning of Correlation Energy

We have chosen hydrogen transfer reactions for our calculations. Here, we have a H-donor and a H-acceptor (radical); X – H and X, respectively in the symmetric case shown in Figure 2, which we consider as spatial orbital partition A and B, respectively. The majority of the orbitals could be identified as belonging to either of these two, assuming they remain confined to the same spatial partition during the reaction. It is, therefore, possible to introduce a partitioning in the orbital space corresponding to this spatial partitioning. However, it is also advisable to introduce another partition, for orbitals which do change during a reaction, those that would normally be chosen as (minimal) active space orbitals in an MCSCF calculation, being likely that they are related with strong long range correlation effects. This will be referred to as the reacting or R orbital partition. A schematic representation of R orbitals is found in Figure 3. Here we can interpret the changes of these orbitals as rotations with respect to the initial set of orbitals (or even to the AOs) as we go along the reaction coordinates in a symmetric H abstraction reaction. It can be seen that the orbitals involved are the SOMO and the doubly occupied orbital that is originally the X – H bond that breaks. These characteristics of the R space lead us to assume that orbitals in a CASSCF type wavefunction would ideally satisfy the needs of such partitioning by getting automatically separated into the desired partitions.

Let us now continue with examining the consequences of such an orbital partitioning to correlation energy. Note that although the CASSCF wavefunction would have certain advantages as discussed before, all calculations eventually carried out use the single reference approach. Considering the Hamiltonian in second quantized spin free formalism,[105] working with real spatial orbitals

$$\hat{H} = \sum_{pq} h_{pq}\hat{E}_{pq} + \frac{1}{2}\sum_{pqtu}(pt|qu)\hat{E}_{pt,qu} \tag{31}$$

the correlation energy can be written as

$$E_c = \langle 0|\hat{H}|\bar{\Psi}\rangle = \sum_{ir} h_{ir}\langle 0|\hat{E}_{ir}|\bar{\Psi}\rangle + \frac{1}{2}\sum_{ijrs}(ir|js)\langle 0|\hat{E}_{ir,js}|\bar{\Psi}\rangle = \sum_{i<j}\epsilon_{ij} \tag{32}$$

where

$$\epsilon_{ij} = \delta_{ij}\sum_{r} h_{ir}\langle 0|\hat{E}_{ir}|\bar{\Psi}\rangle + \sum_{rs}(ir|js)\langle 0|\hat{E}_{ir,js}|\bar{\Psi}\rangle \tag{33}$$

with $|\bar{\Psi}\rangle = |\Psi\rangle - |0\rangle$, and intermediate normalisation is assumed between the reference ($|0\rangle$) and correlated (Ψ) wavefunctions $\langle 0|\Psi\rangle = 1$ (consequently $\langle 0|\bar{\Psi}\rangle = 0$). $\hat{H}$ contains one and two electron excitation operators $\hat{E}_{ir}$ and $\hat{E}_{ir,js}$, respectively, which are defined as follows

$$\hat{E}_{ir} = \sum_{\sigma} a^{\dagger}_{r\sigma}a_{i\sigma} \qquad \hat{E}_{ir,js} = \sum_{\sigma\tau} a^{\dagger}_{r\sigma}a^{\dagger}_{s\tau}a_{j\tau}a_{i\sigma} \tag{34}$$

which excite electrons from orbitals labelled i, j to orbitals r, s. The two indexed $a^{\dagger}$ and a are creation and annihilation operators, respectively, generating various excited configurations through destroying orbitals indexed i, j and creating others indexed r, s. Spin indices σ and τ are applied to keep track of different spin variants of excitations from and to the same spatial orbitals. Since the Hamiltonian contains one and two electron operators, the correlation energy can be separated into pairwise contributions ϵ_{ij}, as

shown in (33), of orbitals i, j occupied in the reference configuration. In this convention we include the arising single terms into ϵ_{ij} as it is defined in (33). Since our partitioning is carried out on $|0\rangle$, it follows that E_c will also be split up to partial sums of ϵ_{ij}. As there are three orbital partitions, the correlation energy will be split up into six terms, denoted as follows: E(A-A), E(B-B), E(R-R), E(A-B), E(A-R) and E(B-R). The notations in brackets indicate the partitions to which the orbitals belong. Partitions E(A-A), E(B-B) and E(R-R) are referred to as self(energy) partitions, due to the fact that the orbital pairs involved are of the same partition. Besides, there are fractions describing interactions between fragments: E(A-B), E(A-R) and E(B-R), where E(A-B) can be interpreted as dispersion. The subscript ts and r would refer to the fact that these partial sums belong to the transition states and reactants, respectively. A Δ symbol refers to correlation contributions to barrier heights, as these are the differences of the aforementioned two types of quantities (e.g. ΔE(A-A)=E$_{ts}$(A-A)-E$_r$(A-A), etc.). If partial sums are present without these symbols then the relevant statement is generally valid.

It is our hope that this partitioning could make practical calculations easier. To understand why, first let us consider the E(A-A) and E(B-B) partitions. Since localized orbitals are used, their contribution should be transferable from one reaction to another. Once these are known in one reaction, they need not be calculated in another. This holds trivially for reactants. A similar argument may hold to the E(A-R) and E(B-R) bits, since we are dealing with similar types of reactions, similar changes should occur in the R space during a reaction, and so these partitions should also be transferable to at least some extent. We are left with E(A-B) and E(R-R). E(A-B) can be understood as a dispersion term (and as such it is trivially zero for the distant reactants), and it could be calculated via for instance force field calculations, in which case the only bit that should be treated quantum mechanically would be the E(R-R) bit, which arises only from orbitals responsible for the majority of static correlation. In this manner a correlation method could be approximated in a QM/MM fashion, where the layers would correspond to the orbital space. It should be noted here that, although on a different basis, there are successful uses of localized orbitals for instance in making the MRCI ansatz more feasible,[106] which reassure us about their usage in novel methodologies.

For practical calculations of barriers of reactions, we first need to choose what orbitals to use with the various possible correlation methods at our disposal. With transition

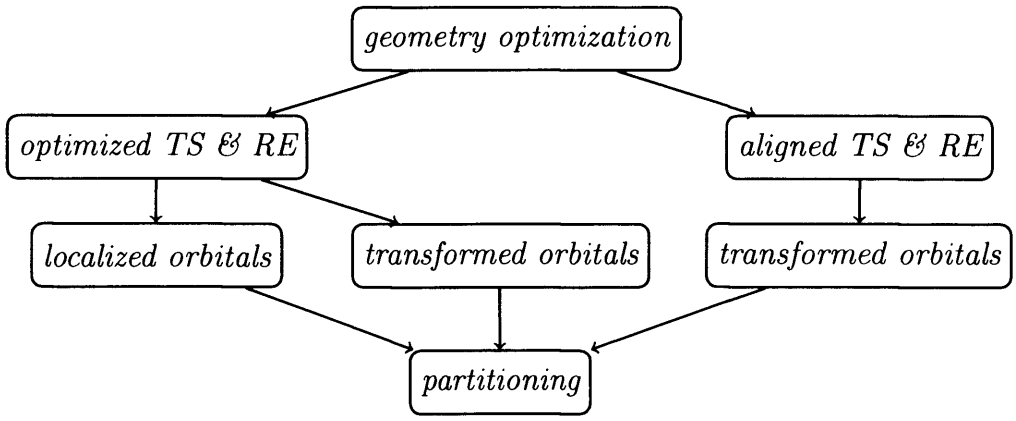

Figure 4: *A summary of the various calculations performed*

state structures, CASSCF orbitals would be an optimum choice, since these account for the static correlation that is expected to become significant if bond breaking occurs, and under these circumstances an adaptive CASSCF procedure is expected to find the right local orbitals. However, since many single reference methods were shown to work well with radical transition states,[92,93] the use of Hartree-Fock orbitals is also possible. Various localization techniques[95–97] may then be used to obtain the proper partitioning. As we will see later, this latter procedure was chosen for our calculations.

The case is similar for the reactants, where single reference methods should be sufficient, and due to the weaker static correlation, even the use of CASSCF would not ensure that one ends up with properly localized orbitals. Performing localization is the simplest way to get the orbitals needed. In our calculations, Boys localization[95] is used. However, since localized transition state orbitals are available, it is also a possibility to come up with a transformation that would make reactant orbitals similar to the corresponding transition state ones. Before going on detailing such a transformation, it should be noted that since reactant and transition state geometries differ, the rotation of p and higher orbitals may represent some problems in such a procedure. To make sure that this effect is dealt with, geometries can be forced to have the same orientation and structure except with larger distance between the segments. All in all, this gives us the choice of using localized, transformed and geometrically aligned and transformed orbitals for our purposes.

In Figure 4, the various calculations performed are summarized; all of these start similarly. After obtaining the transition state structure, the breaking bond distance can be

increased, and the reactant structures can be optimized. The majority of our results involve localized orbitals. For the cases where orbital transformation is applied, the subsequent single point calculations are carried out with the 'noorient' option in MOLPRO with both geometries. This prevents the reorientation of structures based on inertia moments, and since the structures are defined in a similar, analogous manner, it provides a large overlap of reactant and transition state geometries since partition A ends up in nearly the same position with B differing due to the larger distance of the segments in the reactants. As a third possibility, to further minimize the effect of misalignment, we can take the same segment structures only differing in the intersegmental separation, so that partition A would match perfectly in reactants and transition states, and partition B would be geometrically identical but translated relative to each other in the two cases. Thus, we have provided similar geometries, and we have fixed these relative to each other. For the latter two cases, after orbitals are localized, a Löwdin type transformation[107] is performed between corresponding orbitals of transition state and reactant geometries. This will be discussed in the following sections.

Before moving on to orbital transformations, the procedure of sorting the orbitals into partitions should be clarified. There are several possibilities. One is to decide upon the graphical representation of each orbital. Another is to consider the basis function structure of the wavefunction. For this, a threshold for contracted function coefficients in molecular orbitals is needed. This is typically chosen to be 0.15, but never less than 0.1. Next, the centres of relevant AOs are listed for each MO. This list can then be compared to the list of atoms in a partition, revealing the atomic centres to which the orbital is localized around. This procedure is done for transition states and reactants at a large 1000 Bohr separation. Comparing these results makes it possible to assign a partition for each orbital. Usually both methods were applied in our calculations, occasionally accompanied by an analysis of the orbital centroid positions.

3.5 Orbital Transformations

3.5.1 Transformation for Transition State Orbitals

In this chapter of the thesis, Boys localized orbitals are used almost exclusively, as a practical compromise substituting a more general CASSCF orbital set. However, some possibilities were considered in the initial stages of our work, and therefore a short summary of such procedures will be included here, for the sake of completeness.

If such a transformation is carried out, it is carried out before reactant orbitals are transformed, and as we will see, that latter transformation depends on transition state orbitals. Here, we assume that the localized orbitals are mixed up, in particular, as we will see later, the orbitals of partition A with the breaking-bond orbital belonging to partition R. From here on, the R subindex in quantities will refer to this doubly occupied orbital, whereas the SOMO is denoted by r. Since the SOMO is uniquely defined, the minimization of the above mentioned mixing would be achieved by making R "similar" in nature to r, and transforming the A orbitals accordingly, resulting in the separation of orbital partitions A and R. This, in our method, is reached by maximizing the Coulomb interaction between, and therefore the correlation contribution from the transformed R (ψ'_R) and the SOMO ψ_r:

$$\max_{\psi'_R} \frac{\langle \psi'_R \psi_r | r_{12}^{-1} | \psi'_R \psi_r \rangle}{\langle \psi'_R | \psi'_R \rangle} = \max_{c_{Ri}} \frac{\sum_{ij} c_{Ri} c_{Rj} \langle \psi_i \psi_r | r_{12}^{-1} | \psi_j \psi_r \rangle}{\sum_i c_{Ri}^2} \qquad \psi'_R = \sum_i c_{Ri} \psi_i \tag{35}$$

where r_{12}^{-1} denotes interelectronic distance, $\{\psi_i\}$ are orbitals of the direct sum of partition A and the breaking-bond orbital, c_{Ri} are the linear coefficients assuming that the transformed breaking-bond orbital can be expressed as a linear combination of orbitals in the subspace mentioned before. This will lead to an eigenvalue problem:

$$\mathbf{J^r c} = \mathbf{\Lambda c} \qquad J^r_{ij} = \langle \psi_i \psi_r | r_{12}^{-1} | \psi_j \psi_r \rangle \tag{36}$$

where the elements of the matrix $\mathbf{J^r}$ are the three centre two electron integrals J^r_{ij}. The r superindex here denotes dependence on the r orbital. These can be gained by the 'fcidump' function of the FCI module in MOLPRO. The basic idea of this transformation is very similar to that of the Edmiston-Ruedenberg localization procedure described before. Once (36) is solved, the eigenvector belonging to the largest eigenvalue yields the desired transformed breaking-bond orbital, and the rest can be used to set up transformations for the other orbitals involved.

3.5.2 Transformation for Reactant Orbitals

Many possible arrangements for transforming reactant orbitals were considered, although actual calculations discussed in this thesis only utilize one, and as we will see, even the use of that is limited. However, it is perhaps better to show the chain of thoughts (and numerical experiments) by which we have arrived at the one finally used, and in later chapters why transformations were abandoned completely.

Our first idea of a transformation of reactant orbitals was based on the thought that we can make the transition state and reactant orbitals as similar as possible, by maximizing the square of the overlap between the transition state (upper index TS) and the reactant's transformed (indicated by $'$) doubly occupied R space orbital (R lower index). Some other orbitals of the reactant are necessarily transformed as well, and later we will discuss what practical choice we settled with. For now, it is enough to know that any transformed orbital ψ'_i can be written with the untransformed reactant orbitals (RE upper index) as follows:

$$\psi'_i = \sum_j c_{ij} \psi_j^{RE} \tag{37}$$

The function (F) to maximize is:

$$\max F = \max_{\psi'_R} \frac{|\langle \psi_R^{TS} | \psi'_R \rangle|^2}{|\langle \psi'_R | \psi'_R \rangle|} = \max_{c_{Ri}} \frac{\left| \sum_i c_{Ri} \langle \psi_R^{TS} | \psi_i^{RE} \rangle \right|^2}{|\langle \psi'_R | \psi'_R \rangle|} \tag{38}$$

Assuming real orbitals:

$$\max F = \max_{c_{Ri}} \frac{\sum_{ij} c_{Ri}c_{Rj}\langle\psi_R^{TS}|\psi_i^{RE}\rangle\langle\psi_R^{TS}|\psi_j^{RE}\rangle}{\sum_i c_{Ri}^2}$$
$$= \max_{c_{Ri}} \frac{\sum_{ij} c_{Ri}c_{Rj}S_{Ri}S_{Rj}}{\sum_i c_{Ri}^2} = \max_{c_{Ri}} \frac{\sum_{ij} c_{Ri}c_{Rj}P_{ij}}{\sum_i c_{Ri}^2} \tag{39}$$

with the notations $S_{Ri} = \langle\psi_R^{TS}|\psi_i^{RE}\rangle$ and $P_{ij} = S_{Ri}S_{Rj}$, where S_{Ri} are elements of the asymmetric overlap matrix between TS (R in particular) and RE orbitals. For our problem to be well defined, conditions for calculating the overlap matrix $\mathbf{S}$ need to be determined more closely. Let us examine the form of $\mathbf{S}$ at the atomic orbital level:

$$S_{ij} = \langle\psi_i^{TS}|\psi_j^{RE}\rangle = \sum_{kl} p_{ik}q_{jl}\langle\eta_k^{TS}|\eta_l^{RE}\rangle = \sum_{kl} p_{ik}q_{jl}D_{kl} \tag{40}$$

or in a matrix form

$$\mathbf{S} = \mathbf{pDq}^T \tag{41}$$

where $\mathbf{D}$ is the atomic orbital equivalent of $\mathbf{S}$, with molecular orbital coefficient matrices $\mathbf{p}$ and $\mathbf{q}$ for transition states and reactants, respectively. While the latter two matrices can be obtained easily, there are some difficulties with $\mathbf{D}$. To see what these may be, we assume that the atomic orbitals have the following form:

$$\eta_k^G = N_k e^{-\alpha_k|\vec{r}-\vec{A}_{k,G}|^2} \tag{42}$$

where G refers to either TS or RE, $\vec{A}_{k,G}$ is the coordinate vector for the kth AO centroid in the geometry G. In this case, an atomic overlap matrix element between TS and RE geometries may be written as:

$$D_{kl} = \int \mathrm{d}\vec{r}\; N_k N_l e^{-\alpha_k|\vec{r}-\vec{A}_{k,TS}|^2-\alpha_l|\vec{r}-\vec{A}_{l,RE}|^2} \tag{43}$$

$$\approx \int \mathrm{d}\vec{r}\; N_k N_l e^{-\alpha_k|\vec{r}-\vec{A}_{k,TS}|^2-\alpha_l|\vec{r}-\vec{A}_{l,TS}|^2} \tag{44}$$

$$\approx \int \mathrm{d}\vec{r}\; N_k N_l e^{-\alpha_k|\vec{r}-\vec{A}_{k,RE}|^2-\alpha_l|\vec{r}-\vec{A}_{l,RE}|^2} \tag{45}$$

As it can be seen (43) differs from (44) and (45) in replacing the geometry labels for both functions to be either TS or RE instead of the mixed labels of the strict definition. As eventually (45) is used, we now have to discuss what the consequences of this replacement will be, and this should be done by specifying what the relation should be between the coordinates for TS and RE species in our calculations. By using the same label for both species, this means that e.g. in (45) all the transition state AOs are 'transferred' to the RE geometry. It can then be seen that if there were only a translational difference between the two geometries, that would disappear during the integral calculation. For s type orbitals, even rotation is irrelevant due to their spherical symmetry, however, as we have seen before, the 'noorient' option of MOLPRO was used exactly to prevent the ill consequences for p and higher order orbitals. If we consider a rotational variation, these orbitals will not remain invariant, rather, they end up being linear combinations of the original (unrotated) case. This mixing would change overlap values altering our results significantly, and should therefore be avoided. Considering the question whether (44) or (45) are better approximations to (43), our experience shows that they perform equally well, but it is technically simpler to implement (45).

It is also possible to argue that (44) or (45) are actually better choices than (43), for which reason (43) was not eventually coded but approximated as discussed above. It is useful to look into the details of the calculation of the asymmetric matrix $\mathbf{S}$ to see why. As a consequence of eqs (43), (44) and (45), the atomic overlap matrix may be estimated as:

$$\mathbf{D}^{RE} \approx \mathbf{D} \approx \mathbf{D}^{TS} \tag{46}$$

In our calculations, we particularly used the reactant atomic overlap matrix ($\mathbf{D}^{RE}$) cor-

responding to (45) rather than that of the transition state ($\mathbf{D}^{TS}$). Let us examine the structure of these various overlap matrices shown in Figure 5. It should be noted in advance that many of the properties of matrices depicted here are merely assumptions as $\mathbf{D}$ is not calculated, therefore there is only limited numerical evidence to support our conclusions. Nevertheless, numerical data gained from approximations fit into the overall picture proposed below. For simplicity, it is assumed that the molecule has two segments A and B (R is ignored, the SOMO being considered as belonging to B, and the breaking bond orbital to A). In the transition state ($\mathbf{D}^{TS}$) A and B are close, so the overlap matrix has non-zero elements throughout, related overlap minor matrices are denoted as AA, AB, etc. In the reactant case ($\mathbf{D}^{RE}$), the A'B' and B'A' overlaps are zero due to the large separation of reactant partitions A' and B'. If we now superimpose the two, in a way discussed in the previous chapter, then we find that the geometric structure of A nearly coincides with A', B is close to these, and B' is far away. The asymmetric $\mathbf{D}$ describes this case. The consequence of the geometrical arrangement is that it may well be assumed that overlaps AA, A'A and AA' are nearly the same, and so are A'B and AB, since B remains close both to A and A'. Since B' is distant from the rest of the superimposed structure, it means that overlap B'A similarly to B'A' is zero. B'B is also zero, and is unlike BB or B'B'. The conclusion is that for orbitals in partition A all approaches yield approximately the same results. However, since B'B is zero which makes the transformation impossible with (43) i.e. $\mathbf{D}$. As we mentioned AOs are transferred in all the other cases, the corresponding overlaps will not be zero. In our case, this means that the transition state orbitals are transferred to the reactant structures and the overlap is calculated thus, overcoming the difficulty of large intramolecular separation, and making it possible to perform a transformation for all orbitals. However, as we will see, the transformation actually constructed and described later only affects orbitals in segment A, for which all the approximations, as it has been shown, are valid.

Keeping in mind the consequences of the above discussion, we may now go on with our derivation. With the requirement that derivatives according to c_{Ri} vanish, (39) becomes an eigenvalue problem of the matrix $\mathbf{P}$:

$$\mathbf{cP} = \mathbf{\Lambda c} \tag{47}$$

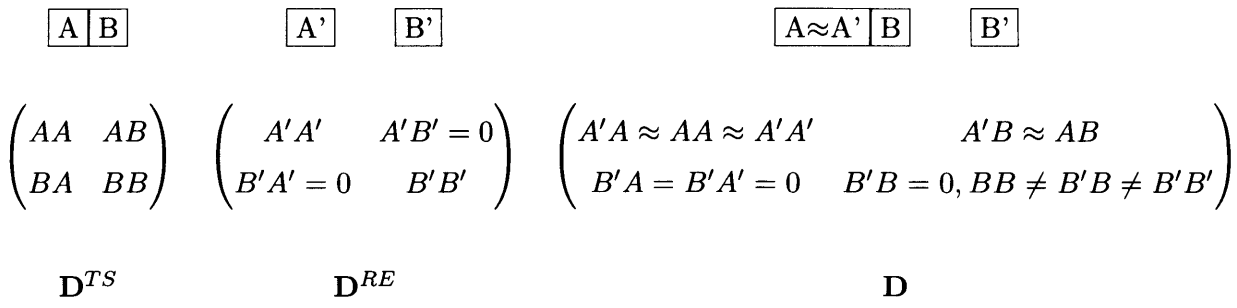

Figure 5: *Structure of atomic overlap matrices of the transition state ($\mathbf{D}^{TS}$), the reactants ($\mathbf{D}^{RE}$) and the two superimposed ($\mathbf{D}$) with schematic representation of the relative positions of the segments. A and B refer to transition state, A' and B' to reactant partitions.*

The eigenvector belonging to the largest eigenvalue is the solution to our problem, the transformed orbital that has the largest overlap with ψ_R^{TS}. In addition, we get in $\mathbf{c}$ a set of other vectors orthogonal to each other, that define a transformation of other orbitals in the subspace. However, except for the largest eigenvalue case, it is not known if this transformation is the best possibility. This is because although ψ'_R is granted to have the largest overlap with ψ_R^{TS}, the rest of the transformed reactant orbitals are such that they are orthogonal to ψ'_R and with that constraint have the largest overlap with ψ_R^{TS}, rather than their own corresponding counterparts among transition state orbitals. This means that there is no guarantee that reactant orbitals other than the breaking-bond orbital have become "more similar" to their transition state equivalents. In the followings, attempts to achieve such a transformation will be introduced. The aim of these discussions is to show the line of thought that brought us to the transformation eventually used, rather than proposing either rigorous or practical alternatives.

The first idea would be to try and maximize the overlap of all orbital pairs involved in the two spaces. If we assume that we have already ordered the orbitals in a way that the ones we want to pair up have the same index in each subspace, then we may use the following trial function to be maximized:

$$\begin{aligned}\max F &= \max_{\psi_i'} \left(\sum_i \langle \psi_i^{TS} | \psi_i' \rangle^2 + \sum_{ij} \lambda_{ij} (\delta_{ij} - \langle \psi_i' | \psi_j' \rangle) \right) \\ &= \max_{c_{ij}} \left(\sum_{ijk} c_{ij} c_{ik} S_{ij} S_{ik} + \sum_{ij} \lambda_{ij} \left(\delta_{ij} - \sum_k c_{ik} c_{jk} \right) \right) \\ &= \max_{c_{ij}} \left(\sum_{ijk} c_{ij} c_{ik} P_{jk}^i + \sum_{ij} \lambda_{ij} \left(\delta_{ij} - \sum_k c_{ik} c_{jk} \right) \right) \end{aligned} \tag{48}$$

where we have chosen the Lagrange multiplier method to express the constraint, $\langle \psi_i' | \psi_j' \rangle = \delta_{ij}$, that the transformed orbitals are orthonormalized. Here P_{jk}^i is the same as in the original transformation (39) and (47), except that it belongs to arbitrary orbital pair i instead of R, which is denoted in the upper index. This yields a similar result to what we had before, since after derivation with a certain parameter c_{ij} only terms corresponding to the chosen i survive, resulting in:

$$\mathbf{cP^i} = \mathbf{\Lambda c} \tag{49}$$

for each orbital pair i. In other words, this defines a separate transformation for each i, where the solutions for the largest eigenvalues are in maximum overlap with their TS pairs. But the rest of the eigenvectors have the same problem as those of the transformation above, and although with the largest eigenvalue cases we have a set of reactant orbitals with granted maximum overlap with the corresponding transition state orbitals, they are not orthonormal, as they are the products of unrelated transformations. We will try to move on for a more global solution.

Our last thought in this line is to try to not only maximize the overlap for a certain TS-RE pair i, but also to minimize it for $i \neq j$. In this case our function takes the following form:

$$\max F = \max_{\psi_i'} \left(\sum_i \langle \psi_i^{TS} | \psi_i' \rangle^2 - \sum_{\substack{ij \\ i \neq j}} \langle \psi_i^{TS} | \psi_j' \rangle^2 + \sum_{ij} \lambda_{ij} (\delta_{ij} - \langle \psi_i' | \psi_j' \rangle) \right)$$

$$= \max_{c_{ij}} \left(\sum_{ijk} c_{ij} c_{ik} P_{jk}^i - \sum_{\substack{ijk \\ i \neq j}} c_{jk} c_{jl} P_{kl}^i + \sum_{ij} \lambda_{ij} \left(\delta_{ij} - \sum_k c_{ik} c_{jk} \right) \right) \tag{50}$$

Now, after derivation one gets the following equation:

$$\mathbf{c} \left(\sum_i \mathbf{P^{i,x}} \right) = \mathbf{\Lambda c} \qquad \mathbf{P^{i,x}} = \begin{cases} -\mathbf{P^i} & \text{if i}\neq\text{x} \\ +\mathbf{P^i} & \text{if i=x} \end{cases} \tag{51}$$

for each orbital pair x. These solutions now have a relation to others, but still they are to be defined for each pair separately, so still suffer from the original problem.

Now let us try a different approach. Assume that by solving (49) for each i, we get a set of solutions belonging to the largest eigenvalues. The problem with these is that they are not necessarily orthonormal. One would think that we could try orthonormalizing these, with a method that transforms them to the least possible extent while doing so, which would be Löwdin's symmetric procedure. Unfortunately, this above mentioned set of transformed orbitals seems to be overdetermined, which prevents the actual application of the method.

However, the *idea* of the Löwdin procedure still appears useful, since it is a method which changes the orbitals to the least possible extent. If one follows this trail, one should take a look at how the procedure was derived. The "similarity" mentioned before is defined in the least squares sense, that is, minimizing the following functional of the initial non-orthonormal orbitals ($\psi_{0,i}$) and the resulting ones (ψ_i):

$$\sum_i \int |\psi_i - \psi_{0,i}|^2 \mathrm{d}\tau = \sum_i \langle \psi_i - \psi_{0,i} | \psi_i - \psi_{0,i} \rangle \tag{52}$$

This should be minimized using the constraint that the resulting orbitals are orthonormal. The idea here in relation to our problem is that TS and transformed RE orbitals should be most similar in the least squares sense:

$$\sum_i \langle \psi_i' - \psi_i^{TS} | \psi_i' - \psi_i^{TS} \rangle \tag{53}$$

We can now expand the above term, and since we know that both the overlap of ψ_i^{TS} and ψ_i' with themselves is constant (we want them to be 1), the minimization only affects the mixed terms, which appear with a negative sign. For this reason, these terms need to be maximized. With the introduction of the orthonormality constraint to the transformed orbitals, the function which we need to maximize is:

$$\begin{aligned} \max F &= \max_{\psi_i'} \left(\sum_i (\langle \psi_i^{TS} | \psi_i' \rangle + \langle \psi_i' | \psi_i^{TS} \rangle) + \sum_{ij} \lambda_{ij} (\delta_{ij} - \langle \psi_i' | \psi_j' \rangle) \right) \\ &= \max_{c_{ij}} \left(2 \sum_{ij} c_{ij} S_{ij} + \sum_{ij} \lambda_{ij} \left(\delta_{ij} - \sum_k c_{ik} c_{jk} \right) \right) \end{aligned} \tag{54}$$

where we have used the fact that we are working with real orbitals, and S_{ij} is the familiar asymmetric overlap between TS and RE orbitals. After derivation, we get the following result:

$$\mathbf{S} = \mathbf{\Lambda c} \tag{55}$$

This result looks like a matrix decomposition. Given that $\mathbf{c}$ is unitary (orthogonal), and $\mathbf{\Lambda}$ is Hermitian (symmetric) because of the symmetry of the constraint, this (assuming Λ is positive semidefinite) is equivalent to the polar decomposition of a matrix. Proceeding accordingly, one gets:

$$\mathbf{\Lambda} = \mathbf{S c}^{-1} \qquad \text{and} \qquad \mathbf{\Lambda}^T = (\mathbf{c}^T)^{-1} \mathbf{S}^T = \mathbf{\Lambda} \tag{56}$$

from which

$$\mathbf{\Lambda}^2 = \mathbf{S}(\mathbf{c}\mathbf{c}^T)^{-1}\mathbf{S}^T = \mathbf{S}\mathbf{S}^T \qquad \Rightarrow \qquad \mathbf{\Lambda} = (\mathbf{S}\mathbf{S}^T)^{\frac{1}{2}} \tag{57}$$

If we express $\mathbf{c}$ from (55), and insert (57), we get:

$$\mathbf{c} = (\mathbf{S}\mathbf{S}^T)^{-\frac{1}{2}}\mathbf{S} \tag{58}$$

The idea was based on the usefulness of the Löwdin transformation with respect to our problem, and followed the concepts of a basic derivation of it.[108] The fact that this transformation - similar to (54) - maximizes the sum of pairwise overlaps of the two sets involved was also obtained by Löwdin's own proof.[107] A more general proof of the symmetric orthonormalization, leading to the Carlson-Keller theorem[109] does not require that the two sets span the same subspace, which is closer to our situation. More importantly, it has also been shown that the minimization in the least squares sense yields a global extremum of the overlap.[110] All these results have been used for orbital transformations related to diabatic states, for example in those using the effective Hamiltonian method. In a related article, also the role of polar decomposition is hinted at.[111]

Similarly, based on previous results with 2×2 Jacobi rotations[112] of orbitals with similar criteria, and a more general one[113] based on (58) was also implemented in MOLPRO as the first step of the DDR[114] procedure. This procedure aims at the determination of the mixing of states beyond the Born-Oppenheimer approximation, i.e. diabatic states, in adiabatic wavefunctions. This requires the determination of two wavefunctions, the one belonging to the geometry of interest and the one where the states are assumed diabatic for instance because they cannot mix due to symmetry reasons. More explicitly, the procedure requires the calculation of the following quantity:

$$\langle\Psi_1(R)|\frac{\partial}{\partial R}|\Psi_2(R)\rangle = \frac{\langle\Psi_1(R)|\Psi_2(R+\Delta R) - \Psi_2(R)\rangle}{\Delta R} \tag{59}$$

where the R and $R+\Delta R$ refer to the geometrical parameters of the (diabatic) reference (Ψ_1) and target wavefunctions (Ψ_2). The wavefunction is then written as a CI expansion, introducing orbitals, and eventually, yields eq (44) or (45). In this manner, the math-

ematical formulation of the problem bears resemblance to ours, and although we may not need the whole procedure, as our transformation uses the same equations - relying only on the overlap matrix - the same code is therefore applicable to our case. See the implementation in the MOLPRO routine `mudiab.f` for (58).

There is one point of interest left for discussion. It is worth comparing transformations (49) and (58). They both start off similarly: equation (49) from maximizing the sum of squares of pairwise overlaps, whereas (58) can be traced back to maximizing that of plain pairwise overlap. However, while (49) needs a separate equation for each orbital pair i, and on top of that, these are not related, equation (58) provides a unique orthonormal set of orbitals where the sum of overlaps of orbital pairs i are maximized. The reason for this difference is that maximizing the pairwise overlaps of orbitals is based on the two indexed quantity $\mathbf{S}$, and as a consequence of taking the square of the overlaps, in (49), although the quantities are formally also two indexed, a third index is introduced through the matrix $\mathbf{P^i}$:

$$\mathbf{P^i} = \underline{S}^i \otimes \underline{S}^i \qquad (\underline{S}^i)_j = S_{ij} \tag{60}$$

where $\underline{S}^i$ is the overlap vector belonging to orbital pair i, and as such, a vector in $\mathbf{S}$; and as $\mathbf{P^i}$ is an outer product of these, the previously discussed dependence of this transformation on the choice of orbital pairs is obvious.

3.6 Results and Conclusions

Considering the reasons mentioned before, and keeping computational feasibility in mind, our calculations were carried out at the RCCSD/aug-cc-pVDZ level, which should be able to supply all the relevant information we are looking for. Geometry optimizations are rather crude, but our purpose was not to achieve the best accuracy but to analyse partitioning for which any sensible geometry would do. B3LYP/6-31G(d) geometry optimisations were carried out in most of the cases; for some of the cases the HF model was used, a fact that will be referred to at the appropriate places. All calculations were carried out with the MOLPRO[115] and Gaussian[116] program packages. In Figure 6 some

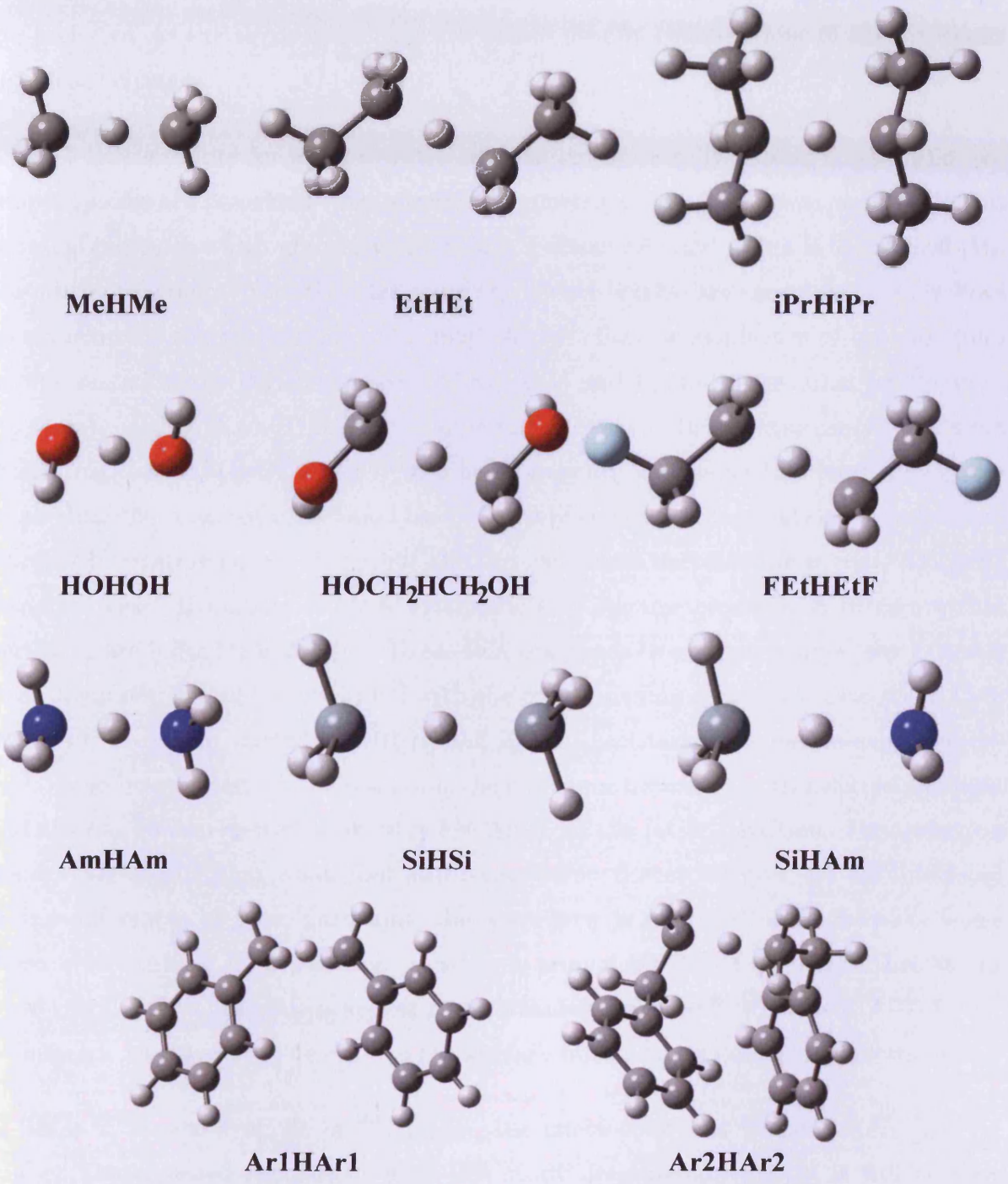

Figure 6: *Some of the species (mainly symmetric) involved in the discussion of partitioning with the codes used to refer to them later. They will be explained at the appropriate places.*

species are shown that are involved in the subsequent discussion to make their identification and the recognition of their geometrical arrangements easier. Symmetric species are preferred, as the asymmetric ones are gained usually combinations of these without significant changes.

Table 1 shows results for alkanes which seem to behave ideally. Methyl, ethyl and isopropyl species are presented with selected asymmetric cases. The position of '+' in this notation indicates which group has the H as a reactant (A) and which is the radical (B). Results for reactants, transition states and the barrier heights are shown, using only Boys localization for the partitioning. One may observe that the symmetry of the molecules is represented in the partitions (see e.g. E_{ts}(A-A) and E_{ts}(B-B) are equal for symmetric cases), or that E_r(A-B) is zero as expected due to the large separation. It is more interesting that ΔE(B-B) is nearly zero here, meaning that using this methodology orbitals from the H acceptor partition have relatively negligible contribution to the barrier. Another interesting note is that while all other partitions decrease the barrier, ΔE(A-A) shows increase. If we now compare symmetric and asymmetric cases, it turns out that partitions are indeed transferable. To see this, one needs to compare values (say E(A-A)) in an asymmetric case (as iPrH+Et) with the corresponding symmetric case (E(A-A) in iPrHiPr). In general, excluding E(R-R) and E(A-B) partitions, for reasons explained before, it can be said that for the reactants, the difference between the transferred partition and the one to be replaced is about a few 0.01% of the latter partition. For transition states, this is somewhat worse, but still below 0.5%. Since, however, we are interested in the differences of these partitions, the agreement is expected still somewhat worse there; errors arising from partition transfers is around 2.5-3% of the barrier height. In these cases, E(R-R) partitions appear to be transferable as well, whereas E(A-B) for an asymmetric case seems to be close to the average of the corresponding symmetric ones.

In Table 2, on the example of CH_3HCH_3, the effect of orbital transformation will be studied.The expected conclusions from this study are twofold. Firstly, it will be seen how orbital transformations influence partitioning calculations, and if there are any improvements compared to the use of untransformed localied orbitals. Ifso, the transformed partitions should get closer to zero. Secondly, the effect of geometrical differences on the results of the transformation are also investigated. All geometry optimizations here were carried out at the RHF/aug-cc-pVDZ level of theory, transition state structures were par-

Table 1: *Partitioning of correlation energies of methyl (Me), ethyl (Et) and isopropyl (iPr) species. A '+' indicates how separation occurs in reactants in asymmetric cases. All results are in Hartree units.*

	MeHMe	EtHEt	iPrHiPr	iPrH+Et	iPr+HEt
			Reactants		
E_r(RHF)	-79.761209	-157.839519	-235.920392	-196.878512	-196.881371
E_r(A-A)	-0.124223	-0.281651	-0.440806	-0.440810	-0.281648
E_r(A-B)	0.000000	0.000000	0.000000	0.000000	0.000000
E_r(A-R)	-0.038786	-0.040398	-0.042169	-0.042167	-0.040400
E_r(B-B)	-0.120351	-0.276930	-0.434998	-0.276930	-0.434992
E_r(B-R)	-0.039033	-0.042842	-0.046099	-0.042848	-0.046119
E_r(R-R)	-0.028479	-0.028610	-0.028709	-0.028709	-0.028612
$E_{c,r}$	-0.350872	-0.670432	-0.992782	-0.831465	-0.831771
			Transition States		
E_{ts}(RHF)	-79.705837	-157.784792	-235.865902	-196.825582	-196.825582
E_{ts}(A-A)	-0.120348	-0.276161	-0.435379	-0.435579	-0.276796
E_{ts}(A-B)	-0.002582	-0.004275	-0.006852	-0.005297	-0.005297
E_{ts}(A-R)	-0.044734	-0.048254	-0.051617	-0.051376	-0.048335
E_{ts}(B-B)	-0.120348	-0.276161	-0.435379	-0.276796	-0.435579
E_{ts}(B-R)	-0.044734	-0.048255	-0.051617	-0.048335	-0.051376
E_{ts}(R-R)	-0.042782	-0.042147	-0.042043	-0.042110	-0.042110
$E_{c,ts}$	-0.375528	-0.695253	-1.022885	-0.859493	-0.859493
			Barriers		
ΔE(RHF)	0.055372	0.054727	0.054490	0.052930	0.055789
ΔE(A-A)	0.003875	0.005490	0.005427	0.005231	0.004852
ΔE(A-B)	-0.002582	-0.004275	-0.006852	-0.005297	-0.005297
ΔE(A-R)	-0.005948	-0.007857	-0.009447	-0.009209	-0.007935
ΔE(B-B)	0.000003	0.000769	-0.000380	0.000134	-0.000587
ΔE(B-R)	-0.005701	-0.005412	-0.005518	-0.005486	-0.005257
ΔE(R-R)	-0.014302	-0.013537	-0.013334	-0.013401	-0.013499
ΔE_c	-0.024656	-0.024821	-0.030103	-0.028028	-0.027723
Total	0.030716	0.029906	0.024387	0.024902	0.028066

tially optimised with breaking and forming C-H bond lengths restricted to be the same. In the first column as a comparison the untransformed, localized orbital partitioning is shown. Next to it are the results after a Löwdin type transformation on the reacting orbitals. To account for the misalignment and corresponding rotational discrepancy between transition state and reacting orbitals, which may have a bad effect on the transformation, a third column is included showing transformation carried out at a reactant geometry that is gained by freezing the segment structure of A of the transition state at large separation. Here only the A segment is aligned, since we would expect ΔE(A-A) to vanish similarly as ΔE(B-B) in the untransformed case, since this transformation results in reactant orbitals very similar to the transition state ones. The transformation only affects A and R orbitals following this expectation. Also, the fact that ΔE(B-B) and, therefore, spatial segment B (i.e. the radical) orbitals can be left out of the transformation means that all of the approximations of the overlap matrix discussed earlier would be suitable for our purposes, but using $\mathbf{D}^{RE}$ is the most convenient. The only R orbital that we include is that of the breaking bond, which is reasonable given the fact that that orbital is spatially closely related to partition A. However, the ΔE(A-A) partitions still remain non-zero after the transformation, regardless of what geometry was used. Indeed, there is very little difference between the transformed and aligned cases. For this reason we can exclude orbital rotation as a cause for failure. The likely reason would be that during the transformation, the orbitals delocalize, and therefore the partitioning is ruined. In Figure 7, **A** denotes the doubly occupied R orbital in the localized case, and **B** is that of the transformed case. It can be seen that whereas the first orbital is neatly localized around the required bond centres, the back lobe of the transformed orbital is spread over H atoms belonging to the A partition satisfying perhaps the requirement of greater similarity with transition state orbitals, but violating the separation between partitions. Since it is hard to find a transformation that maximizes similarity and yet respects localization, we decided to simply use localized orbitals calculated at the optimized reactant and transition state structures in all our calculations. The somewhat surprising success of this choice of Boys localization over some elaborate transformations perhaps follows from the fact that it respects partitioning. It should also be noted that Pipek-Mezey localization was also tested, and yielded similarly good results. This is important, as for larger systems the Pipek-Mezey procedure is more feasible, and it also respects the σ/π symmetries.

Table 2: *Partitioning of correlation energies of the CH_3HCH_3 species, without transformation (localized) and with transformation with and without geometry alignment (transformed and aligned, respectively). All results are in Hartree units.*

	localized	transformed	aligned
	Reactants		
E_r(RHF)	-79.761298	-79.761298	-79.759540
E_r(A-A)	-0.124041	-0.109443	-0.109303
E_r(A-B)	0.000000	0.000000	0.000000
E_r(A-R)	-0.038791	-0.060414	-0.059985
E_r(B-B)	-0.120112	-0.120111	-0.120113
E_r(B-R)	-0.038998	-0.038998	-0.038998
E_r(R-R)	-0.028416	-0.021392	-0.022086
$E_{c,r}$	-0.350358	-0.350358	-0.350485
	Transition States		
E_{ts}(RHF)	-79.705977	-79.705977	-79.705977
E_{ts}(A-A)	-0.120191	-0.120191	-0.120191
E_{ts}(A-B)	-0.002614	-0.002614	-0.002614
E_{ts}(A-R)	-0.044718	-0.044718	-0.044718
E_{ts}(B-B)	-0.120186	-0.120186	-0.120186
E_{ts}(B-R)	-0.044712	-0.044712	-0.044712
E_{ts}(R-R)	-0.042579	-0.042579	-0.042579
$E_{c,ts}$	-0.374999	-0.374999	-0.374999
	Barriers		
ΔE(RHF)	0.055321	0.055321	0.053563
ΔE(A-A)	0.003850	-0.010748	-0.010888
ΔE(A-B)	-0.002614	-0.002614	-0.002614
ΔE(A-R)	-0.005927	0.015696	0.015267
ΔE(B-B)	-0.000074	-0.000074	-0.000073
ΔE(B-R)	-0.005714	-0.005714	-0.005714
ΔE(R-R)	-0.014162	-0.021187	-0.020493
ΔE_c	-0.024641	-0.024641	-0.024514
Total	0.030680	0.030680	0.029049

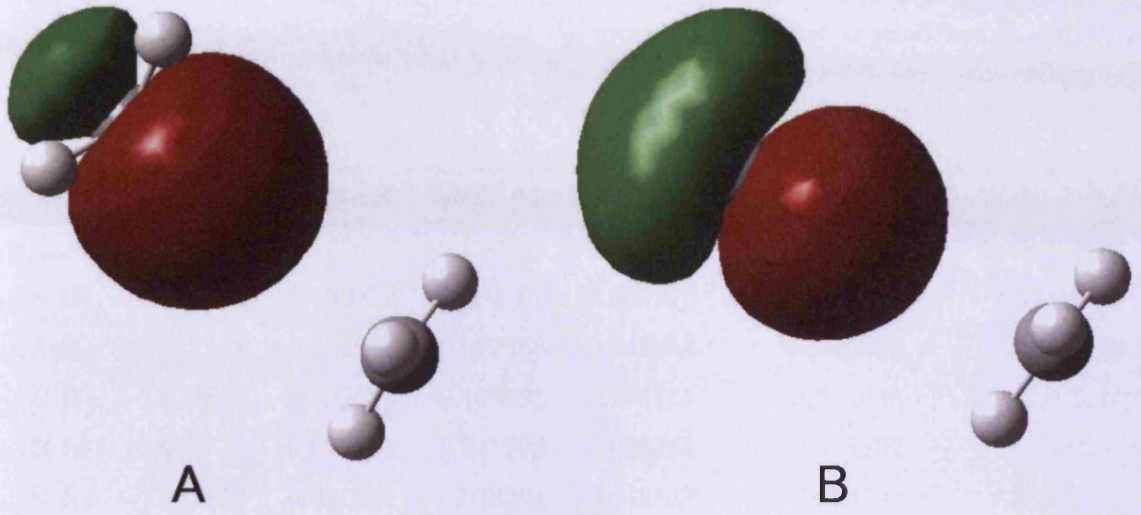

Figure 7: *Comparison of doubly occupied R orbitals in cases when only localization was used (***A***) and when a Löwdin type orbital transformation was also carried out (***B***).*

Next in Table 3, partitions are shown for cases other than hydrocarbons. From this point on, transition state results will be shown, as significant new developments are expected with these, as there is a larger interaction between orbitals in these cases. In the followings, asymmetric species will be labelled in a manner that partition A comes first, and partition B last, the two separated by the H. For example, in MeHOH Me is partition A and OH is partition B, results are to be found in the tables accordingly. The first significant difference is that here ΔE(B-B) is no longer zero with these species. Examining the asymmetric cases shows that the partition transfer causes a larger error than with alkanes, especially with partitions connected with R (E_{ts}(A-R), E_{ts}(B-R)). For MeHOH, the error goes near 6% for E_{ts}(B-R), which might cause an even larger error in the barriers. The cause here is what would be called traditionally the delocalization of the SOMO over atoms which have lone pairs like O. This delocalization is numerically indicated by the large AO coefficients of the SOMO belonging to O. Since the SOMO is uniquely defined, this cannot really be dealt with via localization. Consequently, there will be a violation of partitions since an R orbital mixes with those of other partitions. This would explain why E_{ts}(A-R) and E_{ts}(B-R) are more affected. If we take the other asymmetric case, where delocalization is less likely due to the fact that there the O atom is not the one directly associated with H transfer (being attached to the carbon atom that

Table 3: *Partitioning of correlation energies for some species containing oxygen as a heteroatom. In asymmetric cases the group written first is partition A. All results are in Hartree units.*

	MeHMe	HOHOH	MeHOH	EtHEt	$HOCH_2HCH_2OH$	$EtHCH_2OH$
E_{ts}(A-A)	-0.120348	-0.129679	-0.120758	-0.276161	-0.311845	-0.276613
E_{ts}(A-B)	-0.002582	-0.009122	-0.004107	-0.004275	-0.004046	-0.004038
E_{ts}(A-R)	-0.044734	-0.072698	-0.042256	-0.048254	-0.046802	-0.047899
E_{ts}(B-B)	-0.120348	-0.129681	-0.130991	-0.276161	-0.311845	-0.312417
E_{ts}(B-R)	-0.044734	-0.072701	-0.071292	-0.048255	-0.046802	-0.046267
E_{ts}(R-R)	-0.042782	-0.041013	-0.043304	-0.042147	-0.043199	-0.042525
$E_{c,ts}$	-0.375528	-0.454894	-0.412707	-0.695253	-0.764540	-0.729759

is), the errors shrink back to around 1%. This seems to support the above theory. Why there is still a significant discrepancy, there are some factors to be considered. Firstly, that delocalization is still possible. Secondly, that even if there would be no delocalization possible, electrostatic differences between the H donor and acceptor groups could have a role (different electronegativities). In the remainder of our tables, we seek to gain some further insight into the role of these factors.

Let us, therefore, consider some further species without lone pairs: SiH_3, NH_3^+, Me and their asymmetric combinations as shown in Table 4. These geometries were obtained at the HF/6-31G(d) level, since B3LYP had some convergence problems. Interesting to note that SiH_3 and NH_3^+ groups are eclipsed if we take the Si-H-N line as an axis in $SiH_3HNH_3^+$. This might alter with the use of diffuse funcions. Although it is evident from the optimisations obtained with these species that neither HF nor B3LYP are good choices for accurate calculations, as mentioned before, we do not persue accuracy in our calculations as in theory those should work for any geometries. For this reason we settle with the fast HF or DFT methods, whichever one fits better to the actual system. Among the latter species NH_3^+ seems to have the largest electron withdrawing effect due to its positive charge, then in decreasing order Me and SiH_3. It can also be observed that self partitions (E(A-A) and E(B-B)) are affected to an extent of 0.3-1.5% during the transfer to the asymmetric species. Interestingly, the correlation energy involved in these partitions (meaning its absolute value) decreases in the asymmetric case compared

Table 4: *Partitioning of correlation energies for species containing the groups CH_3 (Me), NH_3^+ (Am) and SiH_3 (Si). In asymmetric cases the group written first is partition A. All results are in Hartree units.*

	MeHMe	AmHAm	SiHSi	AmHMe	SiHMe	SiHAm
E_{ts}(A-A)	-0.119894	-0.126924	-0.104265	-0.125708	-0.104659	-0.105806
E_{ts}(A-B)	-0.002490	-0.002895	-0.000956	-0.002897	-0.001478	-0.001868
E_{ts}(A-R)	-0.044630	-0.054572	-0.024510	-0.057316	-0.023338	-0.018545
E_{ts}(B-B)	-0.119894	-0.126924	-0.104265	-0.120529	-0.119493	-0.125428
E_{ts}(B-R)	-0.044630	-0.054571	-0.024510	-0.041197	-0.045203	-0.057183
E_{ts}(R-R)	-0.043010	-0.048632	-0.040967	-0.042135	-0.042207	-0.040515
$E_{c,ts}$	-0.374549	-0.414518	-0.299474	-0.389782	-0.336377	-0.349344

to the symmetric for the group with the larger withdrawing effect and increases for the other. The opposite tendency can be observed for E_{ts}(A-R) and E_{ts}(B-R). Here, however, the difference is 5-10% for the majority of the cases. SiHMe is the least affected as there is relatively little difference in withdrawing between the two groups. The largest difference is in SiHAm where E_{ts}(A-R) is differing to that in SiHSi with 32% (-0.018545 vs. -0.024510 Hartree, respectively), as a result of matching the strongest and weakest electron withdrawing groups. It is, therefore, evident that electrostatic effects have a bad influence on the transferability of partitions through changing the electron density associated with a group in different molecular environments, while leaving the orbitals localized and the orbital partitioning unharmed.

Intrinsic delocalization of π-electrons over the partition boundaries may be another source of problems. If a phenyl group is present on both the donor and the acceptor atoms, as in the case of benzyl groups, then these may cause the delocalization of R orbitals over the whole molecule that remains even after localization (Ar1 case). By inserting a CH_2 group between the phenyl group and the donor/acceptor carbon atom, the delocalization may be reduced (Ar2 case). The single point calculations here were carried out at the RCCSD/6-31+G(d) level of theory considering the size of the species, and are shown in Table 5. The reason why E_{ts}(A-A) and E_{ts}(B-B) partitions are not equal in the Ar2HAr2 "symmetric" case is because the Ar2 groups are not completely equal owing to the rotational freedom

Table 5: *Partitioning of correlation energies for aromatic species, including benzyl (Ar1) and a derivative group (Ar2 = Ar1-CH_2). In asymmetric cases the group written first is partition A. All results are in Hartree units.*

	MeHMe	Ar1HAr1	Ar2HAr2	Ar1HMe	Ar2HMe	Ar1HAr2
E_{ts}(A-A)	-0.120191	-0.866969	-1.016990	-0.869596	-1.016671	-0.869458
E_{ts}(A-B)	-0.002614	-0.007793	-0.008185	-0.003150	-0.003462	-0.005414
E_{ts}(A-R)	-0.044718	-0.051706	-0.044289	-0.047605	-0.044065	-0.049024
E_{ts}(B-B)	-0.120186	-0.866969	-1.017747	-0.100310	-0.100152	-1.015386
E_{ts}(B-R)	-0.044712	-0.051706	-0.044286	-0.037464	-0.038625	-0.045015
E_{ts}(R-R)	-0.042579	-0.032400	-0.034705	-0.032232	-0.034711	-0.032603
$E_{c,ts}$	-0.374999	-1.877542	-2.166203	-1.090357	-1.237685	-2.016900

arising from the insertion of CH_2. It can be observed that the transferability of self partitions (E_{ts}(A-A) or E_{ts}(B-B)) belonging to aromatic groups are relatively little affected (0.3%). The transferability ruining effect on the E_{ts}(A-R) partition in Ar2HMe (0.5%), or even E_{ts}(B-R) in Ar1HAr2 (1.6%) both belonging to Ar2 are small compared to that of Ar1 in Ar1HMe (5-9%) which is due to the reduced delocalization in species with an Ar2 group in line with our expectations. Somewhat surprisingly, the transferability of the Me group is badly affected in Ar1HMe and Ar2HMe in both E_{ts}(B-B) and E_{ts}(B-R) terms (15-20%). This could not be explained with delocalization, and perhaps it is due to the strong electron withdrawing effect of the aromatic groups. This seems to suggest that delocalization may have an undesired effect on partitioning and transferability, but this seems to be milder than what arises through electrostatic differences, which affects even groups with strongly local orbitals (like those of Me), possibly by decreasing the electron density associated with them.

We originally set out to establish a local methodology that would make quantum chemical calculations more feasible, by saving the effort of having to calculate all the contributions to correlation energy, provided they are known from a previous calculation. This rests heavily upon the condition of transferability of these partitions, which was examined here in detail. It was found that alkanes behave ideally from this point of view, and that such procedures would probably be successful with them. We have found numerous

obstacles, however, with species that contained different functional groups. Electrostatic effects were identified as being very problematic in cases where there is a large change in the molecular environment from this respect in the species between which the transfer is to be made. Delocalization of π and lone pair orbitals has a bad effect as well. For all these reasons, a methodology envisaged on the above basis, would be of lesser use, but still there is a possibility that it would yield reasonable approximations. We have also seen that in calculating barrier heights, localized orbitals performed better than specially constructed transformations to make reactant orbitals similar to transition state ones, because the transformed orbitals violate partition boundaries.

What remains yet to be seen, reveals little new developments, but serves to further illustrate some points already discussed. All the remaining cases are optimized at the RHF/aug-cc-pVDZ level of theory with symmetric bond lengths restricted to be the same. Some further examples of hydrogen transfer from alkane derivatives are examined via results shown in Table 6. If we direct our attention to the energy differences between reactants and transition states, it is again very obvious that ΔE(B-B) is very nearly zero in all cases. If we now consider, say EtHMe in Table 6, it can also be concluded that other partitions except for E(A-B) are virtually the same as corresponding partitions in EtHEt and MeHEt. E(A-B) being a dispersion-like term, seems to be the average of terms from EtHEt and MeHMe. Similar arguments hold for MeHEt (where the Me group is the H donor), where situation is analogous, except that A and B related terms are exchanged, as expected, compared to the previous case. It can be concluded that these partitions are transferable. Interesting to note that E(R-R) seems nearly constant in these cases, and E(R-R) of an asymmetric case is perhaps closer to that of the H-donor. Furthermore, the effect of F in the substituted cases on transferability is negligible, which is in line with results in Table 3; notably those of the OH containing species. It was observed there that the effect of the OH group decreases as it gets farther away from partition R (or more precisely from the donor/acceptor carbon atom). Here, F is even further away, with a CH_2 group inserted between the head atom and itself, which is proven enough to restore transferability.

In Figure 8, the contribution of partitions to the total correlation correction to the barrier height is shown as a percentage of this correction for various symmetric cases. Figure 8 shows some results if one performs partitioning after localization without any transfor-

Table 6: *Results for symmetric H abstraction cases with methyl, ethyl and fluoro-ethyl groups, and for some of the asymmetric cases obtained from combining these. Results are in Hartree units at the RCCSD(T)/aug-cc-pVDZ//RHF/auc-cc-pVDZ level of theory. In the asymmetric cases, the groups written first belong to the A partition (i.e. those are the H donors).*

	MeHMe	EtHEt	FEtHEtF	EtHMe	MeHEt	FEtHEt	EtHEtF
E_r(RHF)	-79.761298	-157.840054	-355.570315	-118.798942	-118.802410	-256.706563	-256.703806
E_r(A-A)	-0.124041	-0.281207	-0.468956	-0.281203	-0.124045	-0.468959	-0.281208
E_r(A-B)	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
E_r(A-R)	-0.038791	-0.040387	-0.041230	-0.040389	-0.038789	-0.041227	-0.040387
E_r(B-B)	-0.120112	-0.276657	-0.463245	-0.120112	-0.276657	-0.276658	-0.463244
E_r(B-R)	-0.038998	-0.042127	-0.043877	-0.038998	-0.042126	-0.042123	-0.043874
E_r(R-R)	-0.028416	-0.028535	-0.028399	-0.028536	-0.028415	-0.028398	-0.028535
$E_{c,r}$	-0.350358	-0.668912	-1.045706	-0.509238	-0.510032	-0.857366	-0.857247
E_{ts}(RHF)	-79.705977	-157.784992	-355.515248	-118.745625	-118.745625	-256.650803	-256.650803
E_{ts}(A-A)	-0.120191	-0.276628	-0.463985	-0.276634	-0.120128	-0.463807	-0.276665
E_{ts}(A-B)	-0.002614	-0.004262	-0.003882	-0.003374	-0.003374	-0.004105	-0.004105
E_{ts}(A-R)	-0.044718	-0.048114	-0.049532	-0.047980	-0.044869	-0.049985	-0.047843
E_{ts}(B-B)	-0.120186	-0.276615	-0.463945	-0.120128	-0.276634	-0.276665	-0.463807
E_{ts}(B-R)	-0.044712	-0.048110	-0.049530	-0.044869	-0.047980	-0.047843	-0.049985
E_{ts}(R-R)	-0.042579	-0.042159	-0.042239	-0.042332	-0.042332	-0.042118	-0.042118
$E_{c,ts}$	-0.374999	-0.695889	-1.073113	-0.535316	-0.535316	-0.884523	-0.884523
ΔE(RHF)	0.055321	0.055062	0.055066	0.053317	0.056784	0.055760	0.053003
ΔE(A-A)	0.003850	0.004579	0.004970	0.004569	0.003917	0.005152	0.004542
ΔE(A-B)	-0.002614	-0.004262	-0.003882	-0.003374	-0.003374	-0.004105	-0.004105
ΔE(A-R)	-0.005927	-0.007727	-0.008302	-0.007591	-0.006080	-0.008758	-0.007456
ΔE(B-B)	-0.000074	0.000042	-0.000700	-0.000015	0.000023	-0.000007	-0.000563
ΔE(B-R)	-0.005714	-0.005984	-0.005654	-0.005871	-0.005854	-0.005719	-0.006111
ΔE(R-R)	-0.014162	-0.013625	-0.013840	-0.013797	-0.013917	-0.013720	-0.013583
ΔE_c	-0.024641	-0.026977	-0.027407	-0.026078	-0.025284	-0.027157	-0.027276
Total	0.030680	0.028086	0.027659	0.027239	0.031500	0.028603	0.025726

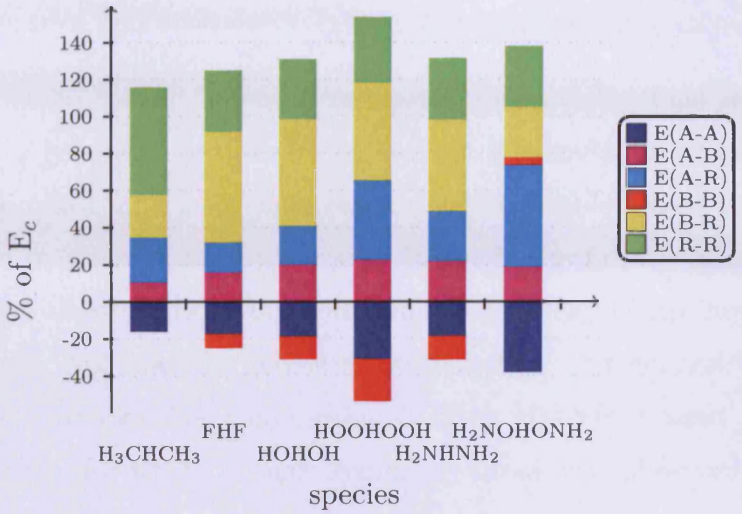

Figure 8: *Partitioning of barrier heights for different species as a percentage of the total barrier.*

mation. ΔE(B-B) is only zero for MeHMe, otherwise it has a negative contribution in this Figure. It is because of the exceptional behaviour of E(B-B) with alkyl groups that the idea of transforming A orbitals so that ΔE(A-A)=0 originally arose. This, of course means, that such a transformation is expected to be the most useful with alkyl groups. Since the correlation energy difference is negative, that is, it decreases barrier heights, a negative value in Figure 8 means a contribution that actually increases the barrier. It can be observed that in general self partitions ΔE(A-A) and ΔE(B-B) are partitions that increase, while changes in reactant partition E(R-R) and the mixed partitions (including the dispersion term E(A-B) and mixed terms with R) decrease the barrier. In the case of $H_2NOHONH_2$ ΔE(B-B) is positive, but nothing conclusive can be said about this other than noting the fact. The results here should be treated with care, as we have seen, these results can in theory be biased by a transformation like the one described before, as long as their sum (the total correlation energy) remains the same. A table with the data of Figure 8 is included in the Appendix.

3.7 A Summary of the Procedure

Here a brief summary of the partitioning procedure discussed so far will follow with highlights on some possible areas for interest:

1. Initial geometries and orbitals

   - RHF wavefunction: possible symmetry break.[83,84] This makes the convergence of geometry optimizations with such calculations hard except for well defined cases, and even then results maybe problematic. A typical example for the manifestation of symmetry break is the failure transition state optimization once the breaking and forming bond legths are not restricted to be the same. Not ideal for geometry optimization, but generally works with correlation methods (as a compromise, often B3LYP is used to get optimized geometries, for which no such symmetry break was observed).
   - MCSCF wavefunction. Geometry optimizations converge, but the method would make it necessary to use some multireference correlation theories. However, it could be used for geometry optimization in itself, although ideally a method for such purpose should contain at least some of the dynamic correlation energy.

2. Localization to replace local correlation methods[101] for open shell systems

   - Boys:[95] it seems to work well, might be improved by adjusting tolerance and convergence criteria
   - Possible uses of Pipek-Mezey localization (e.g. in cases where π orbital symmetry is significant)

3. Partitioning - splitting up the orbital space to orbitals which change during the reaction (R for reacting orbitals) and those remaining localized on the H carrier (A) and on the radical (B); partitioning pair energies accordingly

   - Possibly done via numerical or graphical means
   - In the numerical case, the number of R orbitals may depend on the parameters set in this procedure, therefore, the graphical solution is preferred, if sometimes accompanied by the numerical method

4. Transition state transformation of doubly occupied R and A orbitals so that the Coulomb interaction between R and the SOMO are maximized. This is a possibility we have not considered here

- Other possible transformations, though the one used seems to work well according to some experimental calculations, though not better than localized orbitals

5. Reactant transformation - maximizing overlap related quantities between the transition state and the reactant R and A orbitals, transforming reactant R and A accordingly

   - Special geometrical arrangements are needed, the A segments should overlap as our transformations operate in the A⊕R space. Partial geometrical alignment is possible, but does not improve results much. The 'noorient' option in MOLPRO should be used for calculations
   - Transition state and reactant orbitals need to be paired up, since the transformations depend on orbital pairs. This can again be done numerically or graphically, and again, the latter is preferred
   - Approximation of overlap matrix **S**. With the present geometrical arrangement, and conditions for transformation, this could either be approximated with reactant or transition state overlap matrices. The former one is chosen here
   - Transformations using the sum of the squared absolute value of pairwise overlaps, suffer from dependence on the orbital pair the transformation is based on (i.e. they are not unique). If separate orbital pair based transformations are carried out, an extra orthonormalization step would be required
   - Transformations based on the sum of pairwise overlaps. This transformation is unique and yields a global extremum and a set of orthonormal orbitals
   - For these reasons, the second type of transformations is preferred, although as it turns out, transformations ruin partitioning through delocalization, and should therefore be completely neglected. However, there may exist more ideal solutions, e.g. Mayer's maximally localized orbitals[117] that may work well with problems like ours.

All this results in the conclusion that localized orbitals should be used due to the lack of other similarity transformations respecting partitioning. Results have been presented,

and the use of localized orbitals are recommended for similar procedures with the notes given above.

Finally, some possibilities are suggested how despite the above mentioned difficulties this method may still evolve into practical applications. As it was seen, the main hindrance for transferability had been electrostatic effects. This could be explained with assuming that in different molecular environments different electron densities are associated with a chosen transferred group, and so, the correlation energy partition associated with the group also changes. If in some manner, the partitioning could be corrected with the change in electron density, which is still easily calculated compared to high level correlation methods, transferability may be restored, resulting in the provisional simplifications described in the beginning of this chapter. Several possibilities may be open for such a correction, and this chapter will be closed by mentioning a few of these on a purely speculative basis. The simplest possibility that comes to one's mind is a sort of semiempirical set of parameters specific to groups and environments that would yield a better agreement with more accurate methods. It should also be possible to come up with a modification of localized methods that are invariant to density effects. Finally, it might also be possible to account for the varying environment in the SCF procedure in a manner similar to DFT using parametrized external potentials.

4 The Case of Propene + OH

4.1 Overview

An important special case of asymmetric H-transfer reactions plays an important role among hydrocarbon oxidation mechanisms. These are important in many areas of science, from understanding and reducing pollutant formation in combustion to describing partial oxidation in fuel cells.[118,119] It is widely accepted that the most common initial reaction of hydrocarbons in the atmosphere[120] and in all hydrocarbon frames[121] is the attack by an OH radical. Since propene can be a prototype of 1-alkenes, it is essential to characterize its relevant reactions to understand the chemical behaviour of 1-alkenes with the OH radical.

It is well-known that OH is able to attack the double bond of alkenes in terminal (**T**) and central (**C**) positions (Figure 9). These addition reactions take place via a van der Waals complex (vdW-complex), a so-called π-complex (**R**). However, the importance of the hydrogen transfers such as the consecutive ($\mathbf{A_{con}}$) and direct ($\mathbf{A_{dir}}$) allylic H-abstractions in the case of alkenes + OH reactions has been recognized only recently.[122]

Although the propene and hydroxyl radical system has been studied previously and reported in several theoretical papers[123–128] allylic H-abstraction channels were neglected in most cases. Earlier studies[123–126] have focused mostly on the ratio of the terminal and central addition reaction rates. Although Cvetanović reported in his work that 65% of the additions occur at the terminal carbon atom,[129] theoretical calculations at both the MP2/6-31+G(d)[126] and MP4(SDTQ)/6-31G(d,p)//MP2/6-311G(d,p)[123] levels of theory showed that central addition is preferred. However, it is emphasized in both theoretical works that the energy and entropy differences of the terminal and central transition states

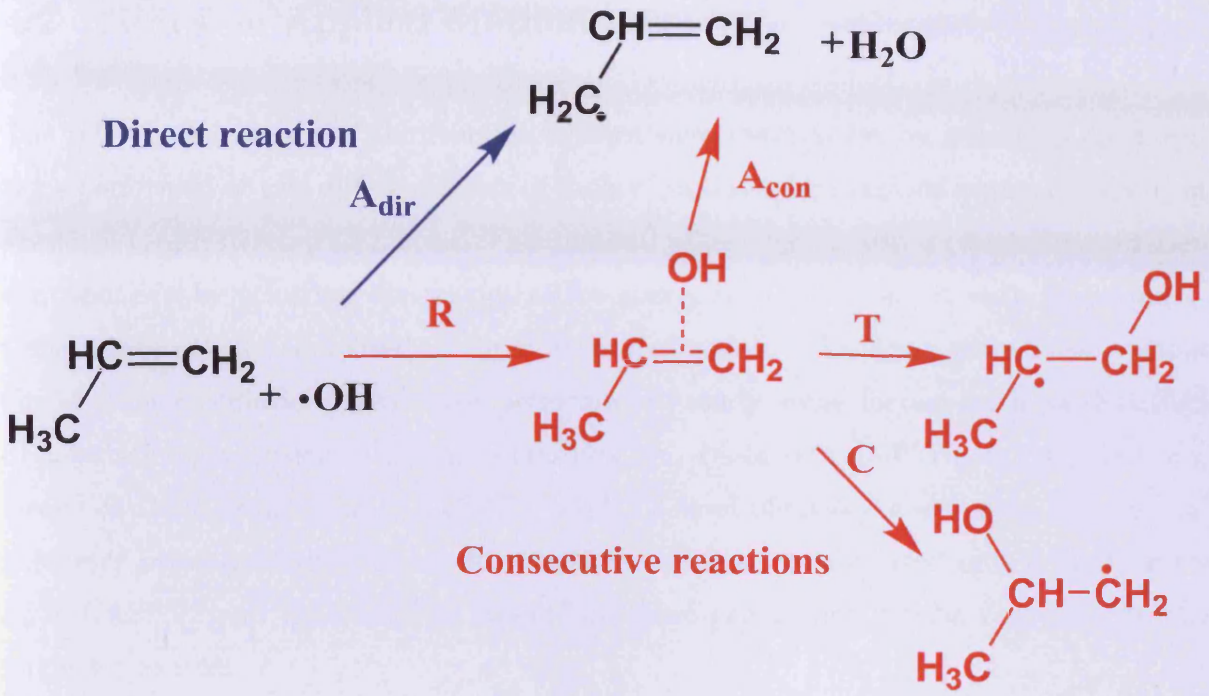

Figure 9: *The studied direct and consecutive reactions of the propene + OH system.*

are quite small.[124] The interest in the kinetic behaviour of the 1-alkene + OH system is also shown by papers published very recently[127,128] in this field, papers in which theoretical calculations offered a mainly kinetic description of the system. These works are mostly based on PMP2/aug-cc-pVQZ//MP2/cc-pVTZ[127] and CCSD(T)/cc-pVDZ//B3LYP/cc-pVTZ[128] methods. We have found relevant discrepancy between these latter two potential energy surfaces, although both state that their results are good descriptions of the overall kinetics. These results will be discussed to some extent later on. All this has led us to determine the accurate energetics of transition states corresponding to the energetically favoured reaction channels with small difference in their energetics. Based on this set of calculations we are able to provide a highly accurate framework for kinetic modelling, as well as a procedure for the logical choice for the active space in such asymmetric species. On the other hand, our aim was also to provide highly reliable results from benchmark ab initio calculations for further tests with density functional methods for larger alkene homologues.

4.2 Notes on Applied Methods

The relevant structures of the reaction system were determined by geometry optimizations performed at two different levels of theory. Initial optimizations were carried out at the [5, 5]-CASPT2/cc-pVTZ level. This method allows for choosing relevant correlation contributions by selecting the proper active space, and, therefore, it reduces computational requirements compared to more accurate models. On these geometries, various single point calculations were then performed to study some factors such as the effect of multireference treatment, spin contamination, basis sets and triples contributions. Based on these results, the UCCSD(T)/cc-pVTZ level of theory was chosen for the final geometry optimizations. Harmonic vibrational analysis was carried out at both levels ([5,5]-CASPT2 and UCCSD(T)). Results on these geometries will be compared in the following section.

For the [5, 5]-CASPT2/cc-pVTZ level of theory, the active space should involve the SOMO in all cases. For **C** and **T** the π bond must be involved, since it participates in the C-O bond formation, and in the remainder of cases, this orbital corresponds to the most mobile electrons out of doubly occupied orbitals (highest orbital energy in RHF reference). For $\mathbf{A_{dir}}$ and $\mathbf{A_{con}}$ the breaking C-H bond must also be involved. For consistency, a C-H bond is involved in the active space for all cases. This makes the treatment balanced since the active space contains contributions for all non-hydrogen atoms for all species. Reactants (1-propene and OH) were treated in the supermolecular approach with a 1000 Å separation, and share the same active space structure. This results in an active space of 5 electrons placed in 5 orbitals, 2 of which are unoccupied in the Hartree-Fock configuration. The occupied orbitals are chosen by first localizing the initial RHF orbitals, then after analysing the basis function contributions, the relevant orbitals may be identified. Similar procedures have been discussed in the literature, addressing the difficulty of choosing a balanced active space resulting in a correct correlation treatment.[85,130] Local orbitals simplify the choice of occupied orbitals, however the difficulty of choosing the right virtual orbitals still remains. Here only the active occupied orbitals are preselected, and the virtuals are chosen purely on the basis of energetic order from the RHF reference. This procedure seems sufficient, since after the MCSCF optimization the active virtuals are the π^*(C-C) and the σ^*(C-H) antibonding orbitals as desired, see Figure 10. It can be

seen that all active orbitals are well localized. This remains the case even if augmented basis sets are used. For further details see the following section.

The multiconfiguration nature of the wavefunction assures that the wavefunction is qualitatively correct, the long range static correlation effects having been considered - avoiding the dissociation related problems of single reference methods. The choice of CASPT2 method ensures that the most relevant short range dynamic electron correlation effects important for geometry optimization are also considered. In the optimizations, the active space described above was used in all cases, for consistency, even in the cases, where the C-H bond remains intact. The removal of this orbital and a corresponding virtual one from the active space; or the choice of an alternative C-H orbital, however, does not influence the result of the optimization significantly. Neither does the use of a basis set augmented with diffuse functions results in any relevant change. In both the case of the modified active space and the basis set augmentation, the resulting change is of the order of a few 0.01 Å in bond lengths, a few 0.1 degrees in angles, corresponding to perhaps a few 0.01 kJ/mol in the calculated energies due to reoptimization.

For a further improvement in our results various single point calculations were carried out with different basis sets, and high level correlated methods. Various kinds of multireference (MR) methods were used, beginning with CASPT2 and CASPT3 perturbative methods. Although in some ways CASPT3 is an improvement over CASPT2, for barrier heights it definitely seems inferior.[131] These methods are cheaper alternatives of the more expensive MRCI method, namely in our calculations the internally contracted MRCI with singles and doubles (ICMRCISD[63]) was used. With the MRCI results, denoted as Q1 and Q2, the Davidson corrected energies for fixed and relaxed references respectively are given in an attempt to make the wavefunction size consistent by adding approximate quadruple corrections. As observed in the literature,[64] Q2 usually yields poorer agreement with FCI, and should be used only in special cases. Following this, iterative size consistency correction methods follow, namely MRACPF and MRAQCC, which are two variants of an approximate MRCC. Both have a tendency to overshoot the correlation energy, the first one more than the latter.[65] For some further details about these methods, see e.g. reference 65.

Various single reference (SR) methods are also presented, the reliability of which depends

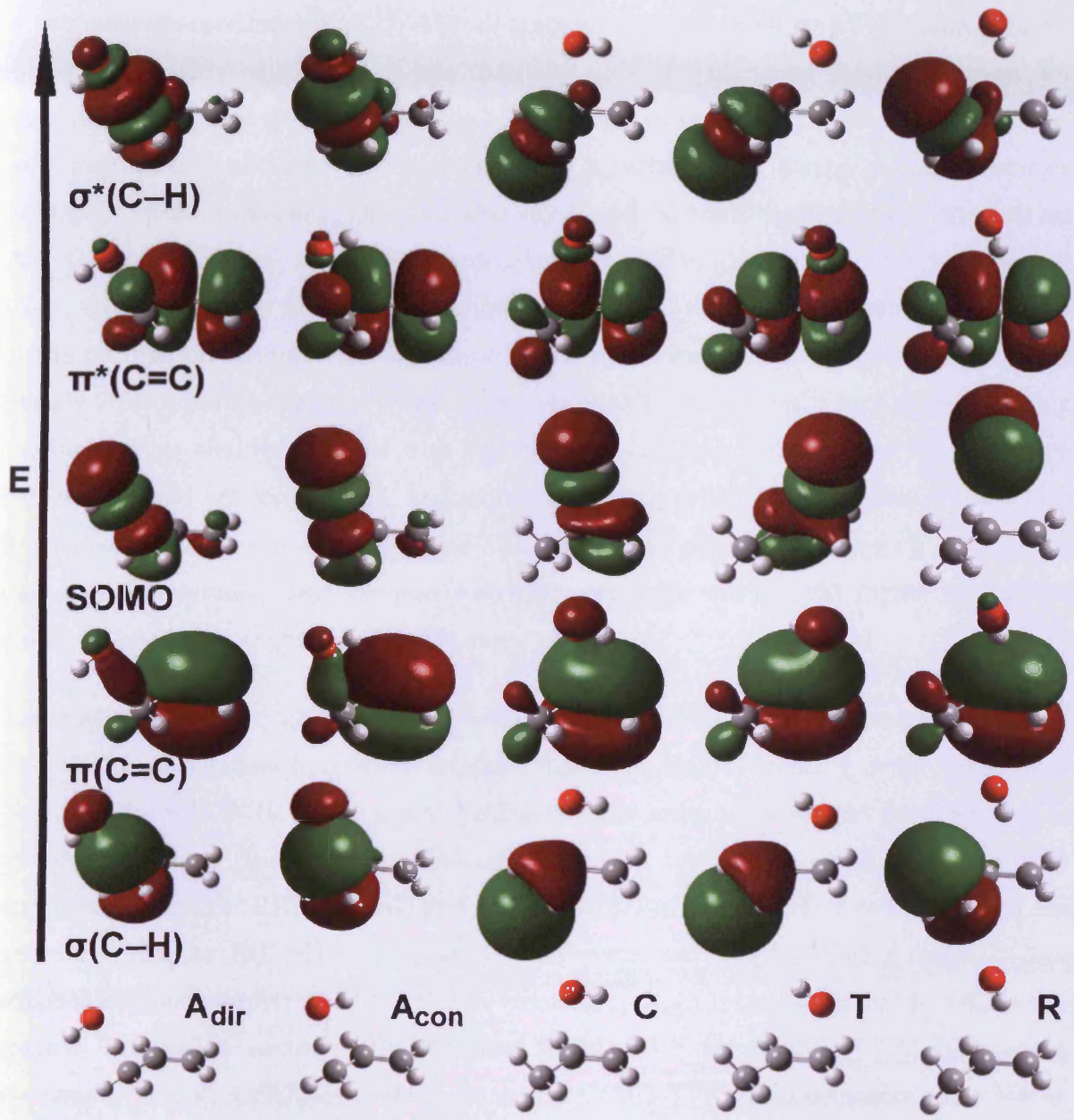

Figure 10: *[5,5]-CASSCF/cc-pVTZ active orbitals for the π-complex (**R**) as well as transition state structures of direct ($\mathbf{A_{dir}}$) and consecutive ($\mathbf{A_{con}}$) allylic H-abstraction; terminal (**T**) and central (**C**) OH-addition reaction channels.*

on whether the wavefunction is dominated by one configuration during the calculation. To answer this, one can first take the T1 diagnostics (from CCSD calculations) which indicates the significance of higher excitations and therefore the possibility of a need for a multireference treatment.[132] For all transition states, with cc-pVQZ basis, the T1 values are roughly equivalent or less than the critical 0.02 value (and well below for minima). As will be seen later on, the contribution of connected triple excitations has an important role, and with that included using a perturbative ansatz, a single reference treatment seems sufficient. One can also say based on multiconfiguration calculations that the weight of the ground state configuration is dominating (about 0.97) over all the rest (about 0.02 or less) at all examined geometries, and this dominance shows itself in the occupation numbers as well, those being quite close to the reference state values (nearly 2 for bonding orbitals, around 1 for the SOMO and about 0 for the antibonding orbitals). This slightly changes with relaxation in MR calculations (see the difference between Q1 and Q2 corrections), and more significantly with the expansion of the active space (the dominance is still conserved although less evident). All this well justifies the use of single reference methods, and a further advantage will be that higher excitations are more feasible to include in the SR case.

The RMP2 values are gained as intermediate values in the coupled cluster procedure. The MP2 model suffers from some artefacts due to its lack of treating single excitations (for a study with FHF see Fox and Schlegel;[84] the arguments should hold for any H transfer with lone pairs close to the radical centre). A variety of CC methods were also used, these are: RHF-RCCSD and RHF-UCCSD models, RHF here referring to the reference orbitals. RCCSD is the partially spin restricted coupled cluster method (spin adapted in linear terms, which results in virtually no spin contamination[50]). Triples are treated in a variety of ways: the standard CCSD(T),[49] the simpler CCSD[T] missing the usually important singles contributions and CCSD-T[133] which considers some higher order perturbation terms compared to CCSD(T).

To approximate the non-relativistic limit, extrapolations were carried out based on the cc-pVXZ bases[29,31] (X=D,T,Q), where the three point exponential formula of Feller[32] was used for HF and MCSCF results, and the two point X^{-3} function form[34] was used for correlation energies with X=T,Q. This latter choice is usually not too different from X=D,T, the most significant difference being with Davidson corrected energies, which

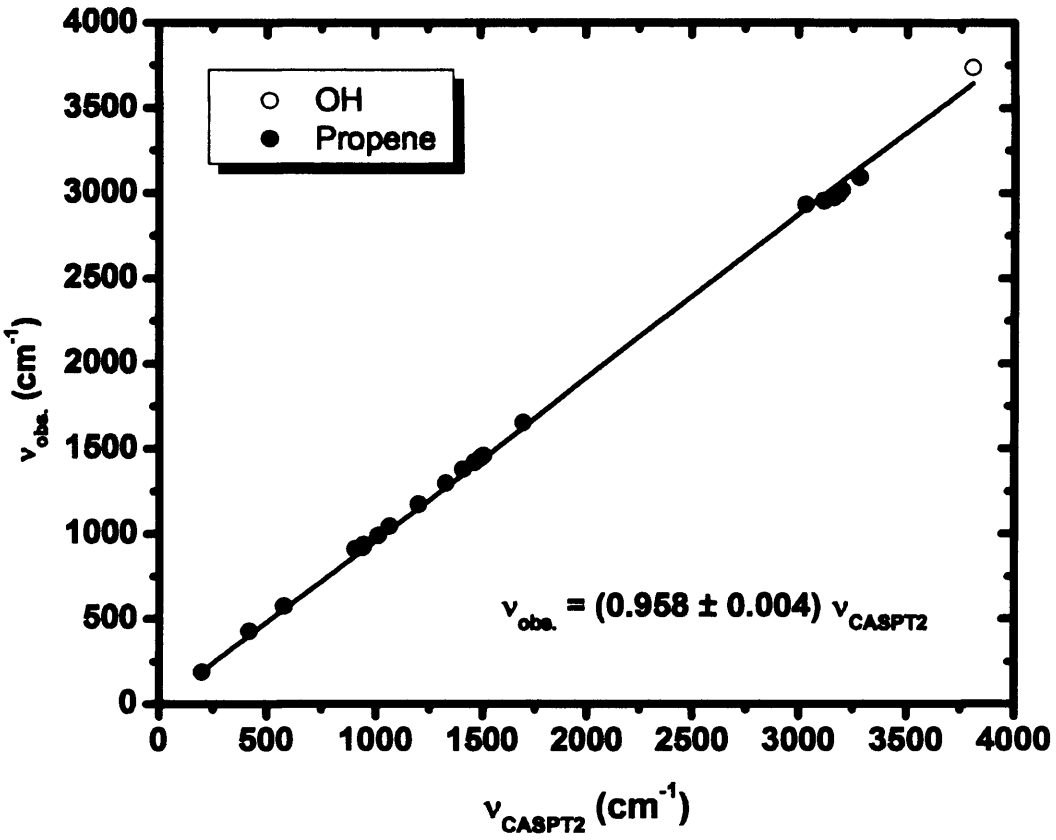

Figure 11: *Fitting the [5,5]-CASPT2/cc-pVTZ frequencies of Propene and OH against experiment.*

show a slower convergence. In the multiconfiguration calculations the choice of the RHF or MCSCF wavefunctions as reference in the Davidson correction does not introduce significant differences – which might be the case if there was a significant amount of dynamic correlation in the active space. In the RHF case, the extrapolation was checked against cc-pV5Z results, and it was found that the difference in predicted RHF barriers is less than 0.01 kJ/mol. The effect of augmented bases were also studied using aug-cc-pVXZ bases[30,31] with X=D,T for correlation energies, and for the references an additional X=Q was calculated.

Finally, some additional calculations were carried out in a less systematic fashion with smaller basis sets due to their computational cost. These include calculations with extended active spaces and some UHF-UCCSD(T) calculations for comparison. The explicitly correlated model UCCSD(T)-F12a[134] with the recommended basis (AVTZ) was

calculated on the UCCSD(T) geometries. Similarly, the extrapolated UCCSD(T) energies for these structures were determined. In some cases, restricted active space (RASSCF) calculations, and the corresponding correlated methods[135] prefixed with RAS were used. This means that we only allow certain excitations in certain regions of the active space. Here in the case of an extended [9,9] active space only double excitations are allowed from the lower two occupied, and to the upper two unoccupied orbitals, which reduces computational cost (singles are eliminated due to numerical reasons rather than due to their quantity).

Most of the calculations were carried out with the MOLPRO program package of Werner and Knowles.[115] For the CCSDT and UHF-UCCSD(T) calculations, the MRCC package of Kállay[136] was used. Vibrational frequencies were calculated on the optimized geometries at the [5,5]-CASPT2/cc-pVTZ and UCCSD(T)/cc-pVTZ levels. To ensure a better agreement with experiment, the scale factor 0.958±0.004 was determined for [5,5]-CASPT2 by fitting the calculated frequencies against the experimental values for propene[137] and OH,[138] see Figure 11. For the UCCSD(T) frequencies the scale factor used is 0.975±0.0021.[139] For comparative purposes, a variation of G3MP2B3 procedure[80] was also carried out where the B3LYP/6-31G(d) geometry optimization and normal mode analysis were replaced by the BH&HLYP/6-31G(d) level of theory due to the fact that the B3LYP functional is not able to characterize the transition state for the consecutive allylic H-abstraction.[122] The BH&HLYP harmonic frequencies were scaled by 0.935.[122] In analogy to G3MP2B3, we term this method G3MP2BH&H, and refer to its earlier use in our publications.[140] All the DFT results were obtained using the Gaussian program package.[116] All enthalpy values are relative to that of the level of propene and OH. Additional details of calculations like geometries, frequencies and other energetic and thermodynamic data are to be found in the Appendix.

4.3 Results and Discussion

First, let us discuss the single point results at the [5,5]-CASPT2 geometries. In Table 7 relative enthalpy results extrapolated from the cc-pVXZ basis sets are shown. The first obvious observation is that there is a significant difference between the multireference and

Table 7: *Method dependence of relative standard enthalpy values (in kJ/mol) obtained by extrapolation of cc-pVXZ basis sets for the π-complex (**R**), transition states of direct (**A**$_{dir}$) and consecutive (**A**$_{con}$) allylic H-abstractions as well as terminal (**T**) and central (**C**) OH-additions.*

	A_{dir}	A_{con}	C	T	R
[5,5]-CASPT2	4.62	2.44	-9.00	-7.45	-10.37
[5,5]-CASPT3	22.10	20.63	6.66	7.08	-8.61
[5,5]-MRCI	28.88	28.19	11.70	11.50	-6.26
[5,5]-MRCI+Q1	15.36	13.97	2.39	6.44	-8.95
[5,5]-MRCI+Q2	16.64	15.20	3.60	7.22	-9.28
[5,5]-MRACPF	16.67	15.19	1.50	2.11	-8.78
[5,5]-MRAQCC	18.49	17.13	3.05	3.54	-8.41
RMP2	15.94	13.95	-3.91	-3.69	-10.96
RCCSD	16.60	15.29	-0.28	-0.19	-8.32
RCCSD[T]	5.69	3.91	-9.41	-8.60	-10.08
RCCSD-T	7.21	5.44	-8.12	-7.44	-10.00
RCCSD(T)	7.20	5.44	-8.19	-7.54	-10.03
UCCSD	14.35	13.04	-2.05	-1.84	-8.38
UCCSD[T]	2.89	1.15	-11.54	-10.60	-10.26
UCCSD-T	4.60	2.87	-10.16	-9.33	-9.96
UCCSD(T)	4.79	3.06	-10.02	-9.24	-10.01

the single reference results, especially when comparing the size consistency corrected MR and the triple corrected CC results, the ones that can be considered as the most reliable from the corresponding sets; this difference is approximately 10 kJ/mol.

Let us first analyse the SR results. The RMP2 result agrees with the CCSD results best, which is not surprising. The CCSD models show a considerable difference between results with and without triples corrections in both the restricted and unrestricted cases. Since the triples are important (see Table 7), the most reliable result must be among the corrected results. The CCSD[T] model as described above does not account for some important contributions, and is therefore inferior to the others, while the other

two methods agree very well, again in accordance with general experience. Therefore the choice of the standard CCSD(T) model is justified. One can still observe some difference between RHF-RCCSD and RHF-UCCSD results. We will return to this later, here we only say that in accordance with previous recommendations,[92,93] we choose the RHF-UCCSD results as the most reliable single reference ones, and will use this for comparison.

The MR results appear somewhat divergent. The CASPT2 results agree well with the UCCSD(T) single point ones within 1-2 kJ/mol, which supports the choice of the inexpensive CASPT2 method for geometry optimization. The CASPT3 results seem to overestimate the barrier heights, and so do the MRCI results because of the size consistency error. It should be noted that in the supermolecular approach size consistency is already approximately dealt with, but the inclusion of higher excitations may still be important. For this reason, the theoretically most reliable results here are the ones with some kind of a correction for the latter error. These (MRCI+Q1, MRCI+Q2, MRACPF and MRAQCC) give results within a broad 3 kJ/mol range. In all cases the difference between these is significantly smaller compared to that with the MRCI results, indicating the importance of higher excitations, and also the fact that the active space may be too small to involve all significant higher excitations. Indeed, if one compares these with the CCSD results (that is without triples correction), one finds a good agreement, showing that the MR calculations with the present active space is comparable with considering only SD excitations. We will come back to this later.

In Table 8, we present some results coming from extrapolation using augmented basis sets for selected methods. In general, there is a good agreement between the two extrapolations, they mostly differ for $\mathbf{A_{dir}}$ and $\mathbf{A_{ind}}$ in MR calculations, and for $\mathbf{R}$ in general. $\mathbf{R}$ being a weakly bound π-complex, longer range interactions are usually more important, which the augmented basis sets handle better (diffuse functions). The augmented basis sets also show a faster convergence. For all these reasons we will prefer results with augmented bases in the followings, and refer to the extrapolations from these as the complete basis set (CBS) limit (see Table 9).

Table 9 provides data to compare some issues of the various calculations carried out at different geometries. These include the effects of the choice of active space and spin

Table 8: *Method dependence of relative standard enthalpy values (in kJ/mol) obtained by extrapolation of aug-cc-pVXZ basis sets for the π-complex (**R**), transition states of direct ($\mathbf{A_{dir}}$) and consecutive ($\mathbf{A_{con}}$) allylic H-abstractions as well as terminal (**T**) and central (**C**) OH-additions.*

	$\mathbf{A_{dir}}$	$\mathbf{A_{con}}$	**C**	**T**	**R**
[5,5]-CASPT2	4.66	2.55	-9.01	-7.58	-10.93
[5,5]-MRCI	28.73	28.07	11.45	11.19	-6.85
[5,5]-MRCI+Q1	16.93	15.59	2.17	11.19	-9.06
[5,5]-MRCI+Q2	18.33	17.01	3.33	4.62	-9.03
UCCSD	14.38	13.16	-2.21	5.56	-9.16
UCCSD[T]	3.05	1.43	-11.51	-10.66	-10.96
UCCSD-T	4.70	3.08	-10.18	-9.43	-10.64
UCCSD(T)	4.90	3.28	-10.02	-9.34	-10.68

contamination for CASPT2 geometries, and convergence tests to judge how far our results are from the complete basis set and from the FCI limit. It begins with the [9,9]-RAS-MRCI+Q1 results. This [9,9] active space is the [5,5] extended with the two C-C bonds, and two unoccupied orbitals (and with excitations restricted from/to these extensions). This extension of the active space improves the agreement with single reference methods on the same double zeta basis, except for **R**. In the next step, we improve the basis set by augmentation, and the active space by removing the restriction of double excitations. The resulting [9,9]-MRCI+Q1 values are now even comparable with the extrapolated UCCSD(T) results but in the case of **R** there is no improvement. If we now take the [5,5]-MRCI+Q1 results (Table 8), it is obvious, that the major differences between SR and MR methods are in the case of the transition states, in the case of **R** there is actually a rather good agreement (differing by 1 kJ/mol only). Furthermore, for **R**, the triples contribution yields a contribution of 1 kJ/mol only in CCSD indicating that a consistent treatment of the triples does not change the result significantly. This would explain why [5,5]-MRCI+Q1, which was described above to have an overall SD quality agrees well with UCCSD(T) for the π-complex. In case of the transition state structures, the extension of the active space brings the desired improvement, indicating that the chosen active space gives a consistent treatment of important higher order excitations. With

Table 9: *Method dependence of relative standard enthalpy values (in kJ/mol) obtained at several levels of theory for the π-complex (***R***), transition states of direct (*$\mathbf{A_{dir}}$*) and consecutive (*$\mathbf{A_{con}}$*) allylic H-abstractions as well as terminal (***T***) and central (***C***) OH-additions. [5,5]-CASPT2 (scale factor 0.958), UCCSD(T) (scale factor 0.975) and BH&HLYP (scale factor 0.935) optimized geometries and frequencies are included.*

[5, 5]-CASPT2 geoms&freqs	$\mathbf{A_{dir}}$	$\mathbf{A_{con}}$	**C**	**T**	**R**
[9, 9]-RAS-MRCI+Q1/cc-pVDZ	12.30	10.81	-1.64	-1.54	-16.84
[9, 9]-MRCI+Q1/aug-cc-pVDZ	5.36	4.33	-10.15	-8.80	-18.22
RCCSD(T)/aug-cc-pVDZ	5.92	4.74	-11.65	-9.99	-12.97
UCCSD(T)/aug-cc-pVDZ	3.96	2.83	-13.13	-11.39	-12.82
UHF-UCCSD(T)/aug-cc-pVDZ	4.36	3.17	-12.59	-10.68	-12.99
UCCSD(T) geoms&freqs	$\mathbf{A_{dir}}$	$\mathbf{A_{con}}$	**C**	**T**	**R**
UCCSD(T)/cc-pVDZ	15.17	13.67	2.15	3.82	-11.03
CCSDT/cc-pVDZ	14.56	13.10	1.20	2.96	-11.04
UCCSD(T)/CBS	3.21	1.67	-9.93	-9.84	-10.56
UCCSD(T)-F12a/AVTZ	3.03	1.55	-10.54	-10.34	-10.47
BH&HLYP geoms&freqs	$\mathbf{A_{dir}}$	$\mathbf{A_{con}}$	**C**	**T**	**R**
G3MP2BH&H	0.74	-0.92	-6.60	-5.35	-8.43

R this does not seem to be the case, that is some important higher order contributions are included, whereas others are left out in the extended active space, which causes an unbalanced, inconsistent treatment. To recover consistency, one should change the active space. However, as it was described earlier, this is not an easy task,[85] which seems only necessary for **R**.

How to control which orbitals go into the active space? We have control over the occupied orbitals, but the virtual ones are harder to choose. Firstly, there is no symmetry condition which could help. One could perhaps see that orbitals with large contributions from the transferring H are involved, but even so, care should be taken to choose such that are only related to the transfer directly, and not to other interactions. This is a hard task in the case of a π-complex, where there are several competing non-covalent interactions. If we decide on not manipulating the virtual orbitals, one could try to change the occupied orbitals and hope that the MCSCF optimization will result in the desired virtuals. There are many possibilities to do this, here we only note that a [9,9] active space where the two C-C bonds and the C-H bond is replaced with the O-H bond, and the two lone pairs of the oxygen yields no better results (-22.77 kJ/mol for **R**). Since the [9,9] results did not bring improvement, one could try to increase or decrease the active space of **R**. Increasing the active space further is not feasible, neither is a larger basis set. A decrease would take us back to the already discussed [5,5] space, which indeed seems an improvement in consistency, which due to the less emphasized importance of triples contributions with **R**, shows itself as a good agreement with UCCSD(T). Since **R** is a minimum structure, it is less likely to have a multiconfigurational nature, so the UCCSD(T) result can be taken as the final word. This seems to be also the case with the transition states, since the above mentioned not too high T1 values seem to be taken care of by the triples correction, and also because the [9,9]-MRCI+Q1 results seem to converge there anyway (if we could allow the use of larger bases). From all this, our conclusion is that consistent MRCI+Q1 values with large enough active space and UCCSD(T) results agree well, and the latter should be chosen for computational and methodological ease.

The remainder of the first section of Table 9 addresses some spin related issues. Besides the RHF based treatments, here some UHF-UCCSD(T) results are also included. As pointed out earlier, there appears to be a roughly 2 kJ/mol difference between restricted and unrestricted CC results based on an RHF reference.[92,93] A somewhat smaller dif-

ference (a few tenths of kJ/mols) is observed between methods based on UHF and RHF orbitals.[92,93] It is difficult to reach firm conclusions on the nature of the spin contamination effects arising from the UHF or the UCCSD(T) procedures from these data. It is worth mentioning that the spin contamination belonging to these methods is around 0.3-0.4. Here, the UCCSD(T) model will be chosen[92,93] with RHF reference[94] as a preferred method for the final geometry optimizations and vibrational frequency calculations. Although the energetic difference between the RCCSD(T) and UCCSD(T) single points remains an open question, the issue may be addressed from a geometrical point of view. As discussed below, **T** (and also **C**) seems to be the most sensitive to correlation methods used for optimization. If one takes this species and optimizes the structure with RCCSD(T) and UCCSD(T) with 6-31G(d) basis, one gets quite similar geometries: the most sensitive parameter, the C$\cdots$O distance is 2.13 Å with the unrestricted, and 2.09 Å with the restricted method. If we now perform a UCCSD(T) single point calculation on the RCCSD(T) geometry (or vice versa) and compare it with the UCCSD(T) optimized value, there is only a slight 0.24 kJ/mol difference. On the other hand, the difference between the optimized energies is 2.55 kJ/mol, which corresponds to the above 2 kJ/mol gap between RCCSD(T) and UCCSD(T). This suggests that the choice of restricted or unrestricted CCSD(T) models has only a negligible effect on geometry optimizations in these cases, despite the energetic difference between the two. This is assumed to hold for all species and bases discussed here.

Having chosen the UCCSD(T) method, geometry optimizations and vibrational frequency analysis were performed with the cc-pVTZ basis. It is interesting to compare the [5,5]-CASPT2 and UCCSD(T) geometries to emphasize the good performance of the much cheaper CASPT2 method. In Figure 12, UCCSD(T) results are indicated first, then in brackets the CASPT2 ones follow. The most significant difference is the C$\cdots$O distance in **T** and **C** (about 0.12 Å) which is probably due to some neglected correlation contributions rather than spin contamination effects (see above). However, this only yields a difference of about 0.3 kJ/mol between UCCSD(T) and UCCSD(T)//[5,5]-CASPT2 barriers with cc-pVTZ basis. Comparing the extrapolated UCCSD(T) energies at 0 K, they agree within 1 kJ/mol, which is an excellent agreement. At 298.15 K, the maximum difference in enthalpies is a somewhat larger 1.5 kJ/mol, since in this case differences in frequencies also play a role. Allylic abstraction barriers are 1.5 kJ/mol lower at the

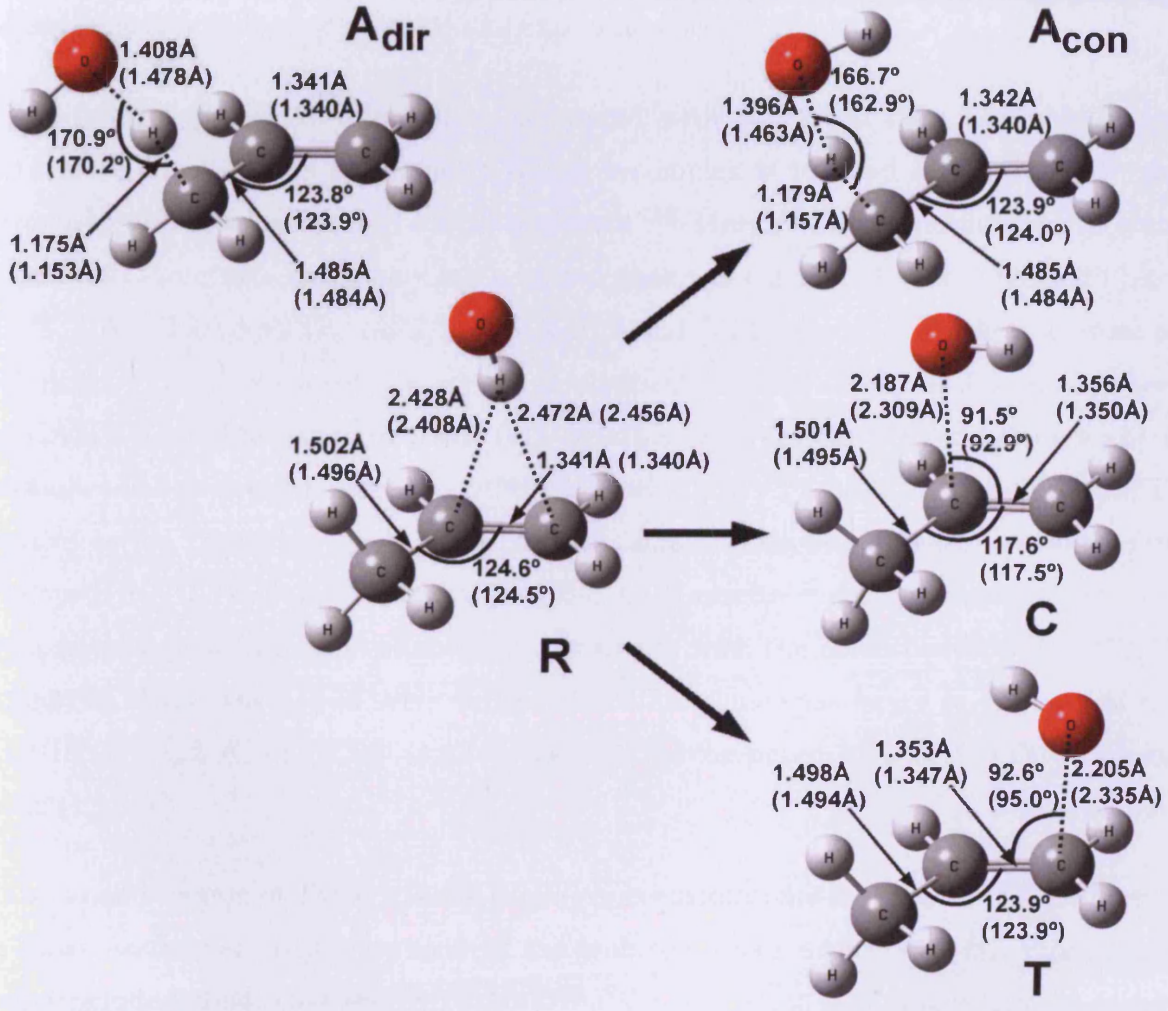

Figure 12: *UCCSD(T)/cc-pVTZ geometry parameters followed by [5,5]-CASPT2/cc-pVTZ ones in brackets for the π-complex (**R**) as well as transition state structures of direct ($\mathbf{A_{dir}}$) and consecutive ($\mathbf{A_{con}}$) allylic H-abstraction; terminal (**T**) and central (**C**) OH-addition reaction channels.*

UCCSD(T) geometries, and the addition barriers are also much closer to each other compared to the results with CASPT2 (0.09 vs. 0.65 kJ/mol difference). From now on, UCCSD(T) geometries will be used by default. Finally, it is also worth mentioning that if we take the extrapolated CASPT2 enthalpy barriers rather than the extrapolated UCCSD(T)//CASPT2 values as above and compare it with optimized UCCSD(T) values, the agreement is still very good (1-2 kJ/mol difference).

In the followings, our results will be compared with structural data available in the literature. The UCCSD(T) geometry of the π-complex is in good agreement with the previously published MP2/6-31+G(d) geometry.[126] Most geometry parameters of transition state structures corresponding addition channels calculated with UCCSD(T)/cc-pVTZ, CASPT2/cc-pVTZ, MP2/6-31+G(d)[126] and MP2/cc-pVTZ[127] are also close to each other. The only exceptions are the bonds being formed (C-O) in **T** and **C**, where the CASPT2 bond lengths are about 0.13 Å larger, whereas the MP2 results are about the same value shorter compared to UCCSD(T) with cc-pVTZ basis. In general CASPT2 predicts earlier transition states than those obtained by single reference methods. Our previous BH&HLYP/6-31G(d) and CCSD/6-31G(d) results[122] on the transition states of allylic hydrogen abstraction channels are consistent with the corresponding UCCSD(T) geometries. Here, the bonds being broken (C-H) are somewhat larger in the case of the BH&HLYP (1.22 Å) or CCSD (1.23 Å) geometries compared to UCCSD(T) transition states (1.18 Å).

In the second section of Table 9, some single point calculations on UCCSD(T) geometries are shown to further investigate some of the problems which occurred so far. Since it has been concluded above that the triples contribution is of great importance, results gained at the (unrestricted) CCSDT/cc-pVDZ level are included, together with the perturbative CCSD(T) results for comparison. The good agreement (within 1 kJ/mol) between CCSDT and UCCSD(T) results suggests that we can indeed rely on the latter as a good approximation for triples contribution.

Finally, the convergence of the basis set extrapolation using the CCSD(T) results is tested. An explicitly correlated theory, UCCSD(T)-F12a[134] – as implemented in MOLPRO – was utilized with the recommended AVTZ basis set. This improves basis set convergence, so that we can obtain accurate results with relatively small bases (differences from

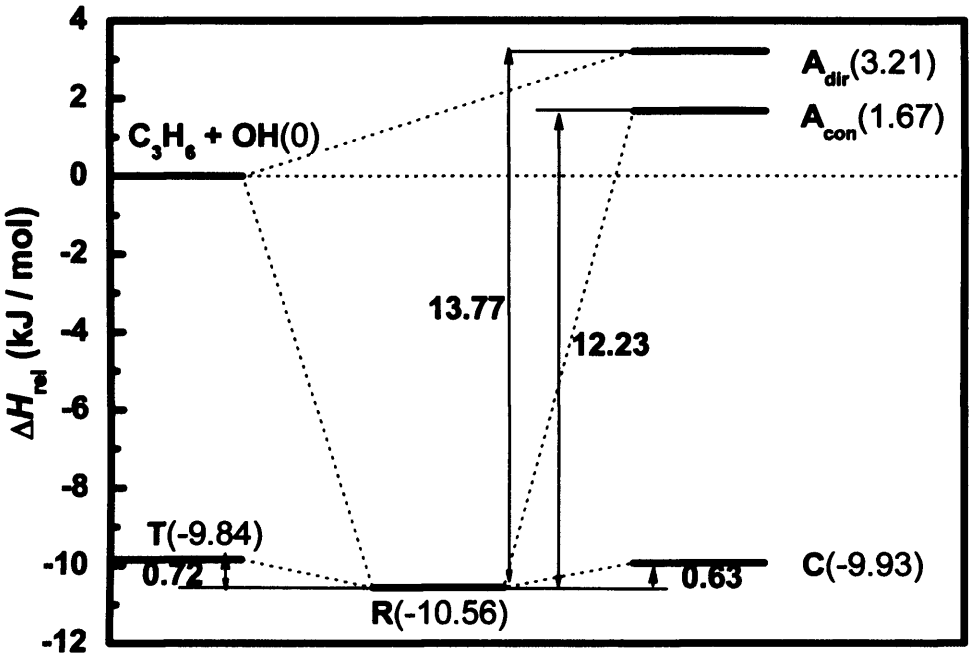

Figure 13: *Energetics of the examined structures at the UCCSD(T)/CBS// UCCSD(T)/cc-pVTZ level of theory.*

UCCSD(T)/CBS are within 0.7 kJ/mol). This procedure yields very similar results from what we had from extrapolation. It is also worth noting that using the G3MP2BH&H extrapolation scheme also gives quasi-quantitative answers, the differences being in the worst cases around 3-4 kJ/mol.

The UCCSD(T)/CBS result for **R** as shown in Figure 13 is -10.56 kJ/mol relative to the infinitely separated species. The addition barriers – as suggested by earlier theoretical works lie very close to each other, actually our calculations show that they are within 0.09 kJ/mol (the virtual activation enthalpies are -9.93 kJ/mol for **C** and -9.84 kJ/mol for **T**). The allylic H-abstraction enthalpy barriers are (about 12 kJ/mol) higher and have a larger difference, 1.54 kJ/mol, making the consecutive reaction energetically favoured (enthalpy barriers: 3.21 kJ/mol for $\mathbf{A_{dir}}$ and 1.67 kJ/mol for $\mathbf{A_{con}}$). Conventional transition state theory (cTST) might provide a rough estimation for the rate constants of these channels by means of our UCCSD(T) results. This approach is the most problematic with addition

barriers. For cTST to be valid, there should be an equilibrium between reactants and transition state. As there is a low lying π-complex to which the TS lies very close, while the reactants slide down on the potential energy surface forming the complex, their kinetic energy increases, since the total energy should remain constant. If as a result of this, the reactants pass over the TS rapidly, there is no time for an equilibrium to take place that would allow for the distribution of the energy to other modes. However, the attractive nature of the PES allows for such ballistic effects that make the distribution of energy and reaching an equilibrium possible. We assume a fourfold electronic degeneracy (g) for OH (g=4, ignoring spin-orbit splitting), g=2 for the transition states and g=1 for propene. A twofold degeneracy of reaction paths were considered in the case of the addition transition states, since they have non-superimposable mirror images (the OH can come from either side of the propene plain). In the case of direct H-abstraction, a threefold rotational degeneracy is assumed (although the conformer with the OH in the propene plain is expected to be energetically a bit different), whereas in the consecutive case the same degeneracy is two. Propene also has a threefold conformational degeneracy because of the methyl group (this only affects the abstraction cases). At room temperature, the rate constants are 4.86×10^{-14} (for $\mathbf{A_{dir}}$) 4.49×10^{-14} (for $\mathbf{A_{con}}$) 4.23×10^{-12} (for **C**) 8.33×10^{-12} (for **T**) in $cm^3molecule^{-1}s^{-1}$ which in total (1.27×10^{-11} $cm^3molecule^{-1}s^{-1}$) agrees within a factor of 2 with the value recommended by IUPAC at this temperature (3.02×10^{-11} $cm^3molecule^{-1}s^{-1}$).[141] This latter experimental value is derived from a temperature dependent formula recommended by IUPAC, and is partially based on work of Zellner and Lorenz,[142] who suggest a value of $(3.0\pm0.5)\times10^{-11}$ $cm^3molecule^{-1}s^{-1}$ at 298 K. Similarly, in Atkinson's review,[143] the suggested 2.63×10^{-11} $cm^3molecule^{-1}s^{-1}$ with a 15% error is again in good agreement with our results. The above mentioned factor of two corresponds to a 1.7 kJ/mol change in the barriers. However, in our estimation, the error of the energy calculations is only around 1 kJ/mol, or perhaps even less in some cases (see Table 9 for CCSDT benchmarks for triples errors, and also for basis set convergence). The rest of the discrepancy must come from sources like the choice of cTST for our estimations, the quality of the calculated frequencies (e.g. ignoring anharmonicity), and some other issues which a more thorough kinetic study should deal with. Since this was not our goal here, we consider our results in good agreement with experiment. To further support this point, the branching ratio for terminal addition (**T**, 65.8%) was calculated, and was found to be in near perfect agreement with Cvetanović's data (65%)[129] with a calculated

contribution of 0.4% for direct and consecutive allylic abstraction channels each. These results may also prove the accuracy of our quantum chemical results. It is also worthy to note that cTST overall rate constant obtained from UCCSD(T)/CBS//CASPT2/cc-pVTZ values (2.26×10^{-11} $cm^3molecule^{-1}s^{-1}$) is closer to Atkinson's experimental one, whereas the calculated branching ratio for **T** (57.8%) is in a less good agreement with the above data (Cvetanović).

In their recent paper, Zhou et. al.[127] have explored the propene + OH potential energy surface with projected MP2 methodology using the PMP2/aug-cc-pVQZ//MP2/cc-pVTZ level of theory, and CCSD(T) methodology at the same geometries using the 6-311+G(3df,2p) basis, and extrapolation from cc-pVDZ and cc-pVTZ bases. Their aim was to give an overall kinetic description at a broad temperature range, whereas our report focuses on species relevant around room temperature. In at least some of their cases, Zhou finds that PMP2 results are closer to experimental values than CCSD(T) ones. It might be justified to choose their PMP2 methodology over an elaborate CCSD(T) optimization with so many species examined. However, the inclusion of higher excitations are known to be important with radical transition states,[92,93] which is particularly true for the studied system as we pointed out in the previous discussion. The good results with the PMP2 methodology are probably due to a cancellation of errors,[125] and the inferior behaviour of CCSD(T) observed by these authors might be a problem of inadequate extrapolation, and simply the fact that CCSD(T) single points are not calculated at their optimized geometries. Both of these issues have been addressed here by using larger bases, using CCSD(T) optimized geometries and by comparing those results with ones from a wider choice of ab initio models. The authors were also able to reasonably reproduce the kinetic behaviour of the system based on weak collision master equation/microcanonical variational RRKM theory by lowering the barrier heights of central OH-addition (TS11) and terminal OH-addition (TS12) with 1 kcal/mol. However, in Figure 11. of Zhou's article, the branching ratio for these is around 50-50% at room temperature versus the experimental 65% preference for the terminal case, which is well predicted by our CCSD(T) model (65.8%). If cTST branching is calculated with their results it turns out to be 54% for the central case, indicating that the difference between their calculations and the ones presented here (and also the experimental results) is not due to the choice of the kinetic model, but to the fact that the optimized CCSD(T) results yield better and more

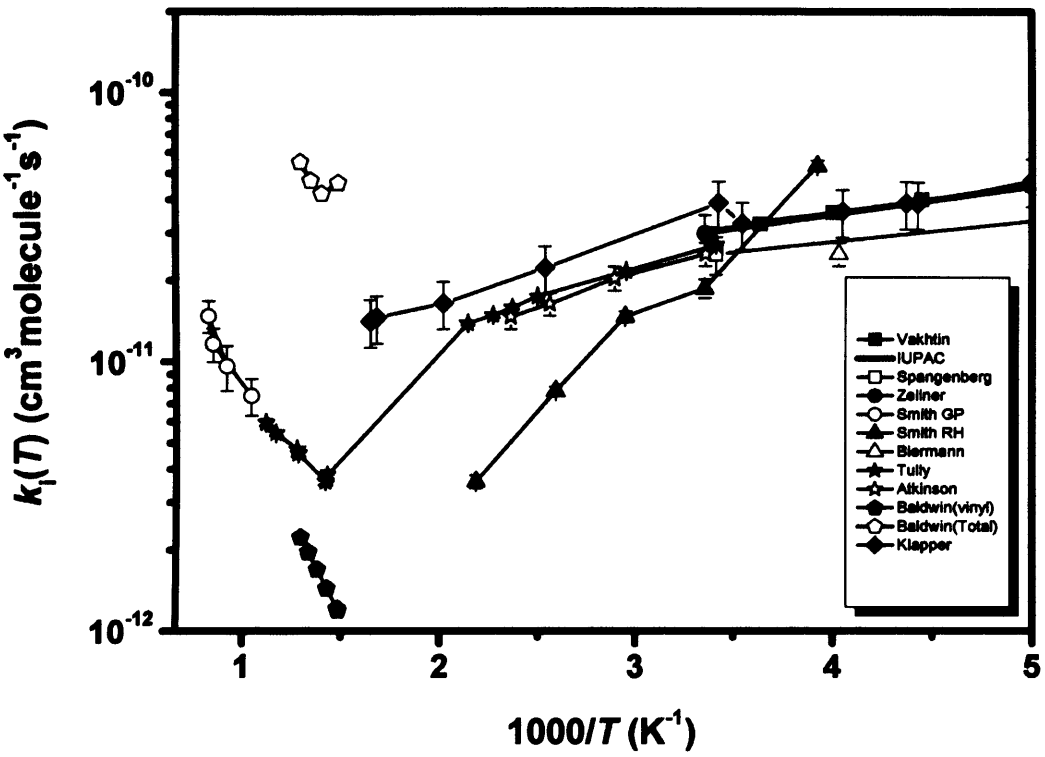

Figure 14: *Experimental measurements for the total rate constant of the Propene+OH reaction.*

consistent results compared to PMP2/aug-cc-pVQZ//MP2/cc-pVTZ.

Another recent paper, Huynh et al.[128] describes the kinetics of the enol formation based on CCSD(T)/cc-pVDZ//B3LYP/cc-pVTZ results. In our experience, results with cc-pVDZ basis are still roughly 10 kJ/mol away from the CBS limit. In addition, the pathological behaviour of the B3LYP in relevant cases is also known for a while.[122]

4.4 Some Further Kinetic Aspects

We have so far discussed the subsystem of the Propene+OH reaction that is depicted in Figure 9. These species were expected to be the ones governing the behaviour of the system at temperatures close to 298.15 K. In Figure 14, the temperature dependence of the total rate constant is shown as measured in different experiments. If we want to give

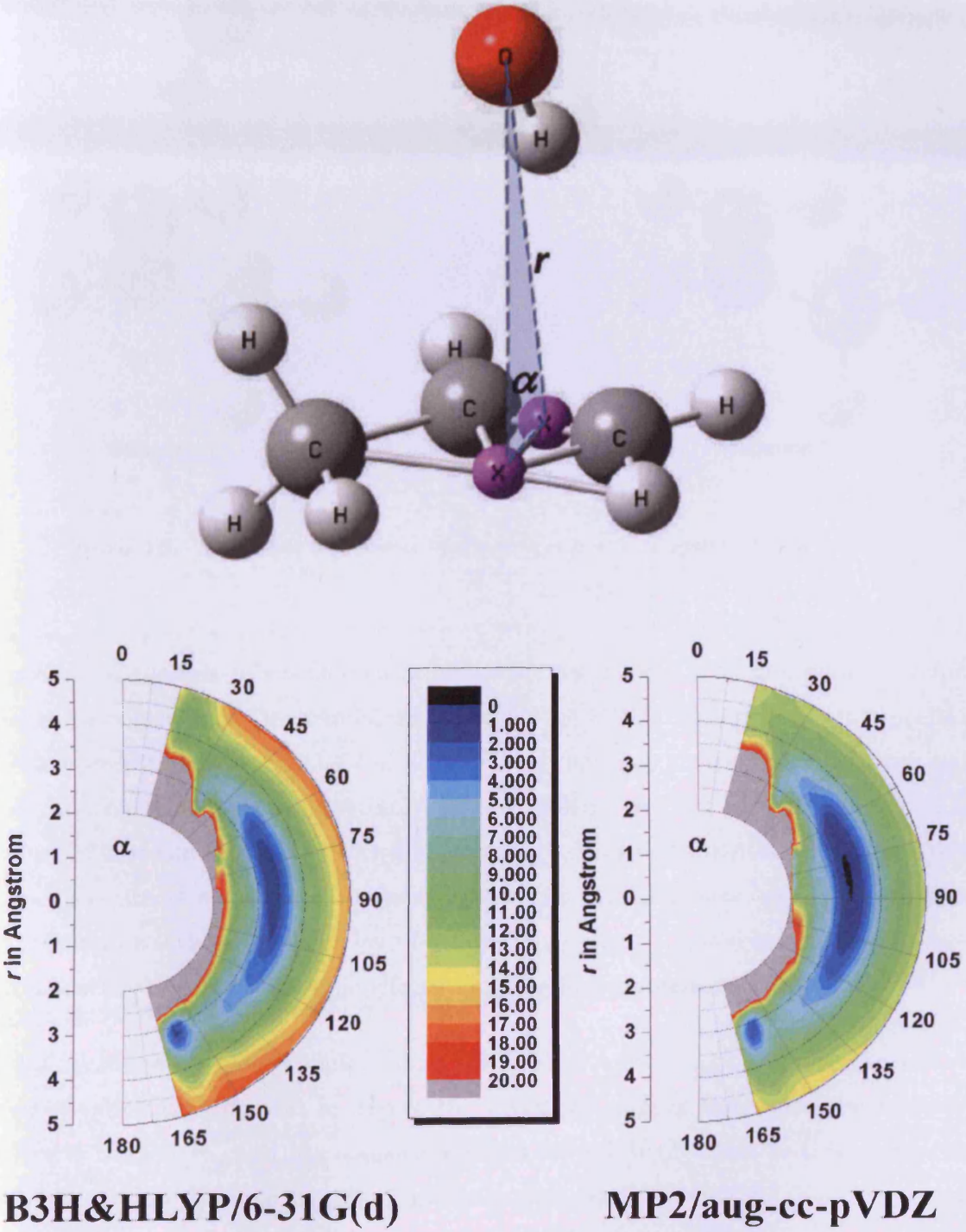

Figure 15: *The potential energy profile of the propene + OH system around a circle perpendicular to the axis of the double bond, as a function of the radius and angle at two different theoretical levels.*

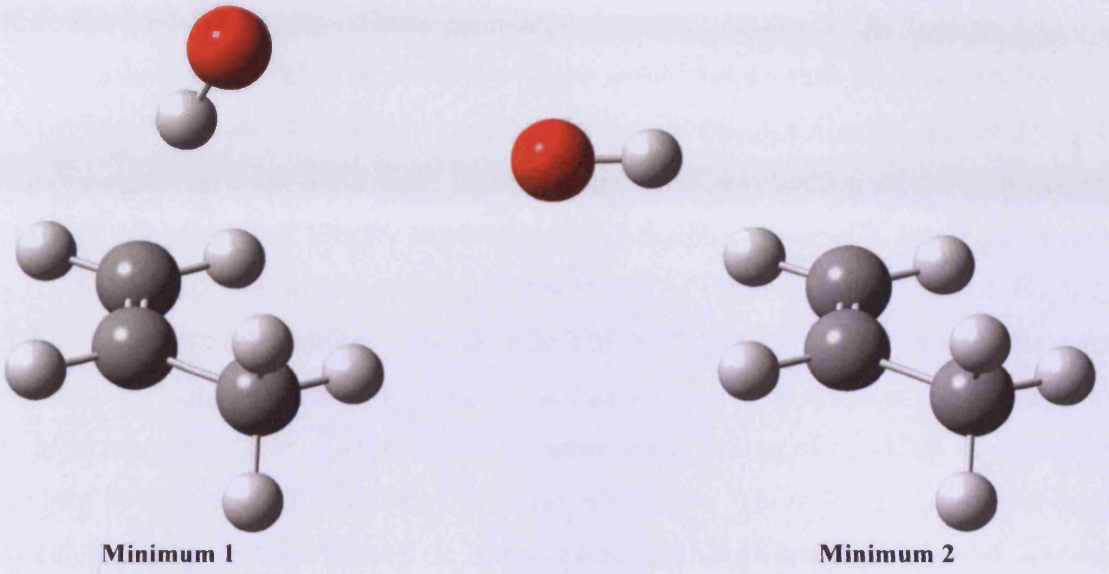

Figure 16: *Minimum structures of the propene + OH system found in Figure 15.*

an account of the system's behaviour at higher temperatures, then other possible channels need to be considered. One candidate would be the H-abstraction from vinyl positions. Another possibility is to consider consecutive steps that may have a major influence on the total rate constant, like the dissociation of the addition product into methyl radical and ethanol. These can be treated with the previously described methodology, although the active space structure should be reconsidered in the multireference cases, but this should not represent a serious problem. Possibly, other more sophisticated kinetic models should also be used, or perhaps dynamic effects could be investigated.

In line with these latter thoughts some preliminary calculations were carried out, the result of which is to be seen in Figure 15. Problems arise in the addition cases, where there is a π-complex, and the transition states are relatively close to this. These may easily necessitate dynamic treatment in the system. In Figure 15, we tried to gain some information about the potential energy surface, by rotating the OH group around the centroid of the double bond with various radii at the BH&HLYP/6-31G(d) and MP2/aug-cc-pVDZ levels of theory. There seem to be two minima on both of these potential energy surfaces, one probably corresponding to the π-complex. However, one should know more

about the variation of the energy with respect to other geometry parameters before further conclusions could be drawn. These calculations merely estimate the cost of a dynamic study, or, rather, what other possibilities there would be so that such a very expensive calculation could be avoided. Other possibilities would require a more precise knowledge of greater segments of the PES than just extrema. The application of the more complex RRKM and other related kinetic models is also possible, especially for the abstraction cases. In Figure 16, the minimum structures found in Figure 15 are shown. Both MP2 and DFT yield similar results, and so only the MP2 results are shown. Minimum 1 corresponds to the π-complex, and has a rather extensive minimum region around it, unlike Minimum 2, which might even be an error of the fitting of the PES as there are no grid points precisely at that very small minimum region where it can be found (a close grid point is shown in the Figure). It is also possible that it is a higher order stationary point that appears as a minimum in this partial scan of the energy, or even a possible artefact arising from the rigid treatment of almost all geometrical parameters during the scan since all of these were frozen except for the three describing the position of the H atom in OH. To reach any firm conclusions further calculations should be considered.

5 Summary

In the first part of the thesis, an attempt was made to arrive at a partitioning of correlation energy based on that of localized orbitals. The transferability of such partitions was examined for functional groups in different environments. Our results may be summarized in the followings:

1. Alkanes behave ideally with respect to transferability. It should be possible to predict the correlation energy belonging to larger species once that of the segments are known.

2. In predicting barrier height contributions, localized orbitals perform better than orbitals transformed to make reactant and transition state orbitals similar. This is because transformed orbitals delocalize and ruin partitioning.

3. The inclusion of oxygen as a heteroatom resulted in the decrease of transferability, due to electrostatic reasons and delocalization of lone pairs.

4. Species without lone pairs but with large electrostatic difference were examined. It turned out that electrostatic effects can ruin transferability if there is a large enough change in the molecules between which the transition is to be made; probably through decreasing/increasing densities even of groups with well localized orbitals.

5. Effects of π orbital delocalization were also considered. These may also be significant if partition boundaries are violated, but apparently electrostatic changes have a larger effect.

In the second half of the thesis, barrier heights for different possible reaction paths were calculated for the propene + OH system with the most accurate models available for

general use. The results could be summarized in the following points:

1. The use of single reference methods are sufficient and accurate in this case, and in fact they yield more accurate results than multireference methods due to computational limitations for the latter. On the other hand, the advantage of using a multireference CASPT2 in this case is that if active orbitals are carefully selected, it is able to approximate UCCSD(T) within 1-2 kJ/mol with considerably less computational effort.

2. The RHF-UCCSD(T)/CBS method is expected to yield the most accurate results. Triples contributions are substantial (typically around 10 kJ/mol for barriers). The restricted coupled cluster variant exhibits a slight difference to these results (around 2 kJ/mol), which is, however, unlikely to affect geometry optimizations.

3. G3MP2BH&H yields a result within 3-4 kJ/mol to the extrapolated UCCSD(T). As another way of approximating the complete basis set limit, the explicitly correlated UCCSD(T)-F12a model was utilized giving results within 0.7 kJ/mol maximum difference compared to extrapolated values.

4. Consecutive allylic abstraction and addition mechanisms go through a π-complex (**R**), which lies at -10.56 kJ/mol with respect to the enthalpy level of the infinite separation of the species.

5. The addition mechanisms have negative enthalpy barriers relative to infinite separation (-9.93 kJ/mol for **C** and -9.84 kJ/mol for **T**). There is only a marginal 0.09 kJ/mol energetic difference between the two.

6. The allylic abstraction mechanisms have slightly positive enthalpy barriers relative to infinite separation (3.21 kJ/mol for $\mathbf{A_{dir}}$ and 1.67 kJ/mol for $\mathbf{A_{con}}$), with the consecutive mechanism favoured by 1.54 kJ/mol. Although they have significantly higher barriers, they may contribute to the overall reaction system at higher temperatures.

7. Using conventional transition state theory, our UCCSD(T) results were able to reproduce the experimental overall high pressure rate constant within a factor of two. Calculated branching ratios show the preference of **T** (65.8%) in good agreement

with experiment. Allylic abstraction channels have a small contribution of 0.4% each. UCCSD(T)/CBS//CASPT2/cc-pVTZ values show similar good agreement supporting its use as an alternative to more expensive methods.

8. For higher 1-alkene homologues, where UCCSD(T) becomes too demanding to compute, a CASPT2 with similar active space structure may still be an option. Another possibility is to use the G3MP2BH&H method, which is found to be somewhat less accurate compared to CASPT2, but does not require constructing an active space.

References

[1] E. Hückel, *Z. Phys.*, 1931, **70**, 204.

[2] T. Helgaker, W. Klopper and D. P. Tew, *Mol. Phys.*, 2008, **106**, 2107.

[3] R. Izsák, M. Szőri, P. J. Knowles and B. Viskolcz, *J. Chem. Theory Comput.*, 2009, **5**, 2313.

[4] M. R. McGillen, C. J. Percival, D. E. Shallcross and J. N. Harvey, *Phys. Chem. Chem. Phys.*, 2007, **9**, 4349.

[5] R. X. Fernandes, H. Hippler and M. Olzmann, *Proc. Combust. Inst.*, 2005, **30**, 1033.

[6] S. S. Brown, *Chem. Rev.*, 2003, **103**, 5219.

[7] A. H. Zewail and R. B. Bernstein, *Chem. Eng. News*, 1988, **66**, 24.

[8] I. M. Waller, T. N. Kitsopoulos and D. M. Neumark, *J. Phys. Chem.*, 1990, **94**, 2240.

[9] M. J. Kurylo and V. L. Orkin, *Chem. Rev.*, 2003, **103**, 5049.

[10] R. Atkinson and J. Arey, *Chem. Rev.*, 2003, **103**, 4605.

[11] J. J. Orlando, G. S. Tyndall and T. J. Wallington, *Chem. Rev.*, 2003, **103**, 4657.

[12] D. A. Good and J. S. Francisco, *Chem. Rev.*, 2003, **103**, 4999.

[13] A. B. Vakhtin, J. E. Murphy and S. R. Leone, *J. Phys. Chem. A*, 2003, **107**, 10055.

[14] L. K. Huynh, K. Barriger and A. Violi, *J. Phys. Chem. A*, 2008, **112**, 1436.

[15] C. Brackholtz, T. A. Brackholtz and C. M. Hadad, *J. Phys. Chem. A*, 2001, **105**, 140.

[16] S. S. Rajindar and R. Weindruch, *Science*, 1996, **273**, 59.

[17] M. Frenette and J. C. Scaiano, *J. Am. Chem. Soc.*, 2008, **130**, 9634.

[18] J. A. Gerlt, in *Hydrogen-Transfer Reactions, Vol 4.*, 2007, 1107.

[19] R. Izsak, B. Jojart, I. G. Csizmadia and B. Viskolcz, *J. Chem. Inf. Model.*, 2006, **46**, 2527.

[20] M. Born and J. R. Oppenheimer, *Ann. Phys.*, 1927, **84**, 457.

[21] M. Born and K. Huang, *Dynamical Theory of Crystal Lattices*, OUP, Oxford, 1954.

[22] J. C. Slater, *Phys. Rev.*, 1929, **34**, 1293.

[23] D. R. Hartree, *Proc. Cambridge Phil. Soc.*, 1928, **24**, 89.

[24] V. Fock, *Z. Phys.*, 1930, **61**, 126.

[25] D. R. Hartree, *Molecular Spectra and Molecular Structure. IV. Constants of Diatomic Molecules*, Wiley, New York, 1957.

[26] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry*, Dover Publications, Inc., New York, 1996, p. 120.

[27] C. C. J. Roothaan, *Rev. Mod. Phys.*, 1951, **23**, 69.

[28] G. G. Hall, *Proc. Roy. Soc.: (London), Ser. A*, 1951, **205**, 541.

[29] T. H. Dunning, *J. Chem. Phys.*, 1989, **90**, 1007.

[30] R. A. Kendall, T. H. Dunning and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796.

[31] D. E. Woon and T. H. Dunning, *J. Chem. Phys.*, 1993, **98**, 1358.

[32] D. Feller, *J. Chem. Phys.*, 1992, **96**, 6104.

[33] J. Kobus, L. Laaksonen and D. Sundholm, *Comp. Phys. Commun.*, 1996, **98**, 346.

[34] T. Helgaker, W. Klopper, H. Koch and J. Noga, *J. Chem. Phys.*, 1997, **106**, 9639.

[35] R. T. Pack and W. B. Brown, *J. Chem. Phys.*, 1966, **45**, 556.

[36] P. J. Knowles, M. Schütz and H.-J. Werner, in *Modern Methods and Algorithms of Quantum Chemistry, NIC Series, Vol. 1*, 2000, 69.

[37] J. A. Pople, M. Head-Gordon and K. Raghavachari, *J. Chem. Phys.*, 1987, **87**, 5968.

[38] C. Møller and M. S. Plesset, *Phys. Rev.*, 1934, **46**, 618.

[39] R. J. Bartlett and D. M. Silver, *Int. J. Quant. Chem. Symp.*, 1974, **8**, 271.

[40] J. S. Brinkley and J. A. Pople, *Int. J. Quant. Chem.*, 1975, **9**, 229.

[41] W. Meyer, *J. Chem. Phys.*, 1976, **64**, 2901.

[42] E. R. Davidson, *J. Comp. Phys.*, 1975, **17**, 87.

[43] B. O. Roos and P. E. M. Siegbahn, in *Methods of Electronic Structure Theory*, Plenum, New York, 1977.

[44] S. R. Langhoff and E. R. Davidson, *Int. J. Quant. Chem.*, 1974, **8**, 61.

[45] R. Ahlrichs, P. Scharf and C. Ehrhardt, *J. Chem. Phys.*, 1985, **82**, 890.

[46] J. A. Pople, M. Head-Gordon and K. Raghavachari, *J. Chem. Phys.*, 1987, **87**, 69.

[47] J. Cizek, *J. Chem. Phys.*, 1966, **45**, 4256.

[48] J. Paldus, J. Cizek and I. Shavitt, *Phys. Rev. A*, 1972, **5**, 50.

[49] J. Watts, J. Gauss and R. J. Bartlett, *J. Chem. Phys.*, 1993, **98**, 8718.

[50] P. J. Knowles, C. Hampel and H.-J. Werner, *J. Chem. Phys.*, 1993, **99**, 5219.

[51] R. Pauncz, *Spin Eigenfunctions, Construction and Use*, Plenum Press, 1979.

[52] J. Hinze, *J. Chem. Phys.*, 1973, **59**, 6424.

[53] H.-J. Werner and W. Meyer, *J. Chem. Phys.*, 1980, **73**, 382.

[54] H.-J. Werner and P. J. Knowles, *J. Chem. Phys.*, 1985, **82**, 5053.

[55] K. Ruedenberg, L. M. Cheung and S. T. Elbert, *Int. J. Quant. Chem.*, 1979, **16**, 1069.

[56] B. O. Roos, P. R. Taylor and P. E. M. Siegbahn, *Chem. Phys.*, 1980, **48**, 157.

[57] R. B. Murphy and R. P. Messmer, *Chem. Phys. Lett.*, 1991, **183**, 443.

[58] K. Andersson, P.-A. Malmqvist and B. O. Roos, *J. Chem. Phys.*, 1992, **96**, 1218.

[59] R. B. Murphy and R. P. Messmer, *Chem. Phys. Lett.*, 1991, **183**, 443.

[60] R. J. Buenker and S. D. Peyerimhoff, *Theor. Chim. Acta*, 1974, **35**, 33.

[61] P. E. M. Siegbahn, *Int. J. Quant. Chem.*, 1980, **18**, 1229.

[62] H. Lischka, R. Shepard, F. B. Brown and I. Shavitt, *Int. J. Quant. Chem. Symp.*, 1981, **15**, 91.

[63] H.-J. Werner and P. J. Knowles, *J. Chem. Phys.*, 1988, **89**, 5803.

[64] H.-J. Werner, M. Kállay and J. Gauss, *J. Chem. Phys.*, 2008, **128**, 034305.

[65] P. G. Szalay, *Chem. Phys.*, 2008, **349**, 121.

[66] D. Mukherjee, R. K. Moitra and A. Mukhopadhyay, *Mol. Phys.*, 1977, **33**, 955.

[67] M. Kallay, P. G. Szalay and P. R. Surjan, *J. Chem. Phys.*, 2002, **117**, 980.

[68] J. J. P. Stewart, in *Reviews in Computational Chemistry, Vol. 1*, 1990, 45.

[69] N. L. Allinger and U. Burkert, *Molecular Mechanics*, ACS Monograph 177, American Chemical Society, Washington, DC, 1982.

[70] P. Hohenberg and W. Kohn, *Phys. Rev.*, 1964, **136**, B864.

[71] W. Kohn and L. J. Sham, *Phys. Rev.*, 1965, **140**, A1133.

[72] I. N. Levine, *Quantum Chemistry*, Prentice Hall, Upper Saddle River, New Jersey, 2000, p. 573.

[73] R. G. Parr and W. Yang, *Density Functional Theory of Atoms and Molecules*, OUP, New York, 1989, p. 150.

[74] A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648.

[75] A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372.

[76] J. A. Pople, M. Head-Gordon, D. J. Fox and K. Raghavachari, *J. Chem. Phys.*, 1989, **90**, 5622.

[77] J. M. L. Martib and G. de Oliveira, *J. Chem. Phys.*, 1999, **111**, 1843.

[78] J. W. Ochterski, G. A. Peterson and J. A. Montgomery, Jr., *J. Chem. Phys.*, 1996, **104**, 2598.

[79] A. Tajti, P. G. Szalay, A. G. Császár, M. Kállay, J. Gauss, E. F. Valeev, B. A. Flowers, J. Vázquez and J. F. Stanton, *J. Chem. Phys.*, 2004, **121**, 11599.

[80] A. G. Baboul, L. A. Curtiss, P. C. Redfern and K. Raghavachari, *J. Chem. Phys.*, 1999, **110**, 7650.

[81] G. C. Schatz, D. Sokolovski and J. N. L. Connor, *Advances in Molecular Vibrations and Collision Dynamics*, JAI Press, Greenwich, CT, 1994, vol. 2B, pp. 1–26.

[82] D. K. Bondi, J. N. L. Connor, J. Manz and J. Römelt, *Mol. Phys.*, 1983, **50**, 467.

[83] A. J. Dobbyn, J. N. L. Connor, N. A. Besley, P. J. Knowles and G. C. Schatz, *Phys. Chem. Chem. Phys.*, 1999, **1**, 957.

[84] G. L. Fox and H. B. Schlegel, *J. Am. Chem. Soc.*, 1993, **115**, 6870.

[85] K. Luth and S. Scheiner, *Int. J. Quant. Chem. Chem. Symp.*, 1992, **26**, 817.

[86] A. A. Zavitsas and C. Chatgilialoglu, *J. Am. Chem. Soc.*, 1995, **117**, 10645.

[87] C. Isborn, D. A. Hrovat, W. T. Borden, J. M. Mayer and B. K. Carpenter, *J. Am. Chem. Soc.*, 2005, **127**, 5794.

[88] G. Chałasiński and M. M. Szczęśniak, *Chem. Rev.*, 2000, **100**, 4227.

[89] F. van Duijneveldt, G. C. J. M. van Duijneveldt-van de Rijdt and J. H. van Lenthe, *Chem. Rev.*, 1994, **94**, 1873.

[90] F. G. Kalatzis and I. N. Demetropoulos, *Mol. Phys.*, 2007, **105**, 2335.

[91] M. S. G. LeRoy and A. Tinant, *Can. J. Chem.*, 1985, **63**, 1447.

[92] K. A. Peterson and T. H. Dunning, *J. Phys. Chem. A*, 1997, **101**, 6280.

[93] Y.-Y. Chuang, E. L. Coitio and D. G. Truhlar, *J. Phys. Chem. A*, 2000, **104**, 446.

[94] P. G. Szalay, J. Vazquez, C. Simmons and J. F. Stanton, *J. Chem. Phys.*, 2004, **121**, 7624.

[95] S.F. Boys, in *Quantum Theory of Atoms, Molecules, and the Solid State*, 1966, 253.

[96] C. Edmiston and K. Ruedenberg, in *Quantum Theory of Atoms, Molecules, and the Solid State*, 1966, 263.

[97] J. Pipek and P. G. Mezey, *J. Chem. Phys.*, 1989, **90**, 4916.

[98] J. W. Boughton and P. Pulay, *J. Comput. Chem.*, 1993, **14**, 736.

[99] P. Pulay, *Chem. Phys. Lett.*, 1983, **100**, 151.

[100] S. Saebø and P. Pulay, *Chem. Phys. Lett.*, 1985, **113**, 13.

[101] C. Hampel, H.-J. Werner, *J. Chem. Phys.*, 1996, **104**, 6286.

[102] S. Humbel, S. Sieber and K. Morokuma, *J. Chem. Phys.*, 1996, **105**, 1959.

[103] P. Celani, H. Stoll and H.-J. Werner, *Mol. Phys.*, 2004, **102**, 2369.

[104] F. Jensen, *Introduction to Computational Chemistry*, John Wiley & Sons Ltd, 2007, p. 304.

[105] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory*, John Wiley & Sons Ltd, 2000, p. 62.

[106] L. Bytautas and K. Ruedenberg, *Chem. Phys.*, 2009, **356**, 64.

[107] P.-O. Löwdin, *Adv. Quant. Chem.*, 1970, **5**, 185.

[108] I. Mayer, *Int. J. Quant. Chem.*, 2002, **90**, 63.

[109] D.C. Carlson, J.H. Keller, *Phys. Rev.*, 1955, **105**, 102.

[110] J.G. Aiken, J.A. Erdos, J.A. Goldstein, *Int. J. Quant. Chem.*, 1980, **18**, 1101.

[111] F.X. Gadéa, M. Pélissier, *J. Chem. Phys.*, 1990, **93**, 545.

[112] W. Domcke, C. Woywod, *Chem. Phys. Lett.*, 1993, **216**, 362.

[113] T. Pacher, L.S. Cederbaum, H. Köppel, *J. Chem. Phys.*, 1988, **89**, 7367.

[114] H.-J. Werner and B. Follmeg and M. H. Alexander, *J. Chem. Phys.*, 1988, **89**, 3139.

[115] H.-J. Werner, P. J. Knowles, R. Lindh, F. R. Manby, M. Schütz, P. Celani, T. Korona, A. Mitrushenkov, G. Rauhut, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, G. Hetzer, T. Hrenar, G. Knizia, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, P. Palmieri, K. Pflüger, R. Pitzer, M. Reiher, U. Schumann, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang and A. Wolf, *MOLPRO, version 2008.1, a package of ab initio programs*, 2008, see http://www.molpro.net.

[116] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, J. A. Montgomery, Jr., T. Vreven, K. N. Kudin, J. C. Burant, J. M. Millam, S. S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, 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, M. Klene, X. Li, J. E. Knox, H. P. Hratchian, 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, P. M. W. Gill, B. Johnson, W. Chen, M. W. Wong, C. Gonzalez and J. A. Pople, *Gaussian 03, Revision C.02*, Gaussian, Inc., Wallingford, CT, 2004.

[117] I. Mayer, *Chem. Phys. Lett.*, 1995, **242**, 499.

[118] C. A. Taatjes, N. Hansen, A. McIlroy, J. A. Miller, J. P. Senosiain, S. J. Klippenstein, F. Qi, L. S. Sheng, Y. M. Zhang, T. A. Cool, J. Wang, P. R. Westmoreland, M. E. Law, T. Kasper and K. Kohse-Hoinghaus, *Science*, 2005, **208**, 1887.

[119] J. A. Miller, M. J. Pilling and J. Troe, *Proc. Combust. Inst.*, 2005, **30**, 43.

[120] D. Poppe, T. Brauers, H.-P. Dorn, M. Karl, T. Mentel, E. Schlosser, R. Tillmann, R. Wegener and A. Wahner, *J. Atmos. Chem.*, 2007, **57**, 203.

[121] C. A. Taatjes, N. Hansen, J. A. Miller, T. A. Cool, J. Wang, P. R. Westmoreland, M. E. Law, T. Kasper and K. Kohse-Hinghaus, *J. Phys. Chem. A*, 2006, **110**, 3254.

[122] M. Szőri, C. Fittschen, I. G. Csizmadia and B. Viskolcz, *J. Chem. Theory Comp.*, 2006, **2**, 1575.

[123] J. R. Alvarez-Idaboy, I. Diaz-Acosta and J. R. Vivier-Bunge, *J. Comp. Chem.*, 1998, **8**, 811.

[124] I. Diaz-Acosta, J. R. Alvarez-Idaboy and J. Vivier-Bunge, *Int. J. Chem. Kin.*, 1999, **31**, 29.

[125] J. R. Alvarez-Idaboy, N. Mora-Diez and A. Vivier-Bunge, *J. Am. Chem. Soc.*, 2000, **122**, 3715.

[126] C. Selcuki and V. Aviyente, *J. Mol. Model.*, 2001, **11**, 398.

[127] C.-W. Zhou, Z.-R. Li and X.-Y. Li, *J. Phys. Chem. A*, 2009, **113**, 2372.

[128] L. K. Huynh, H. R. Zhang, S. Zhang, E. Eddings, A. Sarofim, M. E. Law, P. R. Westmoreland and T. N. Truong, *J. Phys. Chem. A*, 2009, **113**, 3177.

[129] R. J. Cvetanovic, *12th International Symposium on Free Radicals*, 1976, Laguna Beach, CA.

[130] O. Tishchenko, J. Zheng and D. G. Truhlar, *J. Chem. Theory Comput.*, 2008, **4**, 1208.

[131] H.-J. Werner, *Mol. Phys.*, 1996, **89**, 645.

[132] T. J. Lee, A. P. Rendell and P. R. Taylor, *J. Phys. Chem.*, 1990, **94**, 5463.

[133] M. J. O. Deegan and P. J. Knowles, *Chem. Phys. Lett.*, 1994, **227**, 321.

[134] T. B. Adler, G. Knizia and H.-J. Werner, *J. Chem. Phys.*, 2007, **127**, 221106.

[135] J. Olsen, B. O. Roos, P. Jørgensen and H. J. A. Jensen, *J. Chem. Phys.*, 1988, **89**, 2185.

[136] M. Kállay, *MRCC, a string-based quantum chemical program suite*, see also M. Kállay, P. R. Surján, J. Chem. Phys., 115, 2945 (2001) as well as: www.mrcc.hu.

[137] B. Silvi, P. Labarbe and J. P. Perchard, *Spectrochim. Acta A*, 1973, **29**, 263.

[138] K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure. IV. Constants of Diatomic Molecules*, Van Nostrand Reinhold Co., 1979.

[139] NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101 Release 14, Sept 2006, Editor: Russell D. Johnson III http://srdata.nist.gov/cccbdb.

[140] M. Szőri, T. Abou-Abdo, C. Fittschen, I. G. Csizmadia and B. Viskolcz, *Phys. Chem. Chem. Phys.*, 2007, **9**, 1931.

[141] R. Atkinson, D. L. Baulch, R. A. Cox, J. N. Crowley, R. F. Hampson, R. G. Hynes, M. E. Jenkin, J. A. Kerr, M. J. Rossi and J. Troe, Summary of Evaluated Kinetic and Photochemical Data for Atmospheric Chemistry IUPAC Subcommittee on Gas Kinetic Data Evaluation for Atmospheric Chemistry Web Version February 2006 (http://www.iupac-kinetic.ch.cam.ac.uk.

[142] R. Zellner and K. Lorenz, *J. Phys. Chem.*, 1984, **88**, 984.

[143] R. Atkinson, *Chem. Rev.*, 1985, **85**, 69.

Appendix

Table 10: *Partitioning of symmetric species shown in Figure 8. All results are shown in Hartree units.*

	FHF	HOHOH	HOOHOOH	H_2NHNH_2	$H_2NOHONH_2$
			Reactants		
E_r(RHF)	-199.410885	-151.441423	-301.002063	-111.776386	-261.428251
E_r(A-A)	-0.129195	-0.136719	-0.244915	-0.134227	-0.234235
E_r(A-B)	0.000000	0.000000	0.000000	0.000000	0.000000
E_r(A-R)	-0.069082	-0.062349	-0.125695	-0.052187	-0.119460
E_r(B-B)	-0.124420	-0.134108	-0.241777	-0.132237	-0.217477
E_r(B-R)	-0.046117	-0.046304	-0.121503	-0.044038	-0.130400
E_r(R-R)	-0.026822	-0.026851	-0.076635	-0.027215	-0.084471
$E_{c,r}$	-0.395636	-0.406332	-0.810524	-0.389906	-0.786042
			Transition States		
E_{ts}(RHF)	-199.321733	-151.375283	-300.929289	-111.715952	-261.359022
E_{ts}(A-A)	-0.120080	-0.128648	-0.231946	-0.127575	-0.219166
E_{ts}(A-B)	-0.008424	-0.008741	-0.009485	-0.007913	-0.007472
E_{ts}(A-R)	-0.077573	-0.071463	-0.143718	-0.061899	-0.141175
E_{ts}(B-B)	-0.120407	-0.128665	-0.232161	-0.127573	-0.219168
E_{ts}(B-R)	-0.077735	-0.071480	-0.143663	-0.061887	-0.141189
E_{ts}(R-R)	-0.044436	-0.041081	-0.091680	-0.039244	-0.097588
$E_{c,ts}$	-0.448656	-0.450078	-0.852653	-0.426092	-0.825758
			Barriers		
ΔE(RHF)	0.089152	0.066140	0.072774	0.060435	0.069229
ΔE(A-A)	0.009115	0.008071	0.012970	0.006652	0.015068
ΔE(A-B)	-0.008424	-0.008741	-0.009485	-0.007913	-0.007472
ΔE(A-R)	-0.008491	-0.009114	-0.018024	-0.009712	-0.021715
ΔE(B-B)	0.004013	0.005443	0.009615	0.004664	-0.001691
ΔE(B-R)	-0.031619	-0.025176	-0.022160	-0.017849	-0.010789
ΔE(R-R)	-0.017614	-0.014229	-0.015046	-0.012029	-0.013117
ΔE_c	-0.053020	-0.043746	-0.042129	-0.036186	-0.039716
Total	0.036131	0.022393	0.030644	0.024249	0.029514

Table 11: *Cartesian coordinates of direct ($\mathbf{A_{dir}}$) and consecutive ($\mathbf{A_{con}}$) allylic H-abstraction transition states in Å units at the [5,5]-CASPT2/cc-pVTZ level of theory.*

species	atom	X	Y	Z
$\mathbf{A_{dir}}$	C	0.000000	0.000000	0.000000
	C	0.000000	0.000000	1.339786
	H	0.921504	0.000000	-0.561946
	H	-0.925798	0.002005	-0.558230
	H	0.945032	-0.008992	1.868036
	C	-1.231435	-0.017968	2.167978
	H	-2.136291	0.087293	1.575405
	H	-1.213540	0.721840	2.965344
	H	-1.297689	-1.043792	2.690504
	O	-1.489035	-2.231243	3.549423
	H	-2.010559	-1.765867	4.221649
$\mathbf{A_{con}}$	C	0.000000	0.000000	0.000000
	C	0.000000	0.000000	1.340398
	H	0.921005	0.000000	-0.562935
	H	-0.926334	-0.010257	-0.557713
	H	0.946157	0.009198	1.867201
	C	-1.229729	0.024126	2.170413
	H	-1.200049	-0.680810	2.998271
	H	-2.135964	-0.109444	1.586375
	H	-1.321463	1.054900	2.688663
	O	-1.255390	2.488923	2.969937
	H	-0.873475	2.693654	2.101744

Table 12: *Cartesian coordinates of central (**C**) and terminal (**T**) OH-addition transition states in Å units at the [5,5]-CASPT2/cc-pVTZ level of theory.*

species	atom	X	Y	Z
C	C	0.000000	0.000000	0.000000
	H	0.000000	0.000000	1.081429
	C	1.184456	0.000000	-0.646875
	H	2.117482	-0.064215	-0.108097
	H	1.226882	0.050541	-1.726531
	C	-1.240395	-0.041119	-0.833230
	H	-1.004891	-0.257270	-1.873253
	H	-1.924121	-0.801239	-0.464812
	H	-1.757943	0.923958	-0.791981
	O	0.185220	-2.226050	0.582902
	H	0.798192	-2.464609	-0.128190
T	C	0.000000	0.000000	0.000000
	H	0.000000	0.000000	1.083983
	C	1.179838	0.000000	-0.649534
	H	2.115620	0.045134	-0.114575
	H	1.220105	0.049464	-1.727131
	C	-1.328476	-0.032628	-0.683708
	H	-1.207449	-0.046603	-1.764055
	H	-1.897879	-0.915303	-0.393481
	H	-1.930128	0.841342	-0.410336
	O	1.207245	-2.305627	-1.018281
	H	0.844502	-2.486421	-0.138893

Table 13: *Cartesian coordinates of the π-complex (**R**), and the initial propene and OH geometries in Å units at the [5,5]-CASPT2/cc-pVTZ level of theory. Although the supermolecular approach was used, here the reactant structures are given separately.*

species	atom	X	Y	Z
R	C	0.000000	0.000000	0.000000
	C	0.000000	0.000000	1.339986
	H	0.921173	0.000000	-0.563366
	H	-0.926226	-0.014377	-0.558680
	H	0.950679	0.006784	1.860438
	C	-1.232008	-0.002973	2.187891
	H	-1.246356	-0.872446	2.844049
	H	-2.130195	-0.019191	1.574509
	H	-1.264076	0.886422	2.825628
	O	-0.397375	3.244879	1.020974
	H	-0.213382	2.326495	0.756719
Propene	C	0.000000	0.000000	0.000000
	C	0.000000	0.000000	1.337450
	H	0.920829	0.000000	-0.563337
	H	-0.925964	0.001876	-0.558392
	H	0.949030	-0.000778	1.860099
	C	-1.234686	0.001518	2.181955
	H	-1.263598	-0.871498	2.833575
	H	-2.130209	-0.001506	1.564441
	H	-1.265101	0.884856	2.828867
OH	O	0.000000	0.000000	0.000000
	H	0.000000	0.000000	0.968441

Table 14: *Cartesian coordinates of direct ($\mathbf{A_{dir}}$) and consecutive ($\mathbf{A_{con}}$) allylic H-abstraction transition states in Å units at the UCCSD(T)/cc-pVTZ level of theory.*

species	atom	X	Y	Z
$\mathbf{A_{dir}}$	C	0.000000	0.000000	0.000000
	C	0.000000	0.000000	1.341394
	H	0.924097	0.000000	-0.563968
	H	-0.928366	0.004675	-0.560600
	H	0.946779	-0.013112	1.872627
	C	-1.233614	-0.014517	2.168414
	H	-2.142878	0.128413	1.584459
	H	-1.197723	0.679826	3.009093
	H	-1.326647	-1.082104	2.650589
	O	-1.523695	-2.254435	3.405062
	H	-1.949722	-1.816939	4.159556
$\mathbf{A_{con}}$	C	0.000000	0.000000	0.000000
	C	0.000000	0.000000	1.341818
	H	0.923548	0.000000	-0.564949
	H	-0.928815	-0.011762	-0.560167
	H	0.947536	0.011887	1.872006
	C	-1.232167	0.019071	2.170687
	H	-1.193408	-0.656117	3.026578
	H	-2.143159	-0.130208	1.591719
	H	-1.337540	1.078761	2.676763
	O	-1.340228	2.437768	2.996473
	H	-1.007105	2.713772	2.127046

Table 15: *Cartesian coordinates of central (**C**) and terminal (**T**) OH-addition transition states in Å units at the UCCSD(T)/cc-pVTZ level of theory.*

species	atom	X	Y	Z
C	C	0.000000	0.000000	0.000000
	H	0.000000	0.000000	1.083757
	C	1.190464	0.000000	-0.648690
	H	2.123302	-0.098434	-0.109845
	H	1.233998	0.065080	-1.731077
	C	-1.246367	-0.034201	-0.834975
	H	-1.019387	-0.314968	-1.865393
	H	-1.959599	-0.749275	-0.425286
	H	-1.721935	0.950674	-0.845805
	O	0.238751	-2.102407	0.554128
	H	0.561400	-2.409131	-0.307976
T	C	0.000000	0.000000	0.000000
	H	0.000000	0.000000	1.087152
	C	1.185405	0.000000	-0.651886
	H	2.122605	0.071828	-0.116281
	H	1.226351	0.063763	-1.731335
	C	-1.331305	-0.056909	-0.685201
	H	-1.209129	-0.064167	-1.768901
	H	-1.876693	-0.959637	-0.398018
	H	-1.951159	0.799890	-0.406395
	O	1.158212	-2.189266	-0.910212
	H	1.059361	-2.374064	0.036695

Table 16: *Cartesian coordinates of the π-complex (**R**), and the initial propene and OH geometries in Å units at the UCCSD(T)/cc-pVTZ level of theory. Although the supermolecular approach was used, here the reactant structures are given separately.*

species	atom	X	Y	Z
R	C	0.000000	0.000000	0.000000
	C	0.000000	0.000000	1.340520
	H	0.923618	0.000000	-0.565407
	H	-0.928692	-0.015533	-0.561143
	H	0.952982	0.007407	1.863247
	C	-1.235980	-0.002945	2.193698
	H	-1.244748	-0.869988	2.858779
	H	-2.136892	-0.030024	1.578246
	H	-1.272976	0.890771	2.821664
	O	-0.439546	3.256058	1.028234
	H	-0.243768	2.342278	0.750565
Propene	C	0.000000	0.000000	0.000000
	C	0.000000	0.000000	1.338046
	H	0.923527	0.000000	-0.565162
	H	-0.928282	0.001427	-0.561212
	H	0.951542	-0.002100	1.862874
	C	-1.238315	0.003137	2.188068
	H	-1.267906	-0.874284	2.839733
	H	-2.137098	0.002112	1.568848
	H	-1.266698	0.884124	2.835031
OH	O	0.000000	0.000000	0.000000
	H	0.000000	0.000000	0.971116

Table 17: *RHF reference and UCCSD(T) energies used for extrapolation on [5,5]-CASPT2/cc-pVTZ optimized geometries. An is a shorthand notation for aug-cc-pVnZ bases.*

level	$\mathbf{A_{dir}}$	$\mathbf{A_{con}}$	**C**	**T**	**R**	*Propene + OH*
RHF/A2	-192.458085	-192.457718	-192.473409	-192.474049	-192.488437	-192.485824
RHF/A3	-192.505602	-192.505138	-192.520514	-192.521316	-192.536033	-192.533928
RHF/A4	-192.517379	-192.516906	-192.532233	-192.533072	-192.547795	-192.545825
RHF/CBS	-192.521260	-192.520790	-192.536114	-192.536965	-192.551655	-192.549735
UCCSD(T)/A2	-193.157166	-193.157683	-193.167071	-193.166496	-193.167771	-193.161120
UCCSD(T)/A3	-193.331503	-193.332125	-193.340747	-193.340461	-193.341834	-193.335777
UCCSD(T)/CBS	-193.400559	-193.401261	-193.409639	-193.409457	-193.410706	-193.404869

Table 18: *RHF reference and UCCSD(T) energies used for extrapolation on UCCSD(T)/cc-pVTZ optimized geometries. An is a shorthand notation for aug-cc-pVnZ bases.*

level	$\mathbf{A_{dir}}$	$\mathbf{A_{con}}$	**C**	**T**	**R**	*Propene + OH*
RHF/A2	-192.451343	-192.451174	-192.466494	-192.467221	-192.488514	-192.485757
RHF/A3	-192.498578	-192.498321	-192.513193	-192.514136	-192.535914	-192.533645
RHF/A4	-192.510323	-192.510056	-192.524863	-192.525842	-192.547663	-192.545547
RHF/CBS	-192.514210	-192.513946	-192.528751	-192.529734	-192.551535	-192.549484
UCCSD(T)/A2	-193.157523	-193.158027	-193.167696	-193.167320	-193.168349	-193.161626
UCCSD(T)/A3	-193.331533	-193.332138	-193.341019	-193.340980	-193.341985	-193.335874
UCCSD(T)/CBS	-193.400544	-193.401222	-193.409893	-193.409943	-193.410757	-193.404918

Table 19: *Thermochemical data for all species calculated from [5,5]-CASPT2/cc-pVTZ geometries and frequencies scaled by a factor of 0.958. The UCCSD(T)/CBS total energies (E_{tot}), zero point vibrational energies ($ZPVE$) and thermal corrections at $T = 298.15K$ for energies ($E_{therm}(T)$), enthalpies ($H_{therm}(T)$) and free energies ($G_{therm}(T)$) are in Hartree units. Energies relative to the separate Propene+OH (E_{rel}) are in kJ/mol, and entropies ($S(T)$) are in J/molK units.*

species	E_{tot}	E_{rel}	$ZPVE$	$E_{therm}(T)$	$H_{therm}(T)$	$G_{therm}(T)$	S(T)
$\mathbf{A_{dir}}$	-193.400559	11.32	0.083565	0.090089	0.091033	0.052406	340.15
$\mathbf{A_{con}}$	-193.401261	9.47	0.083831	0.090174	0.091119	0.053491	331.34
C	-193.409639	-12.52	0.087122	0.093486	0.094430	0.057408	326.02
T	-193.409447	-12.02	0.087089	0.093553	0.094498	0.056911	330.98
R	-193.410706	-15.32	0.087042	0.094302	0.095246	0.055330	351.50
$Propene + OH$	-193.404869	0.00	0.085023	0.091589	0.093478	0.042364	450.11

Table 20: *Thermochemical data for all species calculated from UCCSD(T)/cc-pVTZ geometries and frequencies scaled by a factor of 0.975. The UCCSD(T)/CBS total energies (E_{tot}), zero point vibrational energies ($ZPVE$) and thermal corrections at $T = 298.15K$ for energies ($E_{therm}(T)$), enthalpies ($H_{therm}(T)$) and free energies ($G_{therm}(T)$) are in Hartree units. Energies relative to the separate Propene+OH (E_{rel}) are in kJ/mol, and entropies ($S(T)$) are in J/molK units.*

species	E_{tot}	E_{rel}	$ZPVE$	$E_{therm}(T)$	$H_{therm}(T)$	$G_{therm}(T)$	S(T)
$\mathbf{A_{dir}}$	-193.400544	11.48	0.083787	0.090136	0.091080	0.053391	331.89
$\mathbf{A_{con}}$	-193.401222	9.70	0.083953	0.090228	0.091172	0.053761	329.44
C	-193.409893	-13.06	0.088374	0.094477	0.095421	0.059176	319.18
T	-193.409943	-13.19	0.088305	0.094559	0.095503	0.058586	325.09
R	-193.410757	-15.33	0.087857	0.095100	0.096045	0.056152	351.30
$Propene + OH$	-193.404918	0.00	0.085815	0.092344	0.094232	0.043166	449.68

Table 21: *Unscaled vibrational frequencies and rotational constants for the various species at the [5,5]-CASPT2/cc-pVTZ level of theory.*

species	B (cm^{-1})	vibrational frequencies (cm^{-1})
$\mathbf{A_{dir}}$	0.598, 0.090, 0.085	663i, 44, 81, 111, 324, 418, 577, 722, 910, 945, 968, 1011, 1106, 1177, 1204, 1328, 1363, 1448, 1483, 1600, 1692, 3115, 3186, 3191, 3210, 3287, 3785
$\mathbf{A_{con}}$	0.448, 0.109, 0.097	870i, 63, 116, 166, 305, 423, 565, 773, 913, 941, 955, 990, 1018, 1197, 1321, 1338, 1383, 1456, 1492, 1535, 1691, 3119, 3185, 3194, 3209, 3286, 3776
C	0.283, 0.193, 0.131	217i, 105, 155, 167, 230, 421, 574, 582, 885, 938, 950, 1006, 1055, 1203, 1314, 1403, 1457, 1488, 1507, 1644, 3031, 3134, 3182, 3184, 3222, 3287, 3794
T	0.371, 0.146, 0.119	199i, 72, 135, 164, 219, 423, 559, 612, 922, 941, 947, 1004, 1060, 1199, 1316, 1408, 1460, 1488, 1503, 1655, 3030, 3119, 3172, 3192, 3202, 3300, 3800
R	0.273, 0.124, 0.094	61, 70, 123, 206, 312, 387, 422, 597, 923, 941, 950, 1025, 1069, 1201, 1327, 1410, 1463, 1491, 1505, 1688, 3039, 3125, 3170, 3182, 3198, 3282, 3717
Propene	1.552, 0.312, 0.273	199, 422, 581, 908, 942, 948, 1013, 1064, 1199, 1327, 1408, 1463, 1492, 1505, 1696, 3033, 3118, 3167, 3184, 3200, 3283
OH	18.956, 18.956, 0.000	3805

Table 22: *Unscaled vibrational frequencies and rotational constants for the various species at the UCCSD(T)/cc-pVTZ level of theory.*

species	B (cm^{-1})	vibrational frequencies (cm^{-1})
$\mathbf{A_{dir}}$	0.577, 0.093, 0.088	860i, 81, 96, 117, 355, 413, 582, 723, 928, 937, 960, 988, 1011, 1125, 1195, 1318, 1321, 1436, 1464, 1492, 1680, 3071, 3138, 3143, 3163, 3234, 3751
$\mathbf{A_{con}}$	0.462, 0.108, 0.097	1020i, 69, 120, 154, 329, 420, 568, 788, 926, 933, 948, 961, 1015, 1169, 1218, 1318, 1320, 1429, 1449, 1482, 1679, 3074, 3139, 3145, 3163, 3233, 3745
C	0.288, 0.205, 0.138	244i, 135, 160, 208, 253, 418, 580, 674, 886, 927, 944, 1005, 1055, 1196, 1303, 1402, 1451, 1486, 1501, 1628, 3035, 3102, 3130, 3139, 3182, 3241, 3746
T	0.379, 0.156, 0.127	222i, 93, 137, 173, 242, 419, 630, 667, 930, 940, 941, 994, 1057, 1193, 1308, 1406, 1452, 1483, 1497, 1641, 3030, 3090, 3121, 3145, 3162, 3251, 3752
R	0.271, 0.123, 0.094	59, 72, 120, 208, 297, 368, 419, 598, 929, 938, 945, 1025, 1074, 1194, 1320, 1411, 1456, 1487, 1501, 1688, 3033, 3099, 3116, 3136, 3150, 3230, 3681
Propene	1.547, 0.310, 0.271	200, 418, 583, 925, 931, 942, 1014, 1068, 1192, 1320, 1408, 1456, 1488, 1502, 1696, 3029, 3089, 3112, 3138, 3151, 3230
OH	18.852, 18.852, 0.000	3744